



HAL
open science

Stochastic control and applications: Mean field types and dynamic utilities

Chefia Ziri

► **To cite this version:**

Chefia Ziri. Stochastic control and applications: Mean field types and dynamic utilities. Probability [math.PR]. Le Mans Université; Université de Tunis El Manar, 2022. English. NNT: 2022LEMA1023 . tel-03869744

HAL Id: tel-03869744

<https://theses.hal.science/tel-03869744>

Submitted on 24 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE DE DOCTORAT DE

LE MANS UNIVERSITE

COMUE UNIVERSITE BRETAGNE LOIRE

En cotutelle avec l'Université de Tunis El Manar

ECOLE DOCTORALE N° 601

*Mathématiques et Sciences et Technologies
de l'Information et de la Communication*

Spécialité : Mathématiques

Par

« Chefia Ziri »

**« Stochastic control and applications : Mean field types and
dynamic utilities »**

Thèse présentée et soutenue à Le Mans, le 04 octobre 2022

Unité de recherche : Laboratoire Manceau de Mathématiques

Thèse N° : 2022LEMA1023,

Rapporteurs avant soutenance :

Boualem Djehiche, Professeur, KTH Stockholm, Suède.

Jean-François Chassagneux : Professeur, Université de Paris Cité, France

Composition du Jury :

Président :

Examineurs : Amel Ben Abda, Professeur, Université de Tunis El Manar, Tunisie

Caroline Hillairet, Professeur, ENSAE-Paris, France

Marina Kleptsyna, Professeur, Le Mans Université, France

Dir. de thèse : Anis Matoussi, Professeur, Le Mans Université, France

Co-dir. de thèse : Mohamed Mnif, Professeur, Université de Tunis El Manar, Tunisie

Co-encadrant : Mohamed Mrad, Maître de conférence, HDR, Université Sorbonne, Paris-Nord, France

Remerciements

Je tiens tout d'abord à remercier profondément mon directeur de thèse Anis Matoussi de m'avoir accueilli au sein de son équipe, qui a su me diriger tout au long de cette thèse avec patience, rigueur, enthousiasme et bonne humeur. Sa patience et ses approches toujours présent à la fois scientifique et pratique ainsi que ses qualités humaines m'ont énormément appris.

J'adresse toute ma gratitude envers mon co-directeur de thèse Mohamed Mnif qui a été présent par ses conseils et ses remarques pour accompagner mes premiers pas en recherche depuis le master.

Je remercie profondément Mohamed Mrad, mon co-encadrant de thèse, d'une part pour avoir accepté de faire partie de cette thèse et d'autre part pour ses conseils, ses remarques, sa générosité avec laquelle il partageait son savoir et sa grande patience.

Je tiens aussi à remercier Jean-Francois Chassagneux et Boualem Djehiche de m'avoir fait l'honneur de rapporter ma thèse. Je suis également très reconnaissante à l'encontre de Marina Kleptsyna, Amal Ben Abda et Caroline Hillairet pour avoir accepté de participer à ce jury en tant qu'examinatrices.

Je remercie également tous les membres du laboratoire Manceau de mathématiques ainsi que le département des mathématiques. Je suis reconnaissante pour leur accueil, leurs conseils et leur sympathie, sans oublier les chères secrétaires Brigitte, Sandie, Pauline et Marie-Charlotte. Avec une pensée très particulière à Cristina et Irène.

Je remercie aussi mes collègues de bureau, doctorants et anciens doctorants: Rym, Tingshu, Manal, Marius, Achraf, Lilit, Guillaume et Dmytro avec qui j'ai passé des moments agréables. Sans oublier Arij qui m'a supporté, soutenu tout au long de cette thèse.

Je remercie aussi tous les membres et les doctorants du laboratoire Lamsin, avec qui j'ai passé des moments agréables et sympathiques pendant les déjeuners et les pauses café, avec une pensée particulière à Mejed et la chère Raoudha.

Je remercie tous mes amis avec qui la vie est bien plaisante. Je pense particulièrement à Safa, Rabeb, Mariem, Mouna, Oussema, Sahar et Marcia je les remercie pour leurs bonnes humeurs, les moments conviviaux partagés et leurs soutiens pendant la rédaction confinée du manuscrit. Une pensée particulière à Yosra, mon amie depuis toujours, avec qui j'ai partagé cette aventure depuis le lycée et plein d'autres aventures que j'oublierai jamais.

Un remerciement particulier à Tata Mounira, Tonton Nouredine, Sana, Sabri, Julien et Alexa, qui étaient ma famille en France, pour leur amour, leurs soutiens et leurs conseils. Je n'oublierai jamais nos soirées jeux, les bons plats et les super gâteaux amoureuxment préparés. Avec une grande pensée à mon petit Aïdan et à ma petite boule d'amour et d'énergie Ambre.

Mes remerciements vont enfin vers ma petite famille. Je remercie mon cher père Kacem, ma chère mère Moufida, ma sœur Boutheina, mes frères Mehdi, Fehmi et Akrem et mon beau-frère Mohsen qui m'ont aimé, aidé, soutenu et encouragé tout au long de cette thèse, dans les hauts et les bas. Cette thèse est leur dédiée. Sans oublier, mes plus beaux garçons, mes neveux: Ghassen, Bassem, Ahmed et Anas, que j'adore, pour l'amour, le bonheur et la joie qu'ils m'apportent à chaque fois.

TABLE OF CONTENTS

1	Introduction	7
1.1	Mean field control problem and application to production of an exhaustible resource	7
1.1.1	An overview of mean field stochastic differential equations	7
1.1.2	Mean field control approach	11
1.1.3	Main results and contributions	16
1.2	Numerical approach for a class of mean field games	20
1.2.1	Mean field game approach	20
1.2.2	Numerical methods	24
1.2.3	Main results and contributions	27
1.3	Learning dynamic utilities	30
1.3.1	Dynamic utilities	30
1.3.2	Main results and contributions	39
2	Mean field control problem and application to production of an exhaustible resource	51
2.1	Introduction	51
2.2	Framework: Notations and setting	52
2.3	Linear quadratic mean field control problem on finite horizon	57
2.4	Extensions of LQMF problem on the case of common noise	67
2.5	Linear quadratic mean field control problem on infinite horizon	71
2.6	Application to production of an exhaustible resource	81
3	Numerical approach for a class of mean field games	89
3.1	Introduction	89
3.2	Framework: Notations and setting	90
3.3	Numerical Approximation	94
3.3.1	Approximating Markov chain	94
3.3.2	Discretization setting	95
3.3.3	Construction of Markov chain associated to MFG problem	96
3.3.4	Discrete MFG equilibrium	99
3.4	Existence and uniqueness results of discrete MFG solution	100
3.4.1	Existence result	100

TABLE OF CONTENTS

3.4.2	Uniqueness result	102
3.5	Convergence results	104
3.6	Numerical simulation	113
3.6.1	Linear quadratic model	113
3.6.2	Explicit solution	114
3.6.3	Numerical results	115
4	Learning dynamic utilities	123
4.1	Introduction	123
4.2	Preliminary results on the numerical approximation of dynamic utility	125
4.2.1	The model	126
4.2.2	Euler schemes	128
4.3	Machine Learning methods	132
4.3.1	Machine Learning programs using Python	133
4.3.2	Support Vector Machine	135
4.3.3	Learning utility with Neural Network	145
4.4	Learning utility with missing data	152
4.4.1	Learning with SVR model	152
4.4.2	Learning with NuSVR model	153
4.4.3	Learning with Neural Network model	154
4.5	Comparison of machine learning models	155
	Bibliography	169

INTRODUCTION

This thesis is a cotutelle thesis between Le Mans University, LMM-IRA and Tunis El-Manar University (UTM), LAMSIN-ENIT. I realized during this thesis, a set of 3 works on the stochastic mean field control problem and dynamic progressive utility. The works that are part of this thesis are the following:

- A. Matoussi, M. Mnif and C. Ziri. Linear quadratic control problems for mean field stochastic differential equation with Jumps. Preprint (2021), submitted for publication [88].
- S. Ben Aziza and C. Ziri. Numerical approximation for a class of mean field games. Forthcoming paper [14].
- M. Mrad and C. Ziri. Learning dynamic utilities. Forthcoming paper [94].

1.1 Mean field control problem and application to production of an exhaustible resource

The first part of this thesis is dedicated to the study of linear quadratic control problems for mean field stochastic differential equation with jumps (LQ-MF SDE in short) and application to the production of an exhaustible resource. This chapter is based on co-work with Matoussi and Mnif [88].

1.1.1 An overview of mean field stochastic differential equations

Historically, the mean-field stochastic differential equations, also known under the name of McKean Vlasov SDE, can be traced back to the works by Kac [65] and McKean [90]. These equations are SDEs whose drift and diffusion coefficients depend not only on the state of the unknown process, but also on its law of probability. Recently, this type of equation has been the subject of renewed interest in the context of mean field game theory and mean field control problems. So, we will dedicate the following part to introduce the main results of mean field BSDEs in the framework with and without jumps.

Mean field backward stochastic differential equations: Continuous setting Recently in 2009, Buckdahn et al. [109] have introduced a new type of backward stochastic differential equations and they have called them the mean field backward stochastic differential equations (MF BSDEs in short). They showed that the BSDE can be obtained by an approximation involving N-particle systems with weak interaction. Afterwords, Buckdahn, Li and Peng [23] generalized their previous results on mean field BSDEs in which the initial data (t, x) were frozen in the law variable of the coefficients, they investigated the associated nonlocal PDE.

More precisely, the authors [23] assumed the following Assumptions:

Assumption 1.1.1. • *There exists a constant $C \geq 0$ such that, \mathbb{P} -a.s., for all $t \in [0, T]$, $y_1, y_2, y'_1, y'_2 \in \mathbb{R}$, $z_1, z_2, z'_1, z'_2 \in \mathbb{R}^d$,*

$$|f(t, y'_1, z'_1, y_1, z_1) - f(t, y'_2, z'_2, y_2, z_2)| \leq C(|y'_1 - y'_2| + |z'_1 - z'_2| + |y_1 - y_2| + |z_1 - z_2|).$$

- $f(\cdot, 0, 0, 0, 0) \in \mathcal{H}_{\mathbb{F}}^2(0, T; \mathbb{R})$.

The authors established the following existence result:

Theorem 1.1.1. (*Buckdahn, Li and Peng [23]*)

Under Assumptions 1.1.1. For any random variable $\xi \in L^2(\omega, \mathcal{F}_T, \mathbb{P})$, the mean field BSDE

$$Y_t = \xi + \int_t^T \mathbb{E}'[f(s, Y_s, Y'_s, Z_s, Z'_s)]ds + \int_t^T Z_s dW_s, \quad (1.1.1)$$

where \mathbb{E}' is an operator defined by

$$\mathbb{E}'(\Gamma(\cdot, \omega)) := \int_{\Omega} \gamma(\omega', \omega) \mathbb{P}(d\omega'), \quad \forall \gamma \in L^1(\Omega^{1+1}, \mathcal{F} \otimes \mathcal{F}, \mathbb{P} \otimes \mathbb{P}),$$

has a unique solution $(Y, Z) \in \mathcal{S}_{\mathbb{F}}^2(0, T; \mathbb{R}) \times \mathcal{H}_{\mathbb{F}}^2(0, T; \mathbb{R}^d)$.

$$Y_t^1 \leq Y_t^2, \quad t \in [0, T], \mathbb{P}\text{-a.s.}$$

In 2016, Li, sun and Xiong [84] was the first to study the stochastic linear quadratic problem for mean field BSDEs. They have established the result of existence and uniqueness for the following linear mean field BSDE which depends on the laws of the state Y and the control α :

$$\begin{cases} dY_t &= (A_t Y_t + \hat{A}_t \mathbb{E}[Y_t] + B_t \alpha_t + \hat{B}_t \mathbb{E}[\alpha_t]) + C_t Z_t + \hat{C}_t \mathbb{E}[Z_t] dt + Z_t dW_t \quad t \in [0, T], \\ Y_T &= \xi, \end{cases} \quad (1.1.2)$$

where ξ is an \mathcal{F}_T -measurable random vector and $A, \hat{A}, B, \hat{B}, C, \hat{C}$ are given deterministic matrix-valued functions that satisfied

$$A, \hat{A} \in L^1(0, T : \mathbb{R}^{n \times n}), \quad B, \hat{B} \in L^2(0, T : \mathbb{R}^{n \times m}), \quad C, \hat{C} \in L^2(0, T : \mathbb{R}^{n \times n}).$$

Moreover, there exist a constant $K > 0$, independent of ξ and α , such that

$$\mathbb{E} \left[\sup_{t \leq s \leq T} |Y_s|^2 + \int_t^T |Z_s|^2 ds \right] \leq K \mathbb{E} \left[\sup_{t \leq s \leq T} |\xi|^2 + \int_t^T |\alpha_s|^2 ds \right].$$

Recently, Li [80] introduced a class of mean field reflected BSDEs which makes the connection between the results of classical reflected BSDEs (RBSDEs in short) [100] and those of MF BSDEs [23, 109]. Later, Djehiche, Elie and Hamadéne [42] deepen the mean field reflected forward backward SDEs by adding the dependence on the distribution of the component of the solution. Since MF BSDE has important application in Mathematical finance as optimal control, we refer the reader for example to [7, 22, 42, 43, 64, 80, 81, 84, 117, 118].

Mean field backward stochastic differential equations : Discontinuous setting One of the directions who attracted many researchers is the case involving a discontinuity in the dynamics of the state solution. In this setting, the mean field BSDE is no longer driven by Brownian motion, but also by a random jump measure. Roughly speaking, the mean field backward stochastic equation with jumps has the following form on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ on which the filtration $\mathbb{F} = (\mathcal{F}_t)_t$ satisfies the usual conditions of completeness and right continuity

$$\begin{cases} dY_s &= -f(s, Y_s, Z_s, R_s, \mathbb{P}_{(Y_s, Z_s, R_s)})ds + Z_s dW_s + \int_{\chi} R_s(e) \tilde{\pi}(ds, de), \\ Y_T &= \xi, \end{cases} \quad (1.1.3)$$

where

- $W = (W_t)_{0 \leq t \leq T}$ is a Brownian motion with respect to its natural filtration under \mathbb{P} .
- $\pi(dt, de)$ is a random jump measure defined on $\mathbb{R}_+ \times \chi$ where $\chi \subset \mathbb{R}^l \setminus \{0\}$ is non-empty open set equipped with its Borel field $\mathcal{B}(\chi)$.

To be more precisely, $\tilde{\pi}$ is the compensated random measure of π defined as following

$$\tilde{\pi}(dt, de) = \pi(dt, de) - \eta(dt, de).$$

where the associated compensator η is absolutely continuous with respect to $\lambda \otimes dt$ (i.e)

$$\eta(dt, de) = \lambda(de)dt.$$

for a σ -finite measure λ on $(\chi, \mathcal{B}(\chi))$ satisfying

$$\int_E (1 \wedge |e|^2) \lambda(de) < \infty.$$

In [83], Li were the first interested by a mean field BSDE driven by a Browian motion and an independent Poisson random measure. The author study the existence and uniqueness of MF BSDE with jumps under following assumptions

Assumption 1.1.2. (i) $f(\cdot, \cdot, 0, 0, 0, \delta_0) \in \mathcal{H}_{\mathbb{F}}^2(0, T)$.

(ii) f is Lipschitz with respect to $(y, z, r, \mu) \in \mathbb{R} \times \mathbb{R}^2 \times L^2(\chi, \mathbf{B}(\chi), \lambda) \times \mathcal{P}_2(\mathbb{R}^{d+1+d} \times L^2(\chi, \mathbf{B}(\chi), \lambda))$, uniformly with respect to (s, w) .

(iii) $\xi \in L^2(\mathcal{F}_T)$.

Theorem 1.1.2. (Li [83])

Under the Assumption 1.1.2, the mean field BSDE with jumps (1.1.3) has a unique solution $(Y, Z, R) \in \mathcal{S}_{\mathbb{F}}^2(0, T : \mathbb{R}) \times \mathcal{H}_{\mathbb{F}}^2(0, T : \mathbb{R}^d) \times \mathbb{H}_{\nu}^2(0, T : \mathbb{R})$.

In 2016, Tang and Meng [114] study linear quadratic optimal control problem for mean field backward stochastic differential equations driven by a Poisson random measure and a Brownian motion which depend on the law of the state X and the control α . They established the existence and uniqueness of linear quadratic BSDE with jumps in the following form

$$\begin{cases} dY_t = \left(A_t Y_t + \hat{A}_t \mathbb{E}[Y_t] + B_t \alpha_t + \hat{B}_t \mathbb{E}[\alpha_t] + C_t Z_t + \hat{C}_t \mathbb{E}[Z_t] \right. \\ \quad \left. + \int_{\chi} D_t R_t(e) \lambda(de) + \int_{\chi} \hat{D}_t \mathbb{E}[R_t(e)] \lambda(de) \right) dt + Z_t dW_t + \int_{\chi} R_t(e) \tilde{\pi}(de, dt), \\ Y_T = \xi, \end{cases} \quad (1.1.4)$$

where ξ is an \mathcal{F}_T -measurable random variable and $A, \hat{A}, B, \hat{B}, C, \hat{C}, D, \hat{D}$ satisfy assumption 1.1.3:

Assumption 1.1.3. The matrix valued functions $A, \hat{A}, B, \hat{B}, C, \hat{C}, D, \hat{D}$ are uniformly bounded measurable functions

Under the above assumptions, the existence result is stated as follows

Theorem 1.1.3. (Tang and Meng [114])

Let Assumption 1.1.3 be satisfied, then for any $(\xi, \alpha) \in L^2(\mathcal{F}_T) \times \mathcal{A}$, the state equation (1.1.4) has a unique solution $(Y, Z, R) \in \mathcal{S}_{\mathbb{F}}^2(0, T : \mathbb{R}) \times \mathcal{H}_{\mathbb{F}}^2(0, T : \mathbb{R}^d) \times \mathbb{H}_{\nu}^2(0, T : \mathbb{R})$.

In 2016, Min and Li [82] considered a new type of BSDEs namely MF BSDEs with jumps coupled with the value function. The authors proved that this type of BSDEs admits a unique solution, and established a comparison theorem as well as a dynamic programming principal.

1.1.2 Mean field control approach

Stochastic optimal control has been widely studied in recent decades due to its applications to mathematical finance, insurance, economics, engineering, etc. Recently, stochastic optimal control problems for mean field stochastic differential equations (in short MF SDEs) have attracted more attention. The control problems of mean field type are stochastic optimal control problems where the cost function and the parameters of the dynamics depend on the law of the controlled stochastic process. In general, such problems correspond to the control of a large number of agents by a global planner. This theory is referred to mean field control (e.g., by Bensoussan, Frehse, and Yam [17]) or control of McKean Vlasov dynamics (e.g., by Carmona and Delarue [34]). It has found applications such as risk management, portfolio management, insurance, economics, engineering, see [6, 11, 31, 57, 87, 91].

To be more precisely, let us consider the following mean field optimal control, with the state following equation

$$\begin{cases} dX_t &= b_t(X_t, \mathbb{P}_{X_t}, \alpha_t, \mathbb{P}_{\alpha_t})dt + \sigma_t(X_t, \mathbb{P}_{X_t}, \alpha_t, \mathbb{P}_{\alpha_t})dW_t, \\ X_0 &= x, \end{cases} \quad (1.1.5)$$

and the following cost functional

$$J(\alpha) := \mathbb{E} \left[\int_0^T e^{-\rho t} f_t(X_t, \mathbb{P}_{X_t}, \alpha_t, \mathbb{P}_{\alpha_t}) dt + e^{-\rho T} g(X_T, \mathbb{P}_{X_T}) \right]. \quad (1.1.6)$$

The main goal in the control problem is to minimise the cost functional J over the set of admissible controls \mathcal{A} defined for $\rho \geq 0$ as,

$$\mathcal{A} := \{ \alpha : \Omega \times [0, T] \rightarrow \mathbb{R}^m \text{ s.t. } \alpha \text{ is } \mathbb{F}\text{-adapted and } \int_0^T e^{-\rho t} \mathbb{E}[|\alpha_t|^2] dt < \infty \}.$$

An admissible control $\alpha^* \in \mathcal{A}$ is said the optimal control if

$$J(\alpha^*) = \inf_{\alpha \in \mathcal{A}} J(\alpha).$$

In general there is two possible approach to solve a stochastic optimal control: The dynamic programming, developed by Richard Bellman in the 1950's and the stochastic maximum principle formulated by Pontryagin and his group in the 1950's also.

Stochastic maximum principle : The pioneering works on the stochastic maximum principle were written by Kushner [68, 69]. Since then there have been a lot of works on this subject, like

Bensoussan [15], Peng [104]. The original version of Pontryagin's maximum principle was first introduced for deterministic control problems in the 1960's by Pontryagin et al. as in classical calculus of variation. The basic idea is to perturb 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.

In 2010, Buckdahn, Djehiche and Li [20] were the first to be interested in the optimal control for stochastic differential equations of mean field type. They introduce a problem in which the coefficients depend on the expected value of the state of the solution process. The authors were establish a Peng-type general stochastic maximum principle, specifying a necessary conditions for optimality, for a general action space, using the classical spike perturbation. This maximum principle differs from the classical one in the sense that here the first order adjoint equation turns out to be a linear mean-field backward SDE, however, the second order adjoint equation remains the same as in Peng's stochastic maximum principle.

Buckdahn et al. [21] extended the approach of Peng in the classic framework by considering the second order variational equations and the corresponding second order adjoint process in the mean field setting. Many authors have made contributions on optimal control problems for mean field SDEs, see for example the works of Anderson and Djehiche [7] and Carmona and Delarue [34] under some convexity assumptions. We can also refer the reader to Alasseur et al. [6] where the authors develop a model for the optimal management of energy storage and distribution in a smart grid system through an extended MFG.

Recently, motivated by economic examples, Graber [57] studied mean field control problem for a linear stochastic differential equation with common noise. The author characterize the solution both in the term of a stochastic maximum principle and a Riccati equation. He reformulated the Nash equilibrium for MFG as an optimal control problem in order to discuss the case where a mean field Nash equilibrium is also the solution to an optimal control problem.

Dynamic programming approach Other approach to solve the mean field control problem is the dynamic programming approach. The classical dynamic programming principle is a fundamental principle in the theory of stochastic control. The basic idea of the approach is to consider a family of control problems by varying the initial state values, and to derive some relations between the associated value functions. This principle, called the dynamic programming principle were initiated in the 1950s by Bellman. This approach yields a certain partial differential equation (PDE in short), of second order and nonlinear, called Hamilton-Jacobi-Bellman (HJB in short). When this PDE can be solved by the explicit or theoretical achievement of a smooth solution, the verification theorem validates the optimality of the candidate solution to

the HJB equation. This method is a weak formulation of the standard martingale optimality principle, where the verification theorem for stochastic control problem is a crucial step (see e.g EL Karoui [47] and Pham [105]).

For mean field control problems, stochastic dynamic programming requires adaptation. Optimality conditions for mean field control problem are derived either by stochastic calculus of variation (see e.g [7]) or by stochastic dynamic programming in the quadratic case (see e.g [17]), but not in the general case for the fundamental reason that Bellman's principle does not apply in its original form on the stochastic trajectories of say X_t if those depend also on its law. Therefore, the idea which is proposed by several researchers (Lions [85], Bensoussan et al. [18]), is to formulate the problem without any such restriction if we work with the probability measure of X_t and use the Fokker-Planck equation.

In 2014, Lauriere and Pironneau [79] under some density assumption on the probability law of the state process, show that Bellman's principle applies to the dynamic programming value function $V(\tau, \rho_\tau)$, where the dependency on ρ_τ is functional as in Lion's analysis of mean field games. Later, Pham and wei [106] derive also the Bellman equation for mean field stochastic control problem, and they prove a verification theorem in a McKean Vlasov framework. The authors give an explicit solutions to the Bellman equation for the linear quadratic mean field control problem, with applications to the mean-variance portfolio selection and a systemic risk model.

Recently, Pham et al. [11, 91] propose an other approach for solving linear quadratic mean field stochastic control problems. The authors introduce the martingale formulation for verification theorem for solving a linear quadratic mean field stochastic problems. The optimal control involves the solution to a system of Riccati ordinary differential equations and to a linear mean field backward stochastic differential equation.

The verification theorem, the key of this approach, is introduced in the following form :

Theorem 1.1.4. *(Finite horizon Verification Theorem [11])*

Let $\{W_t^\alpha, t \in [0, T], \alpha_t \in \mathcal{A}\}$ be a family of \mathbb{F} -adapted process in this form $W_t^\alpha = w_t(X_t^\alpha, \mathbb{E}[X_t^\alpha])$ for some \mathbb{F} -adapted random field $\{w_t(x, \bar{x}), t \in [0, T], x, \bar{x} \in \mathbb{R}^d\}$ satisfying

$$w_t(x, \bar{x}) \leq C(\mathcal{X}_t + |x|^2 + |\bar{x}|^2), \quad t \in [0, T], \quad x, \bar{x} \in \mathbb{R}^d,$$

for some positive constant C , and non-negative process \mathcal{X} with $\sup_{t \in [0, T]} \mathbb{E}[|\mathcal{X}_t|] < \infty$, and such that

(i) $w_T(x, \bar{x}) = g(x, \bar{x}), \quad x, \bar{x} \in \mathbb{R}^d.$

(ii) The map $t \in [0, T] \mapsto \mathbb{E}[\mathbf{S}_t^\alpha]$, with

$$\mathbf{S}_t^\alpha = e^{-\rho t} W_t^\alpha + \int_0^t e^{-\rho s} f_s(X_s^\alpha, \mathbb{E}[X_s^\alpha], \alpha_t, \mathbb{E}[\alpha_t]) ds,$$

is non-decreasing for all $\alpha \in \mathcal{A}$.

(iii) The map $t \mapsto \mathbb{E}[\mathbf{S}_t^{\alpha^*}]$ is constant for some $\alpha^* \in \mathcal{A}$.

Then, α^* is an optimal control and $\mathbb{E}[w_0(X_0, \mathbb{E}[X_0])]$ is the value of the LQ MF control problem (1.1.6)

$$V_0 = \mathbb{E}[w_0(X_0, \mathbb{E}[X_0])] = J(\alpha^*).$$

Moreover, any other optimal control satisfies the condition (iii).

In [11], Basei and Pham interested to solve a linear quadratic control problem. Roughly speaking, they consider a quadratic cost functional of the following type, where $\alpha \in \mathcal{A}$,

$$J(\alpha) := \mathbb{E} \left[\int_0^T e^{-\rho t} f_t(X_t^\alpha, \mathbb{E}[X_t^\alpha], \alpha_t, \mathbb{E}[\alpha_t]) dt + e^{-\rho T} g(X_T, \mathbb{E}[X_T]) \right], \quad (1.1.7)$$

where for each $t \in [0, T]$, $x, \bar{x} \in \mathbb{R}^d$ and $a, \bar{a} \in \mathbb{R}^m$

$$\begin{aligned} f_t(x, \bar{x}, a, \bar{a}) &:= (x - \bar{x})^t Q_t (x - \bar{x}) + \bar{x}^T (Q_t + \tilde{Q}_t) \bar{x} + 2a^T I_t (x - \bar{x}) \\ &+ 2\bar{a}^T (I_t + \tilde{I}_t) \bar{x} + (a - \bar{a})^T N_t (a - \bar{a}) + \bar{a}^T (N_t + \tilde{N}) \bar{a} + 2M_t^T x + 2H_t^T a, \end{aligned} \quad (1.1.8)$$

and

$$g(x, \bar{x}) := (x - \bar{x})^T P (x - \bar{x}) + \bar{x}^T (P + \tilde{P}) \bar{x} + 2L^T x, \quad (1.1.9)$$

subject to

$$\begin{cases} dX_t^\alpha &= (\beta_t + A_t x + \tilde{A}_t \bar{x} + B_t a + \tilde{B}_t \bar{a}) dt + (\gamma_t + C_t x + \tilde{C}_t \bar{x} + D_t a + \tilde{D}_t \bar{a}) dW_t, \\ X_0 &= x, \end{cases} \quad (1.1.10)$$

under the following Assumptions :

Assumption 1.1.4. The coefficients in equation (1.1.10) satisfy:

- (i) $\beta, \gamma \in L^2_{\mathbb{F}}(\Omega \times [0, T], \mathbb{R}^d)$,
- (ii) $A, \tilde{A}, C, \tilde{C} \in L^\infty([0, T], \mathbb{R}^{d \times d})$, $B, \tilde{B}, D, \tilde{D} \in L^\infty([0, T], \mathbb{R}^{d \times m})$.

Assumption 1.1.5. The coefficients in equation (1.1.8)-(1.1.9) satisfy:

- (i) $Q, \tilde{Q} \in L^\infty([0, T], \mathbb{S}^d)$, $P, \tilde{P} \in \mathbb{S}^d$, $N, \tilde{N} \in L^\infty([0, T], \mathbb{S}^m)$, $I, \tilde{I} \in L^\infty([0, T], \mathbb{R}^{m \times d})$,

(ii) $M \in L^2_{\mathbb{F}}(\Omega \times [0, T], \mathbb{R}^d)$, $H \in L^2_{\mathbb{F}}(\Omega \times [0, T], \mathbb{R}^m)$, $L \in L^2_{\mathbb{F}}(\mathbb{R}^d)$,

(iii) There exists $\xi > 0$ such that, for each $t \in [0, T]$,

$$N_t \geq \xi \mathbb{I}_m, \quad P \geq 0, \quad Q_t + I_t^T N_t^{-1} I_t \geq 0.$$

(iv) There exists $\xi > 0$ such that, for each $t \in [0, T]$,

$$N_t + \tilde{N}_t \geq \xi \mathbb{I}_m, \quad P + \tilde{P} \geq 0, \quad (Q_t + \tilde{Q}_t) - (I_t + \tilde{I}_t)^T (N_t + \tilde{N}_t)^{-1} (I_t + \tilde{I}_t) \geq 0.$$

Remark 1.1.5. The uniform positive definite assumption on N and $N + \tilde{N}$ is a standard and natural coercive condition when dealing with linear-quadratic control problems.

Thanks to the weak formulation of verification theorem (see, Theorem 1.1.6) and some techniques, Basei and Pham have proved the existence of an optimal control:

Theorem 1.1.6. (Basei and Pham [11])

Under Assumptions 1.1.4-1.1.5, there exists an optimal control α^* for LQMF problem (1.1.6) giving by the following explicit form

$$\alpha_t^* = -S_t^{-1} U_t (X_t^* - \bar{X}_t^*) - S_t^{-1} (\epsilon_t - \bar{\epsilon}_t) - S_t^{-1} (V_t \bar{X}_t^* + \Theta_t), \quad \forall t \in [0, T] \quad \mathbb{P} - a.s.,$$

where $X^* = X^{\alpha^*}$ is the state process where the α^* is the optimal control and the deterministic coefficients S , \hat{S} , U and V and the stochastic coefficients ϵ_t and Θ_t , are defined by

$$\begin{cases} S_t := N_t + D_t^\top K_t D_t, \\ \hat{S}_t := \hat{N}_t + \hat{D}_t^\top K_t \hat{D}_t, \\ U_t := I_t + D_t^\top K_t C_t + B_t^\top K_t, \\ V_t := \hat{I}_t + \hat{D}_t^\top K_t \hat{C}_t + \hat{B}_t \Lambda_t, \\ \Theta_t := \bar{H}_t + \hat{D}_t^\top K_t \bar{\gamma}_t + \hat{D} \bar{Z}_t^Y. \end{cases} \quad (1.1.11)$$

And $(K, \Lambda, Y, Z^Y, \varphi) \in L^\infty([0, T], \mathbb{S}^d) \times L^\infty([0, T], \mathbb{S}^d) \times \mathcal{S}_{\mathbb{F}}^2(\Omega \times [0, T], \mathbb{R}^d) \times L^2_{\mathbb{F}}(\Omega \times [0, T], \mathbb{R}^d) \times \chi, \mathbb{R}^d) \times L^\infty([0, T], \mathbb{S}^d)$, is the unique solution to the decoupled system (1.1.20):

$$\begin{cases} dK_t = \dot{K}_t dt, \quad 0 \leq t \leq T, \quad K_T = P, \\ d\Lambda_t = \dot{\Lambda}_t dt, \quad 0 \leq t \leq T, \quad \Lambda = \hat{P}, \\ dY_t = \dot{Y}_t dt + Z_t^Y dW_t, \quad 0 \leq t \leq T, \quad Y_T = L, \\ d\varphi_t = \dot{\varphi}_t dt, \quad 0 \leq t \leq T, \quad \varphi_T = 0. \end{cases} \quad (1.1.12)$$

In [91], Pham and Miller propose an approach based on weak martingale optimality principle together with a fixed point argument in the space of controls for solving a stochastic differential game with linear McKean-Vlasov dynamics and quadratic cost functional. They characterize the Nash equilibria by a terms of systems of Riccati ordinary differential equations and linear mean field backward stochastic differential equations.

1.1.3 Main results and contributions

Motivation In this work, we are motivated by an energy market to produce energy from an exhaustible resource such as oil. Roughly speaking, each producer optimizes production rate that drives her revenue, as well as exploration effort to replenish her reserve. This exploration activity is modeled by a jumps process that leads to stochastic increments to reserves level. This model, represented by a model of production of exhaustible resource with accumulating or maintaining a level of reserves, is inspired by a series of works extended from the Hotelling's model [63]. In the classic Hotelling's model, the dynamic market evolution is driven by the use of existing reserves of an exhaustible reserves to produce energy without possibility to exploration and/or discovery of new reserves. However, many studies have made it possible to ensure that there are still resource bases to be explored over time, that is to say that the reservation rate can be increased. Particularly we can see the paper of Pindyck [107] that considered a deterministic model of exploring for a non-renewable resource. The author assumed exploration to be incremental and represented as a deterministic reserve addition. A series of works extended Prindyck's model [107] for the case of stochastic discoveries such as: Deshmukh et al. [41], Arrow and Chang [9] and Keller et al. [60]. This increase in reserve discoveries occurs stochastically via the Point process, it should be noted that this increase is smaller, and it is the reason that the resources always remain exhaustible. In this work, we are particularly motivated by a continuous model of the production of an exhaustible resource recently introduced by Guèant et al. [59] where they have considered a mean field version of a this model with a quadratic cost function. Roughly speaking, the authors consider a large number of oil producers, which can be viewed either as wells or from a more macro standpoint as oil companies. Each oil producers i initially has a reserve x_0^i where are distributed according to an initial distribution. These reserves X^i will of course contribute to production α^i such that, for any specific producer i , his reserve dynamic is given by the following stochastic differential equation

$$\begin{cases} dX_t^i &= -\alpha_t^i dt + \sigma X_t^i dW_t^i, \\ X_0^i &= x_0^i, \end{cases}$$

where W^i is Brownian motion specific to each producer. Each production need to optimize a profit criterion of the following form

$$\max_{\alpha^i} \mathbb{E} \left(\int_0^\infty e^{-r t} [\alpha_t^i p_t - C(\alpha_t^i)] dt \right), \text{ s.t. } \alpha \geq 0, X_t \geq 0, \quad (1.1.13)$$

where

- p is the selling price for producer i follows. The price p of producer i is related to his production and also to the production of all other producers.
- C is the cost functions of production, defined as

$$C(\alpha) := \delta\alpha + \beta \frac{\alpha^2}{2}.$$

In this work, we introduce a model of production of exhaustible resource with accumulating or maintaining a level of reserves. Reserves level X_t^i of each producer decreases at a controlled production rate $\alpha^i \geq 0$, and also has random discrete increment due to exploration. We use a Poisson point process $\tilde{\pi}^i$ to model the new discoveries. To be more precisely, $\tilde{\pi}^i$ has intensity $\lambda(t)a_t^i$, where is the effort of the exploration controlled by player i . So the reserve dynamic of each producer is given by the following stochastic differential equation

$$\begin{cases} dX_t^i &= -\alpha_t^i dt + \sigma X_t^i dW_t^i + \eta \int_{\mathcal{X}} X_t^i \tilde{\pi}^i(de, dt), \\ X_0^i &= x_0^i, \end{cases} \quad (1.1.14)$$

where x_0^i is the initial reserves level and η is the unit amount of discoveries . Each player i wants to maximize his own profit (1.1.15)

$$J(\alpha) := \mathbb{E} \left(\int_0^\infty e^{-\rho t} [\alpha_t^i p_t(\alpha_t^i) - C_p(\alpha_t^i) - C_{\text{ex}}(\alpha_t^i, X_t^i)] dt \right), \quad (1.1.15)$$

where

- P_t^i is the selling price for producer i . It follows a linear inverse demand rule, defined as:

$$P_t^i := P_t(\alpha_t^i) = P_t^0 - \delta \alpha_t^i - \varepsilon \int_0^t \frac{1}{N} \sum_{j=1}^N \alpha_s^j ds.$$

Here δ, ε are positive constants and P^0 is a deterministic function. It will be the same for all producers. The price P^i of producer i is related to his production and also to the production of all other producers.

- C_p is the cost functions of production, defined as:

$$C_p(\alpha_t^i) := c_1 \text{Var}(\alpha_t^i).$$

- C_{ex} is the cost functions of exploration defined as:

$$C_{\text{ex}}(\alpha_t^i, X_t^i) := c_2 \alpha_t^i \left(\frac{x_0^i - X_t^i}{x_0^i} \right).$$

with c_1 and c_2 are positive constants and represent respectively the cost of production and the cost of extraction where are the same for the all producers. In the following, we will propose a explicit method to solve this problem.

Our main result : Our first contribution presented in this thesis to study the solvability of linear quadratic stochastic control problem for mean field stochastic differential equation with jumps in the both finite and infinite horizon. The mean field LQ optimal control problem was initially studied by Yong [117] and was generalized later by Huang, Li and Yong [64], Li, Sun and Xiong [84] to variant cases. Along with the development of mean field LQ optimal control problems, the LQ differential game for mean field SDEs have also attracted extensive research, among which, we would like to mention Bensoussan et al. [17] and Graber [57].

Here, we propose an approach to solve the linear quadratic control problems for mean field stochastic differential equation with jump defined as :

$$V(\alpha^*) = \inf_{\alpha} J(\alpha), \quad (1.1.16)$$

where

$$\begin{aligned} J(\alpha) := & \mathbb{E} \left[\int_0^T e^{-\rho t} \left((x - \bar{x})^t Q_t (x - \bar{x}) + \bar{x}^T (Q_t + \tilde{Q}_t) \bar{x} + 2a^T I_t (x - \bar{x}) \right. \right. \\ & + 2\bar{a}^T (I_t + \tilde{I}_t) \bar{x} + (a - \bar{a})^T N_t (a - \bar{a}) + \bar{a}^T (N_t + \tilde{N}) \bar{a} + 2M_t^T x + 2H_t^T a, \left. \left. dt \right. \right. \\ & \left. \left. + e^{-\rho T} \left((x - \bar{x})^T P (x - \bar{x}) + \bar{x}^T (P + \tilde{P}) \bar{x} + 2L^T x \right) \right], \end{aligned} \quad (1.1.17)$$

subject to

$$\begin{cases} dX_t^\alpha &= (\beta_t + A_t x + \tilde{A}_t \bar{x} + B_t a + \tilde{B}_t \bar{a}) dt + (\gamma_t + C_t x + \tilde{C}_t \bar{x} + D_t a + \tilde{D}_t \bar{a}) dW_t \\ &+ \int_x (\delta_t + F_t(e) x + \tilde{F}_t(e) \bar{x} + G_t(e) a + \tilde{G}_t(e) \bar{a}) \tilde{\pi}(de, dt), \\ X_0 &= x, \end{cases} \quad (1.1.18)$$

under Assumptions 1.1.4-1.1.5 by adding also the following one:

Assumption 1.1.6. *The coefficients in equation (1.1.18) satisfy:*

- (i) $\delta \in L^2_{\mathbb{F}}(\Omega \times [0, T], \mathbb{R}^d)$,
- (ii) $G(\cdot), \tilde{G}(\cdot) \in \mathcal{K}^2([0, T] \times \chi, \mathbb{R}^{d \times d})$, $F(\cdot), \tilde{F}(\cdot) \in \mathcal{K}^2([0, T] \times \chi, \mathbb{R}^{d \times m})$.

Our approach is a weak formulation of the standard martingale optimality principle used in the verification theorem for stochastic control problems. Roughly speaking, using the Verification Theorem 1.1.6 introduced in [91], we get an explicit form of the optimal control:

Theorem 1.1.7. *Under Assumptions (1.1.4)-(1.1.5)-(1.1.6), there exists an optimal control α^* for LQMF problem (2.3.14) giving by the following explicit form*

$$\alpha_t^* = -S_t^{-1}U_t(X_t^* - \bar{X}_t^*) - S_t^{-1}(\epsilon_t - \bar{\epsilon}_t) - S_t^{-1}(V_t\bar{X}_t^* + \Theta_t), \quad \forall t \in [0, T] \quad \mathbb{P} - a.s.,$$

where $X^* = X^{\alpha^*}$ is the state process where the α^* is the optimal control and the deterministic coefficients S, \hat{S}, U and V and the stochastic coefficients ϵ_t and Θ_t , are defined by

$$\begin{cases} S_t := N_t + D_t^T K_t D_t + \int_{\chi} G_t^T(e) K_t G_t(e) \lambda(de), \\ \hat{S}_t := \hat{N}_t + \hat{D}_t^T K_t \hat{D}_t + \int_{\chi} \hat{G}_t^T(e) K_t \hat{G}_t(e) \lambda(de), \\ U_t := I_t + D_t^T K_t C_t + B_t^T K_t + \int_{\chi} G_t^T(e) K_t F_t(e) \lambda(de), \\ V_t := \hat{I}_t + \hat{D}_t^T K_t \hat{C}_t + \hat{B}_t^T \Lambda_t + \int_{\chi} \hat{G}_t^T(e) K_t \hat{F}_t(e) \lambda(de), \\ \Theta_t := \bar{H} + \hat{D}_t^T K_t \bar{\gamma}_t + \hat{B}_t^T \bar{Y}_t + \hat{D}_t \bar{Z}_t^Y + \int_{\chi} \hat{G}_t^T(e) K_t \bar{\delta}_t \lambda(de) + \int_{\chi} \hat{G}_t^T \bar{R}_t^Y(e) \lambda(de), \\ \epsilon_t := H_t + D_t^T K_t \gamma_t + B_t^T Y_t + D_t^T Z_t^Y + \int_{\chi} G_t^T(e) K_t \delta_t \lambda(de) + \int_{\chi} G_t^T(e) R_t^Y(e) \lambda(de). \end{cases} \quad (1.1.19)$$

And $(K, \Lambda, Y, Z^Y, R^Y, \varphi) \in L^\infty([0, T], \mathbb{S}^d) \times L^\infty([0, T], \mathbb{S}^d) \times \mathbb{S}_{\mathbb{F}}^2(\Omega \times [0, T], \mathbb{R}^d) \times L^2_{\mathbb{F}}(\Omega \times [0, T], \mathbb{R}^d) \times \mathcal{K}^2(\Omega \times [0, T] \times \chi, \mathbb{R}^d) \times L^\infty([0, T], \mathbb{S}^d)$, is the unique solution to the following decoupled system:

$$\begin{cases} dK_t = \dot{K}_t dt, \quad 0 \leq t \leq T, \quad K_T = P, \\ d\Lambda_t = \dot{\Lambda}_t dt, \quad 0 \leq t \leq T, \quad \Lambda = \hat{P}, \\ dY_t = \dot{Y}_t dt + Z_t^Y dW_t + \int_{\chi} R_t^Y(e) \tilde{\pi}(de, dt), \quad 0 \leq t \leq T, \quad Y_T = L, \\ d\varphi_t = \dot{\varphi}_t dt, \quad 0 \leq t \leq T, \quad \varphi_T = 0. \end{cases} \quad (1.1.20)$$

By the theory developed above, we will solve a control problem for a model of production of exhaustible resource with accumulating or maintaining a level of reserves.

In our work, we assume that the common price $P_t^0 = p^0$, for all $t \geq 0$ where p^0 is interpreted as a substitute price for the exhaustible resource. In the next proposition, we gives an explicit solution to the problem (1.1.15).

Proposition 1.1.8. *We assume that $p^0 = c_2 + \varepsilon x_0$ for all $t \geq 0$, x_0 is large enough and $1 > \rho^2 + \frac{2\varepsilon}{\delta}$. Then the solution of (1.1.15) is given by*

$$\alpha_t^* = K_{c_1, c_2} X_{t^-}^* + (\Lambda_\varepsilon - K_{c_1, c_2}) \bar{X}_{t^-}^*.$$

Finally, we discussed two different cases: The first one, when $p^0 = c_2 + \varepsilon x_0$ i.e. p^0 coincides with $c_2 + \varepsilon x_0$ which is the cost of extraction for the last unit of resource. The second case when $p^0 < c_2 + \varepsilon x_0$ (i.e.) the Hotelling rent is negative. It is not obvious to check the positivity of the the state process and the optimal strategy. We study the stationary level of the reserve and the optimal production rate in mean.

1.2 Numerical approach for a class of mean field games

The second part of this thesis is dedicated to the numerical resolution of a mean field game problem. We will introduce an approach, called Markov chain approximation (MCA in short) where was developed by H. Kushner and P.G. Dupuis [70]. This method applies a Markov chain approximation to continuous time and continuous space state control problem. This chapter is based on co-work with Ben Aziza [14].

1.2.1 Mean field game approach

The theory of mean field games (MFG in short) is a recent branch of dynamic games that aims at modeling and analyzing complex decision processes involving a large number of agents, which have individually a small influence on the overall system, and are influenced by the behavior of other agents. The MFG is presented by a control problem that approximate a large number of players in interaction. It's devoted to mathematically analyze a problem where there are large number of small players, who have similar references. The theory was introduced since 2006 in series of seminal papers, by Lasry and Lions [73, 75, 77], Caines et al. [25]. Since its inception, the mean field games theory has expanded tremendously, and has become an important tool in the study of dynamical and equilibrium behavior of large systems.

Let us consider a game of a N players who act in terms of the strategies of the other players. The players are in interaction taking into account the global behaviors of the flow. Each player takes his decision in a strategic way considering the decisions of the others presented through the law distribution of the field.

Let X^i be the state of player i , $i = 1, \dots, N$, which satisfies the following stochastic differential

equation

$$\begin{cases} dX_t^i &= -\alpha_t^i dt + \sqrt{2}\sigma dB_t^i, \\ X_0^i &= x^i, \end{cases}$$

where σ is a fixed positive constant, $(B_t^i)_{1 \leq i \leq N}$ are N Brownian Motion independents and α^i the control associated to the player i . Each individual player i will need to choose his strategy α^i to minimize his cost functional J^i over the period $[0, T]$, which is influenced by the state of other players:

$$J^i(\alpha^i, (\alpha^j)_{i \neq j}) := \mathbb{E} \left[\int_0^T (L(s, X_s^i, \alpha_s^i, m_s^i) ds + G(X_T^i, m_T^i)) \right], \quad (1.2.21)$$

where g is the terminal cost and L is measurable and locally bounded functions which satisfies the coercivity condition

$$\lim_{|\alpha| \rightarrow \infty} \frac{|L(t, X_t, \alpha_t, m_t)|}{|\alpha|} = +\infty,$$

with $m_t^i = \sum_{i \neq j} \delta_{X_t^j}$ is the probability measure of other players and δ is the Dirac measure.

For a fixed number N of players, the player i solves his optimization problem assuming that the other players have already made their decisions, in order to find the Nash equilibrium.

Definition 1.2.1. (Nash equilibrium)

$(\alpha^1, \dots, \alpha^N)$ is a Nash equilibrium of the system of N players if

$$J(\alpha^i, (\alpha^j)_{i \neq j}) \leq J(\alpha, (\alpha^j)_{i \neq j}), \text{ for all } i \in \{1, \dots, N\} \text{ and } \alpha \in \mathcal{A}.$$

The Nash equilibrium is thus such that no player regrets his choice in view of the choice of the others, he could not have done better. Often Nash equilibrium is presented as a situation where each player adopts the best answer "given" the choice of the others.

Let us introduce a Hamiltonian system of N players satisfy the following equation :

$$H(x, m, p, m) = \sup_{\alpha \in \mathbb{R}^d} \{-p\alpha - L(t, x, \alpha, m)\},$$

and define the Nash system of N players by the following equation

$$\begin{cases} -\partial_t u^i - \nu \Delta u^i + H(x^i, m^i, Du^i) = f(x^i, m_t^i), \\ u(x^i, T) = \phi(x^i, m_T^i). \end{cases} \quad (1.2.22)$$

When a solution to the above system exists and is regular, a Nash equilibrium is given by the control functions

$$\alpha^i = -H_p^i(t, x^i, \nabla_{x^i} u^i, m^i).$$

Then, when players use these controls, u^i is called the value function of player i , and represents the lowest average cost that player i can achieve between times t and T if its state is given by x at time t . The system (1.2.22) is a system of N coupled Hamilton-Jacobi equations. We refer the reader to the work of Crandall and Lions, for a thorough study of the Hamilton-Jacobi equations and the framework in which they are well posed the theory of viscosity solutions and to the work of A. Bensoussan and J. Frehse [16, 17]. The main difficulty in the study of the system (1.2.22) consists in the fact that a player, when he changes his strategy, affects all the other players. Thus, the existence and uniqueness of solutions of (1.2.22) are in general complicated problems. However, when the number of players is large, any qualitative or quantitative study of the behavior of the solutions of (1.2.22) becomes impossible because of the number of equations and their strong coupling. Contrary to what the previous remark suggests, the structure of Nash equilibrium should simplify when the number of players becomes very large since the effect of each player on the system becomes negligible. This is the founding idea of the theory of mean field games, developed at the same time in the initial independent works of Lasry and Lions [74, 76, 85] and of Huang, Caines and Malhamé [25].

The MFG problem is mathematically described through a system of two non linear partial differential equations introduced by a following system

$$\begin{cases} -\partial_t u - \nu \Delta u + H(x, m, Du) = f(x, m), & \text{in } \mathbb{T}^d \times (0, T), \\ \partial_t m - \nu \Delta m - \operatorname{div}(D_p H(x, m, Du)m) = 0, & \text{in } \mathbb{T}^d \times (0, T), \\ m(0) = m_0, \\ u(x, T) = \phi(x, m(T)), \end{cases} \quad (1.2.23)$$

where ν is a non-negative parameter, and m_0 and m_t are the probability or the density. The Hamiltonian $H : (x, p) \in \mathbb{R}^d \times \mathbb{R} \rightarrow H(x, p) \in \mathbb{R}$ is convex function with respect to p defined by

$$H(x, m, p) = \sup_{\alpha} \{-p\alpha - L(x, \alpha)\}.$$

This is a forward-backward system. The first equation is called a Hamilton-Jacobi equation where is associated with an optimal control problem. This equation describes the reaction of players to the mass. It presents the decisions based on where you want to be in the future. The second one is called a transport or Fokker-Planck equation for the distribution of the agents. The equation describes the aggregation of the action of all players. It presents where the population behavior actually ends up based on the initial distribution. The existence of solutions of HJB-FP system (1.2.23) is treated in many works [16, 25, 26, 32, 33, 58, 74]. A uniqueness criterion for the solution of the field game system was presented by J.M. Lasry and P.L. Lions in their first

work [74] on the theory of mean field games called Lasry-Lions monotonicity condition defined as :

$$\int (f(x, m) - f(x, m')) d(m - m')(x) \geq 0 \quad \forall m, m' \in \mathcal{P},$$

and

$$\int (\phi(x, m) - \phi(x, m')) d(m - m')(x) \geq 0 \quad \forall m, m' \in \mathcal{P}.$$

The interpretation of the above Monotonicity conditions is that the players dislike congested regions and prefer configurations in which they are scattered.

Many works have treated this system in the periodic setting, for example [27, 28, 74, 78, 85]. They establish the existence result of classical solution under a wide range of sufficient conditions on H , f and ϕ , and the uniqueness result under the uniform convexity of the Hamiltonian and Lasry-Lions monotonicity condition. Here, we consider a mean field game model where any players controls its private state X_t at time t , by taking an action α_t in a closed convex subset \mathcal{A} , where the dynamic of state is driven by the following the stochastic differential equation

$$dX_t = \alpha_t dt + \sqrt{2\sigma} dB_t, \quad (1.2.24)$$

where B_t is a standard Brownian motion, σ is a fixed positive constant and \mathcal{A} is the set of admissible control.

Each player will need to choose his strategy to minimize his cost functional (1.2.25) over the period $[0, T]$, which is influenced by the state of other players:

$$J(x, t, \alpha, m) = \mathbb{E}\left[\int_t^T (L(X_s^\alpha, \alpha_s) + F(X_s^\alpha, m_s)) ds + \Phi(X_T^\alpha, m_T)\right]. \quad (1.2.25)$$

The problem of differential games with an infinite number of agents is articulated in the following form:

- (i) for each fixed deterministic flow of probability measures m on \mathbb{R}^d , solve the standard stochastic control problem:

$$V(x, t) := \inf_{\alpha \in \mathcal{A}} J(x, t, \alpha, m). \quad (1.2.26)$$

- (ii) Find a flow m such that $\mathcal{L}(\hat{X}_t) = m_t$ for all $t \in [0, T]$, where \hat{X}_t is a solution to the above SDE (1.2.24).

We can see that the first step provides the best response of a given player interacting with the statistical distribution of the states of the other players if this distribution m is assumed to be given by t , when the second step solves a fixed point problem in the goal of the search for fixed

points of the best response function. Once these two steps have been taken successfully, if the fixed-point optimal control $\hat{\alpha}$ identified in step (ii) is in feedback form $\alpha_t^* = \phi(t, X_t^\alpha, m_t)$, for some deterministic function ϕ on $[0, T] \times \mathbb{T}^d \times \mathcal{P}(\mathbb{R}^d)$. Then, if the players use α_t^* , the mean field equilibrium $\{m_t\}_{0 \leq t \leq T}$ with value in $\mathcal{P}(\mathcal{K})$ should coincide with the law of X and satisfy the Fokker-Planck equation.

The literature MFG has grown considerably: many authors work on this subject, see : Cardaliaguet [27], Bensoussan et al. [17], Carmona and Delarue [33], Gomes et al. [52–54]. A lot of MFG problems have explicit or semi-explicit solutions, but it's not always obvious or easy to solve a coupled MFG PDE system. Therefore, numerical simulations of MFGs play a crucial role to solve this models.

1.2.2 Numerical methods

Numerical resolution has been interested by many authors and there is still an active on going research dealing with the numerical approximation of different types of MFG systems most of them are based on the PDE system introduced by Lions and Lasry in [77]. Finite difference methods has been the first method of approximation the PDE system whose proposed by Achdou et al. in [3]. They propose a finite difference scheme with monotone approximations of the Hamiltonian and a discrete weak formulation of the Fokker-Planck equation, both in finite and infinite horizon.

To introduce the approximation of HJB-FP system (1.2.23) proposed by Achdou et al. [3], let us consider two positive integers N_T and N_h and $\Delta t = T/N_T$ and $h = 1/N_h$ and $t_n = n \times \Delta t$, $x_i = i \times h$ for $(n, i) \in \{0, \dots, N_T\} \times \{0, \dots, N_h\}$. We approximate u and m respectively by vectors U and M such that for each $(n, i) \in \{0, \dots, N_T\} \times \{0, \dots, N_h\}$,

$$u(t_n, x_i) \approx U_{i,n}^h, \quad \text{and,} \quad m(t_n, x_i) \approx M_{i,n}^h.$$

Remark 1.2.2. *To simplify the notation, we present the scheme in the framework of a one dimension.*

Finite difference operators : Let us introduce the elementary finite difference operators:

$$\begin{aligned} D_t U_{i,n}^h &:= \frac{U_{i,n+1}^h - U_{i,n}^h}{\Delta t}, \quad \text{for } 0 \leq i \leq N_T - 1, \\ D U_{i,n}^h &:= \frac{U_{i+1,n}^h - U_{i,n}^h}{h}, \quad \text{for } 1 \leq i \leq N_h - 1, \\ \Delta_h U_{i,n}^h &= -\frac{2U_{i,n}^h - U_{i+1,n}^h - U_{i-1,n}^h}{h^2}, \quad \text{for } 1 \leq i \leq N_h - 1, \\ \nabla_h U_{i,n}^h &:= ((DU)_i, (DU)_{i-1})^T, \quad \text{for } 0 \leq k \leq N_h - 1. \end{aligned} \tag{1.2.27}$$

We consider \tilde{H} the discrete Hamiltonian and assume that it satisfy the following Assumptions:

Assumption 1.2.1. Let $(x, p_1, p_2) \mapsto \tilde{H}(x, p_1, p_2)$ be a discrete Hamiltonian, assumed to satisfy the following properties

1. *Monotonicity* : for any $x \in \mathbb{T}$, \tilde{H} is nonincreasing in p_1 and nondecreasing in p_2 .
2. *Consistency* : for any $x \in \mathbb{T}$, $p \in \mathbb{R}$, $\tilde{H}(x, p, p) = H(x, p)$.
3. *Differentiability* : for any $x \in \mathbb{T}$, \tilde{H} is of class C^1 in p_1 and p_2 .
4. *Convexity* : for any $x \in \mathbb{T}$, $(p_1, p_2) \mapsto \tilde{H}(x, p_1, p_2)$ is convex.

The continuous Hamiltonian H will be approximated by \tilde{H} at (x_i, t_k) . Then, the discret HJB equation will be in the following form:

$$\begin{cases} -D_t U_{i,n}^h - \nu \Delta_h U_{i,n}^h + \tilde{H}(x_i, \nabla_h U_{i,n}^h) = f^h(x_i, M_{i,n+1}^h), \\ U_{i,N_t}^h = \phi^h(M_{i,N_t}^h), \end{cases}$$

where the operators f^h satisfies the following Assumption:

Assumption 1.2.2. 1. We assume that f^h is continuous on \mathcal{K}_h .

2. The numerical cost f^h is monotone in the following sense:

$$(f^h(m_k) - f^h(\bar{m}_k), m_k - \bar{m}_k)_2 \leq 0, \Rightarrow f^h(m_k) = f^h(\bar{m}_k). \quad (1.2.28)$$

3. There exists a constant C independent of h such that for all times t and for all grid function $m \in \mathcal{P}(\mathcal{K}_h)$,

$$\|f^h[m]\|_\infty \leq C, \quad \text{and} \quad |(f^h[m])_{i,j} - (f^h[m])_{k,l}| \leq C d_{\mathbb{T}}(x_{i,j}, x_{k,l}),$$

where $d_{\mathbb{T}}(x, y)$ is the distance between the two points x and y in the torus \mathbb{T}^2 .

4. For all sequence m_h , there exists a continuous and bounded function $\omega : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that $\omega(0) = 0$ and

$$\|f^h[m] - f^h[m_h]\|_{L^\infty(\mathcal{G}^h)} \leq \omega(\|m - m_h\|_{L^1(\mathbb{T}^2)}).$$

The main idea in [4] is to discretise also the Fokker-Planck equation as the HJB equation based on a weak formulation. Let us consider a smooth test function $\Phi \in C^\infty([0, T] \times \mathbb{T}^2)$ which involve the following expression

$$- \int_{\mathbb{T}} \partial_x (H_p(x, \partial_x u(t, x)) m(t, x)) \Phi(t, x) dx = \int_{\mathbb{T}} H_p(x, \partial_x u(t, x)) \partial_x \Phi(t, x) dx, \quad (1.2.29)$$

where they used an integration by parts and the periodic boundary conditions. In view of that precedes, it is quite natural to propose the following discrete version of the right hand side of (1.2.29)

$$h \sum_{i=0}^{N_h-1} M_{i,n+1}^h \left(\tilde{H}(x_i, \nabla_h U_{i,n}^h) \frac{\Phi_{i+1,n} - \Phi_{i,n}}{h} + \tilde{H}(x_i, \nabla_h U_{i,n}^h) \frac{\Phi_{i,n} - \Phi_{i-1,n}}{h} \right). \quad (1.2.30)$$

Performing a discrete integration by parts, we obtain

$$-h \sum_{i=0}^{N_h-1} \mathcal{T}_i(U_{i,n}^h, M_{i,n+1}^h) \Phi_{i,n},$$

where \mathcal{T}_i is a discrete transport operator

$$\begin{aligned} \mathcal{T}_i(U^h, M^h) = & \frac{1}{h} (M_{i,n}^h \tilde{H}_{p1}(x_i, \nabla_h U_{i,n}^h) - M_{i-1,n}^h \tilde{H}_{p1}(x_{i-1}, \nabla_h U_{i-1,n}^h) \\ & + M_{i+1,n}^h \tilde{H}_{p2}(x_{i+1}, \nabla_h U_{i+1,n}^h) - M_{i,n}^h \tilde{H}_{p2}(x_i, \nabla_h U_{i,n}^h)). \end{aligned}$$

Then, the Fokker-Planck equation can be approximated in a short way by:

$$\begin{cases} -D_t M_{i,n}^h - \nu \Delta_h M_{i,n+1}^h + \mathcal{T}_i(U_{i,n}^h, M_{i,n+1}^h) = 0, \\ M_{i,0}^h = \bar{m}_0(x_i). \end{cases}$$

Summarizing, the discrete problem is to look for two grid functions U and M on $G_{h,t}$:

$$\begin{cases} -D_t U_{i,n}^h - \nu \Delta_h U_{i,n}^h + \tilde{H}(x_i, \nabla_h U_{i,n}^h) = f^h(x_i, M_{i,n+1}^h), \\ -D_t M_{i,n}^h - \nu (\Delta_h M_{i,n+1}^h)_i + \mathcal{T}_i(U_{i,n}^h, M_{i,n+1}^h) = 0, \\ U_i^{N_t} = \phi^h(M_i^{N_t}), \\ M_{i,0}^h = \bar{m}_0(x_i). \end{cases} \quad (1.2.31)$$

The uniqueness for the discretized problems can be obtained by similar arguments as those used in the continuous case, using Lasry-Lions monotonicity condition

Theorem 1.2.3 (Achdou and Capuzzo-Dolcetta [4]).

We assume furthermore that H is convex w.r.t. to (q_1, q_2) and that the operator f^h and ϕ^h are strictly monotone (i.e) satisfies Lasry-Lions monotonicity condition:

$$\begin{aligned} (f^h(M_k) - f^h(\bar{M}_k), M_k - \bar{M}_k)_2 \leq 0, & \Rightarrow f^h(M_k) = f^h(\bar{M}_k), \\ (\phi^h(M_k) - \phi^h(\bar{M}_k), M_k - \bar{M}_k)_2 \leq 0, & \Rightarrow \phi^h(M_k) = \phi^h(\bar{M}_k). \end{aligned}$$

Then the problem 1.2.31 has a unique solution.

The convergence of this scheme was discussed by Achdou et al. in [3] where they consider a particular case. They considered an Hamiltonian has the following form $H(x, p) = \mathbf{H}(x) + |p|^\beta$ for $\beta > 1$ and \mathcal{H} is continuous function. The first result of convergence theorems [1, 4] make some assumption that the MFG system of PDEs has a unique classical solution and a strong versions of Lasry-Lions monotonicity assumptions. Under this assumptions, the solution of the discrete system converges towards the classical solution as the grid parameters tend to zero. In [5], Achdou and Porretta showed that the solutions of the discrete system converges to a weak solution of the forward-backward system. In [72], Lachapelle, Salomon and Turinici provided an iterative scheme using a discrete Markov decision problem. They proved that as the grid steps tend to zero, the solution of the discretized MFG system converges to a weak solution.

Recently, Chassagneux, Carmona, Delarue and the others [8] have introduced two algorithms for numerically solving FBSDEs of McKean Vlasov type, which can be used to formulate the solutions to mean field game problems. The first algorithm, which is based on the paper of Chassagneux, Crisan, and Delarue [39], relies on a tree structure to represent the pathwise law of the solution. The second algorithm takes the algorithm presented in the paper of Delarue and Menozzi [40] for solving FBSDEs and extends it to the mean field framework.

In recent work, Bayraktar, Budhiraja, and Cohen tebayraktar2017numerical used the Markov chain approximation method to construct an approximation for the MFG with reflecting barriers. They formulate the MFG model in terms of a controlled reflected diffusion with a cost function and study the convergence of their numerical scheme is established for a small time horizon T . The idea is that for a fixed measure on the path space, they define a Markov decision problem after a time and space discretization. The authors proved that the discrete numerical scheme is an almost contraction. Roughly speaking, the solution of the discrete MFG considered can be seen as the solution of a fixed point problem on the space of probability measures on certain path space. The Markov chain approximation is recently used by Ben Aziza and Toumi [13] to solve the system based on the construction of a discrete mean field game where the controlled state is a Markov chain approximating the controlled SDE. Our work can be seen as the generalization of [13]. For more results on the numerical methods for mean field games, we refer the reader to [2, 3, 8, 29, 30, 58].

1.2.3 Main results and contributions

Let us consider a mean field game model where any players controls its private state X_t at time t , by taking an action α_t in a closed convex subset \mathcal{A} , where the dynamic of state is driven by

the following the stochastic differential equation

$$\begin{cases} dX_t^\alpha &= b(t, X_t^\alpha, \alpha_t)dt + \sigma dB_t, \\ X_0^\alpha &\sim m_0, \end{cases} \quad (1.2.32)$$

where B_t is a standard Brownian motion, σ is a fixed positive constant and \mathcal{A} is the set of admissible control control.

In this work, we introduce a Markov chain approximation method. This approach was inspired from the works of Kushner et al. [70, 71]. It's based on the approximation of the stochastic control problem by a control problem for a discrete time and discrete space controlled Markov chain. An important advantage of this method is that the Markov chain approximation facilitates convergence proofs for the numerical methods in terms of probabilistic argument. The idea behind this approach is to construct a Markov chain ξ^h and its state space. This chain admits a continuous interpolation which approximates the state process X_t . Roughly speaking, for positive k , we define $(\xi_k^h)_k$ the Markov chain on the state space where his transition probabilities are defined by

$$p(\xi_k^h, \eta | \alpha_k) = \mathbb{P}(\xi_{k+1}^h = \eta | \xi_j^h, \alpha_j, j \leq k), \quad (1.2.33)$$

for transitions of the Markov chain from stage ξ_k^h to the stage ξ_{k+1}^h under control policy α_k . These transition probabilities must satisfy the non-negativity $p > 0$ and conservation $\sum p = 1$. The Markov chain transition probabilities $p(\xi_k^h, \eta | \alpha_k)$ are constructed from the finite difference approximation of HJB equation. Its law ν_n at time n is defined by $\nu_n := \nu_{n+1}p$. By iterating this equality, we obtain for any $n \geq 1$,

$$\nu_n = P^n \nu_0,$$

where ν_0 is the initial distribution. Roughly speaking, we start by approximate the HJB equation using difference finite method as follows:

$$-\left(\frac{U_{i,k+1} - U_{i,k}}{\Delta t}\right) - \frac{1}{2}\sigma^2\left(\frac{U_{i+1,k} - 2U_{i,k} + U_{i-1,k}}{h^2}\right) + H\left(x, \frac{U_{i+1,k} - U_{i-1,k}}{2h}\right) = F(M_{i,k}^h, x_i), \quad (1.2.34)$$

with

$$H\left(x_i, \frac{U_{i+1,k} - U_{i-1,k}}{2h}\right) = \inf_{\alpha} \{b(t_k, x_i, \alpha_{i,k}) \cdot \left(\frac{U_{i+1,k} - U_{i-1,k}}{2h}\right) + L(x_i, \alpha_{i,k})\}, \quad (1.2.35)$$

where we assume the following boundary conditions

$$U_{-1,k} = U_{0,k}, \quad U_{N_x^h+1,k} = U_{N_x^h,k}, \quad 1 \leq i \leq N_x^h. \quad (1.2.36)$$

Therefore, it leads to the following approximation of HJB equation

$$\begin{cases} U_{i,k-1} = \inf_{\alpha \in \mathcal{A}^h} \left[\left(L^h(x_i, \alpha_{i,k}) + F^h(x_i, M_{i,k}^h) \right) dt + U_{i,k} \left(1 - \sigma^2 \frac{\Delta t}{h^2} \right) \right. \\ \quad \left. + U_{i-1,k} \left(\frac{\sigma^2 \Delta t}{2 h^2} + \frac{\Delta t}{2h} b(x_i, \alpha_{i,k}) \right) + U_{i+1,k} \left(\frac{\sigma^2 \Delta t}{2 h^2} - \frac{\Delta t}{2h} b(x_i, \alpha_{i,k}) \right) \right], \\ U_{i,N_t^h} = \Phi^h(x_i, M_{i,N_t^h}^h), \quad 0 \leq i \leq N_x^h. \end{cases} \quad (1.2.37)$$

Then, from the finite differential approximation (1.2.37), we define a transition probabilities such that at each time t_k , $p^h(t_k, \nu_{i,k}, \alpha_{i,k}; k, k+1) = P^h(k, \alpha_{i,k})$ is a matrix of transition probabilities for the approximated Markov Chain defined as follows:

$$\begin{aligned} P_{i,i-1}^h(k, \alpha_{i,k}) &= \frac{\sigma^2 \Delta t}{2 (h)^2} + \frac{\Delta t}{2h} b(x_i, \alpha_{i,k}), \quad 1 \leq i \leq N_x^h - 1, \\ P_{i,i}^h(k, \alpha_{i,k}) &= 1 - \sigma^2 \frac{\Delta t}{h^2}, \quad 1 \leq i \leq N_x^h - 1, \\ P_{i,i+1}^h(k, \alpha_{i,k}) &= \frac{\sigma^2 \Delta t}{2 h^2} - \frac{\Delta t}{2h} b(x_i, \alpha_{i,k}), \quad 1 \leq i \leq N_x^h - 1. \end{aligned} \quad (1.2.38)$$

In this work, we define the law associated to discrete Markov chain problem M^h from its transition probabilities as follows

$$\begin{cases} M_{*,0}^h = \pi_0^h, \\ M_{*,k}^h = (P^h(k, \alpha_{*k}))^t M_{*,k-1}^h, \quad \text{for } 1 \leq k \leq N_t^h, \end{cases} \quad (1.2.39)$$

with $(P^h(k, \alpha_{*k}))^t$ is the transposed matrix of $(P^h(k, \alpha_{*k}))$. Then, we will show that it converges weakly to the solution of the Fokker Planck equation.

The first main result of this part is Theorem 1.2.4 given below, where we prove an existence and uniqueness results of the solution of our discrete version of mean field system (1.2.37)-(1.2.39):

Theorem 1.2.4. *Under some assumptions and if F and Φ satisfy the Monotonicity condition. Then there exist a unique solution of the system (1.2.37)-(1.2.39).*

Based on the results of Billingsley [19], our second main result in this part is the following convergence Theorem 1.2.5:

Theorem 1.2.5. *Let $(M^h)_h$ be the discrete MFG equilibrium and ξ^h being its associated optimally controlled Markov chain. Let $(m^h)_h \in \mathcal{P}(\mathcal{D}[0, T])$ be the law of the continuous-time-space piecewise constant interpolated process $\tilde{\xi}^h$.*

Then, the sequence $(m^h)_{h \geq 0}$ converges weakly to some \bar{m} in $\mathcal{P}(\mathcal{D}[0, T])$ where \bar{m} is an equilibrium of MFG problem.

To prove the convergence result, we need the results in Kushner and Dupuis [70] for weak

convergence of time-continuous controlled Markov chains on the space of càdlàg processes $D[0; T]$ which is a separable and complete for the Skorohod topology. We need to add some boundary condition, to force the state variable X to stay in the domain. In fact, with the local reflection direction, if the chain tries to leave the state space then it is returned immediately in space.

It will be also necessary to introduce a new class of control, the relaxed control which provide a very powerful tool in the study of the convergence properties of sequences of optimally controlled processes. The idea of relaxed control is to replace the \mathcal{A} -valued process α_t with a $P(\mathcal{A})$ -valued process α_t^R , where $P(\mathcal{A})$ is the space of probability measures equipped with the topology of weak convergence. According to the result of Kushner and Dupuis in [70] and for any $m \in P$, we get

$$\inf_{\alpha^{R,*}} J(\alpha^R, m^*) = \inf_{\alpha} J(\alpha, m^*).$$

The last part of this work is dedicated to a comparison between an almost explicit simulated solution and the numerical solution corresponding to the convergent discrete MFG solution for a linear model.

1.3 Learning dynamic utilities

The third part considers the task of learning dynamic preferences of an agent from observed sequence of his decisions at discrete dates (possibly random) $(\tau_i)_i$. There are many settings in which this capability is very useful. In cooperative setting, we want to help the agent make good decisions. This is the principle of robo-advisors which are online investment management services that employ mathematical algorithms to provide financial advice with minimal human intervention. Based on online questionnaires to obtain information about the client's degree of risk-aversion, financial status, and desired return on investment, the robo-advisors use different algorithms to manage and allocate client assets in the most possible efficient way according to his preferences. In a more competitive setting, we want to predict the agent's actions in order to optimize better our own payoffs, this is the case of two-players A and B, where A is trying to learn the utility of B, to optimize his/her own payoff which also depends on the actions and utility of B. This chapter based on co-work with Mrad [94].

1.3.1 Dynamic utilities

From a theoretical and application point of view this problem of recovering the utility of an agent from observations is an old inverse problem: Same questions were addressed by the economist Samuelson in the 40's [111, 112], with the theory of "revealed preference" where the observable

is the consumption process. Since then, the theory has grown in interest under the pressure of the economic reality, which created news incentives for different approaches, see Chambers & Echenique [36]. An example is the evolutionary economics by Arthur [10]. Economies are considered as complex evolutionary systems, where the agents try to predict the outcomes of their actions, and how the market would be modified by their decisions.

Beyond the fields of economics and finance, in the last few decades this forward-looking viewpoint has been also considered in many optimization problems as a consequence of the expansion of computer science and e-commerce. The forward modeling allows anticipations on the future values of observables. But, the agents also need to adjust their (random) preferences over time, following an "inverse thinking" approach as has been suggested by Gomez-Ramirez [55].

In the field of machine learning, we refer for example to Friedman and Sandow [49], Chajewska, Koller, and Ormoneit [35]: and Stahl [113] and Hibbard [62]. As well, reasoning with preferences has been recognized as a particularly promising research direction for artificial intelligence see Nielsen and Jensen. [101] and Qi, Xu, and Lafferty. [108]. For e-commerce, see [116]: The principal idea is to build the utility function of a user from his search history and purchases on the Internet. This utility will then be used to target the user by proposing products or baskets of goods that we know are likely to be of interest to him and at specific prices evaluated from this utility function.

Before presenting our learning approach that allow the robo-advisor to predict the agent's utility, we first introduce the concept of dynamic utilities and recall some of the main results in this theory which we will use extensively in this work.

In the classical framework, the utilities are chosen deterministic independently of the the investment universe (very dynamic). In fact, in this classical portfolio optimization program, the investor set at time $t = 0$ his utility U for maturity T , he then optimizes his expected utility over all admissible strategies and find the associated optimal policy and then the optimal portfolio. The problem is that this optimal process is strongly dependent on the maturity T and it is found independently of the market evolution and do not take into account the new investment constraints imposed on this agent. Actually, this is not consistent with the periods of crisis that have been quite frequent in recent years and that can significantly influence the choices and preferences of any agent with regard to products that become much too risky or more profitable. Moreover, this optimal portfolio is not time consistent: if the investor decides to continue investing until a new maturity $T' > T$, the two portfolios obtained for two different maturities do not usually coincide at T when they should (because they represent the same preferences). In

order to rectify this inconsistency, we have no choice: by the principle of dynamic programming, the utility must be a stochastic process to allow a flexibility to modify our strategy if it is necessary and to readjust the agent preferences over time.

In 2002, Musiela and Zariphopoulou [95] introduced the notion of "Forward utility" which are dynamic, progressive utilities, consistent with a given financial market if they satisfy some given conditions (see Definition 1.3.2 below) . A dynamic utility is a random field $U(t, z, \omega)$ adapted to the available information, and at each time t , it is a standard utility function. See [95, 97, 98] for more detailed discussions and motivations. A progressive utility gives an adaptative way to include new information on environment evolution available to economic agents. Since these utilities are stochastic, time dependent and moving forward, we consider them as a family of semi-martingales depending on a parameter, the wealth of the agent in the economic context.

A dynamic utility should represent, possibly changing over time, individual preferences of an agent starting with a today's specification of his utility, $U(0; z) = u(z)$. The preferences are affected over time by the available information represented by the filtration $(F_t)_{t \geq 0}$ defined on the probability space $(\Omega, \mathbb{P}, \mathcal{F})$. The filtered probability space $(\Omega, \mathbb{P}, (F_t)_{t \geq 0})$ is assumed to satisfy usual conditions of right continuity and completeness. The filtration \mathcal{F}_0 is not necessarily assumed to be trivial, so that the initial condition $U(0, z)$ is not necessarily a deterministic function.

We start by introducing the definition of a progressive utility as a progressive random field with concavity property:

Definition 1.3.1 (Progressive Utility). *A progressive utility is a càdlàg progressive random field on $\mathbb{U} = \{U(t, z); t \geq 0, z > 0\}$ such that,*

- UTILITY PROPERTY: \mathbb{U} is strictly concave, strictly increasing, and non-negative.
- REGULARITY PROPERTY: \mathbb{U} is a \mathcal{C}^2 -random field, with continuous first and second derivatives random fields \mathbb{U}_z and \mathbb{U}_{zz} .
- INADA CONDITIONS: \mathbb{U} goes to 0 when x goes to 0 and the derivative \mathbb{U}_z goes to ∞ when z goes to 0, and to 0 when z goes to ∞ .

The notion of progressive utility is very general and should be specified so as to represent more realistically the dynamic evolution of the individual preferences of an investor in a given financial market. Once his consistent progressive utility is defined, an investor can then turn to a portfolio optimization problem in a larger financial market or to calculate indifference prices.

Musiela and Zariphopoulou [96, 98] introduced the following definition of a consistent dynamic utility:

Definition 1.3.2 (Consistent dynamic utility). *A consistent dynamic utility $U = \{U(t, z); t \geq 0, z > 0\}$ is a progressive utility with the following additional properties:*

Consistency with the test-class: *For any admissible wealth process X ,*

$$\mathbb{E}(U(t, X_t)/\mathcal{F}_s) \leq U(s, X_s), \quad \forall s \leq t \text{ a.s.}$$

Existence of optimal wealth: *For any initial wealth $x > 0$, there exists an optimal wealth process X^* such that $X_0^* = x$, and for all $s \leq t$,*

$$U(s, X_s^*) = \mathbb{E}(U(t, X_t^*)/\mathcal{F}_s) \quad \forall s \leq t \text{ a.s.}$$

In short for any admissible wealth X , $U(\cdot, X)$ is a positive supermartingale and a martingale for the optimal-benchmark wealth X^ .*

Motivated by Zariphopoulou et al [95], El Karoui and Mrad [45] study and fully characterize the consistent dynamic utility in an Itô incomplete market. The key idea of [45] is to study in details the marginal utility $U_x(t, x)$ instead of $U(t, x)$ and to consider at the same time the dual problem. They establish the fully non-linear stochastic partial differential equation (SPDE) satisfied by the marginal utility and link it with the SDEs that satisfy the optimal primal and dual processes. Under some regularity assumption, the authors give sufficient conditions ensuring concavity, monotonicity, differentiability both for random field and for solution of SDE. They establish a dynamics of consistent utilities and their convex conjugate and a fully characterization of the marginal utility. In 2020 [46], the authors study the problem of revealed utility from observed data and then give a consistent way to generate the utility of the agent from the informations given by observing his choices over time. The result of this last paper are established in a very general framework (no regularity assumptions on time and without reference to any market) which makes it applicable to a continuous or discrete framework for vast and varied field like e-commerce, Robo-Advisor, artificial intelligence...

In 2021, Matoussi and Mrad [89] study consistent dynamic utilities in an incomplete Itô's market driven by a multidimensional brownian motion and a general Lévy measure.

Let us briefly present the idea and different steps of the revealed dynamic utility problem as in [46]. By definition, the observable is a so-called dynamic positive characteristic process $\{X_t(x)\}$ considered for different values of its initial condition $X_0(x) = x > 0$, and assumed to be increasing in x (to be coherent with the expected utility criterium). The goal is to recover,

from a given initial utility function u , a stochastic dynamic utility $\mathbf{U} = \{U(t, z, \omega), z > 0\}$, "revealed optimally" in the sense that at any (stopping) time τ , the preference for the observable process is in mean equal to its value at time 0, $\mathbb{E}(U(\tau, X_\tau(x))) = u(x)$: from the probabilistic dynamic view point, on a given filtered probability space, the performance process " $U(t, X_t(x))$ is a martingale". Focusing on the concavity of the utility criterium \mathbf{U} , tools of convex analysis play a key role, especially the invertible decreasing marginal utility $U_z(t, z)$. Its allows to define the convex Fenchel-Legendre transform $\tilde{U}(t, y)$ of $U(t, x)$, by $U(t, z) - zU_z(t, z) = \tilde{U}(t, U_z(t, z))$, but also to linearize the recovery problem, by using the one to one correspondence between the class of dynamic utilities \mathbf{U} (revealed by X) and the adjoint processes Y candidate to play the role of $U_z(t, X_t)$, more precisely $\{Y_t(u_z(x)) = U_z(t, X_t(x))\}$. Since the characteristic process is invertible, the triplet $\{(u, X_t(x), U(t, z))\}$ is in one to one correspondence with the triplet $\{(u, X_t(x), Y_t(u_z(x)))\}$. In fact, denoting by $\{\mathcal{X}_t(z) := (X_t)^{-1}(z)\}$ the inverse flow of X , the authors have the characterization

$$U_z(t, z) = Y_t(u_z(\mathcal{X}_t(z))), \quad \text{and} \quad U(t, z) = U(t, z_0) + \int_{z_0}^z Y_t(u_z(\mathcal{X}_t(x)))dz.$$

Note that in the continuous semimartingale [45] and semimartingale with jumps framework [89], the authors have shown that if U is a regular dynamic utility, consistent with a given financial market, it is necessary in this form with X being the optimal portfolio and Y being the optimal state price density process.

To illustrate some of the main results of dynamic utilities [89], let us consider a filtered probability space $(\Omega, \mathbb{F}, \mathbb{P})$ supports a q -dimensional standard Brownian motion $B = (B^1, \dots, B^q)$ on and an independent q' -dimensional Poisson random measure π on $[0, \infty[\times \mathbb{R}^{q'}$ with constant time dependent intensity measure $\lambda dt \times \nu(de)$ defined on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. λ is the intensity of jumps and ν is a finite positive measure on $\mathbb{R}^{q'}$, satisfying standard integrability assumptions. We also denote by $\tilde{\pi}$ the compensated version of π :

$$\tilde{\pi}(dt, de) = \pi(dt, de) - \nu(de)\lambda(t)dt.$$

Let us consider two \mathbb{R}^d -valued stochastic processes X and Y , solutions of the following stochastic differential equations (SDE for short)

$$dX_t(x) = \mu(t, X_t(x))dt + \sum_{i=1}^q \sigma_i(t, X_t(x))dB_t^i + \int_{\mathbb{R}^{q'}} h(t, X_{t-}(y), e)\tilde{\pi}(dt, de), \quad X_0(x) = x, \quad (1.3.40)$$

$$dY_t(y) = b(t, Y_t(y))dt + \sum_{i=1}^q \gamma_i(t, Y_t(y))dB_t^i + \int_{\mathbb{R}^{q'}} g(t, Y_{t-}(y), e)\tilde{\pi}(dt, de), \quad Y_0(y) = y, \quad (1.3.41)$$

where $\mu, b, \sigma_i, \gamma_i$ are deterministic functions from $\mathbb{R}_+ \times \mathbb{R}^d$ into \mathbb{R}^d and h, g are deterministic functions from $\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^{q'}$ into $\mathbb{R}^{d \times q'}$, globally Lipschitz in space to ensure the existence of a unique strong solution under the following assumptions :

Assumption 1.3.1. • **(HP1)** *There exist C^Y and $C^Y(e)$ s.t. for any $t \in [0, T]$, $x, y \in \mathbb{R}^d$ and $e \in \mathbb{R}^{q'}$*

$$\begin{cases} |b(t, x) - b(t, y)| \leq C^Y |x - y|, & |\mu(t, 0)| \leq C^Y, \\ |\gamma(t, x) - \gamma(t, y)| \leq C^Y |x - y|, & |\gamma(t, 0)| \leq C^Y, \\ |g(t, x, e) - g(t, y, e)| \leq C^Y(e) |x - y|, & |g(t, 0, e)| \leq C^Y(e), \end{cases} \quad (1.3.42)$$

where $C^Y(e)$ satisfies $\int_{\mathbb{R}^{q'}} [C^Y(e)]^p \nu(de) < \infty, \forall p \geq 2$.

- **(HP2)** *b, γ and g are continuously space-differentiable functions s.t. $\nabla_y b, \nabla_y \gamma$ and $\nabla_y g$ satisfy **(HP1)** with $(C^{Y, \nabla}, C^{Y, \nabla}(e))$ in place of $(C^Y, C^Y(e))$.*
- **(HP3)** *b, γ and g are α^Y -Hölder continuous in time, locally in space,*

$$\begin{cases} |b(t, x) - b(s, x)| + |\gamma(t, x) - \gamma(s, x)| \leq C^Y (1 + |x|) |t - s|^{\alpha^Y}. \\ |g(t, x, e) - g(s, x, e)| \leq C^Y(e) (1 + |x|) |t - s|^{\alpha^Y}. \end{cases} \quad (1.3.43)$$

- **(HP4)** *$\nabla_y b, \nabla_y \gamma$ and $\nabla_y g$ satisfy **(HP3)** with $(C^{Y, \nabla}, C^{Y, \nabla}(e))$ in place of $(C^Y, C^Y(e))$.*

Assumption 1.3.2. **(HP1)** and **(HP3)** are satisfied for μ, σ and h (instead of b, γ and g) with coefficient $C^X, C^X(e)$ and α^X (instead of $C^Y, C^Y(e)$ and α^Y).

The novelty of [89] is that the authors have succeeded in linking the solutions of these SPDE to those of two SDEs. They have shown that the marginal utility is represented as the compound of the solutions of two SDEs. This explains the need to study the composition of the approximations of the two SDEs instead of trying to discretise the SPDEs directly, which is not an easy task. To be clearer, the dynamic utility of an investor is a random map $(t, z, \omega) \rightarrow U(t, z, \omega)$, depending on time t and wealth. Let us assume that it is of type Itô driven by a Lévy process,

$$dU(t, z) = \beta(t, z)dt + \gamma(t, z)dB_t + \int_{\mathbb{R}} H(t, z, e)\tilde{\pi}(dt, de),$$

where B is a d -dimensional Brownian motion and $\tilde{\pi}$ is the compensated version of independent Poisson random measure π on $[0, \infty] \times \mathbb{R}$ with intensity measure $\lambda(t)dt \times \nu(de)$. The triplet (β, γ, H) denotes the local characteristics of U with values in $\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}$. Denotes by U_z and U_{zz} the first and the second derivative of U with respect to z . Then, the dynamic utility U , associated to the investor's portfolio optimization, is consistent if it solves the second-order fully

nonlinear SPDE driven by Lévy noise and of HJB type

$$\begin{aligned} dU(t, z) &= \left(-zU_z(t, z)r_t + \int_{\mathbb{R}} (U(t, z) + H(t, z, e))\nu(de) - \mathcal{Q}(t, z, \kappa^*) \right) dt \\ &+ \gamma(t, z)dW_t + \int_{\mathbb{R}} H(t, z, e)\tilde{\pi}(dt, de), \end{aligned} \quad (1.3.44)$$

where

$$\begin{aligned} \mathcal{Q}(t, z, \kappa^*) &:= \int_{\mathbb{R}} (U + H)(t, z(1 + \kappa_t^* h^S(t, e)))\nu(de) \\ &+ \frac{1}{2}U_{zz}(t, z)\|z\kappa_t^* + \frac{\gamma_z(t, z) + U_z(t, z)(\eta_t - \alpha_t)}{U_{zz}(t, z)}\|^2 \\ &+ \frac{1}{2}U_{zz}(t, z)\|\frac{\gamma_z(t, z) + U_z(t, z)(\eta_t - \alpha_t)}{U_{zz}(t, z)}\|^2. \end{aligned}$$

With r and η denotes the interest rate and the risk premium of the market. And $z\kappa_t^*(z)$ denotes the strategy of the investor defined as

$$z\kappa_t^*(z) = -\frac{\gamma_z^{\mathbb{R}}(t, z) + U_z(t, z)(\eta_t - \alpha_t)}{U_{zz}(t, z)} - \frac{\int_{\mathbb{R}} (U_z + H_z)(t, z(1 + \kappa_t^*(z)h^S(t, e)))h^S(t, e)\nu(de)}{U_{zz}(t, z)}. \quad (1.3.45)$$

In this work, we wish to apply these results [46, 89] in the field of machine learning for a financial model with defects, in other word a financial model with jumps. Therefore, we will dedicate the following section to introduce the main result of Matoussi and Mrad [89]:

Let us consider a stochastic process depending on a spatial parameter x . This parameter x , motivated by economic considerations, represents the initial wealth of an investor, taking no negative values in \mathbb{R}_+ . It can be viewed as a progressive random fields, in other word, it is a random variable measurable w.r.t. $\mathcal{F}_\infty \otimes \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R}_+^0)$, which is a collection of progressive processes $t \mapsto X(t, x)$. This approach is mainly based on stochastic flows technics. The idea in [89], is to established the forward dynamics of the inverse flow \mathcal{X} of a regular semi-martingale X , monotone with respect to its initial condition. In [48], Authors study the regularity of semi-martingales random field from the regularity of its local characteristics (ν, σ, h) and conversely ([48, Theorem 1.3]). They establish in particular the backward dynamic of the inverse flow \mathcal{X} of a regular semimartingale X , monotonic with respect to its initial condition. In Matoussi and Mrad [89] under some sufficient conditions, the authors are concerned by the forward dynamic of the inverse flow of an Itô random field. This is a key result to establish the main results of the paper and particularly to show that the marginal utility is a compound of two stochastic flow satisfying a SDEs with jumps. So as in El Karoui and Mrad [46], the regularity of the coefficients of SDEs is a fundamental.

Therefore, we start by introducing a certain notion of differentiable random fields called $\mathbb{K}_{loc}^{m,\delta}$ -regular as follows:

- (i) X is said to be $\mathbb{K}_{loc}^{m,\delta}$ -regular (resp. $\overline{\mathbb{K}}_{loc}^{m,\delta}$ -regular) if X is $\mathcal{C}^{m,\delta}$ -random fields such that for any compact $K \subset]0, +\infty[$, and any T ,

$$\int_0^T \|\phi\|_{m,\delta;K}(t, \omega) < \infty, \quad (\text{resp. } \int_0^T \|\psi\|_{m,\delta;K}^2(t, \omega) dt < \infty).$$

- (ii) X is said to be $\widehat{\mathbb{K}}_{loc,\nu}^{m,\delta}$ -regular if X is $\mathcal{C}^{m,\delta}$ -random fields such that for any compact $K \subset]0, +\infty[$, and any T ,

$$\int_0^T \int_{\mathbb{R}} \|\psi\|_{m,\delta;K}^2(t, e, \omega) \nu(de) dt < \infty.$$

- (iii) We use the notations $\mathbb{K}_b^m, \overline{\mathbb{K}}_b^m, \mathbb{K}_b^{m,\delta}, \overline{\mathbb{K}}_b^{m,\delta}, \widehat{\mathbb{K}}_{b,\nu}^m$ and $\widehat{\mathbb{K}}_{b,\nu}^{m,\delta}$, when these different norms are well-defined on the whole space $]0, +\infty[$ where the derivatives (up to a certain order) are bounded in the spatial parameter, with integrable (resp. square integrable) in time random bound.

As we have previously mentioned, the key point in [45, 46, 89] is to show that the marginal utility as the compound maps of Y and the inverse flow \mathcal{X} . So we need to find the forward dynamics of the inverse flow \mathcal{X} of a regular semi-martingale X which is monotonic with respect to its initial condition.

Theorem 1.3.3. (Matoussi and Mrad [89])

Let X be a solution of the following SDE(μ, σ, h),

$$dX_t(x) = \mu(t, X_t(x))dt + \sigma(t, X_t(x))dW_t + \int_{\mathbb{R}} h(t, X_{t-}(x), e)\tilde{\pi}(dt, de). \quad (1.3.46)$$

Let $\delta \in [0, 1[$ and assume $(\mu, \sigma, h) \in \mathbb{K}_b^{2,\delta} \times \overline{\mathbb{K}}_b^{2,\delta} \times \widehat{\mathbb{K}}_{b,\nu}^{2,\delta}$.

- (i) Then, the SDE(μ, σ, h) admits a unique strong solution $X(z)$, starting from z at time $t = 0$.
- (ii) If in addition, the maps $z \mapsto z + h(t, z, e)$ are homeomorphic, the map $z \rightarrow X_t(z)$ is strictly increasing with inverse flow $\mathcal{X}(t, x)$ satisfying the following second order SPDE,

$$\begin{aligned} d\mathcal{X}(t, x) &= \left[\mathcal{X}_x(t, x) \left(-\mu(t, x) + \int_{\mathbb{R}} h(t, x, e)\nu(de) \right) + \frac{1}{2} \partial_x \left(\mathcal{X}_x(t, x) \|\sigma(t, x)\|^2 \right) \right. \\ &\quad \left. - \mathcal{X}(t, x) + \int_{\mathbb{R}} \psi(t, x, e)\nu(de) \right] dt - \mathcal{X}_x(t, x)\sigma(t, x)dW_t \\ &\quad + \int_{\mathbb{R}} \left(-\mathcal{X}(t^-, x) + \mathcal{X}(t^-, \psi(t, x, e)) \right) \tilde{\pi}(dt, de), \end{aligned} \quad (1.3.47)$$

where $\psi(t, z, e)$ is the inverse map of $z \mapsto z + h(t, z, e)$.

Remarks 1. For the proof in [89], the authors use an extension of Itô-Ventzel's formula which was established by Oksendal and Zhang [102].

As mentioned above, see details in [45, 46, 89], we study the marginal utility along the optimal portfolio $U_z(t, X_t)$ and show that $U_z(t, X_t)$ is a price density process solution of some specific SDE (the optimal SDE of the dual problem satisfied by the optimum Y), regularity on the coefficients of this SDE show the existence and uniqueness and then leads to the main result: $U_z(t, z) = Y \circ U_z(\mathcal{X}_t(z))$.

Using the above theorem and the Itô-Ventzel's formula, Matoussi and Mrad [89] introduce the following results in order to gives the dynamics of the compound map $Y \circ \mathcal{X}$,

Corollary 1.3.4 (Matoussi and Mrad [89]). Let Y be a $\mathbb{K}_{loc}^{2,\delta}$ -semimartingale and X a $\mathbb{K}_{loc}^{3,\delta}$ -semimartingale for some $\delta \in]0, 1[$, satisfying the dynamics

$$\begin{cases} dX_t(x) = \mu(t, X_t(x))dt + \sigma(t, X_t(x))dW_t + \int_{\mathbb{R}} h(t, X_{t-}(x), e)\tilde{\pi}(dt, de), \\ dY_t(y) = b(t, Y_t(y))dt + \gamma(t, Y_t(y))dB_t^i + \int_{\mathbb{R}^{q'}} g(t, Y_{t-}(y), e)\tilde{\pi}(dt, de). \end{cases} \quad (1.3.48)$$

Assume X to be monotonic with respect to its initial condition and let \mathcal{X} denotes its inverse flow. Then the compound map $\Phi(t, x) := Y_t(\mathcal{X}_t(x))$ satisfies the following second order SPDE

$$\begin{aligned} d\Phi(t, x) = & \left(\mu^Y(t, \Phi(t, x)) - \Phi_x(t, x)[\mu(t, x) + \sigma_y^Y(t, \Phi(t, x))\sigma(t, x)] \right. \\ & \left. + \frac{1}{2}\partial_x \left(\Phi_x(t, x)\|\sigma(t, x)\|^2 \right) \right) dt \\ & + \int_{\mathbb{R}} \left(H(t, x, e) - h^Y(t, \Phi(t, x), e) + \Phi_x(t, x)h(t, x, e) \right) \nu(de) dt \\ & + (\sigma^Y(t, \Phi(t, x)) - \Phi_x(t, x)\sigma(t, x))dW_t + \int_{\mathbb{R}} H(t, x, e)\tilde{\pi}(dt, de), \end{aligned}$$

with $H(t, x, e) := \Phi(t^-, \psi(t, x, e)) + h^Y(t, \Phi(t^-, \psi(t, x, e)), e) - \Phi(t^-, x)$.

According to [46, 89], the marginal utility is written as $U_z = Y \circ u_z(\mathcal{X}_t(z))$, under Monotonicity assumptions and the following necessary and sufficient condition:

Proposition 1.3.5. In the particular case where $U(t, z) = zY_t$ with

$$dY_t = \mu^Y(t, Y_t)dt + \sigma^Y(t, Y_t)dW_t + \int_{\mathbb{R}} h^Y(t, Y_{t-}, e)\tilde{\pi}(dt, de). \quad (1.3.49)$$

The product XY follows the dynamics

$$\begin{aligned}
d(X_t Y_t) &= \left(X_t \mu^Y(t, Y_t) + Y_t \mu^X(t, X_t) + \sigma^X(t, X_t) \sigma^Y(t, Y_t) \right) dt + \left(X_t \sigma^Y(t, Y_t) + Y_t \sigma^X(t, X_t) \right) dW_t \\
&+ \int_{\mathbb{R}} h^Y(t, Y_t, e) h^X(t, X_t, e) \nu(de) dt \\
&+ \int_{\mathbb{R}} \left(X_t h^Y(t, Y_{t-}, e) + Y_t h^X(t, X_{t-}, e) + h^Y(t, Y_{t-}, e) h^X(t, X_{t-}, e) \right) \tilde{\pi}(dt, de).
\end{aligned} \tag{1.3.50}$$

Therefore, the process $(X_t Y_t)_t$ is a local martingale if and only if

$$X_t \mu^Y(t, Y_t) + Y_t \mu^X(t, X_t) + \sigma^X(t, X_t) \sigma^Y(t, Y_t) + \int_{\mathbb{R}} h^Y(t, Y_t, e) h^X(t, X_t, e) \nu(de) = 0, \quad dt \otimes \mathbb{P}.$$

Note that this result also characterizes a class of solutions of the SPIDE (1.3.44), making a connection with those of two SDEs (1.3.40) and (1.3.41). From a numerical point of view we can then propose a simple scheme to solve the SPIDE, by combining two Euler schemes to solve the optimal SDEs, as established in [51, 93].

On the field of forward utilities, the agent needs to adjust his preferences over time, then the idea is to use the Forward approach to anticipate the future value observable and apply the inverse problem using this prediction to deduct the value of the parameters that characterize the system and this leads us to a family of forward model solutions consistent with the data and not with a prediction.

As mentioned above, this forward thinking is suitable for the study of revealed preference problems in vast and varied field like E-commerce, Robo-Advisor, artificial intelligence. It is a learning approach based on the observation of the behavior of an agent. Roughly speaking, the objective is always to learn about the utility of an agent (player, investor, e-commerce customer) by observing its behavior in front of the important change of its fundamental parameters. So this work is a direct application of ideas developed in El Karoui and Mrad [45] to the framework of robo-advisors.

1.3.2 Main results and contributions

Motivation Automated investment managers or robo-advisors, were imposed these last years in the financial and economic world to replace or alternate with traditional human advisors. The robo-advisors are online platforms that provide financial advice or portfolio management to an investor or client. The performance of the robo-advisor strongly depends on its ability to accurately assess the investor's risk tolerance. They have many advantages over the investment services offered by traditional management consultants. It knows the investor better than a

traditional asset manager. Because of this better knowledge, the robo-advisor may propose a more appropriate asset allocation. In addition, robo-advisor performs the task in a systematic way and implements an automated rebalancing process. From the client's perspective, the biggest advantage of robo-Advisors is that it provides a low-cost alternative to traditional investing.

The idea of this platforms is to develop a new investment framework in which the robo-advisor not only manages the client's portfolio, but also reacts regularly with him to obtain updated information on his performance over time. More precisely, the client communicates his preferences to the robo-advisor only at specific update dates $(\tau_i)_i$. So between two update times, the robo-advisor does not receive any information from the customer. Therefore, it develops a framework to learn the investor's preferences between two update times. In other words, by observing the data communicated by the client at interaction dates, robo-advisor approximates the preferences of the client to make investment decisions and it interacts repeatedly with the client to update its information by asking the investor his new data to avoid making decisions based on stale information. In general, the robo-advisor solicits this information through targeted online questionnaires, and transfers it to digital data.

We assume that we have observations for several different initial conditions $(z_j)_j$. Such informations are not necessarily collected in the questionnaires of the Robo-Advisors, but maybe it is necessary because we know very well that the choices of an agent are strongly conditioned by its initial position (its initial wealth). Moreover, without these observations, we have a little chance to find the dynamic utility of the agent, see the papers [44–46, 89].

Contribution In this work, we are interested in learning the dynamic utility. In other words, even the robo-advisor builds the client's utility U at specific update dates τ_i (where is an increasing sequence of jumps times), it wants to approximate it all the time t . Therefore, we set the following steps to achieve our goal:

1. Simulate the inverse flows \mathcal{X} of the wealth process defined as a pure jumps process.
2. Build the price density process Y .
3. Build the dynamics of marginal utility $U_z(t, z) = Y_t(u_z(\mathcal{X}_t(z)))$.
4. Build the client's utility $U(\tau_i, \cdot)$ at τ_i update times.
5. Approximate utility $U(t, \cdot)$ at an intermediate date $t \in]\tau_i, \tau_{i+1}[$ to be able to make decisions for him.

In this work, we have two main parts, the first to simulate the utility of the client based on the results of the Mrad and its co-authors [46, 89, 93] and the second one is dedicated to the utility learning between two update times. Therefore, in the following we propose algorithms to approximate the inverse flow of X and Y in order to construct the marginal utility.

Simulation of inverse flow of (1.3.40) as a backward in time SDE :

In 2018, Gobet and Mrad [51] have discussed different approaches to the inverse flow of SDE solution in the continuous framework (see Section 4 in [51]). Among these approaches, we are interested in algorithm of inverse flow as a backward in time SDE which it has recently been extended by Mrad [93] to the discontinuous case (see section 4 [93]). In the following we introduce some scheme to approximate the inverse flow of X .

Let us consider $X_{s,t}(x)$ the solution, starting from x at time s , of the SDE with coefficients (μ, σ, h) . Denote by $\mathcal{X}_{s,t}$ the inverse map of $Y_{s,t}$ which can also be defined by $\mathcal{X}_{s,t}(x) := \mathcal{X}_t(X_s(x))$. This approach asks to consider the dynamics of $\mathcal{X}_{s,t}(x)$ in the variable s : doing so, we aim at computing the inverse of Y backward in time instead of forward in time.

Simulating the inverse flow as a solution of a stochastic differential equation backward in time was employed in [51] in a continuous case. With this result in hand, the approximation of $\mathcal{X}_{s,t}$ is made possible simply using a standard Euler scheme like for Y^n . Using similar notations as below:

Scheme 1.3.1. *Based on this key result, the approximation of the inverse flow \mathcal{X} is achieved by the following steps.*

- Set $\mathcal{X}_{t,t}^N(z) = z$. If $t = \theta_k$ for some $k \in \llbracket 1, J \rrbracket$, set $\mathcal{X}_{t^-,t}^N(z) = z - k(t, z, E_k)$ else $\mathcal{X}_{t^-,t}^N(z) = z$.

- For $s \in]\tau_t, t]$, set

$$\begin{aligned} \mathcal{X}_{s,t}^N(z) &= \mathcal{X}_{t^-,t}^N - \sigma(t, \mathcal{X}_{t^-,t}^N(z)) \cdot (B_t - B_s) \\ &\quad - [\mu(t, \mathcal{X}_{t^-,t}^N(z)) - \partial_z \sigma(t, \mathcal{X}_{t^-,t}^N(z)) \cdot \sigma(t, \mathcal{X}_{t^-,t}^N(z)) - \int_{\mathbb{R}} \mathcal{X}(t, \mathcal{X}_{t^-,t}^N(z), e) \nu(de)](t - s). \end{aligned} \quad (1.3.51)$$

- For $l \in \llbracket 0, N + J \rrbracket$ satisfying $t_l < \tau_t$,

- if $t_l = \theta_k$ for some $k \in \llbracket 1, J \rrbracket$, set $\mathcal{X}_{t_l,t}^N(z) = \mathcal{X}_{t_l,t}^N(z) - k(t_l, \mathcal{X}_{t_l,t}^N(z), E_k)$ else $\mathcal{X}_{t_l,t}^N(z) = \mathcal{X}_{t_l,t}^N(z)$

- and for $s \in [t_{(l-1)}, t_l[$

$$\begin{aligned} \mathcal{X}_{s,t}^N(z) &= \mathcal{X}_{t_l^-,t}^N(z) - \sigma(t_l, \mathcal{X}_{t_l^-,t}^N(z)) \cdot (B_{t_l} - B_s) \\ &\quad - [\sigma(t_l, \mathcal{X}_{t_l^-,t}^N(z)) - \partial_z \sigma(t_l, \mathcal{X}_{t_l^-,t}^N(z)) \cdot \sigma(t_l, \mathcal{X}_{t_l^-,t}^N(z)) - \int_{\mathbb{R}} h(t_l, \mathcal{X}_{t_l^-,t}^N(z), e) \nu(de)](t_l - s). \end{aligned} \quad (1.3.52)$$

Simulation of the price density process (1.3.41) using Euler schema

Let us consider a sequence e_1, e_2, \dots of independent random variables with common exponential distribution with parameter 1. We define

$$\Lambda(t) = \int_0^t \lambda(s) ds, \quad t \in [0, T].$$

To simplify, we take $\lambda(t) = 1$. The number of jumps of the random Poisson measure $N(dt, de)$ in an interval $[0, t]$ is determined as

$$J(t) = \max\{k : \sum_{j=1}^k e_j \leq \Lambda(t)\}.$$

We denote by $J = J(T)$ the total number of jumps in $[0, T]$ and The jump times $(\tau_k)_k$ of the Poisson measure defined by

$$\tau_k = \Lambda^{-1}\left(\sum_{j=1}^k e_j\right), \quad k \in \{1, \dots, J\},$$

where Λ^{-1} is the right continuous inverse of Λ . The jump times can be computed recursively by

$$e_k = \int_{\tau_{k-1}}^{\tau_k} \lambda(s) ds, \quad k \in \{1, \dots, J\}.$$

Once the jump times are computed, we proceed to sample the marks $\{E_k\}$, that, conditionally on the values of the jumps times, are independent random variables distributed respectively according to $\{\nu(de)\}$. The random measure with intensity $\lambda(t) dt \ddot{\nu}(de)$ can then be constructed as

$$\pi(dt, de) = \sum_{k=1}^J \delta_{(\tau_k, E_k)}(dt, de).$$

Consequently, the stochastic intergal with respect to the Poisson random measure can be defined as

$$\int_0^t \int_{\mathbb{R}^d} h(s, X_{s-}, e) \pi(dt, de) = \sum_{k=1}^{J(t)} h(\tau_k, X_{\tau_k-}, E_k), \quad t \in [0, T].$$

Let $N \geq 1$ and let us consider the discretization family $\{\bar{t}_i := i\frac{T}{N}, i \in \llbracket 0, N \rrbracket\}$ of $[0, T]$. We consider also the jump times $\{\tau_k, k \in \llbracket 1, J \rrbracket\}$ with corresponding marks $\{E_k, k \in \llbracket 1, J \rrbracket\}$. Then, we consider the argued partition given by the union

$$\mathcal{P}_N = \{t_l, l \in \llbracket 0, N + J \rrbracket\} = \{\bar{t}_i := i\frac{T}{N}, i \in \llbracket 0, N \rrbracket\} \cup \{\tau_k, k \in \llbracket 1, J \rrbracket\}.$$

Thus, using Euler scheme, the discretization of Y (1.3.41) is given by the following steps

Scheme 1.3.2. *The Euler scheme to approximate the solution Y of the SDE (1.3.41) is given by the following steps*

- Set $Y_0^N(x) = x$.
- For $k = 0, \dots, N + J - 1$, if $t \in (t_k, t_{k+1}]$, then

$$Y_{t^-}^N(x) = Y_{t_k}^N(x) + b(t_k, Y_{t_k}^N(x))(t - t_k) + \gamma(t_k, Y_{t_k}^N(x))(B_t - B_{t_k}).$$

- If $t_{k+1} = \tau_l$ for some $l \in \llbracket 1, J \rrbracket$, then we introduce a correction due to jumps discontinuities

$$Y_{t_{k+1}}^N(x) = Y_{t_k}^N(x) + g(t_{k+1}, Y_{t_{k+1}}^N(x), E_l).$$

Simulation of dynamic utility U^N : According to Proposition 1.3.5, we must ensure that the product XY remains a local martingale for the building of process X and Y . To ensure this constraint, we must take care that the parameters of the jumps V and W (which help us to build Y and X) are independent.

Once we have simulated X and Y , we can construct the marginal utility U_z as follows

$$U^N(t, z) = \sum_{i=1}^n Y_t^N(u_z(\mathcal{X}_t^N(z_i)))(z_{i+1} - z_i)$$

which converges to

$$U^N(t, z) = \int_{z_0}^z Y_t^N(u_z(\mathcal{X}_t^N(x)))dx. \quad (1.3.53)$$

Remark 1.3.6. *It should be mentioned that the approximated utility must be concave with respect to its wealth z .*

This approach is based on the work of Mrad and his co-authors [45, 46, 89], it allows us to estimate the agent's utility at dates and interaction points. This is a very important step as until now there was no idea how to quantify the preferences of an agent from these choices. The goal then, starting from the values obtained $(U(t_i, z_j))_{i,j}$, is to estimate $U(t, z)$ for all (t, z) and

to predict its future decisions. For this we will use different machine learning algorithms and we will compare them.

Learning utilities method

In the second part, we propose different methods to learn client's utilities. Let's assume that the client has enough sufficient interaction with the machine. Can we approximate/learn his utility until the next interaction date so that the machine can make decisions for him while remaining consistent with his preferences and past decisions ?

The aim in the following is to propose an algorithm for learning the preferences of the agent. As we do not have an access to a real data, we will simulate our own data, the client's utility from the observations $(X)_{i,j}$ and $(Y)_{i,j}$ for $(i, j) \in \llbracket 1, J \rrbracket \times \llbracket 1, n \rrbracket$ using the above procedure. The generated values $U(\tau_i, \mathfrak{z}_j)$ will play the role of starting data for our learning algorithms. This kind of algorithm could clearly be used on robo-advisor platforms, since it can "predict" (over a short period of time) the preferences of an agent and therefore can make decisions for him/her.

Machine Learning algorithms automatically learn to perform a task or make predictions from data and improve their performance over time. Once trained, the algorithm will be able to find patterns in new data. Depending on the information available during the learning phase, learning is qualified in different ways. If the data are labeled (i.e., the response to the task is known for these data), it is supervised learning. It is called classification if the labels are discrete, or regression if they are continuous. If the model is learned in an iterative way, according to a reward received by the program for each action taken, it is called reinforcement learning. In the most general case, without labels, we try to determine the underlying structure of the data (which can be a probability density) and it is then called unsupervised learning.

Here, we are interested in supervised learning. Our task is to predict a target value which is the value of utility at fixed time τ_i , denoted by $(\mathcal{U}_j^i)_{j \in \llbracket 1, n \rrbracket}$ a vector of the matrix $(\mathbf{U}(\tau_i, \mathfrak{z}_j))$, for $(i, j) \in \llbracket 1, N \rrbracket \times \llbracket 1, n \rrbracket$ from different initial wealth values $(\mathfrak{z}_j)_{j \in \llbracket 1, n \rrbracket}$ and fixed value τ_i :

$$\mathcal{U}_j^i := (\mathbf{U}(\tau_i, \mathfrak{z}_j))_{j \in \llbracket 1, n \rrbracket}.$$

In Machine Learning, all starts from a Dataset which contains our data. In supervised learning, the Dataset contains the questions $(Z_j)_{j \in \llbracket 1, n \rrbracket}$ and answers $(\mathcal{U}_j^i)_{j \in \llbracket 1, n \rrbracket}$ to the problem that the machine must solve.

In this work, we have chosen to use **Scikit-learn library** of **Python** contains all the math-

emathical functions which are necessary for learning problems. **Scikit-learn** is an open source machine learning library that supports supervised and unsupervised learning. It also provides various tools for model fitting, data preprocessing, model selection and evaluation and many other utilities. In a few lines, we can develop a learning models as follows:

- Create the **Dataset**. (Z^{in}, \mathcal{U}^i) such that
 - Z^{in} are the input datas. It contains different initial values of wealth $(\mathfrak{z}_j)_{j \in \llbracket 1, n \rrbracket}$, (i.e)

$$(Z_j^i)_{j \in \llbracket 1, n \rrbracket} = (\mathfrak{z}_j)_{j \in \llbracket 1, n \rrbracket}.$$
 - \mathcal{U}^i are the target. It contains a vector of the approximated utility matrix $(\mathbf{U}(\tau_i, \mathfrak{z}_j))$ for $(i, j) \in \llbracket 1, N \rrbracket \times \llbracket 1, n \rrbracket$.
- Define our **model**: choose the best model to fit our data set among the predefined models in the package.
- Fitting **SKlearn.fit** (Z^{in}, \mathcal{U}^i) : The objective is fitted the model, (i.e), the machine must learn from the model.
 1. The samples matrix Z^{in} . The size of Z^{in} is typically (n_samples, n_features), which means that samples are represented as lines and features are represented as columns.
 2. The target values \mathcal{U}^i is usually 1d array where the i-th entry corresponds to the target of the i th sample of Z^{in} .
- Predicting: **SKlearn.predict** (Z^{in}) : Once the estimator is fitted, it can be used for predicting target values of new data. For this step we have as output :

$$U^{p,i} := \text{SKlearn.predict}(Z^{in}),$$

such that U^p represents the value predicted with our model.

- Evaluation: **SKlearn.score** (Z^{in}, \mathcal{U}^i) : We can now observe the accuracy of our model using the function **Model.score**. This function returns the coefficient of determination R^2 of the prediction. The coefficient R^2 is defined as

$$R^2 = 1 - \frac{(\mathcal{U}^i - U^{p,i})^2}{(\mathcal{U}^i - \bar{U})^2},$$

with $\bar{U} := \frac{1}{N} \sum_{j=1}^N \mathcal{U}_j^i$ represents the mean of $(\mathcal{U}_j^i)_j$.

In this work, we propose to use two packages from **Scikit-learn library** create our model and we will compare the results:

- **Support vector Machine** : In machine learning, **support-vector machines** (SVM in short) are supervised learning models with associated learning algorithms that analyze data for classification and regression analysis, developed by Vladimir Vapnik et Alexey Chervonenkis. The problem of Support Vector regression (SVR) is to find a function that approximates mapping from an input domain of real numbers on the basis of a training sample. The objective, when we are moving on with SVR, is to basically consider the points that are within the decision boundary line. The best fit line is the hyperplane of the equation $a_j Z_j^{in} + b, \forall j \in \llbracket 1, n \rrbracket$ that has a maximum number of points. Thus, to summarize, we are looking for a hyperplane such that most of train observation are inside the marge ϵ and we need to add a supplementary constraint

$$U_j^i - (a_j z_j + b) \leq \epsilon + \xi_j, \text{ if the point is above the margin,} \quad (1.3.54)$$

and

$$(a_j z_j + b) - U_j^i \leq \epsilon + \xi_j^*, \text{ if the point is below the margin,} \quad (1.3.55)$$

where

1. $Z^{in} = (Z_j^{in})_{j \in \llbracket 1, n \rrbracket}$ represents the input defined by different initial values of the wealth.
2. (a, b) represent the parameters of weigth $(a_i)_i$ and biais b : we choose it randomly at the beginning then we try to find the optimal values in the algorithm.
3. $U^i = (U_j^i)_j := (\mathbf{U}(\tau_i, z_j))_j$ represents the target: the value which we try to predict. We remain that the target is defined by a column vector of the matrix $\mathbf{U}(\tau_i, z_j)_{i \in \llbracket 1, N \rrbracket, j \in \llbracket 1, n \rrbracket}$ at fixed τ_i .
4. ϵ is the width of the margin of hyperplan.
5. $(\xi_j)_j, (\xi_j^*)_j \geq 0$ represent the error of the observations outside the hyperplane.

This problem allow some observations to fall on the wrong side of the margin, but will penalized them by the parameter C . Finally, we can introduce our optimization problem:

$$\min_{a \in \mathbb{R}^n, b \in \mathbb{R}} \left\{ \frac{1}{2} \|a\|^2 + C \sum_{j=1}^l (\xi_j + \xi_j^*) \right\}, \quad (1.3.56)$$

where l is the number of the points outside the margin, under the constraint [1.3.54-1.3.55](#).

Remarks 2. The optimization problem (4.3.10)-(4.3.11) can be solved in a numerical way very efficiently by the SVR model described above.

Basically, the idea of SVR model is to minimize the sum of squared weights by taking into account the misclassified observations in order to maximize the margin.

The second idea of support vector machine is the kernel method. It is an extension to non-linear cases. It is quite simple and relies on the projection of the data in a higher dimensional space in which the problem becomes linear. The kernel method is a mathematical trick to solve this problem elegantly, by defining the transformations ϕ and scalar products via a kernel $k(a, b) = \phi(a)^T \phi(b)$ such that the equation of hyperplan is as follow

$$U_j^p := a_j \phi(Z_j^{in}) + b.$$

Therefore, using the kernel method, the optimization problem (1.3.57)-(1.3.58) becomes

$$\min_{a \in \mathbb{R}^n, b \in \mathbb{R}} \left\{ \frac{1}{2} \|a\|^2 + C \sum_{i=1}^l (\xi_j + \xi_j^*) \right\}, \quad (1.3.57)$$

where l is the number of the points outside the margin and we remains that $\xi_j = \mathcal{U}_j^i - (a_j \phi(\mathfrak{z}_j) + b - \epsilon)$ and $\xi_j^* = (a_j \phi(\mathfrak{z}_j) + b + \epsilon) - \mathcal{U}_j^i$, under the following constraint

$$\begin{cases} \mathcal{U}_j^i - (a_j \phi(Z_j^{in}) + b) \leq \epsilon + \xi_j, \\ (a_j \phi(Z_j^{in}) + b) - \mathcal{U}_j^i \leq \epsilon + \xi_j^*, \\ \xi_j, \xi_j^* \geq 0, \forall j. \end{cases} \quad (1.3.58)$$

We also chose to use the NuSVR model of the SVM class: for this model, we add a new parameter ν which controls the number of support vectors and allows to compute automatically ϵ .

Then, for $\nu \in (0, 1]$, the optimization problem becomes for fixed time t :

$$\min_{a \in \mathbb{R}^n, b \in \mathbb{R}} \left\{ \frac{1}{2} \|a\|^2 + C(\nu \epsilon + \frac{1}{l} \sum_{i=1}^l (\xi_i + \xi_i^*)) \right\}, \quad (1.3.59)$$

where l is the number of the points outside the margin and we remains that $\xi_j = \mathcal{U}_j^i - (a_j \phi(\mathfrak{z}_j) + b - \epsilon)$ and $\xi_j^* = (a_j \phi(\mathfrak{z}_j) + b + \epsilon) - \mathcal{U}_j^i$, under the same constraint.

- **Neural Network model** : A neural network is an artificial neural network with multiple layers between the input and output layers. In general, it is made up of three types of layers: the first layer is called the input layer, the last layer is called the output layer and possi-

bly other layers called hidden layers. we use the Multi-layer Perceptron model of Neural Network class of SKlearn library. Multi-layer Perceptron (MLP) is a supervised learning algorithm that learns a function \mathcal{U}^i by training on a dataset. Given a set of features Z^{in} and a target \mathcal{U}^i , it can learn a non-linear function approximator for either classification or regression. It is different from logistic regression, in that between the input and the output layer, there can be one or more non-linear layers, called hidden layers.

The idea of Machine Learning is to let the machine find the parameters $\theta_k^{l,i} := (a, b)$ for the l-th and k-th node layer of the model that minimize the cost function J . In order to do this, we need to use a back-propagation learning algorithm.

Actually, the differences between these outputs U^p and the desired outputs U^t form errors that are corrected via back-propagation, the weights of the neural network are then changed. By applying this step several times, the error tends to decrease and the network offers a better prediction.

Then for a fixed neural network structure, number of layers, number of neurons in each layer and fixed activation functions, the program therefore amounts to determine the set of parameters such that

$$\theta^* := \arg \min_{\theta} J(\theta).$$

In this work, we choose the **LBFGR solver** as Back-propagation algorithm of the optimization problem. **LBFGR** is named from the initials of the mathematicians Broyden, Fletcher, Goldfarb and Shanno, who discovered it independently in the late 1960s. This learning algorithm is based on Newton's method which fits the weights using additionally the Hessian matrix H gives the second derivatives of the cost function J with respect to the weights in the following way:

$$\theta_{k,j}^{l,i} = \theta_{k,j}^{l,i} - \alpha H^{-1} \frac{\partial J(\theta_k^{l,i})}{\partial \theta_{k,j}^{l,i}}, j = 0, \dots, n_l + 1, \text{ for the } l\text{-th hidden layer and the } k\text{-th node.}$$

At the end of this work, we implemented these methods in order to predict the dynamic utility of an agent and compared these different learning methods implemented.

Finally, we can deduce that the neural network model is the best prediction model for our dataset. Beside the other models used in this work are not so bad. Unfortunately, there are no clear rules to follow to choose the model to use from the beginning. In fact, it is recommended to start with the easiest and simplest examples, if these models do not work then consider using a more sophisticated model.

MEAN FIELD CONTROL PROBLEM AND APPLICATION TO PRODUCTION OF AN EXHAUSTIBLE RESOURCE

2.1 Introduction

Our aim in this chapter is to study a linear quadratic optimal control for mean field stochastic differential equation with jumps. We will use a weak formulation of the standard martingale optimality principle to solve our control problem. This approach is used in the verification theorem for stochastic control problem which is a crucial step in the classical approach to dynamic programming (see e.g. [47, 66]). We prove the existence of a unique optimal control on finite horizon, by using this method under some assumptions on the coefficients. The optimality of the control is derived by the stability of a decoupled system of Riccati equation and backward stochastic differential equation with jumps. We also consider the problem with common noise adding up some assumptions on the coefficients. On the other hand, we introduce also the linear quadratic problem on the infinite case. This work can be regarded as an extension of the work of Basei and Pham [11] to the case with jumps.

The main motivation behind this work comes from a stochastic model of production of an exhaustible resource, such as oil. It was inspired by Pindyck [107] from Hotelling's work [63] and it was adapted later on many works [37, 38, 41, 86]. In the model, the reserves of an exhaustible resource are a depleted resource, but there is a possibility of exploration and discovery of new reserves which ensure the accumulating or maintaining a level of this reserves. We modelled the new discoveries occurring according to a jump process with intensity given by the exploration effort. The outline of this chapter is as follows. After recalling some notation, we give in Section 2.3 a detailed description of the techniques used to show the existence of the unique optimal control, we then use the weak formulation of the verification theorem introduced in the previous section. In Section 2.4, we extend our results to the case where a common noise is present, so we add a new set of assumptions for the new parameters.

We adapt also our results to the infinite-horizon case, in Section 2.5. Here, we kept the similar

steps as the case of the finite horizon to apply verification theorem on this case, except that we should look for the stability of decoupled system of Riccati equation on infinite horizon.

Finally, in Section 2.6, we introduce an application of production of exhaustible resource with accumulating or maintaining a level of reserves where this increments of reserves is represented via a Poisson process. This work is concertized in the preprint [88].

2.2 Framework: Notations and setting

Let $T > 0$ be a given time horizon and $(\Omega, \mathcal{F}, \mathbb{F} := (\mathcal{F}_t)_{t \leq T}, \mathbb{P})$ be a stochastic basis such that \mathcal{F}_0 contains all the P -null sets of \mathcal{F} , $\mathcal{F}_{t+} = \bigcap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon} = \mathcal{F}_t$, and we suppose that the filtration is generated by the two following mutually independent processes :

- (i) a standard real Brownian motion $B := (B_t)_{0 \leq t \leq T}$ and
- (ii) a Poisson random measure π on $\mathbb{R}^+ \times \chi$, where $\chi \subset \mathbb{R} \setminus \{0\}$ is equipped with its Borel field $\mathcal{B}(\chi)$. Throughout this paper the measure $\lambda(\cdot)$ is assumed to be finite on $(\chi, \mathcal{B}(\chi))$ i.e. $\int_{\chi} \lambda(de) < \infty$. Let $\eta(dt, de) = \lambda(de)dt$ be its compensated process, i.e., $\{\tilde{\pi}([0, t] \times A) = (\pi - \eta)([0, t] \times A)\}_{t \leq T}$ is a martingale for every $A \in \mathcal{B}(\chi)$.

Let $\rho \geq 0$ be the discount factor and define \mathcal{A} the set of admissible controls as follows:

$$\mathcal{A} := \{\alpha : \Omega \times [0, T] \rightarrow \mathbb{R}^m \text{ s.t. } \alpha \text{ is } \mathbb{F}\text{-predictable and } \int_0^T e^{-\rho t} \mathbb{E}[|\alpha_t|^2] dt < \infty\}.$$

Let \mathbb{S}^d be the set of symmetric matrices and $(\mathbb{H}, |\cdot|)$ a normed space. We define the following sets:

- $L^\infty([0, T], \mathbb{H}) := \left\{ \phi : [0, T] \rightarrow \mathbb{H} \text{ s.t. } \phi \text{ is measurable and } \sup_{0 \leq s \leq T} |\phi_s| < \infty \right\},$
- $\mathcal{K}^\infty([0, T] \times \chi, \mathbb{H}) := \left\{ \phi : [0, T] \rightarrow \mathbb{H} \text{ s.t. } \phi \text{ is measurable and } \sup_{0 \leq s \leq T, e \in \chi} |\phi_s(e)| < \infty \right\},$
- $L^2([0, T], \mathbb{H}) := \left\{ \phi : [0, T] \rightarrow \mathbb{H} \text{ s.t. } \phi \text{ is measurable and } \mathbb{E} \left[\int_0^T e^{-\rho s} |\phi_s|^2 ds \right] < \infty \right\},$
- $L^2_{\mathcal{F}_T}(\mathbb{H}) := \left\{ \phi : \Omega \rightarrow \mathbb{H} \text{ s.t. } \phi \text{ is } \mathcal{F}_T\text{-measurable and } \mathbb{E}[|\phi_s|^2] < \infty \right\},$
- $S^2_{\mathbb{F}}(\Omega \times [0, T], \mathbb{H}) := \left\{ \phi : \Omega \times [0, T] \rightarrow \mathbb{H} \text{ s.t. } \phi \text{ is } \mathbb{F}\text{-progressively measurable and } \mathbb{E}[ess \sup_{s \in [0, T]} |\phi_s|^2] < \infty \right\},$
- $L^2_{\mathbb{F}}(\Omega \times [0, T], \mathbb{H}) := \left\{ \phi : \Omega \times [0, T] \rightarrow \mathbb{H} \text{ s.t. } \phi \text{ is } \mathbb{F}\text{-progressively measurable and } \int_0^T e^{-\rho s} \mathbb{E}[|\phi_s|^2] ds < \infty \right\},$

- $\mathcal{K}_{\mathbb{F}}^2(\Omega \times [0, T] \times \chi, \mathbb{H}) := \{K : \Omega \times [0, T] \times \chi \rightarrow \mathbb{H} \text{ s.t. } K \text{ is } \mathcal{P} \otimes \mathcal{B}(\chi)\text{-measurable process}$
and $\mathbb{E}[\int_0^T \int_{\chi} |K_s(e)|^2 \lambda(de) ds] < \infty\}$,

where \mathcal{P} denote the σ -field of \mathbb{F} -predictable sets on $\Omega \times [0, T]$.

We define the controlled linear mean field stochastic differential equation in \mathbb{R}^d , for a given \mathcal{F} -measurable random variable X_0 and a control $\alpha \in \mathcal{A}$, by:

$$\begin{cases} dX_t^\alpha &= b_t(X_t^\alpha, \mathbb{E}[X_t^\alpha], \alpha_t, \mathbb{E}[\alpha_t])dt + \sigma_t(X_t^\alpha, \mathbb{E}[X_t^\alpha], \alpha_t, \mathbb{E}[\alpha_t])dW_t \\ &+ \int_{\chi} R_t(X_t^\alpha, \mathbb{E}[X_t^\alpha], \alpha_t, \mathbb{E}[\alpha_t], e) \tilde{\pi}(de, dt), \\ X_0^\alpha &= X_0, \end{cases} \quad (2.2.1)$$

where for each $t \in [0, T]$, $x, \bar{x} \in \mathbb{R}^d$, $a, \bar{a} \in \mathbb{R}^m$ and $e \in \chi$, we set:

$$\begin{aligned} b_t(x, \bar{x}, a, \bar{a}) &:= \beta_t + A_t x + \tilde{A}_t \bar{x} + B_t a + \tilde{B}_t \bar{a}, \\ \sigma_t(x, \bar{x}, a, \bar{a}) &:= \gamma_t + C_t x + \tilde{C}_t \bar{x} + D_t a + \tilde{D}_t \bar{a}, \\ R_t(x, \bar{x}, a, \bar{a}, e) &:= R_t(x, \bar{x}, a, \bar{a})(e) \\ &:= \delta_t(e) + F_t(e)x + \tilde{F}_t(e)\bar{x} + G_t(e)a + \tilde{G}_t(e)\bar{a}. \end{aligned} \quad (2.2.2)$$

Here β, γ are vector-valued \mathbb{F} -progressively processes, δ is vector-valued \mathbb{F} -predictable process, and $A, \tilde{A}, B, \tilde{B}, C, \tilde{C}, D, \tilde{D}, F, \tilde{F}, G, \tilde{G}$ are deterministic matrix-valued functions such that $A, \tilde{A}, C, \tilde{C} : [0, T] \rightarrow \mathbb{R}^{d \times d}$, $B, \tilde{B}, D, \tilde{D} : [0, T] \rightarrow \mathbb{R}^{d \times m}$. In the other hand $F, \tilde{F} : [0, T] \times \Omega \times \chi \rightarrow \mathbb{R}^{d \times d}$ and $G, \tilde{G} : [0, T] \times \chi \rightarrow \mathbb{R}^{m \times d}$ are $\mathcal{P} \otimes \mathcal{B}(\chi)$ -measurable process. We consider a quadratic cost functional to be minimized over $\alpha \in \mathcal{A}$ of the following type:

$$J(\alpha) := \mathbb{E} \left[\int_0^T e^{-\rho t} f_t(X_t^\alpha, \mathbb{E}[X_t^\alpha], \alpha_t, \mathbb{E}[\alpha_t]) dt + e^{-\rho T} g(X_T, \mathbb{E}[X_T]) \right], \quad (2.2.3)$$

where for each $t \in [0, T]$, $x, \bar{x} \in \mathbb{R}^d$ and $a, \bar{a} \in \mathbb{R}^m$

$$\begin{aligned} f_t(x, \bar{x}, a, \bar{a}) &:= (x - \bar{x})^\top Q_t (x - \bar{x}) + \bar{x}^\top (Q_t + \tilde{Q}_t) \bar{x} + 2a^\top I_t (x - \bar{x}) \\ &+ 2\bar{a}^\top (I_t + \tilde{I}_t) \bar{x} + (a - \bar{a})^\top N_t (a - \bar{a}) + \bar{a}^\top (N_t + \tilde{N}) \bar{a} + 2M_t^\top x + 2H_t^\top a, \end{aligned} \quad (2.2.4)$$

and

$$g(x, \bar{x}) := (x - \bar{x})^\top P (x - \bar{x}) + \bar{x}^\top (P + \tilde{P}) \bar{x} + 2L^\top x. \quad (2.2.5)$$

The coefficients M, H, L defined in (2.2.4) and (2.2.5) are vector-valued \mathbb{F} -progressively measurable processes such that $M : [0, T] \times \Omega \rightarrow \mathbb{R}^d$, $H : [0, T] \times \Omega \rightarrow \mathbb{R}^m$ and $L : \Omega \rightarrow \mathbb{R}^d$ and the other coefficients $Q, \tilde{Q}, P, \tilde{P}, N, \tilde{N}, I, \tilde{I}$, are deterministic matrix-valued functions such that $Q, \tilde{Q} : [0, T] \rightarrow \mathbb{R}^{d \times d}$, $P, \tilde{P} \in \mathbb{R}^{d \times d}$, $N, \tilde{N} : [0, T] \rightarrow \mathbb{R}^{m \times m}$ and $I, \tilde{I} : [0, T] \rightarrow \mathbb{R}^{m \times d}$. The

symbol $^\top$ denotes the transpose operator of any vector or matrix.

Now, we assume the following conditions on the coefficients of the problem :

(H1) The coefficients in equations (2.2.2) satisfy:

- (i) $\beta, \gamma \in L^2_{\mathbb{F}}(\Omega \times [0, T], \mathbb{R}^d)$ and $\delta \in \mathcal{K}^2_{\mathbb{F}}(\Omega \times [0, T] \times \chi, \mathbb{R}^d)$,
- (ii) $A, \tilde{A}, C, \tilde{C} \in L^\infty([0, T], \mathbb{R}^{d \times d})$, $B, \tilde{B}, D, \tilde{D} \in L^\infty([0, T], \mathbb{R}^{d \times m})$,
 $G(\cdot), \tilde{G}(\cdot) \in \mathcal{K}^\infty([0, T] \times \chi, \mathbb{R}^{d \times m})$, $F(\cdot), \tilde{F}(\cdot) \in \mathcal{K}^\infty([0, T] \times \chi, \mathbb{R}^{d \times d})$.

(H2) The coefficients in equations (2.2.4)-(2.2.5) satisfy:

- (i) $Q, \tilde{Q} \in L^\infty([0, T], \mathbb{S}^d)$, $P, \tilde{P} \in \mathbb{S}^d$, $N, \tilde{N} \in L^\infty([0, T], \mathbb{S}^m)$, $I, \tilde{I} \in L^\infty([0, T], \mathbb{R}^{m \times d})$,
- (ii) $M \in L^2_{\mathbb{F}}(\Omega \times [0, T], \mathbb{R}^d)$, $H \in L^2_{\mathbb{F}}(\Omega \times [0, T], \mathbb{R}^m)$, $L \in L^2_{\mathbb{F}}(\mathbb{R}^d)$,
- (iii) there exists $\xi > 0$ such that, for each $t \in [0, T]$,

$$N_t \geq \xi \mathbb{I}_m, \quad P \geq 0, \quad Q_t - I_t^\top N_t^{-1} I_t \geq 0,$$

- (iv) there exists $\xi > 0$ such that, for each $t \in [0, T]$,

$$N_t + \tilde{N}_t \geq \xi \mathbb{I}_m, \quad P + \tilde{P} \geq 0, \quad (Q_t + \tilde{Q}_t) - (I_t + \tilde{I}_t)^\top (N_t + \tilde{N}_t)^{-1} (I_t + \tilde{I}_t) \geq 0.$$

In order to simplify, we denote X_t for X_t^α , \bar{X}_t for $\mathbb{E}[X_t^\alpha]$, $\bar{\alpha}$ for $\mathbb{E}[\alpha_t]$ and $b_t(x, \bar{x})$, $\sigma_t(x, \bar{x})$, $R_t(x, \bar{x})(e)$ for $b_t(x, \bar{x}, \alpha, \bar{\alpha})$, $\sigma_t(x, \bar{x}, \alpha, \bar{\alpha})$ and $R_t(x, \bar{x}, \alpha, \bar{\alpha})(e)$. The following lemma gives an estimate for the solution of mean field SDE (2.2.2).

Lemma 2.2.1. *Let Assumption (H1) be satisfied. Then for any admissible control α , the equation (2.2.1) has a unique solution $(X_t)_t \in \mathcal{S}^2_{\mathbb{F}}$. Moreover, we have the following estimate:*

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |X_t|^2 \right] \leq C(1 + \mathbb{E} [|X_0|^2]), \quad (2.2.6)$$

where C is a positive constant which depends on α .

Proof. We will use the fixed point argument of an appropriate map l defined from $L^2_{\mathbb{F}}(\Omega \times [0, T], \mathbb{R}^d)$ into itself such that $(X_t)_t := l((x_t)_t)$ where $(X_t)_t$ is solution of the SDE:

$$dX_t = b_t(x_t, \mathbb{E}[x_t])dt + \sigma_t(x_t, \mathbb{E}[x_t])dW_t + \int_{\chi} R_t(x_t, \mathbb{E}[x_t])(e) \tilde{\pi}(de, dt), \quad (2.2.7)$$

Given $x^1, x^2 \in L^2_{\mathbb{F}}(\Omega \times [0, T], \mathbb{R}^d)$, we define X^1 and X^2 by using equation (2.2.7). We assume that $X_0^1 = X_0^2$. We define $L = \|A\|_{\infty} \vee \|\tilde{A}\|_{\infty} \vee \|C\|_{\infty} \vee \|\tilde{C}\|_{\infty} \vee \|F\|_{\infty} \vee \|\tilde{F}\|_{\infty}$. Then

$$\begin{aligned} \mathbb{E}[|X_t^1 - X_t^2|^2] &\leq 2\mathbb{E}\left[\left|\int_0^t b_s(x_s^1, \mathbb{E}[x_s^1]) - b_s(x_s^2, \mathbb{E}[x_s^2])ds\right|^2\right] \\ &\quad + 2\mathbb{E}\left[\left|\int_0^t \sigma_s(x_s^1, \mathbb{E}[x_s^1]) - \sigma_s(x_s^2, \mathbb{E}[x_s^2])dB_s\right|^2\right] \\ &\quad + 2\mathbb{E}\left[\left|\int_0^t R_s(x_{s-}^1, \mathbb{E}[x_{s-}^1])(e) - R_s(x_{s-}^2, \mathbb{E}[x_{s-}^2])(e)\tilde{\pi}(de, ds)\right|^2\right]. \end{aligned}$$

From the definition of b_s and using Cauchy Scharwz's inequality, we have:

$$\begin{aligned} |b_s(x_s^1, \mathbb{E}[x_s^1]) - b_s(x_s^2, \mathbb{E}[x_s^2])|^2 &= |A_s(x_s^1 - x_s^2) + \tilde{A}_s(\mathbb{E}[x_s^1] - \mathbb{E}[x_s^2])|^2 \\ &\leq 2L^2(|x_s^1 - x_s^2|^2 + \mathbb{E}[|x_s^1 - x_s^2|^2]). \end{aligned}$$

Using again Cauchy Scharwz's inequality, we have:

$$\mathbb{E}\left[\left|\int_0^t b_s(x_s^1, \mathbb{E}[x_s^1]) - b_s(x_s^2, \mathbb{E}[x_s^2])ds\right|^2\right] \leq 4L^2T \int_0^t \mathbb{E}[|x_s^1 - x_s^2|^2]ds. \quad (2.2.8)$$

For the stochastic integral terms, we have:

$$\mathbb{E}\left[\left|\int_0^t \sigma_s(x_s^1, \mathbb{E}[x_s^1]) - \sigma_s(x_s^2, \mathbb{E}[x_s^2])dB_s\right|^2\right] \leq 4L^2 \int_0^t \mathbb{E}[|x_s^1 - x_s^2|^2]ds \quad (2.2.9)$$

$$\begin{aligned} &\mathbb{E}\left[\left|\int_0^t R_s(x_{s-}^1, \mathbb{E}[x_{s-}^1])(e) - R_s(x_{s-}^2, \mathbb{E}[x_{s-}^2])(e)\tilde{\pi}(de, ds)\right|^2\right] \\ &\leq 4L^2\lambda(\chi) \int_0^t \mathbb{E}[|x_s^1 - x_s^2|^2]ds. \end{aligned} \quad (2.2.10)$$

Plugging inequalities (2.2.8)-(2.2.10) into (2.2.7), we obtain:

$$\mathbb{E}[|X_t^1 - X_t^2|^2] \leq 4L^2(1 + T + \lambda(\chi)) \int_0^t \mathbb{E}[|x_s^1 - x_s^2|^2]ds.$$

For a positive constant c , we define the norm $\|x\|_c^2 := \mathbb{E}[\int_0^T e^{-cs} |x_s|^2 ds]$. Then, we have:

$$\begin{aligned}
\|X^1 - X^2\|_c^2 &:= \int_0^T \mathbb{E}[|X_t^1 - X_t^2|^2] e^{-ct} dt \\
&\leq 4L^2(1 + T + \lambda(\chi)) \int_0^T \int_0^t \mathbb{E}[|x_s^1 - x_s^2|^2] ds e^{-ct} dt \\
&\leq 4L^2(1 + T + \lambda(\chi)) \int_0^T \mathbb{E}[|x_s^1 - x_s^2|^2] \int_s^T e^{-ct} dt ds \\
&\leq \frac{4L^2(1 + T + \lambda(\chi))}{c} \int_0^T \mathbb{E}[|x_s^1 - x_s^2|^2] e^{-cs} ds \\
&= \frac{4L^2(1 + T + \lambda(\chi))}{c} \|x^1 - x^2\|_c^2.
\end{aligned}$$

For c large enough, we have $\frac{4L^2(1+T+\lambda(\chi))}{c} < 1$. Therefore l is a contraction mapping on $L_{\mathbb{F}}^2(\Omega \times [0, T], \mathbb{R}^d)$ into itself, and so l has a unique fixed point $X \in L_{\mathbb{F}}^2(\Omega \times [0, T], \mathbb{R}^d)$ which is the unique solution to the mean field SDE (2.2.1).

We turn to prove estimate (2.2.6). For $n \in \mathbb{N}$, we define $\tau_n := \inf\{t \geq 0, |X_t - X_0| \geq n\}$ and $f_n(t) := \mathbb{E}[\sup_{s \leq t} |X_{s \wedge \tau_n} - X_0|^2]$. For all $t \leq T$, we have:

$$\begin{aligned}
&\sup_{t \leq T} |X_{t \wedge \tau_n} - X_0|^2 \leq 2(\sup_{t \leq T} |\int_0^{t \wedge \tau_n} b_s(X_s, \mathbb{E}[X_s], \alpha_s, \mathbb{E}[\alpha_s]) ds|^2 \\
&+ \sup_{t \leq T} |\int_0^{t \wedge \tau_n} \sigma_s(X_s, \mathbb{E}[X_s], \alpha_s, \mathbb{E}[\alpha_s]) ds|^2 \\
&+ \sup_{t \leq T} |\int_0^{t \wedge \tau_n} \int_{\mathcal{X}} R_s(X_{s-}, \mathbb{E}[X_{s-}], \alpha_s, \mathbb{E}[\alpha_s], e) \tilde{\pi}(de, ds)|^2).
\end{aligned}$$

Using Cauchy Scharzw's inequality and Burkholder-Davis-Gundy Inequality, we have:

$$\begin{aligned}
&\mathbb{E}[\sup_{t \leq T} |X_{t \wedge \tau_n} - X_0|^2] \leq C(T \mathbb{E}[\int_0^T |b_{s \wedge \tau_n}(X_{s \wedge \tau_n}, \mathbb{E}[X_{s \wedge \tau_n}], \alpha_{s \wedge \tau_n}, \mathbb{E}[\alpha_{s \wedge \tau_n}])|^2 ds \\
&+ \mathbb{E}[\int_0^T |\sigma_{s \wedge \tau_n}(X_{s \wedge \tau_n}, \mathbb{E}[X_{s \wedge \tau_n}], \alpha_{s \wedge \tau_n}, \mathbb{E}[\alpha_{s \wedge \tau_n}])|^2 ds] \\
&+ \mathbb{E}[\int_0^T \int_{\mathcal{X}} |R_{s \wedge \tau_n}(X_{s \wedge \tau_n}, \mathbb{E}[X_{s \wedge \tau_n}], \alpha_{s \wedge \tau_n}, \mathbb{E}[\alpha_{s \wedge \tau_n}], e)|^2 \lambda(de) ds], \tag{2.2.11}
\end{aligned}$$

where C is a positive constant which could change from line to line. For $\alpha \in \mathcal{A}$ and using the definition of the drift of the state process X (See Equation(2.2.1)), under (H1), we have:

$$\begin{aligned}
&\mathbb{E}[\int_0^T |b_{s \wedge \tau_n}(X_{s \wedge \tau_n}, \mathbb{E}[X_{s \wedge \tau_n}], \alpha_{s \wedge \tau_n}, \mathbb{E}[\alpha_{s \wedge \tau_n}])|^2 ds] \leq \int_0^T C \mathbb{E}[1 + |X_{s \wedge \tau_n}|^2 + \mathbb{E}[|X_{s \wedge \tau_n}|^2]] \\
&\leq \int_0^T C(1 + \mathbb{E}[|X_0|^2] + f_n(s)) ds \tag{2.2.12}
\end{aligned}$$

Similarly, we have:

$$\begin{aligned}
& \mathbb{E} \left[\int_0^T |\sigma_{s \wedge \tau_n}(X_{s \wedge \tau_n}, \mathbb{E}[X_{s \wedge \tau_n}], \alpha_{s \wedge \tau_n}, \mathbb{E}[\alpha_{s \wedge \tau_n}])|^2 ds \right] \\
& + \mathbb{E} \left[\int_0^T \int_{\mathcal{X}} |R_{s \wedge \tau_n}(X_{s \wedge \tau_n}, \mathbb{E}[X_{s \wedge \tau_n}], \alpha_{s \wedge \tau_n}, \mathbb{E}[\alpha_{s \wedge \tau_n}], e)|^2 \lambda(de) ds \right] \\
& \leq \int_0^t C(1 + \mathbb{E}[|X_0|^2] + f_n(s)) ds.
\end{aligned} \tag{2.2.13}$$

Plugging inequalities (2.2.12) and (2.2.13) into (2.2.11) and using Gronwall's lemma, we obtain

$$\mathbb{E}[\sup_{s \leq T} |X_{s \wedge \tau_n} - X_0|^2] \leq C(1 + \mathbb{E}[|X_0|^2]).$$

Sending n to infinity, by monotone convergence theorem, we obtain $\mathbb{E}[\sup_{s \leq T} |X_s - X_0|^2] \leq C(1 + \mathbb{E}[|X_0|^2])$ and so the estimate (2.2.6) is obtained.

2.3 Linear quadratic mean field control problem on finite horizon

To solve a linear quadratic mean field control problem (LQMF control problem in short), we have to find a strategy $\alpha^* \in \mathcal{A}$, such that

$$V := J(\alpha^*) = \inf_{\alpha \in \mathcal{A}} J(\alpha), \tag{2.3.14}$$

where the criterion J is defined by (2.2.3). By Assumption (H2) and the estimate (2.2.6), the LQMF control problem (2.3.14) is well defined. The aim of this section is to solve this control problem on finite horizon by proving a suitable verification theorem.

We are going to use a weak formulation of the standard martingale optimality principle, see e.g. [47, 66].

Lemma 2.3.1. *(Finite horizon Verification Theorem, Lemma 3.1 in [11])*

Let $\{\mathcal{W}_t^\alpha, t \in [0, T], \alpha_t \in \mathcal{A}\}$ be a family of \mathbb{F} -adapted process in this form $\mathcal{W}_t^\alpha = w_t(X_t^\alpha, \mathbb{E}[X_t^\alpha])$ for some \mathbb{F} -adapted random field $\{w_t(x, \bar{x}), t \in [0, T], x, \bar{x} \in \mathbb{R}^d\}$ satisfying

$$w_t(x, \bar{x}) \leq C(\mathcal{X}_t + |x|^2 + |\bar{x}|^2), \quad t \in [0, T], \quad x, \bar{x} \in \mathbb{R}^d, \tag{2.3.15}$$

where C is a positive constant, \mathcal{X} is a non-negative process satisfying $\sup_{t \in [0, T]} \mathbb{E}[|\mathcal{X}_t|] < \infty$, and

$$(i) \quad w_T(x, \bar{x}) = g(x, \bar{x}), \quad x, \bar{x} \in \mathbb{R}^d,$$

(ii) the map $t \in [0, T] \mapsto \mathbb{E}[\mathbf{S}_t^\alpha]$, with $\mathbf{S}_t^\alpha = e^{-\rho t} \mathcal{W}_t^\alpha + \int_0^t e^{-\rho s} f_s(X_s^\alpha, \mathbb{E}[X_s^\alpha], \alpha_t, \mathbb{E}[\alpha_t]) ds$, is non-decreasing for all $\alpha \in \mathcal{A}$,

(iii) the map $t \mapsto \mathbb{E}[\mathbf{S}_t^{\alpha^*}]$ is constant for some $\alpha^* \in \mathcal{A}$.

Then, α^* is an optimal control and $\mathbb{E}[w_0(X_0, \mathbb{E}[X_0])]$ is the value of the LQMF control problem (2.3.14) i.e.

$$V_0 = \mathbb{E}[w_0(X_0, \mathbb{E}[X_0])] = J(\alpha^*).$$

Moreover, any other optimal control satisfies the condition (iii).

Proof. From the growth condition (2.3.15) and the estimate (2.2.6), for all $t \in [0, T]$ and $\alpha \in \mathcal{A}$, $\mathbb{E}[\mathbf{S}_t^\alpha]$ is well defined. From condition (i), we have $\mathbb{E}[e^{-\rho T} \mathcal{W}_T^\alpha] = \mathbb{E}[e^{-\rho T} g(X_T, \bar{X}_T)]$, which implies $\mathbb{E}[\mathbf{S}_T^\alpha] = J(\alpha)$.

From condition (ii), we have

$$\mathbb{E}[\mathcal{W}_0^\alpha(X_0, \bar{X}_0)] = \mathbb{E}[\mathbf{S}_0^\alpha] \leq \mathbb{E}[\mathbf{S}_T^\alpha] = J(\alpha).$$

In the other side, for $\alpha = \alpha^*$ and using condition (iii), we have

$$\mathbb{E}[\mathcal{W}_0^{\alpha^*}(X_0, \bar{X}_0)] = \mathbb{E}[\mathbf{S}_0^{\alpha^*}] = \mathbb{E}[\mathbf{S}_T^{\alpha^*}] = J(\alpha^*),$$

which ensures the optimality of control α^* .

Moreover, we consider an another optimal control $\tilde{\alpha} \in \mathcal{A}$, then we have

$$\mathbb{E}[\mathcal{W}_0^{\tilde{\alpha}}(X_0, \bar{X}_0)] = \mathbb{E}[\mathbf{S}_0^{\tilde{\alpha}}] = \mathbb{E}[\mathbf{S}_0^{\alpha^*}] = \mathbb{E}[\mathbf{S}_T^{\alpha^*}] = J(\alpha^*) = J(\tilde{\alpha}).$$

Since the map $t \rightarrow \mathbb{E}[\mathbf{S}_t^{\tilde{\alpha}}]$ is non-decreasing, so it is constant, and we conclude the proof.

We start by introducing the following notations:

$$\begin{aligned} \hat{A} &:= A + \tilde{A}, \quad \hat{B} := B + \tilde{B}, \quad \hat{C} := C + \tilde{C}, \quad \hat{D} := D + \tilde{D}, \quad \hat{F} := F + \tilde{F}, \\ \hat{G} &:= G + \tilde{G}, \quad \hat{I} := I + \tilde{I}, \quad \hat{N} := N + \tilde{N}, \quad \hat{Q} := Q + \tilde{Q}, \quad \hat{P} := P + \tilde{P}. \end{aligned}$$

Our aim is to find the expression of the random field $w_t(x, \bar{x})$ which satisfies the assumptions of Lemma (2.3.1).

From the quadratic form of the cost functional f_t in (2.2.4) and the terminal cost g_t in (2.2.5), we guess the quadratic form of random field $\{w_t(x, \bar{x}), t \in [0, T], x, \bar{x} \in \mathbb{R}^d\}$ i.e.

$$w_t(x, \bar{x}) = (x - \bar{x})^\top K_t (x - \bar{x}) + \bar{x}^\top \Lambda_t \bar{x} + 2Y_t^\top x + \varphi_t, \quad (2.3.16)$$

where (K, Λ, Y, φ) are valued in $\mathbb{S}^d \times \mathbb{S}^d \times \mathbb{R}^d \times \mathbb{R}$ and solution of the following system:

$$\begin{cases} dK_t = \dot{K}_t dt, & 0 \leq t \leq T, & K_T = P, \\ d\Lambda_t = \dot{\Lambda}_t dt, & 0 \leq t \leq T, & \Lambda = \hat{P}, \\ dY_t = \dot{Y}_t dt + Z_t^Y dW_t + \int_{\mathcal{X}} R_t^Y(e) \tilde{\pi}(de, dt), & 0 \leq t \leq T, & Y_T = L, \\ d\varphi_t = \dot{\varphi}_t dt, & 0 \leq t \leq T, & \varphi_T = 0. \end{cases} \quad (2.3.17)$$

The deterministic functions $(\dot{K}, \dot{\Lambda}, \dot{\varphi})$ are valued in $\mathbb{S}^d \times \mathbb{S}^d \times \mathbb{R}$, the processes (\dot{Y}, Z^Y) are \mathbb{F} -adapted valued in $\mathbb{R}^d \times \mathbb{R}^d$ and R^Y is \mathbb{F} -predictable process valued in \mathbb{R}^d . For $\alpha \in \mathcal{A}$, we consider S^α defined as:

$$S_t^\alpha = e^{-\rho t} w_t(X_t, \bar{X}_t) + \int_0^t e^{-\rho s} f_s(X_s, \bar{X}_s, \alpha_t, \bar{\alpha}_t) ds, \quad t \in [0, T]. \quad (2.3.18)$$

Then $\mathbb{E}[S_t^\alpha]$ satisfies the following ODE:

$$d\mathbb{E}[S_t^\alpha] = e^{-\rho t} \mathbb{E}[\mathcal{D}_t^\alpha] dt,$$

where the drift $\mathbb{E}[\mathcal{D}_t^\alpha]$ is defined as follows:

$$\mathbb{E}[\mathcal{D}_t^\alpha] = \mathbb{E} \left[-\rho w_t(X_t, \bar{X}_t) + \frac{d}{dt} \mathbb{E}[w_t(X_t, \bar{X}_t)] + f_t(X_t, \bar{X}_t, \alpha_t, \bar{\alpha}_t) \right].$$

From the dynamics of X_t (See equation (2.2.1)), we have:

$$d\bar{X}_t = [\bar{\beta}_t + \hat{A}_t \bar{X}_t + \hat{B}_t \bar{\alpha}_t] dt, \quad (2.3.19)$$

and

$$\begin{aligned} d(X_t - \bar{X}_t) &= [(\beta_t - \bar{\beta}_t) + A_t(X_t - \bar{X}_t) + B_t(\alpha_t - \bar{\alpha}_t)] dt \\ &+ [\gamma_t + C_t(X_t - \bar{X}_t) + \hat{C}_t \bar{X}_t + D_t(\alpha_t - \bar{\alpha}_t) + \hat{D}_t \bar{\alpha}_t] dW_t \\ &+ \int_{\mathcal{X}} [\delta_t + F_t(e)(X_{t-} - \bar{X}_{t-}) + \hat{F}_t(e) \bar{X}_{t-} + G_t(e)(\alpha_t - \bar{\alpha}_t) + \hat{G}_t(e) \bar{\alpha}_t] \tilde{\pi}(de, dt). \end{aligned} \quad (2.3.20)$$

We apply the Itô's formula to $w_t(X_t, \bar{X}_t)$, we use the quadratic form of the running cost f_t and the dynamics of equations (2.3.19), (2.3.20) and (2.2.1), we obtain :

$$\begin{aligned} \mathbb{E}[\mathcal{D}_t^\alpha] &= \mathbb{E} \left[(X_t - \bar{X}_t)^\top (\dot{K}_t + \phi_t) (X_t - \bar{X}_t) + \bar{X}_t^\top (\dot{\Lambda}_t + \psi_t) \bar{X}_t \right. \\ &\quad \left. + 2(\dot{Y}_t + \Delta_t)^\top X_t + \dot{\varphi}_t - \rho \varphi_t + \bar{\Gamma}_t + \kappa_t(\alpha) \right], \end{aligned} \quad (2.3.21)$$

where, for $t \in [0, T]$,

$$\left\{ \begin{array}{l} \phi_t := -\rho K_t + K_t A_t + A_t^\top K_t + C_t^\top K_t C_t + Q_t + \int_{\mathcal{X}} F_t^\top(e) K_t F_t(e) \lambda(de) = \phi_t(K_t), \\ \psi_t := -\rho \Lambda_t + \Lambda_t \hat{A}_t + \hat{A}_t^\top \Lambda_t + \hat{C}_t^\top K_t \hat{C}_t + \hat{Q}_t + \int_{\mathcal{X}} \hat{F}_t^\top(e) K_t \hat{F}_t(e) \lambda(de) = \psi_t(K_t, \Lambda_t), \\ \Delta_t := -\rho Y_t + A_t^\top (Y_t - \bar{Y}_t) + \hat{A}_t^\top \bar{Y}_t + C_t^\top (Z_t^Y - \bar{Z}_t^Y) + \hat{C}_t^\top \bar{Z}_t^Y \\ \quad + K_t(\beta_t - \bar{\beta}_t) + \Lambda_t \bar{\beta}_t + C_t^\top K_t(\gamma_t - \bar{\gamma}_t) + \hat{C}_t^\top K_t \bar{\gamma}_t + M_t \\ \quad + \int_{\mathcal{X}} F_t^\top(e) (R_t^Y(e) - \bar{R}_t^Y(e)) \lambda(de) + \int_{\mathcal{X}} \hat{F}_t^\top(e) \bar{R}_t^Y(e) \lambda(de) \\ \quad + \int_{\mathcal{X}} F_t^\top(e) K_t (\delta_t(e) - \bar{\delta}_t(e)) \lambda(de) + \int_{\mathcal{X}} \bar{F}_t^\top(e) K_t \bar{\delta}_t(e) \lambda(de) \\ \quad = \Delta_t(K_t, \Lambda_t, Y_t, \bar{Y}_t, Z_t^Y, \bar{Z}_t^Y, R_t^Y, \bar{R}_t^Y), \\ \Gamma_t := \gamma_t^\top K_t \gamma_t + 2\beta_t^\top Y_t + 2\gamma_t^\top Z_t^Y + \int_{\mathcal{X}} \delta_t(e)^\top K_t \delta_t(e) \lambda(de) + 2 \int_{\mathcal{X}} \delta_t(e)^\top R_t^Y(e) \lambda(de) \\ \quad = \Gamma_t(K_t, Y_t, Z_t^Y, R_t^Y), \end{array} \right.$$

and

$$\kappa_t(\alpha) := (\alpha_t - \bar{\alpha}_t)^\top S_t (\alpha_t - \bar{\alpha}_t) + \bar{\alpha}_t^\top \hat{S}_t \bar{\alpha}_t + 2(U_t(X_t - \bar{X}_t) + V_t \bar{X}_t + \Theta_t + \epsilon_t - \bar{\epsilon}_t) \alpha_t. \quad (2.3.22)$$

Here $S_t, \hat{S}_t, U_t, V_t, \Theta_t$ which appear in (2.3.22), are defined as follows

$$\left\{ \begin{array}{l} S_t := N_t + D_t^\top K_t D_t + \int_{\mathcal{X}} G_t^\top(e) K_t G_t(e) \lambda(de), \\ \hat{S}_t := \hat{N}_t + \hat{D}_t^\top K_t \hat{D}_t + \int_{\mathcal{X}} \hat{G}_t^\top(e) K_t \hat{G}_t(e) \lambda(de), \\ U_t := I_t + D_t^\top K_t C_t + B_t^\top K_t + \int_{\mathcal{X}} G_t^\top(e) K_t F_t(e) \lambda(de), \\ V_t := \hat{I}_t + \hat{D}_t^\top K_t \hat{C}_t + \hat{B}_t^\top \Lambda_t + \int_{\mathcal{X}} \hat{G}_t^\top(e) K_t \hat{F}_t(e) \lambda(de), \\ \Theta_t := \bar{H}_t + \hat{D}_t^\top K_t \bar{\gamma}_t + \hat{B}_t^\top \bar{Y}_t + \hat{D}_t^\top \bar{Z}_t^Y + \int_{\mathcal{X}} \hat{G}_t^\top(e) K_t \bar{\delta}_t(e) \lambda(de) + \int_{\mathcal{X}} \hat{G}_t^\top(e) \bar{R}_t^Y(e) \lambda(de), \end{array} \right. \quad (2.3.23)$$

and

$$\left\{ \begin{array}{l} \epsilon_t := H_t + D_t^\top K_t \gamma_t + B_t^\top Y_t + D_t^\top Z_t^Y + \int_{\mathcal{X}} G_t^\top(e) K_t \delta_t(e) \lambda(de) + \int_{\mathcal{X}} G_t^\top(e) R_t^Y(e) \lambda(de), \\ \bar{\epsilon}_t := \bar{H}_t + D_t^\top K_t \bar{\gamma}_t + B_t^\top \bar{Y}_t + D_t^\top \bar{Z}_t^Y + \int_{\mathcal{X}} G_t^\top(e) K_t \bar{\delta}_t(e) \lambda(de) + \int_{\mathcal{X}} G_t^\top(e) \bar{R}_t^Y(e) \lambda(de). \end{array} \right. \quad (2.3.24)$$

We notice that the matrices S_t and \hat{S}_t are definite positive in \mathbb{S}^m . This follows from the non-negativity of the matrix K , conditions (iii)-(iv) in (H2) and the non-negativity of the integral $\int_{\mathcal{X}} G_t^\top(e) K_t G_t(e) \pi(dt, de)$. In this case, one could find a deterministic $\mathbb{R}^{m \times m}$ -valued Υ such

that for all $t \in [0, T]$

$$\Upsilon_t S_t \Upsilon_t^\top = \hat{S}_t,$$

for all $t \in [0, T]$. This implies, that we can rewrite the expectation of $\kappa_t(\alpha)$ as

$$\mathbb{E}[\kappa_t(\alpha)] = \mathbb{E} \left[(\alpha_t - \bar{\alpha}_t + \Upsilon_t^\top \bar{\alpha}_t - \eta_t)^\top S_t (\alpha_t - \bar{\alpha}_t + \Upsilon_t^\top \bar{\alpha}_t - \eta_t) - \zeta_t \right],$$

where

$$\eta_t = a_t^0(X_t, \bar{X}_t) + \Upsilon_t^\top a_t^1(\bar{X}_t) dt \otimes d\mathbb{P} \text{ a.e.}, \quad (2.3.25)$$

with $a_t^0(X_t, \bar{X}_t)$ a centred random variable and a_t^1 a deterministic function

$$a_t^0(x, \bar{x}) := -S_t^{-1} U_t (x - \bar{x}) - S_t^{-1} (\epsilon_t - \bar{\epsilon}_t), \quad a_t^1(\bar{x}) := -\hat{S}_t^{-1} (V_t \bar{x} + \Theta_t),$$

and

$$\begin{aligned} \zeta_t := & (X_t - \bar{X}_t)^\top (U_t^\top S_t^{-1} U_t) (X_t - \bar{X}_t) + \bar{X}_t^\top (V_t \hat{S}_t^{-1} V_t) \bar{X}_t + 2(U_t^\top S_t^{-1} (\epsilon_t - \bar{\epsilon}_t) \\ & + V_t^\top \hat{S}_t^{-1} \Theta_t)^\top X_t + (\epsilon_t - \bar{\epsilon}_t)^\top S_t^{-1} (\epsilon_t - \bar{\epsilon}_t) + \Theta_t^\top \hat{S}_t^{-1} \Theta_t. \end{aligned}$$

It yields that one could write (2.3.21) in the following form :

$$\begin{aligned} \mathbb{E}[\mathcal{D}_t^\alpha] = & \mathbb{E} \left[(X_t - \bar{X}_t)^\top (\dot{K}_t + \phi_t^0) (X_t - \bar{X}_t) + \bar{X}_t^\top (\dot{\Lambda}_t + \psi_t^0) \bar{X}_t \right. \\ & + 2(\dot{Y}_t + \Delta_t^0)^\top X_t + \dot{\varphi}_t - \rho \varphi_t + \bar{\Gamma}_t^0 \\ & + (\alpha_t - a_t^0(X_t, \bar{X}_t) - \bar{\alpha}_t + \Upsilon_t^\top (\bar{\alpha}_t - a_t^1(\bar{X}_t)))^\top S_t \\ & \left. (\alpha_t - a_t^0(X_t, \bar{X}_t) - \bar{\alpha}_t + \Upsilon_t^\top (\bar{\alpha}_t - a_t^1(\bar{X}_t))) \right], \end{aligned}$$

where

$$\begin{cases} \phi_t^0 := \phi_t - U_t^\top S_t^{-1} U_t = \phi_t^0(K_t), \\ \psi_t^0 := \psi_t - V_t^\top \hat{S}_t^{-1} V_t = \psi_t^0(K_t, \Lambda_t), \\ \Delta_t^0 := \Delta_t - U_t^\top S_t^{-1} (\epsilon_t - \bar{\epsilon}_t) - V_t^\top \hat{S}_t^{-1} \Theta_t = \Delta_t^0(K_t, \Lambda_t, Y_t, \bar{Y}_t, Z_t^Y, \bar{Z}_t^Y, R_t^Y, \bar{R}_t^Y), \\ \Gamma_t^0 := \Gamma_t - (\epsilon_t - \bar{\epsilon}_t)^\top S_t^{-1} (\epsilon_t - \bar{\epsilon}_t) - \Theta_t^\top \hat{S}_t^{-1} \Theta_t = \Gamma_t^0(K_t, Y_t, Z_t^Y, R_t^Y). \end{cases} \quad (2.3.26)$$

By choosing (K, Λ, Y, φ) s.t. for all $t \in [0, T]$, we have

$$\dot{K}_t + \phi_t^0 = 0, \quad \dot{\Lambda}_t + \psi_t^0 = 0, \quad \dot{Y}_t + \Delta_t^0 = 0, \quad \dot{\varphi}_t - \rho \varphi_t + \bar{\Gamma}_t^0 = 0, \quad (2.3.27)$$

we obtain

$$\mathbb{E}[\mathcal{D}_t^\alpha] = \mathbb{E} \left[(\alpha_t - a_t^0(X_t, \bar{X}_t) - \bar{\alpha}_t + \Upsilon_t^\top (\bar{\alpha}_t - a_t^1(\bar{X}_t)))^\top S_t (\alpha_t - a_t^0(X_t, \bar{X}_t) - \bar{\alpha}_t + \Upsilon_t^\top (\bar{\alpha}_t - a_t^1(\bar{X}_t))) \right], \quad (2.3.28)$$

which is non-negative for all $t \in [0, T]$, $\alpha \in \mathcal{A}$. This shows that \mathbf{S}^α satisfies the condition (ii) of the verification theorem (See Lemma 2.3.1).

In the following lemma, we study the existence of a unique solution to the system of ODEs-BSDE (2.3.17) satisfying (2.3.27).

Lemma 2.3.2. *The decoupled system*

$$\begin{cases} dK_t = -\phi_t^0(K_t)dt, & 0 \leq t \leq T, & K_T = P, \\ d\Lambda_t = -\psi_t^0(K_t, \Lambda_t)dt, & 0 \leq t \leq T, & \Lambda_T = \hat{P}, \\ dY_t = -\Delta_t^0(K_t, \Lambda_t, Y_t, \bar{Y}_t, Z_t^Y, \bar{Z}_t^Y, R_t^Y, \bar{R}_t^Y)dt + Z_t^Y dW_t, \\ \quad + \int_{\mathcal{X}} R_t^Y(e) \tilde{\pi}(de, dt), & 0 \leq t \leq T, & Y_T = L, \\ d\varphi_t = (\rho\varphi_t - \mathbb{E}[\Gamma_t^0(K_t, Y_t, Z_t^Y, R_t^Y)])dt, & 0 \leq t \leq T, & \varphi_T = 0, \end{cases} \quad (2.3.29)$$

where the processes ϕ^0 , ψ^0 , Δ^0 and Γ^0 are defined in (2.3.26), admits a unique solution $(K, \Lambda, Y, Z^Y, R^Y, \varphi)$ in $L^\infty([0, T], \mathbb{S}^d) \times L^\infty([0, T], \mathbb{S}^d) \times \mathcal{S}_{\mathbb{F}}^2(\Omega \times [0, T], \mathbb{R}^d) \times L_{\mathbb{F}}^2(\Omega \times [0, T], \mathbb{R}^d) \times \mathcal{K}_{\mathbb{F}}^2(\Omega \times [0, T] \times \mathcal{X}, \mathbb{R}^d) \times L^\infty([0, T], \mathbb{S}^d)$.

Proof.

- We start by introducing the following multi-dimensional Riccati-type equations

$$\begin{cases} \dot{K}_t + Q_t - \rho K_t + K_t A_t + A_t^\top K_t + C_t^\top K_t C_t + \int_{\mathcal{X}} F_t^\top(e) K_t F_t(e) \lambda(de) \\ \quad - \left(I_t + D_t^\top K_t C_t + B_t^\top K_t + \int_{\mathcal{X}} G_t^\top(e) K_t F_t(e) \lambda(de) \right)^\top \\ \quad \left(N_t + D_t^\top K_t D_t + \int_{\mathcal{X}} G_t^\top(e) K_t G_t(e) \lambda(de) \right)^{-1} \\ \quad \left(I_t + D_t^\top K_t C_t + B_t^\top K_t + \int_{\mathcal{X}} G_t^\top(e) K_t F_t(e) \lambda(de) \right) = 0, \\ K_T = P, \end{cases} \quad (2.3.30)$$

Where the unknown is the matrix K . It is known that the equations (2.3.30) are related to the linear quadratic stochastic control problem:

$$\begin{aligned} v_t(x) &:= \inf_{\alpha \in \mathcal{A}} \mathbb{E} \left[\int_t^T e^{-\rho s} \left((\tilde{X}_s^{\alpha, x})^\top Q_s \tilde{X}_s^{\alpha, x} + 2\alpha_s^\top I \tilde{X}_s^{\alpha, x} + \alpha_s^\top N \alpha_s \right) ds \right. \\ &\quad \left. + e^{-\rho T} (\tilde{X}_T^{\alpha, x})^\top P \tilde{X}_T^{\alpha, x} \right], \end{aligned}$$

where for $\alpha \in \mathcal{A}$, the process \tilde{X} is solution of the following SDE:

$$\begin{cases} d\tilde{X}_s = (A\tilde{X}_s + B\alpha_s)ds + (C\tilde{X}_s + D\alpha_s)dW_s \\ \quad + \int_{\mathcal{X}} (F_s(e)\tilde{X}_s^{\alpha,x} + G_s(e)\alpha_s)\tilde{\pi}(de, ds), \\ \tilde{X}_0 = x. \end{cases} \quad (2.3.31)$$

A straightforward extension of Yong and Zhou ([118], Theorem 7.2 p.320) to the jump diffusion case, and under the assumptions (H1) and (H2) (i)-(iii), there exists a unique solution $K \in L^\infty([0, T], \mathbb{S}^d)$ to the equation (2.3.30) and the solution satisfies $K_t \geq 0$.

- Given K , we consider the following equation for Λ :

$$\begin{cases} \dot{\Lambda}_t + \hat{Q}_t - \rho\Lambda_t + \Lambda_t \hat{A}_t + \hat{A}_t^\top \Lambda_t + \hat{C}_t^\top K_t \hat{C}_t + \int_{\mathcal{X}} \hat{F}_t^\top(e) K_t \hat{F}_t(e) \lambda(de) \\ \quad - \left(\hat{I}_t + \hat{D}_t^\top K_t \hat{C}_t + \hat{B}_t \Lambda_t + \int_{\mathcal{X}} \hat{G}_t^\top(e) K_t \hat{F}_t(e) \lambda(de) \right)^\top \\ \quad \left(\hat{N}_t + \hat{D}_t^\top K_t \hat{D}_t + \int_{\mathcal{X}} \hat{G}_t^\top(e) K_t \hat{G}_t(e) \lambda(de) \right)^{-1} \\ \quad \left(\hat{I}_t + \hat{D}_t^\top K_t \hat{C}_t + \hat{B}_t \Lambda_t + \int_{\mathcal{X}} \hat{G}_t^\top(e) K_t \hat{F}_t(e) \lambda(de) \right) = 0, \\ \Lambda_T = \hat{P}. \end{cases} \quad (2.3.32)$$

We rewrite this multi-dimensional Riccati equations as follows:

$$\begin{cases} \dot{\Lambda}_t + \hat{Q}_t^K - \rho\Lambda_t + \Lambda_t \hat{A}_t + \hat{A}_t^\top \Lambda_t - (\hat{I}_t^K + \hat{B}_t^\top \Lambda_t)^\top (\hat{N}_t^K)^{-1} (\hat{I}_t^K + \hat{B}_t^\top \Lambda_t) = 0 \\ \Lambda_T = \hat{P}, \end{cases} \quad (2.3.33)$$

where $\hat{Q}_t^K, \hat{I}_t^K, \hat{N}_t^K$ are defined by

$$\begin{cases} \hat{Q}_t^K & := \hat{Q}_t + \hat{C}_t^\top K_t \hat{C}_t + \int_{\mathcal{X}} \hat{F}_t^\top(e) K_t \hat{F}_t(e) \lambda(de), \\ \hat{I}_t^K & := \hat{I}_t + \hat{D}_t^\top K_t \hat{C}_t + \int_{\mathcal{X}} \hat{G}_t^\top(e) K_t \hat{F}_t(e) \lambda(de), \\ \hat{N}_t^K & := \hat{N}_t + \hat{D}_t^\top K_t \hat{D}_t + \int_{\mathcal{X}} \hat{G}_t^\top(e) K_t \hat{G}_t(e) \lambda(de). \end{cases}$$

From Assumption (H2)(iv), we have $\hat{P} \geq 0$, and $\hat{N}_t \geq \xi \mathbb{I}_m$. As $K \in L^\infty([0, T], \mathbb{S}^d)$ and $K_t \geq 0$ for all $t \in [0, T]$, then $\hat{N}_t^K \geq \xi \mathbb{I}_m$. It remains to check that for all $t \in [0, T]$, we have

$\hat{Q}_t^K - \hat{I}_t^K (\hat{N}_t^K)^{-1} \hat{I}_t^K \geq 0$. In fact, we have

$$\begin{aligned}
& \hat{Q}_t^K - (\hat{I}_t^K)^\top (\hat{N}_t^K)^{-1} \hat{I}_t^K \tag{2.3.34} \\
&= \hat{Q}_t + \hat{C}_t^\top K_t \hat{C}_t + \int_{\mathcal{X}} \hat{F}_t^\top(e) K_t \hat{F}_t(e) \lambda(de) - (\hat{I}_t + \hat{D}_t^\top K_t \hat{C}_t + \int_{\mathcal{X}} \hat{G}_t^\top(e) K_t \hat{F}_t(e) \lambda(de))^\top \\
&\quad (\hat{N}_t + \hat{D}_t^\top K_t \hat{D}_t + \int_{\mathcal{X}} \hat{G}_t^\top(e) K_t \hat{G}_t(e) \lambda(de))^{-1} (\hat{I}_t + \hat{D}_t^\top K_t \hat{C}_t + \int_{\mathcal{X}} \hat{G}_t^\top(e) K_t \hat{F}_t(e) \lambda(de)) \\
&= \hat{Q}_t - (\hat{I}_t)^\top (\hat{N}_t)^{-1} \hat{I}_t + (\hat{C}_t - \hat{D}_t \hat{N}_t^{-1} \hat{I}_t)^\top K (\hat{C}_t - \hat{D}_t \hat{N}_t^{-1} \hat{I}_t) \\
&\quad - (\hat{D}_t^\top K (\hat{C}_t - \hat{D}_t \hat{N}_t^{-1} \hat{I}_t))^\top (\hat{N}_t + \hat{D}_t^\top K \hat{D}_t)^{-1} (\hat{D}_t^\top K_t (\hat{C}_t - \hat{D}_t \hat{N}_t^{-1} \hat{I}_t)) \\
&\quad + \int_{\mathcal{X}} (\hat{F}_t(e) - \hat{G}_t(e) \hat{N}_t^{-1} \hat{I}_t)^\top K (\hat{F}_t(e) - \hat{G}_t(e) \hat{N}_t^{-1} \hat{I}_t) \lambda(de) \\
&\quad - \int_{\mathcal{X}} (\hat{G}_t(e)^\top K_t (\hat{F}_t(e) - \hat{G}_t^\top(e) (\hat{N}_t)^{-1} \hat{I}_t))^\top \lambda(de) (\hat{N}_t^K)^{-1} \\
&\quad \int_{\mathcal{X}} (\hat{G}_t(e)^\top K (\hat{F}_t(e) - \hat{G}_t^\top(e) (\hat{N}_t)^{-1} \hat{I}_t)) \lambda(de)
\end{aligned}$$

As $K_t \geq 0$, we have $\hat{N}_t^K \geq \hat{N}_t \geq \xi \mathbb{I}_m > 0$, then

$$\begin{aligned}
& (\hat{D}_t^\top K (\hat{C}_t - \hat{D}_t \hat{N}_t^{-1} \hat{I}_t))^\top (\hat{N}_t^K)^{-1} (\hat{D}_t^\top K_t (\hat{C}_t - \hat{D}_t \hat{N}_t^{-1} \hat{I}_t)) \tag{2.3.35} \\
&\leq (\hat{D}_t^\top K (\hat{C}_t - \hat{D}_t \hat{N}_t^{-1} \hat{I}_t))^\top (\hat{D}_t^\top K \hat{D}_t)^{-1} (\hat{D}_t^\top K_t (\hat{C}_t - \hat{D}_t \hat{N}_t^{-1} \hat{I}_t))
\end{aligned}$$

As $K \in L^\infty([0, T], \mathbb{S}^d)$ and $K_t \geq 0$, then by Cauchy Schwarz inequality, we have

$$\begin{aligned}
& \int_{\mathcal{X}} (\hat{F}_t(e) - \hat{G}_t(e) \hat{N}_t^{-1} \hat{I}_t)^\top K_t (\hat{F}_t(e) - \hat{G}_t(e) \hat{N}_t^{-1} \hat{I}_t) \lambda(de) \tag{2.3.36} \\
&\geq \int_{\mathcal{X}} (\hat{F}_t(e) - \hat{G}_t(e) \hat{N}_t^{-1} \hat{I}_t)^\top \lambda(de) K_t \int_{\mathcal{X}} (\hat{F}_t(e) - \hat{G}_t(e) \hat{N}_t^{-1} \hat{I}_t) \lambda(de)
\end{aligned}$$

From inequalities (2.3.34)-(2.3.36), by algebraic manipulations, we deduce

$$\hat{Q}_t^K - (\hat{I}_t^K)^\top (\hat{N}_t^K)^{-1} \hat{I}_t^K \geq \hat{Q}_t + \hat{I}_t^\top (\hat{N}_t)^{-1} \hat{I}_t \geq 0$$

As in the first step and according to [114] and to Yong and Zhou [118], we deduce that equation (2.3.33) admits a unique solution $\Lambda \in L^\infty([0, T], \mathbb{S}^d)$ with $\Lambda_t \geq 0$.

- Given (K, Λ) , we consider the following mean field BSDE

$$\begin{cases} dY_t = -\Delta_t^0(K_t, \Lambda_t, Y_t, \bar{Y}_t, Z_t^Y, \bar{Z}_t^Y, R_t^Y, \bar{R}_t^Y) dt + Z_t^Y dW_t + \int_{\mathcal{X}} R_t^Y(e) \tilde{\pi}(de, dt) \\ Y_T = L. \end{cases} \tag{2.3.37}$$

It could be written in the following form

$$\left\{ \begin{array}{l} dY_t = (c_t + \theta_t^\top (Y_t - \mathbb{E}[Y_t]) + \hat{\theta}_t^\top \mathbb{E}[Y_t] + \vartheta_t^\top (Z_t^Y - \mathbb{E}[Z_t^Y]) + \hat{\vartheta}_t^\top \mathbb{E}[Z_t^Y] \\ \quad + \int_{\mathcal{X}} \varrho_t^\top(e) (R_t^Y(e) - \mathbb{E}[R_t^Y(e)]) \lambda(de) + \int_{\mathcal{X}} \hat{\varrho}_t^\top(e) \mathbb{E}[R_t^Y(e)] \lambda(de)) dt \\ \quad + Z_t^Y dW_t + \int_{\mathcal{X}} R_t^Y(e) \tilde{\pi}(de, dt), \\ Y_T = L, \end{array} \right. \quad (2.3.38)$$

where the stochastic process $(c_t)_t \in L^2_{\mathbb{F}}(\Omega \times [0, T], \mathbb{R}^d)$ is defined by, $\forall t \in [0, T]$ \mathbb{P} p.s

$$\begin{aligned} c_t := & -M_t - K_t(\beta_t - \bar{\beta}_t) - \Lambda_t \bar{\beta}_t - C_t^\top K_t(\gamma_t - \bar{\gamma}_t) - \hat{C}_t^\top K_t \bar{\gamma}_t \\ & - \int_{\mathcal{X}} F_t^\top(e) K_t(\delta_t(e) - \bar{\delta}_t(e)) \lambda(de) - \int_{\mathcal{X}} \hat{F}_t^\top(e) K_t \bar{\delta}_t(e) \lambda(de) \\ & + U_t^\top S_t^{-1} \left(H_t - \bar{H}_t + D_t^\top K_t(\gamma_t - \bar{\gamma}_t) + \int_{\mathcal{X}} G_t(e)^\top K_t(\delta_t(e) - \bar{\delta}_t(e)) \lambda(de) \right) \\ & + V_t^\top S_t^{-1} (\bar{H}_t + \hat{D}_t^\top K_t \bar{\gamma}_t + \int_{\mathcal{X}} \hat{G}_t^\top(e) K_t \bar{\delta}_t(e) \lambda(de)), \end{aligned}$$

and the deterministic coefficients $\theta, \hat{\theta}, \vartheta, \hat{\vartheta} \in L^\infty([0, T], \mathbb{R}^{d \times d})$, and $\varrho, \hat{\varrho} \in \mathcal{K}^\infty([0, T] \times \mathcal{X}, \mathbb{R}^{d \times d})$ are defined by

$$\begin{aligned} \theta_t &:= \rho \mathbb{I}_d - A_t + B_t S_t^{-1} U_t, & \hat{\theta}_t &:= \rho \mathbb{I}_d - \hat{A}_t + \hat{B}_t \hat{S}_t^{-1} V_t, \\ \vartheta_t &:= -C_t + D_t S_t^{-1} U_t, & \hat{\vartheta}_t &:= -\hat{C}_t + \hat{D}_t \hat{S}_t^{-1} V_t, \\ \varrho_t(e) &:= -F_t(e) + G_t(e) S_t^{-1} U_t, & \hat{\varrho}_t(e) &:= -\hat{F}_t(e) + \hat{G}_t(e) \hat{S}_t^{-1} V_t. \end{aligned}$$

By Tang and Meng [114] (see Lemma 2.2.), under (H1) the BSDE (2.3.37) admits a unique solution $(Y, Z^Y, R^Y) \in \mathcal{S}_{\mathbb{F}}^2(\Omega \times [0, T], \mathbb{R}^d) \times L_{\mathbb{F}}^2(\Omega \times [0, T], \mathbb{R}^d) \times \mathcal{K}_{\mathbb{F}}^2(\Omega \times [0, T] \times \mathcal{X}, \mathbb{R}^d)$.

- Given $(K, \Lambda, Y, Z^Y, R^Y)$, the linear ordinary differential equation for φ :

$$d\varphi_t = (\rho\varphi_t - \mathbb{E}[\Gamma_t^0])dt, \quad \varphi_T = 0, \quad (2.3.39)$$

where Γ_t^0 is defined by (2.3.26), admits a unique explicit solution given by:

$$\varphi_t = \int_t^T e^{-\rho(s-t)} l_s ds.$$

The deterministic function l is defined, for all $t \in [0, T]$ by:

$$l_t := -\mathbb{E} \left[\gamma_t^\top K_t \gamma_t + 2\beta_t^\top Y_t + 2\delta_t^\top Z_t^Y + 2 \int_{\mathcal{X}} \delta_t^\top R_t^Y(e) \lambda(de) \right. \\ \left. + \int_{\mathcal{X}} \delta_t^\top K_t \delta_t \lambda(de) - \epsilon_t^\top S_t^{-1} \epsilon_t \right] - \bar{\epsilon}_t^\top S_t^{-1} \bar{\epsilon}_t + \Theta_t^\top \hat{S}_t^{-1} \Theta_t.$$

The next theorem gives a connection between the solution to the decoupled system (2.3.29) and the solution to the LQMF problem (2.3.14).

Theorem 2.3.3. *Under Assumptions (H1)-(H2), there exists an optimal control α^* for LQMF problem (2.3.14) giving by the following explicit form*

$$\alpha_t^* = -S_t^{-1} U_t (X_{t-}^* - \bar{X}_{t-}^*) - S_t^{-1} (\epsilon_{t-} - \bar{\epsilon}_{t-}) - S_t^{-1} (V_t \bar{X}_{t-}^* + \Theta_t), \quad dt \otimes \mathbb{P} - a.e.$$

where $X^* = X^{\alpha^*}$ is the state process, and α^* is the optimal control. The corresponding value of the problem is

$$V_0 = J(\alpha^*) = \mathbb{E} \left[(X_0 - \bar{X}_0)^T K_0 (X_0 - \hat{X}_0) + \bar{X}_0^T \Lambda_0 \bar{X}_0 + 2Y_0^T X_0 + \varphi_0 \right].$$

Proof. We proved in Lemma 2.3.2 the existence of $(K, \Lambda, Y, Z^Y, R^Y, \varphi)$ solution to the system (2.3.29). We consider the candidate $\{w_t(x, \bar{x})\}$ given by (2.3.16). As K, Λ, R are bounded and Y satisfies a square-integrability condition, then the growth condition of the random field $\{w_t(x, \bar{x})\}$ i.e condition (2.3.15) in the verification theorem (see Lemma 2.3.2) is satisfied. Thanks to the choice of $(K, \Lambda, Y, Z^Y, R^Y, \varphi)$, the terminal conditions satisfied $\{w_T(x, \bar{x})\} = g(x, \bar{x})$, and so Assumption (i) of Lemma 2.3.2 is satisfied. From equation (2.3.28), we deduce that $\mathbb{E}[\mathcal{D}_t^\alpha]$ is non-negative for all $\alpha \in \mathcal{A}$. It yields that Assumption (ii) of the verification in Lemma 2.3.2 is satisfied. Moreover, one could prove that $\mathbb{E}[\mathcal{D}_t^{\alpha^*}] = 0$ for some $\alpha = \alpha^*$ if and only if

$$\alpha_t^* - a_t^0(X_t^*, \bar{X}_t^*) - \bar{\alpha}_t^* + \Upsilon_t^\top (\bar{\alpha}_t^* - a_t^1(\bar{X}_t^*)) = 0 \quad dt \otimes d\mathbb{P} \text{ a.e.} \quad (2.3.40)$$

Taking the expectation and using the fact that $\mathbb{E}[a_t^0(X_t^*, \bar{X}_t^*)] = 0$ and Υ_t is an invertible matrix, we get $\bar{\alpha}_t^* - a_t^1(\bar{X}_t^*) = 0$. Then we obtain $\alpha_t^* = a_t^0(X_t^*, \bar{X}_t^*) + a_t^1(\bar{X}_t^*)$. As the strategy α^* must be \mathbb{F} -predictable and the number of jumps of the state process X is finite a.s. over the time interval $[0, T]$, then the optimal strategy α^* satisfies

$$\alpha_t^* = -S_t^{-1} U_t (X_{t-}^* - \bar{X}_{t-}^*) - S_t^{-1} (\epsilon_{t-} - \bar{\epsilon}_{t-}) - S_t^{-1} (V_t \bar{X}_{t-}^* + \Theta_t).$$

Under Assumptions (H1)-(H2), $S^{-1}, \hat{S}^{-1}, U, V$ are bounded and Θ, ϵ are square-integrable respectively in $L^2([0, T], \mathbb{R}^m)$ and $L_{\mathbb{F}}^2(\Omega \times [0, T], \mathbb{R}^m)$. As X^* satisfies the square integrability condition (See inequality (2.2.6)), then $\alpha^* \in \mathcal{A}$. We proved that there exists a random field

$w_t(x, \bar{x})$ that satisfies the assumptions of Lemma 2.3.2. Then, by the verification theorem, we conclude that α^* is the optimal control for the LQMF problem (2.3.14).

2.4 Extensions of LQMF problem on the case of common noise

In this section, we extend the results of Theorem 2.3.3, to the case with a common noise. Let W and W^0 be two independent real Brownian motions and $\tilde{\pi}$ be an Poisson random measure defined on the same filtered probability space $(\Omega, \mathcal{F}_t, (\mathbb{P}, \mathbb{P}^0))$ where $\mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}$ be the filtration generated by the $(W, W^0, \tilde{\pi})$ and we denote by $\mathbb{F}^0 = \{\mathcal{F}_t^0\}_{t \in [0, T]}$ the filtration generated by W^0 . As in section (2.3), for any r.v. X_0 and $\alpha \in \mathcal{A}$, the controlled process X_t^α is defined by:

$$\begin{cases} dX_t^\alpha = b_t(X_t^\alpha, \mathbb{E}[X_t^\alpha | \mathcal{F}_t^0], \alpha_t, \mathbb{E}[\alpha_t | W_t^0])dt + \sigma_t(X_t^\alpha, \mathbb{E}[X_t^\alpha | \mathcal{F}_t^0], \alpha_t, \mathbb{E}[\alpha_t | \mathcal{F}_t^0])dW_t + \\ \sigma_t^0(X_t^\alpha, \mathbb{E}[X_t^\alpha | \mathcal{F}_t^0], \alpha_t, \mathbb{E}[\alpha_t | \mathcal{F}_t^0])dW_t^0 + \int_{\mathcal{X}} R_t(X_t^\alpha, \mathbb{E}[X_t^\alpha | \mathcal{F}_t^0], \alpha_t, \mathbb{E}[\alpha_t | \mathcal{F}_t^0])(e)\tilde{\pi}(de, dt), \\ X_0^\alpha = X_0, \end{cases} \quad (2.4.41)$$

where for each $t \in [0, T]$, $x, \bar{x} \in \mathbb{R}^d$ and $a, \bar{a} \in \mathbb{R}^m$ we set:

$$\begin{aligned} b_t(x, \bar{x}, a, \bar{a}) &:= \beta_t + A_t x + \tilde{A}_t \bar{x} + B_t a + \tilde{B}_t \bar{a}, \\ \sigma_t(x, \bar{x}, a, \bar{a}) &:= \gamma_t + C_t x + \tilde{C}_t \bar{x} + D_t a + \tilde{D}_t \bar{a}, \\ \sigma_t^0(x, \bar{x}, a, \bar{a}) &:= \gamma_t^0 + C_t^0 x + \tilde{C}_t^0 \bar{x} + D_t^0 a + \tilde{D}_t^0 \bar{a}, \\ R_t(x, \bar{x}, a, \bar{a})(e) &:= \delta_t(e) + F_t(e)x + \tilde{F}_t(e)\bar{x} + G_t(e)a + \tilde{G}_t(e)\bar{a}. \end{aligned} \quad (2.4.42)$$

For this case, we assume that

- (H3) (i) $A, \tilde{A}, B, \tilde{B}, C, \tilde{C}, D, \tilde{D}, C^0, \tilde{C}^0, D^0, \tilde{D}^0, F, \tilde{F}, G, \tilde{G}$ are essentially bounded \mathbb{F}^0 -adapted processes,
(ii) β, γ, γ^0 are square-integrable \mathbb{F} -adapted processes and δ is square-integrable \mathbb{F} -predictable process.

The LQMF control problem is to find $\alpha^* \in \mathcal{A}$ s.t.

$$V_0 := J(\alpha^*) = \inf_{\alpha \in \mathcal{A}} J(\alpha), \quad (2.4.43)$$

where

$$J(\alpha) := \mathbb{E} \left[\int_0^T e^{-\rho t} f_t(X_t^\alpha, \mathbb{E}[X_t^\alpha | \mathcal{F}_t^0], \alpha_t, \mathbb{E}[\alpha_t | \mathcal{F}_t^0])dt + e^{-\rho T} g(X_T^\alpha, \mathbb{E}[X_T^\alpha | \mathcal{F}_T^0]) \right], \quad (2.4.44)$$

and the coefficients f, g defined in (2.2.4)-(2.2.5), satisfy the following assumptions:

- (H4) (i) $Q, \tilde{Q}, I, \tilde{I}, N, \tilde{N}$ are essentially bounded \mathbb{F}^0 -adapted processes,
(ii) P, \tilde{P} are essentially bounded \mathcal{F}_T^0 -measurable random variables,
(iii) M, H are square-integrable \mathbb{F} -adapted processes,
and L is a square-integrable \mathcal{F}_T -measurable random variables.

As in the previous section, we guess a quadratic form for $w_t(x, \bar{x})$ i.e.

$$w_t(x, \hat{x}) = (x - \hat{x})^T K_t (x - \hat{x}) + \bar{x}^T \Lambda_t \bar{x} + 2Y_t^T x + \varphi_t,$$

where the processes K, Λ, Y and φ are to be determined later. As the quadratic terms in f_t and g are \mathbb{F}^0 -adapted, we guess that K and Λ are \mathbb{F}^0 -adapted. Since the affine coefficients in b_t, σ_t and σ_t^0 and the linear coefficients in f_t and g are \mathbb{F} -adapted, then Y is \mathbb{F} -adapted i.e. depends on W, W^0 and $\tilde{\pi}$. We look for processes (K, Λ, Y, φ) valued in $\mathbb{S}^d \times \mathbb{S}^d \times \mathbb{R}^d \times \mathbb{R}$ and satisfy the following system:

$$\begin{cases} dK_t = \dot{K}_t dt + Z_t^K dW_t^0, & 0 \leq t \leq T, & K_T = P, \\ d\Lambda_t = \dot{\Lambda}_t dt + Z_t^\Lambda dW_t^0, & 0 \leq t \leq T, & \Lambda_T = \hat{P}, \\ dY_t = \dot{Y}_t dt + Z_t^Y dW_t + Z_t^{Y,0} dW_t^0 + \int_{\mathcal{X}} R_t^Y(e) \tilde{\pi}(de, dt), & 0 \leq t \leq T, & Y_T = L, \\ d\varphi_t = \dot{\varphi}_t, & 0 \leq t \leq T, & \varphi_T = 0, \end{cases} \quad (2.4.45)$$

for some \mathbb{F}^0 -adapted processes $\dot{K}, \dot{\Lambda}, Z^K, Z^\Lambda$ valued in \mathbb{S}^d and some \mathbb{F} -adapted processes $\dot{Y}, Z^Y, Z^{Y,0}$ valued in \mathbb{R}^d , an \mathbb{F} -predictable process R^Y valued in \mathbb{R}^d and a continuous function $\dot{\varphi}$ valued in \mathbb{R} .

We keep the notations of section 2.3 and we add the following notations:

$$\hat{C}_t^0 = C_t^0 + \tilde{C}_t^0, \quad \hat{D}_t^0 = D_t^0 + \tilde{D}_t^0.$$

For $\alpha \in \mathcal{A}$ and $t \in [0, T]$, we set \mathbf{S}^α in the following form:

$$\mathbf{S}_t^\alpha = e^{-\rho t} W_t^\alpha + \int_0^t e^{-\rho s} f_s(X_s^\alpha, \mathbb{E}[X_s^\alpha | \mathcal{F}_s^0], \alpha_s, \mathbb{E}[\alpha_s | \mathcal{F}_s^0]) ds. \quad (2.4.46)$$

As in Section 2.3, we compute

$$d \mathbb{E}[\mathbf{S}_t^\alpha] = e^{-\rho t} \mathbb{E}[\mathcal{D}_t^\alpha] dt,$$

where

$$\mathbb{E}[\mathcal{D}_t^\alpha] = \mathbb{E} \left[-\rho w_t(X_t^\alpha, \mathbb{E}[X_t^\alpha | \mathcal{F}_t^0]) + \frac{d}{dt} \mathbb{E}[w_t(X_t^\alpha, \mathbb{E}[X_t^\alpha | \mathcal{F}_t^0])] + f_t(X_t^\alpha, \mathbb{E}[X_t^\alpha | \mathcal{F}_t^0], \alpha_t, \mathbb{E}[\alpha_t | \mathcal{F}_t^0]) \right].$$

As in the previous section, we denote by \bar{X}_t for $\mathbb{E}[X_t^\alpha | \mathcal{F}_t^0]$ and $\bar{\alpha}_t$ for $\mathbb{E}[\alpha_t | \mathcal{F}_t^0]$. By applying Itô's formula to S_t^α and using the quadratic form of the cost f_t and the dynamics of \bar{X}_t and $X_t - \bar{X}_t$, we obtain:

$$d\bar{X}_t = [\beta_t + \hat{A}_t \bar{X}_t + \hat{B}_t \bar{\alpha}_t] dt, \quad (2.4.47)$$

and

$$\begin{aligned} d(X_t - \bar{X}_t) &= [(\beta_t - \bar{\beta}_t) + A_t(X_t - \bar{X}_t) + B_t(\alpha_t - \bar{\alpha}_t)] dt \\ &+ [\gamma_t + C_t(X_t - \bar{X}_t) + \hat{C}_t \bar{X}_t + D_t(\alpha_t - \bar{\alpha}_t) + \hat{D}_t \bar{\alpha}_t] dW_t \\ &+ [\gamma_t^0 + C_t^0(X_t - \bar{X}_t) + \hat{C}_t^0 \bar{X}_t + D_t^0(\alpha_t - \bar{\alpha}_t) + \hat{D}_t^0 \bar{\alpha}_t] dW_t^0 \\ &+ \int_{\mathcal{X}} [\delta_t + F_t(e)(X_{t-} - \bar{X}_{t-}) + \hat{F}_t(e) \bar{X}_{t-} + G_t(e)(\alpha_t - \bar{\alpha}_t) + \hat{G}_t(e) \bar{\alpha}_t] \tilde{\pi}(de, dt). \end{aligned} \quad (2.4.48)$$

Using the same arguments as in the previous section, we obtain:

$$\begin{aligned} \mathbb{E}[\mathcal{D}_t^\alpha] &= \mathbb{E} \left[(X_t - \bar{X}_t)^T (\dot{K}_t + \phi_t)(X_t - \bar{X}_t) + \bar{X}_t^\top (\dot{\Lambda}_t + \psi_t) \bar{X}_t \right. \\ &\quad \left. + 2(\dot{Y}_t + \Delta_t) X_t + \dot{\varphi}_t - \rho \varphi_t + \bar{\Gamma}_t + \kappa_t(\alpha) \right], \end{aligned} \quad (2.4.49)$$

where

$$\left\{ \begin{aligned} \phi_t &:= -\rho K_t + K_t A_t + A_t^\top K_t + Z_t^K C_t^0 + (C_t^0)^\top Z_t^K + C_t^\top K_t C_t + (C_t^0)^\top K_t C_t^0 + Q_t \\ &+ \int_{\mathcal{X}} F_t^\top(e) K_t F_t(e) \lambda(de) = \phi_t(K_t, Z_t^K), \\ \psi_t &:= -\rho \Lambda_t + \Lambda_t \hat{A}_t + \hat{A}_t^\top \Lambda_t + Z_t^\Lambda C_t^0 + (C_t^0)^\top Z_t^\Lambda + \hat{C}_t^\top K_t \hat{C}_t + (C_t^0)^\top \Gamma_t C_t^0 + \hat{Q}_t \\ &+ \int_{\mathcal{X}} \hat{F}_t^\top(e) K_t \hat{F}_t(e) \lambda(de) = \psi_t(K_t, \Lambda_t), \\ \Delta_t &:= -\rho Y_t + A_t^\top Y_t + \tilde{A}_t^\top \bar{Y}_t + C_t^\top Z_t^Y + (C_t^0)^\top Z_t^{Y,0} + (\tilde{C}_t^0)^\top \overline{Z_t^{Y,0}} + \tilde{C}_t^\top \overline{Z_t^Y} \\ &+ K_t(\beta_t - \bar{\beta}_t) + \Lambda_t \bar{\beta}_t + C_t^\top K_t(\gamma_t - \bar{\gamma}_t) + \hat{C}_t^\top K_t \bar{\gamma}_t + M_t + Z_t^K(\gamma_t^0 - \bar{\gamma}_t^0) + Z_t^\Lambda \bar{\gamma}_t^0 \\ &+ (C_t^0)^\top K_t(\gamma_t^0 - \bar{\gamma}_t^0) + (\hat{C}_t^0)^\top \Lambda_t \bar{\gamma}_t^0 + \int_{\mathcal{X}} F_t^\top(e)(R_t^Y(e) - \bar{R}_t^Y(e)) \lambda(de) \\ &+ \int_{\mathcal{X}} \hat{F}_t(e)^\top \bar{R}_t^Y(e) \lambda(de) + \int_{\mathcal{X}} F_t^\top(e)(\delta_t(e) - \bar{\delta}_t(e)) \lambda(de) \\ &+ \int_{\mathcal{X}} \hat{F}_t(e)^\top \bar{\delta}_t(e) \lambda(de) = \Delta_t(K_t, Z_t^K, \Lambda_t, Z_t^\Lambda, Y_t, \bar{Y}_t, Z_t^Y, \overline{Z_t^Y}, R_t^Y, \overline{R_t^Y}), \\ \Gamma_t &:= \gamma_t^\top K_t \gamma_t + (\gamma_t^0 - \bar{\gamma}_t^0)^\top K_t(\gamma_t^0 - \bar{\gamma}_t^0) + 2(\delta_t)^\top Z_t^{Y,0} + (\bar{\gamma}_t^0)^\top \Lambda_t(\bar{\gamma}_t^0) + 2\beta_t^\top Y_t \\ &+ 2\gamma_t^\top Z_t^Y + \int_{\mathcal{X}} \delta_t^\top(e) K_t \delta_t(e) \lambda(de) + 2 \int_{\mathcal{X}} \delta_t(e)^\top R_t^Y(e) \lambda(de) \\ &= \Gamma_t(K_t, Y_t, Z_t^Y, R_t^Y), \end{aligned} \right.$$

and

$$\kappa_t(\alpha) := (\alpha_t - \bar{\alpha}_t)^\top S_t(\alpha_t - \bar{\alpha}_t) + \bar{\alpha}_t^\top \hat{S}_t \bar{\alpha}_t + 2(U_t(X_t - \bar{X}_t) + V_t \bar{X}_t + \Theta_t + \epsilon_t - \bar{\epsilon}_t)\alpha_t. \quad (2.4.50)$$

Here $S_t, \hat{S}_t, U_t, V_t, \Theta_t$ are defined as follows:

$$\begin{cases} S_t := N_t + D_t^\top K_t D_t + (D_t^0)^\top K_t D_t^0 + \int_{\mathcal{X}} G_t^\top(e) K_t G_t(e) \lambda(de), \\ \hat{S}_t := \hat{N}_t + \hat{D}_t^\top K_t \hat{D}_t + (\hat{D}_t^0)^\top K_t \hat{D}_t^0 + \int_{\mathcal{X}} \hat{G}_t^\top(e) K_t \hat{G}_t(e) \lambda(de), \\ U_t := I_t + D_t^\top K_t C_t + B_t^\top K_t + (D_t^0)^\top K_t C_t^0 + (D_t^0)^\top Z_t^K + \int_{\mathcal{X}} G_t^\top(e) K_t F_t(e) \lambda(de), \\ V_t := \hat{I}_t + \hat{D}_t^\top K_t \hat{C}_t + \hat{B}_t \Lambda_t + (D_t^0)^\top Z_t^\Lambda + (D_t^0)^\top K_t (C_t^0) + \int_{\mathcal{X}} \hat{G}_t(e) K_t \hat{F}_t(e) \lambda(de), \\ \Theta_t := \bar{H}_t + \hat{D}_t^\top K_t \bar{\gamma}_t + (\hat{D}_t^0)^\top K_t \bar{\gamma}_t + \hat{B}_t^\top \bar{Y}_t + (\hat{D}_t^0)^\top \bar{Z}_t^{Y,0} + \hat{D}_t^\top \bar{Z}_t^Y \\ + \int_{\mathcal{X}} \hat{G}_t^\top(e) K_t \bar{\delta}_t \lambda(de) + \int_{\mathcal{X}} \hat{G}_t^\top(e) \bar{R}_t^Y(e) \lambda(de), \end{cases} \quad (2.4.51)$$

$$\begin{cases} \epsilon_t := H_t + D_t^\top K_t \gamma_t + B_t^\top Y_t + D_t^\top Z_t^Y + (D_t^0)^\top Z_t^{Y,0} + (D_t^0)^\top K_t (\gamma_t) \\ + \int_{\mathcal{X}} G_t^\top(e) K_t \delta_t \lambda(de) + \int_{\mathcal{X}} G_t^\top(e) R_t^Y(e) \lambda(de), \\ \bar{\epsilon}_t := \bar{H}_t + D_t^\top K_t \bar{\gamma}_t + B_t^\top \bar{Y}_t + D_t^\top \bar{Z}_t^Y + (D_t^0)^\top \bar{Z}_t^{Y,0} + (D_t^0)^\top K_t (\bar{\gamma}_t) \\ + \int_{\mathcal{X}} G_t^\top(e) K_t \bar{\delta}_t \lambda(de) + \int_{\mathcal{X}} G_t^\top(e) \bar{R}_t^Y(e) \lambda(de). \end{cases} \quad (2.4.52)$$

Completing the square in the expression of $\mathbb{E}[D_t^\alpha]$ (See equation (2.4.49)) and getting rid of the terms which not depend in α , the computations lead to the following decoupled system,

$$\begin{cases} dK_t = -\phi_t^0(K_t, Z^k)dt + Z^K dW_t^0, \quad 0 \leq t \leq T, \quad K_T = P, \\ d\Lambda_t = -\psi_t^0(K_t, \Lambda_t, Z^\Lambda)dt + Z^\Lambda dW_t^0, \quad 0 \leq t \leq T, \quad \Lambda_T = \hat{P}, \\ dY_t = -\Delta_t^0(K_t, Z_t^K, \Lambda_t, Z_t^\Lambda, Y_t, \bar{Y}_t, Z_t^Y, \bar{Z}_t^Y, R_t^Y, \bar{R}_t^Y)dt + Z_t^Y dW_t \\ + Z_t^{Y,0} dW_t^0 + \int_{\mathcal{X}} R_t^Y(e) \tilde{\pi}(de, dt), \quad 0 \leq t \leq T, \quad Y_T = L, \\ d\varphi_t = (\rho\varphi_t - \mathbb{E}[\Gamma_t^0])dt, \quad 0 \leq t \leq T, \quad \varphi_T = 0, \end{cases} \quad (2.4.53)$$

with

$$\begin{cases} \phi_t^0 := \phi_t - U_t^\top S_t^{-1} U_t, \\ \psi_t^0 := \psi_t - V_t^\top \hat{S}_t^{-1} V_t, \\ \Delta_t^0 := \Delta_t - V_t^\top \hat{S}_t^{-1} \Theta_t - U_t^\top S_t^{-1} (\epsilon_t - \bar{\epsilon}_t), \\ \Gamma_t^0 := \Gamma_t - (\epsilon_t - \bar{\epsilon}_t)^\top S_t^{-1} (\epsilon_t - \bar{\epsilon}_t) - \Theta_t^\top \hat{S}_t^{-1} \Theta_t. \end{cases} \quad (2.4.54)$$

One could prove that (K, Λ) is the unique solution to the Stochastic Backward Riccati Equation

in (2.4.53). Given (K, Λ) , we have also the existence of a unique solution $(Y, Z^Y, Z^{Y,0}, R^Y)$ of the mean field backward stochastic differential equation in (2.4.53). Then from Lemma 2.3.1, we have the following proposition which gives the structure of the optimal control α^* .

Proposition 2.4.1. *Under Assumptions (H3)-(H4), the optimal control α^* for optimal problem (2.4.43)-(2.4.44) is defined as follows*

$$\alpha_t^* = -S_t^{-1}U_t(X_{t-}^* - \mathbb{E}[X_{t-}^* | \mathcal{F}_t^0]) - S_t^{-1}(\epsilon_{t-} - \mathbb{E}[\epsilon_{t-} | \mathcal{F}_t^0]) - \hat{S}_t^{-1}(V_t - \mathbb{E}[X_{t-}^* | \mathcal{F}_t^0] + \Theta_t).$$

Proof. The proof is similar to that of Theorem 2.3.3.

2.5 Linear quadratic mean field control problem on infinite horizon

Let us consider the infinite horizon case. For $\rho > 0$, we define the set of admissible controls as follows:

$$\mathcal{A} := \{\alpha : \Omega \times \mathbb{R}^+ \rightarrow \mathbb{R}^m \text{ s.t } \alpha \text{ is } \mathbb{F}\text{-predictable and } \int_0^\infty e^{-\rho t} \mathbb{E}[|\alpha_t|^2] dt < \infty\}.$$

The controlled process is defined on $[0, \infty)$ by:

$$\begin{cases} dX_t^\alpha &= b_t(X_t^\alpha, \mathbb{E}[X_t^\alpha], \alpha_t, \mathbb{E}[\alpha_t])dt + \sigma_t(X_t^\alpha, \mathbb{E}[X_t^\alpha], \alpha_t, \mathbb{E}[\alpha_t])dW_t \\ &+ \int_{\mathcal{X}} R_t(X_{t-}^\alpha, \mathbb{E}[X_{t-}^\alpha], \alpha_t, \mathbb{E}[\alpha_t])(e) \tilde{\pi}(de, dt), \\ X_0^\alpha &= X_0, \end{cases} \quad (2.5.55)$$

where for each $t \geq 0, x, \bar{x} \in \mathbb{R}^d$, we use the same formulation as in Section 2.3:

$$\begin{aligned} b_t(x, \bar{x}, a, \bar{a}) &:= \beta_t + Ax + \tilde{A}\bar{x} + Ba + \tilde{B}\bar{a}, \\ \sigma_t(x, \bar{x}, a, \bar{a}) &:= \gamma_t + Cx + \tilde{C}\bar{x} + Da + \tilde{D}\bar{a}, \\ R_t(x, \bar{x}, a, \bar{a})(e) &:= \delta_t(e) + F(e)x + \tilde{F}(e)\bar{x} + G(e)a + \tilde{G}(e)\bar{a}. \end{aligned} \quad (2.5.56)$$

For the infinite horizon case, the coefficients of the linear terms are constant vectors, and the coefficients β, γ and δ are stochastic processes. The control problem on infinite horizon is formulated as:

$$V_0 := \inf_{\alpha \in \mathcal{A}} J(\alpha), \quad (2.5.57)$$

where

$$J(\alpha) := \mathbb{E} \left[\int_0^\infty e^{-\rho t} f_t(X_t^\alpha, \mathbb{E}[X_t^\alpha], \alpha_t, \mathbb{E}[\alpha_t]) dt \right], \quad (2.5.58)$$

and, for each $t \geq 0$, $x, \bar{x} \in \mathbb{R}^d$ and $a, \bar{a} \in \mathbb{R}^m$, we define f as in Section 2.2 i.e.

$$\begin{aligned} f_t(x, \bar{x}, a, \bar{a}) := & (x - \bar{x})^\top Q(x - \bar{x}) + \bar{x}^\top (Q + \tilde{Q})\bar{x} + 2a^\top I(x - \bar{x}) \\ & + 2\bar{a}^\top (I + \tilde{I})\bar{x} + (a - \bar{a})^\top N(a - \bar{a}) + \bar{a}^\top (N + \tilde{N})\bar{a} + 2M_t^\top x + 2H_t^\top a. \end{aligned} \quad (2.5.59)$$

Here, we note that the coefficients of the quadratic terms are constant matrices, and the coefficients H and M may be stochastic processes.

Let $(\mathbb{H}, |\cdot|)$ be a normed space, and $\hat{\rho}$ a positive constant. We define the new spaces:

- $L^\infty(\mathbb{R}^+, \mathbb{H}) := \left\{ \phi : \mathbb{R}^+ \rightarrow \mathbb{H} \text{ s.t. } \phi \text{ is measurable and } \sup_{t \geq 0} |\phi_t| < \infty \text{ a.e.} \right\}$,
 - $\mathcal{K}^\infty(\chi, \mathbb{H}) := \left\{ K : \chi \rightarrow \mathbb{H}, K \text{ is } \mathcal{B}(\chi)\text{-measurable and } \sup_{e \in \chi} |K(e)| < \infty \right\}$,
 - $L_{\mathbb{F}}^{2, \hat{\rho}}(\Omega \times \mathbb{R}^+, \mathbb{H}) := \left\{ \phi : \Omega \times \mathbb{R}^+ \rightarrow \mathbb{H} \text{ s.t. } \phi \text{ is } \mathbb{F}\text{-progressively measurable and } \int_0^\infty e^{\hat{\rho}t} \mathbb{E}[|\phi_t|^2] dt < \infty \right\}$
 - $\mathcal{K}_{\mathbb{F}}^{2, \hat{\rho}}(\Omega \times \mathbb{R}^+ \times \chi, \mathbb{H}) := \left\{ K : \Omega \times \mathbb{R}^+ \times \chi \rightarrow \mathbb{H}, K \text{ is } \mathcal{P} \otimes \mathcal{B}(\chi)\text{-measurable processes and } \mathbb{E} \left[\int_0^\infty \int_\chi e^{\hat{\rho}t} |K_t(e)|^2 \lambda(de) dt \right] < \infty \right\}$,
- where \mathcal{P} denote the σ -field of \mathbb{F} -predictable sets on $\Omega \times \mathbb{R}^+$.

We consider the following assumptions on the coefficients of the problem in the infinite horizon case.

(H1') The coefficients in equation (2.5.55) satisfy:

- (i) $\beta, \gamma \in L_{\mathbb{F}}^{2, \hat{\rho}}(\Omega \times \mathbb{R}^+, \mathbb{R}^d)$ and $\delta \in \mathcal{K}_{\mathbb{F}}^{2, \hat{\rho}}(\Omega \times \mathbb{R}^+ \times \chi, \mathbb{R}^d)$,
- (ii) $A, \tilde{A}, C, \tilde{C} \in \mathbb{R}^{d \times d}$ and $B, \tilde{B}, D, \tilde{D}, J, \tilde{J} \in \mathbb{R}^{d \times m}$,
- (iii) $F, \tilde{F} \in \mathcal{K}^\infty(\chi, \mathbb{R}^{d \times d})$ and $G, \tilde{G} \in \mathcal{K}^\infty(\chi, \mathbb{R}^{d \times m})$.

(H2') The coefficients in equation (2.5.59) satisfy:

- (i) $Q, \tilde{Q} \in \mathbb{S}^d, N, \tilde{N} \in \mathbb{S}^m, I, \tilde{I} \in \mathbb{R}^{m \times d}$,
- (ii) $M \in L_{\mathbb{F}}^{2, \hat{\rho}}(\Omega \times \mathbb{R}^+, \mathbb{R}^d), H \in L_{\mathbb{F}}^{2, \hat{\rho}}(\Omega \times \mathbb{R}^+, \mathbb{R}^m)$,
- (iii) $N > 0, \quad Q + I^T N^{-1} I \geq 0$,
- (iv) $N + \tilde{N} > 0, \quad (Q + \tilde{Q}) - (I + \tilde{I})^T (N + \tilde{N})^{-1} (I + \tilde{I}) \geq 0$.

(H3') $\rho > 2 \max\{|A| + |C|^2 + \int_\chi |F(e)|^2 \lambda(de), |\hat{A}|\}$.

Proposition 2.5.1. *Under (H1') and (H3'), the following estimate holds for each square-integrable variable X_0 and $\alpha \in \mathcal{A}$,*

$$\int_0^\infty e^{-\rho t} \mathbb{E}[|X_t^\alpha|^2] dt \leq C_\alpha (1 + \mathbb{E}[|X_0|^2]), \quad (2.5.60)$$

where C_α is a positive constant.

Proof. By Itô's formula and Young's inequality, we have the following estimate

$$\begin{aligned} & \frac{d}{dt} \mathbb{E}[e^{-\rho t} |\bar{X}_t|^2] \\ &= e^{-\rho t} (-\rho |\bar{X}_t|^2 + 2\bar{b}_t^\top \bar{X}_t) \\ &\leq e^{-\rho t} (-\rho |\bar{X}_t|^2 + 2(|\bar{\beta}_t| |\bar{X}_t| + |\hat{B}| |\bar{\alpha}_t| |\bar{X}_t| + \bar{X}_t^\top \hat{A} \bar{X}_t)) \\ &\leq e^{-\rho t} [(-\rho + 2|\hat{A}| + \epsilon) |\bar{X}_t|^2 + c_\epsilon (|\bar{\beta}_t|^2 + |\bar{\alpha}_t|^2)], \end{aligned}$$

where c_ϵ is a positive constant. We define:

$$\zeta_\epsilon := |\bar{X}_0|^2 + c_\epsilon \int_0^\infty e^{-\rho t} \mathbb{E}[|\beta_t^2| + |\alpha_t^2|], \quad \eta_\epsilon := \rho - 2|\hat{A}| - \epsilon.$$

Under (H1') and as $\alpha \in \mathcal{A}$, we have $\zeta_\epsilon < \infty$ and for ϵ small enough, we have $\eta_\epsilon > 0$. By Gronwall's lemma, we get:

$$\int_0^\infty e^{-\rho t} |\bar{X}_t|^2 dt \leq c_{\alpha, \epsilon} (1 + |\bar{X}_0|^2), \quad (2.5.61)$$

where $c_{\alpha, \epsilon}$ is positive constant. Using again Itô's formula and Young's inequality, we have the following estimate:

$$\begin{aligned} & \frac{d}{dt} \mathbb{E}[e^{-\rho t} |X_t - \bar{X}_t|^2] \\ &= e^{-\rho t} \mathbb{E}[-\rho |X_t - \bar{X}_t|^2 + 2(b_t - \bar{b}_t)^\top (X_t - \bar{X}_t) + |\sigma_t|^2 + \int_{\mathcal{X}} |R_t(e)|^2 \lambda(de)] \\ &\leq e^{-\rho t} \mathbb{E}[-\rho |X_t - \bar{X}_t|^2 + 2(|\beta_t - \bar{\beta}_t| |X_t - \bar{X}_t| + |B| |\alpha_t - \bar{\alpha}_t| |X_t - \bar{X}_t| + (X_t - \bar{X}_t)^\top A (X_t - \bar{X}_t)) \\ &\quad + 2(|\gamma_t|^2 + |C|^2 |X_t - \bar{X}_t|^2 + |\hat{C}|^2 |\bar{X}_t|^2 + |D|^2 |\alpha_t|^2 + |\tilde{D}|^2 |\bar{\alpha}_t|^2) \\ &\quad + 2\left(\int_{\mathcal{X}} |\delta_t(e)|^2 \lambda(de) + \int_{\mathcal{X}} |F(e)|^2 \lambda(de) |X_t - \bar{X}_t|^2 + \int_{\mathcal{X}} |\hat{F}(e)|^2 \lambda(de) |\bar{X}_t|^2 + \int_{\mathcal{X}} |G(e)|^2 \lambda(de) |\alpha_t|^2\right. \\ &\quad \left. + \int_{\mathcal{X}} |\tilde{G}(e)|^2 \lambda(de) |\bar{\alpha}_t|^2\right) \\ &\leq e^{-\rho t} \mathbb{E}[(-\rho + 2|A| + 2|C|^2 + 2 \int_{\mathcal{X}} |F(e)|^2 \lambda(de) + \epsilon) |X_t - \bar{X}_t|^2 \\ &\quad + c'_\epsilon (|\beta_t|^2 + |\gamma_t|^2 + \int_{\mathcal{X}} |\delta_t(e)|^2 \lambda(de) + |\alpha_t|^2 + |\bar{X}_t|^2)], \end{aligned}$$

where c_ϵ is a positive constant. We define:

$$\begin{aligned}\zeta'_\epsilon &:= c'_\epsilon \int_0^\infty e^{-\rho t} \mathbb{E}[|\beta_t|^2 + |\gamma_t|^2 + \int_{\mathcal{X}} |\delta_t(e)|^2 \lambda(de)] + |\alpha_t|^2 + |\bar{X}_t|^2 dt, \\ \eta'_\epsilon &:= \rho - 2|A| - 2|C|^2 - 2 \int_{\mathcal{X}} |F(e)|^2 \lambda(de) - \epsilon.\end{aligned}$$

Under (H1'), inequality (2.5.61) and as $\alpha \in \mathcal{A}$, we have $\zeta'_\epsilon < \infty$ and for ϵ small enough, we have $\eta'_\epsilon > 0$. By Gronwall's lemma, we get

$$\mathbb{E}[\int_0^\infty e^{-\rho t} |X_t - \bar{X}_t|^2 dt] \leq c'_{\alpha, \epsilon} (1 + |\bar{X}_0|^2), \quad (2.5.62)$$

for a suitable $c'_{\alpha, \epsilon} > 0$. From inequalities (2.5.61) and (2.5.62), we obtain (2.5.60).

From assumption (H2') and the estimate (2.5.60), the problem (2.5.57) is well defined. The following lemma is a verification theorem based on the weak martingale approach in the infinite horizon case.

Lemma 2.5.2. *Let $\{\mathcal{W}_t^\alpha, t \geq 0, \alpha \in \mathcal{A}\}$ be a family of \mathbb{F} -adapted process in the form $\mathcal{W}_t^\alpha = w_t(X_t^\alpha, \mathbb{E}[X_t^\alpha])$ for some \mathbb{F} -adapted random field $\{w_t(x, \bar{x}), t \in [0, T], x, \bar{x} \in \mathbb{R}^d\}$ satisfying*

$$w_t(x, \bar{x}) \leq C(\chi_t + |x|^2 + |\bar{x}|^2), \quad t \in [0, \infty), \quad x, \bar{x} \in \mathbb{R}^d,$$

for some positive constant C , and non-negative process χ s.t. $e^{-\rho t} \mathbb{E}[\chi_t]$ converge to zero as $t \rightarrow \infty$, and such that:

- (i) the map $t \in \mathbb{R}^+ \mapsto \mathbb{E}[\mathbf{S}_t^\alpha]$, with $\mathbf{S}_t^\alpha = e^{-ct} W_t^\alpha + \int_0^t e^{-cs} f_s(X_s^\alpha, \mathbb{E}[X_s^\alpha], \alpha_t, \mathbb{E}[\alpha_t]) ds$, is non-decreasing for all $\alpha \in \mathcal{A}$,
- (ii) the map $t \in \mathbb{R}^+ \mapsto \mathbb{E}[\mathbf{S}_t^{\alpha^*}]$ is constant for some $\alpha^* \in \mathcal{A}$.

Then, α^* is an optimal control and $\mathbb{E}[w_0(X_0, \mathbb{E}[X_0])]$ is the value of the LQMKV control problem (2.5.57) i.e.

$$V_0 = \mathbb{E}[w_0(X_0, \mathbb{E}[X_0])] = J(\alpha^*).$$

Moreover, any other optimal control satisfies the condition (ii).

We extend the results in Theorem 2.3.3 to the infinite horizon case, where we kept the steps similar to the finite horizon case. We prove our result by applying the lemma 2.5.2. We should look for the stability of decoupled system on infinite horizon.

We adopt the same approach as in the finite time horizon. We consider a candidate for the random field $\omega_t(x, \bar{x})$ in the form:

$$w_t(x, \hat{x}) = (x - \bar{x})^T K_t (x - \bar{x}) + \bar{x}^T \Lambda_t \bar{x} + 2Y_t^T x + \varphi_t,$$

where (K, Λ, Y, φ) valued in $\mathbb{S}^d \times \mathbb{S}^d \times \mathbb{R}^d \times \mathbb{R}$ satisfying the following system:

$$\begin{cases} dK_t = -\phi_t^0 dt, & t \geq 0, \\ d\Lambda_t = -\psi_t^0 dt, & t \geq 0, \\ dY_t = -\Delta_t^0 dt + Z_t^Y dW_t + \int_{\mathcal{X}} R_t^Y(e) \tilde{\pi}(de, dt), & t \geq 0, \\ d\varphi_t = (\rho\varphi_t - \mathbb{E}[\Gamma_t^0]) dt, & t \geq 0. \end{cases} \quad (2.5.63)$$

The maps $\phi^0, \psi^0, \Delta^0, \Gamma^0$ are defined by (2.3.26), where in this case the coefficients $A, \tilde{A}, B, \tilde{B}, C, \tilde{C}, D, \tilde{D}, Q, \tilde{Q}, N, \tilde{N}, I, \tilde{I}, F(\cdot), \tilde{F}(\cdot), G(\cdot), \tilde{G}(\cdot)$ are constant i.e. independent of time.

We note that there are no terminal conditions in the system, as we are in the infinite horizon case. We need to show the existence of a solution to the system (2.5.63).

Lemma 2.5.3. *We assume (H1')-(H2'). Then, the system (2.5.63) admits a solution*

$(K, \Lambda, Y, Z^Y, R^Y, \varphi) \in L^\infty([0, \infty], \mathbb{S}^d) \times L^\infty([0, \infty], \mathbb{S}^d) \times L_{\mathbb{F}}^{2, \hat{\rho}}(\Omega \times [0, \infty], \mathbb{R}^d) \times L_{\mathbb{F}}^{2, \hat{\rho}}(\Omega \times [0, \infty], \mathbb{R}^d) \times \mathcal{K}^{2, \hat{\rho}}(\Omega \times [0, \infty] \times \mathcal{X}, \mathbb{R}^d) \times L^\infty([0, \infty], \mathbb{S}^d)$.

Proof. We prove the existence of a solution to the decoupled system (2.5.63).

- We introduce the following Riccati-type equation:

$$\begin{aligned} & Q - \rho K + KA + A^\top K + C^\top KC + \int_{\mathcal{X}} F^\top(e) KF(e) \lambda(de) \\ & - \left(I + D^\top KC + B^\top K + \int_{\mathcal{X}} G^\top(e) KF(e) \lambda(de) \right)^\top \\ & \left(N + D^\top KD + \int_{\mathcal{X}} G^\top(e) KG(e) \lambda(de) \right)^{-1} \\ & \left(I + D^\top KC + B^\top K + \int_{\mathcal{X}} G^\top(e) KF(e) \lambda(de) \right) = 0. \end{aligned} \quad (2.5.64)$$

We prove the existence of a solution to (2.5.64) by relating it to a suitable infinite-horizon linear-quadratic control problem. For $T \in \mathbb{R}_+ \cup \{\infty\}$ and $x \in \mathbb{R}^d$, we consider the following control problem:

$$V^T(x) := \inf_{\alpha \in \mathbb{A}_T} \mathbb{E} \left[\int_0^T e^{-\rho t} \left((\tilde{X}_t^{\alpha, x})^\top Q \tilde{X}_t^{\alpha, x} + 2\alpha_t^\top I \tilde{X}_t^{\alpha, x} + \alpha_t^\top N \alpha_t \right) dt \right],$$

where \mathbb{A}_T is defined by:

$$\mathbb{A}_T := \{ \alpha : \Omega \times \mathbb{R}^+ \rightarrow \mathbb{R}^m \text{ s.t } \alpha \text{ is } \mathbb{F} \text{ predictable and } \int_0^T e^{-\rho t} \mathbb{E}[|\alpha_t|^2] dt < \infty \},$$

and for $\alpha \in \mathbb{A}_T$, the process $\tilde{X} := \tilde{X}^{\alpha, x}$ is solution of the following SDE:

$$d\tilde{X}_t = (A\tilde{X}_t + B\alpha_t)dt + (C\tilde{X}_t + D\alpha_t)dW_t + \int_{\mathcal{X}} (F(e)\tilde{X}_t + G(e)\alpha_t)\tilde{\pi}(de, dt), \quad \tilde{X}_0 = 0. \quad (2.5.65)$$

Thanks to the integrability condition for $\alpha \in \mathbb{A}_T$, we have $\int_0^T e^{-\rho t} \mathbb{E}[|\tilde{X}_t^{\alpha, x}|^2] dt < \infty$, and so the problems V^T are well-defined for any $T \in \mathbb{R}^+ \cup \{\infty\}$. If $T < \infty$, as already recalled in the finite-horizon case, (H1')-(H2') imply that there exists a unique symmetric solution $(K_t^T)_{t \in [0, T]}$ to Riccati equations:

$$\begin{cases} \frac{d}{dt} K_t^T + Q & -\rho K_t^T + K_t^T A + A^\top K_t^T + C^\top K_t^T C + \int_{\mathcal{X}} F^\top(e) K_t^T F(e) \lambda(de) \\ & - \left(I + D^\top K_t^T C + B^\top K_t^T + \int_{\mathcal{X}} G^\top(e) K_t^T F(e) \lambda(de) \right)^\top \\ & \left(N + D^\top K_t^T D + \int_{\mathcal{X}} G^\top(e) K_t^T G(e) \lambda(de) \right)^{-1} \\ & \left(I + D^\top K_t^T C + B^\top K_t^T + \int_{\mathcal{X}} G^\top(e) K_t^T F(e) \lambda(de) \right) = 0, \\ K_T^T & = 0, \end{cases} \quad (2.5.66)$$

and that for every $x \in \mathbb{R}^d$ we have $V^T(x) = x^\top K_0^T x$. It is easy to check from the definition of V^T that $V^T(x) \rightarrow V^\infty(x)$ as T goes to infinity, from which we deduce that

$$V^\infty(x) = \lim_{T \rightarrow \infty} x^\top K_0^T x = x \left(\lim_{T \rightarrow \infty} K_0^T \right) x, \quad \text{for all } x \in \mathbb{R}^d.$$

This implies the existence of the limit $K = \lim_{T \rightarrow \infty} K_0^T$. By passing to the limit in T in ODE (2.5.66) at $t = 0$, we obtain by standard arguments (see Lemma 2.8 in [?]), that K satisfies (2.5.64). Moreover, $K \in \mathbb{S}^d$ and $K \geq 0$.

- Given K, we consider the following equation of Λ :

$$\hat{Q}^K - \rho\Lambda + \Lambda\hat{A} + \hat{A}^\top\Lambda - (\hat{I}^K + \hat{B}^\top\Lambda)^\top (\hat{N}^K)^{-1} (\hat{I}^K + \hat{B}^\top\Lambda) = 0, \quad (2.5.67)$$

where

$$\begin{cases} \hat{Q}^K & := \hat{Q} + \hat{C}^\top K \hat{C} + \int_{\mathcal{X}} \hat{F}^\top(e) K \hat{F}(e) \lambda(de), \\ \hat{I}^K & := \hat{I} + \hat{D}^\top K \hat{C} + \int_{\mathcal{X}} \hat{G}^\top(e) K \hat{F}(e) \lambda(de), \\ \hat{N}^K & := \hat{N} + \hat{D}^\top K \hat{D} + \int_{\mathcal{X}} \hat{G}^\top(e) K \hat{G}(e) \lambda(de). \end{cases}$$

Existence of a solution to (2.5.67) is obtained by the same arguments used for (2.5.64)

under (H2').

- Given (K, Λ) , we consider the following mean field BSDE with jumps on infinite horizon:

$$\begin{aligned}
dY_t &= \left(c_t + (\rho \mathbb{I}_d + \theta)^\top (Y_t - \mathbb{E}[Y_t]) + (\rho \mathbb{I}_d + \hat{\theta})^\top \mathbb{E}[Y_t] + \vartheta^\top (Z_t^Y - \mathbb{E}[Z_t^Y]) + \hat{\vartheta}^\top \mathbb{E}[Z_t^Y] \right. \\
&\quad + \int_{\mathcal{X}} \varrho^\top(e) (R_t^Y(e) - \mathbb{E}[R_t^Y(e)]) \lambda(de) + \int_{\mathcal{X}} \hat{\varrho}^\top(e) \mathbb{E}[R_t^Y(e)] \lambda(de) \Big) dt \\
&\quad + Z_t^Y dW_t + \int_{\mathcal{X}} R_t^Y(e) \tilde{\pi}(de, dt), \tag{2.5.68}
\end{aligned}$$

where the stochastic process $c \in L_{\mathbb{F}}^{2, \hat{\rho}}(\Omega \times [0, \infty], \mathbb{R}^d)$ is defined by:

$$\begin{aligned}
c_t &:= -M_t - K(\beta_t - \bar{\beta}_t) - \Lambda \bar{\beta}_t - C^\top K(\gamma_t - \bar{\gamma}_t) - \hat{C}^\top K \bar{\gamma}_t \\
&\quad - \int_{\mathcal{X}} F^\top(e) K(\delta_t(e) - \bar{\delta}_t(e)) \lambda(de) - \int_{\mathcal{X}} \bar{F}^\top(e) K \bar{\delta}_t(e) \lambda(de) \\
&\quad + U^\top S^{-1} \left(H_t - \bar{H}_t + D^\top K(\gamma_t - \bar{\gamma}_t) + \int_{\mathcal{X}} G(e)^\top K(\delta_t(e) - \bar{\delta}_t(e)) \lambda(de) \right) \\
&\quad + V^\top \hat{S}^{-1} \left(\bar{H}_t + \hat{D}^\top K \bar{\gamma}_t + \int_{\mathcal{X}} \hat{G}^\top(e) K \bar{\delta}_t(e) \lambda(de) \right),
\end{aligned}$$

the coefficients $\theta, \hat{\theta}, \vartheta, \hat{\vartheta}$ are constant in \mathbb{R}^d and are defined by:

$$\begin{aligned}
\theta &:= -A + B S^{-1} U, \\
\hat{\theta} &:= -\hat{A} + \hat{B} \hat{S}^{-1} V, \\
\vartheta &:= -C + D S^{-1} U, \\
\hat{\vartheta} &:= -\hat{C} + \hat{D} \hat{S}^{-1} V,
\end{aligned}$$

and $\varrho, \hat{\varrho} : \mathcal{X} \rightarrow \mathbb{R}^d$ are defined by:

$$\begin{aligned}
\varrho(e) &:= -F(e) + G(e) S^{-1} U, \quad \forall e \in \mathcal{X}, \\
\hat{\varrho}(e) &:= -\hat{F}(e) + \hat{G}(e) \hat{S}^{-1} V, \quad \forall e \in \mathcal{X}.
\end{aligned}$$

To simplify the notations let us denote:

$$\begin{aligned}
-f(t, Y_t, Z_t, R_t, \mathbb{E}[Y_t], \mathbb{E}[Z_t], \mathbb{E}[R_t]) &= c_t + (\rho \mathbb{I}_d + \theta)^\top (Y_t - \mathbb{E}[Y_t]) + (\rho \mathbb{I}_d + \hat{\theta})^\top \mathbb{E}[Y_t] \\
&\quad + \vartheta^\top (Z_t - \mathbb{E}[Z_t]) + \hat{\vartheta}^\top \mathbb{E}[Z_t] + \int_{\mathcal{X}} \varrho^\top(e) (R_t(e) - \mathbb{E}[R_t(e)]) \lambda(de) + \int_{\mathcal{X}} \hat{\varrho}^\top(e) \mathbb{E}[R_t(e)] \lambda(de).
\end{aligned}$$

We prove that the following linear BSDE with jumps defined by: for $t \geq 0$,

$$\begin{cases} dY_t = -f(t, Y_t, Z_t, R_t, \mathbb{E}[Y_t], \mathbb{E}[Z_t], \mathbb{E}[R_t])dt + Z_t dW_t + \int_{\chi} R_t(e) \tilde{\pi}(de, dt) \\ \lim_{t \rightarrow \infty} Y_t = 0. \end{cases} \quad (2.5.69)$$

has a solution (Y, Z, R) in $L_{\mathbb{F}}^{2, \hat{\rho}}(\Omega \times \mathbb{R}^+, \mathbb{R}^d) \times L_{\mathbb{F}}^{2, \hat{\rho}}(\Omega \times \mathbb{R}^+, \mathbb{R}^d) \times \mathcal{K}^{2, \hat{\rho}}(\Omega \times \mathbb{R}^+ \times \chi, \mathbb{R}^d)$, where $\hat{\rho}$ is a positive constant which will be fixed later.

Existence: Let (Y^n, Z^n, R^n) be a solution on $[0, n]$ of the following BSDE

$$\begin{aligned} Y_t^n &= \int_t^n f^n(s, Y_s^n, Z_s^n, R_s^n, \mathbb{E}[Y_s^n], \mathbb{E}[Z_s^n], \mathbb{E}[R_s^n]) ds - \int_t^n Z_s^n dB_s \\ &\quad - \int_t^n \int_{\chi} R_s^n(e) \tilde{\pi}(de, ds), \quad t \in [0, n], \end{aligned}$$

where

$$\begin{aligned} f^n(t, Y_t^n, Z_t^n, R_t^n, \mathbb{E}[Y_t^n], \mathbb{E}[Z_t^n], \mathbb{E}[R_t^n]) &= ct \mathbf{1}_{[0, n]}(t) + (\rho \mathbb{I}_d + \theta)^\top (Y_t^n - \mathbb{E}[Y_t^n]) \\ + (\rho \mathbb{I}_d + \hat{\theta})^\top \mathbb{E}[Y_t^n] + \vartheta^\top (Z_t^n - \mathbb{E}[Z_t^n]) &+ \hat{\vartheta}^\top \mathbb{E}[\Delta Z_t^n] + \int_{\chi} \varrho^\top(e) (R_t^n(e) - \mathbb{E}[R_t^n(e)]) \lambda(de) \\ + \int_{\chi} \hat{\varrho}^\top(e) \mathbb{E}[R_t^n(e)] \lambda(de), \end{aligned}$$

and we take $(Y_t^n, Z_t^n, R_t^n) = (0, 0, 0)$ on (n, ∞) .

We fix $m > n$. Applying Itô's formula to $e^{\hat{\rho}t} |Y_t^m - Y_t^n|^2$, we get for all $t \geq 0$,

$$\begin{aligned} &e^{\hat{\rho}T} |Y_T^m - Y_T^n|^2 - e^{\hat{\rho}t} |Y_t^m - Y_t^n|^2 \\ &= \int_t^T e^{\hat{\rho}s} \left(\hat{\rho} |Y_s^m - Y_s^n|^2 + |Z_s^m - Z_s^n|^2 + \int_{\chi} |(R_s^m - R_s^n)(e)|^2 \lambda(de) \right) ds \\ &\quad - 2 \int_t^T e^{\hat{\rho}s} (Y_s^m - Y_s^n)^\top \Delta^{n, m} f_s ds \\ &\quad + 2 \int_t^T e^{\hat{\rho}s} (Y_s^m - Y_s^n)^\top (Z_s^m - Z_s^n) dB_s + \int_{\chi} e^{\hat{\rho}s} (Y_s^m - Y_s^n)^\top (R_s^m - R_s^n)(e) \tilde{\pi}(de, ds), \end{aligned}$$

where

$$\Delta^{n, m} f_s := f^m(s, Y_s^m, Z_s^m, R_s^m, \mathbb{E}[Y_s^m], \mathbb{E}[Z_s^m], \mathbb{E}[R_s^m]) - f^n(s, Y_s^n, Z_s^n, R_s^n, \mathbb{E}[Y_s^n], \mathbb{E}[Z_s^n], \mathbb{E}[R_s^n]).$$

We focus on the dependence in Y , we obtain:

$$\begin{aligned}
& e^{\hat{\rho}T} |Y_T^m - Y_T^n|^2 - e^{\hat{\rho}t} |Y_t^m - Y_t^n|^2 \\
= & \int_t^T e^{\hat{\rho}s} \left(\hat{\rho} |Y_s^m - Y_s^n|^2 + |Z_s^m - Z_s^n|^2 + \int_{\mathcal{X}} |(R_s^m - R_s^n)(e)|^2 \lambda(de) \right) ds \\
& - 2 \int_t^T e^{\hat{\rho}s} ((Y_s^m - Y_s^n) - \mathbb{E}[Y_s^m - Y_s^n])^\top (\rho \mathbb{I}_d + \theta) (Y_s^m - Y_s^n) ds \\
& - 2 \int_t^T e^{\hat{\rho}s} \mathbb{E}[Y_s^m - Y_s^n]^\top (\rho \mathbb{I}_d + \hat{\theta}) (Y_s^m - Y_s^n) ds - 2 \int_t^T e^{\hat{\rho}s} (Y_s^m - Y_s^n)^\top \Delta^{n,m} f_s^0 ds \\
& + 2 \int_t^T e^{\hat{\rho}s} (Y_s^m - Y_s^n)^\top (Z_s^m - Z_s^n) dB_s + \int_{\mathcal{X}} e^{\hat{\rho}s} (Y_s^m - Y_s^n)^\top (R_s^m - R_s^n)(e) \tilde{\pi}(de, ds).
\end{aligned}$$

where

$$\Delta^{n,m} f_s^0 := f^m(s, 0, Z_s^m, R_s^m, 0, \mathbb{E}[Z_s^m], \mathbb{E}[R_s^m]) - f^n(s, 0, Z_s^n, R_s^n, 0, \mathbb{E}[Z_s^n], \mathbb{E}[R_s^n]).$$

Taking the expectation, the contribution of the stochastic integrals vanishes. Using the Young's inequality $2ab \leq \epsilon a^2 + \frac{1}{\epsilon} b^2$, where $\epsilon > 0$, and Cauchy Schwarz's inequality, we obtain:

$$\begin{aligned}
& 2\mathbb{E} \left[\int_t^T e^{\hat{\rho}s} ((Y_s^m - Y_s^n) - \mathbb{E}[Y_s^m - Y_s^n])^\top (\rho \mathbb{I}_d + \theta) (Y_s^m - Y_s^n) ds \right] \\
& + 2\mathbb{E} \left[\int_t^T e^{\hat{\rho}s} \mathbb{E}[Y_s^m - Y_s^n]^\top (\rho \mathbb{I}_d + \hat{\theta}) (Y_s^m - Y_s^n) ds \right] \\
\leq & 2\rho \mathbb{E} \left[\int_t^T e^{\hat{\rho}s} |Y_s^m - Y_s^n|^2 \right] ds + 2 \int_t^T e^{\hat{\rho}s} |\theta| E \left[|Y_s^m - Y_s^n|^2 \right] ds \\
& + 2 \int_t^T e^{\hat{\rho}s} |\theta| E \left[|Y_s^m - Y_s^n|^2 \right] ds + 2 \int_t^T e^{\hat{\rho}s} |\hat{\theta}| E \left[|Y_s^m - Y_s^n|^2 \right] ds \\
\leq & 2\rho \mathbb{E} \left[\int_t^T e^{\hat{\rho}s} |Y_s^m - Y_s^n|^2 \right] ds + 2 \int_t^T e^{\hat{\rho}s} (2|\theta| + |\hat{\theta}|) E \left[|Y_s^m - Y_s^n|^2 \right] ds.
\end{aligned}$$

Similar calculus for the term $\int_t^T e^{\hat{\rho}s} (Y_s^m - Y_s^n)^\top \Delta^{n,m} f_s^0 ds$, shows that:

$$\begin{aligned} & \mathbb{E} \left[\int_0^T e^{\hat{\rho}s} \left(\hat{\rho} |Y_s^m - Y_s^n|^2 + |Z_s^m - Z_s^n|^2 + \int_{\mathcal{X}} |(R_s^m - R_s^n)(e)|^2 \lambda(de) \right) ds \right] \\ & \leq \mathbb{E} [e^{\hat{\rho}T} |Y_T^m - Y_T^n|^2] + \left(\delta + \frac{1}{\epsilon} + 2\rho + C \right) \mathbb{E} \left[\int_0^T e^{\hat{\rho}s} |Y_s^m - Y_s^n|^2 ds \right] \\ & + C\epsilon \left(\mathbb{E} \left[\int_0^T e^{\hat{\rho}s} |Z_s^m - Z_s^n|^2 ds \right] + \mathbb{E} \left[\int_0^T \int_{\mathcal{X}} e^{\hat{\rho}s} |(R_s^m - R_s^n)(e)|^2 \lambda(de) ds \right] \right. \\ & \left. + \frac{1}{\delta} \mathbb{E} \left[\int_0^T e^{\hat{\rho}s} |c_s|^2 \mathbf{1}_{[n,m]}(s) ds \right] \right). \end{aligned}$$

where $C := 2|\theta| + |\hat{\theta}| + 2|\vartheta| + |\hat{\vartheta}| + 2\int_{\mathcal{X}} |\varrho(e)|^2 \lambda(de) + \int_{\mathcal{X}} |\hat{\varrho}(e)|^2 \lambda(de)$. Under (H1')(iii), C is finite. By choosing $0 < \epsilon < \frac{1}{2C}$, $\hat{\rho}$ and $\delta > 0$ s.t. $\hat{\rho} > \delta + \frac{1}{\epsilon} + 2\rho + C + \frac{1}{2}$, we deduce:

$$\begin{aligned} & \mathbb{E} \left[\int_0^T e^{\hat{\rho}s} \left(|Y_s^m - Y_s^n|^2 + |Z_s^m - Z_s^n|^2 + \int_{\mathcal{X}} |(R_s^m - R_s^n)(e)|^2 \lambda(de) \right) ds \right] \\ & \leq \frac{2}{\delta} \mathbb{E} \left[\int_0^T e^{\hat{\rho}s} |c_s|^2 \mathbf{1}_{[n,m]}(s) ds \right]. \end{aligned}$$

Sending T to infinity, by the monotone convergence theorem, we obtain:

$$\begin{aligned} & \mathbb{E} \left[\int_0^\infty e^{\hat{\rho}s} \left(|Y_s^m - Y_s^n|^2 + |Z_s^m - Z_s^n|^2 + \int_{\mathcal{X}} |(R_s^m - R_s^n)(e)|^2 \lambda(de) \right) ds \right] \\ & \leq \frac{2}{\delta} \mathbb{E} \left[\int_0^\infty e^{\hat{\rho}s} |c_s|^2 \mathbf{1}_{[n,m]}(s) ds \right]. \end{aligned}$$

As $c \in L_{\mathbb{F}}^{2,\hat{\rho}}(\Omega \times [0, \infty], \mathbb{R}^d)$, then $|c_s|^2 \mathbf{1}_{[n,m]}(s) \rightarrow 0$, $dt \otimes d\mathbb{P}$ a.e., $s \geq 0$ when n goes to infinity. By using the dominated convergence theorem for the right hand side, we deduce that the sequence (Y^n, Z^n, K^n) is a Cauchy sequence in $L_{\mathbb{F}}^{2,\hat{\rho}}(\Omega \times [0, \infty], \mathbb{R}^d) \times L_{\mathbb{F}}^{2,\hat{\rho}}(\Omega \times [0, \infty], \mathbb{R}^d) \times \mathcal{K}^{2,\hat{\rho}}(\Omega \times [0, \infty] \times \mathcal{X}, \mathbb{R}^d)$ and that the limit (Y, Z, K) is a solution of a MF BSDE with jumps (2.5.69).

- Given $(K, \Lambda, Y, Z^Y, R^Y)$, the linear ordinary differential equation for φ

$$d\varphi_t = (\rho\varphi_t - \mathbb{E}[\Gamma_t^0])dt. \quad (2.5.70)$$

where $\Gamma_t^0 = \Gamma_t - (\epsilon_t - \bar{\epsilon}_t)^\top S_t^{-1} (\epsilon_t - \bar{\epsilon}_t) - \Theta_t^\top \hat{S}_t^{-1} \Theta_t$ is defined in (2.3.26), admits a unique explicit solution given by: $\varphi_t = \int_t^\infty e^{-\rho(s-t)} \mathbb{E}[\Gamma_s^0] ds$.

Then, we proved the existence of solutions of the decoupled system (2.5.63).

The following theorem gives the structure of the optimal control for LQMF problem (2.5.57).

Theorem 2.5.4. *Under Assumptions (H1')-(H2'), the optimal control for LQMF problem (2.5.57) is given by*

$$\alpha_t^* = -S^{-1}U(X_{t-}^* - \bar{X}_{t-}^*) - S^{-1}(\epsilon_{t-} - \bar{\epsilon}_{t-}) - S^{-1}(V\bar{X}_{t-}^* + \Theta_t),$$

where $X^* = X^{\alpha^*}$ is the state process where the α^* is the optimal control and the deterministic coefficients S, \hat{S}, U and V and the stochastic coefficients ϵ_t and Θ_t are defined in Section 2.3.

2.6 Application to production of an exhaustible resource

In this section, we study a model of production of exhaustible resource with accumulating or maintaining a level of reserves, inspired by a series of works extended from the Hotelling's model [63]. In the classic Hotelling's model, the dynamics market's evolution is driven by the use of existing reserves of an exhaustible reserves to produce energy without possibility to exploration and/or discovery of new reserves. But many studies have made it possible to ensure that there are still resources to be explored over time, that is to say that the reservation rate can be increased. We can refer to the series of works extended from Prindyck's model [107], Deshmukh et al. [41], Arrow and Chang [9], and Keller et al. [60]. The increase in reserve discoveries occurs stochastically via the Poisson process. It should be noted that this increase is smaller, and it is the reason that the resources always remain exhaustible.

We consider an energy market with N producers (players). Each producer uses exhaustible resources, such as oil, to produce energy. The quantity X_t^i represents the reserve's level of player i , at time t , $i = 1, \dots, N$. It takes values in the set \mathbb{R}^+ . The reserve level X_t^i decreases at a controlled production rate $\alpha_t^i \geq 0$ $dt \otimes dP$ a.e., and also has random discrete increment due to exploration. We use N independent Poisson point process π^i , $i = 1 \dots N$ to model the new discoveries and we denote by $\lambda^i(de)dt$ the associated compensator. We assume that the dynamics of the reserve has a noise which is proportional to the current level of the reserve. The reserve's dynamics of each producer i is given by the following stochastic differential equation:

$$\begin{cases} dX_t^i &= -\alpha_t^i dt + \sigma X_t^i dW_t^i + \eta \int_{\mathcal{X}} X_{t-}^i \tilde{\pi}^i(de, dt), \\ X_0^i &= x_0^i, \end{cases} \quad (2.6.71)$$

where x_0^i is the initial reserve's level of player i , $\sigma > 0$, W^i , $i = 1 \dots N$ are independent standard Brownian motion and independent of π^i , $i = 1 \dots N$, and $\eta > 0$ is the rate of new discoveries.

The cost functional for producer i is given by:

$$J^i(\alpha^1, \dots, \alpha^n) := \mathbb{E} \left(\int_0^\infty e^{-\rho t} \left[-\alpha_t^i P_t(\alpha_t^i) + C_p(\alpha_t^i) + C_{ex}(\alpha_t^i, X_t^i) \right] dt \right), \quad (2.6.72)$$

where

- P_t^i is the selling price for producer i . It follows a linear inverse demand rule, defined as:

$$P_t^i := P_t(\alpha_t^i) = P_t^0 - \delta \alpha_t^i - \varepsilon \int_0^t \frac{1}{N} \sum_{j=1}^N \alpha_s^j ds.$$

Here δ, ε are positive constants and P^0 is a deterministic function. It will be the same for all producers. The price P^i of producer i is related to his production and also to the production of all other producers.

- C_p is the cost functions of production, defined as:

$$C_p(\alpha_t^i) := c_1 \text{Var}(\alpha_t^i).$$

- C_{ex} is the cost functions of exploration defined as:

$$C_{\text{ex}}(\alpha_t^i, X_t^i) := c_2 \alpha_t^i \left(\frac{x_0^i - X_t^i}{x_0^i} \right).$$

The constants c_1 and c_2 are positive and represent respectively the cost of production and the cost of extraction. They are the same for the all producers. From the theory of propagation of chaos, the individual level of reserve X^i and the price process P^i , $i = 1, \dots, N$, become independent and identically distributed, when N goes to infinity, with a common distribution given by the law of the solution (X, P) to the stochastic McKean-Vlasov equation

$$dX_t = -\alpha_t dt + \sigma X_t dW_t + \eta \int_{\mathcal{X}} X_t - \tilde{\pi}(de, dt), X_0 = x_0, \quad (2.6.73)$$

$$P_t = P_t^0 - \delta \alpha_t - \varepsilon \int_0^t \bar{\alpha}_s ds, \quad (2.6.74)$$

where W is a Brownian motion and $\alpha_t = a(t, X_t)$, $t \geq 0$, for some measurable function a on $\mathbb{R}_+ \times \mathbb{R}$. We reduce the problem to a representative producer with initial reserve $x_0 > 0$. The state process is given X (see (2.6.73)). The control processes are given by (α, P) , where P satisfies (2.6.74). The aim of the representative producer is to minimise the cost functional given by:

$$J(\alpha) := \mathbb{E} \left[\int_0^\infty e^{-\rho t} \alpha_t \left(-P_t^0 + \delta \alpha_t + \varepsilon \int_0^t \bar{\alpha}_s ds \right) + c_1 \text{Var}(\alpha_t) + c_2 \alpha_t \left(\frac{x_0 - X_t}{x_0} \right) dt \right], \quad (2.6.75)$$

under the constraints that $\alpha_t \geq 0$ and $X_t \geq 0$ \mathbb{P} a.s. for all $t \geq 0$. As $\bar{X}_t = x_0 - \int_0^t \bar{\alpha}_s ds$, then

$$J(\alpha) = \mathbb{E} \left[\int_0^\infty e^{-\rho t} \alpha_t (-P_t^0 + \delta \alpha_t + \varepsilon(x_0 - \bar{X}_t)) + c_1 \text{Var}(\alpha_t) + c_2 \alpha_t \left(\frac{x_0 - X_t}{x_0} \right) dt \right],$$

and we are in the framework of Section 2.5 with $d = m = 1$ (one-dimensional state variable and control), the coefficients of the state process and the cost functional are given by

$$B = -1, \quad C = \sigma, \quad F(e) = \eta, \quad \text{for all } e \in \chi,$$

and

$$N + \tilde{N} = \delta, \quad N = \delta + c_1, \quad I + \tilde{I} = -\frac{c_2 + \varepsilon x_0}{2x_0}, \quad I = -\frac{c_2}{2x_0}, \quad H_t = \frac{c_2 + \varepsilon x_0 - P_t^0}{2},$$

where the other coefficient are equal to zero. We define $\lambda(\chi) := \int_\chi \lambda(de)$. Notice that under the assumption $\rho > \sigma^2 + \lambda(\chi)\eta^2$, (H1') and (H3') are satisfied. By following the approach developed in section 2.5, the optimal control is given explicitly. We have to solve the decoupled system of Riccati equations and BSDEs with jumps (2.5.63). The Riccati equations (2.5.64) for K and Λ (2.5.67) are given by:

$$\frac{\left(K + \frac{c_2}{2x_0}\right)^2}{\delta + c_1} + (\rho - \sigma^2 - \lambda(\chi)\eta^2)K = 0, \quad (2.6.76)$$

$$\frac{\left(\Lambda + \frac{c_2 + \varepsilon x_0}{2x_0}\right)^2}{\delta} + \rho\Lambda - (\sigma^2 + \lambda(\chi)\eta^2)K = 0. \quad (2.6.77)$$

Let us also remark that the condition (H2') is not satisfied, but we have the existence of a solution (K, Λ) to (2.6) such that $K_{c_1, c_2} := \frac{K + \frac{c_2}{2x_0}}{\delta + c_1} > 0$ and $\Lambda_\varepsilon := \frac{\Lambda + \frac{c_2 + \varepsilon x_0}{2x_0}}{\delta} > 0$, and given by:

$$K_{c_1, c_2} = \frac{-(\rho - \sigma^2 - \lambda(\chi)\eta^2) + \sqrt{(\rho - \sigma^2 - \lambda(\chi)\eta^2)^2 + 2c_2 \frac{\rho - \sigma^2 - \lambda(\chi)\eta^2}{x_0(\delta + c_1)}}}{2} > 0,$$

and

$$\Lambda_\varepsilon = \frac{-\rho + \sqrt{\rho^2 + 2 \frac{\rho(c_2 + \varepsilon x_0) + 2(\sigma^2 + \lambda(\chi)\eta^2)K}{\delta x_0}}}{2} > 0.$$

Therefore, we can write the linear BSDE (2.5.69) with jumps as:

$$\begin{aligned} -dY_t &= \left(\frac{\Lambda_\varepsilon}{2} (c_2 + \varepsilon x_0 - P_t^0) - (\rho + K_{c_1, c_2})Y_t + (K_{c_1, c_2} - \Lambda_\varepsilon)\bar{Y}_t \right. \\ &\quad \left. - \sigma Z_t^Y + \eta \int_\chi R_t^Y(e) \lambda(de) \right) dt - Z_t^Y dW_t - \int_\chi R_t^Y(e) \tilde{\pi}(de, dt). \end{aligned} \quad (2.6.78)$$

One could check that a solution of the BSDE (2.6.78) is given by:

$$(Y, Z^Y, R^Y) = \left(\int_t^\infty e^{-(\rho+\Lambda_\varepsilon)(s-t)} \left(\Lambda_\varepsilon \frac{c_2 + \varepsilon x_0 - P_s^0}{2} \right) ds, 0, 0 \right)_{0 \leq t \leq T}. \quad (2.6.79)$$

In the remaining part of the paper, we assume that $P_t^0 = p^0$ for all $t \geq 0$: p^0 is interpreted as a substitute price for the exhaustible resource. We study two cases. The first one, when $p^0 = c_2 + \varepsilon x_0$ i.e. p^0 coincides with $c_2 + \varepsilon x_0$ which is the cost of extraction for the last unit of resource. In other words, the Hotelling rent $H_r := p^0 - c_2 - \varepsilon x_0$ is equal to zero. The second case when $p^0 < c_2 + \varepsilon x_0$ i.e. the Hotelling rent is negative. The next proposition gives an explicit solution to the problem (2.6.75) when $H_r = 0$.

Proposition 2.6.1. *We assume that $p^0 = c_2 + \varepsilon x_0$ for all $t \geq 0$, x_0 is large enough and $1 > \rho^2 + \frac{2\varepsilon}{\delta}$. Then the solution of (2.6.75) is given by*

$$\alpha_t^* = K_{c_1, c_2} X_{t-}^* + (\Lambda_\varepsilon - K_{c_1, c_2}) \bar{X}_{t-}^*.$$

Proof. Since (H1') and (H3') are satisfied and the Riccati equations have a solution, then, by Theorem (2.5.4), the optimal control is then given by:

$$\alpha_t^* = K_{c_1, c_2} (X_{t-}^* - \bar{X}_{t-}^*) + \Lambda_\varepsilon \bar{X}_{t-}^* - \frac{1}{2\delta} (c_2 + \varepsilon x_0 - P_t^0 - Y_t).$$

As $P_t^0 = p^0 = c_2 + \varepsilon x_0$ for all $t \geq 0$, then the solution of the BSDE (2.6.79) satisfies $Y_t = 0$ for all $t \geq 0$ which yields:

$$\alpha_t^* = K_{c_1, c_2} (X_{t-}^* - \bar{X}_{t-}^*) + \Lambda_\varepsilon \bar{X}_{t-}^*.$$

It remains to show that the optimal strategy satisfies the constraint $\alpha_t^* \geq 0$ \mathbb{P} a.s. for all $t \geq 0$. As x_0 is large, by using Taylor's formula, we have

$$\begin{aligned} 2 \quad (\Lambda_\varepsilon - K_{c_1, c_2}) &= -\rho + \sqrt{\rho^2 + \frac{2\varepsilon}{\delta}} \left(1 + \frac{\rho c_2 + 2(\sigma^2 + \lambda(\chi)\eta^2)K}{\delta x_0 (\rho^2 + \frac{2\varepsilon}{\delta})} \right) - \frac{c_2 \rho}{x_0 (\delta + c_1)} + o\left(\frac{1}{x_0}\right) \\ &= -\rho + \sqrt{\rho^2 + \frac{2\varepsilon}{\delta}} + \frac{c_2 \rho}{x_0} \left(\frac{1}{\delta \sqrt{\rho^2 + \frac{2\varepsilon}{\delta}}} - \frac{1}{\delta + c_1} \right) + \frac{2(\sigma^2 + \lambda(\chi)\eta^2)K}{\delta x_0 \sqrt{\rho^2 + \frac{2\varepsilon}{\delta}}} + o\left(\frac{1}{x_0}\right). \end{aligned}$$

As $1 > \rho^2 + \frac{2\varepsilon}{\delta}$, then

$$\Lambda_\varepsilon - K_{c_1, c_2} \geq 0. \quad (2.6.80)$$

We define the stopping time τ^* as follows:

$$\tau^* := \inf\{t \geq 0 \text{ s.t. } X_t^* \leq 0\}.$$

Then on the set $\{t < \tau^*\}$, from inequality (2.6.80), we have $\alpha_t^* \geq 0$ \mathbb{P} a.s. On the set $\{t = \tau^*\}$, the state process $X_{\tau^*}^* = 0$, which implies $\alpha_{\tau^*}^* = 0$. Since the drift, the diffusion and the jump terms of the state process are equal to zero, then the process X^* remains at the level 0 for all $t \geq \tau^*$ and the optimal strategy α^* is the null strategy for all $t \geq \tau^*$.

In the second case, we assume that $p^0 < c_2 + \epsilon x_0$. It is not obvious to check the positivity of the the state process and the optimal strategy. We study the stationary level of the reserve and the optimal production rate in mean. From the definition of X^* , we have

$$\begin{aligned} \bar{X}_t^* &= x_0 - \int_0^t \alpha_s^* ds = x_0 - \int_0^t \Lambda_\epsilon \bar{X}_s^* ds \\ &+ \int_0^t \frac{1}{2\delta} \left(c_2 + \epsilon x_0 - p^0 - \int_s^\infty e^{-(\rho + \Lambda_\epsilon)(u-s)} \left(\Lambda_\epsilon \frac{c_2 + \epsilon x_0 - p^0}{2} \right) du \right) ds \end{aligned}$$

which implies

$$\begin{aligned} \bar{X}_t^* &= x_0 + \frac{(2\rho + \Lambda_\epsilon)(c_2 + \epsilon x_0)}{2\delta} \frac{1 - e^{-\Lambda_\epsilon t}}{2\Lambda_\epsilon(\rho + \Lambda_\epsilon)} \\ &- \frac{p^0}{2\delta} \int_0^t e^{-\Lambda_\epsilon(t-s)} \left(1 - \int_s^\infty \frac{\Lambda_\epsilon}{2} e^{(\rho + \Lambda_\epsilon)(u-s)} du \right) ds. \end{aligned}$$

It yields that $\lim_{t \rightarrow \infty} \bar{X}_t^* = \frac{(2\rho + \Lambda_\epsilon)(c_2 + \epsilon x_0 - p^0)}{4\delta\Lambda_\epsilon(\rho + \Lambda_\epsilon)}$. As the Hotelling rent $H_r := p^0 - c_2 - \epsilon x_0$ is negative, then $\lim_{t \rightarrow \infty} \bar{X}_t^*$ exists and is positive. As $\lim_{t \rightarrow \infty} \bar{X}_t^* = x_0 - \int_0^\infty \bar{\alpha}_s^* ds$, then $\lim_{t \rightarrow \infty} \bar{\alpha}_t^* = 0$. It means that when we switch to substitute good, there is a remaining resource and we stop the production of exhaustible resource.

NUMERICAL APPROACH FOR A CLASS OF MEAN FIELD GAMES

3.1 Introduction

In this work, we want to apply this method for a MFG problem where the drift depends not only on α but also on X . Therefore, our mean field game problem is written in the following form:

$$\inf_{\alpha} J(x, t, \alpha, m) := \inf_{\alpha} \mathbb{E} \left[\int_t^T f(X_s^{\alpha}, \alpha_s, m_s) ds + \Phi(X_T^{\alpha}, m_T) \right], \quad (3.1.1)$$

subject to

$$\begin{cases} dX_t^{\alpha} &= b(t, X_t^{\alpha}, \alpha_t) dt + \sigma dB_t \\ X_0^{\alpha} &\sim m_0. \end{cases}$$

In this work, we introduce the Markov chain approximation approach to solve a mean field game problem. This approach was inspired from the work of Kushner and Dupuis [70]. It's based on the approximation of the stochastic control problem by a control problem for a discrete time and discrete space controlled Markov chain. In recent work, E. Bayraktar, A. Budhiraja, and A. Cohen [12] used the Markov chain approximation method to construct an approximation for the MFG with reflecting barriers. They formulate the MFG model in terms of a controlled reflected diffusion with a cost function and study the convergence of their numerical scheme is established for a small time horizon T . The idea is that for a fixed measure on the path space, we define a Markov decision problem after a time and space discretization. Roughly speaking, the solution of the discrete MFG considered can be seen as the solution of a fixed point problem on the space of probability measures on certain path space. Then the discrete scheme is interpreted as a formulation of a discrete mean field game where the controlled state-variable is described by a Markov chain chosen to be consistent with the controlled stochastic differential equation associated to the original problem, [70]. The Markov chain corresponds to an optimally controlled one by some representative agent for which we define the discrete control problem. Existence and uniqueness of solution of the discrete MFG equilibrium is proved using an attractive probabilistic formulation. This work is concertized in the preprint [14].

3.2 Framework: Notations and setting

Let us consider a mean field game model where each players i controls its private state X_t^i at time t , by taking an action α_t in a closed convex subset \mathbb{A} , where the dynamic of state is driven by the following the stochastic differential equation

$$\begin{cases} dX_t^i &= b(t, X_t^i, \alpha_t)dt + \sigma dB_t, \\ X_0^i &\sim m_0, \end{cases} \quad (3.2.2)$$

where B_t is a standard Brownian motion, σ is a fixed positive constant and \mathcal{A} is the set of admissible control,

$$\mathcal{A} := \left\{ \alpha : \Omega \times [0, T] \rightarrow \mathbb{A} \text{ s.t } \alpha \text{ is } \mathbb{F}\text{-adapted and } \int_0^t |\alpha_t| dt < \infty \right\}.$$

We consider \mathcal{K} a non empty compact of \mathbb{R}^d and let us assume that the function b satisfy the following assumptions 3.2.1 :

Assumption 3.2.1. *The function b is defined on $[0, T] \times \mathcal{K} \times \mathcal{A}$ in \mathbb{R} as follows*

$$\begin{aligned} b &: [0, T] \times \mathcal{K} \times \mathcal{A} &\rightarrow & \mathbb{R} \\ (t, x, \alpha) &&\mapsto & b(t, x, \alpha). \end{aligned} \quad (3.2.3)$$

We assume that the function has the linear form : $b(t, x, \alpha) = b_1(t, x) + b_2(t)\alpha$, where

(b_{h1}) *The mapping b_1 is measurable and bounded function on $[0, T] \times \mathcal{K}$.*

(b_{h2}) *The mapping b_2 is measurable and bounded function .*

Each player will need to choose his strategy to minimize his cost functional J over the period $[0, T]$, which is influenced by the state of other players. The problem of differential games with an infinite number of agents is articulated in the following form:

- (i) for each fixed flow of probability measures m on \mathbb{R}^d , solve the standard stochastic control problem:

$$u(x, t) := \inf_{\alpha \in \mathcal{A}} J(t, x, \alpha, m), \quad (3.2.4)$$

subject to

$$\begin{cases} dX_t^\alpha &= b(t, X_t^\alpha, \alpha_t)dt + \sigma dB_t \\ X_0^\alpha &\sim m_0. \end{cases}$$

- (ii) Find m such that $\mathcal{L}(X_t^\alpha) = m_t$ for all $t \in [0, T]$, where X_t^α is a solution of the above stochastic differential equation.

We can see that the first step provides the best response u of a given player interacting with the stochastical distribution of the states of the other players. If this distribution m is assumed to be given, when the second step solves a fixed point problem in the goal of the search for fixed points m_t of the best response function. Once these two steps have been taken successfully, if the fixed-point optimal control α^* identified in step (i) is in feedback form $\alpha_t^* = \phi(t, X_t^\alpha, m_t)$ for some deterministic function ϕ on $[0, T] \times \mathbb{T}^d \times \mathcal{P}(\mathcal{K})$. Then if the players use α_t^* , the mean field equilibrium $\{m_t\}_{0 \leq t \leq T}$ with value in $\mathcal{P}(\mathcal{K})$ should coincide with the law of X and satisfy the Fokker Planck equation.

Definition 3.2.1. (MFG equilibrium)

We say that $m^* \in \mathcal{P}(\mathcal{K})$ is an equilibrium for the MFG associated to the initial probability density m_0 if and only if there exist $\alpha^* \in \mathcal{A}$ such that

$$J(t, X_t^\alpha, \alpha^*, m^*) = \inf_{\alpha \in \mathcal{A}} J(t, X_t^\alpha, \alpha, m^*), \quad (3.2.5)$$

$$\text{and } m_t^* = \mathcal{L}(X_t^\alpha), \quad \forall t \in [0, T]. \quad (3.2.6)$$

The objective is to minimize a cost function of the following form :

$$J(x, t, \alpha, m) = \mathbb{E} \left[\int_t^T (f(X_s^\alpha, \alpha_s, m_s) ds + \Phi(X_T^\alpha, m_T)) \right], \quad (3.2.7)$$

where f is a measurable function defined in this form that we will adapt for the rest of the paper:

$$f(X_s^\alpha, \alpha_s, m_s) := L(X_s^\alpha, \alpha_s) + F(X_s^\alpha, m_s).$$

with F and L are a measurable functions satisfied assumptions (3.2.2)-(3.2.3) respectively introduced below:

We introduce first the Wasserstein distance \mathcal{W}_1 as

$$\mathcal{W}_1(\nu, \nu') = \min \left\{ \int |x - y| d\mu \text{ such that } \mu \in \Pi(\nu, \nu') \right\},$$

where

$$\Pi(\nu, \nu') := \left\{ \pi \in \mathcal{P}(\mathcal{K} \times \mathcal{K}), \pi(A \times \mathcal{K}) = \nu(A), \pi(\mathcal{K} \times A) = \nu'(A) \text{ for each } A \in \mathcal{B}(\mathcal{K}) \right\}.$$

Assumption 3.2.2. The function F is defined on $\mathcal{P} \times \mathcal{K}$ in \mathbb{R} as:

$$\begin{aligned} F & : \mathcal{P} \times \mathcal{K} \rightarrow \mathbb{R} \\ (m, x) & \mapsto F(m, x). \end{aligned}$$

(\mathbf{F}_{h1}) F is K_F -Lipschitz and K_F is independent of m , i.e.

$$|F(m, x) - F(m, y)| \leq K_F |x - y|, \quad \forall m \in \mathcal{P} \text{ and } (x, y) \in \mathcal{K} \times \mathcal{K}.$$

(\mathbf{F}_{h2}) F is Lipschitz from the space \mathcal{P} endowed with \mathcal{W}_1 to the space $\mathcal{C}(\mathcal{K})$, endowed with the uniform norm, i.e. there exist $C_F > 0$, such that

$$\|F(m_1, x) - F(m_2, x)\| \leq C_F \mathcal{W}_1(m_1, m_2), \quad \forall m_1, m_2 \in \mathcal{P}(\mathcal{K}) \text{ and } x \in \mathcal{K}.$$

(\mathbf{F}_{h3}) $\partial_x F$ exists and is $C_{\partial F}$ -Lipschitz from the space $\mathcal{P}(\mathcal{K})$ endowed with \mathcal{W}_1 to the space $\mathcal{C}(\mathcal{K})$, endowed with the uniform norm. i.e. there exist

$L_{\partial F} > 0$, such that

$$\|\partial_x F(m_1, x) - \partial_x F(m_2, x)\| \leq C_{\partial F} \mathcal{W}_1(m_1, m_2), \quad \forall m_1, m_2 \in \mathcal{P}(\mathcal{K}) \text{ and } x \in \mathcal{K}.$$

Assumption 3.2.3. The function L is defined on $\mathcal{K} \times \mathcal{A}$ in \mathbb{R} as follows

$$\begin{aligned} L & : \mathcal{K} \times \mathcal{A} \rightarrow \mathbb{R} \\ (x, \alpha) & \mapsto L(x, \alpha). \end{aligned}$$

(\mathbf{L}_{h1}) L is K_L -Lipschitz and K_L is independent of α , i.e.

$$|L(x_1, \alpha) - L(x_2, \alpha)| \leq K_L |x_1 - x_2|, \quad \forall (x_1, x_2) \in \mathcal{K} \times \mathcal{K} \text{ and } \alpha \in \mathcal{A}.$$

(\mathbf{L}_{h2}) L is strictly convex with respect to α , i.e.

$$\partial_\alpha L(x, \alpha_1)(\alpha_2 - \alpha_1) \leq L(x, \alpha_2) - L(x, \alpha_1), \quad \forall \alpha_1, \alpha_2 \in \mathcal{A} \text{ and } x \in \mathcal{K}.$$

(\mathbf{L}_{h3}) For all $x \in \mathcal{K}$, the function $\alpha \mapsto \partial_\alpha L(x, \alpha)$ is continuous function.

Assumption 3.2.4. The function Φ is defined on $\mathcal{K} \times \mathcal{P}$ in \mathbb{R} as:

$$\begin{aligned} \Phi & : \mathcal{P} \times \mathcal{K} \rightarrow \mathbb{R} \\ (m, x) & \mapsto \Phi(m, x). \end{aligned}$$

(Φ_{h1}) Φ is K_Φ -Lipschitz and K_Φ is independent of m ,

$$|\Phi(x, m) - \Phi(y, m)| \leq K_\Phi |x - y|, \quad \forall m \in \mathcal{P}(\mathcal{K}) \text{ and } (x, y) \in \mathcal{K} \times \mathcal{K}.$$

(Φ_{h2}) Φ is C_Φ -Lipschitz from the space \mathcal{P} endowed with \mathcal{W}_1 to the space of continuous functions $\mathcal{C}(K)$, endowed with the uniform norm .i.e. there exists $C_\Phi > 0$, such that $\forall m_1, m_2 \in \mathcal{P}$

$$\|\Phi(x, m_1) - \Phi(x, m_2)\| \leq C_\Phi \mathcal{W}_1(m_1, m_2).$$

(Φ_{h3}) $\partial_x \phi[m]$ exists and is $C_{\partial\phi}$ -Lipschitz on \mathcal{P} space endowed with \mathcal{W}_1 to the space $\mathcal{C}(\mathcal{K})$, endowed with the uniform norm .i.e. there exists $C_{\partial\phi} > 0$, such that

$$\|\partial_x \phi(x, m_1) - \partial_x \phi(x, m_2)\| \leq C_{\partial\phi} \mathcal{W}_1(m_1, m_2), \quad \forall m_1, m_2 \in \mathcal{P}(\mathcal{K}) \text{ and } x \in \mathcal{K}.$$

According to [27], the control problem is mathematically described through a system of two non linear partial differential equations introduced by a following system

$$\begin{cases} -\partial_t u - \nu \Delta u + H(x, m, Du) = f(x, m), & \text{in } \mathbb{T}^d \times (0, T), \\ \partial_t m - \nu \Delta m - \operatorname{div}(D_p H(x, m, Du)m) = 0, & \text{in } \mathbb{T}^d \times (0, T), \\ m(0) = m_0, \\ u(x, T) = \phi(x, m(T)), \end{cases} \quad (3.2.8)$$

where σ is a non-negative given parameter, $T > 0$ is a fixed time horizon, \mathcal{K} is a compact subset of \mathbb{R} , \mathcal{P} is the space probability measures on \mathcal{K} and m_0 is a measure on $\mathcal{P}(\mathcal{K})$. $H : (x, p) \in \mathcal{K} \times \mathbb{R} \rightarrow H(x, p) \in \mathbb{R}$ is the Hamiltonian, a convex function with respect to p defined by

$$H(x, m, p) = \sup_{\alpha} \{-p\alpha - L(x, \alpha)\}.$$

Then we will introduce a numerical method to solve the coupled system of forward-backward partial differential equations, with variables u and m .

We will refer to the system forward-backward PDEs system (3.2.8) as the mean field game PDEs system (MF PDE in short). The first equation is backward in time. It is called a Hamilton-Jacobi equation where is associated with an optimal control problem. This equation describes the reaction of players to the mass. It presents the decisions based on where you want to be in the future. The second one is forward in time. This is called a transport or Fokker-Planck equation for the distribution of the agents. The equation describes the aggregation of the action of all players. It presents where the population behavior actually ends up based on the initial distribution. The existence of solutions of HJB-FP system (3.2.8) is treated in many works

[16, 25, 26, 32, 33, 58, 74]. A uniqueness criterion for the solution of the field game system was presented by J.M. Lasry and P.L. Lions in their pioneer work [74] on the theory of mean field games called Lasry-Lions Monotonicity condition.

3.3 Numerical Approximation

In this section, we propose a numerical method to approximate mean field control problem. We introduce Markov chain Approximation developed by Kushner [70]. The approximating Markov chain is chosen such that certain local properties of the approximating chain are "similar" to those of the original controlled process. A cost function for the Markov chain model which is an appropriate analogue to that for the original model is then found.

This method applies a Markov chain approximation to continuous time, continuous state stochastic control problems by renormalizing finite differences forms as proper Markov chain transition probabilities. These transition probabilities arise when deriving finite difference versions of the dynamic programming equation. An important advantage of this method is that the Markov chain approximation facilitates convergence proofs for the numerical methods in terms of probabilistic argument.

3.3.1 Approximating Markov chain

The idea behind this approach is to construct a discrete time space Markov chain ξ^h and its state space. This chain admits a continuous interpolation which approximates the state process X_t . Roughly speaking, we define a Markov chain ξ^h , on the state space where his transition probabilities denoted by $p(t, \nu, \alpha; x, x + h)$. These transition probabilities must be positive $p > 0$ and satisfies

$$\sum_{y \in \{x \pm h\}} p(t, \nu, \alpha; x, y) = 1.$$

According to the construction of a Markov chain, its probability measure ν_n at time n is defined by

$$\nu_n = \nu_{n+1} p(t, \nu, \alpha; x, y).$$

By iterating this equality, we obtain for any $n \geq 1$,

$$\nu_n = (p(t, \nu, \alpha; x, y))^n \nu_0,$$

where $(p)^n$ is the transition matrix and ν_0 is the initial distribution.

3.3.2 Discretization setting

In this part, we consider T a terminal time $T > 0$, and $\mathcal{X} = [-R, R]$ with $R > 0$. We define $h = \frac{2R}{N_x^h}$ and $\Delta t = \frac{T}{N_t^h}$ where N_x^h and N_t^h as the mesh step and the time step for our discrete scheme, for two positives integers N_x^h and N_t^h . Denote $x_i := -R + ih$, $t_k := k\Delta t$, we consider the space grid \mathcal{G}_h and the time-space grid $\mathcal{G}_{h,dt}$ defined by

$$\mathcal{G}_h := \{x_i, 0 \leq i \leq N_x^h\} \quad \text{and} \quad \mathcal{G}_{h,dt} = \{(x_i, t_k), 0 \leq i \leq N_x^h, 0 \leq k \leq N_t^h\}.$$

We denote the approximation of u and m respectively by vectors U and M such for each (i, k) ,

$$u(t_k, x_i) \approx U_{i,k}, \quad \text{and} \quad M(t_k, x_i) \approx M_{i,k}^h.$$

Then for the discretization of the HJB equation, we begin by introducing the finite difference schemes. So for the approximation of $\partial_x u$ at $(x_i, t_k) \in \mathcal{G}_{h,dt}$ we define the following discrete operators:

$$DU_{i,k} := \frac{U_{i+1,k} - U_{i-1,k}}{2h}, \quad \text{for } 1 \leq i \leq N_x^h - 1, \quad (3.3.9)$$

For the approximation of Δu we will use the standard operator

$$\Delta_h U_{i,k} = \frac{U_{i+1,k} - 2U_{i,k} + U_{i-1,k}}{h^2}, \quad \text{for } 1 \leq i \leq N_x^h - 1, \quad (3.3.10)$$

and for approximate $\partial_t u$ at $(x_i, t_k) \in \mathcal{G}_{h,dt}$ we introduce the following operator

$$\mathcal{D}U_{i,k} := \frac{U_{i,k} - U_{i,k-1}}{\Delta t}, \quad \text{for } 1 \leq k \leq N_t^h - 1. \quad (3.3.11)$$

Let us consider also L^h , F^h and Φ^h be approximated operators respectively of L , F and Φ as:

$$L^h : \mathcal{G}_h \times \mathcal{A}_h \rightarrow \mathbb{R}, \quad F^h : \mathcal{P}(\mathcal{G}_h) \times \mathcal{G}_h \rightarrow \mathbb{R}, \quad \Phi^h : \mathcal{P}(\mathcal{G}_h) \times \mathcal{G}_h \rightarrow \mathbb{R},$$

where \mathcal{A}^h is the set of control α^h .

If $\xi_{k'}^h = x_i$, $\forall M^h \in \mathcal{P}(\mathcal{G}_{h,dt})$, $\forall \alpha^h \in \mathcal{A}^h$

$$[L^h(\alpha^h)]_{i,k'} := L(x_i, \alpha_{i,k'}^h),$$

$$[F^h(M^h)]_{i,k'} := F(x_i, M_{*k'}^h),$$

$$[\Phi^h(M^h)]_{i,k'} := \Phi(x_i, M_{t_{k'}}^h).$$

3.3.3 Construction of Markov chain associated to MFG problem

Discret problem Using the implicit scheme (3.3.11) for $\partial_t u$, the explicit scheme (3.3.10) and the centered scheme (3.3.9), we discretize the HJB equation in (3.2.8) as

$$-\left(\frac{U_{i,k+1} - U_{i,k}}{\Delta t}\right) - \frac{1}{2}\sigma^2\left(\frac{U_{i+1,k} - 2U_{i,k} + U_{i-1,k}}{h^2}\right) + H\left(x, \frac{U_{i+1,k} - U_{i-1,k}}{2h}\right) = F(x_i, M_{i,k}^h), \quad (3.3.12)$$

with

$$H\left(x_i, \frac{U_{i+1,k} - U_{i-1,k}}{2h}\right) = \inf_{\alpha} \{b(t_k, x_i, \alpha_{i,k}) \cdot \left(\frac{U_{i+1,k} - U_{i-1,k}}{2h}\right) + L(x_i, \alpha_{i,k})\}. \quad (3.3.13)$$

Therefore, it leads to the following discretized form of HJB equation for $(x, t) \in \mathcal{G}$:

$$\begin{cases} U_{i,k-1} = \inf_{\alpha \in \mathcal{A}^h} \left[\left(L^h(x_i, \alpha_{i,k}) + F^h(x_i, M_{i,k}^h) \right) \Delta t + U_{i,k} \left(1 - \sigma^2 \frac{\Delta t}{h^2} \right) \right. \\ \quad \left. + U_{i-1,k} \left(\frac{\sigma^2 \Delta t}{2 h^2} - \frac{\Delta t}{2h} b(t_k, x_i, \alpha_{i,k}) \right) + U_{i+1,k} \left(\frac{\sigma^2 \Delta t}{2 h^2} + \frac{\Delta t}{2h} b(t_k, x_i, \alpha_{i,k}) \right) \right], \\ U_{i,N_t^h} = \Phi^h(x_i, M_{i,N_t^h}^h), \quad 0 \leq i \leq N_x^h, \end{cases} \quad (3.3.14)$$

and for $x \in \partial(\mathcal{G} \times [0, T])$, discretized form of HJB equation satisfies the following boundary conditions:

$$\begin{cases} U_{0,k-1} = \left(L(x_0, \alpha_{0,k}) + F^h(x_0, M_{i,k}^h) \right) \Delta t + U_{0,k} \left(1 - \frac{\sigma^2 \Delta t}{2 h^2} + \frac{\Delta t}{2h} b(t_k, x_0, \alpha_{0,k}) \right) \\ \quad + U_{1,k} \left(\frac{\sigma^2 \Delta t}{2 h^2} - \frac{\Delta t}{2h} b(t_k, x_0, \alpha_{0,k}) \right), \\ U_{N_x^h, k-1} = \left(L(x_{N_x^h}, \alpha_{N_x^h, k}) + F^h(x_{N_x^h}, M_{i,k}^h) \right) \Delta t + U_{N_x^h, k} \left(1 - \frac{\sigma^2 \Delta t}{2 h^2} + \frac{\Delta t}{2h} b(t_k, x_{N_x^h}, \alpha_{N_x^h, k}) \right) \\ \quad + U_{N_x^h-1, k} \left(1 - \frac{\sigma^2 \Delta t}{2 h^2} - \frac{\Delta t}{2h} b(t_k, x_{N_x^h}, \alpha_{N_x^h, k}) \right). \end{cases} \quad (3.3.15)$$

We assume also the following boundary conditions

$$U_{-1,k} = U_{0,k}, \quad U_{N_x^h+1,k} = U_{N_x^h,k}, \quad 1 \leq i \leq N_x^h. \quad (3.3.16)$$

Transition probabilities for discrete Markov chain Once we discretize the HJB equation in this form (3.3.14), then, we define $(\xi_k^h)_{0 \leq k \leq N_x^h}$ a controlled Markov chain characterized by its transition probabilities determined explicitly from the HJB approximation (3.3.14). In other

words, from the finite differential approximation, we define the transition probabilities as follows $p(t_k, \nu_{i,k}, \alpha_{i,k}; k, k+1) = P^h(k, \alpha_{i,k})$ such that at each time t_k , $(P_{i,j}^h(k, \alpha_{i,k}))_{0 \leq i, j \leq N_x^h}$ is the matrix of transition probabilities for the controlled Markov Chain, defined by

$$\begin{aligned} P_{i,i-1}^h(k, \alpha_{i,k}) &= \frac{\sigma^2 \Delta t}{2 h^2} - \frac{\Delta t}{2h} b(t_k, x_i, \alpha_{i,k}), \quad 1 \leq i \leq N_x^h - 1, \\ P_{i,i}^h(k, \alpha_{i,k}) &= 1 - \sigma^2 \frac{\Delta t}{h^2}, \quad 1 \leq i \leq N_x^h - 1, \\ P_{i,i+1}^h(k, \alpha_{i,k}) &= \frac{\sigma^2 \Delta t}{2 h^2} + \frac{\Delta t}{2h} b(t_k, x_i, \alpha_{i,k}), \quad 1 \leq i \leq N_x^h - 1. \end{aligned} \quad (3.3.17)$$

With some transition probability with respect to boundary condition

$$\begin{aligned} P_{0,0}^h(k, \alpha_{0,k}) &= 1 - \frac{\sigma^2 \Delta t}{2 h^2} + \frac{\Delta t}{2h} b(t_k, x_0, \alpha_{0,k}), \\ P_{0,1}^h(k, \alpha_{0,k}) &= \frac{\sigma^2 \Delta t}{2 h^2} - \frac{\Delta t}{2h} b(t_k, x_0, \alpha_{0,k}), \\ P_{N_x^h, N_x^h}^h(k, \alpha_{N_x^h, k}) &= \frac{\sigma^2 \Delta t}{2 h^2} - \frac{\Delta t}{2h} b(t_k, x_{N_x^h}, \alpha_{N_x^h, k}), \\ P_{N_x^h, N_x^h-1}^h(k, \alpha_{N_x^h, k}) &= \frac{\sigma^2 \Delta t}{2 h^2} + \frac{\Delta t}{2h} b(t_k, x_{N_x^h}, \alpha_{N_x^h, k}). \end{aligned} \quad (3.3.18)$$

The $P_{i,j}^h(k, \alpha_{i,k})$ represents the conditional probability that the state of the chain ξ^h at time t_{k+1} is x_j given the state x_i at time t_k , and the control action α_{*k} , given by $P_{ij}^h(k, \alpha_{i,k})$. Therefore, the symmetric finite difference approximation HJB (3.3.14) can be written in the following form

$$U_{i,k} = \inf_{\alpha_{i,k}} \left\{ (L(x_i, \alpha_{i,k}) + F(x_i, M_{*k})) dt + \sum_{j=0}^{N_x^h} P_{ij}^h(k, \alpha_{i,k}) U_{j,k+1} \right\}. \quad (3.3.19)$$

Let us denote by \mathcal{A}^h the set of all policies (α^h, ξ^h) for the controlled Markov chain $(m_0^h, \alpha^h, \{P^h(k, \alpha_{*k}^h), 0 \leq k \leq N_t^h\})$,

$$\mathcal{A}^h := \mathcal{A}^h \left(m_0^h; \alpha; (P^h(k, \alpha_{*k}^h))_{0 \leq k \leq N_t^h} \right).$$

Remark 3.3.1. *An important advantage of this method is that the Markov chain approximation facilitates proofs of convergence of numerical methods in terms of probabilistic argument. Roughly speaking, by approximating the state variable process by a Markov chain, we obtain the appropriate transition probabilities of the Markov chain, we will show that converges to the solution of the Fokker Planks equation.*

According to Kushner and Dupuis [70], the Markov chain approximation variation $\Delta \xi_k^h = \xi_{k+1}^h - \xi_k^h$ with $|\Delta \xi_k^h| = o(h)$ must satisfy the local consistency conditions. This means that from a local point of view, the conditional mean and covariance of the changes in state of the chain are

proportional to the local mean drift and covariance for the original process.

Definition 3.3.2. (Local consistency Kushner and Dupuis, 2013, [70]).

We say that $\{P^h(k, \alpha_{*k}), 0 \leq k \leq N_t^h\}$ is locally consistent with the state process (X_t) defined by (3.2.2) if for any admissible policy (μ^h, ξ^h) the following local consistency condition holds

$$\mathbb{E}_{i,k}^\alpha[\Delta \xi_k^h] = b(t_k, X_{t_k}, \alpha_{t_k})dt + \mathbf{o}(h), \quad (3.3.20)$$

and

$$\mathbb{E}_{i,k}^\alpha[|\Delta \xi_k^h - \mathbb{E}_{i,k}^\alpha[\Delta \xi_k^h]|^2] = 2\sigma^2 dt + \mathbf{o}(h), \quad (3.3.21)$$

where $\Delta \xi_k^h := \xi_{k+1}^h - \xi_k^h$ and $\mathbb{E}_{i,k}^\alpha$ is the conditional expectation given $\xi_k^h = x_i$.

Remark 3.3.3. According to Kushner and Dupuis [70], under a local consistency condition the minimal cost function associated to the approximated Markov chain converge to the minimal cost function associated to the original problem.

Lemma 3.3.4. Let $(P^h(k, \alpha_{*k}))_{0 \leq k \leq N_t^h}$ be given by (3.3.17)-(3.3.18). For all k , we choose h and Δt such that

$$\Delta t \leq \frac{h^2}{\sigma^2} \quad \text{and} \quad h \leq \frac{\sigma^2}{\sup_{i,k}\{b(t_k, x_i, \alpha)\}}, \quad (3.3.22)$$

then $(P^h(k, \alpha_{*k}))_{0 \leq k \leq N_t^h}$ is a family of transition probabilities matrices.

Proof. We must check that $0 \leq P_{i,i}^h \leq 1$. Then

$$0 \leq 1 - \sigma^2 \frac{\Delta t}{h^2} \leq 1 \Leftrightarrow 0 \leq \sigma^2 \frac{\Delta t}{h^2} \leq 1 \Leftrightarrow \Delta t \leq \frac{h^2}{\sigma^2}.$$

And we check that $0 \leq P_{i,i-1}^h \leq 1$.

$$0 \leq \frac{\sigma^2}{2} \frac{\Delta t}{h^2} - \frac{\Delta t}{2h} \sup_{i,k} \{b(t_k, x_i, \alpha_{i,k})\} \leq 1 \Leftrightarrow 0 \leq \frac{\Delta t}{h^2} \left(\frac{1}{2} - h \frac{\sup_{i,k} \{b(t_k, x_i, \alpha_{i,k})\}}{2\sigma^2} \right) \leq 1.$$

Then

$$0 \leq \left(\frac{1}{2} - h \frac{\sup_{i,k} \{b(t_k, x_i, \alpha_{i,k})\}}{2\sigma^2} \right) \leq 1 \Leftrightarrow h \leq \frac{\sigma^2}{\sup_{i,k} \{b(t_k, x_i, \alpha_{i,k})\}}.$$

The law of the controlled Markov Chain is defined through the matrix of transition probabilities (3.3.17)-(3.3.18). The idea in this work is instead of discretizing the Fokker Planck equation,

we define the probability measure of the controlled discrete Markov chain using its transition probability matrix,

$$\begin{cases} M_{*,0}^h &= m_0^h, \\ M_{*,k}^h &= (P^h(k, \alpha_{*k}))^t M_{*,k-1}^h, \quad \text{for } 1 \leq k \leq N_t^h, \end{cases} \quad (3.3.23)$$

where we denote by $M_{*,k}^h$ the vector $(M_{i,k}^h)_{0 \leq i \leq N_x^h}$, for all the i and fixed k , and by $(P^h(k, \alpha_{*k}))^t$ the transposed matrix of $(P^h(k, \alpha_{*k}))$.

Remark 3.3.5. In [2], Achdou and Capuzzo-Dolcetta introduce a numerical method using the finite difference methods. They propose a finite difference scheme with monotone approximations of the Hamiltonian and also of the discrete formulation of the Fokker-Planck equation. For our approach, we do not look to discretize the Fokker-Planck equation. From finite difference approximation of HJB equation (3.3.14), we define the law m^h of the Markov chain from its transition probabilities, and we will show that it converges weakly to the solution of the Fokker Planck equation.

3.3.4 Discrete MFG equilibrium

Now, we can define the MFG equilibrium for the Markov chain control problem (3.3.19) as follow

Definition 3.3.6. (Discrete MFG equilibrium)

We say that $M^h \in \mathcal{P}(\mathcal{G}_{h,dt})$ is an equilibrium for the discrete MFG control problem if there exists an admissible strategy (α^h, ξ^h) such that

$$(\alpha^h, \xi^h) = \arg \min_{(\alpha, \xi) \in \mathcal{A}^h} J^h(i, k, \alpha, \xi, M), \quad (3.3.24)$$

where

$$J^h(i, k, \alpha^h, \xi^h, M^h) = \mathbb{E} \left[\sum_{k' \geq k} (L^h(\xi_{k'}^h, \alpha_{*k'}^h) + F(\xi_{k'}^h, M_{t_{k'}})) dt + \Phi^h(\xi_{N_t^h}^h, M_T) \mid \xi_\ell^h, \alpha_{*\ell}^h, \ell \leq k; \xi_k^h = x_i \right], \quad (3.3.25)$$

and M^h is the discrete distribution law of the discrete-time-space controlled Markov chain $(m_0^h, \alpha^h, \{P^h(k, \alpha_{*k}^h), 0 \leq k \leq N_t^h\})$ and satisfies

$$\begin{cases} M_{*0}^h &= m_0^h, \\ M_{*k}^h &= (P^h(k, \alpha_{*k}^h))^t M_{*k-1}^h = \left(\prod_{\ell \leq k-1} (P^h(\ell, \alpha_{*\ell}^h))^t \right) m_0^h, \quad 1 \leq k \leq N_t^h. \end{cases} \quad (3.3.26)$$

3.4 Existence and uniqueness results of discrete MFG solution

In this section, we will study the existence and uniqueness of discrete MFG equilibrium.

3.4.1 Existence result

To show the existence of a solution of discrete MFG problem, we use the result of Brouwer's fixed point theorem. The idea is to show that a function Γ^h admits at least a fixed point using this theorem. Therefore, we define the function Γ^h , for a given $h > 0$, as follows

$$\begin{aligned} \Gamma^h &: \mathcal{P}(\mathcal{G}_{h,dt}) \rightarrow \mathcal{P}(\mathcal{G}_{h,dt}) \\ m^h &\mapsto \Gamma^h(m^h) = M^h, \end{aligned} \quad (3.4.27)$$

such that $M^h = \Gamma^h(m^h)$ is the law of the discrete-time-space controlled Markov chain $(m_0^h, \alpha^h, \{P^h(k, \alpha_{*k}^h), 0 \leq k \leq N_t^h\})$ according to the following equation:

$$\begin{cases} M_{*,0}^h &= m_0^h, \\ M_{*,k}^h &= (P^h(k, \alpha_{*k}^h))^t M_{*,k-1}^h = \left(\prod_{\ell \leq k-1} (P^h(\ell, \alpha_{*\ell}^h))^t \right) m_0^h, \quad \text{for } 1 \leq k \leq N_t^h, \end{cases} \quad (3.4.28)$$

with α^h is associated to the optimal admissible policy

$$(\alpha^h, \xi^h) = \arg \min_{(\nu^h, \xi^h) \in \mathcal{A}^h} J^h(i, k, \nu^h, \xi^h, M^h). \quad (3.4.29)$$

Proposition 3.4.1. *Under assumptions (3.2.1)–(3.2.3), there exists at least one solution of the system (3.3.19)–(3.3.26).*

Proof. As we have that the set $\mathcal{P}(\mathcal{G}_{h,dt})$ being convex compact, indeed, according to Billingsley ([19], Theorem 1.3), the space $\mathcal{P}(\mathcal{G}_{h,dt})$ is a compact set since $\mathcal{G}_{h,dt}$ is also compact set. So we need Brouwer's fixed point theorem to prove that the function Γ^h is continuous function on $\mathcal{P}(\mathcal{G}_{h,dt})$ and in consequence prove the existence of an equilibrium of the discrete MFG control problem (3.3.19)–(3.3.26).

Let us consider that

$$M^{h,n} = \Gamma(m^{h,n}) \quad \text{and} \quad M^h = \Gamma(m^h),$$

where $m^{h,n}, m^h \in \mathcal{P}(\mathcal{G}_{h,dt})$, such that $M^{h,n}$ (respectively M^h) satisfies equation (3.3.26) associated to $\alpha^{h,n}$ (respectively α^h).

In other hand, we can write the above equations in the following form

$$M_{i,k+1}^{h,n} = P_{i,i}^h(k, \alpha_{i,k}^{h,n}) M_{i,k}^{h,n} + P_{i+1,i}^h(k, \alpha_{i,k}^{h,n}) M_{i-1,k}^{h,n} + P_{i-1,i}^h(k, \alpha_{i,k}^{h,n}) M_{i+1,k}^{h,n},$$

and

$$M_{i,k+1}^h = P_{i,i}^h(k, \alpha_{i,k}^h) M_{i,k}^h + P_{i+1,i}^h(k, \alpha_{i,k}^h) M_{i-1,k}^h + P_{i-1,i}^h(k, \alpha_{i,k}^h) M_{i+1,k}^{h,n},$$

with P^h are the probability transitions defined by (3.3.17)-(3.3.18).

Then, we take the difference between $M_{i,k+1}^{h,n}$ and $M_{i,k+1}^h$, for all $0 \leq i \leq N_t^h$,

$$\begin{aligned} M_{i,k+1}^{h,n} - M_{i,k+1}^h &= \left(1 - \sigma^2 \frac{\Delta t}{h^2}\right) (M_{i,k}^{h,n} - M_{i,k}^h) + \frac{\sigma^2 \Delta t}{2 h^2} (M_{i-1,k}^{h,n} - M_{i-1,k}^h) \\ &+ \frac{\sigma^2 \Delta t}{2 h^2} (M_{i+1,k}^{h,n} - M_{i+1,k}^h) M_{i,k-1}^h + \frac{\Delta t}{2h} b(t_k, x_{i-1,k}, \alpha_{i-1,k}^{h,n}) M_{i+1,k}^{h,n} \\ &+ \frac{\Delta t}{2h} b(t_k, x_{i+1,k}, \alpha_{i+1,k}^h) M_{i-1,k}^h - \frac{\Delta t}{2h} b(t_k, x_{i+1,k}, \alpha_{i+1,k}^{h,n}) M_{i-1,k}^{h,n} - \frac{\Delta t}{2h} b(t_k, x_{i-1,k}, \alpha_{i-1,k}^h) M_{i+1,k}^h. \end{aligned}$$

We know that b is bounded, then we get

$$\begin{aligned} M_{i,k+1}^{h,n} - M_{i,k+1}^h &\leq \left(1 - \sigma^2 \frac{\Delta t}{h^2}\right) (M_{i,k}^{h,n} - M_{i,k}^h) + \frac{\sigma^2 \Delta t}{2 h^2} (M_{i-1,k}^{h,n} - M_{i-1,k}^h) \\ &+ \frac{\sigma^2 \Delta t}{2 h^2} (M_{i+1,k}^{h,n} - M_{i+1,k}^h) + \frac{\Delta t}{2h} C_b (M_{i+1,k}^{h,n} - M_{i+1,k}^h) - \frac{\Delta t}{2h} C_b (M_{i-1,k}^{h,n} - M_{i-1,k}^h). \end{aligned}$$

Then, we have

$$\begin{aligned} |M_{i,k+1}^{h,n} - M_{i,k+1}^h| &\leq \left(1 - \sigma^2 \frac{\Delta t}{h^2}\right) |M_{i,k}^{h,n} - M_{i,k}^h| \\ &+ \left(\frac{\sigma^2 \Delta t}{2 h^2} + \frac{\Delta t}{2h} C_b\right) |M_{i-1,k}^{h,n} - M_{i-1,k}^h| + \left(\frac{\sigma^2 \Delta t}{2 h^2} + \frac{\Delta t}{2h} C_b\right) |M_{i+1,k}^{h,n} - M_{i+1,k}^h|. \end{aligned}$$

We take the maximum on i in the both side, we obtain

$$\begin{aligned} \max_i |M_{i,k+1}^{h,n} - M_{i,k+1}^h| &\leq \left(1 - \sigma^2 \frac{\Delta t}{h^2}\right) \max_i |M_{i,k}^{h,n} - M_{i,k}^h| \\ &+ \left(\frac{\sigma^2 \Delta t}{2 h^2} + \frac{\Delta t}{2h} C_b\right) \max_i |M_{i,k}^{h,n} - M_{i,k}^h| + \left(\frac{\sigma^2 \Delta t}{2 h^2} + \frac{\Delta t}{2h} C_b\right) \max_i |M_{i,k}^{h,n} - M_{i,k}^h| \\ &\leq \left(1 + \frac{\Delta t C_b}{2h^2}\right) \max_i |M_{i,k}^{h,n} - M_{i,k}^h| \end{aligned}$$

From the discret Gronwall's lemme, if (v_n) satisfy $v_{n+1} \leq (1 + \eta)v_n + k$, then

$$v_n \leq \exp(n\eta)v_0 + k \sum_{i=1}^{n-1} \exp(i\eta),$$

Hence, we obtain

$$\left\| M_{*,k+1}^{h,n} - M_{*,k+1}^h \right\| \leq C \left\| m_0^h - m_0^h \right\| = 0, \text{ where } C = \exp\left(\frac{\Delta t C_b}{2h^2}\right).$$

Finally, we proved the mapping Γ^h , defined by (3.4.27), is continuous on a convex compact set $\mathcal{P}(\mathcal{G}_{h,\Delta t})$. Then from Brouwer's fixed point theorem, we get the existence of equilibrium of the discrete MFG control problem.

3.4.2 Uniqueness result

Using the same argument of Lions-Lasry Monotonicity method ([74], Theorem 4.1), we will prove the uniqueness of discrete MFG equilibrium.

Proposition 3.4.2. *Under assumptions (3.2.2)–(3.2.4) and if F^h and Φ^h satisfy the Monotony condition, (i.e.),*

$$(F^h(M_{*k}) - F^h(\overline{M}_{*k}), M_{*k} - \overline{M}_{*k})_2 \leq 0 \text{ IR } F^h(M_{*k}) = F^h(\overline{M}_{*k}).$$

$$(\Phi^h(M_{*k}) - \Phi^h(\overline{M}_{*k}), M_{*k} - \overline{M}_{*k})_2 \leq 0 \Rightarrow \Phi^h(M_{*k}) = \Phi^h(\overline{M}_{*k}).$$

Then there exist at least one solution of the system (3.3.19)–(3.3.26).

Proof. Let M and $\overline{M} \in \mathcal{P}(\mathcal{G}_h)$ be two equilibrium for the discrete MFG control problem (α, ξ) (resp. $(\overline{\alpha}, \overline{\xi})$) and U (resp. \overline{U}) be respectively the optimal policy and the discrete value function associated to M (resp. \overline{M}) satisfying the following equations respectively:

$$\begin{cases} U_{i,k} &= \inf_{\alpha_{i,k} \in \mathcal{A}^h} \left\{ (L^h(x_i, \alpha_{i,k}) + F^h(x_i, M_{i,k})) dt + \sum_{j=0}^{N_t^h} P_{ij}^h(k, \alpha_{i,k}) U_{j,k+1} \right\}, \\ U_{*,N_t^h} &= \Phi^h(M_{*,N_t^h}), \end{cases} \quad (3.4.30)$$

and

$$\begin{cases} \overline{U}_{i,k} &= \inf_{\alpha_{i,k} \in \mathcal{A}^h} \left\{ (L^h(x_i, \alpha_{i,k}) + F^h(x_i, \overline{M}_{i,k})) dt + \sum_{j=0}^{N_t^h} P_{ij}^h(k, \alpha_{i,k}) \overline{U}_{j,k+1} \right\}, \\ \overline{U}_{*,N_t^h} &= \Phi^h(\overline{M}_{*,N_t^h}), \end{cases} \quad (3.4.31)$$

Taking the scalar product of the difference between the equations satisfied by U and \overline{U} multiplied by the difference $M - \overline{M}$, we obtain

$$\begin{aligned} & \sum_{k=0}^{N_t^h} (U_{*,k} - \overline{U}_{*,k}, M_{*,k} - \overline{M}_{*,k})_2 = (U_{*,N_t^h} - \overline{U}_{*,N_t^h}, M_{*,N_t^h} - \overline{M}_{*,N_t^h})_2 \\ & + \sum_{k=0}^{N_t^h-1} (P^h(k, \alpha_{*,k}) U_{*,k+1}, M_{*,k} - \overline{M}_{*,k})_2 - \sum_{k=0}^{N_t^h-1} \langle P^h(k, \overline{\alpha}_{*,k}) \overline{U}_{*,k+1}, M_{*,k} - \overline{M}_{*,k} \rangle_2 \\ & + \Delta t \sum_{k=0}^{N_t^h-1} (L^h(x, \alpha_{*,k}) - L^h(x, \overline{\alpha}_{*,k}), M_{*,k} - \overline{M}_{*,k})_2 + \Delta t \sum_{k=0}^{N_t^h-1} (F^h(x, M_{*,k}) - F^h(x, \overline{M}_{*,k}), M_{*,k} - \overline{M}_{*,k})_2. \end{aligned} \quad (3.4.32)$$

Using the fact that M and \bar{M} are two solutions of (3.3.26), then satisfy the same initial condition, i.e.

$$M_{*0} = m_0^h = \bar{M}_{*0}.$$

Therefore, we have the following equality

$$\sum_{k=0}^{N_t^h} (U_{*,k} - \bar{U}_{*,k}, M_{*,k} - \bar{M}_{*,k})_2 = \sum_{k=0}^{N_t^h-1} (U_{*,k+1} - \bar{U}_{*,k+1}, M_{*,k+1} - \bar{M}_{*,k+1})_2.$$

On the other hand, we know that

$$(P^h(k, \alpha_{*,k})U_{*,k+1}, M_{*,k+1})_2 = (U_{*,k+1}, P^h(k, \alpha_{*,k})^t M_{*,k+1})_2 = (U_{*,k+1}, M_{*,k+1})_2.$$

Then, the equation (3.4.32) becomes

$$\begin{aligned} & (U_{*,N_t^h} - \bar{U}_{*,N_t^h}, M_{*,N_t^h} - \bar{M}_{*,N_t^h})_2 + \Delta t \sum_{k=0}^{N_t^h-1} (F^h(x, M_{*,k}) - F^h(x, \bar{M}_{*,k}), M_{*,k} - \bar{M}_{*,k})_2 \\ &= \Delta t \sum_{k=0}^{N_t^h-1} (L^h(x, \alpha_{*,k}) - L^h(x, \bar{\alpha}_{*,k}), \bar{M}_{*,k})_2 - \Delta t \sum_{k=0}^{N_t^h-1} (L^h(x, \bar{\alpha}_{*,k}) - L^h(x, \alpha_{*,k}), M_{*,k})_2 \\ &+ \sum_{k=0}^{N_t^h-1} ((P^h(k, \alpha_{*,k}) - P^h(k, \bar{\alpha}_{*,k}))U_{*,k+1}, \bar{M}_{*,k})_2 - \sum_{k=0}^{N_t^h-1} ((P^h(k, \alpha_{*,k}) - P^h(k, \bar{\alpha}_{*,k}))U_{*,k+1}, M_{*,k})_2. \end{aligned}$$

Using the explicit expression of the transition probabilities, we calculate $(P^h(k, \alpha_{*,k}) - P^h(k, \bar{\alpha}_{*,k}))$, for all $0 \leq k \leq N_t^h - 1$, we have :

$$\begin{aligned} P_{0,0}^h(k, \alpha_{0,k}) - P_{0,0}^h(k, \bar{\alpha}_{0,k}) &= -\frac{\Delta t}{2h} (b(t_k, x_0, \alpha_{0,k}) - b(t_k, x_0, \bar{\alpha}_{0,k})), \\ P_{0,1}^h(k, \alpha_{0,k}) - P_{0,1}^h(k, \bar{\alpha}_{0,k}) &= \frac{\Delta t}{2h} (b(t_k, x_0, \alpha_{0,k}) - b(t_k, x_0, \bar{\alpha}_{0,k})), \\ P_{N_x^h, N_x^h}^h(k, \alpha_{N_x^h, k}) - P_{N_x^h, N_x^h}^h(k, \bar{\alpha}_{N_x^h, k}) &= \frac{\Delta t}{2h} (b(t_k, x_{N_x^h}, \alpha_{N_x^h, k}) - b(t_k, x_{N_x^h}, \bar{\alpha}_{N_x^h, k})), \\ P_{N_x^h, N_x^h-1}^h(k, \alpha_{N_x^h, k}) - P_{N_x^h, N_x^h-1}^h(k, \bar{\alpha}_{N_x^h, k}) &= -\frac{\Delta t}{2h} (b(t_k, x_{N_x^h}, \alpha_{N_x^h, k}) - b(t_k, x_{N_x^h}, \bar{\alpha}_{N_x^h, k})), \end{aligned}$$

and

$$\begin{aligned} P_{i,i-1}^h(k, \alpha_{i,k}) - P_{i,i-1}^h(k, \bar{\alpha}_{Ni,k}) &= \frac{\Delta t}{2h} (b(t_k, x_i, \alpha_{i,k}) - b(t_k, x_i, \bar{\alpha}_{i,k})), \\ P_{i,i}^h(k, \alpha_{i,k}) - P_{i,i}^h(k, \bar{\alpha}_{i,k}) &= -\frac{\Delta t}{2h} (b(t_k, x_i, \alpha_{i,k}) - b(t_k, x_i, \bar{\alpha}_{i,k})), \\ P_{i,i+1}^h(k, \alpha_{i,k}) - P_{i,i+1}^h(k, \bar{\alpha}_{Ni,k}) &= \frac{\Delta t}{2h} (b(t_k, x_i, \alpha_{i,k}) - b(t_k, x_i, \bar{\alpha}_{i,k})). \end{aligned}$$

Then by iterating, we get

$$(P^h(k, \alpha_{*,k}) - P^h(k, \bar{\alpha}_{*,k}))U_{*,k+1} \leq \frac{\Delta t}{h} (b(x_i, \alpha_{i,k}) - b(x_i, \bar{\alpha}_{i,k}))DU_{i,k+1}.$$

Then, using Assumption 3.2.1 and the fact that $\partial_\alpha L = -b_2(t) DU$, we get,

$$\begin{aligned} & (U_{*,N_t^h} - \bar{U}_{*,N_t^h}, M_{*,N_t^h} - \bar{M}_{*,N_t^h})_2 + \Delta t \sum_{k=0}^{N_t^h-1} \langle F^h(x, M_{*,k}) - F^h(x, \bar{M}_{*,k}), M_{*,k} - \bar{M}_{*,k} \rangle_2 \\ & \leq -\Delta t \sum_{k=0}^{N_t^h-1} (L^h(x, \alpha_{*,k+1}) - L^h(x, \bar{\alpha}_{*,k+1}), \bar{M}_{*,k})_2 + \Delta t \sum_{k=0}^{N_t^h-1} (L^h(x, \bar{\alpha}_{*,k+1}) - L^h(x, \alpha_{*,k+1}), M_{*,k})_2 \\ & + C \sum_{k=0}^{N_t^h-1} ((\alpha_{i,k+1} - \bar{\alpha}_{i,k+1}) \partial_\alpha L(x, \alpha_{i,k+1}), \bar{M}_{*,k})_2 - C \sum_{k=0}^{N_t^h-1} ((\alpha_{i,k+1} - \bar{\alpha}_{i,k+1}) \partial_\alpha L(x, \bar{\alpha}_{i,k+1}), M_{*,k})_2. \end{aligned}$$

Moreover, we have that L is strictly convex with respect to α , i.e.

$$\partial_\alpha L(x, \alpha_1)(\alpha_2 - \alpha_1) \leq L(x, \alpha_2) - L(x, \alpha_1), \quad \forall \alpha_1, \alpha_2 \in \mathcal{A}.$$

Therefore,

$$(U_{*,N_t^h} - \bar{U}_{*,N_t^h}, M_{*,N_t^h} - \bar{M}_{*,N_t^h})_2 + \Delta t \sum_{k=0}^{N_t^h-1} (F^h(x, M_{*,k}) - F^h(x, \bar{M}_{*,k}), M_{*,k} - \bar{M}_{*,k})_2 \leq 0.$$

Then F^h and Φ^h satisfy the Monotony condition. Therefore, we obtain the uniqueness of the solution.

3.5 Convergence results

This Markov Chain Approximation approach was initiated by Dupuis and Kushner [70] to construct numerical solution for the HJB equation. The authors prove the existence of approximating consistent Markov chain which converges to an stochastic process \tilde{X} .

In order to study the convergence of our discrete-MFG equilibrium (3.3.24) to the continuous

one, we need the results of Kushner and Dupuis [70] for weak convergence of time-continuous controlled Markov chains on the space of Càdlàg processes $D[0; T]$ which is a separable and complete for the Skorohod topology. We will use the same arguments used by Ben Aziza et al. in [13].

Remark 3.5.1. *From the results of Billingsley ([19], Theorem 6.8 page 73), since the space of Càdlàg processes $D[0; T]$, then the space of probability measures for Càdlàg processes $\mathcal{P}(D[0; T])$ is also separable and complete. Moreover, the relative compactness of any family of probability measures is equivalent to the tightness ([70], Theorem 9.1.1)*

Before announcing the main results of this work, we must introduce some mathematical notions and tools that will be useful to prove this convergence.

To force the state process X to stay in the domain, we need to add the following Neumann boundary condition

$$\forall x \in \partial\mathcal{K}, \forall t \in [0, T], \quad \partial_x u(x, t) \cdot r(x) = 0, \quad (3.5.33)$$

and

$$\forall x \in \partial\mathcal{K}, \forall t \in [0, T], \quad \partial_x m(x, t) = 0, \quad (3.5.34)$$

where $r(x)$ is the reflection direction at x and $\partial\mathcal{K}$ is the boundary set of \mathcal{K} .

In the following we propose a concise description of the limit Mean Field Game related to the system (3.2.8) with boundary conditions (3.5.33)-(3.5.34).

Let $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F} = \{\mathcal{F}_t\})$ be a filtered probability space, and let us consider a population of players where a representative player is described by a state variable $X^{\alpha, \epsilon} \in \mathcal{K}$, we recall that \mathcal{K} is a compact set.

$$\begin{cases} dX_t^\alpha &= b(t, X_t^\alpha, \alpha_t)dt + \sigma dB_t \\ X_0^\alpha &\sim \mu_0, \end{cases} \quad (3.5.35)$$

where B is a standard Brownian motion, α an admissible control process .

In this case, we must deal with the convergence of the controls, in addition to the convergence properties of the processes. Therefore, we need to define a new class of control, the relaxed control which provide a very powerful tool in the study of the convergence properties of sequences of optimally controlled processes. This is due to the fact that under general conditions, arbitrary sequences of relaxed controls have compact closure. This is not true of ordinary controls. The relaxed control is a class of control were first introduced by Yong to establish existence of a minimizing control for problem for the calculus of variations. This class were later used especially to prove the convergence properties of the approximation numerical problem. (See [70] chapter 9 for more details).

The idea of relaxed control is to replace the \mathcal{A} -valued process α_t with a $\mathcal{P}(\mathcal{A})$ -valued process α_t^R , where $\mathcal{P}(\mathcal{A})$ is the space of probability measures equipped with the topology of weak convergence.

Definition 3.5.2. *An admissible relaxed control is a measure α^R on $\mathcal{B}(\mathcal{A} \times [0, T])$ such that, for all $t \in [0, T]$*

$$\alpha^R(\mathcal{A} \times [0, t]) = t.$$

For $B \in \mathcal{B}(\mathcal{A})$, we define

$$\alpha_t^R(B) = \lim_{\delta \rightarrow 0} \frac{\alpha^R(B \times [t - \delta, t])}{\delta},$$

then we can write

$$\alpha^R(du, dt) = \alpha_t^R(du)dt.$$

Let us denote \mathcal{R} the set of relaxed controls.

We consider then the relaxation of the stochastic control problem where the drift function b is replaced by the function $b^R : [0, T] \times \mathcal{X} \times \mathcal{R}$ defined as

$$b^R(t, x, \alpha^R) := \int_{\mathcal{A}} b(s, x, u) \alpha_s^R(du),$$

and the function L is replaced by the function $L^R : \mathcal{X} \times \mathcal{R}$ defined as

$$L^R(x, \alpha^R) := \int_{\mathcal{A}} L(x, u) \alpha_s^R(du)$$

For any relaxed control α^R , we associate a cost function giving by

$$J_R(\alpha^R, m, t) = \mathbb{E} \left[\int_t^T \left[\int_{\mathcal{A}} L(X_s^{\alpha^R}, u) \alpha_s^R(du) + F(X_s^{\alpha^R}, m_s) \right] ds + \Phi(X_T^{\alpha^R}, m_T) \right], \quad (3.5.36)$$

where $X_t^{\alpha^R}$ is solution of the following relaxed model

$$\begin{cases} dX_t^{\alpha^R} &= \int_{\mathcal{A}} b(s, X_s^{\alpha^R}, u) \alpha_s^R(du) ds + \sigma dB_t, \\ X_0^{\alpha^R} &\sim m_0. \end{cases} \quad (3.5.37)$$

Definition 3.5.3. *(Relaxed MFG Equilibrium)*

We say that $m^* \in P[0, T]$ is a relaxed equilibrium for the MFG associated to the initial probability density $\mu_0 \in P$ if and only if it exist $\alpha^{R,*} \in \mathcal{R}$ such that

$$J(t, x \alpha^{R,*}, m) = \inf_{\alpha^R \in \mathcal{R}} J(t, x, \alpha^R, m), \quad \text{and} \quad m^* = \mathcal{L}(X_t^{\alpha^{R,*}}), \quad \forall t \in [0, T]. \quad (3.5.38)$$

Proposition 3.5.4. *According to the result of Kushner and Dupuis in [70] and for any $m \in P$,*

we get

$$\inf_{\alpha^R \in \mathcal{R}} J(t, x, \alpha^R, m) = \inf_{\alpha \in \mathcal{A}} J(t, x, \alpha, m). \quad (3.5.39)$$

Proof. To prove (3.5.39), see the result of Kushner et al. [70], Section 4.6, for more details. The idea is that, given a relaxed control $\alpha^R \in \mathcal{R}$, there exists a sequence $(\alpha_n)_n \in \mathcal{A}$ of ordinary controls such that the random measure $\delta_{\alpha_n(t)}(du)dt$ converges weakly to $\alpha^R(du)dt$ \mathbb{P} -a.s.

Remark 3.5.5. *The class of relaxed controls will be of importance in the convergence proof only. They do not enter into the numerical algorithms.*

In this section, we will give the main result on this section, using the intermediate tool introduced above.

Theorem 3.5.6. *Let M^h be the discrete MFG equilibrium (3.3.24) and ξ^h being its associated optimally controlled Markov chain defined by (3.3.17). Let $m^h \in \mathcal{P}(\mathcal{D}[0, T])$ be the law of the continuous-time-space piecewise constant interpolated process $\tilde{\xi}^h$.*

Then, the sequence $\{m^h\}_{h \geq 0}$ converges weakly to some \bar{m} in $\mathcal{P}(\mathcal{D}[0, T])$ where \bar{m} is an equilibrium of the MFG (3.5.38).

Moreover, let $\alpha^{R,h}$ be the relaxed control process associated to the time-space piecewise interpolated process $\alpha^{R,h}$, then $\alpha^{R,h}$ converges weakly to some relaxed process α^R such that

$$J_R(t, \bar{X}_t, \alpha^{R,*}, \bar{m}) = \min_{\alpha^R \in \mathcal{R}} J_R(t, \bar{X}_t, \alpha^R, \bar{m}) \quad \text{and} \quad \bar{m}_t = \mathcal{L}(\bar{X}_T^{\alpha^R}), \quad \forall t \in [0, T],$$

where the process $X_t^{\alpha^R}$ is a weak solution of

$$\begin{cases} d\bar{X}_t^{\alpha^R} &= \int_{\mathcal{A}} b(s, \bar{X}_s^{\alpha^R}, u) \alpha_s^R(du) ds dt + \sigma dB_t + dz(t), \\ \bar{X}_0^{\alpha^R} &\sim m_0. \end{cases} \quad (3.5.40)$$

Remark 3.5.7. *In this theorem, we had to change the frame and add a local reflection direction to force the process \bar{X} to return to the space if it leaves it. where z being a reflecting term on \mathcal{K} associated to reflecting directions $r(\partial\mathcal{K})$, such that there exist a process $\gamma : [0, T] \rightarrow \mathbb{R}$, $\gamma(s) \in r(X_s^{\alpha, \epsilon})$ almost surely with respect to the random measure induced by $|z(\cdot)|$:*

$$|z|(\cdot) = \int_0^t \mathbf{1}_{\mathcal{K}}(\bar{X}_s^{\alpha, \epsilon}) d|z|(s), \quad (3.5.41)$$

and

$$z(t) = \int_0^t \gamma(s) d|z|(s). \quad (3.5.42)$$

*In fact, for the discretization of our HJB equation, we have imposed boundary conditions ** ** to force the result to remain in the set, which can avoid the discretization of the reflection term.*

Alternatively, we can also present a discretization of the HJB by taking the reflection term into account, by agreeing on Kushner's approximation [70], Section 11 in the multidimensional case \mathbb{R}^d , where they has to consider with 2 cases:

- if $x \in \mathcal{K}$, u satisfy the same discrete equation

$$U_{i,k} = \inf_{\alpha_{i,k}} \left\{ (L(x_i, \alpha_{i,k}) + F(x_i, M_{*k})) dt + \sum_{j=0}^{N_x^h} P_{ij}^h(k, \alpha_{i,k}) U_{j,k+1} \right\},$$

- if $x \in \partial\mathcal{K}$, they consider a bounded and differentiable \mathcal{C} s.t. $\partial_x \mathcal{C}(x)\gamma > 0$, for any $\gamma \in r(x)$, $x \in \partial\mathcal{K}$, where u satisfy the following equation

$$U_{i,k} = \inf_{\alpha_{i,k}} \left\{ \sum_{j=0}^{N_x^h} P_{ij}^h(k, \alpha_{i,k}) U_{j,k+1} + \partial_x \mathcal{C}(x_i) \Delta z^h(x_i) \right\}.$$

with the state $\xi_n^h = x$ is in \mathcal{K} as $\Delta z^h(x) = \mathbb{E}_n^h[\Delta \xi_n^h]$.

It should be mentioned that, according to [70], the Markov chain in this case is always locally consistent with the reflected diffusion (3.5.35)-(3.5.41)-(3.5.42), if it is locally consistent with (3.2.2) in \mathcal{K} and is also locally consistent with the reflection direction $r(x)$.

The proof of this theorem is done in many steps, we start by studying the convergence of the approximating Markov Chain to stochastic controlled process \bar{X} and its law m converges in the space $\mathcal{P}(\mathcal{D}[0, T])$ to some \bar{m} . Then by injecting this limit \bar{X} in the control problem (3.5.36)-(3.5.37), we show that $\mathcal{L}(\bar{X})$ associated to \bar{X} converges to the same limit \bar{m} .

Thurs, the first step is to show that the controlled Markov chain $\{\xi^h, h > 0\}$ converges to a Càdlàg process on $\mathcal{D}[0, T]$ and the law m^h of the chain converges to $\bar{m} = \mathcal{L}(\xi^h)$ on $\mathcal{P}(\mathcal{D}[0, T])$. So we start by study the tightness of $\{\xi^h, h > 0\}$ which will give us the convergence of Markov chain.

Remark 3.5.8. By our construction, the Markov chain $\{\xi^h\}_{h>0}$ defined by (3.3.17)-(3.3.18) satisfies the local consistency conditions (3.3.20)-(3.3.21).

First, we need to announced the following theorem:

Theorem 3.5.9. (Kushner and Dupuis [70])

Consider an arbitrary collection of processes $\{x^n, n \in \mathbb{N}\}$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in $\mathcal{D}[0, \infty[$. Assume that for each $t \in [0, \infty[$ and $\delta > 0$ there exists compact $\mathcal{K}_{t,\delta}$ such that

$$\sup_n \mathbb{P}[x^n(t) \notin \mathcal{K}_{t,\delta}] \leq \delta. \quad (3.5.43)$$

Define \mathcal{F}_t^n to be the σ -algebra generated by $x^n, n \leq t$. Let \mathcal{T}_T^n be the set of \mathcal{F}_t^n -stopping times which are less than or equal to T with probability one, and assume for each $T \in [0, \infty[$ that

$$\lim_{\delta \rightarrow 0} \sup_n \sup_{\tau \in \mathcal{T}_T^n} \mathbb{E}[1 \wedge |x^n(\tau + \delta) - x^n(\tau)|] = 0. \quad (3.5.44)$$

Then $\{x^n, n \in \mathbb{N}\}$ is tight.

Now we can announced the following result:

Proposition 3.5.10. *The collection $\{\xi^h, h > 0\}$ is tight.*

Proof. We will show that the conditions (3.5.43)-(3.5.44) hold for the process $\{\hat{\xi}^h, h > 0\}$ associated to the transition probabilities (3.3.18)-(3.3.17).

Step 1: In fact, for $x_i \in \mathcal{G}_h$, we have

$$\xi^h(t) - x = \xi_k^h - \xi_0 = \sum_{i=0}^{N_t^h-1} \Delta \xi_{i+1}^h = \sum_{i=0}^{N_t^h-1} \left[\mathbb{E}_{i,k}^\alpha[\Delta \xi_k^h] + \Delta \xi_{i+1}^h - \mathbb{E}_{i,k}^\alpha[\Delta \xi_k^h] \right].$$

We apply Young's inequality , we get

$$\begin{aligned} |\xi^h(t) - x|^2 &= \left| \sum_{i=0}^{N_t^h-1} \left[\mathbb{E}_{i,k}^\alpha[\Delta \xi_k^h] + \Delta \xi_{i+1}^h - \mathbb{E}_{i,k}^\alpha[\Delta \xi_k^h] \right] \right|^2 \\ &\leq 2 \left\{ \left| \sum_{i=0}^{N_t^h-1} \mathbb{E}_{i,k}^\alpha[\Delta \xi_k^h] \right|^2 + \left| \sum_{i=0}^{N_t^h-1} \Delta \xi_{i+1}^h - \mathbb{E}_{i,k}^\alpha[\Delta \xi_k^h] \right|^2 \right\}. \end{aligned}$$

We take the conditional expectation $\mathbb{E}_{i,k}^\alpha$,

$$\mathbb{E}_{i,k}^\alpha \left[|\xi^h(t) - x|^2 \right] \leq 2 \left\{ \mathbb{E}_{i,k}^\alpha \left[\left| \sum_{i=0}^{N_t^h-1} \mathbb{E}_{i,k}^\alpha[\Delta \xi_k^h] \right|^2 \right] \mathbb{E}_{i,k}^\alpha \left[\left| \sum_{i=0}^{N_t^h-1} \Delta \xi_{i+1}^h - \mathbb{E}_{i,k}^\alpha[\Delta \xi_k^h] \right|^2 \right] \right\}.$$

From Remark 3.5.8, the chain $\{\xi^h\}_h$ satisfies the local consistency conditions (3.3.20)-(3.3.21).

Then we obtain

$$\begin{aligned} \mathbb{E}_{i,k}^\alpha \left[|\xi^h(t) - x|^2 \right] &\leq 2 \left\{ \mathbb{E}_{i,k}^\alpha \left[\left| \sum_{i=0}^{N_t^h-1} b(\xi_k^h) \Delta t \right|^2 \right] + \mathbb{E}_{i,k}^\alpha \left[\sum_{i=0}^{N_t^h-1} |2\sigma \Delta t + \mathbf{o}(\Delta t(\xi_k^h))| \right] \right\} \\ &\leq 2K^2 t^2 + 2(K + \mathbf{o}(h))t, \end{aligned}$$

where K is a bound for $|b(x)| \vee |\sigma(x)|$ for all x . Finally, from Tchebyshev's inequality, the first condition (3.5.43) is hold.

Step 2: Verifying that

$$\lim_{\delta \rightarrow 0} \sup_n \sup_{\tau \in \mathcal{T}_T^n} \mathbb{E}[1 \wedge |x^n(\tau + \delta) - x^n(\tau)|] = 0.$$

Thanks to the local consistency condition, we have

$$\begin{aligned} \mathbb{E}[1 \wedge |\xi^h(\tau + \delta) - \xi^h(\tau)|] &\leq \left(\mathbb{E}[|\xi^h(\tau + \delta) - \xi^h(\tau)|] \right)^{1/2} \\ &\leq \left(2K^2\delta^2 + 2(K + \mathbf{o}(h))\delta \right)^{1/2} \xrightarrow{\delta \rightarrow 0} 0 \end{aligned}$$

Finally, from the result 3.5.9 of Kushner and Dupuis [70], we obtain that the chain $\{\xi^h, h > 0\}$ is tight.

Therefore, we have the following result:

Lemma 3.5.11. *The sequence $(\xi^h)_h$ converges to a Càdlàg process \bar{X} in $D[0, T]$ such that $\mathcal{L}(\bar{X})$ corresponds to the limit of $\mathcal{L}(\xi^h)$.*

Proof. By tightness proved on Proposition 3.5.10, we may extract a weakly convergent subsequence, again referred to as $(\xi^h, h > 0)$.

To prove that ξ^h converges weakly to $\bar{X} \in D[0, T]$ such that $\mathcal{L}(\bar{X})$ corresponds to the limit of $\mathcal{L}(\xi^h)$, we need to use the result of Billingsley, 2013, ([19], theorem 13.4, page 142). Then we should prove the almost sure continuity of the limit process trajectories \bar{X} .

So, thanks to the construction of ξ^h , we get

$$\sup_{0 < t < T} |\xi^h(t) - \xi^h(t_-)| = dx = h \xrightarrow{h \rightarrow 0} 0.$$

Then we conclude that ξ^h converge to a Càdlàg process \bar{X} in $D[0, T]$.

Based also on the tightness of m and the results of Billingsley [19], we get

Lemma 3.5.12. *Let $(m^h)_h$ the sequence of laws associated to ξ^h in $\mathcal{P}(D[0, T])$, as mentioned in previous Theorem 3.5.6. Then, the sequence $(m^h)_h$ is tight and there exists an subsequence, still noted $(m^h)_h$ converging weakly to some \bar{m} in $\mathcal{P}(D[0, T])$.*

Proof. From remark 3.5.1, the space $\mathcal{P}(D[0, T])$ is a complete and separable space. According results of Billingsley ([19], Theorem 1.3) each probability measure on $\mathcal{P}(D[0, T])$ is tight. Thus the sequence $(m^h)_h$ is tight then we can extract a sub-sequence that converge to some limit \bar{m} in $\mathcal{P}(D[0, T])$.

And finally we conclude

Lemma 3.5.13. *For all $(t_i)_i$ a finite sequence in $[0, T]$, we have*

$$\mathcal{L}(\xi^h(t_1), \dots, \xi^h(t_k)) \equiv m_{t_1}^h \otimes \dots \otimes m_{t_k}^h$$

converges to $\bar{m}_{t_1} \otimes \dots \otimes \bar{m}_{t_k}$ when $h \rightarrow 0$.

Proof. To prove this last point of the lemma, thanks again to the results of Billingsley ([19], theorem 2.7, page 21) and the Lemma 3.5.11, we have the almost sure continuity of the limit process trajectories \bar{X} . As it explained in section 13 of Billingsley [19] for tightness in $D[0, T]$, we conclude that .

Thus we inject (\bar{m}, ξ^h) into the control problem (3.4.29) in order to show that the solution of Fokker-Planck and the solution of equation (3.3.26) converge to the same limit.

So let us denote by $(\bar{\alpha}^h, \bar{\xi}^h)$ the optimal policy associated to the control problem

$$(\bar{\alpha}^h, \bar{\xi}^h) = \arg \min_{(\alpha^h, \xi^h) \in \mathcal{A}^h} \tilde{J}^h(i, k, \alpha^h, \xi^h, \bar{m}). \quad (3.5.45)$$

Then, similarly to (3.3.19), we obtain the following equations for the associated discrete value function \bar{U}^h

$$\bar{U}_{i,k}^h = \inf_{\alpha \in \mathcal{A}^h} \left\{ \left(L^h(x_i, \alpha_{*,k}) + F[\bar{m}_{t_k}](x_i) \right) dt + \sum_j P_{ij}^h(k, \alpha_{i,k}) \bar{U}_{j,k+1}^h \right\},$$

with terminal condition

$$\bar{U}_{*,N_t^h}^h = \Phi[\bar{m}_T].$$

Moreover, $\forall 0 \leq k \leq N_t^h - 1$ and $\forall 1 \leq i \leq N_x^h$, the optimal action $\bar{\alpha}$ verifies

$$\bar{\alpha}_{i,k}^h = \arg \min_{\alpha \in \mathcal{A}^h} \left[[L^h(\alpha)]_{i,k} dt + \sum_j P_{ij}^h(k, \alpha) \bar{U}_{j,k+1}^h \right].$$

Let us define \bar{m}^h as the law of the controlled Markov chain $(m_0^h, \bar{\alpha}^h, \{P^h(k, \bar{\alpha}_{*k}^h), 0 \leq k \leq N_t^h\})$ with its piecewise constant interpolation density function $\bar{m}^h(x, t)$.

$$\begin{cases} \bar{m}_{*k}^h &= P^h(k, \bar{\alpha}_{*k}^h)^T \bar{M}_{*k-1}^h = \left(\prod_{\ell \leq k-1} (P^h(\ell, \bar{\alpha}_{*\ell}^h))^T \right) m_0^h, \quad \text{for } 1 \leq k \leq N_t^h, \\ \bar{m}_{*0}^h &= m_0^h, \end{cases} \quad (3.5.46)$$

Proposition 3.5.14. $\bar{\alpha}^{h,R}$ converges weakly to some relaxed process $\bar{\alpha}^R$ as $h \rightarrow 0$ such that

$$J_R(x, t, \bar{\alpha}^R, \bar{m}) = \inf_{\alpha_R \in \mathcal{R}(\mathcal{A} \times [0, T])} J_R(x, t, \alpha^R, \bar{m}), \quad \forall t \in [0, T], \quad (3.5.47)$$

where $X_t^{\bar{\alpha}^R}$ is a weak solution of

$$\begin{cases} dX_t^{\bar{\alpha}^R} = \int_{\mathcal{X}} b(s, X_s^{\bar{\alpha}^R}, u) \bar{\alpha}_s^R(du) dt + \sigma dB_t + dz(t) \\ X_t^{\bar{\alpha}^R} \sim m_0, \end{cases}$$

such that $\mathcal{L}(X_t^{\bar{\alpha}^R}) = \bar{m}^h$, with z is a reflecting process defined in (3.5.41)-(3.5.42).

The last step is to show that \bar{m}^h converges to the same limit as m^h , using the following result of Bayraktar et al [12] :

Lemma 3.5.15. *There exist a $\hat{T} > 0, h_0 > 0$ and $q \in (0, 1)$, such that for every $T \leq \hat{T}$, $m_1, m_2 \in \mathcal{P}(D(0, T), G)$ and $h \in (0, h_0 \wedge \hat{T})$,*

$$\mathcal{W}_1^2(\mathcal{L}(\hat{\xi}^{h, m_1}), \mathcal{L}(\hat{\xi}^{h, m_2})) \leq q \left(h^2 + \mathcal{W}_1^2(m_1, m_2) \right)$$

Lemma 3.5.16. *\bar{m}^h converges weakly to \bar{m} , the weak limit of m^h .*

Proof. To prove this convergence, we will use the Wassertein distance \mathcal{W}_1 . Using the following inequality

$$\mathcal{W}_1^2(\bar{m}^h, \bar{m}) \leq \mathcal{W}_1^2(\bar{m}^h, m^h) + \mathcal{W}_1^2(m^h, \bar{m}).$$

From Lemma 3.5.15, we have that $\mathcal{W}_1^2(m^h, \bar{m}) \xrightarrow{h \rightarrow 0} 0$. So we just need to prove that

$$\mathcal{W}_1^2(\bar{m}^h, m^h) \xrightarrow{h \rightarrow 0} 0.$$

We remind that \bar{m}^h and m^h are respectively the law of the optimal controlled Markov chain $(m_0^h, \bar{\alpha}^h, \{P^h(k, \bar{\alpha}_{*k}^h), 0 \leq k \leq N_t^h\})$ and $(m_0^h, \bar{\alpha}^h, \{P^h(k, \bar{\alpha}_{*k}^h), 0 \leq k \leq N_t^h\})$.

From Lemma 3.5.15, we have

$$\begin{aligned} \mathcal{W}_1^2(\bar{m}^h, m^h) &\leq \mathcal{W}_1^2(\mathcal{L}(\bar{\xi}^h), \mathcal{L}(\xi^h)) \\ &\leq q(h^2 + \mathcal{W}_1^2(m^h, \bar{m})). \end{aligned}$$

Moreover, m^h and \bar{m}^h present the induced measures associated to m^h and \bar{m} respectively. Consequently, we get that \bar{m}^h and m^h converge weakly to the same limit \bar{m} .

Concluding, the previous results collected together lead to the convergence result that we look to prove in this work. In one hand we have that the tow families of probability measures $\{\bar{m}^h, h > 0\}$ and $\{m^h, h > 0\}$ converge weakly to the same limit law \bar{m} . On the other hand, the limit law \bar{m} is a equilibrium of the MFG continuous problem (3.2.5).

Then, the solution of a MFG discrete problem (3.3.24) converges weakly to a solution of a MFG continuous problem (3.2.5).

3.6 Numerical simulation

The theoretical study of the approach proposed in this thesis pushes the numerical validation. Therefore in order to validate our approach, we study the case of Linear Quadratic Mean Field Games (LQ-MFG for short). First we present the class of LQ-MFG. Later we solve the associated MFG PDEs using the Markov chain approximation approach. To compare the results and certify it we applied for a model with explicit solution.

3.6.1 Linear quadratic model

We introduce a Linear Quadratic model of MFG where the cost functional is quadratic in all state variables, control variables and the mean field terms, while the controlled dynamics are linear. The Linear Quadratic model was well studied in Bensoussan et al. [17]. Let us consider the following linear quadratic form for L , F , b and Φ :

$$\begin{aligned} L(x, \alpha) &:= \frac{1}{2} (x^* Q x + \alpha^* R \alpha), \\ F(m, x) &:= \frac{1}{2} \left(x - S \int \varepsilon m(\varepsilon) d\varepsilon \right)^* \bar{Q} \left(x - S \int \varepsilon m(\varepsilon) d\varepsilon \right), \\ b(t, x, \alpha) &:= Ax + B\alpha, \\ \Phi(m, x) &:= \frac{1}{2} \left(x^* Q_T x + \left(x - S_T \int \varepsilon m(\varepsilon) d\varepsilon \right)^* \bar{Q}_T \left(x - S_T \int \varepsilon m(\varepsilon) d\varepsilon \right) \right), \end{aligned}$$

where A , B , R , Q , \bar{Q} , S are given constants.

We remain that we look to solve the following coupled HJB-FP system of PDE

$$\begin{cases} -\partial_t u(t, x) - \frac{1}{2} \sigma^2 \Delta u(t, x) + H(x, \partial_x u(t, x)) = F[m(t, x)](x) \\ \partial_t m(t, x) - \frac{1}{2} \sigma^2 \Delta m(t, x) - \partial_x(m(t, x)) \partial_p H(x, \partial_x u(t, x)) = 0 \\ u(T, x) = \Phi(m(T, x), x), \quad m(0, x) = \mu_0(x). \end{cases} \quad (3.6.48)$$

3.6.2 Explicit solution

In this section, we have to solve the system of HJB-FP equations introduced by 3.6.48 which reads as follows

$$\begin{cases} -\partial_t u(t, x) - \frac{1}{2}\sigma^2 \Delta u(t, x) = -\partial_x u^* BR^{-1}B^* \partial_x u + \partial_x u A x + \\ \quad + \frac{1}{2} \left(x^* Q x + \left(x - S \int \varepsilon m(\varepsilon) d\varepsilon \right)^* \bar{Q} \left(x - S \int \varepsilon m(\varepsilon) d\varepsilon \right) \right), \\ \partial_t m(t, x) - \frac{1}{2}\sigma^2 \Delta m(t, x) - \partial_x (m(Ax - BR^{-1}B^* \partial_x \alpha)) = 0, \\ u(T, x) = \frac{1}{2} \left(x^* Q_T x + \left(x - S_T \int \varepsilon m(\varepsilon) d\varepsilon \right)^* \bar{Q}_T \left(x - S_T \int \varepsilon m(\varepsilon) d\varepsilon \right) \right), \\ m(x, 0) = m_0(x). \end{cases} \quad (3.6.49)$$

We look for a solution $u(t, x)$ of the following quadratic form

$$u(t, x) := \frac{1}{2} x^* P(t) x + x^* r(t) + s(t), \quad (3.6.50)$$

so we get

$$\partial_x u(t, x) = P(t)x + r(t), \quad \Delta u(t, x) = P(t).$$

Then the FP equation be written as

$$\partial_t m(t, x) - \frac{1}{2}\sigma^2 \Delta m(t, x) - \partial_x (m((A - BR^{-1}B^* P(t))x - BR^{-1}B^* r(t))) = 0. \quad (3.6.51)$$

We set $z(t) = \int x m(t, x) dx$ the mean of m . It easy to see that $z(t)$ satisfies the following equation

$$\begin{cases} \frac{dz}{dt} = (BR^{-1}Br(t) - A) z(t) + BR^{-1}Br(t), \\ z(0) = z_0. \end{cases} \quad (3.6.52)$$

Using the quadratic form of u , we get the coupled system of differential equations:

$$\frac{dP}{dt} + AP(t) + P(t)A - P(t)BR^{-1}BP(t) + Q + \bar{Q} = 0 \quad (3.6.53)$$

$$\frac{dr}{dt} = -Ar(t) + 2P(t)BR^{-1}Br(t) + Sz(t)Q, \quad (3.6.54)$$

$$\frac{ds}{dt} = \frac{-\sigma^2}{2} + r(t)BR^{-1}Br(t) - \frac{1}{2}Sz(t)\bar{Q}Sz(t), \quad (3.6.55)$$

$$P(T) = Q + \bar{Q}, \quad r(T) = -\bar{Q}Sz(T), \quad s(T) = \frac{1}{2}z(T)S\bar{Q}Sz(T). \quad (3.6.56)$$

The difficulty of the numerical resolution is the resolution of the coupled forward backward system (3.6.52)-(3.6.55). Thanks to shooting method (see [24, 99]), we can solve it. The shooting method is a method for solving a boundary value problem by reducing it to the system of an initial value problem. Roughly speaking, we shoot our trajectories in different directions until we

find a trajectory that has the desired boundary value.

In other words, by initializing z^0 , we solve the equation (3.6.52) by replacing z by z^0 . Then we inject the solution z^1 in (3.6.55) and solve it by initializing r^0 . We repeat these steps by injecting each time the solution found z^k in the equations for the next step and so on. We repeat this procedure until we find c .

Algorithm 1 Shooting method

Initial (z^0, r^0) .

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

Find z^{k+1} solution of (3.6.52) $\frac{dz}{dt} = (BR^{-1}Br_t^k - A) z_t^k + BR^{-1}Br_t^k$,

Find r^{k+1} (3.6.55) $\frac{dr}{dt} = -Ar_t^k + 2P_tBR^{-1}Br_t^k + Sz_t^{k+1}Q$,

Return (z^N, r^N)

Repeat until found $r_T^k = r_T$.

Therefore, the first step is to solve (3.6.54) and the second one is to solve the coupled system (3.6.52)-(3.6.55) for z and r using shooting method and we end up with find the solution s of (3.6.56). To simplify the calculation, we take the Gaussian probability density defined as follows

$$m(t, x) = \frac{1}{\sqrt{2\pi v(t)}} \exp \frac{(x - z(t))^2}{2v^2(t)}.$$

Moreover, we know that the optimal control α^k is defined as follows

$$\alpha^k = \min_{\alpha} J(x, t, \alpha, m).$$

Thus, in this linear quadratic framework, we obtain this explicit form :

$$\alpha^k := -\frac{B}{R} \partial_x u(t, x). \tag{3.6.57}$$

3.6.3 Numerical results

The numerical solution of the problem is based on a fixed point argument. The distribution is initialized and then an iteration method is applied to the optimal policies. At each iteration the optimal strategy is computed, then the value function and the new distribution until stability.

Algorithm 2 Point fixe algorithm

Initial M^0 and some threshold $\epsilon > 0$.
while $\|M^{k+1} - M^k\| < \epsilon$ and $\|U^{k+1} - U^k\| < \epsilon$ **do**:
 for $k = N - 1$ to 0 **do**:
 calculate U^k (3.3.19),
 calculate M^k (3.3.23),
 calculate α^k (3.6.57).
end while.

Here, we compare both explicit and numerical solutions for the same initial probability density $m_0 = \mathcal{N}(0, 0.5)$ for $A = 0.01$.

Let us start by setting the parameters chosen for the numerical experiments:

$$T = 1, \quad \sigma = 0.35, \quad Q = 2, \quad \bar{Q} = 1, \quad S = 0.5, \quad \mathcal{K} = [-5, 5].$$

Test 1 :

From figures 3.1 and 3.2, we can see that the probability densities are more and more similar when evolving in time. The evolution of the probability distributions seems to be toward a centered Gaussian one.

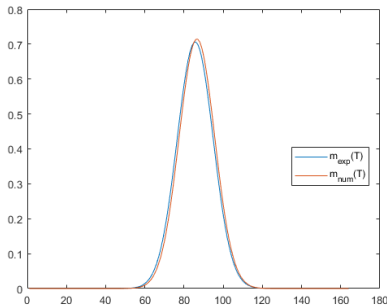


Figure 3.1: Explicit and Numerical probability measure at T

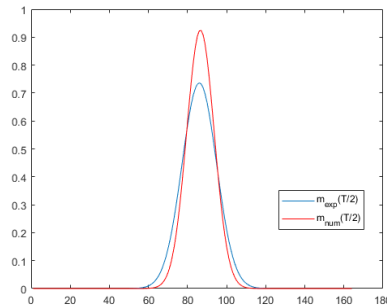


Figure 3.2: The explicit probability measure at $T/2$

The probability measure: FP solution

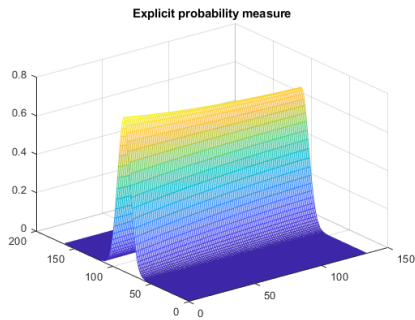


Figure 3.3: The explicit probability measure

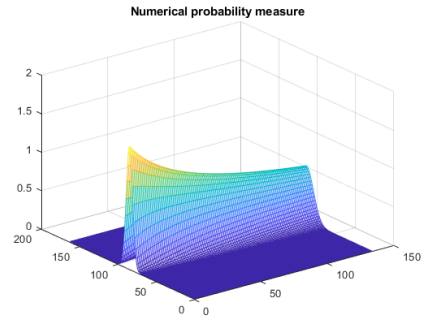


Figure 3.4: The numerical probability measure

As for the probability densities, it is clear from figures 3.5 and 3.6, that the value functions are more and more similar when evolving in time

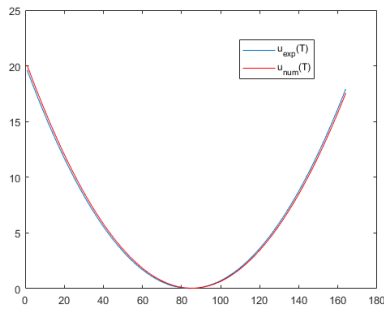


Figure 3.5: Explicit and numerical value function at T .

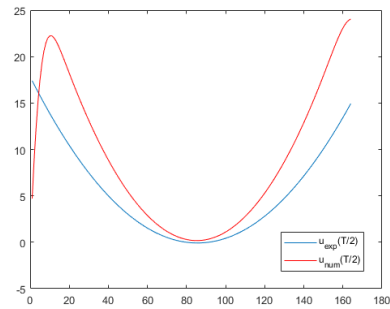


Figure 3.6: Explicit and numerical value function at $T/2$.

The value function : HJB solution

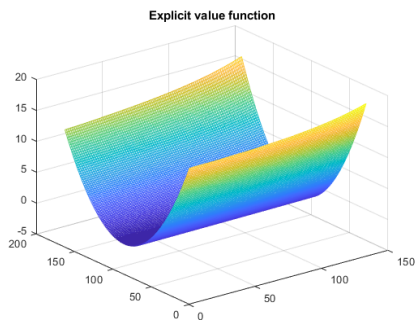


Figure 3.7: The explicit value function

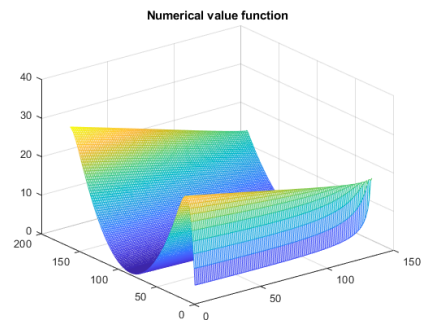


Figure 3.8: The numerical value function

Actually, according to the numerical implementations, once we take a larger value of A , we

lose the convergence of the explicit solution to the numerical solution.

Test 2 : For $A = 0.5$, we get the following numerical result:

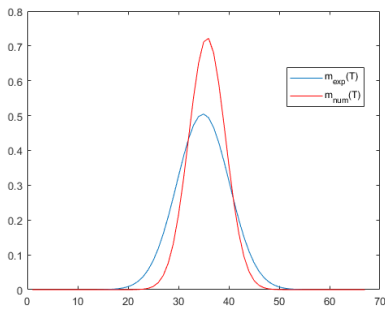


Figure 3.9: Explicit and Numerical probability measure at T

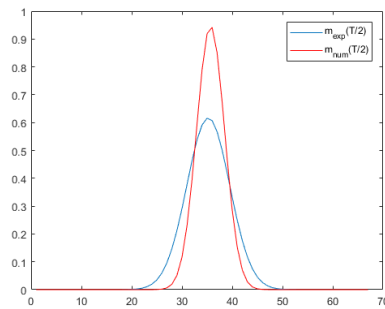


Figure 3.10: The explicit probability measure at $T/2$

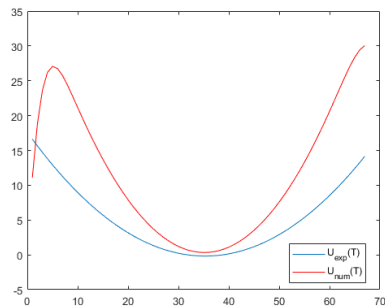


Figure 3.11: Explicit and Numerical value function at T

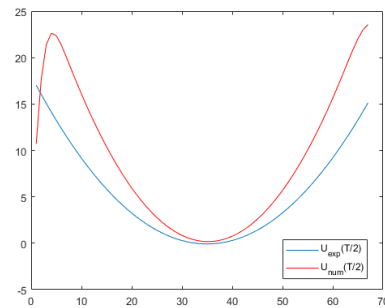


Figure 3.12: The explicit value function at $T/2$

Test 3 : For $A = 1$, we get the following numerical result:

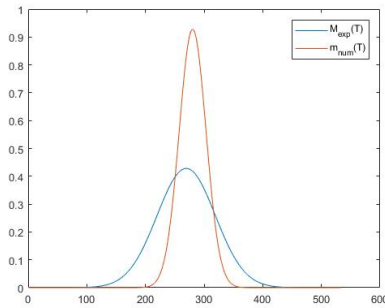


Figure 3.13: Explicit and Numerical probability measure at T

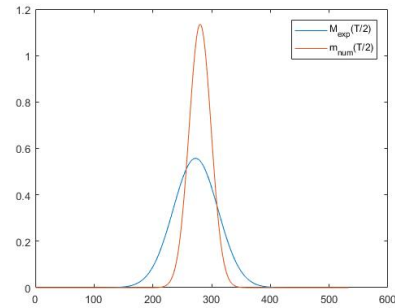


Figure 3.14: The explicit probability measure at $T/2$

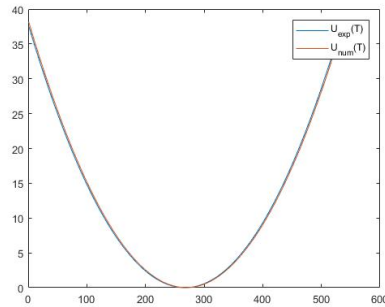


Figure 3.15: Explicit and Numerical value function at T

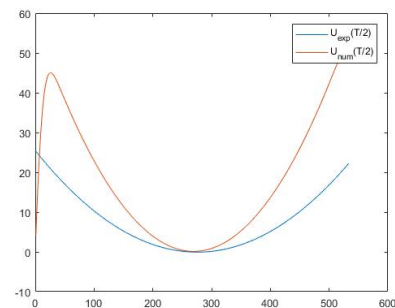


Figure 3.16: The explicit value function at $T/2$

It is clear, from figures (3.9)-(3.10)-(3.13)-(3.14) that we lose the convergence of the m density. In fact, if A is increasingly large then the discretization path dx is increasingly small and therefore the conditions (3.3.22) are not verified. Numerically, we risk to have values very close to zero and we lose the compact criterion of the discretization set $\mathcal{G}_{\{t,h\}}$ and consequently the convergence of the numerical solution is lost.

In the following, we take the same values of the X and Y parameters, however we will expand the time intervals $[0, T] = [0, 10]$ and $\mathcal{K} = [-10, 10]$, we obtain the following numerical results:

Test 4 : We take $A = 0.5$.

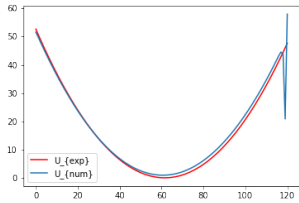


Figure 3.17: Explicit and Numerical value function at t_0

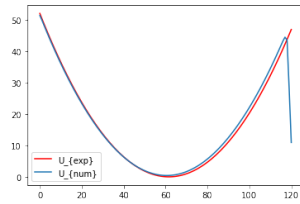


Figure 3.18: Explicit and Numerical value function at $T/2$

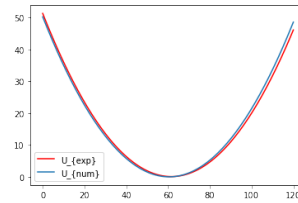


Figure 3.19: Explicit and Numerical value function at T

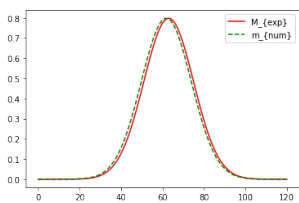


Figure 3.20: Explicit and Numerical probability measure at t_0

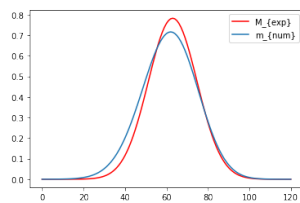


Figure 3.21: Explicit and Numerical probability measure at $T/2$

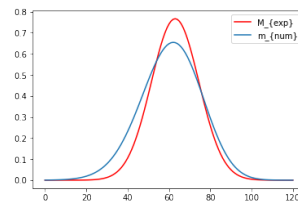


Figure 3.22: Explicit and Numerical probability measure at T

By extending these intervals, we notice, according to the figures above, that the numerical solution of the value function is closer to the explicit solution. On the other hand, the numerical solution of the probability measure, for this test, is also closer to the explicit solution. On the other hand, for a larger value of A , we also lose the convergence of the solutions.

LEARNING DYNAMIC UTILITIES

4.1 Introduction

This chapter aims to propose an algorithm to learn the utility of an agent by observing her decisions at discrete (possibly random) dates $(\tau_i)_i$. There are many settings in which this capability is very useful. In cooperative settings, we may want to help the agent make good decisions. This is the principle of robo-advisors which are online investment management services that employ mathematical algorithms to provide financial advice with minimal human intervention. Based on online questionnaires to obtain information about the client's degree of risk-aversion, financial status, and desired return on investment, the robo-advisors use their algorithms to manage and allocate client assets in the most efficient way possible. In a more competitive setting, we may want to predict the agent's actions so as to better optimize our own payoffs, this is the case of two-player A and B where A is trying to learn the utility's of B, to optimize her own payoff which also depends on B's actions and utility.

From a theoretical and application point of view this problem of recovering the utility of an agent from observations is an old inverse problem: Same questions were addressed by the economist Samuelson in the 40's [111, 112], with the theory of "revealed preference" where the observable is the consumption process. Since then, the theory has been growing in interest under the pressure of the economic reality, which created new incentives for different approaches, see Chambers & Echenique [36]. An example is the evolutionary economics by Arthur [10]. Economies are considered as complex evolutionary systems, where the agents try to predict the outcomes of their actions, and how the market would be modified by their decisions.

Beyond the fields of economics and finance, in the last few decades this forward-looking viewpoint has been also considered in many optimization problems as a consequence of the expansion of computer science and e-commerce. The forward modeling allows anticipations on the future values of observations. But, the agents also need to adjust their (random) preferences over time, following an "inverse thinking" approach as has been suggested by Gomez-Ramirez [55].

In the field of machine learning, we refer for example to Friedman and Sandow [49], Chajewska, Koller, and Ormoneit [35]; and Stahl [113] and Hibbard [62]. As well, reasoning with preferences has been recognized as a particularly promising research direction for artificial intelligence see

Nielsen and Jensen. [101] and Qi, Xu, and Lafferty. [108]. For e-commerce, see [116]: The principal idea is to build the utility function of a user from his search history and purchases on the Internet. This utility will then be used to target the user by proposing products or baskets of goods that we know are likely to be of interest to him and at specific prices evaluated from this utility function.

Since the problem is posed forward in time and given the randomness of the universe, the utility that we seek to construct from the observations must be a dynamic utility coherent in time and consistent with this universe (in a sense that we specify further on). For these reasons, we are concerned, in this work by, the concept of "forward" dynamic utilities which was first introduced and considered by Musiela and Zariphopoulou [95, 97, 98] to model the preferences of an agent in a continuous semimartingale financial market, see also Henderson and Hobson [61], El Karoui and Mrad [45] and El Karoui & al [44] for a model with consumption. See also Matoussi and Mrad [89] for a semimartingale market with Jumps.

Recently in [46], the authors have addressed the problem of revealed utility in a general dynamic framework beyond finance and economics. They establish a necessary and sufficient condition for the existence of at least one solution to the "revealed" utility problem and given a simple way for recovering the utility of an agent provided that the observed process $X(x)$ is monotonic with respect to its initial condition x , which is an assumption that is satisfied in finance and economics, one does not invest more to earn less. In a more general framework, it also remains true as soon as we have some regularity (Lipschitz is enough) on the diffusion coefficients of X .

They are the results of this last paper that we implement in this work. Let us briefly present the idea and different steps of the revealed dynamic utility problem as in [46]. By definition, the observable is a so-called dynamic positive characteristic process $\{X_t(x)\}$ considered for different values of its initial condition $X_0(x) = x > 0$, and assumed to be increasing in x (to be coherent with the expected utility criterium). The goal is to recover, from a given initial utility function u , a stochastic dynamic utility $\mathbf{U} = \{U(t, z, \omega), z > 0\}$, "revealed optimally" in the sense that at any (stopping) time τ , the preference for the observable process is in mean equal to its value at time 0, $\mathbb{E}(U(\tau, X_\tau(x))) = u(x)$: from the probabilistic dynamic view point, on a given filtered probability space, the performance process " $U(t, X_t(x))$ is a martingale". Focusing on the concavity of the utility criterium \mathbf{U} , tools of convex analysis play a key role, especially the invertible decreasing marginal utility $U_z(t, z)$. Its allows to define the convex Fenchel-Legendre transform $\tilde{U}(t, y)$ of $U(t, x)$, by $U(t, z) - zU_z(t, z) = \tilde{U}(t, U_z(t, z))$, but also to linearize the recovery problem, by using the one to one correspondence between the class of dynamic utilities \mathbf{U} (revealed by X) and the adjoint processes Y candidate to play the role of $U_z(t, X_t)$, more precisely $\{Y_t(u_z(x)) = U_z(t, X_t(x))\}$. Since the characteristic process is invertible, the triplet $\{(u, X_t(x), U(t, z))\}$ is in one to one correspondence with the triplet $\{(u, X_t(x), Y_t(u_z(x)))\}$. In

fact, denoting by $\{\mathcal{X}_t(z) := (X_t)^{-1}(z)\}$ the inverse flow of X , we have the characterization

$$U_z(t, z) = Y_t(u_z(\mathcal{X}_t(z))), \quad \text{and} \quad U(t, z) = U(t, z_0) + \int_{z_0}^z Y_t(u_z(\mathcal{X}_t(x))) dz.$$

Note that in the continuous semimartingale [45] and semimartingale with jumps framework [89], the authors have shown that if U is a regular dynamic utility, consistent with a given financial market, it is necessarily in this form with X being the optimal portfolio and Y being the optimal state price density process.

The goal, in the revealed utility problem, is then to identify the class of the adjoint processes $\{Y_t(y)\}$ in bijection with a revealed dynamic utility U , such that $U(t, X_t(x))$ is a martingale. A quasi-necessary condition on the triplet $\{(u, X_t(x), Y_t(y))\}$ is " $\{X_t(x)Y_t(y)\}$ is a supermartingale for any x, y and $\{X_t(x)Y_t(u_z(x))\}$ is a martingale for $y = u_z(x)$ "; this last condition guarantees that the conjugate dynamic process $\{\tilde{U}(t, Y_t(y))\}$ read along the adjoint process $\{Y_t(y)\}$, is a martingale.

Note that the revealed utility problem does not necessarily have a unique solution, there are as many utilities as there are processes Y satisfying this last condition. From a financial point of view, the observable \mathcal{X} is the optimal portfolio of the investor and Y is his optimal state price (pricing kernel). If the market is complete, Y is unique and corresponds, up to a discount factor, to the risk neutral density. However, if the market is incomplete, several choices of Y are possible. This work is concertized in the preprint [94].

4.2 Preliminary results on the numerical approximation of dynamic utility

Automated investment managers or robo-advisors were imposed these last years in the financial and economic world to replace or alternate with traditional human advisors. The robo-advisors are online platforms that provide financial advice or portfolio management to an investor or a client. The performance of the robo-advisor strongly depends on its ability to accurately assess the investor's risk tolerance. They have many advantages over the investment services offered by traditional management consultants. Robo-advisor knows the investor better than a traditional asset manager. Because of this better knowledge, the robo-advisor may propose a more appropriate asset allocation. In addition, it performs the task in a systematic way and implements an automated rebalancing process. From the client's perspective, the biggest advantage of robo-Advisor is that it provides a low-cost alternative to traditional investing.

The idea of these platforms is to develop a new investment framework in which the robo-advisor not only manages the client's portfolio but also reacts regularly with him/her to obtain updated

information on his/her performance during time. More precisely, the client is giving an information about his/her preferences to the robo-advisor only at specific update dates $(\tau_i)_i$. So, between two update dates, the robo-advisor does not receive any information from the customer. Therefore, it develops a framework to learn the investor's preferences between two update dates. In other words, by observing the data communicated by the client at interaction dates, robo-advisor approximates the preferences of the client to make investment decisions and it interacts repeatedly with the client to update its informations by asking from the investor his/her new data in order to avoid making decisions based on stale information. In general, the robo-advisor solicits this information through targeted online questionnaires, and transfers it to digital data. We assume that we have observations for several different initial conditions $(z_j)_j$. Such informations are not necessarily collected in the questionnaires of the robo-Advisors, but maybe it is necessary because we know very well that the choices of an agent are strongly conditioned by its initial position (e.g initial wealth). Moreover, without these observations, we have a little chance to find the dynamic utility of the agent, see the papers [44–46, 89]. In this work the idea is the following

- *Step 1:* From the observations $X_{\tau_i}(z_j)$ and $Y_{\tau_i}(y_j)$ we construct the marginal utility at the points (τ_i, z_j) , using the characterization $U_z(\tau_i, z_j) = Y_{\tau_i}(u_z(\mathcal{X}_{\tau_i}(z_j)))$, where $\mathcal{X}_t(z) = X_t^{-1}(z)$ denotes the inverse maps of $z \rightarrow X_t(z)$.
- *Step 2:* We integrate numerically with respect to the spatial parameter and get $U(\tau_i, z_j)$.
- *Step 3:* This data is used as a starting point to learn the utility U at any point (t, z) , using several artificial intelligence methods, specifically the deep learning method.
- *Step 4:* Finally we try to predict the utility at a reasonable horizon (to be able to make decisions in place of the client) and we conclude with a comparison of these different approaches.

Note, as we do not dispose of observations, we will generate them by ourselves and this by assuming that the processes X and Y evolve according to two stochastic differential equations with jumps.

4.2.1 The model

In this section, we suppose that the filtered probability space $(\Omega, \mathbb{F}, \mathbb{P})$ supports a q -dimensional standard Brownian motion $B = (B^1, \dots, B^q)$ on and an independent q' -dimensional Poisson random measure N on $[0, \infty[\times \mathbb{R}^{q'}$ with constant time dependent intensity measure $\lambda dt \times \nu(de)$ defined on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. λ is the intensity of jumps and ν is a finite

positive measure on $\mathbb{R}^{q'}$, satisfying standard integrability assumptions. We also denote by \tilde{N} the compensated version of N :

$$\tilde{N}(dt, de) = N(dt, de) - \nu(de)\lambda(t)dt.$$

We are concerned by two \mathbb{R}^d -valued stochastic processes X and Y , solutions of the following stochastic differential equations (SDE for short)

$$dX_t(x) = \mu(t, X_t(x))dt + \sum_{i=1}^q \sigma_i(t, X_t(x))dB_t^i + \int_{\mathbb{R}^{q'}} h(t, X_{t-}(y), e)\tilde{N}(dt, de), \quad X_0(x) = x, \quad (4.2.1)$$

$$dY_t(y) = b(t, Y_t(y))dt + \sum_{i=1}^q \gamma_i(t, Y_t(y))dB_t^i + \int_{\mathbb{R}^{q'}} g(t, Y_{t-}(y), e)\tilde{N}(dt, de), \quad Y_0(y) = y, \quad (4.2.2)$$

where $\mu, b, \sigma_i, \gamma_i$ are deterministic functions from $\mathbb{R}_+ \times \mathbb{R}^d$ into \mathbb{R}^d and h, g are deterministic functions from $\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^{q'}$ into $\mathbb{R}^{d \times q'}$, globally Lipschitz in space to ensure the existence of a unique strong solution.

Assumption 4.2.1 (For Y). **(HP1)** There exist C^Y and $C^Y(e)$ s.t. for any $t \in [0, T]$, $x, y \in \mathbb{R}^d$ and $e \in \mathbb{R}^{q'}$

$$\begin{cases} |b(t, x) - b(t, y)| \leq C^Y|x - y|, & |\mu(t, 0)| \leq C^Y, \\ |\gamma(t, x) - \gamma(t, y)| \leq C^Y|x - y|, & |\gamma(t, 0)| \leq C^Y, \\ |g(t, x, e) - g(t, y, e)| \leq C^Y(e)|x - y|, & |g(t, 0, e)| \leq C^Y(e), \end{cases} \quad (4.2.3)$$

where $C^Y(e)$ satisfies $\int_{\mathbb{R}^{q'}} [C^Y(e)]^p \nu(de) < \infty$, $\forall p \geq 2$. **(HP2)** b, γ and g are continuously space-differentiable functions s.t. $\nabla_y b, \nabla_y \gamma$ and $\nabla_y g$ satisfy **(HP1)** with $(C^{Y, \nabla}, C^{Y, \nabla}(e))$ in place of $(C^Y, C^Y(e))$. **(HP3)** b, γ and g are α^Y -Hölder continuous in time, locally in space,

$$\begin{cases} |b(t, x) - b(s, x)| + |\gamma(t, x) - \gamma(s, x)| \leq C^Y(1 + |x|)|t - s|^{\alpha^Y}. \\ |g(t, x, e) - g(s, x, e)| \leq C^Y(e)(1 + |x|)|t - s|^{\alpha^Y}. \end{cases} \quad (4.2.4)$$

(HP4) $\nabla_y b, \nabla_y \gamma$ and $\nabla_y g$ satisfy **(HP3)** with $(C^{Y, \nabla}, C^{Y, \nabla}(e))$ in place of $(C^Y, C^Y(e))$.

Assumption 4.2.2 (For X). **(HP1)** and **(HP3)** are satisfied for μ, σ and h (instead of b, γ and g) with coefficient $C^X, C^X(e)$ and α^X (instead of $C^Y, C^Y(e)$ and α^Y).

Before giving the Euler scheme, we first expose how to simulate the integral with respect to the Poisson measure. We follow the presentation of [92].

Simulation of the integral with respect to the Poisson random measure Let us consider a sequence e_1, e_2, \dots of independent random variables with common exponential distribution with

parameter 1. We define

$$\Lambda(t) = \int_0^t \lambda(s) ds, \quad t \in [0, T].$$

To simplify, we take $\lambda(t) = 1$.

The number of jumps of the random Poisson measure $N(dt, de)$ in an interval $[0, t]$ is determined as

$$J(t) = \max\{k : \sum_{j=1}^k e_j \leq \Lambda(t)\}.$$

We denote by $J = J(T)$ the total number of jumps in $[0, T]$ and the jump times $(\tau_k)_k$ of the Poisson measure defined by

$$\tau_k = \Lambda^{-1}\left(\sum_{j=1}^k e_j\right), \quad k \in \llbracket 1, J \rrbracket,$$

where Λ^{-1} is the right continuous inverse of Λ . The jump times can be computed recursively by

$$e_k = \int_{\tau_{k-1}}^{\tau_k} \lambda(s) ds, \quad k \in \llbracket 1, J \rrbracket.$$

Once the jump times are computed, we proceed to sample the marks $\{E_k\}$, that, conditionally on the values of the jumps times, are independent random variables distributed respectively according to $\{\nu(de)\}$. The random measure with intensity $\lambda(t)dt\tilde{\nu}(de)$ can then be constructed as

$$N(dt, de) = \sum_{k=1}^J \delta_{(\tau_k, E_k)}(dt, de).$$

Consequently, the stochastic integral with respect to the Poisson random measure can be defined as

$$\int_0^t \int_{\mathbb{R}^d} g(s, Y_{s-}, e) N(dt, de) = \sum_{k=1}^{J(t)} G(\tau_k, Y_{\tau_k-}, E_k), \quad t \in [0, T].$$

4.2.2 Euler schemes

Let $N \geq 1$ and let us consider the discretization family $\{\bar{t}_i := i\frac{T}{N}, i \in \llbracket 0, N \rrbracket\}$ of $[0, T]$. We consider also the jump times $\{\tau_k, k \in \llbracket 1, J \rrbracket\}$ with corresponding marks $\{E_k, k \in \llbracket 1, J \rrbracket\}$. Consider the augmented partition given by the union

$$\mathcal{P}_N = \{t_l, l \in \llbracket 0, N + J \rrbracket\} = \{\bar{t}_i := i\frac{T}{N}, i \in \llbracket 0, N \rrbracket\} \cup \{\tau_k, k \in \llbracket 1, J \rrbracket\}.$$

Scheme 4.2.1. *The Euler scheme to approximate the solution Y of the SDE (4.2.2) is given by the following steps*

- Set $Y_0^N(x) = x$.
- For $k = 0, \dots, N + J - 1$, if $t \in (t_k, t_{k+1}]$, then

$$Y_{t^-}^N(x) = Y_{t_k}^N(x) + b(t_k, Y_{t_k}^N(x))(t - t_k) + \gamma(t_k, Y_{t_k}^N(x))(B_t - B_{t_k}).$$

- If $t_{k+1} = \tau_l$ for some $l \in \llbracket 1, J \rrbracket$, then we introduce a correction due to jumps discontinuities

$$Y_{t_{k+1}}^N(x) = Y_{t_k}^N(x) + g(t_{k+1}, Y_{t_{k+1}}^N(x), E_l).$$

Inverse flow as a backward in time SDE One last difficulty remains to be overcome: How to invert X ? There are at least four approaches for computing the inverse flow: (a) as an inverse of a random function; (b) as a forward in time SPDE; (c) as a forward in time SDE with stochastic coefficients; (d) as a backward in time SDE with standard coefficients. In Gobet et Mrad [51], the authors discussed these different approaches in detail (see [51, Section 4]) and they have exposed the main difficulties of each method. The simplest, most efficient and less expensive is the backward method, it allows us to approximate the inverse flow \mathcal{X} using the Euler's scheme. Without going into detail, this is the method we will adopt here.

Let $(X_{s,t}(x))_{t \geq s}$ be the solution, starting from x at time s , of the SDE with coefficients (μ, σ, h) . Denote by $\mathcal{X}_{s,t}$ the inverse map of $X_{s,t}$ which can also be defined by $\mathcal{X}_{s,t}(x) := \mathcal{X}_t(X_s(x))$ where $\mathcal{X}_t(x) \neq \mathcal{X}_{0,t}(x)$, focuses on the dynamics of $\mathcal{X}_{s,t}(x)$ in the variable s : doing so, we aim at computing the inverse of Y backward in time instead of forward in time. This approach relies on the following key result.

Theorem 4.2.1 ([67, Theorems 3.11 and 3.13]). *Suppose the coefficients μ, σ and h of the SDE (4.2.1) satisfy Assumption 4.2.1. Assume further that the maps $\phi(t, \cdot, e) : z \mapsto z + h(t, z, e); \mathbb{R} \rightarrow \mathbb{R}$ are homeomorphic with $1 + \partial_z h(t, z, e)$ being invertible in z for a.e (t, e) .*

- (i) *Then the solution X defines a stochastic flow of \mathcal{C}^1 -diffeomorphisms.*
- (ii) *Let $\psi(t, \cdot, e)$ be the inverse maps of $\phi(t, \cdot, e)$ and $k(t, z, e) = z - \psi(t, z, e)$. Assume $\gamma \in \mathcal{C}^{1,2}$ in (t, x) and $\int_{\mathbb{R}} |k(t, z, e) - h(t, z, e)| \nu(de)$ is bounded. Then the inverse flow \mathcal{X} satisfies the following backward SDE*

$$\begin{aligned} d\mathcal{X}_{s,t}(z) = & [\mu(s, \mathcal{X}_{s,t}(z)) - \partial_z \sigma(s, \mathcal{X}_{s,t}(z)) \cdot \sigma(s, \mathcal{X}_{s,t}(z)) - \int_{\mathbb{R}} h(s, \mathcal{X}_{s,t}(z), e) \nu(de)] \hat{d}s \\ & + \sigma(t, \mathcal{X}_{s,t}(z)) \cdot d\overleftarrow{B}_s + \int_{\mathbb{R}} k(s, \mathcal{X}_{s,t}(z), e) N(\hat{d}s, de), \quad s \leq t. \end{aligned} \quad (4.2.5)$$

$$\mathcal{X}_{t,t}(z) = z.$$

The notations $\hat{d}s$ and \overleftarrow{B}_s are to remind that the integrals must be considered in a backward way¹. With this result in hand, the approximation of $\mathcal{X}_{s,t}$ is made possible simply using a standard Euler scheme like for Y^n . Using similar notations as below,

Scheme 4.2.2. *Based on this key result, the approximation of the inverse flow \mathcal{X} is achieved by the following steps.*

- Set $\mathcal{X}_{t,t}^N(z) = z$. If $t = \theta_k$ for some $k \in \llbracket 1, J \rrbracket$, set $\mathcal{X}_{t^-,t}^N(z) = z - k(t, z, E_k)$ else $\mathcal{X}_{t^-,t}^N(z) = z$.

- For $s \in]\tau_t, t]$, set

$$\begin{aligned} \mathcal{X}_{s,t}^N(z) &= \mathcal{X}_{t^-,t}^N(z) - \sigma(t, \mathcal{X}_{t^-,t}^N(z)) \cdot (B_t - B_s) \\ &\quad - \left[\mu(t, \mathcal{X}_{t^-,t}^N(z)) - \partial_z \sigma(t, \mathcal{X}_{t^-,t}^N(z)) \cdot \sigma(t, \mathcal{X}_{t^-,t}^N(z)) - \int_{\mathbb{R}} \mathcal{X}(t, \mathcal{X}_{t^-,t}^N(z), e) \nu(de) \right] (t - s). \end{aligned} \quad (4.2.6)$$

- For $l \in \llbracket 0, N + J \rrbracket$ satisfying $t_l < \tau_t$,

- if $t_l = \theta_k$ for some $k \in \llbracket 1, J \rrbracket$, set $\mathcal{X}_{t_l^-,t}^N(z) = \mathcal{X}_{t_l,t}^N(z) - k(t_l, \mathcal{X}_{t_l,t}^N(z), E_k)$ else $\mathcal{X}_{t_l^-,t}^N(z) = \mathcal{X}_{t_l,t}^N(z)$

- and for $s \in [t_{(l-1)}, t_l[$

$$\begin{aligned} \mathcal{X}_{s,t}^N(z) &= \mathcal{X}_{t_l^-,t}^N(z) - \sigma(t_l, \mathcal{X}_{t_l^-,t}^N(z)) \cdot (B_{t_l} - B_s) \\ &\quad - \left[\sigma(t_l, \mathcal{X}_{t_l^-,t}^N(z)) - \partial_z \sigma(t_l, \mathcal{X}_{t_l^-,t}^N(z)) \cdot \sigma(t_l, \mathcal{X}_{t_l^-,t}^N(z)) - \int_{\mathbb{R}} h(t_l, \mathcal{X}_{t_l^-,t}^N(z), e) \nu(de) \right] (t_l - s). \end{aligned} \quad (4.2.7)$$

It is well known that the Euler scheme, whether forward (as for Y^N) or backward (as for \mathcal{X}^N), converges, under regularity assumptions (like Lipschitz-type), to the solution of the stochastic differential equation at the rate $\frac{1}{N^{\frac{1}{2}}}$. However, few results give us the rate of convergence of the compound of two approximation schemes. Thanks to the results of the paper [93], we know that our compound process $Y_{0,t}^N(u_z(\mathcal{X}_{0,t}^N(z)))_{t \geq 0}$ converges to the marginal utility $U_z(t, z)$ at the rate $\frac{1}{N^{1/2}}$.

Theorem 4.2.2 (Mrad [93]). *Assume that the coefficients (b, γ, g) of Y satisfy Assumption 4.2.1 and the coefficients $(\mu - \partial_z \sigma \cdot \sigma - \int h \nu(de), \gamma, k)$ of $\mathcal{X}_{\cdot,t}$, equation (4.2.5), satisfy Assumption 4.2.2. Denote by $X_{0,\cdot}^N$ the Euler approximation associated to $X_{0,\cdot}$, with time step T/N , and by $\mathcal{X}_{0,t}^N$ the Euler approximation of the inverse flow $\mathcal{X}_{0,t}$ of X_t , with time step T/N , according to (4.2.6)-(4.2.7).*

Then, for any concave function u with Lipschitz marginal utility u_z , the compound Euler scheme

$$\begin{aligned} \int_s^t \mu(s, X_u) \hat{d}u &:= \lim_{|\Delta| \rightarrow 0} \sum_{k=0}^{n-1} \mu(t_{k+1}, X_{t_{k+1}}) (t_{k+1} - t_k) \\ \int_s^t \sigma(s, X_u) d\overleftarrow{B}_u &:= \lim_{|\Delta| \rightarrow 0} \sum_{k=0}^{n-1} \sigma(t_{k+1}, X_{t_{k+1}}) (B_{t_{k+1}} - B_{t_k}), \end{aligned}$$

where $\Delta = \{s = t_0 < t_1 < \dots < t_n = t\}$ and $|\Delta| = \max_{0 \leq k \leq n-1} (t_{k+1} - t_k)$.

$X_t^N(u_z(\xi^N))$ converges to $U_z(\cdot, \cdot)$ in any L_p -norm, at the order $\beta := \min(\alpha^X, \alpha^Y, \frac{1}{2})$ w.r.t. N : For any $p > 0$, $t \in [0, T]$ and any z ,

$$\|Y_{0,t}^N(u_z(X_{0,t}^N(z))) - U_z(t, z)\|_{L_p} = O(N^{-\beta}).$$

Thus with a simple integration, we can calculate an approximation \mathbf{U} of the utility U at any time t_i , by

$$\mathbf{U}(t_i, z_j) = \sum_{k=0}^{j-1} Y_{t_i}^N(u_z(X_{t_i}^N(z_k)))(z_{k+1} - z_k),$$

since $z_0 = 0$ and we assume $u(t, 0) = 0$ a.s. for any $t \geq 0$.

This approach allows us to generate the utility of the agent at the dates τ_i of interaction with the robo-advisor on a grid (z_j) points. Here are some numerical results obtained by applying this process, scheme 1 for simulating Y and scheme 2 for the inverse of X .

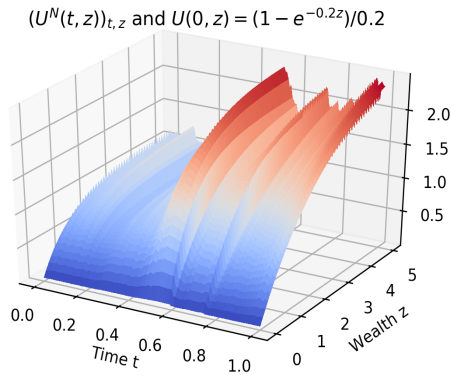


Figure 4.1: Simulation of utility at each time τ .

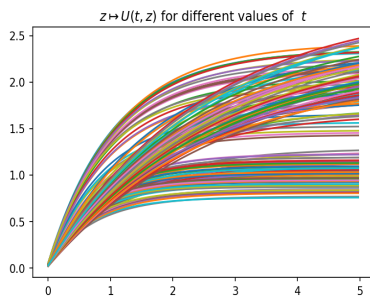


Figure 4.2: Monotonicity and Concavity of approximated utilities $\mathbf{U}(t, z)$

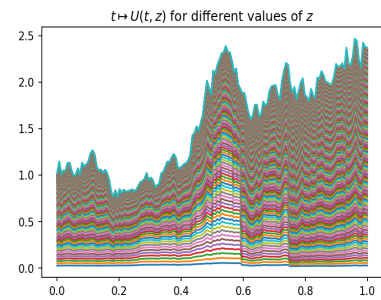


Figure 4.3: Dynamic of approximated utilities $\mathbf{U}(t, z)$

This approach is based on the work of Mrad and his co-authors [45, 46, 89], it allows us to estimate the agent's utility at dates and points of interaction . This is a very important step as until

now there was no idea how to quantify the preferences of an agent from these choices. The goal then, starting from the values obtained $(U(t_i, z_j))_{i,j}$, is to estimate $U(t, z)$ for all (t, z) and to predict its future decisions. For this we will use different machine learning algorithms and we will compare them.

The aim in the following is to propose an algorithm for learning the preferences of the agent. As we do not have access to real data, we will simulate our own data $X_{\tau_i}(z_j)$ and Y_{τ_i} for $(i, j) \in \llbracket 1, J \rrbracket \times \llbracket 1, n \rrbracket$. Once these values have been simulated we proceed to approximate the utility U at the points (t_i, \mathfrak{z}_j) (using the above procedure). The generated values $U(\tau_i, \mathfrak{z}_j)$ will play the role of starting data for our learning algorithms. This kind of algorithm could clearly be used on robo-advisor platforms, since it can "predict" (over a short period of time) the preferences of an agent and therefore can make decisions for him/her.

4.3 Machine Learning methods

Machine Learning algorithms automatically learn to perform a task or make predictions from data and improve their performance over time. Once trained, the algorithm will be able to find patterns in new data. Depending on the information available during the learning phase, learning is qualified in different ways. If the data are labeled (i.e., the response to the task is known for these data), it is supervised learning. It is called classification if the labels are discrete, or regression if they are continuous. If the model is learned in an iterative way according to a reward received by the program for each action taken, it is called reinforcement learning. In the most general case, without labels, we try to determine the underlying structure of the data (which can be a probability density) and it is then called unsupervised learning. Machine learning can be applied to different types of data, such as graphs, trees, curves or more simply feature vectors, which can be continuous or discrete qualitative or quantitative variables.

In this work, we are interested in supervised learning. Our task is to predict a target value which is the value of utility at fixed time τ_i , denoted by $(\mathcal{U}_j^i)_{j \in \llbracket 1, n \rrbracket}$ a vector of the matrix $(\mathbf{U}(\tau_i, \mathfrak{z}_j))_{(i,j) \in \llbracket 1, N \rrbracket \times \llbracket 1, n \rrbracket}$ from different initial wealth values $(\mathfrak{z}_j)_{j \in \llbracket 1, n \rrbracket}$ and fixed value τ_i :

$$\mathcal{U}_j^i := (\mathbf{U}(\tau_i, \mathfrak{z}_j))_{j \in \llbracket 1, n \rrbracket},$$

such that $(\tau_i)_{i \in \llbracket 1, N \rrbracket}$ are the dates we built in the first simulation step.

From this dataset, we create a model, which is nothing else than a mathematical function. Once we have created it, we develop and train our model with new data $(Z_j^i) = (\mathfrak{z}_j)_{j \in \llbracket 1, n \rrbracket}$ in order to predict the target or the answer $(\mathcal{U}_j^i)_j$.

4.3.1 Machine Learning programs using Python

Most of the time, it is not necessary to develop its own algorithm, thanks to **PYTHON**, they are already developed in open source libraries. For example, **Scikit-learn library** contains all the mathematical functions which are necessary for learning problems. **Scikit-learn** is an open source machine learning library that supports supervised and unsupervised learning. It also provides various tools for model fitting, data preprocessing, model selection and evaluation and many other utilities.

With **Scikit-learn**, it is enough to write a few lines to develop models of linear regression and classification:

- Create the **Dataset**. (Z^{in}, \mathcal{U}^i) such that
 - Z^{in} are the input data. It contains different initial values of wealth $(\mathfrak{z}_j)_{j \in \llbracket 1, n \rrbracket}$, (i.e)

$$(Z_j^{in})_{j \in \llbracket 1, n \rrbracket} = (\mathfrak{z}_j)_{j \in \llbracket 1, n \rrbracket}.$$

- \mathcal{U}^i are the target. It contains a vector of the utility matrix $(\mathbf{U}(\tau_i, \mathfrak{z}_j))_{(i,j) \in \llbracket 1, N \rrbracket \times \llbracket 1, n \rrbracket}$ approximated in Section 4.2.2. In other words, \mathcal{U}^i is utility values at a fixed time τ_i and for different values of $(\mathfrak{z}_j)_{j \in \llbracket 1, n \rrbracket}$:

$$(\mathcal{U}_j^i)_{j \in \llbracket 1, n \rrbracket} := (\mathbf{U}(\tau_i, \mathfrak{z}_j))_{j \in \llbracket 1, n \rrbracket}.$$

- Choice of the **model**: choose the best model to fit our data set. Among the predefined models in SKlearn library, we can mention
 - Decision Tree
 - Support vector Machine (SVM)
 - Neural Network
 - K nearest neighbors (KNN).
- Fitting **SKlearn.fit** (Z^{in}, \mathcal{U}^i) : The objective is to fit the model, (i.e), the machine must learn from the model:
 1. The samples matrix Z^{in} . The size of Z^{in} is typically (n_samples, n_features), which means that samples are represented as lines and features are represented as columns.
 2. The target values \mathcal{U}^i is usually 1d array where the i th entry corresponds to the target of the i th sample of Z^{in} .

- Predicting: **SKlearn.predict**(Z^{in}): Once the estimator is fitted, it can be used for predicting a future value $U^{p,i}$ using the following syntax (we use the notation "p" for prediction)

$$U^{p,i} := \text{SKlearn.predict}(Z^{in}).$$

- Evaluation: **SKlearn.score**(Z^{in}, \mathcal{U}^i): The accuracy of our model is obtained using the function **Model.score**. This function returns the coefficient of determination R^2 of the prediction. The coefficient R^2 is defined as

$$R^2 = 1 - \frac{(\mathcal{U}^i - U^{p,i})^2}{(\mathcal{U}^i - \bar{U})^2},$$

with $\bar{U} := \frac{1}{n} \sum_{j=1}^n \mathcal{U}_j^i$ represents the mean of $(\mathcal{U}_j^i)_j$. The best possible score is 1.0 and it can be negative (because the model can be arbitrarily worse).

Remarks 3. *Train the model using all the Data, an also use it to test the result is not a good way to measure the performance of our model and may involve overfitting. So to avoid this, we choose to split the data, usually we make 80/20: **Train set** (80%) and **Test set** (20%).*

*To create a train set and test set, thanks to python again, we use the **train_test_split** function of Sklearn. **SKlearn.train_test_split** is a function in Sklearn model selection for splitting data arrays into two subsets: for training data and for testing data. With this function, we don't need to divide the data manually. As we said above, using the same dataset for both training and testing steps, increases the chances of inaccurate predictions.*

In this section, we use the method described above to learn the utility of an investor. Firstly, we create the target, then we want to implement (separately) two packages from **Scikit-learn library: Support vector Machine** and **Neural Network model** to create the model and we compare the results. To recap, we are looking to implement the following algorithm 3 for SVM and Neural Network models

Algorithm 3 Implementation algorithm in Python

Import sklearn library.

Get Data Z^{in} .

Get Target \mathcal{U}^i .

Create the **model**.

Split into **training and test set**:

$$Z_{\text{train}}^{in}, Z_{\text{test}}^{in}, \mathcal{U}_{\text{train}}^i, \mathcal{U}_{\text{test}}^i = \text{sklearn.train_test_split}(Z^{in}, \mathcal{U}^i, \text{train_size} = 0.8, \text{test_size} = 0.2).$$

Fit the model: **sklearn.fit**($Z_{\text{train}}^{in}, \mathcal{U}_{\text{train}}^i$).

Evaluation the model with training data: **sklearn.score**($Z_{\text{train}}^{in}, \mathcal{U}_{\text{train}}^i$).

Evaluation the model with test data: **sklearn.score**($Z_{\text{test}}^{in}, \mathcal{U}_{\text{test}}^i$).

Predict the model: **sklearn.predict**(Z_{test}^{in}).

We want to apply the learning methods described above to predict the utility of an investor according to the information observed by the Robo-Advisor, the preferences of the customer or the response of the customer to a questionnaire.

4.3.2 Support Vector Machine

SVM approach

In machine learning, **support-vector machines** (SVM in short) are supervised learning models with associated learning algorithms that analyze data for classification and regression analysis, developed by Vladimir Vapnik et Alexey Chervonenkis. Support Vector Regression (SVR) uses the same principle as SVM, but for regression problems. The problem of regression is to find a function that approximates mapping from an input domain to real numbers on the basis of a training sample. The objective, when we are moving on with SVR, is to basically consider the points that are within the decision boundary line. The best fit line is the hyperplane of equation $a_j Z_j^{in} + b = 0$, $\forall j \in \llbracket 1, n \rrbracket$ that has a maximum number of points between the two decision boundary lines.

Thus, to summarize, we are looking for a hyperplane such that most of train observation are inside the marge ϵ satisfying the following constraints

$$\mathcal{U}_j^i - (a_j Z_j^{in} + b) \leq \epsilon, \quad \forall j \in \llbracket 1, n \rrbracket, \text{ if the point is above the margin,} \quad (4.3.8)$$

and

$$(a_j Z_j^{in} + b) - \mathcal{U}_j^i \leq \epsilon, \quad \forall j \in \llbracket 1, n \rrbracket, \text{ if the point is below the margin,} \quad (4.3.9)$$

where

1. $Z^{in} = (Z_j^{in})_{j \in \llbracket 1, n \rrbracket}$ represents the input defined by different initial values of the wealth.
2. (a, b) represent the parameters of weight $(a_i)_i$ and bias b : we choose it randomly at the beginning then we try to find the optimal values in the algorithm.
3. $\mathcal{U}^i = (\mathcal{U}_j^i)_j := (\mathbf{U}(\tau_i, z_j))_j$ represents the target. We remain that the target is defined by a column vector of the matrix $\mathbf{U}(\tau_i, \mathfrak{z}_j)_{i \in \llbracket 1, N \rrbracket, j \in \llbracket 1, n \rrbracket}$ at fixed τ_i .
4. ϵ is the width of the margin of hyperplan.

The idea is to build a plane of equation $ax + b = 0$ with a margin ϵ , in other words, we want to build a hyperplane of width 2ϵ (see figure 4.4) in order to have the maximum of points between the two decision bounds (4.3.8)-(4.3.9) (which is the black points on figure 4.4. The green points on the decision bounds are called the support vectors which play a main role in this problem). By constructing this hyperplane with a margin of tolerance , we can have points outside the

margin (the red points in figure 4.4), so the difference with the classical regression is that Support Vector Regression problem also takes into consideration the points which are outside the margin or the misplaced points (the red points in figure 4.4).

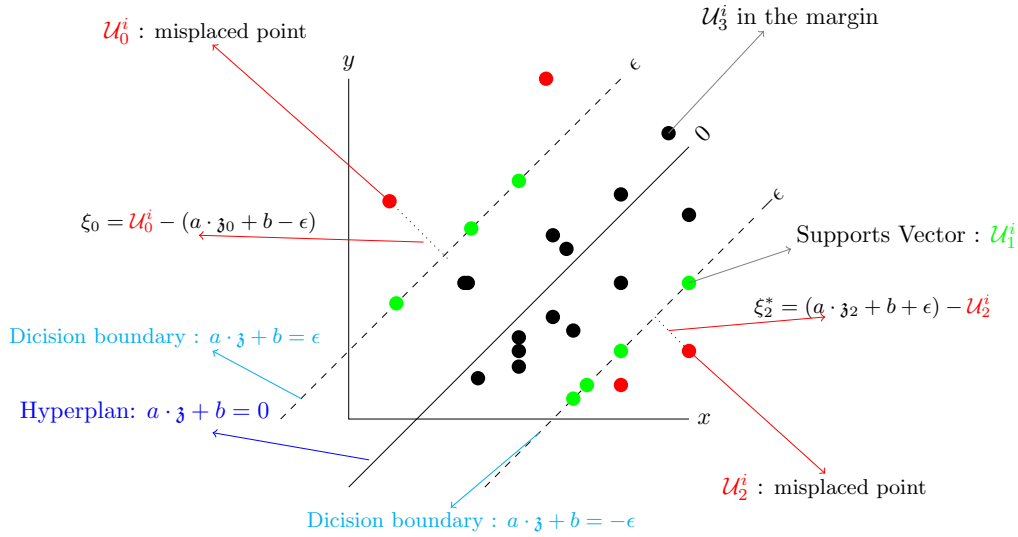


Figure 4.4: Graph illustrating a Support vector regression SVR

Therefore, the constraint can be relaxed by allowing that a point may not be of the same label as the majority of the points as long as it is not too far from the boundary. Any value that lies outside the margin ξ , we can indicate its deviation from the margin as ϵ . In practice, we often cannot have $U_j^i - (a_j z_j + b) > \epsilon$, for all $j \in \llbracket 1, n \rrbracket$, we relax by introducing a positive variables ξ such that

$$U_j^i - (a_j z_j + b) \leq \epsilon + \xi_j, \text{ if the point is above the margin,}$$

and

$$(a_j z_j + b) - U_j^i \leq \epsilon + \xi_j^*, \text{ if the point is below the margin.}$$

The variable $(\xi_j, \xi_j^*)_j$ represent the error of the observations outside the hyperplane,

$$\xi_j := U_j^i - (a_j z_j + b - \epsilon) \geq 0, \text{ for misplaced points above the margin,}$$

and

$$\xi_j^* := (a_j z_j + b - \epsilon) - U_j^i \geq 0, \text{ for misplaced points below the margin.}$$

This problem allow some observations to fall on the wrong side of the margin, but will penalized them by the parameter C . Finally, we can introduce our optimization problem:

$$\min_{a \in \mathbb{R}^n, b \in \mathbb{R}} \left\{ \frac{1}{2} \|a\|^2 + C \sum_{j=1}^l (\xi_j + \xi_j^*) \right\}, \quad (4.3.10)$$

under the following constraint

$$\begin{cases} \mathcal{U}_j^i - (a_j Z_j^{in} + b) \leq \epsilon + \xi_j, \\ (a_j Z_j^{in} + b) - \mathcal{U}_j^i \leq \epsilon + \xi_j^*, \\ \xi_j, \xi_j^* \geq 0, \forall j, \end{cases} \quad (4.3.11)$$

where l is the number of the points outside the margin.

Remarks 4. *The optimization problem (4.3.10)-(4.3.11) can be solved in a numerical way very efficiently by the SVR model described above.*

Basically, the idea of SVR model is to minimize the sum of squared weights by taking into account the misplaced observations in order to maximize the margin.

A second idea is the kernel method. It is an extension to non-linear cases. It is quite simple and relies on the projection of the data in a higher dimensional space in which the problem becomes linear. The kernel method is a mathematical trick to solve this problem elegantly, by defining the transformations ϕ and scalar products via a kernel $k(a, b) = \phi(a)^T \phi(b)$ such that the equation of hyperplan is as follow

$$U_j^p := a_j \phi(Z_j^{in}) + b.$$

Actually, it is far from being easy to identify the transformation function ϕ , so the second problem is rather than choosing ϕ , we choose the kernel function k . A kernel is a function k capable of computing the scalar product $\phi(a)^T \phi(b)$ based only on the original vector a and b without having to compute the transformation ϕ . In practice, this function is used ϕ indirectly, without knowing it. The most used kernels are

- Linear : $k(a, b) = a^T b$.
- Polynomial : $k(a, b) = (\gamma a^T b + r)^d$.
- Gaussian RBF : $k(a, b) = \exp(-\gamma \|a - b\|^2)$.

Nevertheless, there is no rule for choosing the most adapted kernel for a specific problem. This technique is based on distance minimization.

Using the kernel method, the optimization problem (4.3.10)-(4.3.11) becomes

$$\min_{a \in \mathbb{R}^n, b \in \mathbb{R}} \left\{ \frac{1}{2} \|a\|^2 + C \sum_{j=1}^l (\xi_j + \xi_j^*) \right\}, \quad (4.3.12)$$

where l is the number of the points outside the margin and we remain that $\xi_j = \mathcal{U}_j^i - (a_j \phi(z_j) + b - \epsilon)$ and $\xi_j^* = (a_j \phi(z_j) + b + \epsilon) - \mathcal{U}_j^i$, under the following constraint

$$\begin{cases} \mathcal{U}_j^i - (a_j \phi(Z_j^{in}) + b) \leq \epsilon + \xi_j, \\ (a_j \phi(Z_j^{in}) + b) - \mathcal{U}_j^i \leq \epsilon + \xi_j^*, \\ \xi_j, \xi_j^* \geq 0, \forall j. \end{cases} \quad (4.3.13)$$

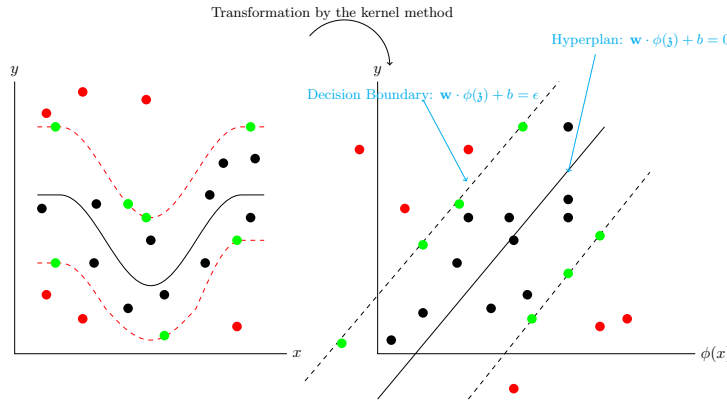


Figure 4.5: Graphic illustration of SVM model using the kernel method.

There are different implementations from the SVR : SVR, NuSVR and LinearSVR. LinearSVR is a SVR with a linear kernel often used for linearly separable data. The SVR and NuSVR models are similar methods, but with slightly different parameter sets and different mathematical formulations. In NuSVR, we introduce a new parameter ν which controls the number of support vectors and margin errors (ξ, ξ^*).

In this work, we implement the algorithm with the two models SVR and NuSVR.

Numerical results of SVR prediction

First, we consider a SVR model, and we take a vector of approximated Utility ($\mathbf{U}(t, z)$) at fixed time t (simulated in previous Section 4.2.2) as a target value \mathcal{U}^i . We wish to predict this target \mathcal{U}^i from the input parameters z .

More precisely, we start by defining our data (Z^{in}, \mathcal{U}^i) . We take as input $Z^{in} := (Z_j^{in})_j = (z_j)_j$

in different initial values of wealth and as target $\mathcal{U}^i = (\mathcal{U}_j^i)_j$ the utility value for a fixed value of time τ_i and different values $(z_j)_j$.

Then, under given parameters $C > 0$ and choosing Gaussian RBF as kernel k , defined by:

$$k(z_i, z_j) = \exp(-\gamma \|z_i - z_j\|^2).$$

This kernel function is typical and very widely used since it allows with an appropriate setting to represent any function. The Gaussian kernel represents a measure of similarity between samples (x_i, x_j) .

We assume that the equation of hyperplan is as follow

$U_j^p := a_j \phi(Z_j^{in}) + b$, where ϕ is the implicit mapping embedded in the RBF kernel $k(x, y)$, see [115].

Then, the equation of the decision boundary becomes

$$a_j \phi(Z_j^{in}) + b = \epsilon \quad , \quad a_j \phi(Z_j^{in}) + b = -\epsilon.$$

Actually, there is no rule to choose the regularization parameter C , we need to just test to find the most adjusted value to keep.

Solving the following optimization problem under the constraint (4.3.11):

$$\min_{a \in \mathbb{R}^n, b \in \mathbb{R}} \left\{ \frac{1}{2} \|a\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) \right\},$$

we find the learned or predicted value denoted by U^p . To evaluate SVR model, we visualize, on figure 4.6, the evolution of the Mean Squared Loss defined by

$$E_{\text{err},k}^i := \frac{1}{2N} \sum_{j=1}^n (\mathcal{U}_j^i - U_j^{p,k})^2, \text{ which is the error at the } k\text{-th iteration.} \quad (4.3.14)$$

We chose to evaluate the error of our learning model for the last vector of the utility matrix in order to have more visibility on the last step of the learning process.

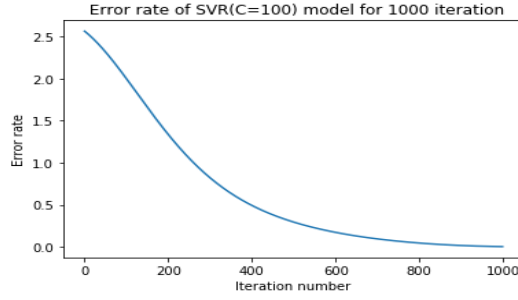


Figure 4.6: Evolution of SVR model's utility learning error at $t_i = t_N$ fixed for 1000 iteration

We can clearly see from Figure 4.6 that the error rate decreases, which ensures that our model has successfully learned. In fact, if the margin of error is small, the results are even more precise and the probability that they are close to the target value is greater. This justifies that the model has learned well.

After implementing Algorithm 3 for the SVR class of the SVM package and taking $C = 100$, we get the following prediction.

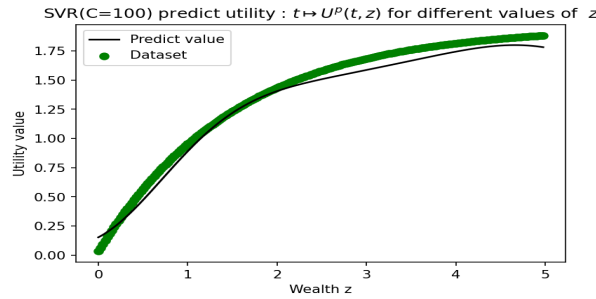


Figure 4.7: Prediction Utility with SVR model for fixed value of τ_i .

Looking to Figure 4.7, we notice that the SVR model is not the best one for our data, even the score $R^2 = 0.8$.

In order to perform some more tests with this learning model, we take, in this time, a column vector of the approximated utility matrix for different values of t and a fixed wealth value z as input data. Will it be possible to predict the dynamics of utility as a function of time t ?

In other words, we try to learn the dynamic of utility. We define our data (T^j, U^j) in following way: as input $T^j := (t_i)_i$ in different values of times between two update τ_i and τ_{i+1} and as target

$$U^j := (\mathbf{U}(\tau_i, \mathfrak{z}_j))_{i \in \llbracket 1, J \rrbracket},$$

where U^j is the j -th vector-line of the generated utility matrix $(\mathbf{U}(\tau_i, \mathfrak{z}_j))_{(i,j) \in \llbracket 1, J \rrbracket \times \llbracket 1, n \rrbracket}$.

Then, with a $R^2 = 0.52$, we can deduce that SVR model is not the best model to predict the dynamic of utility. It is clear from Figure 4.8 that the prediction misses a lot of points which justifies that the value of R^2 is not high.

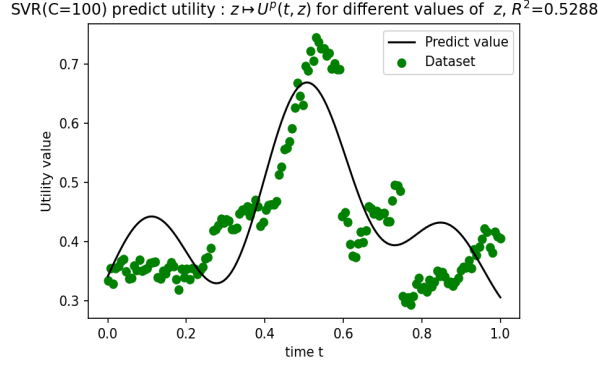


Figure 4.8: Predict dynamic of utility for fixed value of β .

After a multitude of tests, the following table 4.1 gives the values of the various indices of performances concerning the prediction built by SVR. These performances show that is not the best model to predict the agent’s utility. Moreover, the errors (mean squared error, $MSE = 3.38$ and mean absolute error, $MAE = 1.22$) are important which shows that SVR model did not succeed in predicting the utility. In fact, this is due to the linearity of the SVR model which does not allow to find the concavity of our data. In order to have an objective evaluation, we have to calculate the ratio \bar{R}_{test} of the MSE performance of our model and the average of the test data values defined by

$$\bar{R}_{test} := \frac{MSE}{\bar{U}_{test}}, \quad \text{with } \bar{U}_{test} := \frac{1}{n} \sum_{j=1}^n u_{test}. \quad (4.3.15)$$

This ratio must be as close as possible to zero, in other words, the lower the MSE value is compared to the average value, the better it is. It is better that this ratio is lower than 0.4 (0.5 is much higher). Unfortunately this is not the case here. This leads us to conclude that SVR model is not the best fit for our data.

R^2	MSE	MSA	\bar{R}_{test}
0.52	3.38	1.22	0.45

Table 4.1: Performance of SVR model

For this test, we talk about underfitting problem. It is a learning problem that should be avoided. Underfitting means that the predictive model generated during the learning phase does not fit the training set. The SVR model is not able to obtain a sufficiently low error value on the training set.

Numerical results of NuSVR prediction

In this paragraph, we we will test the NuSVR model of SVM package with a regularization parameter $C = 100$.

As before, $Z^{in} := (Z_j^{in})_j = (\mathfrak{z}_j)_j$ is an input data and $\mathcal{U}^i := (\mathbf{U}(\tau_i, \mathfrak{z}_j))_{j \in \llbracket 1, n \rrbracket}$ is the target vector of values. \mathcal{U}^i is the i -th column vector of the utility matrix $(\mathbf{U}(\tau_i, \mathfrak{z}_j))_{(i,j) \in \llbracket 1, J \rrbracket \times \llbracket 1, n \rrbracket}$.

In this model, the parameter ν allows us to control controls the number of support vectors and allows to compute automatically ϵ .

As on SVR model, we take Gaussian RBF as kernel k and for $\nu \in (0, 1]$, we want to solve the following optimization problem for fixed time t :

$$\min_{a \in \mathbb{R}^n, b \in \mathbb{R}} \left\{ \frac{1}{2} \|a\|^2 + C(\nu \epsilon + \frac{1}{l} \sum_{i=1}^l (\xi_i + \xi_i^*)) \right\}, \quad (4.3.16)$$

where l is the number of the points outside the margin where $\xi_j = \mathcal{U}_j^i - (a_j \phi(\mathfrak{z}_j) + b - \epsilon)$ and $\xi_j^* = (a_j \phi(\mathfrak{z}_j) + b + \epsilon) - \mathcal{U}_j^i$, under the following constraint

$$\begin{cases} (a_j \phi(Z_j^{in}) + b) - \mathcal{U}_j^i \leq \epsilon + \xi_j, \\ \mathcal{U}_j^i - (a_j \phi(Z_j^{in}) + b) \leq \epsilon + \xi_j^*, \\ \xi_j, \xi_j^* \geq 0, \forall i. \end{cases} \quad (4.3.17)$$

where we remain that ξ represents the errors of the points outside the hyperplane and ϕ is a transformation maps such that $k(x, y) = \phi(x)^T \phi(y)$.

Remarks 5. • *If ϵ increases, the quantity $\frac{1}{l} \sum_{i=1}^l (\xi_i + \xi_i^*)$ decreases (as less samples outside the margin), the function smoothness $\frac{1}{2} \|a\|$ increases and the accuracy decreases.*

• *If ϵ decreases, the term $\nu \epsilon$ decreases, but the quantity $\frac{1}{l} \sum_{i=1}^l (\xi_i + \xi_i^*)$ increases (as more samples outside the margin), the function $\frac{1}{2} \|a\|$ is less smoothed and the the accuracy increases.*

As for the SVR test model, we visualize the evolution of the error rate, we execute our algorithm in a for loop several times taking the Mean Squared Loss defined by (4.3.14).

After implementing many times, we calculated the error rate each iteration (see figure 4.9), comparing with the error rates for SVR model (see figure 4.6), we can conclude that we succeeded in reducing the error by adding the ν parameter.

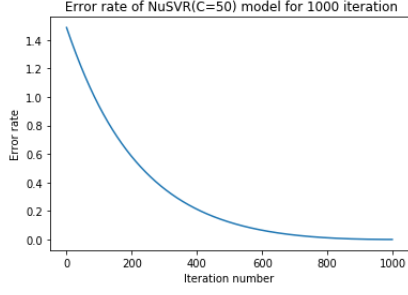


Figure 4.9: Evolution of the NuSVR error at $t_i = T$ fixed for $n = 1000$ iteration

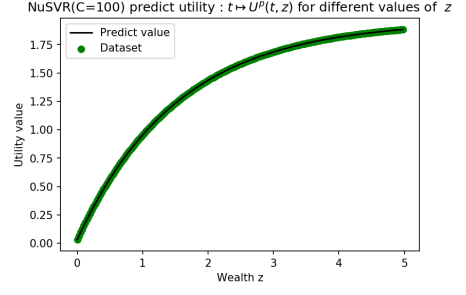


Figure 4.10: Prediction Utility with NuSVR model

By performing several implementations, we succeeded to keep the concavity of the prediction for each one (see Figure 4.10, the impact of the ν parameter is visible which allows to have a small margin of error.

For the following test, we want to learn client's utility at different values of τ . We take different values of wealth \mathfrak{z} in a for loop by changing at each step the values of τ .

Then, with a for loop from $i = 1$ to N , we run the following steps for different values of time τ_i : $(Z_j^{in})_j = (\mathfrak{z})_j$ still the input, the target is $(U_j^i)_j = (U(\tau_i, \mathfrak{z}_j))_j$. Then, we train the model with **Fit** $((Z_j^{in})_j)$ and finally, we test it by **predict** $((Z_j^{in})_j, (U_j^i)_j)$.

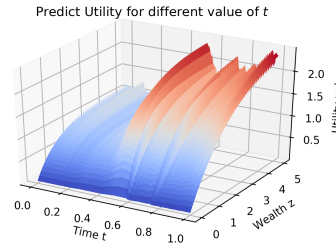


Figure 4.11: Utility's map: $(\tau, \mathfrak{z}) \mapsto U^P(\tau, \mathfrak{z})$

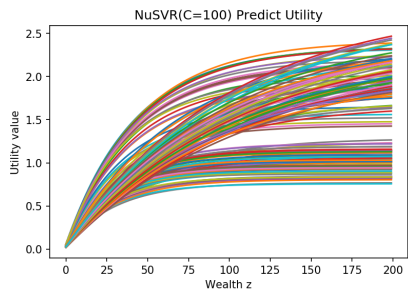


Figure 4.12: $z \mapsto U^P(t, z)$: Concavity of utility for different values of z for $N = 200$.

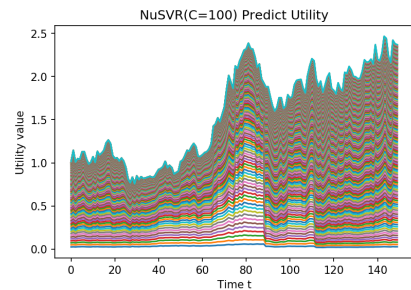


Figure 4.13: Dynamics : $t \mapsto U^P(\tau, \mathfrak{z})$ for different values of t for $n = 150$.

From Figure 4.11, we can see that predict values of utilities for different points of t are concave with respect to z which satisfies the dynamic utility properties. The NuSVR learned better from our data, we find the concavity and Monotonicity of the utility even by making few iterations. It illustrates that by adding the parameter ν , we reduce the learning error (see table 4.2).

Concerning the learning of the utility dynamics with the Neural Network model, we take a data (T^j, \mathcal{U}^j) in the following way: as input $T^j := (t_i)_{i \in \llbracket 1, N \rrbracket}$ for different values of times between two update τ_i and τ_{i+1} and as target

$$\mathcal{U}^j = (\mathcal{U}_i^j)_i = (\mathbf{U}(\tau_i, \mathfrak{z}_j))_{i \in \llbracket 1, N \rrbracket}.$$

We remain that $(\mathbf{U}(\tau_i, \mathfrak{z}_j))_i$ is a line vector with utility value for a fixed value of wealth \mathfrak{z} and different values of $(\tau_i)_{i \in \llbracket 1, N \rrbracket}$.

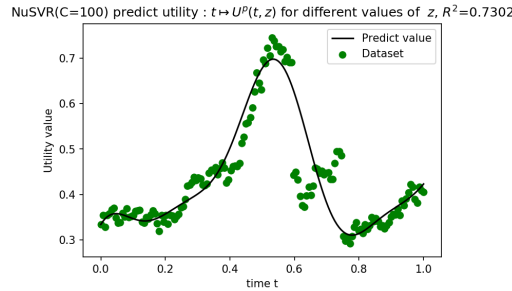


Figure 4.14: Dynamic of Utility for a fixed trajectory at fixed \mathfrak{z} .

We can see from Figure 4.14, that in the prediction there are a lot of missing values, especially for the part $[0, 4, 0, 8]$, the prediction of the dynamics of utility with NuSVR model is not the best model to make the predictions but still it behaves better than SVR model in terms of R^2 coefficient. We get also more important MSE and MAE errors than the SVR model, see Table 4.2. In fact, the presence of ν parameters plays an important role. This parameter represents the lower and upper limit of the number of examples that are on the wrong side of the hyperplane. Essentially, Nu Support Vector regression improves on Support Vector regression by allowing the hyperplan width to automatically fit the data.

R^2	MSE	MSA	\bar{R}_{test}
0.73	2.04	1.06	0.23

Table 4.2: Performance of NuSVR model

In the next section, we will try these learning tests with the Neural Network model, we will see

if it will be better than SVR and NuSVR models.

4.3.3 Learning utility with Neural Network

Neural Network Approach

In this paragraph, we try to learn the investor's utility, using the Neural Network model from the SKlearn library of **Python**. The Neural Network is the most used technique in Machine learning. It is an artificial Neural Network with multiple layers between the input and output. There are different types of Neural Networks and they always consist of the same components: neurons, synapses, weights, biases, and functions. These components functioning similar to the human brains and can be trained like any other Machine Learning algorithm. In general, Neural Network consists of three types of layers: the first one is called the input layer, the last one is called the output layer and the remaining one are called hidden layers.

For the computation implementation, we use the Multi-layer Perceptron classifier from the SKlearn library in **Python**. Multi-layer Perceptron (MLP) is a supervised learning algorithm that learns a function U by training on a dataset. Given a set of features X and a target y , it can learn a non-linear function approximator for either classification or regression. It is different from logistic regression, in that between the input and the output layer, there can be one or more non-linear layers, called hidden layers. MLPRegressor model implements a multi-layer perceptron (MLP) that trains using back-propagation with no activation function in the output layer, which can also be seen as using the identity function as activation function.

For this model, we have three important parameters to set

- Hidden layer sizes, to fix the number of neurons and number of hidden layer.
- Activation function for the hidden layer
 - identity, useful to implement linear bottleneck, returns $f(x) = x$.
 - logistic: the logistic sigmoid function, returns $f(x) = 1/(1 + \exp(-x))$.
 - tanh: the hyperbolic tan function, returns $f(x) = \tanh(x)$.
 - Relu: the rectified linear unit function, returns $f(x) = \max(0, x)$
- The solver for weight optimization:
 - lbfgs is an optimizer in the family of quasi-Newton methods.
 - sgd refers to stochastic gradient descent.
 - adam refers to a stochastic gradient-based optimizer proposed by Kingma, Diederik, and Jimmy Ba.

Example: We develop a simple example of a Neural Network. We introduce a model with three hidden layers and we set the activation function as **Activation=tanh**:

1. **Input layer** : it contains the data $Z^{in} = (Z_j^{in})_j, j = 1, \dots, n_0$.
2. **First hidden layer:** it contains n_1 perceptron having for input $(Z_j^{in})_{j \in \{1, \dots, n_0\}} \in \mathbb{R}^{n_0}$ and for output $Z^{1,i} \in \mathbb{R}^{n_1}$ defined as, for $a^{1,i} \in \mathbb{R}^{n_1, n_0}$ and $b^{1,i} \in \mathbb{R}^{n_1}$,

$$Z_k^{1,i} = \sigma \left(\sum_{j=1, \dots, n_0} Z_j^{in} a_{k,j}^{1,i} + b_k^{1,i} \right), \quad \forall k \in \{1, \dots, n_1\},$$

where we take $\sigma(x) = \tanh(x)$.

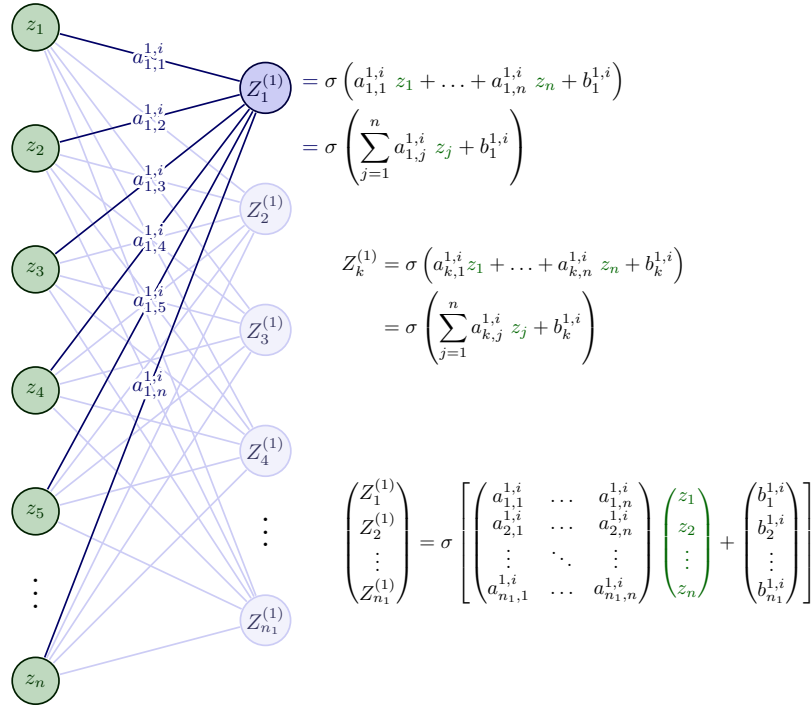


Figure 4.15: Graph illustrating the transition from the input layer to the first hidden layer.

3. **Second hidden layer:** it contains n_2 perceptron having for input $(Z_j^{1,i})_{j \in \{1, \dots, n_1\}} \in \mathbb{R}^{n_1}$ and for output $Z^{2,i} \in \mathbb{R}^{n_2}$ given by , for $a^{2,i} \in \mathbb{R}^{n_2, n_1}$ and $b^{2,i} \in \mathbb{R}^{n_2}$,

$$Z_k^{2,i} = \sigma \left(\sum_{j=1, \dots, n_1} Z_j^{1,i} a_{k,j}^{2,i} + b_k^{2,i} \right), \quad \forall k \in \{1, \dots, n_2\},$$

where we take $\sigma(x) = \tanh(x)$.

4. **Third hidden layer:** it contains n_3 perceptron having for input $(Z_j^{2,i})_{j \in \{1, \dots, n_2\}} \in \mathbb{R}^{n_2}$ and for output $Z_j^{3,i} \in \mathbb{R}^{n_3}$ defined as for $a^{3,i} \in \mathbb{R}^{n_3, n_2}$ and $b^{3,i} \in \mathbb{R}^{n_3}$,

$$Z_j^{3,i} = \sigma \left(\sum_{j=1, \dots, n_2} Z_j^{2,i} a_{k,j}^{3,i} + b_k^{3,i} \right), \quad \forall k \in \{1, \dots, n_3\}$$

where we take $\sigma(x) = \tanh(x)$.

5. **Output layer :** It contains the predict value $U^p \in \mathbb{R}^{n_4}$ given by, for $a^{4,i} \in \mathbb{R}^{n_4, n_3}$ and $b^{4,i} \in \mathbb{R}^{n_4}$,

$$U_k^p = \sigma \left(\sum_{j=1, \dots, n_3} Z_j^{3,i} a_{k,j}^{4,i} + b_k^{4,i} \right) = \sum_{j=1, \dots, n_3} Z_j^{3,i} a_{k,j}^{4,i} + b_k^{4,i}, \quad \forall k \in \{1, \dots, n_4\},$$

with for output layer $\sigma(x) = \text{identity}(x)$.

This corresponds to

$$U_k^p = \sum_{h=1, \dots, n_3} \sigma \left(\sum_{u=1, \dots, n_2} \sigma \left(\sum_{p=1, \dots, n_1} \sigma \left(\sum_{j=1, \dots, n_0} Z_j^{in} a_{p,j}^{1,i} + b_p^{1,i} \right) a_{u,p}^{2,i} + b_u^{2,i} \right) a_{h,u}^{3,i} + b_h^{3,i} \right) a_{k,h}^{4,i} + b_k^{4,i}.$$

where activation function for the hidden layers is $\sigma(x) = \tanh(x)$.

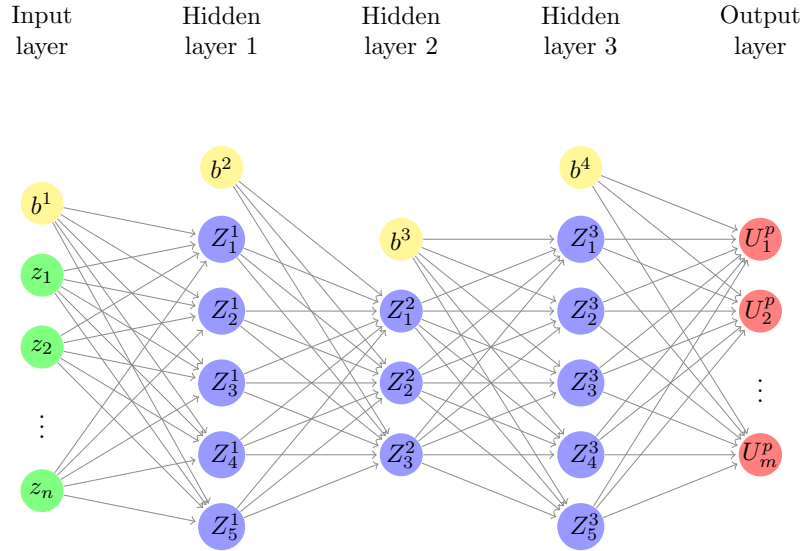


Figure 4.16: Neural Network graph of a 5-layer perceptron with n input units and m output units and 3 hidden layers where $n_1 = 5$, $n_2 = 3$ and $n_3 = 5$.

For the Neural Network, we start from the input Z (first layer) to reach the output U^p (output layer), so that the information flows in a forward way. This is called forward propagation. During training, forward propagation can continue onward until it produces a cost J .

The idea of Machine learning is to let the machine find the parameters $\theta_k^{l,i} := (a, b)$ for the l -th and k -th node layer of the model that minimize the cost function J . In order to do this, we need to use a back-propagation learning algorithm: the most common example of such an algorithm is Gradient Descent. This back-propagation algorithm allows the information from the cost J to flow backward through the network in order to compute the gradient ∇J and to update the set of parameters $\theta_k^{l,i} = (a, b)$ as follows:

$$\theta_{k,j}^{l,i} = \theta_{k,j}^{l,i} - \alpha \frac{\partial J(\theta_k^{l,i})}{\partial \theta_{k,j}^{l,i}}, j = 0, \dots, n_k + 1, \text{ for the } l\text{-th hidden layer and the } k\text{-th node,}$$

where n_l is the number of nodes for the l -th layer and α is a given learning rate.

Actually, the differences between these outputs U^p and the target \mathcal{U} form errors that are corrected via back-propagation, the weights of the Neural Network are then changed. By applying this step several times, the error tends to decrease and the network offers a better fitting or prediction.

Then for a fixed Neural Network structure, number of layers, number of neurons in each layer and fixed activation functions, the program therefore amounts to determine the set of parameters such that

$$\theta^* := \arg \min_{\theta} J(\theta).$$

In this work, we choose the **LBFGS solver** as Back-propagation algorithm of the optimization problem. **LBFGS** is named from the initials of the mathematicians Broyden, Fletcher, Goldfarb and Shanno, who discovered it independently in the late 1960s. This learning algorithm is based on Newton's method which fits the weights using additionally the Hessian matrix H gives the second derivatives of the error function with respect to the weights in the following way:

$$\theta_{k,j}^{l,i} = \theta_{k,j}^{l,i} - \alpha H^{-1} \frac{\partial J(\theta_k^{l,i})}{\partial \theta_{k,j}^{l,i}}, j = 0, \dots, n_k + 1, \text{ for the } l\text{-th hidden layer and the } k\text{-th node.}$$

For this algorithm, we choose to optimize the squared-loss function J defined as

$$J(\theta) := J(a, b) := \frac{1}{2N} \sum_{i=1}^n (\mathcal{U}_i - (a \mathfrak{z}_i + b))^2.$$

in the following way

1. Calculate the gradient and the Hessian of the Cost Function,

$$\frac{\partial J(\theta)}{\partial \theta_i} \quad \text{and} \quad H(J) = \left(\frac{\partial^2 J(\theta)}{\partial \theta_i \partial \theta_j} \right)_{i,j}, \quad \text{with } \theta = (a, b)$$

such that

$$\frac{\partial J(a, b)}{\partial a} = \frac{1}{N} \sum_{i=1}^n (\mathcal{U}_i - (a \mathfrak{z}_i + b)) \mathfrak{z}_i,$$

$$\frac{\partial^2 J(a, b)}{\partial^2 a} = \frac{1}{N} \sum_{i=1}^n (\mathfrak{z}_i)^2,$$

$$\frac{\partial J(a, b)}{\partial b} = \frac{1}{N} \sum_{i=1}^n (\mathcal{U}_i - (a \mathfrak{z}_i + b)),$$

$$\frac{\partial^2 J(a, b)}{\partial^2 b} = 1,$$

2. Update the parameters $\theta_{k,j}^{l,i}$ to this form:

$$\theta_{k,j}^{l,i} = \theta_{k,j}^{l,i} - \alpha H^{-1} \frac{\partial J(\theta_{k,j}^{l,i})}{\partial \theta_{k,j}^{l,i}}, \quad j = 0, \dots, n_k + 1, \quad \text{for the } l\text{-th hidden layer and the } k\text{-th node.}$$

where α is the learning rate and n_k is the number of nodes for the $(l - 1)$ -th layer.

3. Start the forward propagation again, i.e. we repeat the Neural Network algorithm with the new parameters θ and find a new prediction U^p .

4. If we get

$$|U^p - \mathcal{U}| \leq \epsilon, \quad \text{for a given } \epsilon, \quad (4.3.18)$$

we keep the output U^p .

Else we repeat steps 1, 2 and 3 until verifying the condition (4.3.18).

Numerical results of Neural Network prediction

In this section, as the other models, we want to learn client's utility at different points of time τ . We run algorithm 3 for different values of wealth \mathfrak{z} and for different points of time τ in a for loop. Therefore, we take a data set $((Z_j^{in})_j, (\mathcal{U}_j^i)_j)$ such that input set is give by $(Z_j^{in})_j = (\mathfrak{z})_j$ and the target value $(\mathcal{U}_j^i)_j$ is given by

$$(\mathcal{U}_j^i)_j = (\mathbf{U}(\tau_i, \mathfrak{z}_j))_j, \quad \forall j \in \llbracket 1, n \rrbracket.$$

As the other models, machine succeeded the prediction of utility for different points of time t (see Figure 4.17), we see that Neural Network model is better than the others (SVR and NuSVR models), it succeeded in predicting for values of wealth $z \in [0, 0.5]$.

In fact, for the neural network model, we start with random values of weight a and bias b and work backwards, several times, to improve the output. The difference with the other methods tested in this work, that for the Neural Network model we have local predictions for each layer and each forward passage in the loop. Which allows to decrease the error each iteration and give a better prediction or fitting compared to SVR or NuSVR models.

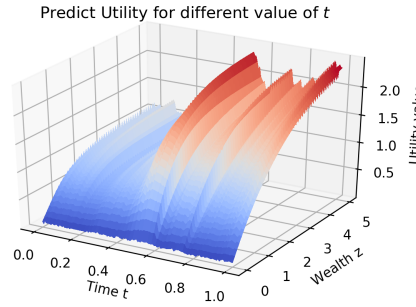


Figure 4.17: Predict utility $(t, z) \rightarrow U^p(t, z)$

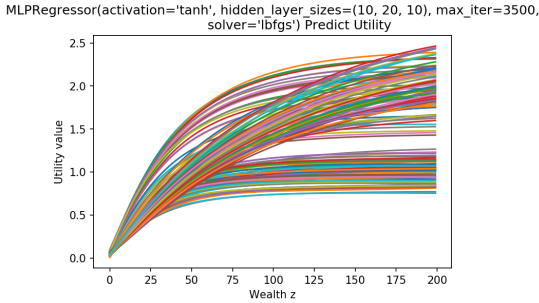


Figure 4.18: $z \mapsto U^p(t, z)$: Concavity of utility for different values of wealth z

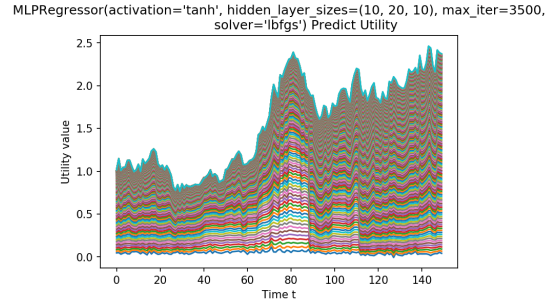


Figure 4.19: Dynamics : $t \mapsto U^p(t, z)$ for different points of time t

Concerning the learning of dynamic of utility with the Neural Network model, we take a data (T^j, \mathcal{U}^j) in the following way: as input $T^j := (\tau_i)_i$ in different values of times and as target

$$\mathcal{U}^j = (\mathcal{U}_i^j)_i = (\mathbf{U}(\tau_i, \mathfrak{z}_j))_i,$$

where $(\mathbf{U}(\tau_i, \mathfrak{z}_j))_i$ is a line vector with utility value for a fixed value of wealth \mathfrak{z}_j and different points of time τ .

We can deduce that the Neural Network prediction is better than the other predictions with

SVR and NuSVR models (see Figure 4.20).

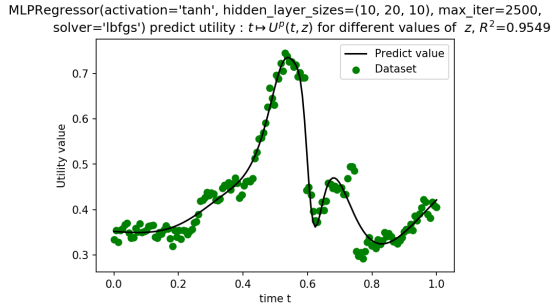


Figure 4.20: Dynamic of Utility for fixed value of z .

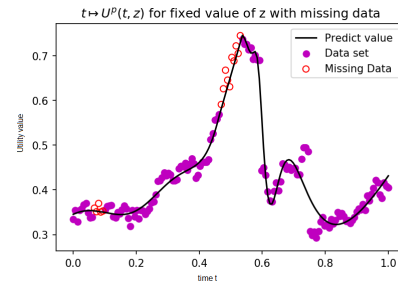


Figure 4.21: Dynamic of Utility for fixed value of z with missing data.

The prediction of this model was more accurate than the other models, with a determination score $R^2 = 0.95$. We can conclude that this is the most successful model for our data. According to Table 4.3, we have a value very close to zero for the rate \bar{R} which is a good indication, that MSE is very low compared to the mean value of the data. Even for making the test with hidden/unseen data the Neural Network model managed to predict better, see Figure 4.21.

R^2	MSE	MSA	\bar{R}_{test}
0.9	0.80	0.56	0.06

Table 4.3: Performance of Neural Network model

The Neural Network model provided good prediction capabilities with respect to the regression problem of SVR models. In fact, with increasing input size the svm model is having lineary increasing number of parameters, on the other hand, Neural network model is not having. Even if we are mainly interested in three hidden layer, Neural Network can have as more than three hidden layers. This implies that a deep Neural Network with the same number of parameters as an SVM is always more complex than the lattest one. Support vector machines effectively use only a subset of a data set as training data. This is because they reliably identify the decision boundary based on only the support vectors. Therefore, for well-separated classes, the number of observations required to train an SVM is not large.

We will dedicate the Section 4.5 to compare the three models used in this work to predict the dynamic utility of an investor.

4.4 Learning utility with missing data

In this part, we want to show the impact of data lack where the agent is not telling all the data to the robo-advisor. We propose a simpler example of the model (4.2.1)-(4.2.2). We consider a pure jump process X whose size depends on its value at the previous time, defined by

$$X_{\tau_{i+1}} = X_{\tau_i} (1 + V_{\tau_i}), \quad \text{for } i = 1, \dots, N, \quad (4.4.19)$$

where N is a random variable representing the jump's number assumed to follow a uniform law on $[0, T]$ and $V = (V_i)_{1 \leq i \leq N}$, a sequence of random variables, represents the jump sizes and $(\tau_n)_n$ a sequence of jump times.

4.4.1 Learning with SVR model

In this section, we take the SVR model. We consider the vector of targeted values \mathcal{U}^i as a fixed value of the dynamic utility $\mathbf{U}(\tau, \mathfrak{z})$ at fixed time τ_i

$$(\mathcal{U}_j^i)_{j \in \llbracket 1, n \rrbracket} := (\mathbf{U}(\tau_i, \mathfrak{z}_j))_{j \in \llbracket 1, n \rrbracket},$$

to solve the optimization problem (4.3.10)-(4.3.11).

Then, for the following test, we hid the part of the data for $z \in [10, 20]$ and $z \in [30, 50]$ in our input data and the target value. We obtain the following prediction with $R^2 = 0.89$.

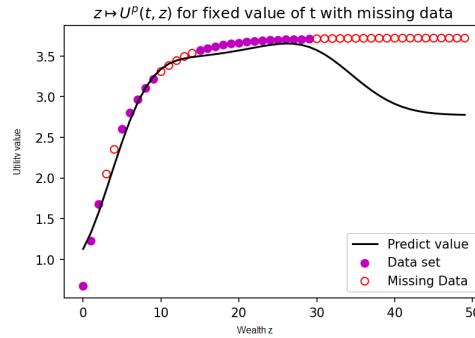


Figure 4.22: Prediction utility with missing data for a fixed t .

In the previous part, we try to learn the dynamic of utility using the SVR model. We define our data (T^j, \mathcal{U}^j) in following way: as input $T^j := (t_i)_i$ in different values of times between two update τ_i and τ_{i+1} and as target $\mathcal{U}^j = (\mathcal{U}_i^j)_i$ the utility value for a fixed value of wealth \mathfrak{z}_j and different values of t :

$$(\mathcal{U}_i^j)_{i \in \llbracket 1, N \rrbracket} := (\mathbf{U}(\tau_i, \mathfrak{z}_j))_{i \in \llbracket 1, N \rrbracket}.$$

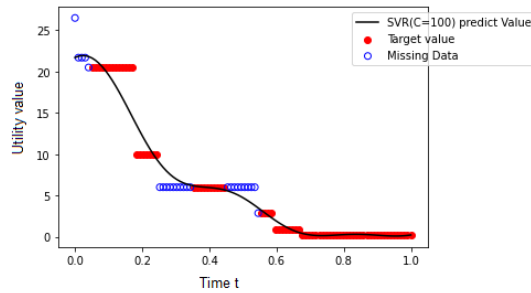


Figure 4.23: Prediction of dynamic utility with missing data for a fixed z .

4.4.2 Learning with NuSVR model

We use the same data (Z^{in}, U^i) to solve the optimization problem (4.3.16)-(4.3.17) associated to NuSVR model.

Similar to the SVR test, we make predictions for the case with missing values for the parts where $z \in [10, 20]$ and $z \in [30, 50]$.

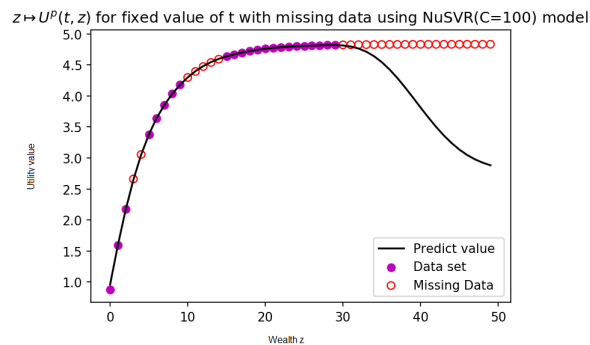


Figure 4.24: Prediction with missing data.

We also test the prediction for dynamic of utility with missing data. According to figure 4.25, we can see that the machine doesn't predict the hidden data. It is clear that the machine misses several values and it is not precise enough.

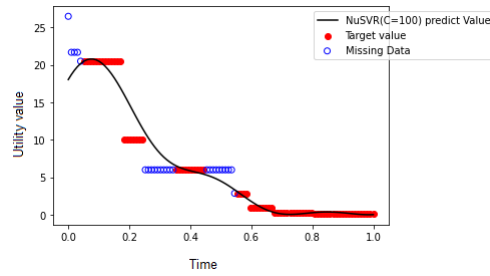


Figure 4.25: Prediction of dynamic utility with missing data.

4.4.3 Learning with Neural Network model

We suppose that the robo-advisor has collected more information than the one communicated by the agent, can it still predict the utility of the client from this data. Roughly speaking, we take a vector of data that contain more samples than the initial one. Then we get the following result, using the same data $(Z^{in}, U^{t,i})$,

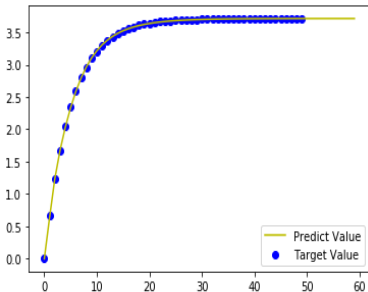


Figure 4.26: Dataset with 60 samples, Implementation time 0.145 seconds

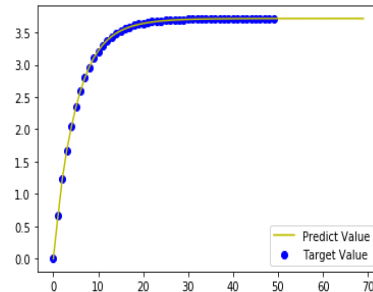


Figure 4.27: Dataset with 70 samples, Implementation time 0.145 seconds

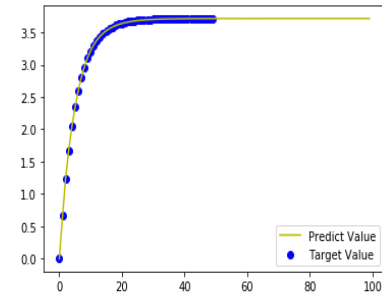


Figure 4.28: Dataset with 100 samples, Implementation time 0.142 seconds

Another case that can be encountered with real data is the missing data. So for the next part, we hide some data, in other words, we don't communicate all the target values to the machine. Then we hid the parts where $z \in [10, 20]$ and $z \in [30, 50]$. Contrary to the other models, we obtain a good prediction for the unseen or missing values, see figure 4.29. The neural network model predicts the hidden values better than the other models, especially the values for $z \in [30, 50]$.

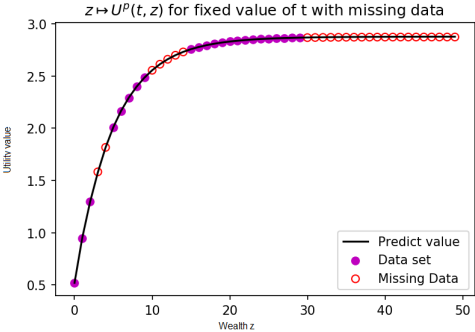


Figure 4.29

For the previous test, we hide the information in the middle and at the end to give the machine the form or appearance of the target value. Thus, we perform this test one more time by hiding the first values of the input and we can realize that the model finds the concavity of utility.

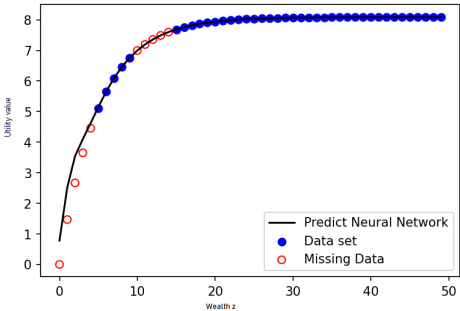


Figure 4.30

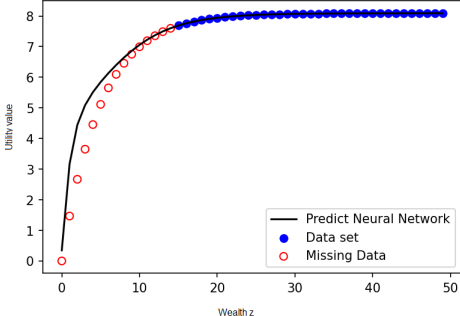


Figure 4.31

4.5 Comparison of machine learning models

In this section, we aim to compare between different methods of machine learning considered in this chapter. Thus, we implemented the three models for the same input values and the same values of targeted utility built the previous section.

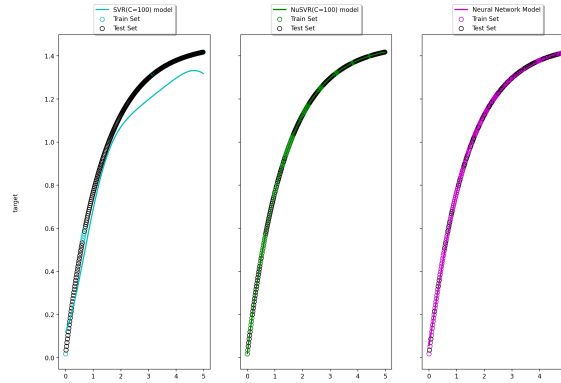


Figure 4.32: Comparison of 3 methods for a prediction of a trajectory for τ_i fixed.

In the following table we represent the values of R^2 obtained for each prediction model for fixed τ_i and the running time of the prediction algorithm for a fixed τ_i to perform the tests. We can see that the values of R^2 are almost the same for the NuSVR and Neural Network models, but that the Neural Network model is faster than the other, it takes only 0.02 seconds to predict utility, compared to 0.40 seconds for the NuSVR model.

	U_{SVR}^p	U_{NuSVR}^p	$U_{Neural\ Network}^p$
R^2 for a fixed τ_i	0.883409	0.999989	0.998356
Prediction's running times	0.054069	0.404845	0.021763

Table 4.4: R^2 and running time for the 3 methods for a fixed trajectory.

For the following test, we implemented the three models for the same input values and the same values of targeted utility, to learn the dynamic of utility for fixed value of z .

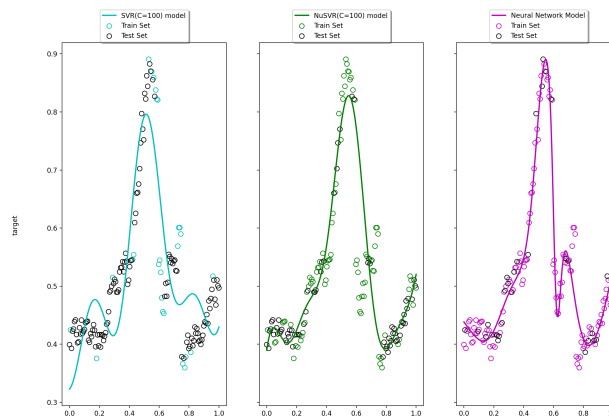


Figure 4.33: Comparison of 3 methods for a prediction of a trajectory at fixed β_j .

In the following table, we represent the values of R^2 obtained for each prediction model

and the compilation time to perform the tests. we can observe that the values of R^2 associated to the Neural Network model is more important than the others and that it is faster than the others, it takes only 0.06 seconds to predict the utility, compared with 0.10 seconds for the NuSVR model.

	U_{SVR}^p	U_{NuSVR}^p	$U_{Neural\ Network}^p$
R^2 for fixed \mathfrak{z}_j	0.580844	0.717010	0.816211
Prediction's Runing times	0.058962	0.101724	0.061751

Table 4.5: R^2 and running time for the 3 methods for a fixed trajectory.

According to Figure 4.33, we can clearly see that the Neural Network class better than the other models. It misses less points when predicting, but the other models clearly miss more points. Actually, the SVR model suffers here from a problem called underfitting, which is a very phenomenon in Machine Learning and that you must absolutely avoid. Underfitting means that the predictive model generated during the learning phase does not fit the training set well. In other words, the predictive model does not even manage to capture the correlations of the training set. Obviously, the predictive model will also not generalize well the unseen data. In fact, we can say that the SVR model is unable to provide accurate predictions. It should be mentioned that the Neural Network model is not an overfitting the data. In fact, the term "overfitting" is used to describe the fact that the model has become too specialized on the data and has lost all sense of generalization. It is clear here that this is not the case, see Figure 4.34.

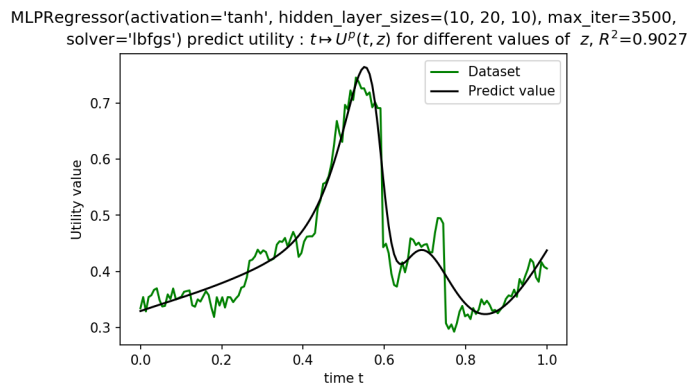


Figure 4.34: Neural Networks prediction vs. target utility.

To compare the prediction of each method, we zoom in on two time interval $[0, 1]$ and $[0.25, 0.6]$ to see which predicted values are closer to the target values.

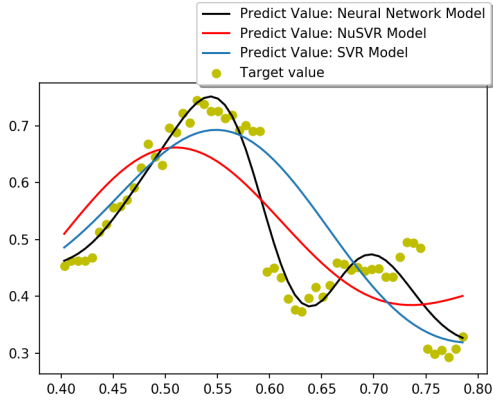


Figure 4.35: Predict utility for $\tau_i \in [0.4, 0.8]$

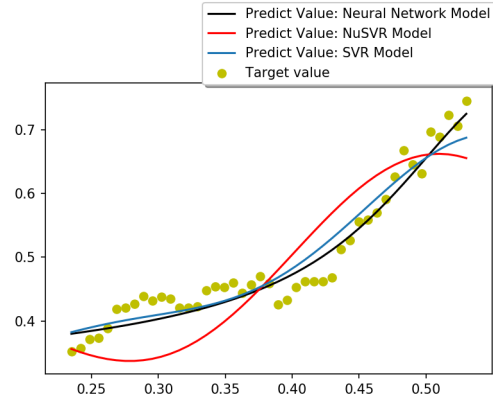


Figure 4.36: Predict utility for $\tau_i \in [0.2, 0.6]$

In the following table, we present some of the predicted values (at randomly selected times) obtained using the 3 methods on the test set:

	$U_{test}^{t,j}$	U_{NuSVR}^p	U_{SVR}^p	$U_{Neural\ Network}^p$
	0.440249	0.427372	0.410366	0.433389
	0.566508	0.541622	0.480728	0.482669
	0.502611	0.694972	0.602699	0.584770
	0.441064	0.416489	0.499655	0.444093
	0.456763	0.430144	0.433427	0.431930
	0.438289	0.436985	0.452408	0.433529
	0.456289	0.431692	0.438223	0.432147
	0.448934	0.471556	0.390601	0.470345
	0.728282	0.754597	0.788956	0.712515
	0.550927	0.526230	0.525660	0.534294
mean value	0.83751	0.83502	0.82775	0.83758

Table 4.6: A sample of the few predicted values via Target value on **Test set** for different models.

The mean of the values predicted by the neural network is very close to the mean of the target values, which reinforces and confirms our conclusions.

Conclusion and perspectives We deduce that the Neural Network model is the best prediction model for our dataset. But the other models are not so bad as well. Unfortunately, there are no clear rules to follow to choose the model to use from the beginning. In fact, it is recommended to start with the easiest and simplest examples, if these models do not work then consider using a more sophisticated model.

Basically, to know which is the best model to use, we have to test all the models and then to choose the one with the best performance, the one, with, a higher determination score R^2 . The data are always very different, we can never define a precise rule to fix the model that works at a hundred percent, that's why we are based on the performance of our models.

In our future work, we will adapt the Neural Network model, the best performing among the tested models, as a learning model in a generalized market model framework. We also wish to generalize our learning algorithms to more advanced models, such as the Recurrent Neural Network and/or Reinforcement Learning.

BIBLIOGRAPHY

- [1] ACHDOU, Y., CAMILLI, F., AND CAPUZZO-DOLCETTA, I. Mean field games: convergence of a finite difference method. *SIAM Journal on Numerical Analysis* 51, 5 (2013), 2585–2612.
- [2] ACHDOU, Y., CAMILLI, F., AND CAPUZZO-DOLCETTA, I. Mean field games: convergence of a finite difference method. *SIAM Journal on Numerical Analysis* 51, 5 (2013), 2585–2612.
- [3] ACHDOU, Y., AND CAPUZZO-DOLCETTA, I. Mean field games: Numerical methods. *SIAM Journal on Numerical Analysis* 48, 3 (2010), 1136–1162.
- [4] ACHDOU, Y., AND CAPUZZO-DOLCETTA, I. Mean field games: Numerical methods. *SIAM Journal on Numerical Analysis* 48, 3 (2010), 1136–1162.
- [5] ACHDOU, Y., AND PORRETTA, A. Convergence of a finite difference scheme to weak solutions of the system of partial differential equations arising in mean field games. *SIAM Journal on Numerical Analysis* 54 (04 2015).
- [6] ALASSEUR, C., BEN TAHAR, I., AND MATOUSSI, A. An extended mean field game for storage in smart grids. *arXiv preprint arXiv:1710.08991* (2017).
- [7] ANDERSON, D., DJEHICHE, B. A maximum principle for sdes of mean-field type. *Appl Math Optim* 63 (2010), 341–356.
- [8] ANGIULI, A., GRAVES, C. V., LI, H., CHASSAGNEUX, J.-F., DELARUE, F., AND CARMONA, R. Numerical Probabilistic Approach to MFG. *ESAIM: Probability and Statistics* (2019).
- [9] ARROW, K.-J., CHANG, S. Optimal pricing, use, and exploration of uncertain resource stocks. *JOURNAL OF ENVIRONMENTAL ECONOMICS AND MANAGEMENT* (1982), 10.
- [10] ARTHUR, W. Complexity and the economy. *Science* 284, 5411 (1999), 107–109.
- [11] BASEI, M., PHAM, H. A weak martingale approach to linear-quadratic mckean vlasov stochastic control problems. *Journal of Optimization Theory and Applications* (2019).

-
- [12] BAYRAKTAR, E., BUDHIRAJA, A., AND COHEN, A. A numerical scheme for a mean field game in some queueing systems based on markov chain approximation method. *arXiv preprint arXiv:1708.08343* (2017).
- [13] BEN AZIZA, S., AND TOUMI, S. A probabilistic numerical method for a class of mean field games. *Stochastics and Dynamics* (2021).
- [14] BEN AZIZA, S., AND ZIRI, C. Numerical approach for mean field games. *In progress* (2022).
- [15] BENSOUSSAN, A. Lectures on stochastic control. *Springer-Verlag 972* (1983), 1–62.
- [16] BENSOUSSAN, A., AND FREHSE, J. Control and nash games with mean field effect. *Chin. Ann. Math. Se* (2013), 34(2):161–192.
- [17] BENSOUSSAN, A., FREHSE, J., AND YAM, P. *Mean field games and mean field type control theory*, vol. 101. Springer, 2013.
- [18] BENSOUSSAN, A., FREHSE, J., AND YAM, S. C. P. The master equation in mean field theory. *Journal de Mathématiques Pures et Appliquées* 103, 6 (2015), 1441–1474.
- [19] BILLINGSLEY, P. *Convergence of probability measures*. John Wiley & Sons, 2013.
- [20] BUCKDAHN, R., DJEHICHE, B., AND LI, J. A general stochastic maximum principle for sdes of mean-field type. *Applied Mathematics Optimization* 64 (10 2011), 197–216.
- [21] BUCKDAHN, R., LI, J., AND MA, J. A stochastic maximum principle for general mean-field systems. *Applied Mathematics Optimization* 74 (12 2016).
- [22] BUCKDAHN, R., LI, J., AND MA, J. A mean-field stochastic control problem with partial observations. *The Annals of Applied Probability* 27 (02 2017).
- [23] BUCKDAHN, R., LI, J., PENG, S. Mean-field backward stochastic differential equations and related partial differential equations. *Stochastic Processes and their Applications*, 119 (2009), 3133–3154.
- [24] BULIRSCH, J. S. *Introduction to Numerical Analysis*, vol. 12. 2002.
- [25] CAINES, HUANG, M. Large-population cost-coupled lqg problems with nonuniform agents: individual-mass behavior and decentralized nash equilibria. *IEEE Trans.* 52, 9 (2006), 1560–1571.
- [26] CAMILLI, F., AND SILVA, F. A semi-discrete approximation for a first order mean field game problem. *Netw. Heterog. Media* (2012), 7(2):263–277.

-
- [27] CARDALIAGUET, P. Notes on mean field games. notes from pl lions' lectures at the college de france, 2012.
- [28] CARDALIAGUET, P., DELARUE, F., LASRY, J.-M., AND LIONS, P.-L. The master equation and the convergence problem in mean field games. *arXiv e-prints* (2015).
- [29] CARLINI, E., AND SILVA, F. J. A fully discrete semi-lagrangian scheme for a first order mean field game problem. 45–67.
- [30] CARLINI, E., AND SILVA, F. J. A semi-lagrangian scheme for a degenerate second order mean field game system. 4269–4292.
- [31] CARMONA, R. Applications of mean field games in financial engineering and economic theory, 2020.
- [32] CARMONA, R. Applications of mean field games in financial engineering and economic theory.
- [33] CARMONA, R., AND DELARUE, F. *Probabilistic Theory of Mean Field Games with Applications I-II*. Springer, 2018.
- [34] CARMONA, R., DELARUE, F. The master equation for large population equilibriums. *Springer International Publishing* (2014), 77–128.
- [35] CHAJEWSKA, U., KOLLER, D., AND ORMONEIT, D. Learning an agent's utility function by observing behavior. In *ICML* (2001), pp. 35–42.
- [36] CHAMBERS, C., AND ECHENIQUE, F. *Revealed preference theory*, vol. 56. Cambridge University Press, 2016.
- [37] CHAN, P., SIRCAR, R. Bertrand and cournot mean field games. *Appl. Math. Optim* 71, 3 (2015), 533–569.
- [38] CHAN, P., SIRCAR, R. Fracking, renewables and mean field games. *SIAM Review* 59, 3 (2017), 588–615.
- [39] CHASSAGNEUX, J.-F., CRISAN, D., AND DELARUE, F. Numerical method for fbsdes of mckean-vlasov type, 2017.
- [40] DELARUE, F., AND MENOZZI, S. A forward–backward stochastic algorithm for quasi-linear PDEs. *The Annals of Applied Probability* 16, 1 (2006), 140 – 184.
- [41] DESHMUKH, S-D., PLISKA, S-R. Optimal consumption and exploration of nonrenewable resources under uncertainty. *Econometrica* 48 (Jan 1980), 177–200.

-
- [42] DJEHICHE, B., ELIE, R., AND HAMADENE, S. Mean-field reflected backward stochastic differential equations. *The Annals of Applied Probability* (2019).
- [43] DJEHICHE, B., AND HAMADÈNE, S. Mean-field backward-forward stochastic differential equations and nonzero sum stochastic differential games. *arXiv preprint arXiv:1904.06193* (2019).
- [44] EL KAROUI, N., HILLAIRET, C., AND MRAD, M. Consistent utility of investment and consumption: a forward/backward spde viewpoint. *Stochastics* 90, 6 (2018), 927–954.
- [45] EL KAROUI, N., AND MRAD, M. An Exact connection between two Solvable SDEs and a Non Linear Utility Stochastic PDEs. *SIAM JOURNAL on Financial Mathematics* 4, 1 (2013), 697–736.
- [46] EL KAROUI, N., AND MRAD, M. Recover Dynamic Utility from Observable Process: Application to the economic equilibrium. working paper or preprint, 2020.
- [47] EL KAROUI, N. Les aspects probabilistes du controle stochastique: In 9th saint flour probability summer school-1979. lecture notes in mathematics. *Springer* 876 (1981), 73–238.
- [48] FUJIWARA, T., KUNITA, H., ET AL. Stochastic differential equations of jump type and lévy processes in diffeomorphisms group. *Journal of mathematics of Kyoto University* 25, 1 (1985), 71–106.
- [49] FÜRNKRANZ, J., AND HÜLLERMEIER, E. *Preference learning*. Springer, 2011.
- [50] GÈRON, A. *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, 1st ed. O'Reilly Media, Inc., 2017.
- [51] GOBET, E., AND MRAD, M. Convergence rate of strong approximations of compound random maps, application to spdes. *Discrete & Continuous Dynamical Systems-B* 23, 10 (2018), 4455.
- [52] GOMES, D. A., ET AL. Mean field games models—a brief survey. *Dynamic Games and Applications* 4, 2 (2014), 110–154.
- [53] GOMES, D. A., PATRIZI, S., AND VOSKANYAN, V. On the existence of classical solutions for stationary extended mean field games. *Nonlinear Analysis: Theory, Methods & Applications* 99 (2014), 49–79.
- [54] GOMES, D. A., PIRES, G. E., AND SÁNCHEZ-MORGADO, H. A-priori estimates for stationary mean-field games. *NHM* 7, 2 (2012), 303–314.

-
- [55] GOMEZ-RAMIREZ, J. Don't blame the economists. it is an inverse problem! *European JOURNAL of Futures Research* 1, 1 (2013), 1.
- [56] GOODFELLOW, I., BENGIO, Y., AND COURVILLE, A. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- [57] GRABER, P-J. Linear quadratic mean field type control and mean field games with common noise, with application to production of an exhaustible resource. *Appl. Math. Optim* 74, 3 (2016), 459–486.
- [58] GUÉANT, O. New numerical method for mean field games with quadratic costs. 315–336.
- [59] GUÉANT, O., LASRY, J-M., LIONS, P-L. Mean field games and applications. paris-princeton lectures 30 on mathematical finance 2010. *Springer* (2011), 205–266.
- [60] HAGAN, P., WOODWARD, D., CAFLISCH, R., KELLER, J. Optimal pricing, use and exploration of uncertain natural resources. *Applied Mathematical Finance* 1 (1994), 87–108.
- [61] HENDERSON, V., AND HOBSON, D. Horizon-unbiased utility functions. *Stochastic Process. Appl.* 117, 11 (2007), 1621–1641.
- [62] HIBBARD, B. Model-based utility functions. *Journal of Artificial General Intelligence* 3, 1 (2012), 1–24.
- [63] HOTELLING, H. The economics of exhaustible resources. *journal of Political Economy* 39, 2 (1931), 137–175.
- [64] HUANG, J., LI, X., YONG, J. A linear-quadratic optimal control problem for mean-field stochastic differential equations in infinite horizon. *Mathematical Control and Related Fields* 5 (2012).
- [65] KAC, M. Foundations of Kinetic Theory. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* 3 (1956), 171–197, University of California Press, Berkeley, Calif.
- [66] KORN, R. The martingale optimality principle : The best you can is good enough. *WILMOTT* (2003).
- [67] KUNITA, H. Stochastic differential equations based on lévy processes and stochastic flows of diffeomorphisms. In *Real and stochastic analysis*. Springer, 2004, pp. 305–373.
- [68] KUSHNER, H. Optimal stochastic control. *IRE Trans. Auto. Control* AC-7 (1962), 120–122.

-
- [69] KUSHNER, H. On the stochastic maximum principle: Fixed time of control. *J. Math. Anal. Appl.* 11 (1965), 78–92.
- [70] KUSHNER, H., AND DUPUIS, P. G. *Numerical methods for stochastic control problems in continuous time*, vol. 24. Springer Science & Business Media, 2013.
- [71] KUSHNER, H. J. Numerical methods for stochastic control in continuous time. *Control and Optimizations SIAM J.* (2013), 28, 999–1048.
- [72] LACHAPELLE, A., SALOMON, J., AND TURINICI, G. Computation of mean field equilibria in economics. *Mathematical Models and Methods in Applied Sciences* 20 (04 2010).
- [73] LASRY, J.-M., AND LIONS, P.-L. Jeux à champ moyen. i—le cas stationnaire. *Comptes Rendus Mathématique* 343, 9 (2006), 619–625.
- [74] LASRY, J.-M., AND LIONS, P.-L. Jeux à champ moyen. i—le cas stationnaire. *Comptes Rendus Mathématique* 343, 9 (2006), 619–625.
- [75] LASRY, J.-M., AND LIONS, P.-L. Jeux à champ moyen. ii—horizon fini et contrôle optimal. *Comptes Rendus Mathématique* 343, 10 (2006), 679–684.
- [76] LASRY, J.-M., AND LIONS, P.-L. Jeux à champ moyen. ii—horizon fini et contrôle optimal. *Comptes Rendus Mathématique* 343, 10 (2006), 679–684.
- [77] LASRY, J.-M., AND LIONS, P.-L. Mean field games. *Japanese journal of mathematics* 2, 1 (2007), 229–260.
- [78] LASRY, J.-M., AND LIONS, P.-L. Mean field games. *Japanese journal of mathematics* 2, 1 (2007), 229–260.
- [79] LAURIERE, M., PIRONNEAU, O. Dynamic programming for mean-field type control, comptes rendus mathematique. *SIAM journal on Control and Optimization* 352 (2014), 707–713.
- [80] LI, J. Reflected mean-field backward stochastic differential equations. approximation and associated nonlinear pdes. *Journal of Mathematical Analysis and Applications* 413, 1 (2014), 47–68.
- [81] LI, J., LIANG, H., AND ZHANG, X. General mean-field bsdes with continuous coefficients. *Journal of Mathematical Analysis and Applications* 466 (06 2018).
- [82] LI, J., AND MIN, H. Controlled mean-field backward stochastic differential equations with jumps involving the value function. *Journal of Systems Science and Complexity* 29 (05 2016).

-
- [83] LI, J. Mean-field forward and backward SDEs with jumps and associated nonlocal quasi-linear integral-PDEs. *Stochastic Processes and their Applications* 128, 9 (2018), 3118 – 3180.
- [84] LI, X., SUN, J., XIONG, J. Linear quadratic optimal control problems for mean-field backward stochastic differential equations. *Appl Math Optim* <https://doi.org/10.1007/s00245-017-9464-7> (Dec 2017).
- [85] LIONS, P.-L. Cours au college de france,. *Stochastics: An International Journal of Probability and Stochastic Processes* 35, 1 (1991), 31–62.
- [86] LUDKOVSKI, M., SIRCAR, R. Exploration and Exhaustibility in Dynamic Cournot Games. *European Journal of Applied Mathematics* 23 (2011), 343–372.
- [87] MATOUSSI, A., MANAI, A., AND SALHI, R. Mean-field backward-forward sde with jumps and storage problem in smart grids. *arXiv preprint arXiv:1906.08525* (2019).
- [88] MATOUSSI, A., MNIF, M., AND ZIRI, C. Mean field control problems and application to production of an exhaustible resource. *Preprint* (2021).
- [89] MATOUSSI, A., AND MRAD, M. Dynamic Utility and related nonlinear SPDE driven by Lévy Noise. working paper or preprint.
- [90] MCKEAN, H. A CLASS OF MARKOV PROCESSES ASSOCIATED WITH NONLINEAR PARABOLIC EQUATIONS. 6, 56 (1966), 1907–1911.
- [91] MILLER, E., PHAM, H. Linear-quadratic mckean-vlasov stochastic differential games. *Modeling, Stochastic Control, Optimization, and Applications* 164. Springer, Cham (2019), 451–481.
- [92] MORDECKI, E., SZEPESSY, A., TEMPONE, R., AND ZOURARIS, G. E. Adaptive weak approximation of diffusions with jumps. *SIAM Journal on Numerical Analysis* 46, 4 (2008), 1732–1768.
- [93] MRAD, M. Solving some stochastic partial differential equations driven by lévy noise using two sdes.
- [94] MRAD, M., AND ZIRI, C. Learning dynamic utilities. *In progress* (2022).
- [95] MUSIELA, M., AND ZARIPHPOULOU, T. Backward and forward utilities and the associated pricing systems: The case study of the binomial model. *Preprint* (2003).

-
- [96] MUSIELA, M., AND ZARIPHOPOULOU, T. *Investment and Valuation Under Backward and Forward Dynamic Exponential Utilities in a Stochastic Factor Model*. 01 2007, pp. 303–334.
- [97] MUSIELA, M., AND ZARIPHOPOULOU, T. Portfolio choice under space-time monotone performance criteria. *SIAM JOURNAL on Financial Mathematics* 1, 1 (2010), 326–365.
- [98] MUSIELA, M., AND ZARIPHOPOULOU, T. Stochastic partial differential equations and portfolio choice. 195–216.
- [99] NA, T. Y. *Computational methods in engineering boundary value problems*, vol. 145. 1979.
- [100] NICOLE EL KAROUÏ, CHRISTOPHE KAPOUDJIAN, E. P. S. P., AND QUENEZ, M.-C. Reflected solutions of backward sde's, and related obstacle problems for pde's. *The Annals of Applied Probability* (1997), 702–737.
- [101] NIELSEN, T., AND JENSEN, F. V. Learning a decision maker's utility function from (possibly) inconsistent behavior. *Artificial Intelligence* 160, 1 (2004), 53–78.
- [102] OKSENDAL, B., ZHANG, T., ET AL. The itô-ventzell formula and forward stochastic differential equations driven by poisson random measures. *Osaka Journal of Mathematics* 44, 1 (2007), 207–230.
- [103] PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., MICHEL, V., THIRION, B., GRISEL, O., BLONDEL, M., PRETTENHOFER, P., WEISS, R., DUBOURG, V., VANDERPLAS, J., PASSOS, A., COURNAPEAU, D., BRUCHER, M., PERROT, M., AND DUCHESNAY, E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* 12 (2011), 2825–2830.
- [104] PENG, S. A general stochastic maximum principle for optimal control problem. *SIAM J. Control Optim.* 4 (1990), 966–979.
- [105] PHAM, H. *Continuous-time stochastic control and optimization with financial applications*. Stochastic Modelling and Applied Probability n°61. Springer, 2009.
- [106] PHAM, H., AND WEI, X. Dynamic programming for optimal control of stochastic mckean-vlasov dynamics.
- [107] PINDYCK, R-S. The optimal exploration and production of nonrenewable resources. *Journal of Political Economy* 86, 5 (1978), 841–861.
- [108] QI, Y., XU, M., AND LAFFERTY, J. Learning high-dimensional concave utility functions for discrete choice models. *Cited on* (2014), 26.

-
- [109] R. BUCKDAHN, B. DJEHICHE, J. L. S. P. Mean-field backward stochastic differential equations: a limit approach. *Statistics & Probability Letters* 37, 4 (2009), 1524–1565.
- [110] SAINT-CIRGUE, G. Apprendre le machine learning en une semaine. <https://machinelearning.com/apprendre-le-machine-learning-en-une-semaine/> (2019).
- [111] SAMUELSON, P. A note on the pure theory of consumer's behavior. *Economica* 5, 17 (1938), 61–71.
- [112] SAMUELSON, P. Consumption theory in terms of revealed preference. *Economica* 15, 60 (1948), 243–253.
- [113] STAHL, A. Approximation of utility functions by learning similarity measures. In *Logic versus Approximation*. Springer, 2004, pp. 150–172.
- [114] TANG, M., MENG, Q. . Linear-quadratic optimal control problems for mean-field backward stochastic differential equations with jumps. *arXiv:1611.06434* (Nov 2016).
- [115] VERT, J.-P., TSUDA, K., AND SCHÖLKOPF, B. A primer on kernel methods. *Kernel methods in computational biology* 47 (2004), 35–70.
- [116] WILLIAM, E., TESAURO, G., KEPHART, J. O., AND DAS, R. Utility functions in autonomic systems. In *1st International Conference on Autonomic Computing (ICAC 2004), 17-19 May 2004, New York, NY, USA* (2004), pp. 70–77.
- [117] YONG, J. Linear-quadratic optimal control problems for mean-field stochastic differential equations. *SIAM journal on Control and Optimization* 51 (2013), 2809–2838.
- [118] YONG, J., ZHOU, X. *Stochastic Controls: Hamiltonian Systems and HJB Equations*. 1999.

Titre : Contrôle stochastique et applications : types Champ moyen et utilités dynamiques

Mots clés : Problème de contrôle de type mean field, énergies épuisables, jeux à champ moyen, utilités dynamiques, machine Learning

Résumé : Cette thèse est consacrée à l'étude des contrôles stochastiques et leurs applications. Dans le premier chapitre, on a étudié un problème de contrôle linéaire quadratique type mean field dans le cadre avec saut. On a utilisé une approche basée sur une formulation faible du principe d'optimalité des martingales standard dans la théorie du contrôle stochastique afin de prouver l'existence d'un contrôle optimal. Nous avons étendu le résultat trouvé pour un modèle de production d'énergies épuisables avec extraction. Dans le deuxième chapitre, on s'est intéressé à la résolution numérique d'un problème des jeux à champ moyen (Mean field Game). La caractérisation des stratégies optimales des agents en interaction est obtenue via l'étude d'un système d'EDP non linéaires couplées permettant la détermination de la caractéristique de l'agent ainsi que la distribution de ces caractéristiques dans la population. Nous avons présenté un schéma numérique pour déterminer le couple

(u, m) associé à la fonction de valeur et à la densité de probabilité, en se basant sur un argument de point fixe du système couplé de l'équation de Hamilton Jacobi Bellman et l'équation Fokker Planck. Nous avons appliqué notre approche pour un exemple linéaire quadratique avec une solution explicite afin de comparer les deux solutions numérique et explicite. Dans le troisième chapitre, on introduit un algorithme pour « prédire » l'utilité dynamique d'un agent en observant ses décisions à des dates discrètes (éventuellement aléatoires). C'est le principe des Robo-advisors qui sont des services de gestion d'investissement en ligne qui utilisent des algorithmes mathématiques pour fournir des services financiers avec un minimum d'intervention humaine. L'idée de ce travail est l'apprentissage de l'utilité dynamique d'un investisseur sur un marché financier avec défauts en utilisant des différentes méthodes de Machine Learning.

Title : Stochastic control and applications : Mean field types and dynamic utilities

Keywords : Mean field control problem, exhaustible energy production, mean field games, Markov chain approximation, dynamic utilities, machine learning

Abstract : This thesis is devoted to the study of stochastic controls and their applications. In the first chapter, we studied a linear quadratic mean field control problem in the framework with jump. We used an approach based on a weak formulation of the optimality principle of standard martingales to prove the existence of an optimal control. We have extended the result found for a model of exhaustible energy production with extraction. In the second chapter, we are interested in the numerical solution of a mean field game problem. The characterization of the optimal strategies of the interacting agents is obtained by studying a system of coupled nonlinear PDEs allowing the determination of the agent's characteristics as well as the distribution of these characteristics in the population. We have presented a numerical

scheme to determine the couple (u, m) associated to the value function and the probability density, based on a fixed point argument of the coupled system of the Hamilton Jacobi Bellman equation and the Fokker Planck equation. We applied our approach for a linear quadratic example with an explicit solution in order to compare the two numerical and explicit solutions. In the third chapter, we introduce an algorithm to "predict" the dynamic utility of an agent by observing its decisions at discrete dates. This is the principle of Robo-advisors which are online investment management services that use mathematical algorithms to provide financial services with minimal human intervention. The idea of this work is to learn the dynamic utility of an investor in a financial market with defaults using different Machine Learning methods.