



**HAL**  
open science

# Jeux différentiels stochastiques non-Markoviens et dynamiques de Langevin à champ-moyen

Kaitong Hu

► **To cite this version:**

Kaitong Hu. Jeux différentiels stochastiques non-Markoviens et dynamiques de Langevin à champ-moyen. Probabilités [math.PR]. Institut Polytechnique de Paris, 2020. Français. NNT : 2020IP-PAX005 . tel-02918519

**HAL Id: tel-02918519**

**<https://theses.hal.science/tel-02918519v1>**

Submitted on 20 Aug 2020

**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.



INSTITUT  
POLYTECHNIQUE  
DE PARIS

NNT : 2020IPPAX005

Thèse de doctorat



# Jeux Différentiels Stochastiques non-Markoviens et Dynamiques de Langevin à Champ-moyen

Thèse de doctorat de l'Institut Polytechnique de Paris  
préparée à l'École Polytechnique

École doctorale n°574 Ecole Doctorale de Mathématiques  
Hadamard (EDMH)  
Spécialité de doctorat : Mathématiques Appliquées

Thèse présentée et soutenue à Palaiseau, le 18 Juin 2020, par

**Kaitong Hu**

Composition du Jury :

Johannes Muhle-Karbe Professeur, Imperial College London	Rapporteur
Stéphane Villeneuve Professeur, Toulouse School of Economics	Rapporteur
Caroline Hillairet Professeur, ENSAE ParisTech	Président du jury
Gilles Pagès Professeur, Université Pierre et Marie Curie	Examineur
Jean-François Chassagneux Professeur, Université Paris Diderot	Examineur
Zhenjie Ren Maître de Conférence, Université Paris Dauphine	Examineur
Nizar Touzi Professeur, Ecole Polytechnique	Directeur de thèse



---

# CONTENTS

<b>Abstract</b>	<b>iv</b>
<b>I Introduction</b>	<b>1</b>
<b>Introduction</b>	<b>1</b>
I.1 Problème de Principal-Agent en Temps Continu . . . . .	1
I.1.1 Prologue : aperçu général . . . . .	1
I.1.2 Résultats de la littérature . . . . .	4
I.1.3 Extension aux cas dégénérés . . . . .	6
I.1.4 Intermède : EDS progressives-rétrogrades dépendantes de la trajectoire . . . . .	7
I.1.5 Le problème de Principal-Agent avec plusieurs Principals . . . . .	9
I.2 La Dynamique de Langevin à Champ-moyen, Applications aux Réseaux de Neurones . . . . .	11
I.2.1 Motivation : apprentissage profond . . . . .	11
I.2.2 La dynamique de Langevin à champ-moyen . . . . .	14
I.2.3 Les systèmes de Langevin à champ-moyen . . . . .	17
<b>II Continuous-time Principal-Agent Problem in Degenerate Systems</b>	<b>21</b>
II.1 Introduction . . . . .	22
II.2 Principal-Agent Problem in Degenerate Systems . . . . .	24
II.2.1 Preliminaries . . . . .	24
II.2.2 Problem Formulation . . . . .	25
II.2.3 Motivation . . . . .	26
II.2.4 Main Results . . . . .	29
II.3 Proofs . . . . .	32
II.4 Applications . . . . .	39
II.4.1 Principal-Agent Problem in Partially Observed Linear System . . . . .	39

II.4.2	Mean Field Interacting Agents in Partially Observed Linear System . . . . .	41
II.4.2.1	The Setting of Mean Field Interacting Agent Problem . . . . .	42
II.4.2.2	The Agent's Problem . . . . .	43
II.4.2.3	Solving the Principal's Problem . . . . .	44
II.5	Appendix . . . . .	46
<b>III</b>	<b>Principal-Agent Problem with Multiple Principals</b>	<b>50</b>
III.1	Introduction . . . . .	51
III.2	Moral hazard problem : $n$ principals and one agent . . . . .	53
III.2.1	Agent's problem . . . . .	54
III.2.2	Principals' problem : time inconsistency . . . . .	57
III.3	Infinite number of principals : mean field approach . . . . .	59
III.3.1	Heuristic analysis . . . . .	59
III.3.2	Mean field game : existence . . . . .	60
III.3.3	Mean field game : approximation . . . . .	64
III.3.3.1	Agent problem : backward propagation of chaos . . . . .	64
III.3.3.2	Principals' problem . . . . .	69
<b>IV</b>	<b>The Wellposedness of Path-dependent Multidimensional Forward-backward SDE</b>	<b>71</b>
IV.1	Introduction . . . . .	72
IV.2	Notations and General Assumptions . . . . .	74
IV.3	Decoupling Field for Multidimensional Path-Dependent FBSDE . . . . .	75
IV.4	Main Results . . . . .	76
IV.4.1	Wellposedness of FBSDE on a Small Time Interval . . . . .	76
IV.4.2	Existence of Regular Decoupling Field . . . . .	77
IV.4.2.1	Variational FBSDE, Characteristic BSDE and Dominating ODE . . . . .	78
IV.4.2.2	Decoupled Path-dependent FBSDE . . . . .	81
IV.4.2.3	When $b = b_t(x, y)$ and $\sigma = \sigma_t(x)$ . . . . .	81
IV.4.2.4	When $\sigma = \sigma_t(x, y)$ . . . . .	82
IV.4.2.5	General Case . . . . .	84
IV.5	Stability Property of Path-dependent Multidimensional FBSDE . . . . .	85
IV.6	Proofs . . . . .	89
<b>V</b>	<b>Mean-Field Langevin Dynamics and Energy Landscape of Neural Networks</b>	<b>94</b>
V.1	Introduction . . . . .	95
V.1.1	Organisation of the Chapter . . . . .	98
V.1.2	Theoretical Contributions and Literature Review . . . . .	98
V.1.3	Calculus on the Space of Probability Measures . . . . .	100

---

V.2	Main Results . . . . .	101
V.3	Application to Gradient Descent of Neural Networks . . . . .	105
V.3.1	Fully connected 1-hidden layer neural network . . . . .	106
V.3.2	Gradient Descent . . . . .	107
V.4	Free Energy Function . . . . .	108
V.5	Mean Field Langevin Equations . . . . .	113
V.6	Convergence to the Invariant Measure . . . . .	115
V.7	Appendix . . . . .	121
<b>VI</b>	<b>Mean-field Langevin System, Optimal Control and Deep Neural Networks</b>	<b>124</b>
VI.1	Introduction . . . . .	125
VI.2	Preliminaries . . . . .	128
VI.2.1	Regularized Relaxed Optimal Control . . . . .	128
VI.2.2	System of Mean-Field Langevin Equations . . . . .	129
VI.2.3	Notation . . . . .	130
VI.3	Main Results . . . . .	130
VI.3.1	Wellposedness of the System of Mean-Field Langevin Equation . . . . .	131
VI.3.2	Convex Case . . . . .	132
VI.3.3	Contraction Case . . . . .	133
VI.4	Application to Deep Neural Networks . . . . .	135
VI.4.1	Discretization of mean-field Langevin dynamics and gradient descent . . . . .	135
VI.4.2	A model for studying the dropout . . . . .	138
VI.5	Wellposedness of the system of mean-field Langevin equations . . . . .	141
VI.5.1	Wellposedness of the System . . . . .	141
VI.5.2	Gradient Flow . . . . .	145
VI.6	Proof for the Convex Case . . . . .	148
VI.6.1	Sufficient First Order Condition . . . . .	148
VI.6.2	Convergence Towards the Invariant Measure . . . . .	149
VI.7	Proofs for the Contraction Case . . . . .	150
	<b>Acknowledgements</b>	<b>156</b>
	<b>Bibliographie</b>	<b>158</b>

## Résumé

Cette thèse se compose de deux parties indépendantes et la première regroupant deux problématiques distinctes.

Dans la première partie, nous étudions d'abord le problème de Principal-Agent dans des systèmes dégénérés, qui apparaissent naturellement dans des environnements à l'observation partielle où l'Agent et le Principal n'observent qu'une partie du système, e.g. [56, 156, 136]. Nous présentons une approche se basant sur le principe du maximum stochastique, dont le but est d'étendre les travaux existants qui utilisent le principe de la programmation dynamique dans des systèmes non-dégénérés, e.g. [139, 41, 36]. D'abord nous résolvons le problème du Principal dans un ensemble des contrats élargi donné par la condition du premier ordre du problème de l'Agent sous forme d'une équation différentielle stochastique progressive-rétrograde (abrégée EDSPR) dépendante de la trajectoire. Ensuite nous utilisons la condition suffisante du problème de l'Agent pour vérifier que le contrat optimal obtenu est bien implémentable. Une étude parallèle est consacrée à l'existence et l'unicité de la solution des EDSPRs dépendantes de la trajectoire dans le chapitre IV. Nous étendons la méthode de champ de découplage, e.g. [39, 163, 110], aux cas où les coefficients des équations peuvent dépendre de la trajectoire du processus forward. Nous démontrons également une propriété de stabilité pour ce genre d'EDSPRs. Enfin, nous étudions le problème de hasard moral avec plusieurs Principals. L'Agent ne peut travailler que pour un seul Principal à la fois et fait donc face à un problème de switching optimal. En utilisant la méthode de randomisation nous montrons que la fonction valeur de l'Agent et son effort optimal sont donnés par un processus d'Itô. Cette représentation nous aide à résoudre ensuite le problème du Principal lorsqu'il y a une infinité de Principals en équilibre selon un jeu à champ-moyen. Nous justifions la formulation à champ-moyen par un argument de propagation de chaos.

La deuxième partie de cette thèse est constituée des chapitres V et VI. La motivation de ces travaux est de donner un fondement théorique rigoureux pour la convergence des algorithmes du type descente de gradient très souvent utilisés dans la résolution des problème non-convexes comme la calibration d'un réseau de neurones.

Pour les problèmes non-convexes du type réseaux de neurones à une couche cachée, l'idée clé est de transformer le problème en un problème convexe en le relevant dans l'espace des mesures. Nous montrons que la fonction d'énergie correspondante admet un unique minimiseur qui peut être caractérisé par une condition du premier ordre utilisant la dérivation dans

l'espace des mesures au sens de Lions. Nous présentons ensuite une analyse du comportement à long terme de la dynamique de Langevin à champ-moyen, qui possède une structure de flot de gradient dans la métrique de 2-Wasserstein. Nous montrons que le flot de la loi marginale induite par la dynamique de Langevin à champ-moyen converge vers une loi stationnaire en utilisant le principe d'invariance de LaSalle, qui est le minimiseur de la fonction d'énergie.

Dans le cas de réseaux de neurones profonds, nous les modélisons à l'aide d'un problème de contrôle optimal en temps continu. Nous donnons d'abord la condition du premier ordre à l'aide du principe de Pontryagin, qui nous aidera ensuite à introduire le système d'équation de Langevin à champ-moyen, dont la mesure invariante correspond au minimiseur du problème de contrôle optimal. Enfin, avec la méthode de couplage par réflexion nous montrons que la loi marginale du système de Langevin à champ-moyen converge vers la mesure invariante avec une vitesse exponentielle.

**Mots-Clés :** contrôle stochastique non-Markovien, principe du maximum stochastique, EDS progressive-rétrograde dépendante de la trajectoire, problème de Principal-Agent, théorie du contrat, switching optimal, jeu à champ-moyen, propagation de chaos, dynamiques de Langevin à champ-moyen, flot du gradient, réseau de neurones



## Abstract

Two independent subjects are studied in this thesis, the first of which consists of two distinct problems.

In the first part, we begin with the Principal-Agent problem in degenerate systems, which appear naturally in a partially observed random environment in which the Agent and the Principal can only observe one part of the system, see e.g. [56, 156, 136]. Our approach is based on stochastic maximum principle, the goal of which is to extend the existing results using dynamic programming principle, see e.g. [139, 41, 36], to the degenerate case. We first solve the Principal's problem in an enlarged set of contracts given by the first order condition of the Agent's problem in form of a path-dependent forward-backward stochastic differential equation (abbreviated FBSDE). Afterwards, we use the sufficient condition of the Agent's problem to verify that the previously obtained optimal contract is indeed implementable. At the same time, a parallel study is devoted to the wellposedness of path-dependent FBSDEs in the chapter IV. We generalize the decoupling random field method, e.g. [39, 163, 110], to the case where the coefficients of the equations can depend on the whole path of the forward process and we show the stability property of this type of FBSDEs. Finally, we study the Principal-Agent problem with multiple Principals. The Agent can only work for one Principal at a time and therefore needs to solve an optimal switching problem. By using randomization, we show that the value function of the Agent's problem and his optimal control are given by an Itô process. This representation allows us to solve the Principal's problem in the mean-field case when there is an infinite number of Principals. We justify the mean-field formulation using an argument of backward propagation of chaos.

The second part of the thesis consists of chapter V and VI. The motivation of this work is to give a rigorous theoretical underpinning for the convergence of gradient-descent type of algorithms frequently used in non-convex optimization problems like calibrating a deep neural network.

For one-layer neural networks, the key insight is to convexify the problem by lifting it to the measure space. We show that the corresponding energy function has a unique minimiser which can be characterized by some first order condition using derivatives in measure space. We present a probabilistic analysis of the long-time behavior of the mean-field Langevin dynamics, which have a gradient flow structure in 2-Wasserstein metric. By using a generalization

of LaSalle's invariance principle, we show that the flow of marginal laws induced by the mean-field Langevin dynamics converges to the stationary distribution, which is exactly the minimiser of the energy function. As for deep neural networks, we model them as some continuous-time optimal control problems. Firstly, we find the first order condition by using Pontryagin maximum principle, which later helps us find the associated mean-field Langevin system, the invariant measure of which is again the minimiser of the optimal control problem. As last, by using the reflection coupling techniques, we show that the marginal distribution of the mean-field Langevin system converges to the unique invariant measure exponentially.

**Keywords :** stochastic control of non-Markovian systems, stochastic maximum principles, path-dependent forward-backward SDEs, Principal-Agent problem, contract theory, optimal switching, mean field games, propagation of chaos, mean-field Langevin dynamics, gradient flow, neural networks



---

---

# CHAPITRE I

---

## INTRODUCTION

### I.1 Problème de Principal-Agent en Temps Continu

#### I.1.1 Prologue : aperçu général

Le problème de Principal-Agent modélise une situation de délégation entre deux agents dans un environnement aléatoire, en établissant les incitations nécessaires afin de contourner le risque de moralité. Un exemple classique du problème de Principal-Agent est comment trouver un contrat optimal fixant les termes d'une rémunération incitative entre les investisseurs d'un fond d'investissement et le gestionnaire du portefeuille. Les applications du problème de Principal-Agent se trouvent dans beaucoup de domaines comme la finance d'entreprise [15, 11] ou la transition énergétique [2, 1].

Le problème consiste à trouver un contrat  $\xi$  qui maximise l'utilité du Principal et motive suffisamment l'Agent à travailler pour le Principal. La manière que le contrat puisse être optimisé dépend cruciallement de l'information disponible pour le Principal. Dans la littérature classique du problème de Principal-Agent, nous distinguons trois cas : le partage du risque (first best), le hasard moral (second best) et la sélection adverse (third best).

Pour mieux illustrer les notions précédentes, regardons un exemple dans le cas discret. Soit  $X_1^a$  la production finale donnée par

$$X_1^a = x_0 + a + B$$

avec  $B$  une variable aléatoire fixée,  $a$  l'effort de l'Agent et  $x_0$  la production initiale. On suppose que les préférences de l'Agent sont caractérisées par un critère d'utilité espérée de fonction d'utilité

$$U_A(\xi - c(a)) := -\frac{1}{\gamma_A} e^{-\gamma_A(\xi - \frac{a^2}{2})}, \quad (I.1.1)$$

où  $c$  est la fonction de coût de l'Agent supposée quadratique et  $\xi$  est la somme que l'Agent reçoit à la maturité du contrat. On dit qu'un contrat  $\xi$  est *admissible* s'il permet de procurer à l'Agent un niveau de satisfaction supérieure à un seuil pré-défini, ce qui est appelé la contrainte de participation. Dans la suite on ne considère que les contrats admissibles et on note  $\Xi$  l'ensemble de ces contrats. Les préférences

du Principal sont également caractérisées par un critère d'utilité espérée avec fonction d'utilité

$$U_P(X_1^a - \xi) := -\frac{1}{\gamma_P} e^{-\gamma_P(X_1^a - \xi)}. \quad (1.1.2)$$

**Partage du risque** Le Principal et l'Agent ont la même information et ils doivent décider comment partager le risque. Le Principal a tout le pouvoir de négociation et il dicte l'action de l'Agent. Mathématiquement, le partage optimal de risque est formulé par le problème d'optimisation s'écrit

$$\sup_{\xi, a} \mathbb{E} \left[ U_P(X_1^a - \xi) + \lambda U_A(\xi - c(a)) \right],$$

où  $\lambda$  est le facteur du partage du risque. Formellement si on dérive par rapport à  $\xi$ , à l'optimum la condition du premier ordre donne

$$\frac{U'_P(X_1 - \xi)}{U'_A(\xi - c(a))} = \lambda.$$

On appelle l'égalité ci-dessus la règle de Borch, voir [17]. Remplaçant les fonctions d'utilité par leur définition (1.1.1) et (1.1.2), pour une action  $a$  donnée, le contrat optimal est le suivant :

$$\xi = \rho \left( \gamma_P X_1^a + \frac{\gamma_A a^2}{2} + \log(\lambda) \right)$$

avec  $\rho := \frac{1}{\gamma_A + \gamma_P}$ . Remarquons que la sensibilité du contrat par rapport à la valeur finale de la production  $X_1$  est égale à  $\frac{\gamma_P}{\gamma_P + \gamma_A}$  qui est inférieure à 1. On peut l'interpréter de la manière suivante : un Agent qui est très risque-averse ne veut pas être exposé à l'incertitude et un Agent qui est risque neutre accepte d'absorber la variance en échange du cash. En dérivant la fonction valeur par rapport à l'effort de l'Agent  $a$ , l'effort optimal pour la valeur collective est donné par  $a^* = 1$ , qui ne dépend ni de la valeur initial  $x_0$ , ni du niveau d'aversion au risque  $\gamma_A$ . Ce phénomène est lié aux fonctions d'utilité exponentielles.

**Aléa moral (moral hazard)** Dans le cas d'aléa moral on suppose que l'effort de l'Agent n'est plus observable par le Principal. La seule source d'information provient du processus de production. Revoyons le cas du fond d'investissement, il est très difficile pour un investisseur, qui joue le rôle de Principal, d'observer la composition exacte du portefeuille géré par l'Agent, *i.e.* le gestionnaire. Deuxièmement, On suppose que l'Agent est rationnel et qu'il choisit l'effort qui optimise sa fonction d'utilité étant donné un contrat admissible  $\xi \in \Xi$ , *i.e.*

$$a^*(\xi) = \operatorname{argmax}_a \mathbb{E} \left[ U_A \left( \xi - \frac{a^2}{2} \right) \right].$$

Le Principal anticipe la réponse optimale de l'Agent puis choisit à son tour un contrat qui maximise sa fonction d'utilité. Son problème s'écrit de la manière suivante :

$$V_P = \sup_{\xi \in \Xi} \mathbb{E} \left[ U_P(X_T^{a^*(\xi)} - \xi) \right].$$

Mirrlees in [119] a montré qu'en général on ne peut espérer l'existence d'un contrat optimal sous les conditions ci-dessus. Pour simplifier on se restreint aux contrats linéaires en fonction de  $X_1^a$ , autrement dit, il existe  $k_0$  et  $k_1$  constants telles que  $\xi = k_0 + k_1 X_1^a$ . Pour ce contrat l'effort optimal de l'Agent est  $a^*(\xi) = k_1$ , proportionnel à la sensibilité du contrat par rapport à la production finale. Soit  $R_0$  l'utilité que l'Agent reçoit étant donné le contrat  $\xi$ . On a

$$R_0 = -\frac{1}{\gamma_A} e^{-\gamma_A \left( k_0 + \frac{(a^*)^2}{2} - \frac{(a^*)^2}{2} \gamma_A \right)}.$$

Avec l'équation ci-dessus, on remarque que l'on peut paramétrer le contrat  $\xi$  en utilisant  $R_0$  et  $k_1$ , et le problème du Principal devient

$$\mathbb{E} [U_P(X_1 - \xi)] = \frac{(-\gamma_A R_0)^{\frac{\gamma_P}{\gamma_A}}}{-\gamma_P} e^{-\gamma_P \left( -\frac{1}{2} \gamma_P (a^* - 1)^2 - a^* (a^* - 1) + \frac{(a^*)^2}{2} - \frac{(a^*)^2}{2} \gamma_A \right)}.$$

En dérivant par rapport à  $a^*$ , on obtient facilement l'action optimale de l'Agent qui maximise la fonction d'utilité du Principal, qui est donnée

$$a^* = \frac{1 + \gamma_P}{1 + \gamma_A + \gamma_P},$$

et par conséquent, le contrat optimal correspondant est donné par

$$\xi = k_0 + \frac{1 + \gamma_P}{1 + \gamma_A + \gamma_P} X_1.$$

La constante  $k_0$  est choisie de telle sorte que le contrat satisfait la contrainte de participation.

**Sélection adverse** Dans beaucoup de situations il est raisonnable de supposer que le Principal ne connaît pas les caractéristiques de l'Agent. C'est par exemple la situation d'un vendeur de voitures face à ses clients. L'information sur la richesse de ses clients n'est pas disponible pour le vendeur, pourtant c'est une information crucial pour optimiser le contrat. Les agents se distinguent ainsi par leurs caractéristiques individuelles, appelées *type*. Les préférences des agents sont encore caractérisées par une utilité espérée de fonction d'utilité  $U_A(\xi - c(a), \theta)$  paramétrée par son type  $\theta$  non-observable par le Principal. Pour remédier à cette asymétrie d'information, le Principal propose un menu de contrats révélateurs paramétrés et l'Agent de type  $\theta$  révèle son type en choisissant le contrat  $\xi(\theta)$ , qui est optimal pour l'Agent parmi tous les contrats dans le menu. Mathématiquement, La contrainte de révélation s'écrit

$$\max_{\tilde{\theta}} V_A(\xi(\tilde{\theta}), \theta) = V_A(\xi(\theta), \theta),$$

où  $V_A(\xi(\tilde{\theta}), \theta) = \max_a \mathbb{E} [U_A(\xi(\tilde{\theta}) - c(a), \theta)]$ . On suppose que la distribution des types d'Agent  $F(\theta)$  est connu par le Principal. Notons  $\Xi_R$  l'ensemble des contrats admissibles qui sont aussi révélateurs. Alors

le problème du Principal s'écrit

$$V_P = \max_{\xi \in \Xi_R} \int \mathbb{E} [U_P(\xi(\theta), a^*(\xi(\theta)))] F(d\theta).$$

Cvitanic et Zhang [40] résolvent le problème par calcul de variation en se restreignant aux contrats linéaire en  $X_1$ , *i.e.*  $\xi(\theta) = k_0(\theta) + k_1(\theta)X_1$  et en supposant que la distribution des types d'Agent est uniforme sur un intervalle  $[\theta_L, \theta_U]$ . On doit aussi mentionner que dans le cas déterministe, Ekeland [50] a montré que le problème de sélection adverse a un lien important avec le problème de transport optimal. Il a montré que la contrainte de révélation correspond au fait que le transport entre les types d'Agent et les contrats soit optimal. Pour plus de détail sur ce sujet, voir [101, 15, 40, 37].

Dans le cas de hasard moral et de la sélection adverse, il est difficile de résoudre le problème du Principal à moins qu'on n'arrive à résoudre le problème de l'Agent explicitement pour chaque contrat comme illustré par les calculs précédents dans le cas discret. Nous allons voir dans la suite qu'en passant en temps continu, le problème de Principal-Agent se simplifie grâce aux outils du calcul stochastique et de la théorie des martingales.

## 1.1.2 Résultats de la littérature

Dans cette thèse nous nous concentrons sur le problème de hasard moral en temps continu. Le premier modèle en temps continu est proposé dans le papier de Holmström et Milgrom [78], dans lequel ils étudient un cas simplifié de hasard moral à horizon fini où l'Agent ne contrôle que le drift du processus d'état. Ils montrent que le contrat optimal doit être linéaire en fonction de la valeur finale du processus d'état quand la fonction d'utilité de l'Agent est exponentielle. Leur travaux ont été généralisés par Schättler et Sung [140, 141], Sung [146, 147], Müller [123, 124], Hellwig et Schmidt [73]. Plus tard, Williams [155], Cvitanic, Wan et Zhang [38] utilisent le principe du maximum stochastique et les équations différentielles stochastiques progressives-rétrogrades (ou FBSDEs, l'abréviation de forward-backward stochastic differential equations en anglais) pour caractériser le contrat optimal pour une famille plus générale de fonction d'utilité.

Beaucoup de modèles en temps continu ont été proposés pendant ces dizaines d'années, notamment, celui introduit par Sannikov dans son papier [139]. Il étudie le problème de hasard moral à horizon infini avec paiement continu. De plus, le Principal peut terminer le contrat et virer l'Agent à tout moment et l'agent peut de même démissionner à un moment de son choix. Mathématiquement parlant, en utilisant le principe de la programmation dynamique Sannikov a mis en évidence la relation contractuelle entre le Principal et l'Agent, sa méthode peut être vue comme une procédure de calcul pour trouver le contrat optimal qui est relativement simple.

Plus récemment, l'idée de Sannikov a été réinterprétée et étendue par Cvitanic, Possamaï et Touzi [36] dans un cadre plus général avec une approche plus simple et plus directe. Nous allons illustrer leur

contribution avec le modèle simplifié. Considérons le problème d'optimisation pour l'Agent suivant

$$\max_{\alpha} \mathbb{E} \left[ \xi(X^{\alpha}) - \int_0^T c(\alpha_t) dt \right], \quad \text{où} \quad dX_t^{\alpha} = dW_t - \alpha_t dt.$$

L'observation cruciale dans [36] est de constater que le contrat  $\xi$  et le contrôle optimal de l'Agent  $\alpha^*[\xi]$  peuvent être tous les deux caractérisés par l'équation différentielle stochastique rétrograde (ou backward stochastic differential equation en anglais, abrégée dans la suite BSDE) suivante

$$dY_t = -c^*(Z_t)dt + Z_t dW_t, \quad Y_T = \xi, \quad \text{où} \quad c^*(z) = \max_a \{az - c(a)\}. \quad (1.1.3)$$

Nous nous référons à [41, 52, 129] pour une présentation détaillée sur la théorie des BSDEs et leurs applications. Dans notre cas, comme  $\xi = Y_T$  et  $\alpha_t^*[Z] = \operatorname{argmax}_a \{aZ_t - c(a)\}$  pour tout  $t \in [0, T]$ , nous obtenons une représentation naturelle du couple  $(\xi, \alpha^*[\xi])$  :

$$\begin{cases} \xi = Y_T^{Y_0, Z} := Y_0 - \int_0^T c^*(Z_t)dt + \int_0^T Z_t dW_t \\ \alpha_t^*[\xi] := \alpha_t^*[Z] = \operatorname{argmax}_a \{aZ_t - c(a)\}, \quad \text{for all } t \in [0, T] \end{cases} \quad \text{pour certain } (Y_0, Z).$$

La contrainte de participation devient simplement  $Y_0 \geq R$  où  $R$  est la niveau minimale de satisfaction de l'Agent. Avec cette représentation, le problème du Principal est devenu un problème de contrôle stochastique classique, *i.e.*

$$\max_{\xi \in \Xi} \mathbb{E} \left[ U \left( X_T^{\alpha^*[\xi]} - \xi \right) \right] = \max_{Y_0 \geq R, Z} \mathbb{E} \left[ U \left( X_T^{\alpha^*[Z]} - Y_T^{Y_0, Z} \right) \right].$$

Cette idée de représentation s'applique naturellement aux études du cas de multiples agents. Initialement le modèle multi-agents à une période a été étudié par Holmström [77], Mookherjee [122], Green et Stokey [65], Harris, Kriebel et Raviv [68], Nalebuff et Stiglitz [125], Demski et Sappington [44]. Des extensions aux modèles en temps continu ont été faites par Koo, Shim et Sung [96]. Plus récemment Elie, Possamaï [55] étudient le cas où les agents sont en compétition. Étant donnés  $N$  contrats  $\xi_1, \dots, \xi_N$  respectivement, les  $N$  agents atteignent un optimum lorsque leurs stratégies forment un équilibre de Nash. Dans le cas d'utilité exponentielle Elie et Possamaï montrent que l'existence d'un équilibre de Nash pour  $N$ -agents est équivalente à l'existence d'une solution d'un système d'EDSR quadratique en utilisant la représentation du couple  $(\xi^i, \alpha_i^*[\xi])$  montrée ci-dessus pour chaque agent  $i = 1, \dots, N$ .

Toutefois, la difficulté apparaît lorsque le nombre d'agents devient grand. Contrairement au cas de dimension 1, l'existence et l'unicité de solution d'un système d'EDSR quadratique ne sont pas toujours garanties, comme montré par Frei et Dos Reis [59], Xing et Žitković [158], Harder et Richou [69]. Dans le papier [54], Elie, Mastrolia, Possamaï proposent une approche basant sur la formulation de jeu à champ-moyen, introduite par Larsy et Lions [102], Huang, Caine et Malhamé [86]. Nous nous référons à Carmona et Delarue [25, 26] pour une description complète de ce sujet. Elie, Mastrolia, Possamaï étudient d'abord



la situation où il y a une infinité d'agents et montrent que les équilibres de Nash sont caractérisés par une EDSR à champ-moyen de dimension 1, dont l'existence de solution ont été étudiée dans Carmona et Lacker [27]. En se restreignant à l'ensemble des contrats représentables par cette EDSR à champ-moyen, le problème du Principal est devenu un problème de contrôle stochastique de type McKean-Vlasov. Enfin, ils montrent que la formulation à champ-moyen permet d'approximer la situation de N-Agents lorsque N est très grand avec les techniques de propagation de chaos.

### 1.1.3 Extension aux cas dégénérés

La méthode décrite précédemment se base sur une hypothèse importante : le système stochastique contrôlé est non-dégénéré. Plus précisément, dans le cœur de l'approche de programmation dynamique se trouve la représentation des contrats par les équations différentielles stochastiques rétrogrades comme montrée dans la section précédente, qui nécessite que le système soit non-dégénéré.

La principale motivation de ce travail est de résoudre le problème de Principal-Agent dans un système à observation partielle. Nous nous retrouvons avec un système dégénéré lorsque nous remplaçons le processus non-observable par le processus filtré qui caractérise la loi conditionnelle sachant le processus observable en utilisant les techniques de filtrage. Pour plus de détail sur le filtrage, nous nous référons à Bensoussan [10]. Un cas très simple est le cas avec un paramètre inconnu. Ce problème a été étudié par Fernandes et Phelan [56], Williams [156], puis par Prat et Jovanovic [136], qui ont inspiré et motivé cette étude. Pour illustrer l'idée, considérons le processus observable suivant :

$$dB_t = (\mu + \beta_t)dt + \sigma dW_t, \quad (1.1.4)$$

où  $\beta$  est le contrôle de l'Agent et  $\mu$  un paramètre inconnu qui représente la productivité invariante. La distribution *a priori* de  $\mu$  est Gaussienne de moyenne  $m_0$  et de variance  $V_0$ , respectivement. En utilisant la formule de Bayes, nous calculons la distribution *a posteriori* de  $\mu$ , qui dépend  $B_t$  et aussi l'effort cumulé de l'Agent  $A_t := \int_0^t \beta_s ds$ . En remplaçons  $\mu$  par son espérance conditionnelle, nous obtenons une autre expression de la dynamique de  $B$  :

$$dB_t = \left( \frac{V_t}{V_0} m_0 + \frac{V_t}{\sigma^2} (B_t - A_t) + \beta_t \right) dt + dI_t,$$

où  $V_t = \frac{\sigma^2 V_0}{\sigma^2 + V_0 t}$  est la variance *a posteriori* et  $(I_t)_{t \geq 0}$  est un mouvement Brownien  $\mathbb{F}^B$ -adapté et est appelé le processus d'Innovation. La dynamique décrite par (1.1.4) est devenue un système dégénéré avec le nouveau processus d'état  $(A_t)_{t \geq 0}$ . Comme le Principal ne peut observer l'effort de l'Agent, il ne peut observer le processus  $A$  et par conséquent il ne peut le mettre dans le contrat. Dans le cas général, pour un système à observation partielle, il nécessite de connaître tout l'historique de l'effort de l'Agent pour pouvoir calculer le filtrage. La difficulté du problème de Principal-Agent dans ce cas-là est de trouver un moyen de décrire l'ensemble des contrats utilisant seulement la partie observable du système, *i.e.* le

processus  $B$  dans l'exemple ci-dessus.

Dans cette première partie de la thèse, nous considérons le problème de Principal-Agent dans un système dégénéré contrôlé général utilisant la formulation faible, permettant l'Agent de contrôler le drift et la volatilité de la partie dégénérée du système correspondante au filtrage. On suppose que seulement le drift du processus observable est contrôlé par l'Agent. Nous donnons la condition du premier ordre pour le contrôle optimal de l'Agent pour chaque contrat donné. Cette condition du premier ordre est décrite par une EDS progressive-rétrograde dépendante de la trajectoire que nous allons étudier dans un autre chapitre. Plus précisément les coefficients de cette EDS progressive-rétrograde dépendent de la trajectoire du processus progressif de l'EDS progressive-rétrograde. Dans la littérature des EDS progressives-rétrogrades, certains auteurs utilisent la même terminologie pour les EDS progressives-rétrogrades dont les coefficients dépendent de la trajectoire du mouvement Brownien. Pour la distinction nous appelons les EDS progressives-rétrogrades avec des coefficients aléatoires les EDS progressives-rétrogrades non-Markoviennes. Revenons au problème de Principal-Agent, l'EDS progressive-rétrograde de la condition du premier ordre peut être vue comme une représentation du contrat et du contrôle optimal de l'Agent étant donné ce contrat. Toutefois, comme l'EDS progressive-rétrograde ne décrit qu'une condition nécessaire du problème de l'Agent, l'ensemble des contrats représenté peut contenir plus de contrats que l'ensemble des contrats admissibles. Notre stratégie est donc de résoudre d'abord le problème du Principal dans cet ensemble élargi des contrats donné par la condition du premier ordre du problème de l'Agent et puis de vérifier l'admissibilité du contrat optimal à l'aide d'une condition suffisante du problème de l'Agent.

La contribution de ce travail est de donner une méthode pour résoudre le problème de Principal-Agent lorsqu'on ne peut utiliser une partie des variables d'état pour construire le contrat optimal, ce qui est crucial quand on rencontre les systèmes à observation partielle. Nous résolvons deux exemples explicites dans la section [II.4](#).

### **I.1.4 Intermède : EDS progressives-rétrogrades dépendantes de la trajectoire**

Dans la section précédente, nous avons présenté comment résoudre le problème de Principal-Agent à l'aide du calcul de variation. La représentation du contrat que l'on utilise dans les systèmes dégénérés est une EDS progressive-rétrograde dépendante de la trajectoire. Plus généralement, les EDS progressives-rétrogrades apparaissent naturellement lors de la résolution d'un problème de contrôle par la méthode de calcul de variation. En supposant l'existence d'un contrôle optimal, le processus progressif décrit le processus d'état sous le contrôle optimal tandis que le processus rétrograde représente le processus adjoint. La théorie des EDS progressives-rétrogrades a été étudiée extensivement pendant ces dernières dizaines d'années, motivée par les applications en finance quantitative et leur lien intrinsèque avec les équations aux dérivées partielles quasi-linéaires. Une EDS progressive-rétrograde est un système sous la

forme

$$\begin{cases} dX_t = b_t(X_t, Y_t, Z_t)dt + \sigma_t(X_t, Y_t, Z_t)dW_t & X_0 = x, \\ dY_t = -f_t(X_t, Y_t, Z_t)dt + Z_t dW_t & Y_T = g(X_T), \end{cases}$$

où  $X$  est appelé le processus progressif et  $(Y, Z)$  le processus rétrograde. Le triplet  $(X, Y, Z)$  peut être multidimensionnel, les coefficients des équations  $b, \sigma, f, g$  peuvent être déterministes ou bien aléatoires. On dit qu'un système est découplé si ni  $b$  ni  $\sigma$  dépend de  $(Y, Z)$ . Ce genre d'équations a été étudié par Pardoux et Peng dans [130]. Sous les hypothèses standards, ces équations peuvent être transformées en équations différentielles stochastiques rétrogrades, dont l'existence et l'unicité peuvent être garanties.

Dans le cas couplé, la situation est beaucoup moins claire. Regardons l'exemple suivant. Soit  $k$  une constante et considérons l'EDS progressive-rétrograde suivante

$$\begin{cases} dX_t = (k + Z_t)dW_t & X_0 = x \\ dY_t = Z_t dW_t & Y_T = X_T. \end{cases}$$

Remarquons que

$$Y_t = X_T - \int_t^T Z_s ds = X_t + k(W_T - W_t). \quad (I.1.5)$$

En particulier, si  $k \neq 0$ , on a  $Y_0 - x = kW_T$ , ce qui est impossible. Dans le cas où  $k = 0$ , l'EDS progressive-rétrograde a une infinité de solutions. En général, on ne peut garantir ni l'existence de la solution ni l'unicité de la solution sur un intervalle quelconque. Si les coefficients sont réguliers, déterministes et l'équation forward est non-dégénérée, la méthode du schéma à quatre étapes est proposée par Ma, Protter, Yong [109], qui peut résoudre les équations sur un intervalle arbitrairement grand. Cette méthode relie les EDS progressives-rétrogrades à coefficients déterministes et réguliers aux équations aux dérivées partielles quasi-linéaire, *i.e.*

$$u_t + \frac{1}{2}\sigma^2(t, x, u)u_{xx} + b(t, x, u, \sigma(t, x, u)u_x)u_x + f(t, x, u, \sigma(t, x, u)u_x) = 0, \quad (I.1.6)$$

$$u(T, x) = g(x).$$

Sous les hypothèses standards, la solution de l'EDS progressive-rétrograde avec les coefficients  $(b, \sigma, f, g)$  est donnée par

$$Y_t = u(t, X_t); \quad Z_t = \sigma(t, X_t, u(t, X_t))u_x(t, X_t).$$

Dans le cas non-Markovien, *i.e.* les coefficients  $(b, \sigma, f, g)$  sont aléatoires, la méthode de continuation introduite par Hu et Peng, [84], Peng et Wu [134], puis développée par Yong [159, 160], est beaucoup utilisée, notamment dans [157], [161]. La méthode dépend crucialement d'une hypothèse de monotonie sur les coefficients, qui est assez restrictive pour les applications. Plus tard en utilisant la notion de champ de découplage, une technique générale a été développée après une série de travaux, notamment Cvitanic et Zhang [40], Zhang [163], Ma, Wu, Zhang, Zhang [110] et elle est utilisée pour généraliser

la méthode de contraction sur les petits intervalles proposée par Antonelli [4], Pardoux et Tang [131] pour construire une solution sur les intervalles arbitrairement grands en collant des solutions locales. Le champ de découplage  $u$  peut être vu comme une extension de la version déterministe de la solution de l'équation différentielle partielle (1.1.6). C'est une fonction aléatoire telle que la solution de l'EDS progressive-rétrograde satisfasse

$$Y_t := u_t(X_t).$$

L'existence et l'unicité du champ de découplage ont d'abord été étudiées dans le cas unidimensionnel dans [110] puis étendu par Zhang [165] dans le cas où le processus rétrograde est multidimensionnel. Fromm et Imkeller dans [60] utilisent la dérivée au sens faible pour généraliser les résultats au cas multidimensionnel général et définissent la notion de l'intervalle maximal pour une EDS progressive-rétrograde.

Nous étudions l'existence et l'unicité des EDS progressives-rétrogrades dépendantes de la trajectoire, plus précisément la trajectoire du processus progressif. Nous considérons les EDS progressives-rétrogrades sous la forme suivante

$$\begin{cases} dX_t = b_t(X, Y_t, Z_t)dt + \sigma_t(X, Y_t, Z_t)dW_t & X_0 = x, \\ dY_t = -f_t(X, Y_t, Z_t)dt + Z_t dW_t & Y_T = g(X). \end{cases}$$

Nous démontrons d'abord l'existence et l'unicité locale pour ce genre d'EDS progressive-rétrograde. Puis en introduisant une nouvelle norme sur l'espace des trajectoires, nous définissons l'équation différentielle stochastique rétrograde caractéristique associée. En suivant les étapes dans [110], nous distinguons des différents cas et donnons des conditions suffisantes pour l'existence et l'unicité de la solution. Enfin, nous donnons aussi un résultat de stabilité pour les EDS progressives-rétrogrades dépendantes de la trajectoire.

### 1.1.5 Le problème de Principal-Agent avec plusieurs Principals

Contrairement au problème de Principal-Agent avec plusieurs Agents, il y a peu de résultats sur le problème de Principal-Agent avec plusieurs Principals. Mastrolia et Ren dans [115] considèrent le cas où plusieurs Principals emploient un Agent commun en même temps, c'est-à-dire l'Agent peut travailler simultanément pour plusieurs Principals sur des projets différents. Mais à notre connaissance, il n'y a pas encore de modèle pour les situations où l'Agent travaille exclusivement pour un seul Principal. Dans un tel modèle, l'Agent fait face à un problème de switching optimal. Considérons un cas simple avec  $N$  Principals et notons  $(I_t)_{t \geq 0}$  le processus à valeurs dans  $\llbracket 1, N \rrbracket$  qui désigne le Principal pour qui l'Agent travaille à temps  $t$ . Le processus d'état contrôlé par l'Agent est

$$X_t = X_0 + \int_0^t e_{I_s} ds + W_t, \quad \text{où } e_i = (0, \dots, 0, 1, 0, \dots, 0)^\top \in \mathbb{R}^N,$$

qui signifie que le processus de production du Principal employant l'Agent a un drift positif tandis que les autres processus de production ne sont que des mouvements Browniens. Pour  $i = 1, \dots, N$ , le Principal  $i$  propose un contrat  $(\xi^i, w^i)$  qui consiste un paiement finalement  $\xi^i$  à temps  $T$  si l'Agent termine son contrat avec lui et un salaire continu  $w^i$ . L'Agent fait donc face au problème d'optimisation suivant :

$$V_0^A(\xi, w) = \sup_I \mathbb{E} \left[ \sum_{i=1}^N \left( \xi^i \mathbf{1}_{\{I_T=i\}} + \int_0^T w_t^i \mathbf{1}_{\{I_t=i\}} du \right) - \text{“coût de switching”} \right],$$

Dans la littérature sur le switching optimal, voir e.g. [135, 67, 85, 31], l'équivalence de la caractérisation du problème de l'Agent par une équation différentielle stochastique rétrograde (1.1.3) est une équation différentielle stochastique rétrograde réfléchie sous la forme

$$dY_t = -f_t(Y_t, Z_t)dt + Z_t dW_t - dK_t, \quad Y_T = \xi,$$

où  $K$  est un processus croissant satisfaisant une condition de Skorokhod rétrograde. La présence du processus  $K$  donne une difficulté supplémentaire pour trouver une bonne représentation pour le couple  $(\xi, \alpha^*[\xi])$  comme dans Cvitanić, Possamaï et Touzi [36].

Nous proposons une méthode alternative pour surmonter cette difficulté. Au lieu de laisser l'Agent contrôler le temps d'arrêt directement pour changer d'employeur, nous étudions la version randomisée du problème de switching optimal où l'Agent influence le temps d'arrêt en contrôlant l'intensité du processus de Poisson associé. Plus précisément, on suppose que le processus  $(I_t)_{t \geq 0}$  est un processus ponctuel de Poisson et que l'Agent contrôle son intensité. Notons  $V_t^{A,i} := V_t^A|_{I_t=i}$  pour  $i = 1, \dots, N$ . Grâce au théorème de Girsanov pour les processus ponctuels multivariés et le principe de la programmation dynamique, le problème de l'Agent en formulation faible devient

$$\begin{aligned} V_t^{A,i} &= \text{esssup}_{\alpha} \mathbb{E}_t^{\mathbb{P}^{\alpha}} \left[ \xi^i \mathbf{1}_{\{\tau_t > T\}} + \sum_{j \neq i} \left( V_{\tau_t}^{A,j} \mathbf{1}_{\{\tau_t \leq T, I_{\tau_t}=j\}} + \int_t^{\tau_t \wedge T} \frac{w_u^i - c((N-1)\alpha_u^j)}{N-1} du \right) \right] \\ &= \text{esssup}_{\alpha} \mathbb{E}_t^{\mathbb{P}^{\alpha}} \left[ \xi^i \beta^{i,\alpha}(t, T) + \int_t^T \beta^{i,\alpha}(t, u) \sum_{j \neq i} \left( \alpha_u^j V_u^{A,j} + \frac{w_u^i - c((N-1)\alpha_u^j)}{N-1} \right) du \right], \end{aligned}$$

où  $\alpha(dt, \cdot) = \sum_{k \in \llbracket 1, N \rrbracket} \alpha_t^k \delta_k dt$  est l'intensité du processus de Poisson contrôlée par l'Agent, la fonction  $c$  désigne la fonction de coût de switching de l'Agent,  $\tau_t$  et  $\beta^{i,\alpha}(t, s)$  sont définis par

$$\tau_t := \inf\{s \geq t : I_s \neq I_t\} \quad \text{et} \quad \beta^{i,\alpha}(t, s) := \exp \left( - \int_t^s \sum_{j \neq i} \alpha_u^j du \right) \quad \text{respectivement.}$$

Des formulations randomisées similaires de switching ont déjà été proposées dans la littérature, e.g. Bouchard [20], Elie et Kharroubi [53] et plus récemment Fuhrman et Morlais [61].

Cette formulation nous permet de caractériser le couple  $(\xi, \alpha^*[\xi])$  par un système d'équations diffé-

rentielles stochastiques rétrogrades sans réflexion. Malheureusement cette caractérisation du problème de l'Agent ne permet pas de résoudre le problème des Principals directement. Nous montrons que le problème d'optimisation des Principals n'est pas consistant en temps, et ainsi ne peut être résolu par le principe de la programmation dynamique et la notion d'optimalité du problème des Principals est mal définie. Un autre problème de cette approche est que les Principals ont besoin de connaître tous les composants du processus d'état  $X$  pour construire les contrats, autrement dit, chaque Principal a besoin de connaître le processus de production des autres Principals.

Pour surmonter ces difficultés, nous proposons d'étudier le problème avec la formulation de jeu à champ-moyen. Tout d'abord nous remarquons que l'inconsistance temporelle dans le problème de  $N$  Principals disparaît dès que le nombre de Principals devient infini. En utilisant la représentation des contrats par une BSDE à champ-moyen, le problème des Principals devient maintenant un problème de jeu à champ-moyen. Avec les techniques du papier de Lacker [99], nous démontrons l'existence d'un équilibre de Nash pour le problème des Principals en passant par la formulation relaxée de jeu à champ-moyen. De plus, pour chaque Principal, le contrat optimal au sens de Nash dépend maintenant seulement de sa propre performance. Enfin, en introduisant une technique que nous appelons la propagation de chaos rétrograde, nous justifions notre formulation à champ-moyen en montrons que dans le cas de  $N$ -Principals pour  $n$  assez grand, la valeur du problème des Principals correspondant au contrat  $\xi^*$  reste proche de  $V_\infty(\xi^*)$ , où  $\xi^*$  est un contrat optimal au sens d'équilibre de Nash dans le cas de champ-moyen et  $V_\infty(\xi^*)$  la valeur de l'équilibre correspondant.

## 1.2 La Dynamique de Langevin à Champ-moyen, Applications aux Réseaux de Neurones

### 1.2.1 Motivation : apprentissage profond

L'apprentissage profond est un des domaines le plus étudié dans le monde de l'intelligence artificielle et il a eu des impacts importants dans beaucoup de domaines d'applications comme la santé, les robotiques, l'ingénierie biomédicale, les systèmes aérospatiales etc. Les réseaux de neurones profonds, qui peuvent avoir éventuellement des milliers de paramètres et pourtant ne nécessitent pas de connaissance d'expertise ou de calibration à la main, arrivent parfois à une performance qui dépasse celle des humains dans des problèmes complexes comprenant la synthèse visuelle [98], le traitement de langage naturel [145], et le planning séquentiel [142]. Toutefois, rien n'est gratuit. Les algorithmes utilisés dans l'entraînement des réseaux de neurones demandent beaucoup de données en pratique. Un manque de données peut rendre les algorithmes inefficaces, conduisant parfois à des résultats médiocres [62]. Les réseaux de neurones profonds sont aussi réputés être peu robustes face à des perturbations de données, voir e.g. [64, 5].

Un fondement théorique rigoureux de la théorie de l'apprentissage profond et des algorithmes d'apprentissage est crucial pour la conception de l'architecture et la calibration des réseaux de neurones. Pour

mieux illustrer les problèmes, nous présentons d'abord l'architecture d'un réseau de neurones profond général. Pour  $l \in \mathbb{N}$  définissons  $\varphi^l : \mathbb{R}^l \rightarrow \mathbb{R}^l$  par  $\varphi^l(z) = (\varphi(z_1), \dots, \varphi(z_l))^T$  pour  $z = (z_1, \dots, z_l)^T$ , où  $\varphi$  est une fonction localement Lipschitzienne  $\mathbb{R} \rightarrow \mathbb{R}$ . Nous fixons  $L \in \mathbb{N}$  qui représente le nombre de couches du réseau de neurones,  $l_k \in \mathbb{N}$  la dimension de l'entrée de la  $k$ -ième couche pour  $k = 0, 1, \dots, L-1$  et  $l_L$  représente la dimension de la sortie du réseau. Un réseau de neurones entièrement connecté est donné par  $\Psi = ((\alpha^1, \beta^1), \dots, (\alpha^L, \beta^L)) \in \Pi$ , où, pour  $k = 1, \dots, L$ , nous avons les matrices  $\alpha^k$  de la taille  $l^k \times l^{k-1}$  et les vecteurs  $\beta^k$   $l^k$ -dimensionnel. Autrement dit,  $\Pi = (\mathbb{R}^{l^1 \times l^0} \times \mathbb{R}^{l^1}) \times (\mathbb{R}^{l^2 \times l^1} \times \mathbb{R}^{l^2}) \times \dots \times (\mathbb{R}^{l^L \times l^{L-1}} \times \mathbb{R}^{l^L})$ .

Nous définissons le réseau de neurones  $\mathcal{R}\Psi : \mathbb{R}^{l^0} \rightarrow \mathbb{R}^{l^L}$  de manière récursive. Pour  $z_0 \in \mathbb{R}^{l^0}$  nous avons

$$(\mathcal{R}\Psi)(z^0) = \alpha^L z^{L-1} + \beta^L, \quad z^k = \varphi^{l^k}(\alpha^k z^{k-1} + \beta^k), \quad k = 1, \dots, L-1.$$

Si pour chaque  $k = 1, \dots, L-1$  nous écrivons  $\alpha_i^k, \beta_i^k$  pour la  $i$ -ième ligne de la matrice  $\alpha^k$  et le vecteur  $\beta^k$ , on peut réécrire la formulation ci-dessus par

$$(\mathcal{R}\Psi)(z^0)_i = \alpha_i^L \cdot z^{L-1} + \beta_i^L, \quad (z^k)_i = \varphi(\alpha_i^k \cdot z^{k-1} + \beta_i^k), \quad k = 1, \dots, L-1. \quad (1.2.7)$$

Remarquons que le nombre de paramètres dans le réseau défini ci-dessus est  $\sum_{i=1}^L (l_{k-1} l_k + l_k)$ . Dans l'apprentissage supervisé, l'objectif est de trouver les paramètres optimaux pour  $\Psi$  tel que le réseau de neurones donne une bonne approximation du problème étudié. Mathématiquement cela veut dire la résolution du problème d'optimisation suivant étant donné un fonctionnel de potentiel  $\Phi$  et des données d'entraînement  $(y^j, z^j)_{j=1}^N, (y_j, z_j) \in \mathbb{R}^d$  :

$$\operatorname{argmin}_{\Psi \in \Pi} \frac{1}{N} \sum_{j=1}^N \Phi(y^j - (\mathcal{R}\Psi)(z^j)).$$

Comme en pratique on travaille très souvent avec un grand nombre de données, il est raisonnable de supposer que les données sont distribuées selon une loi de probabilité grâce à la loi des grands nombres. Le problème s'écrit ainsi

$$\operatorname{argmin}_{\Psi \in \Pi} \int_{\mathbb{R}^d} \Phi(y - (\mathcal{R}\Psi)(z)) \nu(dy, dz). \quad (1.2.8)$$

Ceci est un problème d'optimisation non-convexe de haute dimension, qui est très difficile à résoudre en général. Dans le cas où  $L = 2$ , *i.e.* quand le réseau de neurones a une seule couche cachée, l'existence du minimum est garantie par le théorème d'approximation universelle, dont une preuve élémentaire se trouve dans [79, Théorème 2].

**Theorem 1.2.1** (Théorème d'Approximation Universelle). *Si la fonction d'activation  $\varphi$  est bornée, conti-*

nue et non-constante, alors pour tout compact  $K \subset \mathbb{R}^d$  l'ensemble

$$\{(\mathcal{R}\Psi) : \mathbb{R}^d \rightarrow \mathbb{R} : (\mathcal{R}\Psi) \text{ donnée par (1.2.7) avec } L = 2 \quad n \in \mathbb{N}, \alpha_j^2, \beta_j^1 \in \mathbb{R}, \alpha_j^1 \in \mathbb{R}^d, j = 1, \dots, n\}$$

est dense dans  $C(K)$ .

En pratique on utilise très souvent la descente de gradient stochastique pour entraîner les réseaux de neurones, voir *e.g.* Goodfellow, Bengio, Courville [63, Chapitre 6]. Comme le réseau est défini explicitement de manière récursive, on peut calculer le gradient de la fonction de potentiel qu'on veut minimiser par rapport à chaque paramètre du réseau. Ceci peut être défini de manière récursive et rétrograde, c'est ce qu'on appelle la propagation rétrograde (backward propagation en anglais). L'avantage de cet algorithme est sa faible complexité en temps (en général  $\mathcal{O}\left(\sum_{i=1}^L (l_{k-1}l_k + l_k)\right)$ ). Malgré le fait que le gradient stochastique est très utile en pratique, il manque encore de la théorie mathématique qui explique sa réussite dans des problèmes non-convexes. En effet, comme le problème d'optimisation (1.2.8) n'est pas convexe, l'algorithme trouve *a priori* un minimum local et il n'y a aucune garantie que ce minimum local soit proche du vrai minimum que l'on cherche. Un autre fait intéressant est que l'on observe souvent que la performance du réseau de neurones est insensible au nombre de paramètres dans les couches cachées (étant donné qu'il est grand) et en pratique on travaille souvent avec des réseaux de neurones qui ont plus de paramètres que le nombre des données, voir *e.g.* [70, 8]. On dit dans ce cas que le réseau est surparamétré. Un premier argument qui justifie l'utilisation de ce genre de réseau est le théorème d'approximation universelle qui dit qu'une représentation parfaite d'une fonction continue sur un compact nécessite en général une infinité de paramètres. Dans cette deuxième partie de la thèse, nous explorons les réseaux de neurones surparamétrés avec un point de vue de champ-moyen et nous établissons des liens forts entre le minimiseur du problème d'optimisation (1.2.8) et la mesure invariante du processus/système de Langevin à champ-moyen, qui est la version continue randomisée de l'algorithme de la descente de gradient stochastique utilisé en pratique. Notre première contribution est de montrer l'existence et l'unicité de la mesure invariante du processus/système de Langevin à champ-moyen, qui correspond à la solution stationnaire d'une équation/un système de Fokker-Planck non-linéaire. Nous montrons également que la convergence a lieu à une vitesse exponentielle que l'on peut déterminer de manière explicite. Ensuite, avec un argument variationnel, nous montrons que cette mesure invariante est aussi le minimiseur du problème d'optimisation (1.2.8) plongé dans l'espace des mesures de probabilité. Nous traitons d'abord le cas de réseau de neurones à une couche cachée, qui devient un problème convexe une fois plongé dans l'espace des mesures de probabilité. Nous donnons la formule explicite du flot du gradient en utilisant le calcul différentiel dans l'espace des probabilités. Dans le dernier chapitre nous traitons le cas général de réseau de neurones, qui reste en général un problème non-convexe dans l'espace des mesures de probabilité.



## 1.2.2 La dynamique de Langevin à champ-moyen

Rappelons brièvement la définition classique de l'équation de Langevin en dimension finie. Soit  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  un potentiel Lipschitzien satisfaisant certaines conditions de croissance à l'infini. L'équation de Langevin s'écrit

$$dX_t = -\nabla f(X_t)dt + \sigma dW_t, \quad (1.2.9)$$

où  $\sigma$  est une constante et  $W$  est un mouvement Brownien  $d$ -dimensionnel. On peut voir cette dynamique de deux manières :

1. la solution de (1.2.9) est une diffusion de Markov homogène en temps, il existe une unique mesure invariante  $m^{\sigma,*}$  dont la densité doit être sous la forme

$$m^{\sigma,*}(x) = \frac{1}{Z} \exp\left(-\frac{2}{\sigma^2}f(x)\right), \quad \text{pour tout } x \in \mathbb{R}^d, \quad \text{où } Z := \int_{\mathbb{R}^d} \exp\left(-\frac{2}{\sigma^2}f(x)\right) dx;$$

2. la dynamique (1.2.9) peut être vue comme une version continue randomisée de l'algorithme de la descente du gradient.

Les deux perspectives peuvent être unifiées si on considère la forme variationnelle de la mesure invariante, c'est-à-dire  $m^{\sigma,*}$  comme l'unique minimiseur de la fonction d'énergie libre

$$V^\sigma(m) := \int_{\mathbb{R}^d} f(x)m(dx) + \frac{\sigma^2}{2} \text{Ent}(m)$$

parmi toutes les mesures de probabilité  $m$ , où  $\text{Ent}$  est l'entropie par rapport à la mesure de Lebesgue. Ce point de vue variationnel a été proposé dans Jordan et Kinderlehrer [92] et Jordan, Kinderlehrer et Otto [93]. En outre, on peut aussi observer que s'il existe une mesure uniforme  $U$  sur l'ensemble  $\text{argmin } f$  de minimiseurs de  $f$ , alors la mesure invariante  $m^{\sigma,*}$  concentre sur  $U$  lorsque  $\sigma \rightarrow 0$ . En particulier, si  $f$  admet un unique minimiseur,  $m^{\sigma,*}$  concentre sur  $\delta_{\text{argmin } f}$  sans que la convexité de la fonction  $f$  soit demandée. Ceci construit un lien important entre la théorie d'échantillonnage statistique et l'optimisation et on voit que l'équation de Langevin joue un rôle essentiel dans les problèmes d'optimisation non-convexes. Ce fait a aussi été remarqué dans la communauté d'optimisation numérique et d'apprentissage statistique, e.g. [87, 76, 75].

Nous généralisons cette connexion entre le minimiseur global et la mesure invariante aux cas où le potentiel est une fonction dans l'espace des mesures de probabilité. Rappelons que le théorème d'approximation universelle nous dit que les réseaux de neurones à une seule couche cachée peuvent approximer toute fonction continue définie sur un compact. Toutefois il ne nous dit pas comment trouver les paramètres optimaux. En effet, les paramètres optimaux sont solutions du problème d'optimisation

non-convexe suivant

$$\min_{\beta_{n,i} \in \mathbb{R}, \alpha_{n,i} \in \mathbb{R}^{d-1}} \left\{ \int_{\mathbb{R} \times \mathbb{R}^{d-1}} \Phi \left( y - \frac{1}{n} \sum_{i=1}^n \beta_{n,i} \varphi(\alpha_{n,i} \cdot z) \right) \nu(dy, dz) \right\}, \quad (1.2.10)$$

où  $\Phi$  est une fonction convexe,  $\varphi$  est une fonction d'activation continue bornée non-constante et  $\nu$  est une mesure à support compact représentant la distribution des données. On peut définir la loi empirique des paramètres comme  $m^n := \frac{1}{n} \sum_{i=1}^n \delta_{\{\beta_{n,i}, \alpha_{n,i}\}}$ . Alors le problème (1.2.10) s'écrit

$$\frac{1}{n} \sum_{i=1}^n \beta_{n,i} \varphi(\alpha_{n,i} \cdot z) = \int_{\mathbb{R}^d} \beta \varphi(\alpha \cdot z) m^n(d\beta, d\alpha).$$

Pour simplifier les notations, notons  $\hat{\varphi}(x, z) := \beta \varphi(\alpha \cdot z)$  pour  $x = (\beta, \alpha) \in \mathbb{R}^d$  et notons  $\mathbb{E}^m$  l'espérance de la variable aléatoire  $X$  sous la mesure de probabilité  $m$ . Maintenant, au lieu d'étudier le problème (1.2.10) nous proposons d'étudier le problème de minimisation sur l'espace des mesures de probabilité suivant

$$\min_m F(m), \quad \text{avec} \quad F(m) := \int_{\mathbb{R}^d} \Phi \left( y - \mathbb{E}^m [\hat{\varphi}(X, z)] \right) \nu(dy, dz), \quad (1.2.11)$$

Cette reformulation est cruciale, elle rend le potentiel  $F$  défini maintenant sur l'espace des mesures convexe, *i.e.* pour toute mesure de probabilité  $m$  et  $m'$ , on a

$$F((1 - \alpha)m + \alpha m') \leq (1 - \alpha)F(m) + \alpha F(m') \quad \text{pour tout } \alpha \in [0, 1].$$

Cet exemple nous montre comment transformer un problème de minimisation non-convexe en dimension finie en un problème de minimisation convexe en le plongeant dans l'espace des mesures de probabilité. Pour retrouver la dynamique de Langevin associée, nous ajoutons ensuite l'entropie relative comme régularisation. Contrairement au cas classique, on utilise maintenant l'entropie relative par rapport à une mesure de Gibbs dont la densité est proportionnelle à  $e^{-U(x)}$ . Un choix typique de mesure de Gibbs est la loi normale centrée réduite.

Nous caractérisons le minimiseur de la fonction d'énergie libre

$$V^\sigma := F + \frac{\sigma^2}{2} \text{Ent}$$

à l'aide de la dérivation dans l'espace des mesures de probabilité au sens de Lions, notée  $\frac{\delta}{\delta m}$ , introduite initialement dans le calcul de variation et maintenant largement utilisée dans la théorie de jeu à champ-moyen, voir *e.g.* [24]. Nous donnons d'abord la condition du premier ordre

$$m^* = \arg \min_m V^\sigma(m) \quad \text{si et seulement si} \quad \frac{\delta F}{\delta m}(m^*, \cdot) + \frac{\sigma^2}{2} \log(m^*) + \frac{\sigma^2}{2} U = \text{constant}.$$

Prenant en compte le fait que  $m^*$  soit une mesure de probabilité, la condition du premier ordre nous donne

$$m^*(x) = \frac{1}{Z} \exp \left( -\frac{2}{\sigma^2} \left( \frac{\delta F}{\delta m}(m^*, x) + U(x) \right) \right),$$

où  $Z$  est la constante de normalisation. On remarque aussi que la forme variationnelle de la mesure invariante de l'équation de Langevin classique est aussi une forme particulière de la condition du premier ordre donnée ci-dessus. Formellement, une mesure satisfaisant la condition du premier ordre est une solution stationnaire de l'équation de Fokker-Planck non-linéaire

$$\partial_t m = \nabla \cdot \left( \left( D_m F(m, \cdot) + \frac{\sigma^2}{2} \nabla U \right) m + \frac{\sigma^2}{2} \nabla m \right), \quad (1.2.12)$$

avec  $D_m F$  la dérivée intrinsèque sur l'espace des probabilités, définie par  $D_m F(m, x) := \nabla \frac{\delta F}{\delta m}(m, x)$ . Nous retrouvons facilement la dynamique de Langevin à champ-moyen correspondant à l'équation de Fokker-Planck non-linéaire :

$$dX_t = - \left( D_m F(m_t, X_t) + \frac{\sigma^2}{2} \nabla U(X_t) \right) dt + \sigma dW_t, \quad \text{où } m_t := \mathcal{L}(X_t). \quad (1.2.13)$$

Enfin pour revenir aux réseaux de neurones, il suffit de remarquer que la dynamique de Langevin à champ-moyen correspond à une version continue de l'algorithme de la descente du gradient randomisé et régularisé. Plus précisément, dans le cas où les données  $(y_m, z_m)_{m \in \mathbb{N}}$  sont i.i.d. de la loi  $\nu$  et la fonction de coût  $\Phi$  est quadratique, l'algorithme de la descente du gradient est donné par

$$x_{k+1}^i = x_k^i + 2\tau \left( \left( y_k - \frac{1}{N} \sum_{j=1}^N \hat{\varphi}(x_k^j, z_k) \right) \nabla \hat{\varphi}(x_k^i, z_k) - \frac{\sigma^2}{2} \nabla U(x_k^i) \right) + \sigma \sqrt{\tau} \xi_k^i,$$

où  $(\xi_k^i)$  sont des Gaussiens centrés réduits indépendants. Ceci peut être vu comme une discrétisation de (1.2.13).

Les solutions stationnaires des équations de diffusion non-locales (1.2.12) ont été étudiées depuis longtemps, notamment en physique statistique et théorie cinétique, e.g. Mischler et Mouhot [120]. En particulier, l'approche variationnelle a été développée dans Carrillo, McCann et Villani [28], Otto, [127], Tugaut [151] dans lesquels les auteurs étudient la dissipation de l'entropie pour les équations du milieu granulaire avec un potentiel symétrique du type convolution. Nous référons aussi à Cattiaux, Guillin et Malrieu [29], Veretennikov [153] pour des résultats similaires en utilisant l'approximation des particules, à Eberle, Guillin et Zimmer [49] pour les techniques de couplage et à Butkovsky [23], Bogachev, Röckner et Shaposhnikov [13] pour les techniques de Khasminskii. Tous ces travaux reposent sur des conditions restrictives sur le potentiel et/ou l'hypothèse de faible interaction. En utilisant le principe d'invariance de LaSalle et l'inégalité de HWI d'Otto et Villani dans [128], nous démontrons la convergence de la loi marginale de la dynamique de Langevin à champ-moyen vers la mesure invariante pour tout potentiel

régulier borné sur l'espace des probabilités. Nous donnons également une preuve probabiliste se basant sur le calcul d'Itô pour le flot de gradients de (I.2.12), qui peut être vu comme une extension de [94, Théorème 3.1] de Karatzas, Schachermayer et Tschiderer aux dynamiques de McKean-Vlasov. Pour les calculs dans l'espace des probabilités, nous référons au livre de Carmona et Delarue [26, Chapitre 5] pour une présentation détaillée.

Pendant la collaboration de ce travail, il y a d'autres groupes qui ont développé des explications similaires en utilisant le langage de champ-moyen sur les problèmes d'apprentissage non-convexe, e.g. Mei, Montanari, Nguyen [117], Mei, Misiakiewicz, Montanari [116], Chizat, Bach [34], Rotskoff et Vanden-Eijnden [138], Javanmard, Mondelli et Montanari [90], Sirignano et Spiliopoulos [143]. En particulier, Mei, Misiakiewicz et Montanari [117], Chizat et Bach [34], Rotskoff et Vanden-Eijnden [138] démontrent la convergence des algorithmes du type descente du gradient vers le minimum en utilisant le théorème de flot de gradient dans l'espace de Wasserstein des probabilités, e.g. Ambrosio, Gigli et Savar [3]. Les résultats dans [117] ressemblent le plus aux nôtres mais les preuves sont différentes. Tandis que [117] s'inspirent du papier [93] de Jordan, Kinderlehrer et Otto, nous donnons un point de vue purement probabiliste. Nous généralisons et donnons des preuves complètes sur des résultats clés comme le flot de gradients [117, Lemme 6.1] (notre Théorème V.2.9) et la convergence globale [117, Lemme 6.12] (notre Théorème V.2.11). En particulier, nous démontrons la convergence vers la mesure invariante sous la distance 2-Wasserstein et que la vitesse de convergence est exponentielle pour des problèmes suffisamment régularisés. Notre méthode peut aussi traiter des fonctions de coût générales, comme conjecturé dans Appendix B de [117].

### I.2.3 Les systèmes de Langevin à champ-moyen

La section précédente explore les liens entre la mesure invariante de la dynamique de Langevin à champ-moyen, le minimiseur de la fonction d'énergie libre et l'algorithme de la descente du gradient pour les réseaux de neurones à une couche cachée. En rendant le problème (I.2.10) convexe, on peut approximer le minimiseur grâce à la dynamique de Langevin à champ-moyen (I.2.13). Malheureusement pour les réseaux de neurones profonds (I.2.8), on ne peut plus les rendre convexes simplement en les plongeant dans l'espace des mesures de probabilité. Proposé récemment dans Chen, Rubanova, Bettencourt, Duvenaud [33], Cuchiero, Larsson, Teichmann [35], E, Han, Li [47], Liu, Markowich [107] et également dans le cours de P.L. Lions au Collège de France (les idées similaires datées de LeCun, Touresky, Hinton [104], Pearlmutter [132]), une façon de faire est de modéliser les réseaux de neurones profonds à l'aide d'un problème de contrôle optimal en temps continu. Plus précisément, on peut modéliser un réseau de neurones profond par

$$X_{t+1} = X_t + \delta\varphi(t, \alpha_t, X_t, Z),$$

où  $Z$  représente les données,  $X_t$  est la sortie de la  $t$ -ième couche du réseau et  $(\alpha_t)$  représente les paramètres de la  $t$ -ième couche du réseau, qui peut être vu comme le contrôle du système. En temps

continu, la dynamique devient

$$dX_t = \varphi(t, \alpha_t, X_t, Z)dt.$$

En contrôlant  $\alpha$ , on cherche à minimiser

$$\inf_{\alpha} \mathbb{E} \left[ g(|f(Z) - X_T|) + \int_0^T \frac{\lambda}{2} |\alpha_t|^2 dt \right],$$

où  $\mathbb{E}$  est l'espérance par rapport à la variable  $Z$  représentant les données. Toutefois, à notre connaissance, il n'y a pas de littérature existante qui explique rigoureusement pourquoi l'algorithme de la descente du gradient stochastique peut approcher le minimum global de (1.2.10) sous des conditions pas trop restrictives.

Nous nous intéressons donc au problème de contrôle déterministe général suivant

$$\inf_{\alpha} V^0(\alpha), \quad \text{où} \quad V^0(\alpha) := \int_0^T L(t, X_t^\alpha, \alpha_t) dt + G(X_T^\alpha) \quad \text{et} \quad dX_t = \phi(t, X_t^\alpha, \alpha_t) dt. \quad (1.2.14)$$

En particulier, nous cherchons à donner un algorithme qui arrive à résoudre ce genre de problème lorsque la dimension du processus d'état  $X$  et du contrôle  $\alpha$  sont toutes les deux très grandes, comme dans le cas des réseaux de neurones profonds. Les méthodes classiques comme l'équation de Bellman et le principe du maximum de Pontryagin [16] deviennent inefficace lorsque la dimension du problème devient grande. Rappelons que pour qu'un contrôle  $\alpha^*$  soit optimal pour le problème (1.2.14),  $\alpha^*$  doit satisfaire, d'après le principe du maximum de Pontryagin, le système d'équation différentielle ordinaire suivant :

$$\begin{cases} \alpha_t^* = \operatorname{argmin}_a H(t, X_t^*, a, P_t^*), & \text{où} \quad H(t, x, a, p) := L(t, x, a) + p \cdot \phi(t, x, a), \\ X_t^* = x_0 + \int_0^t \phi(r, X_r^*, \alpha_r^*) dr, \\ P_t^* = \nabla_x G(X_T^*) + \int_0^t \nabla_x H(r, X_r^*, \alpha_r^*, P_r^*) dr. \end{cases}$$

Rappelons également que cette condition nécessaire devient suffisante quand les coefficients du système satisfont certaines conditions de convexité. Pour résoudre ce système, une méthode naïve est d'utiliser un algorithme de point fixe, commençant par un contrôle  $\alpha$  quelconque, de résoudre l'équation progressive puis l'équation rétrograde et enfin calculer un nouveau contrôle  $\bar{\alpha}$  qui minimise l'hamiltonien dans (1.2.14). Sous certaines hypothèses, on peut montrer que la fonction  $\alpha \rightarrow \bar{\alpha}$  est une contraction au moins sur un petit intervalle de temps, voir e.g. Ma, Wu, Zhang, Zhang [110] pour une discussion générale sur les EDS progressives-rétrogrades. Malgré la simplicité apparente de l'algorithme, l'optimisation dans la première ligne de (1.2.14) devient difficile à résoudre dès que la dimension du problème est grande (sauf dans les cas où il y a une formule explicite du minimiseur). C'est pourquoi après la découverte du principe du maximum de Pontryagin, plusieurs auteurs ont étudié les algorithmes liés à la descente du gradient, voir e.g. Bryson, Denham [21], Mitter [121]. Dans un tel algorithme, à chaque étape on met à jour la valeur

du processus de contrôle dans la direction opposée du gradient du Hamiltonien  $\nabla_a H$ , *i.e.*

$$\alpha^{i+1} := \alpha^i - \eta \nabla_a H(t, X_t^i, a, P_t^i), \quad (1.2.15)$$

où  $\eta$  est le taux d'apprentissage,  $X^i, P^i$  représente le processus progressif et rétrograde évalués avec  $\alpha^i$ , respectivement. Pour comprendre la convergence de ce genre d'algorithmes de la descente du gradient, nous considérons la version continue de (1.2.15) en l'horizon infini :

$$\begin{cases} \frac{d\alpha_t^s}{ds} = -\nabla_a H(t, X_t^s, \alpha_t^s, P_t^s) \quad \text{sur } \{s \geq 0\} \quad \text{pour tout } t \in [0, T], \\ X_t^s = x_0 + \int_0^t \phi(r, X_r^s, \alpha_r^s) dr, \\ P_t^s = \nabla_x G(X_T^s) + \int_0^t \nabla_x H(r, X_r^s, \alpha_r^s, P_r^s) dr. \end{cases} \quad (1.2.16)$$

On peut vérifier après une série de calcul que

$$\frac{dV^0(\alpha^s)}{ds} = - \int_0^T \left| \nabla_a H(r, X_r^s, \alpha_r^s, P_r^s) \right|^2 dr.$$

où  $V^0$  est défini dans (1.2.14). Ainsi,  $V^0$  est une fonction de Lyapunov naturelle pour le processus de contrôle  $(\alpha^s)$ , et pour que l'égalité  $\frac{dV^0(\alpha)}{ds} = 0$  soit vraie, le contrôle  $\alpha$  doit vérifier le système d'ODE (1.2.16). Cette analyse montre, au moins intuitivement, pourquoi cet algorithme peut converger. Toutefois, comme tout algorithme de la descente du gradient, sans des structures particulières, il converge *a priori* vers un minimum local car le principe du maximum de Pontryagin n'est simplement qu'une condition nécessaire du premier ordre. Si on impose la convexité du système, le processus d'état  $X$  doit être linéaire en  $\alpha$ , autrement dit la fonction  $\phi$  est linéaire en  $(x, a)$ , ce qui est une hypothèse très restrictive.

Pour aller plus loin que le cas convexe, nous faisons appel à la dynamique de Langevin comme dans le cas du réseau de neurones à une couche cachée. Nous introduisons d'abord la formulation relaxée de (1.2.14). Plus précisément, au lieu de contrôler directement la valeur du processus  $\alpha$ , nous contrôlons la loi du processus. Le problème relaxé s'écrit ainsi

$$X_t = x_0 + \int_0^t \int_{\mathbb{R}^m} \phi(r, X_r, a) \nu_r(da) dr,$$

et nous cherchons à minimiser

$$\inf_{\nu} V(\nu), \quad \text{où } V(\nu) := \int_0^T \int L(t, X_t, a) \nu_t(da) dt + G(X_T).$$

On a immédiatement  $\inf_{\nu} V(\nu) \leq \inf_{\alpha} V^0(\alpha)$ . Par des arguments classiques comme dans Fleming [57], El Karoui, Nguyen, Jeanblanc-Picqué [51], la valeur du minimum des deux formulations coïncident sous certaines conditions standards. Nous introduisons également l'entropie relative par rapport à la

mesure de Lebesgue comme précédemment. Le problème d'optimisation régularisé est donné par

$$\inf_{\nu} V^{\sigma}(\nu), \quad \text{où} \quad V^{\sigma}(\nu) := V(\nu) + \frac{\sigma^2}{2} \int_0^T \text{Ent}(\nu_t) dt. \quad (1.2.17)$$

Nous montrons que le système de Langevin à champ-moyen suivant admet une unique mesure invariante et que le minimiseur global du problème régularisé est encore une fois caractérisé par cette mesure invariante.

$$\begin{cases} d\Theta_t^s = -\nabla_a H(t, X_t^s, \Theta_t^s, P_t^s) ds + \sigma dW_s, & \text{pour } s \in \mathbb{R}^+, \quad \text{pour } t \in [0, T], \quad \text{où} \\ X_t^s = X_0 + \int_0^t \int_{\mathbb{R}^m} \phi(r, X_r^s, a) \nu_r^s(da) dr, & \text{avec } \nu_r^s := \text{loi}(\Theta_r^s), \\ P_t^s = \nabla_x G(X_T^s) + \int_t^T \int_{\mathbb{R}^m} \nabla_x H(r, X_r^s, a, P_r^s) \nu_r^s(da) dr, \end{cases} \quad (1.2.18)$$

Remarquons que les différentes équations sont couplées à travers de la loi marginale  $(\nu_t^s)_{t \in [0, T], s \in \mathbb{R}^+}$ . Nous montrons également que cette caractérisation reste valable même si  $V$  n'est pas convexe et nous donnons une preuve probabiliste pour la convergence de la loi marginale du système (1.2.18) vers l'unique mesure invariante. En particulier, nous calculons explicitement la vitesse de convergence.

Notre système de Langevin à champ-moyen (1.2.18) et son lien avec le problème d'optimisation relaxé (1.2.17) suggèrent une méthode numérique pour calculer le contrôle optimal et nous observons comme dans la section précédente que le schéma d'Euler explicite de (1.2.18) est équivalent à la descente du gradient régularisée et randomisée classique pour beaucoup de réseaux de neurones profonds, e.g. les réseaux résiduels, les réseaux de convolution etc. Enfin, nous conjecturons que la mesure invariante du système (1.2.18) est stable par rapport à la loi des données  $\mathcal{L}(Z)$ , qui se traduit par la robustesse de l'algorithme de descente de gradient randomisé et régularisé pour les réseaux de neurones surparamétrés.

---

---

## CHAPITRE II

---

# CONTINUOUS-TIME PRINCIPAL-AGENT PROBLEM IN DEGENERATE SYSTEMS

**Abstract :** In this chapter we present a variational calculus approach to Principal-Agent problem with a lump-sum payment on finite horizon in degenerate stochastic systems, such as filtered partially observed linear systems. Our work extends the existing methodologies in the Principal-Agent literature using dynamic programming and BSDE representation of the contracts in the non-degenerate controlled stochastic systems. We first solve the Principal's problem in an enlarged set of contracts defined by a forward-backward SDE system given by the first order condition of the Agent's problem using variational calculus. Then we use the sufficient condition of the Agent's problem to verify that the optimal contract that we obtain by solving the Principal's problem is indeed implementable (i.e. belonging to the admissible contract set). Importantly we consider the control problem in a weak formulation. Finally, we give explicit solution of the Principal-Agent problem in partially observed linear systems and extend our results to some mean field interacting Agents case.

**Keywords :** Stochastic Control of non-Markovian systems, Stochastic Maximum Principles, path-dependent Forward-backward SDEs, Principal-Agent Problem, Contract Theory



## II.1 Introduction

Moral hazard is one of the prime risks of systemic instability and inefficiency, as already pointed out by Adam Smith. Finding the optimal contract between two parties - The Principal and the Agent, when the Agent's effort cannot be observed therefore cannot be contracted upon, is a classical moral hazard problem in microeconomics. Applications can be widely found in corporate finance, portfolio management [15] and more recently energy transition [2, 1].

While the research on discrete-time models dated further back, the first continuous-time model was proposed in the seminal work of Holmström and Milgrom [78], in which they study a simplified model with lump-sum payment on finite horizon while assuming the Agent controls only the drift of the state process. They show that the optimal contract should be linear in aggregate output when the Agent has CARA utility functions with a monetary cost of effort. Their work has been extended by Schättler and Sung [140, 141], Sung [146, 147], Müller [123, 124], Hellwig and Schmidt [73]. Later Williams [155], Cvitanić, Wan and Zhang [38] use the stochastic maximum principle and forward-backward stochastic differential equations (abbreviated FBSDE) to characterize the optimal compensation for more general utility functions.

Many different continuous-time models have been proposed in the past few decades, notably the one introduced by Sannikov in his seminal paper [139]. He considered Principal-Agent problem on infinite horizon with continuous payment while allowing Principal to fire or retire the Agent at any time. From a mathematical perspective, his approach sheds light on the contracting relationship in Principal-Agent's problem using dynamic programming and leads to simple computational procedure to find the optimal contract by, in his case, solving an ordinary differential equation.

More recently, Sannikov's idea was reinterpreted and extended by Cvitanić, Possamaï and Touzi [36] to a more general set-up with a more direct and easier approach. We shall illustrate their contribution in the following toy model. Denote by  $\xi$  the contract paid at the end of a finite time interval  $[0, T]$ . Let's consider the following optimization for the Agent :

$$\max_{\alpha} \mathbb{E} \left[ \xi(X^{\alpha}) - \int_0^T c(\alpha_t) dt \right], \quad \text{where} \quad dX_t^{\alpha} = dW_t - \alpha_t dt.$$

The crucial observation in [36] is that both the contract  $\xi$  and the Agent's best response  $\alpha^*[\xi]$  can be characterized by the following backward stochastic differential equation (in short, BSDE, for readers not familiar with BSDE we refer to [129, 52], and in particular to [41] for the applications on the contract theory) :

$$dY_t = -c^*(Z_t)dt + Z_t dW_t, \quad Y_T = \xi, \quad \text{where} \quad c^*(z) = \max_a \{az - c(a)\},$$

namely,  $\xi = Y_T$  and  $\alpha_t^*[Z] = \operatorname{argmax}_a \{aZ_t - c(a)\}$  for all  $t \in [0, T]$ . This induces a natural (forward)

representation of the couple  $(\xi, \alpha^*[\xi])$  :

$$\begin{cases} \xi = Y_T^{Y_0, Z} := Y_0 - \int_0^T c^*(Z_t)dt + \int_0^T Z_t dW_t \\ \alpha_t^*[\xi] := \alpha_t^*[Z] = \operatorname{argmax}_a \{aZ_t - c(a)\}, \text{ for all } t \in [0, T] \end{cases} \quad \text{for some } (Y_0, Z),$$

and this neat representation transforms the once puzzling principal's problem to be a classical stochastic control problem, namely,

$$\max_{\xi} \mathbb{E} \left[ U \left( X_T^{\alpha^*[\xi]} - \xi \right) \right] = \max_{Y_0, Z} \mathbb{E} \left[ U \left( X_T^{\alpha^*[Z]} - Y_T^{Y_0, Z} \right) \right].$$

This idea of representation is further applied to study the case where the Principal can hire multiple Agents e.g. [54], [55] or Agent facing multiple Principals e.g. [115], [83] using the formulation of mean field games (as for the mean field game we refer to the seminal paper [102] and the recent books [25] and [26]).

However, the method relies on an important hypothesis : the stochastic system is non-degenerate. More specifically, at the heart of the dynamic programming approach to Principal-Agent problem lies the BSDE representation of the contract as shown in the above example, which requires the non-degeneracy of the system.

It is well known that degenerate stochastic systems appear naturally in partially observed systems when we replace the non-observable part of the system by the filtered process, i.e. the conditional law of the unobservable given the observable, see e.g. [10]. One of the simplest cases of partial observed system is system with parameter uncertainty. This is studied by Fernandes and Phelan [56], Williams [156], then by Prat and Jovanovic [136], which inspires and motivates this work. To illustrate the idea, consider the following process

$$dB_t = (\mu + \beta_t)dt + \sigma dW_t, \quad (\text{II.1.1})$$

where  $\beta$  is the control of the Agent and  $\mu$  represents a unknown parameter of the system which is called the time-invariant productivity. The common priors on  $\mu$  are normal with mean  $m_0$  and variance  $V_0$ . Using Bayes' formula and replace  $\mu$  by the posteriors, which depend on  $B_t$  and the cumulative effort  $A_t := \int_0^t \beta_s ds$ , we get the following controlled system :

$$dB_t = \left( \frac{V_t}{V_0} m_0 + \frac{V_t}{\sigma^2} (B_t - A_t) + \beta_t \right) dt + dI_t,$$

where  $V_t = \frac{\sigma^2 V_0}{\sigma^2 + V_0 t}$  is the posterior variance and  $(I_t)_{t \geq 0}$  is a  $\mathbb{F}^B$ -adapted Brownian motion called the innovation process. The partially observed state process (II.1.1) becomes now a degenerate controlled system with a new degenerate process  $(A_t)_{t \geq 0}$ . Since the process  $A$  is not observable by the Principal, it cannot be contracted upon. Indeed, to compute the posteriors on  $\mu$  at time  $t$ , one needs to know the Agent's control  $\beta$  up to time  $t$ . The main difficulty is to find a dynamic representation of the contracts

but using only the observable part of the system, namely in the above case the process  $B$ .

In this chapter we consider general degenerate controlled system in a weak formulation, allowing the drift and the volatility of the filtered process (i.e. degenerate part of the system, denoted throughout the chapter by  $X$ ) be controlled by the Agent. We assume that the observable part of the system (i.e. non-degenerate part of the system, denoted throughout the chapter by the process  $B$ ) is controlled by the Agent only via the drift. We give the first order condition of the Agent's optimal control when he is given a contract. The first order condition is described by a path-dependent FBSDE. Contrary to the FBSDE literature, by path-dependent we mean the coefficients of the FBSDE dependent on the path of the forward process. In the Principal-Agent problem, the dependency mainly comes from the Agent's cost function and the contract. The wellposedness of these kinds of FBSDE is of independent interest and is studied in the accompanying paper [80]. Our strategy is to solve the Principal's problem first over the set of contracts described by the FBSDE given by the first order condition of the Agent's problem. The contracts described by the FBSDE may not necessarily be implementable, i.e. the optimal control of the Agent's problem may not exist. The implementability is then checked by the sufficient condition of the optimality of the Agent's problem.

The main contribution of this work is to provide a way to solve the Principal-Agent's problem when only part of the state variables can be contracted upon, which is crucial when the system is partially observed. The rest of the chapter is organised as follow. In section 2 we state our moral hazard problem in degenerated systems and give the main results of the chapter. The proofs of the results will be given in Section 3. In Section 4 we solve the Principal-Agent problem in partially observed linear systems with explicit optimal contract and extend the result to a specific mean-field interacting Agents case.

## II.2 Principal-Agent Problem in Degenerate Systems

### II.2.1 Preliminaries

Let  $\Omega := C([0, T], \mathbb{R}^d)$  be the canonical space and  $B$  the canonical process. Denote by  $\mathbb{F} = (\mathcal{F}_t^B)_{0 \leq t \leq T}$  the associated filtration and  $\mathbb{P}$  the Wiener measure. Let  $A$  be a compact convex subset of some finite dimensional space. Let

$$b : [0, T] \times \Omega \times \mathbb{R}^n \times A \rightarrow \mathbb{R}^d, b(\cdot, a) \text{ } \mathbb{F}\text{-optional for any } a \in A,$$

$$\eta : [0, T] \times \Omega \times \mathbb{R}^n \times A \rightarrow \mathbb{R}^n, \eta(\cdot, a) \text{ } \mathbb{F}\text{-optional for any } a \in A,$$

$$\sigma : [0, T] \times \Omega \times \mathbb{R}^n \times A \rightarrow \mathcal{M}_{n,d}(\mathbb{R}), \sigma(\cdot, a) \text{ } \mathbb{F}\text{-optional for any } a \in A,$$

where  $\mathcal{M}_{n,d}(\mathbb{R})$  denotes the set of  $n \times d$  matrices with real entries. Throughout this chapter we shall be studying control problems in a weak formulation. Denote  $\mathcal{U}$  the set of admissible controls taking values in  $A$ . For any  $\alpha \in \mathcal{U}$ , under the standard global Lipschitz conditions of Assumption II.2.5 below, let  $X$

be the unique strong solution of the following stochastic differential equation

$$dX_t = \eta_t(X_t, \alpha_t)dt + \sigma_t(X_t, \alpha_t)dB_t, \quad (\text{II.2.2})$$

for some given initial data  $X_0$ .

Then we can define  $d\mathbb{P}^\alpha|_{\mathcal{F}_T} := \mathcal{E}_T^\alpha d\mathbb{P}|_{\mathcal{F}_T}$  where

$$\mathcal{E}_T^\alpha := \exp\left(\int_0^T b_t(X_t, \alpha_t) \cdot dB_t + \frac{1}{2} \int_0^T |b_t(X_t, \alpha_t)|^2 dt\right). \quad (\text{II.2.3})$$

In the case where the above process is a positive martingale, by Girsanov theorem we can define  $W_t^\alpha := B_t - B_0 - \int_0^t b_s(X_s, \alpha_s)ds$ , which is a  $d$ -dimensional Brownian motion under  $\mathbb{P}^\alpha$  and which means that the canonical process  $B$  satisfies the following stochastic differential equation

$$dB_t = b_t(X_t, \alpha_t)dt + dW_t^\alpha. \quad (\text{II.2.4})$$

This defines the family of probability measures on  $\Omega$

$$\mathcal{P} := \left\{ \mathbb{P}^\alpha \in \mathcal{P}(\Omega) : \alpha \in \mathcal{U} \right\},$$

where  $\mathcal{P}(\Omega)$  is the set of all probability measures on  $\Omega$ .

## II.2.2 Problem Formulation

The Agent signs a contract, works for the Principal for a predetermined period  $T$  and receives  $\xi$  at time  $T$ . A contract  $\xi$  is a  $\mathcal{F}_T$ -measurable random variable, which represents the payment the Agent receives at time  $T$ . Let  $c : [0, T] \times \Omega \times A \rightarrow \mathbb{R}$  be a measurable function representing the Agent's cost of effort at time  $t$ . Let  $k^A : [0, T] \times \Omega \times A \rightarrow \mathbb{R}$  be a bounded  $\mathbb{F}$ -optional function and define  $\mathcal{K}_T^A := \exp\left(-\int_0^T k_t^A(\alpha_t)dt\right)$  representing the Agent's discount factor. The Agent aims at choosing an optimal effort  $\alpha \in \mathcal{U}$  to optimize his utility when given a contract  $\xi$  proposed by the Principal :

$$V_A(\xi) := \sup_{\alpha \in \mathcal{U}} J_A(\alpha) := \sup_{\alpha \in \mathcal{U}} \mathbb{E}^\alpha \left[ \mathcal{K}_T^A \xi - \int_0^T \mathcal{K}_t^A c_t(\alpha_t) dt \right], \quad (\text{II.2.5})$$

where  $\mathbb{E}^\alpha$  denotes the expectation under the probability measure  $\mathbb{P}^\alpha \in \mathcal{P}$  defined previously. For any contract  $\xi$ , denote  $\mathcal{M}^*(\xi) \subset \mathcal{U}$  the set of optimal controls of the Agent's problem.

The Principal on the other hand takes benefits from the outcome of the controlled process and pays the Agent accordingly. Also, the Agent's participation is conditioned on having his expected return above his reservation utility  $R$ , in other words :  $V_A(\xi) \geq R$ . In all, the Principal choose a contract among the

following set of contracts in order to optimize her own utility :

$$\Xi = \left\{ \xi \text{ } \mathcal{F}_T\text{-measurable} : \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}^{\mathbb{P}} [\xi^2] < \infty \text{ and } V_A(\xi) \geq R \right\}. \quad (\text{II.2.6})$$

We will call the contracts in  $\Xi$  the implementable contracts. In the case where the Agent has more than one optimal control, we follow the standard convention that the Agent chooses the one that is the best for the Principal. Let  $k^P : [0, T] \times \Omega \rightarrow \mathbb{R}$  be a bounded  $\mathbb{F}$ -optional function and define  $\mathcal{K}_T^P := \exp\left(-\int_0^T k_t^P dt\right)$ , which represents the discount factor of the Principal. She aims at solving the following optimization

$$V_P = \sup_{\xi \in \Xi} \sup_{\alpha^* \in \mathcal{M}^*(\xi)} \mathbb{E}^{\alpha^*} [\mathcal{K}_T^P U(B_T - \xi)],$$

where the function  $U : \mathbb{R} \rightarrow \mathbb{R}$  is a given non-decreasing and concave utility function and we use the convention  $\sup \emptyset = -\infty$ .

**Remark II.2.1.** *One of our goals is to generalize the setting of the paper [36] by Cvitanić, Possamai and Touzi, in which they consider the following system*

$$X_t = X_0 + \int_0^t \sigma_r(X, \beta_t)(\lambda_r(X, a_r)dr + dW_r), \quad t \in [0, T].$$

*The drift of the dynamic needs to be in the range of the volatility, which means the system cannot be degenerate.*

### II.2.3 Motivation

The main motivation of our formulation is to tackle the Principal-Agent problem in a partially observed linear system. Using the above notation, the canonical process  $B$  represent the observable process, observed by the Agent and the Principal at the same time. The degenerate part of the system described by the process  $X$  represents the filtered process.

More precisely, let  $\nu = (\alpha, \beta) \in \mathcal{U}$  be an admissible control. Let  $\hat{X}$  be the unique strong solution of the following linear stochastic differential equation representing the unobservable part of the system :

$$d\hat{X}_t = (\eta(t)\hat{X}_t + \alpha_t)dt + \sigma(t)dW_t, \quad \hat{X}_0 = \mu, \quad (\text{II.2.7})$$

where  $\mu$  a unobservable random variable independent of  $W$  assumed to be Gaussian with mean and variance  $m_0$  and  $V_0$ , respectively. In addition, we assume that the Principal and the Agent both know  $m_0$  and  $V_0$ . The functions  $\eta$  and  $\sigma$  are deterministic. Furthermore, the observable part of the system is generated by the noisy signal

$$dB_t = (h(t)\hat{X}_t + \beta_t)dt + dW_t^\nu, \quad (\text{II.2.8})$$

where  $W^\nu$  is a Brownian motion independent of  $W$  and  $h$  a deterministic function.



More generally speaking, one can consider the following partially observed system :

$$\begin{aligned} dB_t &= b_t(\hat{X}_t, \beta_t)dt + dW_t \\ d\hat{X}_t &= \eta_t(\hat{X}_t, \alpha_t)dt + \sigma_t(\hat{X}_t, \alpha_t)d\tilde{W}_t, \end{aligned}$$

where  $B$  and  $\hat{X}$  represent as before the observable process and the unobservable process, respectively. One way to transform the system into an equivalent but fully observable system is to replace the unobservable part by its (unnormalized) conditional law described by Zakai equation. In this chapter, we shall mainly focus on the finite dimensional case.

**Remark II.2.4.** *The moral hazard framework of Principal-Agent problem is also frequently called the second best. The first best corresponds to the case where the Principal and the Agent have the same information. It is typically assumed that the Principal dictates the Agent's actions. Mathematically, the problem becomes a stochastic control problem for a single individual - the Principal. It is well-known that in non-degenerate systems under some mild conditions, the first best coincides with the second best in terms of Principal's value when the Agent's criterion is risk neutral. However, in a partially observed system, the result no longer holds true in general, even in the risk-neutral case. Consider the following example where there is no observable process. The unobservable process follows a SDE :*

$$d\hat{X}_t = \alpha_t dt + dW_t \quad \text{with } \hat{X}_0 \sim \mathcal{N}(m_0, V_0) \text{ independent of } \mathbb{F}^W.$$

Clearly, the filtered process is given by  $X_t = m_0 + \int_0^t \alpha_s ds$ . The Principal aims at optimizing his utility function :

$$V_P = \sup_{\xi \in \Xi} \sup_{\alpha^* \in \mathcal{M}^*(\xi)} \mathbb{E}^{\alpha^*} [\hat{X}_T - \xi],$$

where  $\alpha^*$  is the optimal response of the Agent given the contract  $\xi$ , namely

$$\alpha^* = \operatorname{argmax}_{\alpha \in \mathcal{U}} \mathbb{E}^\alpha \left[ \xi - \frac{1}{2} \int_0^T \alpha_t^2 dt \right].$$

In the second best case, the Principal has no information and can therefore only offer constant contracts. Agent has no incentive to work and therefore  $\alpha \equiv 0$ . Taking into account the participation constraint the optimal contract is a constant payment  $R$  at time  $T$  where  $R$  is Agent's reservation value. The Principal's value in this case is  $m_0 - R$ . In the first best case, the Principal observes the Agent's action and therefore the filtered process can be contracted upon. For simplicity, let's consider only linear contracts, i.e.  $\xi = cX_T + d$ . The Agent's optimal control is  $\alpha^* \equiv c$  with participation constraint  $\frac{Tc^2}{2} + m_0c \geq R - d$ . As for the Principal, for any given linear contract  $\xi$  we have

$$\mathbb{E}^{\alpha^*} [\hat{X}_T - \xi] = \mathbb{E}^{\alpha^*} \left[ \hat{X}_T - \frac{1}{2} \int_0^T (\alpha_t^*)^2 dt \right] - \mathbb{E}^{\alpha^*} \left[ \xi - \frac{1}{2} \int_0^T (\alpha_t^*)^2 dt \right] \leq m_0 + \frac{T}{2} - R.$$

The inequality becomes equality when  $c = 1$  and  $d = R - m_0 - \frac{T}{2}$ , in which case the Principal's value is increased by  $\frac{T}{2}$  comparing to the moral hazard case.

## II.2.4 Main Results

For presentation simplicity, we assume that the Agent's discount factor  $\mathcal{K}_t^A$  is equal to 1 for all  $t \in [0, T]$ . All the main results can be generalised without much difficulty to the case with  $\mathcal{K}_t^A = \exp\left(-\int_0^t k_s^A(\alpha_s) ds\right)$  where  $k^A : [0, T] \times \Omega \times A \rightarrow \mathbb{R}$  is a bounded  $\mathbb{F}$ -optional function. Throughout this chapter, we shall impose the following conditions on these coefficients.

**Assumption II.2.5.** *The maps  $b, \sigma, \eta$  are  $C^2$  in  $x$ . Moreover, there exists a constant  $L > 0$  and a modulus of continuity  $\varpi : [0, \infty) \rightarrow [0, \infty)$  such that for  $\phi = b, \sigma, \eta, \partial_x b, \partial_x \sigma, \partial_x \eta$ ,  $\mathbb{P}$ -almost surely*

$$|\phi_t(x, a) - \phi_t(x', a')| \leq L|x - x'| + \varpi(|a - a'|) \text{ for all } t \in [0, T], x, x' \in \mathbb{R}^n, a, a' \in A$$

and for  $\phi = b, \sigma, \eta$ ,  $\mathbb{P}$ -almost surely

$$|\partial_{xx}\phi_t(x, a) - \partial_{xx}\phi_t(x', a')| \leq \varpi(|x - x'| + |a - a'|) \text{ for all } t \in [0, T], x, x' \in \mathbb{R}^n, a, a' \in A.$$

The following assumption ensures that all weak controls  $\mathbb{P}^\alpha \in \mathcal{P}$  are equivalent and that high order error terms in the variational calculus can be ignored in the non-linear system case.

**Assumption II.2.6.** *The Doléans-Dade exponential  $\mathcal{E}^\alpha$  defined in (II.2.3) is a positive martingale for all  $\alpha \in \mathcal{U}$ . Furthermore, if the controlled system is not linear, namely not in the form of (II.2.7)-(II.2.8), there exists  $p > 2$  such that  $\sup_{\alpha \in \mathcal{U}} \mathbb{E}[(\mathcal{E}_T^\alpha)^p] < \infty$ .*

We aim to solve the Agent's problem (II.2.5) by variational calculus. In the following, for  $p \geq 1$ ,  $\mathbb{P}^\alpha \in \mathcal{P}$  and  $T > 0$ , denote

$$\mathbb{H}_T^p(\mathbb{P}^\alpha) := \left\{ Z \text{ } \mathbb{F}\text{-adapted} : \|Z\|_{\mathbb{H}_T^p(\mathbb{P}^\alpha)} := \mathbb{E}^\alpha \left[ \int_0^T |Z_t|^p dt \right]^{\frac{1}{p}} < +\infty \right\}.$$

For  $(t, \omega, x, x', a, a', p, q, z) \in [0, T] \times \Omega \times (\mathbb{R}^n)^2 \times A^2 \times \mathbb{R}^n \times \mathcal{M}_{n,d}(\mathbb{R}) \times \mathbb{R}^d$ , define the Hamiltonian :

$$H_t(x, x', a, a', p, q, z) := p \cdot (\eta_t(x, a) + \sigma_t(x, a)b_t(x', a')) + \text{Tr}[q\sigma_t^\top(x, a)] + z \cdot b_t(x, a) - c_t(a). \quad (\text{II.2.15})$$

Implicitly, at time  $t$  the Hamiltonian may depend on the path of canonical process  $B_{\cdot \wedge t}$  up to time  $t$ . In the following, we will use frequently the partial derivative  $\partial_a H, \partial_x H$ , which are derivatives with respect to  $x$  and  $a$  but not  $x', a'$ .



**Theorem II.2.7** (Necessary Condition). *Assume that Assumption II.2.5 and II.2.6 hold true. For a contract  $\xi \in \Xi$ , let  $\alpha^* \in \mathcal{M}^*(\xi)$ . Then the following FBSDE*

$$\begin{cases} dY_t = c_t(\alpha_t^*)dt + Z_t dW_t^{\alpha^*} & Y_T = \xi, & \text{(II.2.16)} \\ dP_t = -\partial_x H_t(X_t, X_t, \alpha_t^*, \alpha_t^*, P_t, Q_t, Z_t)dt + Q_t dW_t^{\alpha^*} & P_T = 0, & \text{(II.2.17)} \\ dB_t = \partial_z H_t(X_t, X_t, \alpha_t^*, \alpha_t^*, P_t, Q_t, Z_t)dt + dW_t^{\alpha^*} & B_0 = 0, & \text{(II.2.18)} \\ dX_t = \partial_p H_t(X_t, X_t, \alpha_t^*, \alpha_t^*, P_t, Q_t, Z_t)dt + \sigma_t(X_t, \alpha_t^*)dW_t^{\alpha^*} & X_0 = x_0, & \text{(II.2.19)} \end{cases}$$

where  $\alpha^*$  verifies

$$\alpha_t^* = \operatorname{argmax}_{a \in A} H_t(X_t, X_t, a, \alpha_t^*, P_t, Q_t, Z_t) \text{ for all } t \in [0, T], \mathbb{P}^{\alpha^*} - a.s. \quad \text{(II.2.20)}$$

has an  $\mathbb{F}$ -adapted solution, denoted by

$$(Y^*, Z^*, P^*, Q^*, X^*) \in \mathbb{H}_T^2(\mathbb{P}^{\alpha^*})^5.$$

**Remark II.2.8.** 1. *In the case where  $\eta = \sigma = 0$  and  $b$  does not depend on  $x$ , our formulation is reduced to the framework of Principal-Agent problem studied in [36] in the uncontrolled diffusion case, where the state equations are not degenerate. The FBSDE given in Theorem II.2.7 reduces to the BSDE (II.2.16), which is the same representation of the contract under the optimal probability using the dynamic programming approach. To see that, note that the Hamiltonian in this case becomes*

$$H_t(a, z) = zb_t(a) - c_t(a)$$

and together with the definition of the optimal control given in (II.2.20), the equation (II.2.16) can be written as

$$\begin{aligned} dY_t &= H_t(\alpha_t^*, Z_t)dt + Z_t dB_t \\ &= \max_{a \in A} H_t(a, Z_t)dt + Z_t dB_t, \quad Y_T = \xi, \end{aligned}$$

which is the canonical representation of the contract  $\xi$  given in [36, Definition 3.2].

2. *Since we are working on the weak formulation of stochastic control, it is not surprising that the necessary condition we obtain here does not coincide with the standard stochastic Pontryagin's Maximum Principle (see e.g. [133, 91]). Comparing to the standard result, the forward-backward system has an extra backward equation (II.2.16). The process  $Y$  in (II.2.16) indeed characterizes the value function. The adjoint process (II.2.17) and the forward one (II.2.19) are similar to those appearing in the standard stochastic maximum principle.*
3. *The contract  $\xi$  appearing in the terminal condition of (II.2.16) may be path-dependent. To the best of our knowledge, the well-posedness of such FBSDE has not yet been studied in the literature. For*

further details, we refer the readers to the accompanying paper on the path-dependent FBSDEs [80].

4. Another way to see how to obtain the FBSDE (II.2.16)-(II.2.19) is to conduct formally the variational calculus by introducing a new state process  $L$  representing the change of measure, namely

$$dL_t = b_t(X_t, \alpha_t)dB_t$$

and rewrite the Agent's problem as following :

$$V_A = \sup_{\alpha \in \mathcal{U}} \mathbb{E} \left[ L_T \xi - \int_0^T L_t c_t(\alpha_t) dt \right],$$

where the controlled system becomes

$$\begin{cases} dX_t = \eta_t(X_t, \alpha_t)dt + \sigma_t(X_t, \alpha_t)dB_t & X_0 = x_0, \\ dL_t = L_t b_t(X_t, \alpha_t)dB_t & L_0 = 1. \end{cases}$$

The standard Hamiltonian associated with the above control problem is given by

$$\hat{H}_t(x, l, a, p, q, z) = \eta_t(x, a) \cdot p + \text{Tr}[q\sigma_t^\top(x, a)] + l(zb_t(x, a) - c_t(a)). \quad (\text{II.2.21})$$

The adjoint equations associated with  $X$  and  $L$  are given by

$$\begin{cases} dP_t = -\partial_x \hat{H}_t(X_t, L_t, \alpha_t, P_t, Q_t, Z_t)dt + Q_t dB_t & P_T = 0, \\ dY_t = -\partial_l \hat{H}_t(X_t, L_t, \alpha_t, P_t, Q_t, Z_t)dt + Z_t dB_t & Y_T = \xi. \end{cases} \quad (\text{II.2.22})$$

$$\begin{cases} dP_t = -\partial_x \hat{H}_t(X_t, L_t, \alpha_t, P_t, Q_t, Z_t)dt + Q_t dB_t & P_T = 0, \\ dY_t = -\partial_l \hat{H}_t(X_t, L_t, \alpha_t, P_t, Q_t, Z_t)dt + Z_t dB_t & Y_T = \xi. \end{cases} \quad (\text{II.2.23})$$

The process  $(Y, Z)$  in (II.2.23) is actually the equation (II.2.16) and if we define  $\hat{P}_t = \frac{P_t}{L_t}$  and  $\hat{Q}_t = \frac{Q_t}{L_t} - \frac{P_t b_t(X_t, \alpha_t)}{L_t}$ , then the process  $(\hat{P}, \hat{Q})$  satisfies exactly the equation (II.2.17).

However, there are a few problems in this approach. First of all, since the coefficients of the control system (II.2.22)-(II.2.23) are not Lipschitz, to our best knowledge, one cannot directly apply the stochastic maximum principle. Our method provides a way to overcome this problem. Secondly, the newly introduced state variable  $L$  is not observable, the contract and the optimal control of the Agent should not depend on  $L$ . And finally, note that the convexity of the Hamiltonian (II.2.21) can be satisfied in very few situations due to the term  $lzb_t(x, a)$ , which makes it tricky to get a useful sufficient condition for the Agent's problem.

Now we give a sufficient condition for the Agent's problem. Define the functional

$$\begin{aligned} \mathcal{G}(t, \omega, x, x', a, a', p, q, z) &:= H_t(x, x', a, a', p, q, z) + (x - x')qb_t(x, a) \\ &\quad + p \left( \sigma_t(x, a)b_t(x, a) - \sigma_t(x, a)b_t(x', a') - \sigma_t(x', a')b_t(x, a) \right). \end{aligned}$$

**Theorem II.2.9** (Sufficient Condition). *Let Assumptions II.2.5 and II.2.6 hold true. Let  $(Y^*, Z^*, P^*, Q^*, X^*)$  be a solution to the system (II.2.16)-(II.2.19), where  $\alpha^*$  satisfies (II.2.20). If  $(x, a) \in \mathbb{R}^n \times A \mapsto \mathcal{G}_t(x, X_t^*, a, \alpha_t^*, P_t^*, Q_t^*, Z_t^*)$  is concave,  $\alpha^*$  is an optimal control for the problem  $V_A(Y_T^*)$ .*

We next focus on the Principal's problem. Based on the necessary and the sufficient condition of the Agent's problem, we consider the following sets of contracts :

$$\bar{\Xi} := \{Y_T | Y_0 \geq R, Z \in \bar{\mathcal{Z}}\} \quad \text{and} \quad \Xi := \{Y_T | Y_0 \geq R, Z \in \underline{\mathcal{Z}}\},$$

where  $Y$  is defined by the equation (II.2.16) and

$$\begin{aligned} \bar{\mathcal{Z}} &:= \left\{ Z \in \mathbb{H}_T^2(\mathbb{P}) : \exists \mathbb{P}^\alpha \in \mathcal{P} \text{ s.t. (II.2.17), (II.2.19) have a solution } (P, Q, X) \text{ in } \mathbb{H}^2(\mathbb{P}^\alpha)^3 \right\}, \\ \underline{\mathcal{Z}} &:= \left\{ Z \in \bar{\mathcal{Z}} : Q \text{ satisfies the sufficient condition given in Theorem II.2.9} \right\}. \end{aligned}$$

Clearly we have  $\Xi \subset \bar{\Xi} \subset \bar{\Xi}$ . Further, we introduce the enlarged optimization of the Principal.

$$\bar{V}_P := \sup_{Y_0 \geq R} \sup_{Z \in \bar{\mathcal{Z}}} \bar{J}_P(Y_0, Z) := \sup_{Y_0 \geq R} \sup_{Z \in \bar{\mathcal{Z}}} \mathbb{E}^{\alpha^*} \left[ \mathcal{K}_T^P U(B_T - Y_T) \right]. \quad (\text{II.2.24})$$

**Theorem II.2.10.** *Assume that Assumption II.2.5 and II.2.6 hold true. Then  $V_P \leq \bar{V}_P$ . Moreover, if the control problem  $\bar{V}_P$  has a solution  $Z^*$  which satisfies the sufficient condition given in Theorem II.2.9, then  $V_P = \bar{V}_P$  and  $\xi = Y_T^*$  is an optimal contract for  $V_P$ .*

## II.3 Proofs

Before giving the proof of the necessary condition, we give the following martingale representation theorem which we shall need in a moment.

**Lemma II.3.1** (Martingale Representation). *Let  $(M_t)_{0 \leq t \leq T}$  be a stochastic process with decomposition*

$$dM_t = h_t dt + dW_t,$$

where  $h$  is a  $\mathbb{F}^M$ -adapted process such that  $\mathbb{E} \left[ \exp \left( \int_0^T h_t dW_t - \frac{1}{2} \int_0^T h_t^2 dt \right) \right] = 1$ . Then for any  $\mathbb{F}^M$  martingale  $\xi = (\xi_t)_{0 \leq t \leq T}$ , there exists a  $\mathbb{F}^M$ -adapted process  $f$  such that

$$\forall t \in [0, T], \xi_t = \xi_0 + \int_0^t f_s dW_s, \mathbb{P} - a.s.$$

*Démonstration.* Let  $(\xi_t)_{0 \leq t \leq T}$  be a  $\mathbb{F}^M$ -martingale. For  $t \in [0, T]$ , define

$$L_t := e^{-\int_0^t h_s dW_s - \frac{1}{2} \int_0^t h_s^2 ds}.$$

Clearly  $(L_t)_{0 \leq t \leq T}$  is a positive uniformly integrable martingale. By Girsanov's theorem, denote  $\mathbb{Q}$  the probability under which  $M$  is a Brownian motion, we have  $\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_t^M} = L_t$ . By Bayes' rule,  $\frac{\xi}{L}$  is a  $\mathbb{F}^M$  martingale under  $\mathbb{Q}$ . Since  $M$  is now a Brownian motion under the probability  $\mathbb{Q}$ , by the martingale representation theorem, there exists a  $\mathbb{F}^M$ -adapted process  $k = (k_t, \mathbb{F}_t^M)$ ,  $0 \leq t \leq T$ , such that  $\forall t \in [0, T]$ ,  $\frac{\xi_t}{L_t} = \xi_0 + \int_0^t k_s dM_s$ . Finally, by Ito's formula, we have

$$d\xi_t = d(L_t \cdot \frac{\xi_t}{L_t}) = -\frac{\xi_t}{L_t} L_t h_t dW_t + L_t k_t dM_t - L_t h_t k_t dt = (L_t k_t - \xi_t h_t) dW_t.$$

■

The following lemma gives an estimate on the error of the first order approximation of the difference between the solution of (II.2.2) given two different control  $\alpha$  and  $\alpha'$ .

**Lemma II.3.2.** *Let  $\alpha, \alpha'$  be two control processes and  $X, X'$  be the corresponding solution of (II.2.2). Denote  $\Delta\alpha := \alpha - \alpha'$  and  $\Delta X := X - X'$ . Let  $\Delta\hat{X}$  be the solution of the following linear SDE with initial condition  $\Delta\hat{X}_0 = 0$  :*

$$d\Delta\hat{X}_t = \left( \partial_x \eta_t(X'_t, \alpha'_t) \Delta\hat{X}_t + \partial_a \eta_t(X'_t, \alpha'_t) \Delta\alpha_t \right) dt + \left( \partial_x \sigma_t(X'_t, \alpha'_t) \Delta\hat{X}_t + \partial_a \sigma_t(X'_t, \alpha'_t) \Delta\alpha_t \right) dB_t. \quad (\text{II.3.25})$$

Then for any  $p \geq 2$ , there exists  $K > 0$  depending only on  $p, T$  and the coefficients  $\eta, \sigma$  of the equation such that

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |\Delta\hat{X}_t - \Delta X_t|^p \right] \leq K \mathbb{E} \left[ \int_0^T |\Delta\alpha_t|^{2p} dt \right],$$

where the expectation is taken under the Wiener measure.

*Démonstration.* In this proof, for the simplicity of the notation, we shall always use  $C$  to denote some strictly positive constants, which may vary from line to line. By Assumption II.2.5 and Taylor expansion, we note that

$$|\eta_t(X_t, \alpha_t) - \eta_t(X'_t, \alpha'_t) - \partial_x \eta_t(X'_t, \alpha'_t) \Delta X_t - \partial_a \eta_t(X'_t, \alpha'_t) \Delta\alpha_t| \leq L(|\Delta\alpha_t|^2 + |\Delta X_t|^2),$$

$$|\sigma_t(X_t, \alpha_t) - \sigma_t(X'_t, \alpha'_t) - \partial_x \sigma_t(X'_t, \alpha'_t) \Delta X_t - \partial_a \sigma_t(X'_t, \alpha'_t) \Delta\alpha_t| \leq L(|\Delta\alpha_t|^2 + |\Delta X_t|^2).$$

Therefore, by triangle inequality, we get

$$|\eta_t(X_t, \alpha_t) - \eta_t(X'_t, \alpha'_t) - \partial_x \eta_t(X'_t, \alpha'_t) \Delta\hat{X}_t - \partial_a \eta_t(X'_t, \alpha'_t) \Delta\alpha_t| \leq L(|\Delta\alpha_t|^2 + |\Delta X_t|^2 + |\Delta\hat{X}_t - \Delta X_t|),$$

$$|\sigma_t(X_t, \alpha_t) - \sigma_t(X'_t, \alpha'_t) - \partial_x \sigma_t(X'_t, \alpha'_t) \Delta\hat{X}_t - \partial_a \sigma_t(X'_t, \alpha'_t) \Delta\alpha_t| \leq L(|\Delta\alpha_t|^2 + |\Delta X_t|^2 + |\Delta\hat{X}_t - \Delta X_t|),$$

for some other constant  $L > 0$ .

By Burkholder-Davis-Gundy inequality and Jensen inequality, we have

$$\begin{aligned} \mathbb{E} \left[ \sup_{t \in [0, T]} |\Delta \hat{X}_t - \Delta X_t|^p \right] &\leq C \mathbb{E} \left[ \int_0^T \left( |\Delta \alpha_t|^{2p} + |\Delta X_t|^{2p} + \sup_{t \in [0, T]} |\Delta \hat{X}_t - \Delta X_t|^p \right) dt \right. \\ &\quad \left. + \left( \int_0^T (|\Delta \alpha_t|^2 + |\Delta X_t|^2 + |\Delta \hat{X}_t - \Delta X_t|^2) dt \right)^{p/2} \right] \\ &\leq C \mathbb{E} \left[ \int_0^T \left( |\Delta \alpha_t|^{2p} + |\Delta X_t|^{2p} + \sup_{t \in [0, T]} |\Delta \hat{X}_t - \Delta X_t|^p \right) dt \right]. \end{aligned}$$

By Grönwall inequality, we get

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |\Delta \hat{X}_t - \Delta X_t|^p \right] \leq C \exp(CT) \mathbb{E} \left[ \int_0^T (|\Delta \alpha_t|^{2p} + |\Delta X_t|^{2p}) dt \right]. \quad (\text{II.3.26})$$

Now we shall estimate the moments of  $\Delta X$ . Again, for all  $p \geq 2$ , by Burkholder-Davis-Gundy inequality and Jensen inequality, we have

$$\begin{aligned} \mathbb{E} \left[ \sup_{t \in [0, T]} |\Delta X_t|^p \right] &\leq C \mathbb{E} \left[ \int_0^T (|\Delta \alpha_t|^p + |\Delta X_t|^p) dt + \left( \int_0^T (|\Delta \alpha_t| + |\Delta X_t|)^2 dt \right)^{p/2} \right] \\ &\leq C \mathbb{E} \left[ \int_0^T \left( |\Delta \alpha_t|^p + \sup_{t \in [0, T]} |\Delta X_t|^p \right) dt \right]. \end{aligned}$$

By Grönwall inequality, we get

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |\Delta X_t|^p \right] \leq C \exp(CT) \mathbb{E} \left[ \int_0^T |\Delta \alpha_t|^p dt \right]. \quad (\text{II.3.27})$$

Finally, combing (II.3.26) and (II.3.27), we get

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |\Delta \hat{X}_t - \Delta X_t|^p \right] \leq C \mathbb{E} \left[ \int_0^T |\Delta \alpha_t|^{2p} dt \right],$$

where  $C$  is a constant depending on  $p$ ,  $T$  and the coefficients of the SDE (II.2.2). ■

We shall now give the proof of Theorem II.2.7 in the non-linear system case. Note that in the linear case,  $\Delta X$  and  $\Delta \hat{X}$  defined in the above lemma are the same, which means that there will not be any high order error terms in the variational calculus. Therefore, the estimation of the error terms in step 5 of the following proof will not be necessary for linear system.

**Proof of Theorem II.2.7** For simplicity, we shall only prove the case where  $n = d = 1$ . In higher dimensional case, the proof is essentially the same. Let  $\xi$  be an implementable contract such that there exists an  $\mathbb{F}$ -adapted control  $\alpha^*$  which optimizes the Agent's expected value. We denote the corresponding state variable  $X^*$ .

*Step 1.* Introduce the dynamic version of the value function  $Y_t := \mathbb{E}_t^{\alpha^*} [\xi - \int_t^T c_s(\alpha_s^*) ds]$ . One can check straightforward that  $(M_t)_{0 \leq t \leq T} := (Y_t - \int_0^t c_s(\alpha_s^*) ds)_{0 \leq t \leq T}$  is an  $\mathbb{F}$ -martingale under the probability  $\mathbb{P}^{\alpha^*}$ . By the martingale representation in Lemma II.3.1, there exists an  $\mathbb{F}$ -adapted process  $Z$ , such that for all  $t \in [0, T]$

$$M_t = J_A(\alpha^*) + \int_0^t Z_s dW_s^{\alpha^*}. \quad (\text{II.3.28})$$

By Assumption II.2.5 and [41, Theorem 9.3.5],  $Z \in \mathbb{H}_T^2(\mathbb{P}^{\alpha^*})$ .

*Step 2.* Next we perform a variational calculus around the optimal control  $\alpha^*$ . Let  $\alpha \in \mathcal{U}$  and for all  $t \in [0, T]$  denote  $\Delta\alpha_t = \alpha_t - \alpha_t^*$ . Define  $\alpha_t^\varepsilon := \alpha_t^* + \varepsilon\Delta\alpha_t$  and denote the corresponding state variable by  $X^\varepsilon$ . Since the control set  $\mathcal{U}$  is assumed to be convex, we have  $\alpha^\varepsilon \in \mathcal{U}$ . Denote  $\delta\alpha_t = \varepsilon\Delta\alpha_t$ ,  $\Delta X_t = \frac{X_t^\varepsilon - X_t^*}{\varepsilon}$  and  $\delta X_t = \varepsilon\Delta X_t$ . As  $\xi = Y_T$ , we have

$$\begin{aligned} J_A(\alpha^\varepsilon) &= \mathbb{E}^{\alpha^\varepsilon} \left[ \xi - \int_0^T c_t(\alpha_t^\varepsilon) dt \right] \\ &= \mathbb{E}^{\alpha^\varepsilon} \left[ J_A(\alpha^*) + \int_0^T c_t(\alpha_t^*) dt + \int_0^T Z_t dW_t^{\alpha^*} - \int_0^T c_t(\alpha_t^\varepsilon) dt \right] \\ &= \mathbb{E}^{\alpha^\varepsilon} \left[ J_A(\alpha^*) + \int_0^T -\delta c_t dt + \int_0^T Z_t (dW_t^{\alpha^\varepsilon} + (b_t(X_t^\varepsilon, \alpha_t^\varepsilon) - b_t(X_t^*, \alpha_t^*)) dt) \right] \end{aligned}$$

where  $\delta c_t := c_t(\alpha_t^\varepsilon) - c_t(\alpha_t^*)$ . By the definition of the set of admissible contracts (II.2.6), we know that  $\int_0^T Z_t dW_t^{\alpha^\varepsilon}$  is a true martingale under  $\mathbb{P}^{\alpha^\varepsilon}$ . Denote  $\mathcal{E}^\varepsilon$  the Radon-Nikodym derivative between  $\mathbb{P}^{\alpha^\varepsilon}$  and  $\mathbb{P}^{\alpha^*}$ , namely

$$\mathcal{E}^\varepsilon := \exp \left( \int_0^T (b_t(X_t^\varepsilon, \alpha_t^\varepsilon) - b_t(X_t^*, \alpha_t^*)) dW_t^{\alpha^*} - \frac{1}{2} \int_0^T (b_t(X_t^\varepsilon, \alpha_t^\varepsilon) - b_t(X_t^*, \alpha_t^*))^2 dt \right).$$

Note that since  $b$  is bounded,  $\mathcal{E}^\varepsilon$  is in  $L^p$  for all  $p \geq 1$ . Applying Girsanov theorem and together with Taylor expansion on  $b$ , we have

$$\begin{aligned} &J_A(\alpha^\varepsilon) - J_A(\alpha^*) \\ &= \mathbb{E}^{\alpha^*} \left[ \mathcal{E}^\varepsilon \int_0^T \left( -\delta c_t + Z_t (b_t(X_t^\varepsilon, \alpha_t^\varepsilon) - b_t(X_t^*, \alpha_t^*)) \right) dt \right] \\ &= \varepsilon \mathbb{E}^{\alpha^*} \left[ \mathcal{E}^\varepsilon \int_0^T \left( -\partial_a c_t(\alpha_t^*) \Delta\alpha_t + Z_t (\partial_a b_t(X_t^*, \alpha_t^*) \Delta\alpha_t + \partial_x b_t(X_t^*, \alpha_t^*) \Delta\hat{X}_t) + \mathcal{R}_t^\varepsilon / \varepsilon \right) dt \right], \end{aligned}$$

where  $\Delta\hat{X}$  is the solution of the SDE (II.3.25) with initial condition  $\Delta\hat{X}_0 = 0$  and

$$\begin{aligned} \mathcal{R}_t^\varepsilon := & -\left(\delta c_t - \partial_a c_t(\alpha_t^*)\delta\alpha_t\right) + Z_t\left(b_t(X_t^\varepsilon, \alpha^\varepsilon) - b_t(X_t^*, \alpha_t^*) - \partial_a b_t(X_t^*, \alpha_t^*)\delta\alpha_t + \partial_x b_t(X_t^*, \alpha_t^*)\delta X_t\right) \\ & + \varepsilon Z_t \partial_x b_t(X_t^*, \alpha_t^*)(\Delta X_t - \Delta\hat{X}_t). \end{aligned} \quad (\text{II.3.29})$$

We claim and will prove in *Step 5* that  $\limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^2} \mathbb{E}^{\alpha^*} \left[ \mathcal{E}^\varepsilon \int_0^T |\mathcal{R}_t^\varepsilon| dt \right] < \infty$ . Then,

$$\frac{J_A(\alpha^\varepsilon) - J_A(\alpha^*)}{\varepsilon} = \mathbb{E}^{\alpha^*} \left[ \mathcal{E}^\varepsilon \int_0^T \left( -\partial_a c_t(\alpha_t^*)\Delta\alpha_t + Z_t(\partial_a b_t(X_t^*, \alpha_t^*)\Delta\alpha_t + \partial_x b_t(X_t^*, \alpha_t^*)\Delta\hat{X}_t) \right) dt \right] + \mathcal{O}(\varepsilon). \quad (\text{II.3.30})$$

*Step 3.* In order to continue the variational calculus above, we need to introduce the adjoint process  $P$  and  $Q$  satisfying (II.2.17). Let  $(P, Q)$  be the unique strong solution in  $\mathbb{H}^2(\mathbb{P}^{\alpha^*})$  of the following BSDE :

$$dP_t = -\partial_x H_t(X_t^*, X_t^*, \alpha_t^*, \alpha_t^*, P_t, Q_t, Z_t)dt + Q_t dW_t^{\alpha^*}, \quad P_T = 0,$$

where  $H$  is the Hamiltonian given in (II.2.15). Note that the above BSDE is a linear BSDE and since  $P_T = 0$ ,  $P$  and  $Q$  are in  $\mathbb{H}^p(\mathbb{P}^{\alpha^*})$  for all  $p \geq 2$ .

*Step 4.* We are going to continue the variational calculus above and obtain (II.2.20). By Itô's formula, we have

$$\begin{aligned} d(P_t \Delta\hat{X}_t) = & \left( -Z_t \partial_x b_t(X_t^*, \alpha_t^*)\Delta\hat{X}_t + \Delta\alpha_t \left( P_t(\partial_a b_t + b_t \partial_a \sigma_t)(X_t^*, \alpha_t^*) + Q_t \partial_a \sigma_t(X_t^*, \alpha_t^*) \right) \right) dt \\ & + \left( Q_t \Delta\hat{X}_t + P_t(\partial_x \sigma_t(X_t^*, \alpha_t^*)\Delta\hat{X}_t + \partial_a \sigma_t(X_t^*, \alpha_t^*)\Delta\alpha_t) \right) dW_t^{\alpha^*}. \end{aligned}$$

Using the fact that  $P_T = 0$  and  $\Delta\hat{X}_0 = 0$ , we get

$$\begin{aligned} 0 = & \int_0^T \left( -Z_t b_t(X_t^*, \alpha_t^*)\Delta\hat{X}_t + \Delta\alpha_t \left( P_t(\partial_a b_t + b_t \partial_a \sigma_t)(X_t^*, \alpha_t^*) + Q_t \partial_a \sigma_t(X_t^*, \alpha_t^*) \right) \right) dt \\ & + \int_0^T \left( Q_t \Delta\hat{X}_t + P_t(\partial_x \sigma_t(X_t^*, \alpha_t^*)\Delta\hat{X}_t + \partial_a \sigma_t(X_t^*, \alpha_t^*)\Delta\alpha_t) \right) dW_t^{\alpha^*}. \end{aligned} \quad (\text{II.3.31})$$

Note that

$$\begin{aligned} & \int_0^T \left( Q_t \Delta\hat{X}_t + P_t(\partial_x \sigma_t(X_t^*, \alpha_t^*)\Delta\hat{X}_t + \partial_a \sigma_t(X_t^*, \alpha_t^*)\Delta\alpha_t) \right) dW_t^{\alpha^*} \\ = & \int_0^T \left( Q_t \Delta\hat{X}_t + P_t(\partial_x \sigma_t(X_t^*, \alpha_t^*)\Delta\hat{X}_t + \partial_a \sigma_t(X_t^*, \alpha_t^*)\Delta\alpha_t) \right) (b_t(X_t^\varepsilon, \alpha_t^\varepsilon) - b_t(X_t^*, \alpha_t^*)) dt \\ & + \int_0^T \left( Q_t \Delta\hat{X}_t + P_t(\partial_x \sigma_t(X_t^*, \alpha_t^*)\Delta\hat{X}_t + \partial_a \sigma_t(X_t^*, \alpha_t^*)\Delta\alpha_t) \right) dW_t^{\alpha^\varepsilon}. \end{aligned} \quad (\text{II.3.32})$$

Denote

$$\hat{\mathcal{R}}_t^\varepsilon := \left( Q_t \Delta \hat{X}_t + P_t (\partial_x \sigma_t(X_t^*, \alpha_t^*) \Delta \hat{X}_t + \partial_a \sigma_t(X_t^*, \alpha_t^*) \Delta \alpha_t) \right) (b_t(X_t^\varepsilon, \alpha_t^\varepsilon) - b_t(X_t^*, \alpha_t^*)). \quad (\text{II.3.33})$$

We claim and will prove in *Step 5* that  $\limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E}^{\alpha^*} \left[ \mathcal{E}^\varepsilon \int_0^T |\hat{\mathcal{R}}_t^\varepsilon| dt \right] < \infty$ .

Inserting the equation (II.3.31) and equation (II.3.32) into (II.3.30), we get

$$\begin{aligned} \frac{J_A(\alpha^\varepsilon) - J_A(\alpha^*)}{\varepsilon} &= \mathbb{E}^{\alpha^*} \left[ \mathcal{E}^\varepsilon \int_0^T \left( -\partial_a c_t(\alpha_t^*) \Delta \alpha_t + Z_t \partial_a b_t(X_t^*, \alpha_t^*) \Delta \alpha_t \right. \right. \\ &\quad \left. \left. + \Delta \alpha_t \left( P_t (\partial_a b_t + b_t \partial_a \sigma_t)(X_t^*, \alpha_t^*) + Q_t \partial_a \sigma_t(X_t^*, \alpha_t^*) \right) \right) dt \right] + \mathcal{O}(\varepsilon). \end{aligned}$$

Now, let  $\varepsilon$  goes to 0 and by dominated convergence theorem, we obtain

$$\begin{aligned} 0 \geq \lim_{\varepsilon \rightarrow 0} \left( \frac{J_A(\alpha^\varepsilon) - J_A(\alpha^*)}{\varepsilon} \right) &= \mathbb{E}^{\alpha^*} \left[ \int_0^T \Delta \alpha_t \left( -\partial_a c_t(\alpha_t^*) + Z_t \partial_a b_t(X_t^*, \alpha_t^*) \right. \right. \\ &\quad \left. \left. + P_t (\partial_a b_t + b_t \partial_a \sigma_t)(X_t^*, \alpha_t^*) + Q_t \partial_a \sigma_t(X_t^*, \alpha_t^*) \right) dt \right] \\ &= \mathbb{E}^{\alpha^*} \left[ \int_0^T \partial_a H_t(X_t^*, X_t^*, \alpha_t^*, \alpha_t^*, P_t, Q_t, Z_t) \Delta \alpha_t dt \right] \end{aligned}$$

Since  $\Delta \alpha$  is arbitrary, if  $\alpha^*$  is an optimal control, (II.2.20) holds true.

*Step 5.* Finally we shall complete the proof by proving the claim we made on  $(\mathcal{R}_t^\varepsilon)_{t \in [0, T]}$  and  $(\hat{\mathcal{R}}_t^\varepsilon)_{t \in [0, T]}$ . The first term in (II.3.29) is clearly uniformly bounded by  $C\varepsilon^2$  for some positive constant  $C$ , since the controls are bounded.

By Hölder's inequality, we have

$$\begin{aligned} &\mathbb{E}^{\alpha^*} \left[ \mathcal{E}^\varepsilon \int_0^T Z_t \left( b_t(X_t^\varepsilon, \alpha_t^\varepsilon) - b_t(X_t^*, \alpha_t^*) - \partial_a b_t(X_t^*, \alpha_t^*) \delta \alpha_t + \partial_x b_t(X_t^*, \alpha_t^*) \delta X_t \right) dt \right] \\ &\leq \mathbb{E}^{\alpha^*} \left[ \mathcal{E}^\varepsilon \left( \int_0^T Z_t^2 dt \right)^{\frac{1}{2}} \left( \int_0^T \left( b_t(X_t^\varepsilon, \alpha_t^\varepsilon) - b_t(X_t^*, \alpha_t^*) - \partial_a b_t(X_t^*, \alpha_t^*) \delta \alpha_t + \partial_x b_t(X_t^*, \alpha_t^*) \delta X_t \right)^2 dt \right)^{\frac{1}{2}} \right] \\ &\leq \|\mathcal{E}^\varepsilon\|_p \|Z\|_{\mathbb{H}_T^2(\mathbb{P}^{\alpha^*})} \mathbb{E}^{\alpha^*} \left[ \left( \int_0^T \left( b_t(X_t^\varepsilon, \alpha_t^\varepsilon) - b_t(X_t^*, \alpha_t^*) - \partial_a b_t(X_t^*, \alpha_t^*) \delta \alpha_t + \partial_x b_t(X_t^*, \alpha_t^*) \delta X_t \right)^2 dt \right)^{\frac{q}{2}} \right]^{\frac{1}{q}} \\ &\leq CT \sup_{\varepsilon} \|\mathcal{E}^\varepsilon\|_p \|Z\|_{\mathbb{H}_T^2(\mathbb{P}^{\alpha^*})} \mathbb{E}^{\alpha^*} \left[ \left( \int_0^T \left( \Delta \alpha_t^2 + \Delta X_t^2 \right) dt \right)^{\frac{q}{2}} \right]^{\frac{1}{q}} \varepsilon^2, \end{aligned}$$

where  $\frac{1}{p} + \frac{1}{q} = \frac{1}{2}$  with  $p$  given in Assumption II.2.6. As shown in the proof of Lemma II.3.2,  $\Delta X \in \mathbb{H}_T^p(\mathbb{P}^{\alpha^*})$  for all  $p \geq 2$ , therefore, the second term of the right hand side of (II.3.29) is also bounded by  $C\varepsilon^2$  for some positive constant  $C$ . As for the third term in (II.3.29), again by Lemma II.3.2, using the definition



$\varepsilon \Delta X_t = X_t^\varepsilon - X_t^*$ , we have

$$\begin{aligned} \mathbb{E}^{\alpha^\varepsilon} \left[ \int_0^T \varepsilon |Z_t \partial_x b_t(X_t^*, \alpha_t^*)| (\Delta X_t - \Delta \hat{X}_t) dt \right] &\leq C \mathbb{E} \left[ \hat{\mathcal{E}}^\varepsilon \int_0^T \varepsilon |Z_t (\Delta X_t - \Delta \hat{X}_t)| dt \right] \\ &\leq C \|\hat{\mathcal{E}}^\varepsilon\|_p \|Z\|_{\mathbb{H}_T^2(\mathbb{P})} \mathbb{E} \left[ \sup_{t \in [0, T]} |\varepsilon \Delta X_t - \varepsilon \Delta \hat{X}_t|^q \right]^{\frac{1}{q}} \\ &\leq C \sup_\varepsilon \|\hat{\mathcal{E}}^\varepsilon\|_p \|Z\|_{\mathbb{H}_T^2(\mathbb{P})} \mathbb{E} \left[ \sup_{t \in [0, T]} |\Delta \alpha_t|^{2q} \right]^{\frac{1}{q}} \varepsilon^2, \end{aligned}$$

where  $\frac{1}{p} + \frac{1}{q} = \frac{1}{2}$  with  $p$  given in Assumption II.2.6 and  $\hat{\mathcal{E}}^\varepsilon$  is the Radon-Nikodym derivative between  $\mathbb{P}^{\alpha^\varepsilon}$  and the Wiener measure  $\mathbb{P}$ , namely

$$\hat{\mathcal{E}}^\varepsilon := \exp \left( \int_0^T b_t(X_t^\varepsilon, \alpha_t^\varepsilon) dB_t - \frac{1}{2} \int_0^T b_t^2(X_t^\varepsilon, \alpha_t^\varepsilon) dt \right),$$

which has finite moment uniformly in  $\varepsilon$  by Assumption II.2.6. Finally to complete our proof, we shall prove  $\limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E}^{\alpha^*} \left[ \mathcal{E}^\varepsilon \int_0^T |\hat{\mathcal{R}}_t^\varepsilon| dt \right] < \infty$  where  $\hat{\mathcal{R}}_t^\varepsilon$  is defined by the equation (II.3.33). Note that  $\alpha^\varepsilon - \alpha^* = \varepsilon \Delta \alpha$  and as shown in the proof of Lemma II.3.2,

$$\mathbb{E}^{\alpha^*} \left[ \int_0^T |X_t^\varepsilon - X_t^*|^p dt \right] \leq C \mathbb{E} \left[ \sup_{t \in [0, T]} |X_t^\varepsilon - X_t^*|^p \right] \leq C \mathbb{E} \left[ \int_0^T |\Delta \alpha_t|^p dt \right] \varepsilon^p$$

for some constants  $C$ . Therefore, by Assumption II.2.5, using Cauchy-Schwarz inequality and Hölder inequality, we have

$$\begin{aligned} &\mathbb{E}^{\alpha^*} \left[ \mathcal{E}^\varepsilon \int_0^T |\hat{\mathcal{R}}_t^\varepsilon| dt \right] \\ &\leq C \mathbb{E}^{\alpha^*} \left[ \int_0^T (|P_t|^4 + |Q_t|^4 + |\Delta X_t|^4 + |\Delta \alpha_t|^4) dt \right]^{\frac{1}{2}} \mathbb{E}^{\alpha^*} \left[ \left( \int_0^T (|\Delta X_t|^2 + |\Delta \alpha_t|^2) dt \right)^2 \right]^{\frac{1}{4}} \varepsilon, \end{aligned}$$

for some constant  $C$ . Therefore,  $\limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \mathbb{E}^{\alpha^*} \left[ \mathcal{E}^\varepsilon \int_0^T |\hat{\mathcal{R}}_t^\varepsilon| dt \right] < \infty$ . ■

**Proof of Theorem II.2.9** Let  $\alpha \in \mathcal{U}$  and denote the corresponding state variable by  $X$ . Denote  $\Delta \alpha := \alpha - \alpha^*$  and  $\Delta X = X - X^*$ . By the same computation as in the proof of Theorem II.2.7 and the

concavity of  $\mathcal{G}$ , we have

$$\begin{aligned}
 & J_A(\alpha) - J_A(\alpha^*) \\
 &= \mathbb{E}^\alpha \left[ \int_0^T \left( -\delta c_t + Z_t(b_t(X_t, \alpha_t) - b_t(X_t^*, \alpha_t^*)) \right) dt + P_T \Delta X \right] \\
 &= \mathbb{E}^\alpha \left[ \int_0^T \left( \mathcal{G}_t(X_t, X_t^*, \alpha_t, \alpha_t^*, P_t, Q_t, Z_t) - \mathcal{G}_t(X_t, X_t^*, \alpha_t, \alpha_t^*, P_t, Q_t, Z_t) \right. \right. \\
 &\quad \left. \left. - \partial_x \mathcal{G}_t(X_t^*, X_t^*, \alpha_t^*, \alpha_t^*, P_t, Q_t, Z_t) \Delta X_t \right) dt \right] \leq 0.
 \end{aligned}$$

■

**Proof of Theorem II.2.10** Since the enlarged Principal's problem has a larger admissible contract set, the Principal's value is higher than the initial problem of the Principal. If  $\bar{V}_P$  has an optimal solution  $\xi^*$  such that the solution  $Q^*$  of the FBSDE (II.2.16)-(II.2.19) satisfies the sufficient condition given by Theorem II.2.9, then we know that the control  $\alpha^*$  defined by (II.2.20) is indeed the optimal control for the Agent's Problem given this contract  $\xi^*$ . Therefore,  $\xi^* \in \Xi$  and consequently  $V_P \geq \bar{V}_P$ . Combined with the first statement of the theorem, we get  $V_P = \bar{V}_P$ . ■

## II.4 Applications

We shall make the following assumption throughout this section.

**Assumption II.4.1.** *The Agent's cost function  $(t, b) \mapsto c(t, b)$  is deterministic, strictly convex, increasing and twice differentiable in  $b$ .*

### II.4.1 Principal-Agent Problem in Partially Observed Linear System

We return to the context in Section II.2.3 and aim at solving the Principal-Agent problem (II.2.13)-(II.2.14). In addition, we assume that the unobservable process  $\hat{X}$  is non-affected by the Agent's control  $\alpha$ . We recall that the controlled state variables follow (II.2.9)-(II.2.10) with  $\alpha \equiv 0$  and that the associated Hamiltonian function is given by

$$H(t, x, b, p, q, z) := \left( \eta(t)x - h(t)V(t)(h(t)x + b) \right) p - \left( h(t)x + b \right) z + c(t, b).$$

The necessary condition by Theorem II.2.7 becomes

$$\begin{cases} dY_t = c(t, \beta_t^*) dt + Z_t dI_t & Y_T = \xi, \end{cases} \quad (II.4.34)$$

$$\begin{cases} dP_t = \left( h(t)Z_t - (\eta(t) - V(t)h^2(t))P_t \right) dt + Q_t dI_t & P_T = 0, \end{cases} \quad (II.4.35)$$

$$\begin{cases} dX_t = \eta(t)X_t dt + h(t)V(t)dI_t & X_0 = m_0, \end{cases} \quad (II.4.36)$$

where

$$\beta_t^* = \operatorname{argmin}_{b \in A} H(t, X_t, b, P_t, Q_t, Z_t). \quad (\text{II.4.37})$$

The above system describes the contracts under the optimal response of the Agent. The sufficient condition can also be re-written under simpler form :

**Theorem II.4.2.** *Let  $(Y^*, Z^*, P^*, Q^*, X^*)$  be a solution to the system (II.4.34)-(II.4.35), where  $\beta^*$  satisfies (II.4.37). Define  $\xi := Y_T^*$ . Then  $\beta^*$  is an optimal control if*

$$\frac{Q_t^*}{2h(t)} \in \left[ 0, \inf_{b \in A} \left\{ \partial_{bb}^2 c(t, b) \right\} \right] \quad \text{for all } t \in [0, T]. \quad (\text{II.4.38})$$

In particular, in this case  $\xi \in \Xi$ .

**Remark II.4.3.** *We can also obtain the necessary and sufficient condition in the case of exponential utility Agent, the proof and the hypothesis are slightly different, we shall detail them in the Appendix.*

Together with Theorem II.2.10, we can solve the Principal's problem explicitly.

**Theorem II.4.4.** *Let Assumption II.4.1 hold true. Assume in addition that the function  $b \in B \mapsto b - c(t, b)$  has a maximiser, denoted by  $\bar{\beta}^*(t)$ . Define*

$$Z_t^* = \partial_b c(t, \bar{\beta}^*(t)) + \int_t^T e^{\int_t^s \eta(r) dr} h(s) \partial_b c(s, \bar{\beta}^*(s)) ds.$$

The optimal contract for the Principal's problem is given by

$$\xi^* = R + \int_0^T (c(t, \bar{\beta}^*(t)) - Z_t^*(h(t)\bar{X}_t^* + \bar{\beta}^*(t))) dt + \int_0^T Z_t^* dB_t, \quad (\text{II.4.39})$$

where  $\bar{X}_t^*$  is the unique solution to (II.4.36) with innovation process  $I$  computed with  $\bar{\beta}^*$  and the Principal's optimal expected utility is

$$V_P = -R + \int_0^T (\bar{\beta}^*(t) - c(t, \bar{\beta}^*(t))) dt + \int_0^T h(t)m(t)dt,$$

where  $m(t) := m_0 e^{\int_0^t \eta(s) ds}$  for  $t \in [0, T]$ .

**Remark II.4.5.** *It is immediately verified that the optimal contract (II.4.39) is indeed implementable as it depends only on the observable of  $B$ . In contrast with the classical Principal-Agent literature, the filtered process  $X$  is not observable by the Principal and cannot be contracted upon. In fact, in order to compute the filtered process  $X$ , one needs to observe the Agent's control  $\beta$  to compute the innovation process  $I$  given by (II.2.12) before solving the equation (II.4.36).*

*Démonstration.* Note that for any  $Z \in \mathbb{H}_T^2$ ,

$$\begin{aligned}
 \bar{J}_P(Y_0, Z) &= \mathbb{E}^{\beta^*} [(B_T - Y_T)] = \mathbb{E}^{\beta^*} \left[ \left( \int_0^T (h(t)X_t + \beta_t^*)dt - \int_0^T c(t, \beta_t^*)dt \right) \right] - Y_0 \\
 &= \mathbb{E}^{\beta^*} \left[ \int_0^T (\beta_t^* - c(t, \beta_t^*)) dt \right] - Y_0 + \int_0^T h(t)m(t)dt, \\
 &\leq \mathbb{E}^{\beta^*} \left[ \int_0^T (\bar{\beta}_t^* - c(t, \bar{\beta}^*(t))) dt \right] - R + \int_0^T h(t)m(t)dt.
 \end{aligned} \tag{II.4.40}$$

We are going to show that  $\bar{\beta}^*$  is Agent's optimal control associated to a contract  $\xi^* \in \underline{\Xi}$ . In order for  $\bar{\beta}^*$  to be an optimal control, it follows from (II.4.37) that the processes  $Z^*$  and  $P^*$  in the forward-backward system must satisfy

$$Z_t^* = \partial_b c(t, \bar{\beta}^*(t)) - V(h)h(t)P_t^*. \tag{II.4.41}$$

Inserting the equality above into (II.4.35), we obtain

$$dP_t = (h(t)\partial_b c(t, \bar{\beta}^*(t)) - \eta(t)P_t)dt + Q_t dI^{\bar{\beta}^*}, \quad P_T = 0.$$

This BSDE does have a unique solution given by

$$P_t^* = - \int_t^T e^{\int_t^s \eta(r)dr} h(s)\partial_b c(s, \bar{\beta}^*(s))ds, \quad Q_t^* = 0, \quad t \in [0, T].$$

Note that  $Q$  satisfies (II.4.38). Again by (II.4.41), we have

$$Z_t^* = \partial_b c(t, \bar{\beta}^*(t)) + \int_t^T e^{\int_t^s \eta(r)dr} h(s)\partial_b c(s, \bar{\beta}^*(s))ds,$$

which is a deterministic process. Further, there is a unique  $\mathbb{F}^B$ -adapted solution  $\bar{X}^*$  to the SDE (II.2.9) with  $\alpha \equiv 0$ . Using the process  $Z^*$  and  $\bar{X}^*$ , we immediately find an  $\mathbb{F}^B$ -adapted process  $Y^*$  satisfying (II.4.34). So far, we find an  $\mathbb{F}^B$ -adapted solution  $(Y^*, Z^*, P^*, Q^*, X^*)$  to the forward-backward system. Let  $\xi^*$  be the contract defined by  $\xi^* := Y_T^*$ , in other words,  $\xi^*$  satisfies (II.4.39). Clearly, we have  $\xi^* \in \underline{\Xi}$  and  $\bar{\beta}^*$  is the corresponding Agent's optimal control. Therefore, by (II.4.40) and  $\underline{\Xi} \subset \Xi \subset \bar{\Xi}$  we obtain  $\bar{V}_P = V_P$ . ■

## II.4.2 Mean Field Interacting Agents in Partially Observed Linear System

In our second application, we consider the Principal-Agent problem in the same partially observed system but with  $N$  interacting Agents. As in Section II.2.3, we denote  $\hat{X}$  as the unobservable process and  $B$  as the observable, both assumed to be 1-dimensional for simplicity. Agents  $i = 1, \dots, n$  have

private state processes  $(\hat{X}^i, B^i)$  with dynamics

$$dB_t^i = (h(t)\bar{\lambda}_t^N \hat{X}_t^i + \beta_t^i)dt + dW_t^{\beta^i}, \quad B_0 = 0,$$

where  $\bar{\lambda}_t^N = \frac{1}{N} \sum_1^N B_t^i$  is the mean value of the observables at time  $t$  and  $\hat{X}$  is the unique strong solution of

$$d\hat{X}_t^i = \eta(t)\hat{X}_t^i dt + \sigma(t)dW_t^i, \quad \hat{X}_0 = \mu^i.$$

Here,  $W^1, \dots, W^N, W^{\beta^1}, \dots, W^{\beta^N}$  are independent Brownian motions, and  $\beta^i$  is the control chosen by the Agent  $i$ . We assume that  $\mu^1, \dots, \mu^N$  are independent Gaussian variables with mean  $m_0$  and variance  $V_0$ .

The Principal proposes the same contract  $\xi$  to all the Agents. Then the interacting agents agree on a Nash equilibrium in which each Agent cannot improve his utility by deviating unilaterally.

The problem becomes very difficult in general because of the coupling. Note that the observable process  $B^i$  for Agent  $i$  is no longer Gaussian and the filter of  $X^i$ , namely the conditional law of  $X^i$  given  $B^i$ , is no longer described by a finite-dimensional process.

However, as the number of Agents  $N \rightarrow \infty$ , we expect the system to become decoupled, in which case we may be able to learn something about the Nash equilibrium from the corresponding mean field game (MFG).

#### II.4.2.1 The Setting of Mean Field Interacting Agent Problem

As before, denote  $\hat{X}$  the unique strong solution of the following linear stochastic differential equation, which represents the unobservable part of the system.

$$d\hat{X}_t = \eta(t)\hat{X}_t dt + \sigma(t)dW_t, \quad X_0 = \mu.$$

Let  $\lambda \in C([0, T], \mathbb{R})$  represent the observed population mean, for any  $\beta \in \mathcal{U}$ , define

$$\frac{d\mathbb{P}^\beta}{d\mathbb{P}} \Big|_{\mathcal{F}_T} = \exp \left( \int_0^T (h(t)\lambda_t \hat{X}_t + \beta_t) dB_t + \frac{1}{2} \int_0^T (h(t)\lambda_t \hat{X}_t + \beta_t)^2 dt \right).$$

By Girsanov theorem, the canonical process  $B$  satisfies the following stochastic differential equation under  $\mathbb{P}^\beta$  :

$$dB_t = (h(t)\lambda_t \hat{X}_t + \beta_t)dt + dW_t^\beta, \quad B_0 = 0, \tag{II.4.42}$$

where  $W_t^\beta$  is a Brownian motion under  $\mathbb{P}^\beta$ . Each Agent receives the payment  $\xi$  at the maturity  $T$  of the

contract and aims at optimizing his utility function :

$$V_A(\xi, \lambda) = \sup_{\mathbb{P}^\beta \in \mathcal{P}(\lambda)} \mathbb{E}^\beta \left[ \xi - \frac{1}{2} \int_0^T c(t, \beta_t) dt \right].$$

Denote  $\mathcal{M}^*(\xi, \lambda)$  the set of optimal controls of the Agent's problem given the contract  $\xi$  and the observed population mean  $\lambda$ . Define

$$\Phi(\xi, \lambda) := \{ \tilde{\lambda} : \tilde{\lambda}_t = \mathbb{E}^{\beta^*} [B_t] \text{ for all } t \in [0, T] \text{ and } \beta^* \in \mathcal{M}^*(\xi, \lambda) \}.$$

A mean field equilibrium is a fixed point of the map  $\lambda \mapsto \Phi(\xi, \lambda)$ . Denote  $\Theta(\xi)$  the set of equilibria among the Agents when given the contract  $\xi$  satisfying the participation constraint, namely the Agent's value should be above the reservation utility  $R$ . The Principal's optimization problem is given by

$$V_P = \sup_{\xi \in \Xi} \sup_{\lambda \in \Theta(\xi)} \sup_{\beta^* \in \mathcal{M}^*(\xi, \lambda)} \mathbb{E}^{\beta^*} [B_T - \xi]$$

with the convention  $\sup \emptyset = -\infty$ .

**Remark II.4.6.** *As discussed in Elie, Mastrolia, Possamai [54], we are only going to consider contracts such that the Principal is able to compute the reaction of the Agents. In other words, the set of contracts given which there is at least one Nash equilibrium among the Agents. Indeed, the Principal needs to be able to anticipate how the Agents are going to react to the contracts that he may offer, and will therefore not offer contracts for which Agents cannot agree on an equilibrium.*

### II.4.2.2 The Agent's Problem

For any fixed  $\lambda \in C([0, T], \mathbb{R})$ , the observable process  $B$  in (II.4.42) is a Gaussian process and the Agent can compute the conditional mean and variance of the unobservable process  $(\hat{X}_t)_{t \geq 0}$  using the Kalman filter. Denote as before  $X_t := \mathbb{E}^\beta[\hat{X}_t | \mathcal{F}_t^B]$  and  $V_t = \mathbb{E}^\beta[(\hat{X}_t - X_t)^2 | \mathcal{F}_t^B]$ .

**Proposition II.4.7.** *Let  $\lambda \in C([0, T], \mathbb{R})$  be the observed population mean and  $\beta \in \mathcal{U}$ . We have the following filter system :*

$$\begin{cases} dX_t = \eta(t)X_t dt + h(t)V(t)\lambda_t dI_t^\beta & X_0 = m_0, \\ dV_t = (2\eta(t)V(t) - h^2(t)V^2(t)\lambda_t^2 + \sigma(t))dt & V_0 = V_0, \\ dB_t = (h(t)\lambda_t X_t + \beta_t)dt + dI_t^\beta & B_0 = 0, \end{cases} \quad (\text{II.4.43})$$

where  $I^\beta$  is the innovation process.

Using the filtered system, the Agent's problem becomes fully observable. We now recall the necessary and sufficient condition for the Agent's problem in this specific case in the two following propositions.

**Proposition II.4.8.** *Let  $\lambda \in C([0, T], \mathbb{R})$  be the observed population mean. For a contract  $\xi \in \Xi$ , let  $\beta^* \in \mathcal{M}^*(\lambda, \xi)$ . Then the following FBSDE :*

$$\begin{cases} dY_t = c(t, \beta_t^*)dt + Z_t dI_t^{\beta^*} & Y_T = \xi, \\ dP_t = \left( h(t)\lambda_t Z_t - (\eta(t) - V(t)h^2(t)\lambda_t^2)P_t \right)dt + Q_t dI_t^{\beta^*} & P_T = 0, \\ dX_t = \eta(t)X_t dt + h(t)V(t)\lambda_t dI_t^{\beta^*} & X_0 = m_0, \\ dB_t = (h(t)\lambda_t X_t + \beta_t^*)dt + dI_t^{\beta^*} & B_0 = 0, \end{cases} \quad (\text{II.4.44})$$

has a solution, denoted

$$(Y^*, Z^*, P^*, Q^*, X^*, B^*) \in \left( \mathbb{H}_T^2(\mathbb{P}^{\beta^*}) \right)^6.$$

Besides, for all  $t \in [0, T]$ , for all  $b \in A$ , the optimal control  $\beta^*$  must verify the local incentive constraint :

$$(Z_t^* + V(t)h(t)\lambda_t P_t^* - \beta_t^*)(b - \beta_t^*) \leq 0. \quad (\text{II.4.46})$$

**Proposition II.4.9.** *Let  $(Y^*, Z^*, P^*, Q^*, X^*, B^*)$  be a solution to the system (II.4.44)-(II.4.45) with  $\beta^*$  satisfying (II.4.46). Define  $\xi := Y_T^*$ . Then  $\beta^*$  is an optimal control if for all  $t \in [0, T]$ ,*

$$0 \leq \frac{Q_t^*}{2h(t)\lambda_t} \leq \inf_{b \in A} \left\{ \partial_{bb}^2 c(t, b) \right\}. \quad (\text{II.4.47})$$

### II.4.2.3 Solving the Principal's Problem

Define the following subsets of  $\mathbb{H}_T^2(\mathbb{P})$  :

$$\begin{aligned} \bar{\mathcal{Z}} &:= \left\{ Z : \exists \mathbb{P}^\beta \in \mathcal{P} \text{ s.t. (II.4.44)-(II.4.45) with } \lambda = \mathbb{E}[B] \text{ have a solution } (P, Q, X) \text{ in } \mathbb{H}_T^2(\mathbb{P}^\beta)^3 \right\}, \\ \underline{\mathcal{Z}} &:= \left\{ Z \in \bar{\mathcal{Z}} : Q \text{ satisfies (II.4.47)} \right\}, \end{aligned}$$

and define

$$\bar{\Xi} := \{ Y_T | Y_0 \geq R, Z \in \bar{\mathcal{Z}} \} \quad \text{and} \quad \underline{\Xi} := \{ Y_T | Y_0 \geq R, Z \in \underline{\mathcal{Z}} \},$$

where  $Y$  is defined by the equation (II.4.44). As before, we have  $\underline{\Xi} \subset \Xi \subset \bar{\Xi}$  and the enlarged optimization of the Principal is given by

$$\bar{V}_P := \sup_{Y_0 \geq R} \sup_{Z \in \bar{\mathcal{Z}}} \bar{J}_P(Y_0, Z) := \sup_{Y_0 \geq R} \sup_{Z \in \bar{\mathcal{Z}}} \mathbb{E}^{\beta^*} [B_T - Y_T]. \quad (\text{II.4.48})$$

Denote

$$k(t) := h(t)\mathbb{E}^{\beta^*} [X_t] = h(t)m_0 e^{\int_0^t \eta(s)ds} \quad \text{and} \quad \rho_t = e^{\int_t^T k(s)ds} - 1.$$

**Theorem II.4.10.** *Assume that Assumption II.4.1 holds true and assume that  $\partial_b c(t, \cdot)$  is surjective. Denote  $\bar{\beta}^*(t) = (\partial_b c(t, \cdot))^{-1}(\rho_t + 1)$ . There exists an optimal contract for the Principal's problem in the*

mean field setting, which is given by

$$\xi^* = \int_0^T c(t, \bar{\beta}^*(t)) dt + (\bar{\beta}^*(T) - V(T)h(T)\bar{B}_T P_T)(dB_T - h(T)\bar{B}_T \bar{X}_T dt),$$

where  $(\bar{X}, \bar{B}, V)$  are solutions of the following system of forward equations :

$$\begin{cases} d\bar{X}_t = \eta(t)\bar{X}_t dt + h(t)V(t)\bar{B}_t(dB_t - (h(t)\bar{B}_t \bar{X}_t + \bar{\beta}^*(t))dt) & \bar{X}_0 = m_0, \end{cases} \quad (\text{II.4.49})$$

$$\begin{cases} \frac{d\bar{B}_t}{dt} = k(t)\bar{B}_t + \bar{\beta}_t^* & \bar{B}_0 = 0, \end{cases} \quad (\text{II.4.50})$$

$$\begin{cases} \dot{V}(t) = 2\eta(t)V(t) - h^2(t)V^2(t)\bar{B}_t^2 + \sigma(t) & V_0 = V_0. \end{cases} \quad (\text{II.4.51})$$

*Démonstration.* For notation simplicity, denote

$$\bar{B}_t := \mathbb{E}^{\beta^*}[B_t], \quad \bar{\beta}_t = \mathbb{E}^{\beta^*}[\beta_t^*].$$

By (II.4.45), at equilibrium we have

$$\frac{d\bar{B}_t}{dt} = k(t)\bar{B}_t + \bar{\beta}_t. \quad (\text{II.4.52})$$

Replacing  $B_T$  and  $Y_T = \xi$  in (II.4.48) by the representations in (II.4.44) and (II.4.45), we get

$$\begin{aligned} \bar{V}_P &= -R + \sup_{Z \in \bar{\mathcal{Z}}} \mathbb{E}^{\beta^*} \left[ \int_0^T \left( h(t)\bar{B}_t X_t + \beta_t^* - c(t, \beta_t^*) \right) dt \right] \\ &\leq -R + \sup_{Z \in \bar{\mathcal{Z}}} \left\{ \int_0^T \left( k(t)\bar{B}_t + \bar{\beta}_t - c(t, \bar{\beta}_t) \right) dt \right\}. \end{aligned}$$

Notice that the inequality above becomes equality if and only if  $\beta^*$  is deterministic.

*Step 1.* Solve the deterministic control problem

$$\sup_{\bar{\beta}} \left\{ \int_0^T \left( k(t)\bar{B}_t + \bar{\beta}_t - c(t, \bar{\beta}_t) \right) dt \right\}, \quad (\text{II.4.53})$$

where  $\bar{B}$  satisfies (II.4.52). It follows from the Pontryagin maximum principle that the optimal control  $\bar{\beta}^*$  is :

$$\bar{\beta}^*(t) = (\partial_b c(t, \cdot))^{-1}(\rho_t + 1), \quad (\text{II.4.54})$$

where  $\rho$  is the adjoint process :

$$\frac{d\rho_t}{dt} = -k(t)(\rho_t + 1), \quad \rho_T = 0,$$

*Step 2.*  $\bar{V}_P$  is attained if the optimal control of the Agent is deterministic and is given by (II.4.54), now we need to check if such a contract exists and belongs to  $\bar{\Xi}$ .



Under the control  $\bar{\beta}^*(t)$ , the process  $\bar{B}$  follows :

$$\frac{d\bar{B}_t}{dt} = k(t)\bar{B}_t + (\partial_b c(t, \cdot))^{-1}(\rho_t + 1).$$

In order for  $\bar{\beta}^*(t)$  to be the optimal control of the agent, by (II.4.46) we must have

$$Z_t^* = (\partial_b c(t, \cdot))^{-1}(\rho_t + 1) - V(t)h(t)\bar{B}_t P_t.$$

Inserting the above equality into the equation of  $P$  and we get

$$dP_t = \left( h(t)\bar{B}_t\bar{\beta}^*(t) - \eta(t)P_t \right) dt + Q_t dI_t^{\beta^*},$$

and a trivial solution of the BSDE is given by  $(\bar{P}, 0)$ , where  $\bar{P}$  is the solution of the ODE :

$$\frac{d\bar{P}_t}{dt} = h(t)\bar{B}_t\bar{\beta}^*(t) - \eta(t)\bar{P}_t \quad \text{with} \quad \bar{P}_T = 0.$$

So far we find a solution to the system (II.4.44)-(II.4.45). Define the contract :

$$\xi := \int_0^T c(t, \bar{\beta}^*(t)) dt + (\bar{\beta}^*(t) - V(t)h(t)\bar{B}_t P_t)(dB_t - h(t)\bar{B}_t \bar{X}_t dt),$$

where  $\bar{X}$  is the solution of the following SDE :

$$dX_t = \eta(t)X_t dt + h(t)V(t)\bar{B}_t(dB_t - (h(t)\bar{B}_t X_t + \bar{\beta}^*(t))dt), \quad X_0 = m_0.$$

Clearly,  $\bar{\beta}^*(t)$  is the best effort of the agent under the contract  $\xi$ , and the sufficient condition for the Agent's problem is verified since  $Q = 0$ , i.e.  $\xi \in \Xi$ . We conclude by using Theorem II.2.10. ■ ■

## II.5 Appendix

In this section we shall detail the additional hypothesis and the proof of necessary condition for the Agent's problem in the case of exponential utility in the partially observed linear system. The controlled system is given by (II.2.9)-(II.2.10). The Agent's problem that we consider becomes

$$V_A := \sup_{\nu \in \mathcal{U}} J_A(\nu) := \sup_{\nu \in \mathcal{U}} \mathbb{E}^\nu \left[ -\exp \left( -\lambda \left( \xi - \int_0^T c_t(\nu_t) dt \right) \right) \right], \quad (\text{II.5.55})$$

To guarantee integrability, we shall need additional assumptions on the set of admissible contracts. For convenience, we say that a strictly positive local martingale  $(M_t)_{t \in [0, T]}$  satisfies the condition (H) if it is uniformly integrable and there exists  $p \in (1, \infty)$  and  $K_p > 0$  constant depending only on  $p$  such

that for every stopping time  $\tau$ ,

$$K_p \mathbb{E}[M_T^{1/p} | \mathcal{F}_\tau] \geq M_\tau^{1/p}. \quad (\text{H})$$

We recall that a contract  $\xi$  is implementable if there exists at least one optimal control for the Agent's problem. In the rest of the paper, we constrain our study on the following set of contracts :

$$\Xi := \left\{ \xi : V_A(\xi) \geq R \text{ and } \left( \mathbb{E}^{\nu^*}[\exp(-\lambda\xi) | \mathcal{F}_t^B] \right)_{t \in [0, T]} \text{ satisfied the condition (H)} \right\}, \quad (\text{II.5.56})$$

where  $\nu^* \in \mathcal{M}^*(\xi)$ .

**Remark II.5.1.** 1. Let  $M$  be a martingale and denote  $\mathcal{E}(M)$  the Doléan-Dade exponential. As given in the Corollary 3.4 in [95],  $\mathcal{E}(M)$  satisfying the condition (H) is equivalent to  $M \in BMO$ .

2. Let  $\xi \in \Xi$  and  $\nu^*$  be an optimal control of the Agent's problem. For any control  $\nu \in \mathcal{U}$ , denote

$$M_t := \mathbb{E}^{\nu^*} \left[ \exp \left( -\lambda \left( \xi - \int_0^T c_s(\nu_s) ds \right) \right) \middle| \mathcal{F}_t^B \right].$$

Since  $\nu$  is bounded, so is  $\exp(\lambda \int_0^T c_s(\nu_s) ds)$ . One can show easily that  $(M_t)_{t \in [0, T]}$  satisfies the condition (H).

Define the Hamiltonian function :

$$H_t(x, a, b, p, q, z) := \left( \eta(t)x + a - h(t)V(t)(h(t)x + b) \right) p - \left( h(t)x + b \right) z + c_t(a, b) + \lambda x q z.$$

**Remark II.5.2.** Compared to the definition of the Hamiltonian functional given in (II.2.15), the additional term  $\lambda x q z$  in the above definition is due to the exponential utility in the Agent's problem. If the Agent is risk-neutral, i.e.  $\lambda = 0$ , this term disappears.

Now we can give the necessary condition for the Agent's problem.

**Theorem II.5.3.** For a contract  $\xi \in \Xi$ , let  $\nu^* \in \mathcal{M}^*(\xi)$ . Then the following FBSDE

$$\begin{cases} dY_t = \left( c_t(\nu_t^*) + \frac{1}{2} \lambda Z_t^2 \right) dt + Z_t dI_t^{\nu^*} & Y_T = \xi, \end{cases} \quad (\text{II.5.57})$$

$$\begin{cases} dP_t = -\partial_x H_t(X_t, \nu_t^*, P_t, Q_t, Z_t) dt + Q_t dI_t^{\nu^*} & P_T = 0, \end{cases} \quad (\text{II.5.58})$$

$$\begin{cases} dX_t = \partial_p H_t(X_t, \nu_t^*, P_t, Q_t, Z_t) dt + h(t)V(t) dB_t & X_0 = m_0, \end{cases} \quad (\text{II.5.59})$$

with  $I^{\nu^*}$  defined as in (II.2.10),  $V$  defined as in (II.2.11) and

$$\nu_t^* = (\alpha_t^*, \beta_t^*) = \underset{(a, b) \in A}{\operatorname{argmin}} H_t(X_t, a, b, P_t, Q_t, Z_t),$$

has an  $\mathbb{F}$ -adapted solution, denoted by

$$(Y^*, Z^*, P^*, Q^*, X^*) \in \mathbb{H}^2(\mathbb{P}^{\nu^*}) \times BMO \times \mathbb{H}^2(\mathbb{P}^{\nu^*}) \times \mathbb{H}^2(\mathbb{P}^{\nu^*}) \times \mathbb{H}^2(\mathbb{P}^{\nu^*}).$$

**Lemma II.5.4.** *Let  $\mathbb{P}^{\nu_1}$  and  $\mathbb{P}^{\nu_2}$  be the two weak control of the Agent and  $Z$  a  $\mathbb{F}$ -adapted process. If  $\int_0^\cdot Z_t dI_t^{\nu_1} \in BMO(\mathbb{P}^{\nu_1})$ , then  $\int_0^\cdot Z_t dI_t^{\nu_2} \in BMO(\mathbb{P}^{\nu_2})$ .*

*Démonstration.* Denote  $X_1$  and  $X_2$  as the strong solution of the equation (II.2.9) given  $\nu_1 = (\alpha_1, \beta_1)$  and  $\nu_2 = (\alpha_2, \beta_2)$ , respectively. Denote  $\Delta X := X_2 - X_1$  and  $\Delta\beta = \beta_2 - \beta_1$ . The change of measure between  $\mathbb{P}^{\nu_1}$  and  $\mathbb{P}^{\nu_2}$  are given by

$$d\mathbb{P}^{\nu_2} = \mathcal{E} \left( \int_0^\cdot (h(s)\Delta\hat{X}_s + \Delta\beta_s) dI_s^{\nu_1} \right) d\mathbb{P}^{\nu_1}.$$

In particular, since the controls are bounded, the martingale  $(\int_0^t (h(s)\Delta\hat{X}_s + \Delta\beta_t) dI_s^{\nu_1})_{t \in [0, T]}$  is in BMO. We conclude by using the Theorem 3.8 in [95]. ■ ■

In light of Lemma II.5.4, without ambiguity, we say a  $\mathbb{F}^B$ -adapted process  $Z \in BMO$  if there exists  $\mathbb{P}^\nu \in \mathcal{P}$  such that  $\int_0^\cdot Z_t dI_t^\nu \in BMO(\mathbb{P}^\nu)$ .

Here we introduce a notation : for an  $\mathbb{F}^B$ -adapted process  $Z$  satisfying sufficient integrability condition, we denote

$$\mathcal{E}^\nu(Z) := \mathcal{E} \left( \int_0^\cdot Z_t dI_t^\nu \right).$$

**Proof of Theorem II.5.3** As in the proof of Theorem II.2.7, we start by assuming the existence of an  $\mathbb{F}^B$ -adapted control  $\nu^*$  which optimises the Agent's expected value under a given contract  $\xi$ . We denote the corresponding state variable  $\hat{X}^*$ .

*Step 1.* Introduce the dynamic version of the value function

$$\tilde{Y}_t := \mathbb{E}_t^{\nu^*} \left[ - \exp \left( -\lambda \left( \xi - \int_t^T c_s(\nu_s^*) ds \right) \right) \right].$$

One can check straightforward that  $(M_t)_{0 \leq t \leq T} := \left( \tilde{Y}_t \exp \left( \lambda \int_0^t c_s(\nu_s^*) ds \right) \right)_{0 \leq t \leq T}$  is an  $\mathbb{F}^B$ -martingale under the probability  $\mathbb{P}^{\nu^*}$ , and satisfies the condition (H) according to Remark II.5.1. By the martingale representation in Lemma II.3.1, there exists an  $\mathbb{F}^B$ -adapted process  $\tilde{Z}$ , such that for all  $t \in [0, T]$

$$M_t = V_A(\nu^*) + \int_0^t \tilde{Z}_s dI_s^{\nu^*} = V_A(\nu^*) + \int_0^t M_s Z_s dI_s^{\nu^*}, \quad \text{where } Z_t := -M_t^{-1} \tilde{Z}_t \quad (\text{II.5.60})$$

Let  $Y_t = -\frac{1}{\lambda} \ln(-\tilde{Y}_t)$ . Then (II.5.57) follows from Itô's formula. Further, in view of (II.5.60) we obtain

$$M_t = V_A(\nu^*) \mathcal{E}^{\nu^*}(-\lambda Z).$$

Since  $M$  satisfies the condition (H), we have  $Z \in BMO$  according to Remark II.5.1 and  $\mathcal{E}^{\nu^*}(-\lambda Z)$  is a uniformly integrable  $\mathbb{P}^{\nu^*}$ -martingale.

*Step 2.* Next we perform a variational calculus around the optimal control  $\nu^*$ . We shall use the same notation as in the proof of Theorem II.2.7. Note that  $\Delta\hat{X}$  satisfies the following stochastic differential equation :

$$d\Delta\hat{X}_t = \left[ (\eta(t) - V(t)h^2(t))\Delta\hat{X}_t + \Delta\alpha_t - V(t)h(t)\Delta\beta_t \right] dt.$$

Since  $J_A(\nu^*) = V_A = -e^{-\lambda Y_0}$ , we have

$$J_A(\nu^\epsilon) = \mathbb{E}^{\nu^\epsilon} \left[ J_A(\nu^*) \mathcal{E}^{\nu^\epsilon}(-\lambda Z) \exp \left( -\lambda \left( \int_0^T -\delta c_t(\nu_t^*) + \epsilon Z_t (h(t)\Delta\hat{X}_t + \Delta\beta_t) dt \right) \right) \right] \quad (\text{II.5.61})$$

Define  $d\bar{\mathbb{P}}^{\nu^\epsilon} := \mathcal{E}^{\nu^\epsilon}(-\lambda Z)d\mathbb{P}^{\nu^\epsilon}$  and denote  $\bar{\mathbb{E}}^{\nu^\epsilon}$  the expectation under  $\bar{\mathbb{P}}^{\nu^\epsilon}$ . Applying the Taylor expansion on the exponential function in (II.5.61), we have

$$\frac{J_A(\nu^\epsilon)}{J_A(\nu^*)} = \bar{\mathbb{E}}^{\nu^\epsilon} \left[ 1 - \epsilon \lambda \left( \int_0^T -\partial_\alpha c_t(\nu_t^*)\Delta\alpha_t - \partial_\beta c_t(\nu_t^*)\Delta\beta_t + Z_t (h(t)\Delta\hat{X}_t + \Delta\beta_t) dt \right) + \mathcal{R}_T^\epsilon \right],$$

where

$$\mathcal{R}_T^\epsilon := (\delta c_t(\nu_t^*) - \partial_\alpha c_t(\nu_t^*)\delta\alpha_t - \partial_\beta c_t(\nu_t^*)\delta\beta_t) + \sum_{n=2}^{\infty} \frac{(-1)^n \epsilon^n \lambda^n}{n!} \left( \int_0^T -\delta c_t(\nu_t^*) + Z_t (h(t)\Delta\hat{X}_t + \Delta\beta_t) dt \right)^n. \quad (\text{II.5.62})$$

Define  $d\bar{\mathbb{P}}^{\nu^*} := \mathcal{E}^{\nu^*}(-\lambda Z)d\mathbb{P}^{\nu^*}$ ,  $\bar{\mathbb{E}}^{\nu^*}$  the expectation under  $\bar{\mathbb{P}}^{\nu^*}$ , and denote  $\bar{I}_t^* := I_t^* + \int_0^t \lambda Z_s ds$ . We claim and will prove in *Step 3* that  $\limsup_{\epsilon \rightarrow 0} \bar{\mathbb{E}}^{\nu^*} [|\mathcal{R}_T^\epsilon|] / \epsilon^2 < \infty$ . Then,

$$\frac{J_A(\nu)}{J_A(\nu^*)} - 1 = \epsilon \lambda \bar{\mathbb{E}}^{\nu^\epsilon} \left[ \int_0^T (\partial_\alpha c_t(\nu_t^*)\Delta\alpha_t - (Z_t - \partial_\beta c_t(\nu_t^*))\Delta\beta_t - Z_t h(t)\Delta\hat{X}_t) dt \right] + \mathcal{O}(\epsilon^2).$$

The construction of the costate process  $(P, Q)$  and the variational calculus is the same as step 3 and step 4 of the proof of Theorem II.2.7. We shall leave them to the readers.

*Step 3.* Finally we shall complete the proof by proving the claim we made on  $\mathcal{R}_T^\epsilon$ . The first term in (II.5.62) is clearly uniformly bounded by  $C\epsilon^2$  for some positive constant  $C$ , since the controls are bounded. Since the controls are bounded, using the Hölder inequality and the energy inequalities given in [118, Chapter VII-6], we have

$$\begin{aligned} \bar{\mathbb{E}}^{\nu^*} [|\mathcal{R}_T^\epsilon|] &\leq C\epsilon^2 + \sum_{n=2}^{\infty} \frac{\epsilon^n \lambda^n}{n!} \left( C_1^n + C_2^n \bar{\mathbb{E}}^{\nu^*} \left[ \mathcal{E} \left( \int_0^T \epsilon (h(t)\Delta\hat{X}_t + \Delta\beta_t) d\bar{I}_t^* \right) \left( \int_0^T |Z_t| dt \right)^n \right] \right) \\ &\leq C\epsilon^2 + \sum_{n=2}^{\infty} \frac{\epsilon^n \lambda^n}{n!} \left( C_1^n + C_2^n C_3 \bar{\mathbb{E}}^{\nu^*} \left[ \left( \int_0^T Z_t^2 dt \right)^{n/2} \right] \right) \\ &\leq C\epsilon^2 + \sum_{n=2}^{\infty} \frac{\epsilon^n \lambda^n}{n!} \left( C_1^n + C_2^n C_3 \sqrt{n!} \|Z\|_{BMO(\bar{\mathbb{P}}^{\nu^*})}^n \right) = \mathcal{O}(\epsilon^2). \end{aligned}$$

■

---

---

## CHAPITRE III

---

# PRINCIPAL-AGENT PROBLEM WITH MULTIPLE PRINCIPALS

**Abstract :** We consider a moral hazard problem with multiple principals in a continuous-time model. The agent can only work exclusively for one principal at a given time, so faces an optimal switching problem. Using a randomized formulation, we manage to represent the agent's value function and his optimal effort by an Itô process. This representation further helps to solve the principals' problem in case we have infinite number of principals in the sense of mean field game. Finally the mean field formulation is justified by an argument of propagation of chaos.

**Keywords :** Moral hazard, contract theory, backward SDE, optimal switching, mean field games, propagation of chaos.

### III.1 Introduction

The principal-agent problem is a study of optimizing the incentives, so central in economics. In particular, the optimal contracting between the two parties, principal and agent(s), is called moral hazard, when the agent's effort is not observable by the principal. It has been widely applied in many areas of economics and finance, for example in corporate finance (see [15]). More recently, we also witness the works [1, 2] using the principal-agent formulation to study how to encourage people to embrace the energy transition.

While the research on the discrete-time model can be dated back further, the first paper on continuous-time principal-agent problem is the seminal work by Holmström and Milgrom [78], who study a simple continuous-time model in which the agent gets paid at the end of a finite time interval. They show that optimal contracts are linear in aggregate output when the agent has exponential utility with a monetary cost of effort. The advantage of the continuous-time model is further explored by Sannikov [139]. Not only he considers a new model which allows the agent to retire, but also (and more importantly in the mathematical perspective) he introduces new dynamic insights to the principal-agent problem, and it leads to simple computational procedure to find the optimal contract by solving an ordinary differential equation in his case.

Later, the idea of Sannikov is interpreted and introduced to the mathematical finance community by Cvitanović, Possamaï and Touzi [36]. Let us illustrate their contribution with a toy model. Denote by  $\xi$  the contract paid at the end of a finite time interval  $[0, T]$ . Assume the agent faces the following optimization :

$$\max_{\alpha} \mathbb{E} \left[ \xi(X^{\alpha}) - \int_0^T c(\alpha_t) dt \right], \quad \text{where} \quad dX_t^{\alpha} = dW_t - \alpha_t dt.$$

The crucial observation in [36] is that both the contract  $\xi$  and the Agent's best response  $\alpha^*[\xi]$  can be characterized by the following backward stochastic differential equation (in short, BSDE, for readers not familiar with BSDE we refer to [129, 52], and in particular to [41] for the applications on the contract theory) :

$$dY_t = -c^*(Z_t)dt + Z_t dW_t, \quad Y_T = \xi, \quad \text{where} \quad c^*(z) = \max_a \{az - c(a)\}, \quad (\text{III.1.1})$$

namely,  $\xi = Y_T$  and  $\alpha_t^*[Z] = \operatorname{argmax}_a \{aZ_t - c(a)\}$  for all  $t \in [0, T]$ . This induces a natural (forward) representation of the couple  $(\xi, \alpha^*[\xi])$  :

$$\begin{cases} \xi = Y_T^{Y_0, Z} := Y_0 - \int_0^T c^*(Z_t)dt + \int_0^T Z_t dW_t \\ \alpha_t^*[\xi] := \alpha_t^*[Z] = \operatorname{argmax}_a \{aZ_t - c(a)\}, \quad \text{for all } t \in [0, T] \end{cases} \quad \text{for some } (Y_0, Z),$$

and this neat representation transforms the once puzzling principal's problem into a classical control

problem, namely,

$$\max_{\xi} \mathbb{E} \left[ U \left( X_T^{\alpha^*[\xi]} - \xi \right) \right] = \max_{Y_0, Z} \mathbb{E} \left[ U \left( X_T^{\alpha^*[Z]} - Y_T^{Y_0, Z} \right) \right].$$

This idea of representation is further applied to study the case where the principal can hire multiple agents [55]. Eventually, in [54] the authors follow the same machinery to study the model where the principal hires infinite number of agents using the formulation of mean field games (as for the mean field game we refer to the seminal paper [102] and the recent books [25] and [26]).

There are fewer existing literature considering the model with multiple principals. In [115] the authors consider the case where the several principals hire one common agent *at the same time* (i.e. the agent simultaneously works on different projects for different principals). However, to the best of our knowledge, no one has yet considered a  $n$ -principal/1-agent model where the agent can only exclusively work for one principal at a given time. In such a model, the agent is facing an optimal switching (among the principals) problem. According to the classic literature of optimal switching, see e.g. [135, 67, 85, 31], the counterpart of the BSDE characterization (III.1.1) for the agent's problem would be a system of reflected BSDE in the form :

$$dY_t = -f_t(Y_t, Z_t)dt + Z_t dW_t - dK_t, \quad Y_T = \xi,$$

where  $K$  is an increasing process satisfying some (backward) Skorokhod condition. The presence of the process  $K$  and its constraint make it difficult to find a representation of the couple  $(\xi, \alpha^*[\xi])$  as in [36]. In this paper, we propose an alternative facing this difficulty. Instead of letting the agent choose stopping times to change his employers, we study a randomized optimal switching problem where the switching time is modelled by a random time characterized by a Poisson point process and the agent influences the random time by controlling the intensity of the Poisson point process. It is fair to note that similar randomized formulations of switching have been discussed in the literature, see e.g. [20, 53]. In such framework, we may characterize  $(\xi, \alpha^*[\xi])$  by the solution to a system of BSDEs (without reflection).

Unfortunately, this characterization of the agent's problem and the corresponding representation do not help us to solve simply the principals' problem. We observe that the optimizations the principals face are time-inconsistent, and thus cannot be solved by dynamic programming. However, we also note that this time-inconsistency disappears once the number of principals tends to infinity. In the setting of infinite number of principals, it is natural to adopt the formulation of mean field game. We prove the existence of mean field equilibrium largely based on the recipes in Lacker [99]. Further, in order to justify our mean field game formulation, we introduce a machinery which we name 'backward propagation of chaos' which may carry independent interest itself.

The rest of the paper is organized as follows. In Section III.2, we state our moral hazard problem with  $n$ -principal and one agent, we shall solve the agent's problem under the previously mentioned randomized optimal switching formulation, and observe the time-inconsistency of the  $n$ -principal problem. Then in

section III.3, we shall derive the mean field game among the principals, prove its existence, and justify it using the “backward propagation of chaos” technique.

## III.2 Moral hazard problem : $n$ principals and one agent

In this paper we consider the principal-agent problem on the a finite time horizon  $[0, T]$  for some  $T > 0$ . The main novelty is to introduce the new model and method to allow the agent to choose working for different employers. In this section, we set up the model in which one agent switches among  $n$  different principals.

For the agent, the set of possible regimes is  $\mathbb{I}_n := \{1, 2, \dots, n\}$ . Denote by  $C([0, T]; \mathbb{R}^n)$  the set of continuous functions from  $[0, T]$  to  $\mathbb{R}^n$ , endowed with the supremum norm  $\|\cdot\|_T$ , where

$$\|x\|_t := \sup_{0 \leq s \leq t} |x_s|, \quad t \in [0, T], \quad x \in C([0, T]; \mathbb{R}^n).$$

Denote by  $D([0, T], \mathbb{I}_n)$  the set of càdlàg functions from  $[0, T]$  to  $\mathbb{I}_n$ . We introduce the canonical space  $\Omega := C([0, T]; \mathbb{R}^n) \times D([0, T]; \mathbb{I}_n)$ , denote the canonical process by  $(X, I)$ , and the canonical filtration by  $\mathbb{F} = \{\mathcal{F}_t\}_{t \in [0, T]}$ . The process  $X_t$  represents the outputs of the principals and  $I_t$  records for which principal the agent is working at time  $t$ . Denote by  $\mathbb{F}^X$  the filtration generated by the process  $X$  alone. We also define  $\mathbb{P}_0$  as the Wiener measure on  $C([0, T], \mathbb{R}^n)$ .

In our model, for simplicity, the output process follows the dynamic

$$X_t = X_0 + \int_0^t e_{I_s} ds + W_t, \quad \text{where } e_i = (0, \dots, 0, 1, 0, \dots, 0)^\top \in \mathbb{R}^n, \quad (\text{III.2.2})$$

where  $W$  is a  $n$ -dimensional Brownian motion. That is, the principal employing the agent has a constant-1 drift while outputs of the others just follow the Brownian motion. The agent controls the process  $I$ , that is, he switches from one employer to another so as to maximize his utility :

$$V_0^A(\xi, w) = \sup_I \mathbb{E} \left[ \sum_{i=1}^n \left( \xi^i \mathbf{1}_{\{I_T=i\}} + \int_0^T w_t^i \mathbf{1}_{\{I_t=i\}} du \right) - C_T(I) \right], \quad (\text{III.2.3})$$

where  $\xi^i$  is the reward at the terminal time  $T$ ,  $w^i$  is the wage paid by principal  $i$ , and  $C_T(I)$  denotes the cost of switching  $I$  up to time  $T$  which will be defined later when we analyze the agent's problem in detail. Naturally, the payment  $\xi^i$  should be an  $\mathcal{F}_T^X$ -measurable random variable and  $w^i$  should be an  $\mathbb{F}^X$ -adapted process.

**Remark III.2.1.** *The upcoming method can easily handle the output process :*

$$dX_t = b(t, X_t, I_t)dt + \sigma(t, X_t)dW_t,$$

where  $b, \sigma$  are uniformly Lipschitz in  $x$ . However, for the simplification of notations, we prefer to focus



on the form of (III.2.2).

Providing the contracts  $\{(\xi^i, w^i)\}_{i \in \mathbb{I}_n}$  respectively, the  $n$  principals search among them a Nash equilibrium so as to maximize their profits :

$$V_0^{P,i} = \max_{(\xi^i, w^i)} \mathbb{E} \left[ X_T^{i, \xi^i} - U(\xi^i) \mathbf{1}_{\{I_T=i\}} - \int_0^T U(w_t^i) \mathbf{1}_{\{I_t=i\}} dt \right], \quad (\text{III.2.4})$$

where  $X^{i, \xi^i}$  denotes the optimal output process controlled by the agent.

As in [139] and [36], we are going to provide a representation of the value function of the agent's problem so as to solve the principals' problem by dynamic programming.

### III.2.1 Agent's problem

In our model, instead of allowing the agent to control the process  $I$  in a "singular" way, we assume that  $I$  is driven by a Poisson random measure (p.r.m.) and allow the agent to control its intensity. More precisely, in the weak formulation of stochastic control, the agent aims at choosing his optimal control among :

$$\left\{ \mathbb{P}^\alpha \in \mathcal{P}(\Omega) : \begin{array}{l} \mathbb{P}^\alpha\text{-a.s. (III.2.2) holds true and } dI_t = \int_{\mathbb{I}_n} (k - I_{t-}) \mu(dt, dk), \\ \text{where } \mu \text{ is a p.r.m. with intensity } \alpha(dt, \cdot) = \sum_{k \in \mathbb{I}_n} \alpha_t^k \delta_k dt, \\ \text{for some nonnegative } \mathbb{F}\text{-adapted } (\alpha^k)_{k \in \mathbb{I}_n}, W \text{ and } \mu \text{ are independent} \end{array} \right\},$$

where  $\mathcal{P}(\Omega)$  is the set of all probability measures on  $\Omega$ .

**Remark III.2.2.** Define the first jump time of the process  $I$  :

$$\tau_t := \inf\{s \geq t : I_s \neq I_t\}. \quad (\text{III.2.5})$$

It follows from the Girsanov theorem for multivariate point processes (see e.g. [89]) that for any  $\mathbb{F}^X$ -adapted processes  $\{\theta(i)\}_{i \in \mathbb{I}_n}$  we have

$$\mathbb{E}_t^{\mathbb{P}^\alpha} \left[ \theta_{\tau_t}(I_{\tau_t}) \mathbf{1}_{\{\tau_t \leq s, I_{\tau_t}=j\}} \right] = \mathbb{E}_t^{\mathbb{P}^\alpha} \left[ \int_t^s \beta^{I_t, \alpha}(t, u) \alpha_u^j \theta_u(j) du \right], \quad \text{for all } j \in \mathbb{I}_n \text{ and } j \neq I_t,$$

as well as  $\mathbb{P}^\alpha \left[ \tau_t > s \mid \mathcal{F}_s \right] = \beta^{I_t, \alpha}(t, s)$ , where  $\beta^{i, \alpha}(t, s) := \exp \left( - \int_t^s \sum_{j \neq i} \alpha_u^j du \right)$ .

These results will be useful in the upcoming calculus.

With the intensity  $\alpha$  introduced as above, we can now make it precise how to define the cost of switching in (III.2.3). Given  $n$  contracts  $\{(\xi^i, w^i)\}_{i \in \mathbb{I}_n}$ , the agent solves the following optimal switching problem

$$V_0^A = \sup_{\alpha} \mathbb{E}^{\mathbb{P}^{\alpha}} \left[ \sum_{i=1}^n \xi^i \mathbf{1}_{\{I_T=i\}} + \int_0^T \left( \sum_{i=1}^n w_u^i \mathbf{1}_{\{I_u=i\}} - \frac{1}{n-1} \sum_{i \neq I_u} c((n-1)\alpha_u^i) \right) du \right]. \quad (\text{III.2.6})$$

Intuitively, the intensity process  $\alpha$  describes the hesitation of changing employer. The bigger  $\alpha$  is, the less hesitation the agent has to change his employer. The normalization  $\frac{1}{n-1}c((n-1)\cdot)$  is for later use, as  $n \rightarrow \infty$ .

**Remark III.2.3.** *We note here that the use of weak formulation to model the agent's problem is standard in the literature, as in the moral hazard framework here the principals can only observe the output process  $X$  rather than the noise  $W$ . So it makes more sense to assume that the control  $\alpha$  depends on  $X$  instead of on  $W$ . That is, weak formulation is more appropriate than strong formulation. For more about this issue, we refer the interested reader to the text book of Cvitanić and Zhang [41, Chapter 5].*

As in the classical literature of the optimal switching problems, we shall use the dynamic programming principle to obtain the system of equations characterizing the value function. First, define the dynamic version of (III.2.6) :

$$V_t^A = \text{esssup}_{\alpha} \mathbb{E}_t^{\mathbb{P}^{\alpha}} \left[ \sum_{i=1}^n \xi^i \mathbf{1}_{\{I_T=i\}} + \int_t^T \left( \sum_{i=1}^n w_u^i \mathbf{1}_{\{I_u=i\}} - \frac{1}{n-1} \sum_{i \neq I_u} c((n-1)\alpha_u^i) \right) du \right].$$

Recall  $\tau_t$  defined in (III.2.5). By the dynamic programming, see, e.g., [61, Theorem 5.2], we have

$$V_t^A = \text{esssup}_{\alpha} \mathbb{E}_t^{\mathbb{P}^{\alpha}} \left[ \sum_{i=1}^n \xi^i \mathbf{1}_{\{I_T=i, \tau_t > T\}} + V_{\tau_t}^A \mathbf{1}_{\{\tau_t \leq T\}} + \int_t^{\tau_t \wedge T} \left( \sum_{i=1}^n w_u^i \mathbf{1}_{\{I_u=i\}} - \frac{1}{n-1} \sum_{i \neq I_u} c((n-1)\alpha_u^i) \right) du \right].$$

Further, by defining  $V_t^{A,i} := V_t^A |_{I_t=i}$ , the value of the control problem at time  $t$  given  $I_t = i$ , we obtain

$$\begin{aligned} V_t^{A,i} &= \text{esssup}_{\alpha} \mathbb{E}_t^{\mathbb{P}^{\alpha}} \left[ \xi^i \mathbf{1}_{\{\tau_t > T\}} + \sum_{j \neq i} \left( V_{\tau_t}^{A,j} \mathbf{1}_{\{\tau_t \leq T, I_{\tau_t}=j\}} + \int_t^{\tau_t \wedge T} \frac{w_u^i - c((n-1)\alpha_u^j)}{n-1} du \right) \right] \\ &= \text{esssup}_{\alpha} \mathbb{E}_t^{\mathbb{P}^{\alpha}} \left[ \xi^i \beta^{i,\alpha}(t, T) + \int_t^T \beta^{i,\alpha}(t, u) \sum_{j \neq i} \left( \alpha_u^j V_u^{A,j} + \frac{w_u^i - c((n-1)\alpha_u^j)}{n-1} \right) du \right]. \quad (\text{III.2.7}) \end{aligned}$$

The second equality above is due to the results in Remark III.2.2. In view of (III.2.7), it becomes a classical stochastic control problem. In particular, the value function of this control problem can be

characterized by the BSDE.

**Assumption III.2.4.** Assume that the cost function  $c$  is convex, and  $c$  takes value of  $+\infty$  out of a compact set  $K$ . We also assume that there exists a unique maximizer

$$a^*(y) := \operatorname{argmax}_{a \geq 0} \{ay - c(a)\} \in K, \quad \text{for all } y \in \mathbb{R}.$$

Define the convex conjugate of  $c$

$$c^*(y) := \sup_{a \geq 0} \{ay - c(a)\}, \quad \text{for all } y \in \mathbb{R}.$$

Further assume that  $c^*$  is Lipschitz continuous.

As an example, consider  $c(a) := \frac{1}{2}a^2$  and  $K := [0, \bar{K}]$  for some  $\bar{K} > 0$ . Then,

$$a^*(y) = \begin{cases} 0, & y \leq 0, \\ y, & 0 \leq y \leq \bar{K}, \\ \bar{K}, & y \geq \bar{K}, \end{cases} \quad \text{and} \quad c^*(y) = \begin{cases} 0, & y \leq 0, \\ \frac{1}{2}y^2, & 0 \leq y \leq \bar{K}, \\ \bar{K}y - \frac{1}{2}\bar{K}^2, & y \geq \bar{K}. \end{cases}$$

Obviously, the function  $c^*$  is Lipschitz continuous with constant  $\bar{K}$ .

**Proposition III.2.5.** Under Assumption III.2.4, given  $\xi^i \in \mathbb{L}^2(\mathbb{P}_0)$ ,  $w^i \in \mathbb{H}^2(\mathbb{P}_0)$ <sup>1</sup> for all  $i \in \mathbb{I}_n$ , the following system of BSDEs has a unique solution  $(Y^i, Z^i)_{i \in \mathbb{I}_n}$ :

$$Y_t^i = \xi^i + \int_t^T \left( \frac{1}{n-1} \sum_{j \neq i} c^*(Y_u^j - Y_u^i) + w_u^i + e_i Z_u^i \right) du - \int_t^T Z_u^i \cdot dX_u, \quad i \in \mathbb{I}_n, \quad \mathbb{P}_0\text{-a.s.} \quad (\text{III.2.8})$$

Moreover, we have  $Y^i = V^{A,i}$ ,  $\mathbb{P}_0$ -a.s. In particular, the optimal intensity satisfies

$$\alpha_t^{j,*} = \frac{1}{n-1} a^*(Y_t^j - Y_t^{I_t}), \quad j \neq I_t, \quad \text{for all } t \in [0, T], \quad \mathbb{P}_0\text{-a.s.} \quad (\text{III.2.9})$$

*Démonstration.* In view of the control problem (III.2.7), following the result of [66, 52], see also [164],

---

1. We denote by  $\mathbb{H}^2(\mathbb{P}_0)$  the space of progressively measurable processes  $w$  such that

$$\|w\|_{\mathbb{H}^2(\mathbb{P}_0)}^2 := \mathbb{E}^{\mathbb{P}_0} \left[ \left( \int_0^T |w_u|^2 du \right)^{\frac{1}{2}} \right] < \infty.$$

the corresponding BSDE reads

$$Y_t^i = \xi^i + \int_t^T \left( \frac{1}{n-1} \sum_{j \neq i} \sup_{a^j \geq 0} \left\{ (n-1)a^j (Y_u^j - Y_u^i) - c((n-1)a^j) \right\} + w_u^i + e_i Z_u^i \right) du - \int_t^T Z_u^i \cdot dX_u.$$

Then, (III.2.8) follows from the definition of  $c^*$ . Since all the coefficients are Lipschitz continuous, the wellposedness of the BSDE system and the verification for the control problem is classical, see e.g. [129].

■

### III.2.2 Principals' problem : time inconsistency

In the previous section, we managed to represent the value function of the agent by an Itô process (Proposition III.2.5). As in [139, 36], we expect that this representation would help us to solve the principals' problem by the dynamic programming approach. However, in this model, this approach does not work. We shall explain in the case  $n = 2$  for the simplification of notation.

Consider the set of all contracts

$$\Xi := \left\{ \left\{ (\xi^i, w^i) \right\}_{i=1,2} : \xi^i \in \mathbb{L}^2(\mathbb{P}_0), w^i \in \mathbb{H}^2(\mathbb{P}_0) \text{ and } V^A(\xi, w) \geq R \right\},$$

where  $R$  is the reservation value of the agent for whom only the contracts such that  $V^A(\xi, w) \geq R$  are acceptable. Now define

$$\mathcal{V} := \left\{ \left\{ (Y_0^i, Z^i) \right\}_{i=1,2} : Y_0^i \geq R, Z^i \in \mathbb{H}^2(\mathbb{P}_0), i = 1, 2 \right\}.$$

It follows from Proposition III.2.5 that

$$\begin{aligned} \Xi = \left\{ \left\{ (\xi^i, w^i) \right\}_{i=1,2} : w^i \in \mathbb{H}^2(\mathbb{P}_0) \text{ and } \xi^i = Y_T^{i, Y_0^i, Z^i, w^i}, \text{ where } Y^{i, Y_0^i, Z^i, w^i} \text{ satisfies} \right. \\ \left. Y_T^i = Y_0^i - \int_0^T \left( c^*(Y_t^j - Y_t^i) + w_t^i + e_i Z_t^i \right) dt + \int_0^T Z_t^i \cdot dX_t, \mathbb{P}_0\text{-a.s.}, \right. \\ \left. \text{with } j \neq i \text{ and } \left\{ (Y_0^i, Z^i) \right\}_{i=1,2} \in \mathcal{V} \right\}. \end{aligned}$$

For simplicity of notation, if there is no ambiguity, we write simply  $Y^i$  instead of  $Y^{i, Y_0^i, Z^i, w^i}$ . But just keep in mind that the process  $Y^i$  is controlled by  $Z^i$  and  $w^i$  with initial value  $Y_0^i$ . The corresponding optimal intensity reads  $\alpha_t^* = \left( a^*(Y_t^1 - Y_t^2) \mathbf{1}_{\{I_t=2\}}, a^*(Y_t^2 - Y_t^1) \mathbf{1}_{\{I_t=1\}} \right)$ . Therefore, the principals are

searching for a Nash equilibrium so as to maximize :

$$\begin{aligned} V_0^{P,i} &= \sup_{(\xi^i, w^i)} \mathbb{E}^{\mathbb{P}^{\alpha^*}} \left[ X_T^i - U(\xi^i) \mathbf{1}_{\{I_T=i\}} - \int_0^T U(w_t^i) \mathbf{1}_{\{I_t=i\}} dt \right] \\ &= \sup_{(Y_0^i, Z^i, w^i)} \mathbb{E}^{\mathbb{P}^{\alpha^*}} \left[ X_T^i - U(Y_T^i) \mathbf{1}_{\{I_T=i\}} - \int_0^T U(w_t^i) \mathbf{1}_{\{I_t=i\}} dt \right] = \sup_{Y_0^i \geq R} J_0^i(Y_0^i), \end{aligned}$$

where

$$J_0^i(Y_0) := \sup_{(Z^i, w^i)} \mathbb{E}^{\mathbb{P}^{\alpha^*}} \left[ X_T^i - U\left(Y_T^i, Y_0^i, Z^i, w^i\right) \mathbf{1}_{\{I_T=i\}} - \int_0^T U(w_t^i) \mathbf{1}_{\{I_t=i\}} dt \right].$$

Up to now, we are applying the same strategy as in [36]. Assume the control problem is time-consistent, that is, admits dynamic programming. Fix  $Y_0$  and define the dynamic version of  $J_0^i$  :

$$J_t^i(Y_t^i) := \text{ess sup}_{(Z^i, w^i)} \mathbb{E}_t^{\mathbb{P}^{\alpha^*}} \left[ X_T^i - U(Y_T^i) \mathbf{1}_{\{I_T=i\}} - \int_t^T U(w_u^i) \mathbf{1}_{\{I_u=i\}} du \right].$$

Defining  $J_t^{i,1} := J_t^i(Y_t^i)|_{I_t=1}$  and  $J_t^{i,2} = J_t^i(Y_t^i)|_{I_t=2}$  according to the different regimes, we expect to have for  $i = 1$

$$\begin{aligned} J_t^{1,1} &= \text{ess sup}_{(Z^1, w^1)} \mathbb{E}_t^{\mathbb{P}^{\alpha^*}} \left[ \left( X_T^1 - U(Y_T^1) \right) \mathbf{1}_{\{\tau_t > T\}} + J_{\tau_t}^{1,2} \mathbf{1}_{\{\tau_t \leq T\}} - \int_t^{\tau_t \wedge T} U(w_u^1) du \right] \\ &= \text{ess sup}_{(Z^1, w^1)} \mathbb{E}_t^{\mathbb{P}^{\alpha^*}} \left[ \beta^{1,*}(t, T) \left( X_T^1 - U(Y_T^1) \right) \right. \\ &\quad \left. + \int_t^T \beta^{1,*}(t, u) a^*(Y_u^2 - Y_u^1) \left( J_u^{1,2} - U(w_u^1) \right) du \right] \end{aligned} \quad (\text{III.2.10})$$

where

$$dX_t^1 = dt + dW_t^1, \quad \beta^{1,*}(t, s) := \exp \left( - \int_t^s a^*(Y_u^2 - Y_u^1) du \right),$$

and similarly

$$\begin{aligned} J_t^{1,2} &= \text{ess sup}_{(Z^1, w^1)} \mathbb{E}_t^{\mathbb{P}^{\alpha^*}} \left[ X_T^1 \mathbf{1}_{\{\tau_t > T\}} + J_{\tau_t}^{1,1} \mathbf{1}_{\{\tau_t \leq T\}} \right] \\ &= \text{ess sup}_{(Z^1, w^1)} \mathbb{E}_t^{\mathbb{P}^{\alpha^*}} \left[ \beta^{2,*}(t, T) X_T^1 + \int_0^T \beta^{2,*}(t, u) a^*(Y_u^1 - Y_u^2) J_u^{1,1} du \right] \end{aligned} \quad (\text{III.2.11})$$

where

$$dX_t^1 = dW_t^1, \quad \beta^{2,*}(t, s) := \exp \left( - \int_t^s a^*(Y_t^1 - Y_t^2) du \right).$$

Although it seems promising to solve the system of value functions  $(J^{i,1}, J^{i,2})$  as in the agent problem, note that the admissible controls  $(Z^i, w^i)$  are constrained to be  $\mathbb{F}^X$ -adapted, in particular,  $(Z^i, w^i)$  cannot depend on the regimes. Since the optimizations in (III.2.10) and (III.2.11) are not symmetric,

we cannot expect that the optimal controls  $(Z^i, w^i)$  coincide on the different regimes. Therefore, it is a contradiction to the alleged time-consistency.

This time-inconsistency appears in our model, if we consider a finite number of principals. In the following section, we shall bypass the difficulty in the case of infinite number of principals, using the mean field formulation.

**Remark III.2.6.** *Besides, the approach for solving the  $n$ -principal problem has another drawback. Note that each contract  $\xi^i$  for  $i \in \mathbb{I}_n$  is a function of the entire vector  $X$ , that is, the principal need to know the other companies' performance in order to design his own contract. This problem can be also avoided in the mean field approach, where  $\xi^i$  will be a function of  $X^i$ .*

### III.3 Infinite number of principals : mean field approach

#### III.3.1 Heuristic analysis

Heuristically, as  $n \rightarrow \infty$  the equation (III.2.8), which characterizes the agent's value function, converges to

$$dY_t^i = - \left( \int c^*(y - Y_t^i) p_t(dy) + w_t^i + Z_t^i \right) dt + Z_t^i dX_t^i, \quad p_t = \mathcal{L}(Y_t^i). \quad (\text{III.3.12})$$

Besides, it follows from (III.2.9) and the definition of the discount factor  $\beta$  that the optimal  $\alpha^*$  and  $\beta^*$  converge to

$$\alpha_t^{j,*} \rightarrow 0 \text{ for } j \neq i, \quad \sum_{j \neq i} \alpha_t^{j,*} \rightarrow \alpha_t^* := \int a^*(y - Y_t^i) p_t(dy), \quad (\text{III.3.13})$$

and  $\beta_t^{n,*} \rightarrow \beta_t^* := e^{-\int_0^t \alpha_s^* ds}$ .

**Remark III.3.1.** — *To be fair, the form of the limit equation (III.3.12) does not entirely follow the intuition. Note that we removed the  $e_i$  in front of the term  $Z_t^i dt$  and replaced the stochastic integrator  $dX_t$  by  $dX_t^i$ . At this stage, let us admit that once in the drift of the BSDE there is no longer dependence on other  $Y^j$  ( $j \neq i$ ), the system would be decoupled, that is, the equation of  $Y^i$  no longer needs the information of  $X^j$  ( $j \neq i$ ), and it leads to the limit form (III.3.12). We will justify the mean field formulation in Section III.3.3.*

— *The first observation in (III.3.13) implies that in the limit case once the agent leaves the company  $i$ , he will have null probability to come back due to the fact that we have a continuum of players. Therefore, the principal should be only interested in the optimization for the regime where the agent works for her. Remember that in the  $n$ -principal problem, the time-inconsistency is due to the optimizations in the different regimes. Therefore, it is reasonable that the mean field formulation would bypass this difficulty.*

Recall the principal's problem (III.2.4). It follows from the observation in Remark III.3.1 that

$$V_0^{P,i} = \begin{cases} \mathbb{E}^{\mathbb{P}^{\alpha^*}} [X_0^i + W_T] = X_0^i, & \text{as } I_0 \neq i, \\ \max_{(\xi^i, w^i)} \mathbb{E}^{\mathbb{P}^{\alpha^*}} \left[ X_\tau^i \mathbf{1}_{\{\tau \leq T\}} + (X_T^i - U(\xi^i)) \mathbf{1}_{\{\tau > T\}} - \int_0^{\tau \wedge T} U(w_u^i) du \right], & \text{as } I_0 = i. \end{cases}$$

Therefore, in the upcoming mean field game, we should only keep the nontrivial regime ( $I_0 = i$ ), i.e., denote

$$\begin{aligned} V_0^{P,i} &:= \max_{(\xi^i, w^i)} \mathbb{E}^{\mathbb{P}^{\alpha^*}} \left[ X_\tau^i \mathbf{1}_{\{\tau \leq T\}} + (X_T^i - U(\xi^i)) \mathbf{1}_{\{\tau > T\}} - \int_0^{\tau \wedge T} U(w_u^i) du \right] \\ &= \max_{(\xi^i, w^i)} \mathbb{E}^{\mathbb{P}^{\alpha^*}} \left[ \int_0^T \alpha_u^* \beta_u^* (X_u^i - U(w_u^i)) du + \beta_T^* (X_T^i - U(\xi^i)) \right]. \end{aligned} \quad (\text{III.3.14})$$

Though the rigorous definition of the mean field game will be introduced in the next section, we are ready to give a simple description of the mean field equilibrium we are searching for. In the mean field formulation, we remove the superscript  $i$ , and instead we use the notations  $\bar{X}, \bar{Y}, \bar{Z}$ . Also note that  $(\bar{X}_t - t)_{t \in [0, T]}$  is a  $\mathbb{P}^{\alpha^*}$ -Brownian motion, denoted by  $\bar{W}$ . Following the dynamic programming approach in [139, 36], given  $(p_t)_{t \in [0, T]}$ , consider the contracts in the form :

$$\xi \in \Xi(p) = \left\{ \bar{Y}_T : \bar{Y}_T = \bar{Y}_0 - \int_0^T \left( \int c^*(y - \bar{Y}_u) p_u(dy) + w_u \right) du + \int_0^T \bar{Z}_u d\bar{W}_u \right\}.$$

It follows from (III.3.14) that each principal faces the optimization :

$$V_0^P(p) = \sup_{(\bar{Y}_0, \bar{Z}, w)} \mathbb{E}^{\mathbb{P}^{\alpha^*}} \left[ \int_0^T \alpha_u^* \beta_u^* (\bar{X}_u - U(w_u)) du + \beta_T^* (\bar{X}_T - U(\bar{Y}_T)) \right]. \quad (\text{III.3.15})$$

The maximizer  $(\bar{Y}_0^*, \bar{Z}^*, w^*)$  would define a process  $\bar{Y}^*$ . The law  $(p_t)_{t \in [0, T]}$  is a mean field equilibrium, if  $p_t = \mathcal{L}(\bar{Y}_t^*)$  for all  $t \in [0, T]$ .

### III.3.2 Mean field game : existence

Our rigorous definition of the mean field game and the argument to prove its existence rely largely on the framework in Lacker [99].

In this part of the paper, we denote the canonical space by  $\bar{\Omega} = C([0, T], \mathbb{R}^2)$ , the canonical process

by  $\Theta := (\bar{X}, \bar{Y})$  and the canonical filtration by  $\bar{\mathbb{F}}$ . Given  $p \in \mathcal{P}(\bar{\Omega})$ , define  $p_t := p \circ Y_t^{-1} \in \mathcal{P}(\mathbb{R})$  and

$$\begin{aligned} \mathcal{W}(p) := & \left\{ \mathbb{P}^{\lambda, \eta} \in \mathcal{P}(\bar{\Omega}) : \mathbb{P}^{\lambda, \eta}\text{-a.s. } d\bar{X}_t = dt + dW_t, \quad X_0 = x_0, \right. \\ & d\bar{Y}_t = - \int \left( c^*(y - \bar{Y}_t) p_t(dy) + w_t \right) dt + \sqrt{\eta_t} dW_t, \quad \mathcal{L}(\bar{Y}_0) = \lambda, \\ & \left. \text{for some } \lambda \in \mathcal{I}, \quad \eta \in \mathcal{U}, \quad w \in \mathbb{H}^2(\mathbb{P}^{\lambda, \eta}) \right\} \end{aligned}$$

where, for technical reasons, we define

$$\begin{aligned} \mathcal{I} := & \left\{ \lambda \in \mathcal{P}(\mathbb{R}) : \lambda \text{ with a compact support } K \text{ in } [R, \infty) \right\} \\ \text{and } \mathcal{U} := & \left\{ \eta \text{ } \bar{\mathbb{F}}\text{-adapted} : \eta \text{ takes values in a compact set } \Sigma \text{ in } \mathbb{R}^+ \right\}. \end{aligned}$$

In other words, we will consider a mean field game in which the choice of the initial value and the volatility is constrained in compact sets. Further define

$$J(\mathbb{P}, w; p) := \mathbb{E}^{\mathbb{P}} \left[ \int_0^T \alpha_u^* \beta_u^* (\bar{X}_u - U(w_u)) du + \beta_T^* (\bar{X}_T - U(\bar{Y}_T)) \right],$$

where we recall  $\alpha_t^* = \int a^*(y - \bar{Y}_t) p_t(dy)$  and  $\beta_t^* = \exp\left(-\int_0^t \alpha_s^* ds\right)$ .

**Assumption III.3.2.** *We assume that  $a^* : \mathbb{R} \rightarrow \mathbb{R}$  is Lipschitz continuous.  $U$  is convex and of  $q$ -polynomial growth, i.e. there are constants  $C < C'$  and  $q > 1$  such that*

$$C(|w|^q - 1) \leq U(w) \leq C'(|w|^q + 1).$$

**Theorem III.3.3.** *Under Assumption III.2.4 and III.3.2, there exists  $p \in \mathcal{P}(\bar{\Omega})$ ,  $(\hat{\lambda}, \hat{\eta}) \in \mathcal{I} \times \mathcal{U}$  and an  $\bar{\mathbb{F}}$ -adapted process  $\hat{w}$  such that*

$$\left( \mathbb{P}^{\hat{\lambda}, \hat{\eta}}, \hat{w} \right) \in \underset{\mathbb{P} \in \mathcal{W}(p), w}{\operatorname{argmin}} J(\mathbb{P}, w; p) \quad \text{and} \quad p = \mathbb{P}^{\hat{\lambda}, \hat{\eta}}.$$

**Remark III.3.4.** — *This definition of mean field game via the weak formulation of stochastic control follows the same spirit as those in Carmona & Lacker [27] and Lacker [99]. However, here we also include the control of the initial distribution  $\lambda$  of  $\bar{Y}$ .*

- *It is noteworthy that among the triple of control  $(\lambda, \eta, w)$  of this mean field game,  $\lambda$  takes values of measures, that is, the principals are allowed to play a mixed strategy.*
- *If we constrain  $w$  to take values in a compact set in  $\mathbb{R}$ , then given a bounded function  $U$  which is convex on this compact set, the mean field game exists.*

We realize that Theorem III.3.3 can be proved by essentially the same argument as in [99]. In the rest of the section, we shall outline the strategy of the proof and refer the readers for details to [99].



First, we shall linearize the functional  $J$  using the so-called relaxed control. Denote by  $\mathcal{D}$  the set of measures  $q$  on  $[0, T] \times \Sigma \times \mathbb{R}$ . Instead of controlling via the processes  $\eta$  (taking values in  $\Sigma$ ) and  $w$  (taking values in  $\mathbb{R}$ ), we shall control through the measure  $q$  in the relaxed formulation. The canonical space for the relaxed control becomes  $\widehat{\Omega} := \overline{\Omega} \times \mathcal{D}$ . Denote the canonical process by  $(\overline{X}, \overline{Y}, \Lambda)$ , and the canonical filtration by  $\widehat{\mathbb{F}}$ . Note that one may define a  $\mathcal{P}(\Sigma \times \mathbb{R})$ -valued  $\widehat{\mathbb{F}}$ -predictable process  $(\Lambda_t)_{0 \leq t \leq T}$  such that  $\Lambda(dt, d\eta, dw) = \Lambda_t(d\eta, dw)dt$ . Here we abuse the notations using  $\eta, w$  to represent points in  $(\Sigma, \mathbb{R})$ .

Denote by  $\mathcal{C}_0^\infty(\mathbb{R}^2)$  denote the set of infinitely differentiable functions  $\phi : \mathbb{R}^2 \mapsto \mathbb{R}$  with compact support. Define the generator  $L$  on  $\mathcal{C}_0^\infty(\mathbb{R}^2)$  by

$$L^{p,\eta}\phi(t, x, y) = \left(1, - \int c^*(y' - y)p_t(dy') - w_t\right) \nabla\phi + \frac{1}{2}(\partial_{xx}\phi + \eta\partial_{yy}\phi + 2\sqrt{\eta}\partial_{xy}\phi),$$

for  $(t, x, y, p, \eta) \in [0, T] \times \mathbb{R}^2 \times \mathcal{P}(\overline{\Omega}) \times \Sigma$ . Further define

$$M_t^{p,\phi} := \phi(\overline{X}_t, \overline{Y}_t) - \int_0^t \int_{\Sigma \times \mathbb{R}} L^{p,\eta}\phi(s, \overline{X}_s, \overline{Y}_s)\Lambda_s(d\eta)ds$$

**Definition III.3.5.** Given  $p \in \mathcal{P}(\overline{\Omega})$ , define the set of the controlled martingale problems :

$$\begin{aligned} \mathcal{R}(p) = \left\{ \widehat{\mathbb{P}} \in \mathcal{P}(\widehat{\Omega}) : \widehat{\mathbb{P}}\text{-a.s. } \overline{X}_0 = x_0, \overline{Y}_0 \sim \lambda, \text{ for some } \lambda \in \mathcal{I}, \right. \\ \mathbb{E}^{\widehat{\mathbb{P}}} \left[ \int_0^T |\Lambda_t|^q dt \right] < \infty, \\ \left. M^{p,\phi} \text{ is a } \widehat{\mathbb{P}}\text{-martingale for each } \phi \in \mathcal{C}_0^\infty(\mathbb{R}^2) \right\}, \end{aligned} \quad (\text{III.3.16})$$

where  $|\Lambda_t|^q := \int_{\Sigma \times \mathbb{R}} |(\eta, w)|^q \Lambda_t(d\eta, dw)$ .

Further, in the relaxed formulation the object function of the principals reads :

$$\hat{J}(\widehat{\mathbb{P}}; p) := \mathbb{E}^{\widehat{\mathbb{P}}} \left[ \int_0^T \int_{\mathbb{R}} \alpha_u^* \beta_u^* (\overline{X}_u - U(w)) \Lambda_u(dw) du + \beta_T^* (\overline{X}_T - U(\overline{Y}_T)) \right],$$

and define the set of the optimal control :

$$\mathcal{R}^*(p) := \operatorname{argmax}_{\widehat{\mathbb{P}} \in \mathcal{R}(p)} \hat{J}(\widehat{\mathbb{P}}; p).$$

We say  $\widehat{\mathbb{P}} \in \mathcal{P}(\widehat{\Omega})$  is a relaxed mean field game if  $\widehat{\mathbb{P}} \in \mathcal{R}^*(\widehat{\mathbb{P}} \circ (\overline{X}, \overline{Y})^{-1})$ .

**Theorem III.3.6.** Under Assumption III.2.4 and III.3.2, there exists a relaxed mean field game.

*Démonstration.* Our setting slightly distinguishes from the one in [99], because we allow to control the initial law  $\lambda$  of  $\overline{Y}$ . However, since we constrain the choice of  $\lambda$  among the distribution in  $\mathcal{P}(K)$  where

$K$  is a compact set in  $\mathbb{R}$ , we are still able to prove the tightness of  $\mathcal{R}(p)$  (in the case  $\Lambda$  is truncated), and the rest of the argument would follow the same lines in [99, Section 4 and 5]. ■

It remains to construct a strict mean field game (as in Theorem III.3.3) based on a relaxed one. Again we can follow the classical argument. We shall only provide the sketch of the proof, for more details we refer the readers to [99, Proof of Theorem 3.7].

*Proof of Theorem III.3.3.* Let  $\widehat{\mathbb{P}}$  be a relaxed mean field game. First, by the same argument as in the proof of [99, Proof of Theorem 3.7], we may find an  $\overline{\mathbb{F}}$ -adapted process  $\widehat{q} : [0, T] \times \overline{\Omega} \rightarrow \mathcal{P}(\Sigma \times \mathbb{R})$  such that

$$\widehat{q}(t, \overline{X}, \overline{Y}) = \mathbb{E}^{\widehat{\mathbb{P}}}[\Lambda_t | \overline{\mathcal{F}}_t], \quad \widehat{\mathbb{P}}\text{-a.s.} \quad t \in [0, T]. \quad (\text{III.3.17})$$

Further, for each  $(t, x, y, p) \in [0, T] \times \overline{\Omega} \times \mathcal{P}(\overline{\Omega})$ , since  $\alpha^*, \beta^*$  are always nonnegative and function  $U$  is convex, the following subset

$$K(t, x, y, p) := \left\{ (\eta, w, l) : \eta \in \Sigma, l \leq f(t, x, y, p, w) := \alpha_t^*(y, p)\beta_t^*(y, p)(x_t - U(w)) \right\}$$

is convex. This verifies the ‘‘convexity’’ assumption in [99, Assumption (Convex)]. Therefore, using the measurable selection result in [71, Lemma 3.1], there exist  $\overline{\mathbb{F}}$ -adapted processes  $\widehat{\eta}, \widehat{w}$  and  $\widehat{l} \geq 0$  such that

$$\int (\eta, f(t, x, y, p, w)) \widehat{q}(t, x, y)(d\eta, dw) = \left( \widehat{\eta}(t, x, y), f(t, x, y, p, \widehat{w}(t, x, y)) - \widehat{l}(t, x, y) \right).$$

Further, it would be easy to verify that  $p := \widehat{\mathbb{P}} \circ (\overline{X}, \overline{Y})^{-1} = \mathbb{P}^{\lambda, \widehat{\eta}}$  as well as  $J(\mathbb{P}^{\lambda, \widehat{\eta}}, \widehat{w}; p) \geq \widehat{J}(\widehat{\mathbb{P}}; p)$ . Therefore, we find a mean field game in sense of Theorem III.3.3. ■

From the sketch of proof, we may observe the following.

**Corollary III.3.7.** *Under Assumption III.2.4 and III.3.2, there exists a mean field game such that*

$$\mathbb{E}^{\mathbb{P}^{\lambda, \widehat{\eta}}} \left[ \int_0^T |\widehat{w}_u|^q du \right] < \infty.$$

*Démonstration.* Note that in the proof of Theorem III.3.3, the  $\widehat{w}$  we constructed satisfies

$$C(|\widehat{w}_t|^q - 1) \leq U(\widehat{w}_t) \leq \int_{\mathbb{R}} U(w) \widehat{q}_t(dw),$$

where  $\widehat{q}$  is defined in (III.3.17). Therefore,

$$\begin{aligned} C \mathbb{E}^{\mathbb{P}^{\lambda, \widehat{\eta}}} \left[ \int_0^T |\widehat{w}_u|^q du \right] &\leq C + \mathbb{E}^{\widehat{\mathbb{P}}} \left[ \int_0^T \int_{\mathbb{R}} U(w) \Lambda_u(dw) du \right] \\ &\leq C + \mathbb{E}^{\widehat{\mathbb{P}}} \left[ \int_0^T \int_{\mathbb{R}} C'(|w|^q + 1) \Lambda_u(dw) du \right]. \end{aligned}$$

Then, the desired result follows from (III.3.16). ■

### III.3.3 Mean field game : approximation

In this section, we shall justify our mean field formulation, that is, answer the question why the principals would apply the mean field game we studied in the previous section. Unlike the classical cases, recall that in our  $n$ -player problem the stochastic control problem is time-inconsistent. Therefore, it would be difficult to verify whether a contract provided by the mean field game would be  $\epsilon$ -optimal for the  $n$ -player game. Instead, we shall verify that using a contract provided by the mean field game, the principals would receive, in a  $n$ -player game, a utility  $\epsilon$ -close to the value function calculated in the mean field game.

#### III.3.3.1 Agent problem : backward propagation of chaos

Here we shall analyze the agent's behavior, once he is given  $n$  contracts provided by the mean field game. Note that the mean field contract has the form :

$$\xi = \bar{Y}_0 - \int_0^T \left( \int_{\mathbb{R}} c^*(y - \bar{Y}_u) p_u^*(dy) + w_u + \sqrt{\eta_u} \right) du + \int_0^T \sqrt{\eta_u} d\bar{X}_u. \quad (\text{III.3.18})$$

In this section, we shall assume that the contract  $(\xi, w)$  is  $\mathcal{F}_T^{\bar{X}, \bar{Y}_0}$ -measurable, that is, we may write  $\xi(\bar{X}, \bar{Y}_0)$  and  $w(\bar{X}, \bar{Y}_0)$ . Now, recall the canonical space  $\Omega$  in the setting of  $n$ -principal problem, and the canonical process  $X = \{X^i\}_{i \in \mathbb{I}_n}$  representing the outputs of all principals. Further, let  $\{(Y^{*,i}, \xi^i, w^i, \eta^i)\}_{i \in \mathbb{I}_n}$  be the  $n$  independent copies of  $(\bar{Y}, \xi, w, \eta)$  such that

$$\begin{aligned} Y_t^{*,i} &= \xi^i + \int_t^T \left( \int_{\mathbb{R}} c^*(y - Y_u^{*,i}) p_u^*(dy) + w_u^i + \sqrt{\eta_u^i} \right) du - \int_t^T \sqrt{\eta_u^i} dX_u^i \\ &= \xi^i + \int_t^T \left( \int_{\mathbb{R}} c^*(y - Y_u^{*,i}) p_u^*(dy) + w_u^i + e_i Z_u^{*,i} \right) du - \int_t^T Z_u^{*,i} \cdot dX_u, \end{aligned} \quad (\text{III.3.19})$$

with  $Z^{*,i} := e_i \sqrt{\eta^i} = (0, \dots, 0, \sqrt{\eta^i}, 0, \dots, 0)^\top$ .

In this section, we shall simply use the notation  $\mathbb{E}$  instead of  $\mathbb{E}^{\mathbb{P}^{\lambda, \hat{\eta}}}$ . The following estimate follows directly from Corollary III.3.7.

**Lemma III.3.8.** *Assume  $q \geq 2$  in Assumption III.3.2. Then we have*

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |Y_t^{*,i}|^2 \right] < \infty$$

In particular  $\mathbb{E}[|\xi^i|^2] < \infty$ .

Given such contracts  $\{(\xi^i, w^i)\}_{i \in \mathbb{I}_n}$ , the agent would solve the system of BSDE (III.2.8), namely,

$$\begin{aligned} Y_t^{n,i} &= \xi^i + \int_t^T \left( \frac{1}{n-1} \sum_{\substack{j=1 \\ j \neq i}}^n c^*(Y_u^{n,j} - Y_u^{n,i}) + w_u^i + e_i Z_u^{n,i} \right) du - \int_t^T Z_u^{n,i} \cdot dX_u \\ &= \xi^i + \int_t^T \left( \frac{n}{n-1} \int_{\mathbb{R}} c^*(y - Y_u^{n,i}) p_u^n(dy) - \frac{c^*(0)}{n-1} + w_u^i + e_i Z_u^{n,i} \right) du - \int_t^T Z_u^{n,i} \cdot dX_u, \end{aligned} \quad (\text{III.3.20})$$

where  $p^n$  is the empirical measure  $\frac{1}{n} \sum_{j=1}^n \delta_{Y_u^{n,j}}$  and  $p_u^n := \frac{1}{n} \sum_{j=1}^n \delta_{Y_u^{n,j}}$ . Define  $\Delta Y^i := Y^{n,i} - Y^{*,i}$ ,  $\Delta Z^i := Z^{n,i} - Z^{*,i}$ . Let  $\mathcal{C} := C([0, T]; \mathbb{R})$ . Denote the square of the Wasserstein-2 distance on  $\mathcal{P}^2(\mathcal{C})$  by

$$d_t^2(\mu, \nu) := \inf_{\pi \in \Pi(\mu, \nu)} \int_{\mathcal{C} \times \mathcal{C}} \sup_{t \leq u \leq T} |x_u - y_u|^2 \pi(dx, dy).$$

Here is the main result concerning the agent's problem.

**Proposition III.3.9.** *Assume that the contract  $(\xi, w)$  is  $\mathcal{F}_T^{\bar{X}, \bar{Y}_0}$ -measurable and Assumption III.2.4 and III.3.2 hold true for some  $q \geq 2$ . Then we have*

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ d_0^2(p^n, p^*) \right] = 0. \quad (\text{III.3.21})$$

**Remark III.3.10.** — *There is a certain similarity between the previous result and the one in [22] where the authors also study the convergence from a  $n$ -player BSDE to the mean field limit equation in the form of (III.3.19). To our understanding of their paper, the solution to the  $n$ -player BSDE there is a fixed point of the following map :*

$$p \mapsto Y(p) \mapsto \frac{1}{n} \sum_{i=1}^n \delta_{Y^i(p)},$$

where  $Y(p)$  is the solution of the BSDE

$$Y_t = \xi + \int_t^T \left( \int_{\mathbb{R}} c^*(y - Y_u) p_u(dy) + w_u + Z_u \right) du - \int_t^T Z_u d\bar{X}_u, \quad (\text{III.3.22})$$

and  $\{Y^i(p)\}_{i \in \mathbb{I}_n}$  are the independent copies of  $Y(p)$ . In other words, their formulation of the  $n$ -player problem is via the open loop while ours is via the closed loop. Another crucial difference between the two limit results is that the solution  $(Y, Z)$  to (III.3.22) is adapted in the filtration generated by  $(p, \bar{X})$ , in particular, note that  $\bar{X}$  is of 1-dimension, while the solution  $\{(Y^{n,i}, Z^{n,i})\}_{i \in \mathbb{I}_n}$  to our  $n$ -player BSDE system is adapted in the filtration generated by  $X$ , the  $n$ -dimensional process.

— As we will show, the technique involved to prove Proposition III.3.9 is a combination of the BSDE estimates and the argument for proving the propagation of chaos. That is why we would name this section the backward propagation of chaos.

— Indeed the result of Proposition III.3.9 and the upcoming analysis hold valid for the broader class of BSDE systems in the form of

$$Y_t^{n,i} = \xi^i + \int_t^T F(t, Y^{n,i}, Z^{n,i}, p_t^n) dt - Z_t^{n,i} dX_t, \quad i \in \mathbb{I}_n,$$

where  $F$  is Lipschitz continuous in  $(y, z, p)$ . Apparently, this result could have an independent interest, once one wants to look into the asymptotic behaviour of solutions to such BSDE systems.

Before proving this main result, we first obtain the following estimates through some classical BSDE arguments.

**Lemma III.3.11.** *We have*

$$\mathbb{E} \left[ \sup_{t \leq s \leq T} |\Delta Y_s^i|^2 \right] \leq C \left( \mathbb{E} \left[ \int_t^T d_u^2(p^n, p^*) du \right] + \frac{1}{n^2} \right), \quad (\text{III.3.23})$$

$$\text{and} \quad \mathbb{E} \left[ \int_t^T |\Delta Z_u^i|^2 du \right] \leq C \left( \mathbb{E} \left[ \int_t^T d_u^2(p^n, p^*) du \right] + \frac{1}{n^2} \right). \quad (\text{III.3.24})$$

*Démonstration.* Comparing (III.3.19) and (III.3.20) we obtain

$$\begin{aligned} |\Delta Y_t^i| \leq & \frac{n}{n-1} \int_t^T \left| \int_{\mathbb{R}} c^*(y - Y_u^{n,i}) p_u^n(dy) - \int_{\mathbb{R}} c^*(y - Y_u^{*,i}) p_u^*(dy) \right| du + \frac{|c^*(0)|}{n-1} T \\ & + \frac{1}{n-1} \int_t^T \int_{\mathbb{R}} c^*(y - Y_u^{*,i}) p_u^*(dy) du + \left| \int_t^T \Delta Z_u^i \cdot (dX_u - e_i du) \right|. \end{aligned} \quad (\text{III.3.25})$$

Using the Kantorovich duality and Lipschitz continuity of  $c^*$ , we have

$$\begin{aligned} & \left| \int_{\mathbb{R}} c^*(y - Y_u^{n,i}) p_u^n(dy) - \int_{\mathbb{R}} c^*(y - Y_u^{*,i}) p_u^*(dy) \right| \\ & \leq \left| \int_{\mathbb{R}} c^*(y - Y_u^{n,i}) p_u^n(dy) - \int_{\mathbb{R}} c^*(y - Y_u^{n,i}) p_u^*(dy) \right| \\ & \quad + \left| \int_{\mathbb{R}} c^*(y - Y_u^{n,i}) p_u^*(dy) - \int_{\mathbb{R}} c^*(y - Y_u^{*,i}) p_u^*(dy) \right| \\ & \leq \mathcal{W}_1(p_u^n, p_u^*) + L |\Delta Y_u^i| \\ & \leq \mathcal{W}_2(p_u^n, p_u^*) + L |\Delta Y_u^i|. \end{aligned} \quad (\text{III.3.26})$$

Squaring and taking supremum and expectation on both sides of (III.3.25), and using Jensen and BDG-

inequality, we obtain

$$\begin{aligned}
 \mathbb{E} \left[ \sup_{t \leq s \leq T} |\Delta Y_s^i|^2 \right] &\leq C_T \mathbb{E} \left[ \int_t^T \mathcal{W}_2^2(p_u^n, p_u^*) du \right] + C_{L,T} \mathbb{E} \left[ \int_t^T |\Delta Y_u^i|^2 du \right] + \frac{C_T}{n^2} |c^*(0)|^2 \\
 &\quad + C_T \mathbb{E} \left[ \sup_{t \leq s \leq T} \left| \int_t^s \Delta Z_u^i \cdot (dX_u - e_i du) \right|^2 \right] \\
 &\quad + \frac{C_T}{n^2} \mathbb{E} \left[ \int_t^T \left| \int_{\mathbb{R}} c^*(y - Y_u^{*,i}) p_u^*(dy) \right|^2 du \right] \\
 &\leq C_T \mathbb{E} \left[ \int_t^T \mathcal{W}_2^2(p_u^n, p_u^*) du \right] + C_{L,T} \mathbb{E} \left[ \int_t^T |\Delta Y_u^i|^2 du \right] + \frac{C_T}{n^2} |c^*(0)|^2 \quad (\text{III.3.27}) \\
 &\quad + C_{T,BDG} \mathbb{E} \left[ \int_t^T |\Delta Z_u^i|^2 du \right] + \frac{C_T}{n^2} \mathbb{E} \left[ \int_t^T \left| \int_{\mathbb{R}} c^*(y - Y_u^{*,i}) p_u^*(dy) \right|^2 du \right].
 \end{aligned}$$

Further we shall estimate  $\mathbb{E} \left[ \int_t^T |\Delta Z_u^i|^2 du \right]$ . By Itô's formula,

$$\begin{aligned}
 |\Delta Y_t^i|^2 + \int_t^T |\Delta Z_u^i|^2 du &= \frac{2n}{n-1} \int_t^T \Delta Y_u^i \left( \int_{\mathbb{R}} c^*(y - Y_u^{n,i}) p_u^n(dy) - \int_{\mathbb{R}} c^*(y - Y_u^{*,i}) p_u^*(dy) \right) du \\
 &\quad - \frac{2c^*(0)}{n-1} \int_t^T \Delta Y_u^i du - 2 \int_t^T \Delta Y_u^i \Delta Z_u^i \cdot (dX_u - e_i du) \\
 &\quad + \frac{2}{n-1} \int_t^T \Delta Y_u^i \int_{\mathbb{R}} c^*(y - Y_u^{*,i}) p_u^*(dy) du.
 \end{aligned}$$

Together with (III.3.26) and Young's inequality, we obtain

$$\begin{aligned}
 \mathbb{E} \left[ \int_t^T |\Delta Z_u^i|^2 du \right] &\leq 2 \mathbb{E} \left[ \int_t^T |\Delta Y_u^i| \left( \mathcal{W}_2(p_u^n, p_u^*) + C_L |\Delta Y_u^i| \right) du \right] + \frac{c^*(0)^2}{(n-1)^2} T \\
 &\quad + 2 \mathbb{E} \left[ \int_t^T |\Delta Y_u^i|^2 du \right] + \frac{1}{(n-1)^2} \mathbb{E} \left[ \int_t^T \left| \int_{\mathbb{R}} c^*(y - Y_u^{*,i}) p_u^*(dy) \right|^2 du \right] \\
 &\leq (3 + 2C_L) \mathbb{E} \left[ \int_t^T |\Delta Y_u^i|^2 du \right] + \mathbb{E} \left[ \int_t^T \mathcal{W}_2^2(p_u^n, p_u^*) du \right] \\
 &\quad + \frac{c^*(0)^2}{(n-1)^2} T + \frac{1}{(n-1)^2} \mathbb{E} \left[ \int_t^T \left| \int_{\mathbb{R}} c^*(y - Y_u^{*,i}) p_u^*(dy) \right|^2 du \right]. \quad (\text{III.3.28})
 \end{aligned}$$

We now estimate the last term above,

$$\begin{aligned}
 \mathbb{E} \left[ \int_t^T \left| \int_{\mathbb{R}} c^*(y - Y_u^{*,i}) p_u^*(dy) \right|^2 du \right] &\leq \mathbb{E} \left[ \int_t^T \left( \int_{\mathbb{R}} (L|y| + L|Y_u^{*,i}| + |c^*(0)|) p_u^*(dy) \right)^2 du \right] \\
 &\leq \mathbb{E} \left[ \int_t^T \left( L\mathbb{E}[|Y_u^{*,i}|] + L|Y_u^{*,i}| + |c^*(0)| \right)^2 du \right] \\
 &\leq 6L^2 T \mathbb{E} \left[ \sup_{0 \leq u \leq T} |Y_u^{*,i}|^2 \right] + 3T |c^*(0)|^2,
 \end{aligned}$$

which is bounded from above by a constant  $C_0$ , independent of  $t$  and  $i$ , by the a priori estimate for  $Y^{*,i}$

in Lemma III.3.8. Together with (III.3.27) we obtain

$$\begin{aligned} \mathbb{E} \left[ \sup_{t \leq s \leq T} |\Delta Y_s^i|^2 \right] &\leq (C_T + C_{T,BDG}) \mathbb{E} \left[ \int_t^T \mathcal{W}_2^2(p_u^n, p_u^*) du \right] \\ &\quad + (C_{L,T} + (2C_L + 3)C_{T,BDG}) \int_t^T \mathbb{E} \left[ \sup_{u \leq s \leq T} |\Delta Y_s^i|^2 \right] du \\ &\quad + \frac{1}{n^2} (C_T + 4C_{T,BDG}T) c^*(0)^2 + \frac{1}{n^2} (C_T + 4C_{T,BDG}) C_0. \end{aligned}$$

Applying Grönwall inequality, we get

$$\mathbb{E} \left[ \sup_{t \leq s \leq T} |\Delta Y_s^i|^2 \right] \leq C \left( \mathbb{E} \left[ \int_t^T \mathcal{W}_2^2(p_u^n, p_u^*) du \right] + \frac{1}{n^2} \right) \leq C \left( \mathbb{E} \left[ \int_t^T d_u^2(p^n, p^*) du \right] + \frac{1}{n^2} \right),$$

that is, the estimate (III.3.23), for some constant depending on  $T$ ,  $L$ ,  $C_0$  and the constant from BDG inequality. Finally, the estimate (III.3.24) follows from (III.3.28) and (III.3.23). ■

*Proof of Proposition III.3.9*: Define the empirical measure,

$$\nu^n := \frac{1}{n} \sum_{i=1}^n \delta_{Y^{*,i}}.$$

The empirical measure  $\frac{1}{n} \sum_{i=1}^n \delta_{(Y^{n,i}, Y^{*,i})}$  is a coupling of the empirical measures  $p^n$  and  $\nu^n$ , so

$$d_t^2(p^n, \nu^n) \leq \frac{1}{n} \sum_{i=1}^n \sup_{t \leq u \leq T} |\Delta Y_u^i|, \quad a.s.$$

Together with (III.3.23), we obtain that

$$\mathbb{E} \left[ d_t^2(p^n, \nu^n) \right] \leq C \left( \mathbb{E} \left[ \int_t^T d_u^2(p^n, p^*) du \right] + \frac{1}{n^2} \right).$$

Apply the triangle inequality and the previous inequality to obtain

$$\begin{aligned} \mathbb{E} \left[ d_t^2(p^n, p^*) \right] &\leq 2\mathbb{E} \left[ d_t^2(p^n, \nu^n) \right] + 2\mathbb{E} \left[ d_t^2(\nu^n, p^*) \right] \\ &\leq 2C \mathbb{E} \left[ \int_t^T d_u^2(p^n, p^*) du \right] + 2\mathbb{E} \left[ d_t^2(\nu^n, p^*) \right] + \frac{2C}{n^2}. \end{aligned}$$

Using Grönwall's inequality we obtain

$$\mathbb{E} \left[ d_0^2(p^n, p^*) \right] \leq 2e^{2CT} \left( \mathbb{E} \left[ d_0^2(\nu^n, p^*) \right] + \frac{C}{n^2} \right).$$

Each  $\nu^n$  is the empirical measures of i.i.d. samples from the law  $p^*$ , (III.3.21) follows from the law of

large numbers. ■

### III.3.3.2 Principals' problem

For technical reasons, in this section we would consider the case where  $U$  is bounded and  $w$  taking values in a compact set in  $\mathbb{R}$ . Recall the third point in Remark III.3.4, under the above setting, we still have existence of the mean field games.

**Proposition III.3.12.** *Let all the  $n$  principals offer the contract provided by a mean field game, for principal  $i$  the reward becomes*

$$V_0^{n,i} = \int_K \lambda(dy_0) \mathbb{E}^{\mathbb{P}^{\alpha^*}} \left[ X_T^i - U(\xi^i) \mathbf{1}_{\{I_T=i\}} - \int_0^T U(w_u^i) \mathbf{1}_{\{I_u=i\}} du \right], \quad (\text{III.3.29})$$

where  $\alpha^*$  is the optimal intensity of the agent satisfying (III.2.9). Then, as  $n \rightarrow \infty$ ,  $V_0^{n,i}$  converges to the value of the mean field game.

*Démonstration.* The dynamic version of (III.3.29) reads

$$V_t^{n,i} = \mathbb{E}_t^{\mathbb{P}^{\alpha^*}} \left[ X_T^i - U(\xi^i) \mathbf{1}_{\{I_T=i\}} - \int_t^T U(w_u^i) \mathbf{1}_{\{I_u=i\}} du \right], \quad \text{for } 0 < t \leq T.$$

Since  $U$  is bounded

$$\mathbb{E}^{\mathbb{P}^{\alpha^*}} \left[ \sup_{0 \leq t \leq T} |V_t^{n,i}|^2 \right] < \infty. \quad (\text{III.3.30})$$

We denote  $V_t^{n,i,0} := V_t^{n,i} \Big|_{I_t \neq i}$  and  $V_t^{n,i,1} := V_t^{n,i} \Big|_{I_t = i}$  on different regimes. Recall  $\tau_t^n = \inf\{s \geq t : I_s \neq I_t\}$  and define  $\bar{\tau}_t^n := \inf\{s \geq t : I_s = i\}$ . By the tower property of the conditional expectation, we have

$$\begin{cases} V_t^{n,i,1} = \mathbb{E}_t^{\mathbb{P}^{\alpha^*}} \left[ V_{\tau_t^n}^{n,i,0} \mathbf{1}_{\{\tau_t^n \leq T\}} + (X_T^i - U(\xi^i)) \mathbf{1}_{\{\tau_t^n > T\}} - \int_t^{\tau_t^n \wedge T} U(w_u^i) du \right] \\ V_t^{n,i,0} = \mathbb{E}_t^{\mathbb{P}^{\alpha^*}} \left[ V_{\bar{\tau}_t^n}^{n,i,1} \mathbf{1}_{\{\bar{\tau}_t^n \leq T\}} + X_T^i \mathbf{1}_{\{\bar{\tau}_t^n > T\}} \right] = \mathbb{E}_t^{\mathbb{P}^{\alpha^*}} \left[ (V_{\bar{\tau}_t^n}^{n,i,1} - X_T^i) \mathbf{1}_{\{\bar{\tau}_t^n \leq T\}} \right] + X_t^i. \end{cases} \quad (\text{III.3.31})$$

It follows from (III.2.9) that  $\lim_{n \rightarrow \infty} \mathbb{P}^{\alpha^*}[\bar{\tau}_t^n \leq T] = 0$ . Together with (III.3.30), we obtain

$$\lim_{n \rightarrow \infty} \mathbb{E}^{\mathbb{P}^{\alpha^*}} \left[ |V_t^{n,i,0} - X_t^i|^2 \right] = 0. \quad (\text{III.3.32})$$

Define

$$\tilde{V}_0^{n,i,1} = \int_K \lambda(dy_0) \mathbb{E}^{\mathbb{P}^{\alpha^*}} \left[ \int_0^T \alpha_u^n \beta_u^n (X_u^i - U(w_u^i)) du + \beta_T^n (X_T^i - U(\xi^i)) \right],$$



where  $\alpha_t^n = \frac{1}{n-1} \sum_{j \neq i} a^*(Y_t^{n,j} - Y_t^{n,i})$  and  $\beta_t^n := \exp\left(-\int_0^t \alpha_u^n du\right)$ . Note that  $\alpha^*, \beta^*$  are bounded. By (III.3.32), we have

$$\lim_{n \rightarrow \infty} |\tilde{V}_0^{n,i,1} - V_0^{n,i,1}| = 0.$$

Further, on the regime  $I = i$ , the expectation  $\int_K \hat{\lambda}(dy_0) \mathbb{E}^{\alpha^*}[\cdot]$  coincides with  $\mathbb{E}(=\mathbb{E}^{\hat{\lambda}, \hat{\eta}})$ , so

$$\tilde{V}_0^{n,i,1} = \mathbb{E} \left[ \int_0^T \alpha_u^n \beta_u^n (X_u^i - U(w_u^i)) du + \beta_T^n (X_T^i - U(\xi^i)) \right].$$

Since the function  $a^*$  is bounded and Lipschitz continuous, we have

$$|\beta_t^n - \beta_t^*|^2 + |\alpha_t^n \beta_t^n - \alpha_t^* \beta_t^*|^2 \leq C \left( d_0^2(p^n, p^*) + |Y_t^{n,i} - Y_t^{*,i}|^2 \right).$$

Finally, the convergence result (III.3.21) and the estimate (III.3.23) imply that

$$\lim_{n \rightarrow \infty} \tilde{V}_0^{n,i,1} = \mathbb{E} \left[ \int_0^T \alpha_u^* \beta_u^* (X_u^i - U(w_u^i)) du + \beta_T^* (X_T^i - U(Y_T^{*,i})) \right],$$

and the latter is the value of the mean field game. ■

---

---

# CHAPITRE IV

---

## THE WELLPOSEDNESS OF PATH-DEPENDENT MULTIDIMENSIONAL FORWARD-BACKWARD SDE

**Abstract :** We study in this paper the wellposedness of path-dependent multidimensional forward-backward stochastic differential equations (FBSDE). By path-dependent we mean that the coefficients of the forward-backward SDE at time  $t$  can depend on the whole path of the forward process up to time  $t$ . These kinds of forward-backward SDE appear when solving path-dependent stochastic control problem by means of variational calculus. At the heart of our analysis is the construction of a decoupling random field on the path space. We first prove the existence and the uniqueness of decoupling field on small time interval. Then by introducing the characteristic BSDE, we show that a global decoupling field can be constructed by patching local solutions together as long as the solution of the characteristic BSDE remains bounded. Finally, we show that the solution of a path-dependent forward-backward SDE is stable.

**Keywords :** Decoupling Random Field, Forward-Backward SDE, Characteristic BSDE, Backward Stochastic Riccati Equations, Stability Theorem

## IV.1 Introduction

Forward-backward SDE appears naturally while solving stochastic control problems by means of variational calculus. Provided that an optimal control exists, the forward component describes the optimal state equation, while the backward component stands for the costate. Motivated by various applications in mathematical finance and their close links with the quasi-linear PDE, the wellposedness of forward-backward SDE (FBSDE) has been extensively studied during the past two decades. A FBSDE is a system of the form

$$\begin{cases} dX_t = b_t(X_t, Y_t, Z_t)dt + \sigma_t(X_t, Y_t, Z_t)dW_t & X_0 = x, \\ dY_t = -f_t(X_t, Y_t, Z_t)dt + Z_t dW_t & Y_T = g(X_T), \end{cases}$$

where  $X$  is called the forward process and  $(Y, Z)$  is called the backward process. The triple  $(X, Y, Z)$  can be multidimensional and the above notations represent in general a system of equations. Besides, the coefficients  $(b, \sigma, f, g)$  of the FBSDE can be random or deterministic.

The system is called decoupled if neither  $b$  nor  $\sigma$  depend on  $(Y, Z)$ , which has been studied by Pardoux and Peng in [130]. The decoupled problem reduces to the wellposedness of BSDE, which, under some standard assumptions, can be guaranteed.

In the case of strongly coupled FBSDE, the wellposedness is far more complex and counter-examples under very simple forms can be found : a FBSDE can either do not have any solution on large time interval or do not have any solution at all, either have an infinity of solutions, see Example IV.4.1.1. Fortunately in the case of regular deterministic non-degenerate coefficients, the four steps scheme method proposed by Ma, Protter and Yong in [109] ensures the existence and uniqueness of the solution of the FBSDE while allowing the time horizon  $T$  to be arbitrarily large. Moreover, the method relates the FBSDE to the following PDE

$$u_t + \frac{1}{2}\sigma^2(t, x, u)u_{xx} + b(t, x, u, \sigma(t, x, u)u_x)u_x + f(t, x, u, \sigma(t, x, u)u_x) = 0, \quad (\text{IV.1.1})$$

$$u(T, x) = g(x).$$

Under some mild assumptions, the solution of the FBSDE is explicitly given by

$$Y_t = u(t, X_t); \quad Z_t = \sigma(t, X_t, u(t, X_t))u_x(t, X_t).$$

In the non-Markovian FBSDE case, the method of continuation initially introduced by Hu and Peng [84], Peng and Wu [134], and later developed by Yong [159] and recently [160], has been widely used in various applications, see e.g. [161, 157]. However, the method depends crucially on the monotonicity conditions on the coefficients, which is restrictive in a different way comparing to four steps scheme. Later, using the notion of decoupling field, a general technique has been developed after a series of works of Cvitanic and Zhang [39], Zhang [163], Ma, Wu, Zhang, Zhang [110] and is used to extend the

contraction method proposed by Antonelli [4], Pardoux and Tang [131] to construct solutions on large intervals by patching together local solutions. The decoupling field  $u$  can be seen as an extension of the solution of the quasi-linear PDE (IV.1.1), which is the deterministic version of decoupling field. Generally speaking, the decoupling field is a random function such that the solutions of the FBSDE satisfy

$$Y_t := u_t(X_t). \quad (\text{IV.1.2})$$

The key idea is to transform the wellposedness of FBSDE problem to the existence and the uniqueness of the decoupling field, which are guaranteed by the combination of local wellposedness of FBSDE and the comparison principle of quadratic BSDE. The result was first applied to one-dimensional FBSDE in [110] then extended by Zhang in [165] to FBSDE with high dimensional backward process under even weaker conditions. Fromm and Imkeller in [60] redefined the decoupling field using weak derivatives and applied it to general multidimensional FBSDE and defined the notion of maximal interval of a FBSDE.

Note that all previous works have been focusing on FBSDEs where the coefficients depend only  $X_t$ , which are not sufficient for path-dependent stochastic control problems. For example in Principal-Agent problem, the contract given to the Agent by the Principal is typically path-dependent, see e.g. [41, 82]. The paper is largely inspired by the paper [110] and our main contribution is the extension of the existing results to the path-dependent FBSDE :

$$\begin{cases} dX_t = b_t(X, Y_t, Z_t)dt + \sigma_t(X, Y_t, Z_t)dW_t & X_0 = x, \\ dY_t = -f_t(X, Y_t, Z_t)dt + Z_t dW_t & Y_T = g(X). \end{cases}$$

By introducing a new metric on the path space (see Assumption IV.2.1), we first extend the local existence result of [4] to path-dependent multidimensional FBSDEs. We then define the characteristic BSDE as in the classical case and by the same spirit of [110, 165], we construct the decoupling field on the path space using the notion of dominating ODE and the comparison principle of quadratic BSDE. Finally we give the stability property of path-dependent FBSDEs.

Note that in the general literature, the FBSDE whose coefficients depend on  $\omega \in \Omega$  are referred to as non-Markovian FBSDE. In order to distinguish our framework with the classical ones and avoid confusion, we call throughout the paper FBSDE whose coefficients depend on  $X_{\cdot \wedge t}$  at time  $t$  path-dependent FBSDE. All the results of path-dependent FBSDE can naturally be applied to classical framework.

The paper is structured as follow. In Section IV.2, we shall introduce the setting of our problem, define the notion of decoupling field for path-dependent FBSDEs and discuss some basic properties and show how they lead to the wellposedness of FBSDEs. In Section IV.4, we shall prove the local existence and uniqueness of strongly coupled path-dependent FBSDEs. Then we will study the dynamics of the decoupling field and investigate the global solutions of FBSDEs using dominating ODE. In Section IV.5, we will give a stability result for path-dependent multidimensional FBSDEs.

## IV.2 Notations and General Assumptions

Throughout this paper, we denote  $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$  a filtered probability space on which is defined a  $n$  dimensional Brownian motion  $W = (W_t)_{t \geq 0}$ . Denote  $\mathbb{F} := \{\mathcal{F}_t^W\}_{t \geq 0}$ , the natural filtration generated by  $W$ , augmented by the  $\mathbb{P}$ -null sets of  $\mathcal{F}$ .

For  $t \geq 0$ , denote  $\mathbb{H}_t^2(\mathbb{F}, \mathbb{R}^n)$  the space of continuous  $\mathbb{F}$ -adapted processes on  $[0, t]$  taking values in  $\mathbb{R}^n$  satisfying the integrability condition

$$\|Y\|_{\mathbb{H}_t^2} := \mathbb{E} \left[ \left( \int_0^t |Y_s|^2 ds \right)^{\frac{1}{2}} \right] < +\infty,$$

where  $|\cdot|$  is the Euclidean distance. Let  $T > 0$  be a fixed time horizon. We consider the following fully coupled forward-backward stochastic differential equation (abbreviated FBSDE) on  $[0, T]$  :

$$\begin{cases} dX_t = b_t(X, Y_t, Z_t)dt + \sigma_t(X, Y_t, Z_t)dW_t & X_0 = x, \\ dY_t = -f_t(X, Y_t, Z_t)dt + Z_t dW_t & Y_T = g(X). \end{cases} \quad (\text{IV.2.3})$$

$$(\text{IV.2.4})$$

Throughout the paper, we shall make use of the following standard Lipschitz assumptions.

**Assumption IV.2.1.** (i) Let  $\Theta := \mathcal{C}([0, T], \mathbb{R}^d) \times \mathbb{R}^n \times \mathcal{M}_n(\mathbb{R})$ , the coefficients of the FBSDE

$$b : [0, T] \times \Omega \times \Theta \rightarrow \mathbb{R}^d, f : [0, T] \times \Omega \times \Theta \rightarrow \mathbb{R}^n, \sigma : [0, T] \times \Omega \times \Theta \rightarrow \mathcal{M}_{d,n}(\mathbb{R})$$

are  $\mathbb{F}$ -progressively measurable, for fixed  $\theta = (x, y, z) \in \Theta$ ; the function

$$g : \mathcal{C}([0, T], \mathbb{R}^d) \times \Omega \rightarrow \mathbb{R}^n$$

is  $\mathcal{F}_T$ -measurable, for fixed  $x \in \mathcal{C}([0, T], \mathbb{R}^d)$ .

(ii) Denote  $\xi_t^0 := \xi_t(\bar{0}, 0, 0)$  for  $\xi = b, f, \sigma$  and  $g^0 := g(\bar{0})$ , where  $\bar{0}$  is the function on  $[0, T]$  constantly equals to 0. The following integrability condition holds :

$$\mathbb{E} \left[ \left( \int_0^T (|b_t^0| + |f_t^0|) dt \right)^2 + \int_0^T |\sigma_t^0|^2 dt + |g^0|^2 \right] < \infty. \quad (\text{IV.2.5})$$

(iii) The coefficients  $b, \sigma, f$  satisfy the following Lipschitz condition in the spacial variable : there exists  $K_0 > 0$  such that for  $\theta = (\omega, x, y, z)$  and  $\theta' = (\omega, x', y', z')$ ,

$$|\xi_t(\theta) - \xi_t(\theta')| \leq K_0 (\|x - x'\|_{2,t} + |y - y'| + |z - z'|), \quad (\text{IV.2.6})$$

uniformly in  $\omega \in \Omega$ , where  $\xi = b, \sigma, f$  and  $\|x - x'\|_{2,t}^2 := \int_0^t |x(s) - x'(s)|^2 ds + |x(t) - x'(t)|^2$ .

In particular, denote  $|\nabla_z \sigma|_\infty$  Lipschitz coefficient of the function  $\sigma$  with respect to  $z \in \mathcal{M}_n(\mathbb{R})$ .

(iv) The function  $g$  satisfies the following Lipschitz condition : there exists  $K_1 > 0$  such that for all  $x, x' \in \mathcal{C}([0, T], \mathbb{R}^d)$ , we have

$$|g(x) - g(x')| \leq K_1 \|x - x'\|_{2,T} \quad (\text{IV.2.7})$$

uniformly in  $\omega \in \Omega$ .

**Remark IV.2.2.** In the classical literature, Markovian FBSDEs refer to systems whose the coefficients are deterministic and non-Markovian FBSDEs refer to systems with random coefficients, i.e. coefficients depending on the Brownian motion. In both cases, the coefficients of the FBSDE at time  $t$  depend only on the value at time  $t$  of the forward process  $X$ . It is worth noting that the aforementioned cases are covered by our setting.

### IV.3 Decoupling Field for Multidimensional Path-Dependent FBSDE

A general technique for solving a FBSDE, initiated by Protter, Ma and Yong in [109] then further developed by various authors in [111], [112], [131], [43],[39],[163], [113], [110], consists in finding a decoupling function  $u$  such that the solution of the FBSDE satisfies the relation

$$Y_t := u_t(X_t).$$

In the case where the coefficients of the FBSDE, namely  $b, f, \sigma$  are deterministic and Markovian, it is shown that  $u$  is related to the solution of the quasilinear PDE

$$u_t + \frac{1}{2} \sigma^2(t, x, u) u_{xx} + b(t, x, u, \sigma(t, x, u) u_x) u_x + f(t, x, u, \sigma(t, x, u) u_x) = 0,$$

either in the classical sense or in the viscosity sense. In the case where the coefficients are allowed to be random, the function  $u$  is shown to be the solution of some backward stochastic PDE or is constructed as a random field using localization technique under certain conditions of the coefficients. In this section, we will extend the notion of decoupling field to path-dependent FBSDE. More precisely, we have the following definition of the decoupling field.

**Definition IV.3.1.** An  $\mathbb{F}$ -progressively measurable random field  $u : [0, T] \times \Omega \times \mathcal{C}([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}^n$  with  $u(T, \omega, x) = g(\omega, x)$  is said to be a decoupling field of FBSDE (IV.2.3)-(IV.2.4) if there exists a constant  $\delta > 0$  such that, for any  $0 \leq t_1 < t_2 \leq T$  with  $t_2 - t_1 \leq \delta$  and any  $x \in \mathbb{H}_{t_1}^2(\mathbb{F}, \mathbb{R}^d)$ , the FBSDE (IV.2.3)-(IV.2.4) with initial value  $X_{\cdot \wedge t_1} = x$  and terminal condition  $Y_{t_2} = u(t_2, \cdot)$  has an unique solution that satisfies

$$Y_t = u(t, \omega, X_{\cdot \wedge t}) = u(t, \omega, X), \quad t \in [t_1, t_2], \quad \mathbb{P} - a.s.$$

A decoupling field  $u$  is called regular if it is Lipschitz with respect to  $x$  : there exists  $K > 0$  such that

for all  $x, x' \in \mathcal{C}([0, T], \mathbb{R}^d)$ ,  $\forall t \in [0, T]$ , we have

$$|u(t, \omega, x) - u(t, \omega, x')| \leq K \|x - x'\|_{2,t}, \mathbb{P} - a.s.$$

For notation simplicity, denote  $u_t(X) := u(t, \omega, X)$ .

Note that the existence of decoupling fields implies the well-posedness of FBSDE over a small time interval. The following result shows the implication of the existence of decoupling fields for the well-posedness of FBSDEs over an arbitrary duration.

**Theorem IV.3.2.** *Assume that Assumption IV.2.1 holds, and that there exists a decoupling field  $u$  for the FBSDE (IV.2.3)-(IV.2.4). Then the FBSDE (IV.2.3)-(IV.2.4) has a unique solution  $(X, Y, Z)$  and*

$$\forall t \in [0, T], Y_t = u_t(X).$$

The theorem is a direct generalisation of [110, Theorem 2.3]. For the readers' convenient, we shall detail the proof in Section IV.6.

## IV.4 Main Results

### IV.4.1 Wellposedness of FBSDE on a Small Time Interval

Our starting point to construct a decoupling field will be the local existence and the uniqueness of the solution of FBSDEs. The local existence of non-Markovian FBSDE is a well-understood problem using fixed point approach, see for example in the book of Cvitanic and Zhang [41]. The following Theorem generalizes the local existence result to path-dependent FBSDEs.

**Theorem IV.4.1.** *Under Assumption IV.2.1, if  $K_1 |\nabla_z \sigma|_\infty < 1$ , then there exists  $\delta > 0$  such that for all  $T < \delta$ , the FBSDE (IV.2.3)-(IV.2.4) has an unique solution  $(X, Y, Z)$ .*

Since the proof of Theorem IV.4.1 is quite lengthy and technical, we shall detail it in Section IV.6.

**Example IV.4.1.1.** *One can show that condition  $K_1 |\nabla_z \sigma|_\infty < 1$  cannot be weakened to a larger bound. Generally speaking, neither the uniqueness nor the existence can be guaranteed if  $K_1 |\nabla_z \sigma|_\infty \geq 1$ . To see this, here is an example given in [43] : let  $k$  be a constant and consider the following FBSDE*

$$\begin{cases} dX_t = (k + Z_t)dW_t & X_0 = x \\ dY_t = Z_t dW_t & Y_T = X_T . \end{cases}$$

Notice that

$$Y_t = X_T - \int_t^T Z_s ds = X_t + k(W_T - W_t). \tag{IV.4.8}$$

In particular, if  $k \neq 0$ , we have  $Y_0 - x = kW_T$ , which is impossible. In the case where  $k = 0$ , the corresponding FBSDE has an infinite number of solutions.

- Remark IV.4.2.**
1. In the case where  $\sigma$  does not depend on  $Z$ , the condition  $K_1|\nabla_z\sigma|_\infty < 1$  is automatically satisfied. This corresponds to the control problems where only the drift of the diffusion is controlled.
  2. The length of the interval given by Theorem IV.4.1 depends on the parameters of the FBSDE, more specifically  $K_0$ ,  $K_1$  and the product  $K_1|\nabla_z\sigma|_\infty$ . The larger the Lipschitz constants  $K_0$  and  $K_1$ , the smaller the interval on which Theorem IV.4.1 applies.
  3. If the forward and backward process are one dimensional and if  $z \rightarrow \sigma_t(x, y, z)$  and  $x \rightarrow g(x)$  are both invertible, the local existence of non-Markovian FBSDE in the case  $K_1|\nabla_z\sigma|_\infty > 1$  can be proved by using a time inversion technique, see Theorem 6.2 in [110].
  4. The local existence and uniqueness of the solution of FBSDE given in Theorem IV.4.1 provides a way to construct the decoupling field on a small time interval. More precisely, consider the FBSDE (IV.2.3)-(IV.2.4) on  $[0, T]$  satisfying the assumptions in Theorem IV.4.1. For any  $t \in [0, T]$ , for any  $x \in \mathcal{C}([0, t], \mathbb{R}^d)$ , the same FBSDE on  $[t, T]$  with initial condition  $X_{t_1} = x$  and terminal condition  $Y_T = g(X)$  still has a unique solution. Let  $u_t(x) := Y_t$ . One can check easily that  $u$  is the unique associated decoupling field.

## IV.4.2 Existence of Regular Decoupling Field

In this section, we will try to extend the local existence result to larger time interval. There are two important ingredients in Theorem IV.4.1 which guarantee the local existence : the Lipschitz constant of the terminal condition must be small than  $|\nabla_z\sigma|_\infty^{-1}$  and the length of the time interval has to be smaller than a constant  $\delta_0$  determined by the coefficients of the FBSDE.

The strategy of constructing a decoupling field on a larger time interval is the following : first we construct a decoupling field  $u$  on  $[T - \delta_0, T]$  by solving locally the FBSDE and we estimate the upper bound of the gradient of the decoupling field  $u$  with respect to the forward process, which is denoted  $K_{T-\delta_0}$  and will be used as the Lipschitz constant of the terminal condition when we then try to construct a decoupling field on  $[T - \delta_0 - \delta_1, T - \delta_0]$  for some  $\delta_1$ . If the condition  $K_{T-\delta_0}|\nabla_z\sigma|_\infty < 1$  is still satisfied at  $T - \delta_0$ , we can proceed the same way and extend the local existence and uniqueness step by step until the whole interval is covered.

Notice that obtaining a bound on gradient of the decoupling field  $u$  with respect to the forward process is equivalent to find an upper bound of the corresponding variational FBSDE, which will be defined in the following Section IV.4.2.1. The technique consists in using the comparison principle of quadratic BSDE given by Kobylanski in [97] to find a dominating ODE, the solution of which, if exists on  $[0, T]$ , dominates the variational FBSDE. This method is used in [110] in the context of one-dimensional non-Markovian FBSDE then generalized in [165] to the case where the backward process is multidimensional. In this



subsection, we generalize the existing results to different types of path-dependent FBSDEs. We shall begin by showing the wellposedness of decoupled FBSDE in Section IV.4.2.2 as a direct consequence of the existence of dominating ODE. In Section IV.4.2.3, we shall study the case where  $b_t(X, Y_t, Z_t) = b_t(X, Y_t)$  and  $\sigma_t(X, Y_t, Z_t) = \sigma_t(X)$ . These are the only cases where global existence can be guaranteed for arbitrary duration under some conditions. The case where  $\nabla_z \sigma = 0$  is discussed in Section IV.4.2.4 and the general case at Section IV.4.2.5. In both cases, the corresponding dominating ODE is a Riccati equation and we introduce the notion of maximal interval as in [60].

#### IV.4.2.1 Variational FBSDE, Characteristic BSDE and Dominating ODE

**Notation.** (i) For  $i \in \llbracket 1, d \rrbracket$ , denote  $X_i$  the  $i$ -th component of  $X$ . The corresponding 1-dimensional forward equation is

$$dX_{i,t} = b_{i,t}(X, Y_t, Z_t)dt + \sigma_{i,t}^\top(X, Y_t, Z_t)dW_t,$$

where  $b_{i,t}$  and  $\sigma_{i,t}$  is the  $i$ -th component of the vector  $b_t$  and the  $i$ -th row of the matrix  $\sigma_t$ , respectively.

(ii) For  $i \in \llbracket 1, n \rrbracket$ , denote  $Y_i$  (respectively  $Z_i$ ) the  $i$ -th component of  $Y$  (respectively the  $i$ -th row of  $Z$ ). The corresponding component-wise backward equation is

$$dY_{i,t} = -f_{i,t}(X, Y_t, Z_t)dt + Z_{i,t}^\top dW_t,$$

where  $f_{i,t}$  is the  $i$ -th component of the vector  $f_t$ .

(iii) For  $\theta, \theta' \in \Theta$ , denote  $\Delta\theta = (\Delta x, \Delta y, \Delta z) := \theta - \theta'$ . For  $\xi = b, \sigma, f$ , denote

$$\xi_{x,t}(\theta, \theta') := \frac{\xi(x, y, z) - \xi(x', y', z')}{\|\Delta x\|_{2,t}}; \quad (\text{IV.4.9})$$

$$\xi_{y,t}(\theta, \theta') := \left( \frac{\xi(x', y'_1, \dots, y'_{k-1}, y_k, \dots, z) - \xi(x', y'_1, \dots, y'_k, y_{k+1}, \dots, z)}{y_k - y'_k} \right)_{k \in \llbracket 1, n \rrbracket};$$

$$\xi_{z,t}(\theta_1, \theta_2) := \left( \frac{\xi(x', y', z'_{1,1}, \dots, z'_{k,l}, \dots, z_{n,n}) - \xi(x', y', z'_{1,1}, \dots, z'_{k,l}, \dots, z_{1,n,n})}{z_{k,l} - z'_{k,l}} \right)_{k,l \in \llbracket 1, n \rrbracket};$$

where  $y_k$  is the  $k$ -th component of the vector  $y$  and  $z_{k,l}$  is the value of the component at the position  $(k, l)$  of the matrix  $z$ . Here and in the sequel, for any Lipschitz continuous function  $\xi(x)$ , when  $x = x'$ , we will always take the convention that

$$\frac{\xi(x) - \xi(x)}{x - x} := \lim_{\tilde{x} \rightarrow x} \frac{\xi(\tilde{x}) - \xi(x)}{\tilde{x} - x}.$$

Under the above notation, for  $\xi = b, \sigma, f$ , we have

$$\xi_t(\theta) - \xi_t(\theta') = \xi_{x,t} \|\Delta x\|_{2,t} + \xi_{y,t} \Delta y_t + \text{Tr}[\xi_{z,t} \Delta z_t].$$

For notation simplicity, we shall omit the index  $t$  in the following.

Let  $X_{\cdot, \wedge t_0}, \tilde{X}_{\cdot, \wedge t_0}$  two processes on  $\mathbb{H}_{t_0}^2(\mathbb{F}, \mathbb{R}^d)$ . By Theorem IV.4.1, the FBSDE (IV.2.3)-(IV.2.4) has an unique solution  $(X, Y, Z)_{[t_0, T]}$  (respectively  $(\tilde{X}, \tilde{Y}, \tilde{Z})_{[t_0, T]}$ ) on  $[t_0, T]$  given the initial condition  $X_{\cdot, \wedge t_0}$  (respectively  $\tilde{X}_{\cdot, \wedge t_0}$ ) if  $T - t_0$  is small enough. Denote  $X$  (respectively  $\tilde{X}$ ) the concatenation of  $X_{\cdot, \wedge t_0}$  and  $X_{[t_0, T]}$  (respectively  $\tilde{X}_{\cdot, \wedge t_0}$  and  $\tilde{X}_{[t_0, T]}$ ).

Let  $\mathcal{X} := X - \tilde{X}$ ,  $\mathcal{Y} := Y - \tilde{Y}$  and  $\mathcal{Z} := Z - \tilde{Z}$ . We can check easily that  $(\mathcal{X}, \mathcal{Y}, \mathcal{Z})$  is a solution of the following variational FBSDE associated to the original FBSDE (IV.2.3)-(IV.2.4) on the time interval  $[t_0, T]$ :

$$\begin{cases} d\mathcal{X}_t = (b_x \mathcal{D}_t + b_y \mathcal{Y}_t + \text{Tr}(b_z \mathcal{Z}_t)) dt + (\sigma_x \mathcal{D}_t + \sigma_y \mathcal{Y}_t + \text{Tr}(\sigma_z \mathcal{Z}_t)) dW_t & \mathcal{X}_{t_0} = x_0, & \text{(IV.4.10)} \\ d\mathcal{Y}_t = (f_x \mathcal{D}_t + f_y \mathcal{Y}_t + \text{Tr}(f_z \mathcal{Z}_t)) dt - \mathcal{Z}_t dW_t & \mathcal{Y}_T = \Delta g, & \text{(IV.4.11)} \end{cases}$$

where  $\mathcal{D}_t := \|\mathcal{X}\|_{2,t}$  and  $\Delta g := g(X) - g(\tilde{X})$ . Define

$$\mathcal{H}_t := \frac{\mathcal{Y}_t^\top \mathcal{Y}_t}{\mathcal{D}_t^2}, \alpha_t := \frac{\mathcal{Z}_t}{\mathcal{D}_t}, \beta_t := \frac{\mathcal{X}_t}{\mathcal{D}_t}, P_t := \frac{\mathcal{Y}_t}{|\mathcal{Y}_t|},$$

$$d\tilde{W}_t := dW_t - 2(\sigma_x + \sigma_y P_t \sqrt{\mathcal{H}_t} + \text{Tr}(\sigma_z \alpha_t))^\top \beta_t dt.$$

Then it follows from Itô's formula that

$$d\mathcal{H}_t = -\mathcal{F}_t(\mathcal{H}_t) dt + \mathcal{N}_t d\tilde{W}_t \text{ with } \mathcal{F}_t(h) = A_t h^2 + B_t h^{3/2} + C_t h + D_t h^{1/2} + F_t, \quad \text{(IV.4.12)}$$

$$A_t = \text{Tr}(\sigma_y P_t P_t^\top \sigma_y^\top) - 8\beta_t^\top P_t^\top \sigma_y^\top \sigma_y P_t \beta_t,$$

$$B_t = 2\beta_t^\top b_y P_t + 2\text{Tr}(\sigma_y P_t (\sigma_x + \text{Tr}(\sigma_z \alpha_t))^\top) - 16\beta_t^\top P_t^\top \sigma_y^\top (\sigma_x + \text{Tr}(\sigma_z \alpha_t)) \beta_t,$$

$$C_t = 2P_t^\top f_y P_t + |\beta_t|^2 + 2\beta_t^\top (b_x + \text{Tr}(b_z \alpha_t)) + \text{Tr}(\sigma_x \sigma_x^\top) - 8\beta_t^\top \sigma_x^\top \sigma_x \beta_t,$$

$$D_t = 2P_t f_x + 2P_t^\top \text{Tr}(f_z \alpha_t), F_t = -\text{Tr}(\alpha_t \alpha_t^\top),$$

$$\mathcal{N}_t = 2\mathcal{H}_t^{1/2} P_t^\top \alpha_t - 2\mathcal{H}_t \beta_t^\top (\sigma_x + \sigma_y P_t \mathcal{H}_t^{1/2} + \text{Tr}(\sigma_z \alpha_t)).$$

We call the equation (IV.4.12) the characteristic BSDE of the FBSDE.

**Definition IV.4.3.** Let  $G : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$  be a continuously differentiable function. The ODE

$$\dot{y}_t = -G(t, y_t)$$

is called a dominating ODE of the FBSDE (IV.2.3)-(IV.2.4) if  $G$  satisfies the following conditions :

- (i) for all  $t \in [0, T]$ , we have  $\mathcal{F}_t(\cdot) \leq G(t, \cdot)$ ,  $\mathbb{P}$ -almost surely;
- (ii) for all  $M > 0$  there exists  $l, \hat{l} \in L^1([0, T], \mathbb{R})$  such that for all  $t \in [0, T]$ ,  $h \in [-M, M]$ , we have

$$|G(t, h)| \leq l(t) \text{ and } \left| \frac{\partial G}{\partial h}(t, h) \right| \leq \hat{l}(t).$$

The following proposition is a direct adaptation from [110, Theorem 2.3] motivation of finding a dominating ODE.

**Proposition IV.4.4.** *Assume that Assumption IV.2.1 holds true and that there exists a continuously differentiable function  $G : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$  such that  $\dot{y}_t = -G(t, y_t)$  is a dominating ODE of the FBSDE (IV.2.3)-(IV.2.4). If the ODE has a bounded solution on  $[0, T]$ , then the FBSDE has a unique regular decoupling field on  $[0, T]$  and therefore, it is well-posed.*

**Proof of Proposition IV.4.4.** Let  $T > 0$ . Let  $y$  be the solution of the dominating ODE :

$$\dot{y}_t = -G(t, y_t), \quad y_T = K_1^2. \quad (\text{IV.4.13})$$

Denote  $K_{max}^2 := \max_{t \in [0, T]} y_t$ , the upper bound of  $y$ . Firstly, by the comparison principle, we have  $\mathcal{H}_t \leq y_t \leq K_{max}^2$  for  $t \in [T - \epsilon_0, T]$  where  $\epsilon_0$  is a constant depending on  $K_0, K_1, n$  and  $d$  given by Theorem IV.4.1. This can be reformulated as follow using the decoupling field : for all  $t \in [T - \epsilon, T]$ , for all given initial condition for the forward process  $X_{T-\epsilon_0} = x \in \mathbb{H}_{T-\epsilon_0}^2(\mathbb{F}, \mathbb{R}^d)$ ,  $\mathbb{P}$ -almost surely,

$$|u(t, X) - u(t, X')|^2 \leq y_t \|X - X'\|_{2,t}^2 \leq K_{max}^2 \|X - X'\|_{2,t}^2. \quad (\text{IV.4.14})$$

To finish the proof, we only need to repeat the same procedure at  $T - \epsilon_0$  and so on. Again by Theorem IV.4.1, we can find  $\epsilon_1$  such that the FBSDE has a unique solution on  $[T - (\epsilon_1 + \epsilon_0), T - \epsilon_0]$  and for all  $t \in [T - (\epsilon_1 + \epsilon_0), T - \epsilon_0]$ , for all given initial condition for the forward process  $X_{T-\epsilon_0-\epsilon_1} = x \in \mathbb{H}_{T-\epsilon_0-\epsilon_1}^2(\mathbb{F}, \mathbb{R}^d)$ ,

$$|u(t, X) - u(t, X')|^2 \leq y_t \|X - X'\|_{2,t}^2 \leq K_{max}^2 \|X - X'\|_{2,t}^2. \quad (\text{IV.4.15})$$

Notice since  $K_{max}$  dominates the Lipschitz constants of the decoupling field  $u(t, \cdot)$  for all  $t \in [0, T]$ , we can choose each  $\epsilon_i \geq \bar{\epsilon}$  where  $\bar{\epsilon}$  is a constant given by Theorem IV.4.1 when applied to a FBSDE with Lipschitz constant  $K_0$  and  $K_{max}$ . Therefore, by iterating at most  $T/\bar{\epsilon}$ , we construct a decoupling field for the FBSDE (IV.4.16)-(IV.4.17) on  $[0, T]$  and by Theorem IV.3.2, the FBSDE has an unique solution.

■

### IV.4.2.2 Decoupled Path-dependent FBSDE

Consider the following decoupled path-dependent FBSDE :

$$\begin{cases} dX_t = b_t(X)dt + \sigma_t(X)dW_t & X_0 = x \end{cases} \quad (\text{IV.4.16})$$

$$\begin{cases} dY_t = -f_t(X, Y_t, Z_t)dt + Z_t dW_t & Y_T = g(X) . \end{cases} \quad (\text{IV.4.17})$$

The decoupled FBSDEs are always wellposed under standard Lipschitz assumptions because one can always solve independently the forward process then inject the solution into the backward equation and solve it as a standard BSDE. Another way to prove the wellposedness is to show the existence of a unique decoupling field of the FBSDE, which, in the decoupled case, is guaranteed by Proposition IV.4.4. More precisely, in this case we have  $A_t = B_t = 0$  and the characteristic BSDE (IV.4.12) becomes

$$d\mathcal{H}_t = -\left(C_t\mathcal{H}_t + D_t\mathcal{H}_t^{1/2} + F_t\right)dt + \mathcal{N}_t d\tilde{W}_t.$$

One can find a linear dominating ODE for the FBSDE (IV.4.16)-(IV.4.17), which has a bounded solution on  $[0, T]$ .

**Proposition IV.4.5.** *Consider the decoupled path-dependent FBSDE (IV.4.16)-(IV.4.17). Under Assumptions IV.2.1, for all  $T > 0$ , the equation (IV.4.16)-(IV.4.17) has a unique solution on  $[0, T]$ .*

### IV.4.2.3 When $b = b_t(x, y)$ and $\sigma = \sigma_t(x)$

When we use variational calculus in a stochastic control problem where only the drift of the state process is affected by the control, the optimal control given by the first order condition depends only on the value of the costate, i.e. the process  $Y$  but not  $Z$ . Therefore, the associated FBSDE satisfies  $b = b_t(x, y)$  and  $\sigma = \sigma_t(x)$ . Consider the following coupled path-dependent FBSDE :

$$\begin{cases} dX_t = b_t(X, Y_t)dt + \sigma_t(X)dW_t & X_0 = x \end{cases} \quad (\text{IV.4.18})$$

$$\begin{cases} dY_t = -f_t(X, Y_t, Z_t)dt + Z_t dW_t & Y_T = g(X) . \end{cases} \quad (\text{IV.4.19})$$

In this case,  $A_t = 0$  and the characteristic BSDE (IV.4.12) for the FBSDE (IV.4.18)-(IV.4.19) becomes

$$d\mathcal{H}_t = -\left(B_t\mathcal{H}_t^{3/2} + C_t\mathcal{H}_t + D_t\mathcal{H}_t^{1/2} + F_t\right)dt + \mathcal{N}_t d\tilde{W}_t. \quad (\text{IV.4.20})$$

The following theorem gives a condition under which the existence of a dominating ODE associated with (IV.4.18)-(IV.4.19) on  $[0, T]$  is guaranteed.

**Theorem IV.4.6.** *Let  $T > 0$ . Consider the coupled path-dependent FBSDE (IV.4.18)-(IV.4.19). Define the set*

$$\mathcal{B}_y := \left\{ b_y(t, x, y_1, x, y_2) \in \mathcal{M}_{d,n}(\mathbb{R}) \text{ for } (t, x, y_1, y_2) \in [0, T] \times \mathcal{C}([0, T], \mathbb{R}^d) \times (\mathbb{R}^n)^2, y_1 \neq y_2 \right\}.$$

Under Assumption IV.2.1, if for all  $\theta_i = (x_i, y_i, z_i) \in \mathcal{C}([0, T], \mathbb{R}^d) \times \mathbb{R}^n \times \mathcal{M}_n(\mathbb{R})$ ,  $i = 1, 2$ , for all  $t \in [0, T]$  and  $b_y \in \mathcal{B}_y$ , we have

$$\begin{aligned} & (b_t(\theta_1) - b_t(\theta_2))^\top b_y \Delta y - (f_t(\theta_1) - f_t(\theta_2))^\top b_y^\top \Delta x_t \\ & \quad + \text{Tr}((\sigma_t(\theta_1) - \sigma_t(\theta_2))^\top b_y \Delta z) \geq (g(x_1) - g(x_2))^\top b_y^\top x_T, \end{aligned} \quad (\text{IV.4.21})$$

then the FBSDE (IV.4.18)-(IV.4.19) has an unique solution on  $[0, T]$ .

**Proof of Theorem IV.4.6.** Using the definition of  $\beta$  and  $P$ , by Itô's formula, we have

$$\begin{aligned} \mathcal{H}_t(\beta_t^\top b_y P_t \sqrt{\mathcal{H}_t}) &= \mathcal{H}_t \frac{\mathcal{X}_t^\top b_y \mathcal{Y}_t}{\mathcal{D}_t^2} \\ &= \frac{\mathcal{H}_t}{\mathcal{D}_t^2} \mathbb{E}_t \left[ \Delta X_T^\top b_y \Delta g - \int_t^T (\Delta b_s^\top b_y \mathcal{Y}_s - \mathcal{X}_s^\top b_y \Delta f_s + \text{Tr}(\Delta \sigma_s^\top b_y \mathcal{Z})) ds \right], \end{aligned}$$

which is smaller than 0 by (IV.4.21). Since  $B_t$  is non-positive, one can find two constants  $c$  and  $d$  such that

$$\dot{y}_t = -cy - d$$

is a dominating ODE for the BSDE (IV.4.20). Since the above ODE has a bounded solution on  $[0, T]$  for all  $T > 0$ , by Proposition IV.4.4, the FBSDE (IV.4.18)-(IV.4.19) has a unique solution.  $\blacksquare$

**Remark IV.4.7.** In the one-dimensional case, if  $b$  is increasing in  $y$ , we can have the following sufficient condition for (IV.4.21), which is easier to verify.

$$\begin{cases} (b_t(\theta_1) - b_t(\theta_2)) \Delta y - (f_t(\theta_1) - f_t(\theta_2)) \Delta x_t + (\sigma_t(\theta_1) - \sigma_t(\theta_2)) \Delta z \geq 0; \\ (g(x_1) - g(x_2)) \Delta x_T \leq 0. \end{cases}$$

Similarly, if  $b$  is decreasing in  $y$ , we have

$$\begin{cases} (b_t(\theta_1) - b_t(\theta_2)) \Delta y - (f_t(\theta_1) - f_t(\theta_2)) \Delta x_t + (\sigma_t(\theta_1) - \sigma_t(\theta_2)) \Delta z \leq 0; \\ (g(x_1) - g(x_2)) \Delta x_T \geq 0. \end{cases}$$

This condition shares the same spirit as the monotonicity condition in the continuation method for solving one-dimensional Markovian framework FBSDE introduced by Hu and Peng in [84]. More details on the continuation methods for solving Markovian FBSDE can be found for example in [41, Section 11.4].

#### IV.4.2.4 When $\sigma = \sigma_t(x, y)$

Under Assumption IV.2.1, we know that even in the Markovian case, a FBSDE satisfying the hypothesis in Theorem IV.4.1 may not have a solution on arbitrary time interval.

**Example IV.4.7.1.** Here is an example given in [60]. Consider the following fully coupled FBSDE :

$$\begin{cases} dX_t = Y_t dt & X_0 = x \\ dY_t = Z_t dW_t & Y_T = X_T. \end{cases} \quad \begin{matrix} \text{(IV.4.22)} \\ \text{(IV.4.23)} \end{matrix}$$

We notice that the condition in Theorem IV.4.1 is satisfied in this case. Clearly for  $T < 1$  the problem has a unique decoupling field

$$u(t, x) = \frac{x}{1 - (T - t)}$$

and we have

$$X_t = x \frac{1 - (T - t)}{1 - T}, \quad Y_t = \frac{x}{1 - T} \quad \text{and} \quad Z_t = 0.$$

Notice that when  $x \neq 0$ ,  $u$  tend to infinity in the neighbourhood of 0 when  $T \rightarrow 1$ , thus there is no decoupling field on  $[0, 1]$  for the FBSDE (IV.4.22)-(IV.4.23).

Now let's consider the following path-dependent FBSDE :

$$\begin{cases} dX_t = b_t(X, Y_t, Z_t)dt + \sigma_t(X, Y_t)dW_t & X_0 = x \\ dY_t = -f_t(X, Y_t, Z_t)dt + Z_t dW_t & Y_T = g(X). \end{cases} \quad \begin{matrix} \text{(IV.4.24)} \\ \text{(IV.4.25)} \end{matrix}$$

The condition in Theorem IV.4.1  $K_1 |\nabla_z \sigma|_\infty < 1$  is automatically satisfied. Therefore, there exists  $\epsilon > 0$  such that the FBSDE (IV.4.24)-(IV.4.25) has a unique regular decoupling field  $u$  on  $[T - \epsilon, T]$  with the terminal condition  $u_T(X) = g(X)$ . Denote  $\mathcal{H}$  the solution of the associated characteristic BSDE (IV.4.12). Note that  $\text{esssup } \mathcal{H}_t$  is a Lipschitz constant of  $u_t$  with respect to the path space variable. Therefore, as long as the solution of the characteristic BSDE  $\mathcal{H}$  is bounded on  $[T - \epsilon, T]$ , we can re-apply the local existence result at  $T - \epsilon$  with terminal condition  $Y_{T-\epsilon} = u_{T-\epsilon}(X)$  and so on. Notice that the length of the time interval  $\epsilon$  given by Theorem IV.4.1 will decrease when the Lipschitz constant of the terminal condition of the backward process increases. The Lipschitz constant that we get at time  $T - \epsilon$  is  $\text{esssup } \mathcal{H}_{T-\epsilon}$ , which is always bigger than  $K_1$ . It means that the length of step at which we iterate the procedure decreases. In order to find the maximal time interval on which we can construct a solution by the above procedure, one way is to find a dominating ODE and find the time of the explosion  $T_{max}$  of the ODE. By Proposition IV.4.4, for any  $T < T_{max}$ , the FBSDE has a unique solution. One possible dominating ODE is

$$\dot{y}_t = \left( |A_t|_\infty + \frac{|B_t|_\infty}{2} \right) y_t^2 + \left( |C_t|_\infty + \frac{|B_t|_\infty}{2} + \frac{|D_t|_\infty}{2} \right) y_t + \left( |F_t|_\infty + \frac{|D_t|_\infty}{2} \right),$$

where  $|\cdot|_\infty$  is the essential supremum and the coefficients are given in (IV.4.12). The result is summarized in the following Theorem.

**Theorem IV.4.8.** Assume that Assumption IV.2.1 holds true. Then there exists a dominating Riccati ODE with terminal condition  $y_T = K_1$ . In addition, there exists  $T_{max} > 0$  depending only on the

dimension and the Lipschitz coefficients of the FBSDE such that the dominating ODE has a bounded solution on  $[0, T]$  for all  $T < T_{\max}$  and hence, the FBSDE (IV.4.24)-(IV.4.25) has a unique solution on  $[0, T]$ .

#### IV.4.2.5 General Case

In the general case where  $|\nabla_z \sigma|_\infty \neq 0$ , in order to have the existence on small time interval, we need to have the condition  $K_1 |\nabla_z \sigma|_\infty < 1$ . To use the same technique to extend the existence result on larger interval, we need to maintain the very same condition, i.e.  $|\nabla_x u(t, \cdot)|_\infty |\nabla_z \sigma|_\infty < 1$ , where  $|\nabla_x u(t, \cdot)|_\infty$  is the essential supremum of all the directional derivatives of the decoupling field  $u$  with respect to the path space variable  $x \in \mathcal{C}([0, t], \mathbb{R}^d)$  at time  $t$  as defined in (IV.4.9). We introduce now the notion of maximal interval as in [60].

**Definition IV.4.9.** *The maximal interval  $I_{\max}$  on  $[0, T]$  for the FBSDE  $(b, \sigma, f, g)$  is defined as the union of all intervals of form  $[t, T]$  on which the FBSDE  $(b, \sigma, f, g)$  has a decoupling field  $u$  such that  $|\nabla_x u(s, \cdot)|_\infty |\nabla_z \sigma|_\infty < 1$  for all  $s \in [t, T]$ .*

**Remark IV.4.10.** *Notice that the maximal interval for an FBSDE given  $T$  may very well be open to the left. In this case we say a decoupling field is regular on  $I_{\max}$  if  $u$  restricted to  $[s, T]$  is a regular decoupling field for all  $s \in I_{\max}$ .*

**Proposition IV.4.11.** *Under Assumption IV.2.1, if  $K_1 |\nabla_z \sigma|_\infty < 1$ , let  $I_{\max}$  be the maximal interval associated to the FBSDE  $(b, \sigma, f, g)$ , then there exists a unique decoupling field  $u$  satisfying*

$$|\nabla_x u(t, \cdot)|_\infty |\nabla_z \sigma|_\infty < 1, \quad (\text{IV.4.26})$$

and it is regular.

For the proof of Proposition IV.4.11, we refer to [60, Theorem 2].

**Proposition IV.4.12.** *Under Assumption IV.2.1 and assume that  $K_1 |\nabla_z \sigma|_\infty < 1$ , if the maximal interval associated to the FBSDE  $(b, \sigma, f, g)$  is open on the left, i.e.  $I_{\max} = (t_{\min}, T]$ , then necessarily,*

$$\lim_{t \downarrow t_{\min}} |\nabla_x u(t, \cdot)|_\infty |\nabla_z \sigma|_\infty = 1. \quad (\text{IV.4.27})$$

The proof of Proposition IV.4.12 is the same as in Markovian case given in [60], for readers' convenience, we detail the proof below.

**Proof of Proposition IV.4.12.** Assume that there exist a sequence of  $(t_n)_{n \geq 0} \downarrow t_{\min}$  such that

$$\lim_{t_n \downarrow t_{\min}} |\nabla_x u(t_n, \cdot)|_\infty |\nabla_z \sigma|_\infty < 1.$$

According to Remark IV.4.2, one can construct a small time interval  $\epsilon$  depending only on the Lipschitz coefficient of the FBSDE  $K_0$ ,  $\limsup_{t_n \downarrow t_{\min}} |\nabla_x u(t_n, \cdot)|_\infty$  and the product  $\lim_{t \downarrow t_{\min}} |\nabla_x u(t, \cdot)|_\infty |\nabla_z \sigma|_\infty$  such that for  $n$  large enough, we can construct a decoupling field for the FBSDE on the interval  $[t_n - \epsilon, t_n]$ . Since  $\epsilon$  is independent of  $n$ , one can choose a  $n$  such that  $t_n - \epsilon < t_{\min}$ , contradicting the definition of maximal interval.  $\blacksquare$

## IV.5 Stability Property of Path-dependent Multidimensional FBSDE

In this subsection, we shall give the stability result for the path-dependent FBSDE. Consider the path-dependent FBSDE (IV.2.3)-(IV.2.4). To deduce the stability property, we will need the following estimate on the solutions of FBSDE. Denote  $\mathcal{L}$  the space of all  $\mathbb{F}$ -adapted processes  $(Y, Z)$  such that

$$\|(Y, Z)\|_2 := \sup_{t \in [0, T]} \left\{ \mathbb{E} \left[ |Y_t|^2 + \int_t^T |Z_s|^2 ds \right] \right\} < +\infty.$$

Denote

$$I_0^2 := \mathbb{E} \left[ \left( \int_0^T |f_t^0| + |b_t^0| dt \right)^2 + \int_0^T |\sigma_t^0|^2 dt \right].$$

Recall first the following well-known estimate on SDE, BSDE. The proofs can be found for example in [41, Chapter 9] of Cvitanic and Zhang.

**Lemma IV.5.1** (A Priori Estimate on path-dependent SDE). *Let  $T > 0$ . Consider the following path-dependent SDE :*

$$dX_t = b_t(X)dt + \sigma_t(X)dW_t, \quad X_0 = x. \quad (\text{IV.5.28})$$

Assume that

1.  $x \in L^2(\mathcal{F}_0)$ ,  $b^0, \sigma^0 \in \mathbb{H}_T^2(\mathbb{F}, \mathbb{R}^d)$ ;
2.  $b, \sigma$  are uniformly Lipschitz continuous on  $x \in \mathcal{C}([0, T], \mathbb{R}^d)$  as in the Assumption IV.2.1, namely  $\forall t \in [0, T], \forall x_i \in \mathcal{C}([0, T], \mathbb{R}^d), i = 1, 2$ ,

$$|b_t(x_1) - b_t(x_2)| + |\sigma_t(x_1) - \sigma_t(x_2)| \leq C \|x_1 - x_2\|_{2,t}.$$

If  $X \in \mathbb{H}_T^2(\mathbb{F}, \mathbb{R}^d)$  is a solution of the SDE (IV.5.28), then

$$\mathbb{E}[\|X\|_{2,T}] \leq C \left( |x|^2 + \mathbb{E} \left[ \left( \int_0^T (|b_t^0|) dt \right)^2 + \int_0^T |\sigma_t^0|^2 dt \right] \right).$$

where the constant  $C$  depends on  $T$ , the dimension  $d$  and the Lipschitz constant of  $b$  and  $\sigma$ .



**Lemma IV.5.2** (A Priori Estimate on BSDE). *Let  $T > 0$ . Consider the following BSDE*

$$dY_t = -f_t(Y_t, Z_t)dt + Z_t dW_t, \quad Y_T = \xi, \quad (\text{IV.5.29})$$

Assume that

1.  $\xi \in L^2(\mathcal{F}_T, \mathbb{R}^n)$  and  $f^0 \in \mathbb{H}_T^2(\mathbb{F}, \mathbb{R}^n)$ ;
2.  $f$  is uniformly Lipschitz continuous on  $(y, z)$ , namely  $\forall t \in [0, T], \forall y_i \in \mathbb{R}^n$  and  $\forall z_i \in \mathcal{M}_n(\mathbb{R})$ ,  $i = 1, 2$ ,

$$|f_t(y_1, z_1) - f_t(y_2, z_2)| \leq C [|y_1 - y_2| + |z_1 - z_2|].$$

If  $(Y, Z) \in \mathbb{H}_T^2(\mathbb{F}, \mathbb{R}^n) \times \mathbb{H}_T^2(\mathbb{F}, \mathcal{M}_n(\mathbb{R}))$  is a solution of the BSDE (IV.5.29), then

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |Y_t|^2 + \int_0^T |Z_t|^2 dt \right] \leq C \mathbb{E} \left[ |\xi|^2 + \left( \int_0^T |f_t^0| dt \right)^2 \right],$$

where the constant  $C$  depends on  $T$ , the dimension  $n$  and the Lipschitz constant of  $f$ .

The following lemma generalizes the existing result on the a priori estimate on FBSDE. The techniques are similar with an additional difficulty that the coefficients of the FBSDE can depend on the whole path of the forward process  $X$ .

**Lemma IV.5.3** (A Priori Estimate for FBSDE on Small Time Interval). *Assume that all the hypotheses in the Assumption IV.2.1 are satisfied. Let  $T$  be a small time horizon on which Theorem IV.4.1 applies. If  $(X, Y, Z) \in \mathbb{H}_T^2 \times \mathcal{L}$  are solution of the FBSDE (IV.2.3)-(IV.2.4) on  $[0, T]$ , then there exists a constant  $C$  such that*

$$\sup_{t \in [0, T]} \left\{ \mathbb{E} \left[ \|X\|_{2,t}^2 + |Y_t|^2 + \int_t^T |Z_s|^2 ds \right] \right\} \leq C (|x|^2 + \mathbb{E} [|g^0|^2] + I_0^2).$$

**Proof of Lemma IV.5.3.** Let  $(y, z)$  two progressively measurable processes. Let  $(X, Y, Z)$  be the unique solution of the following decoupled FBSDE on  $[0, T]$  :

$$\begin{cases} dX_t = b_t(X, y_t, z_t)dt + \sigma_t(X, y_t, z_t)dW_t & X_0 = x \end{cases} \quad (\text{IV.5.30})$$

$$\begin{cases} dY_t = -f_t(X, y_t, z_t)dt + Z_t dW_t & Y_T = g(X_{\cdot \wedge T}). \end{cases} \quad (\text{IV.5.31})$$

We have shown that the mapping  $(y, z) \mapsto (Y, Z)$  is a contraction in the proof of Theorem IV.4.1 in the space  $(\mathcal{L}, \|\cdot\|_2)$ . Denote  $(X_0, Y_0, Z_0)$  the solution of the FBSDE (IV.5.30)-(IV.5.31) with  $(y, z) = (0, 0)$ . We have

$$\|(Y - Y_0, Z - Z_0)\|_2 \leq C \|(Y, Z)\|_2,$$

where  $C < 1$ . By the triangle inequality, we get

$$\|(Y, Z)\|_2 \leq \|(Y - Y_0, Z - Z_0)\|_2 + \|(Y_0, Z_0)\|_2 \leq C \|(Y, Z)\|_2 + \|(Y_0, Z_0)\|_2,$$

and therefore, together with Lemma IV.5.2 and Lemma IV.5.1, we have

$$\begin{aligned}
 \|(Y, Z)\|_2 &\leq \frac{1}{1-C} \|(Y_0, Z_0)\|_2 \\
 &\leq C \mathbb{E} \left[ |g(X_0)|^2 + \left( \int_0^T |f(t, X_0(t), 0, 0)| dt \right)^2 \right] \\
 &\leq C \mathbb{E} \left[ K_1^2 \|X_0\|_{2,T}^2 + K_0^2 \int_0^T \|X_0\|_{2,t}^2 dt + |g(0)|^2 + \left( \int_0^T |f_t^0| dt \right)^2 \right] \\
 &\leq C \mathbb{E} \left[ |x|^2 + |g(0)|^2 + \left( \int_0^T |f_t^0| + |b_t^0| dt \right)^2 + \int_0^T |\sigma_t^0|^2 dt \right] \\
 &= C(\mathbb{E} [|x|^2 + |g(0)|^2] + I_0^2),
 \end{aligned}$$

where the constants  $C$  may vary from line to line. Now let's examine the forward process  $X$ . By Lemma IV.5.1, we get

$$\begin{aligned}
 \sup_{0 \leq t \leq T} \mathbb{E}[\|X\|_{2,t}] &\leq C \left( |x|^2 + \mathbb{E} \left[ \left( \int_0^T (|b(t, 0, Y_t, Z_t)| dt) \right)^2 + \int_0^T |\sigma(t, 0, Y_t, Z_t)|^2 dt \right] \right) \\
 &\leq C \left( |x|^2 + \|(Y, Z)\|_2^2 + \mathbb{E} \left[ \left( \int_0^T (|b_t^0| dt) \right)^2 + \int_0^T |\sigma_t^0|^2 dt \right] \right) \\
 &\leq C(\mathbb{E} [|x|^2 + |g(0)|^2] + I_0^2).
 \end{aligned}$$

Combining the above inequalities, we get

$$\sup_{t \in [0, T]} \left\{ \mathbb{E} \left[ \|X\|_{2,t}^2 + |Y_t|^2 + \int_t^T |Z_s|^2 ds \right] \right\} \leq C(\mathbb{E} [|x|^2 + |g(0)|^2] + I_0^2).$$

■

**Theorem IV.5.4** (Stability Property of path-dependent FBSDE). *Assume that  $(b, \sigma, f, g)$  and  $(b', \sigma', f', g')$  satisfy the same condition (i.e. they belong to the same case discussed in the Section 4). Let  $T$  be a time horizon on which both FBSDE have a solution, denoted respectively  $\Xi = (X, Y, Z)$  and  $\Xi' = (X', Y', Z')$ . For  $\phi = b, \sigma, f, g$ , denote  $\Delta\phi := \phi - \phi'$ . Let*

$$\Delta I_0^2 := \mathbb{E} \left[ \left( \int_0^T |\Delta f_t(\Xi'_t)| + |\Delta b_t(\Xi'_t)| dt \right)^2 + \int_0^T |\Delta \sigma_t(\Xi'_t)|^2 dt \right].$$

We have

$$\sup_{t \in [0, T]} \left\{ \mathbb{E} \left[ \|\Delta X\|_{2,t}^2 + |\Delta Y_t|^2 + \int_t^T |\Delta Z_s|^2 ds \right] \right\} \leq C(|\Delta x|^2 + \mathbb{E} [|\Delta g(X')|^2] + \Delta I_0^2).$$

**Proof of Theorem IV.5.4.** We follow the steps of the proof of Theorem 8.1 in the paper [110]. Using the notation described in the Section 4, we have

$$\begin{aligned} d\Delta X_t &= (b_x \|\Delta X\|_{2,t} + b_y \Delta Y_t + \text{Tr}(b_z \Delta Z_t) + \Delta b_t(\Xi'(t)))dt \\ &\quad + (\sigma_x \|\Delta X\|_{2,t} + \sigma_y \Delta Y_t + \text{Tr}(\sigma_z \Delta Z_t) + \Delta \sigma_t(\Xi'))dW_t, \end{aligned} \quad (\text{IV.5.32})$$

$$d\Delta Y_t = \int_t^T (f_x \|\Delta X\|_{2,t} + f_y \Delta Y_t + \text{Tr}(f_z \Delta Z_t) + \Delta f_t(\Xi'(t)))dt - \int_t^T \Delta Z_s dW_t, \quad (\text{IV.5.33})$$

with initial condition  $\Delta X_0 = x - x'$  and terminal condition  $\Delta Y_T = g_x \|\Delta X\|_{2,t} + \Delta g(X')$ .

Since both FBSDE satisfy the same condition, which means there exists  $n \in \mathbb{N}$  and  $0 = t_0 < \dots < t_n = T$  such that on each small interval  $[t_i, t_{i+1}]$ , Lemma IV.5.3 applies to both FBSDE, which means Lemma IV.5.3 applies equally to the above FBSDE (IV.5.32)-(IV.5.33). Denote  $u_1$  and  $u_2$  the two associated decoupling fields. We have

$$\begin{aligned} &\sup_{t \in [t_i, t_{i+1}]} \left\{ \mathbb{E} \left[ \|\Delta X\|_{2,t}^2 + |\Delta Y_t|^2 + \int_t^{t_{i+1}} |\Delta Z_s|^2 ds \right] \right\} \\ &\leq C \mathbb{E} \left[ \|\Delta X\|_{2,t_i}^2 + |\Delta u(t_{i+1}, X')|^2 + \left( \int_{t_i}^{t_{i+1}} |\Delta f_t(\Xi'(t))| + |\Delta b_t(\Xi'(t))| dt \right)^2 + \int_{t_i}^{t_{i+1}} |\Delta \sigma_t(\Xi'(t))|^2 dt \right] \\ &\leq C (\mathbb{E} [\|\Delta X\|_{2,t_i}^2 + |\Delta u(t_{i+1}, X')|^2] + \Delta I_0^2). \end{aligned} \quad (\text{IV.5.34})$$

Apply Lemma IV.5.3 to the above linear forward-backward equation (IV.5.32)-(IV.5.33) with initial condition  $\Delta X_{\cdot \wedge t_i} = 0$  and terminal condition  $u_x \|\Delta X\|_{2,t_{i+1}} + \Delta u(t_{i+1}, X')$  on  $[t_i, t_{i+1}]$ , and note that the difference between the solution of equation  $(b, \sigma, f, g)$  with initial condition  $X_{t_i} = X'_{t_i}$  and terminal condition  $Y_{t_{i+1}} = u(t_{i+1}, X)$  and the solution of equation  $(b', \sigma', f', g')$  with initial condition  $X_{t_i} = X'_{t_i}$  and terminal condition  $Y_{t_{i+1}} = u'(t_{i+1}, X)$  on the interval  $[t_i, t_{i+1}]$  is exactly the very solution, we get

$$\mathbb{E}[|\Delta u(t_i, X')|^2] = \mathbb{E}[|\Delta Y_{t_i}|^2] \leq C \Delta I_0^2 + C \mathbb{E}[|\Delta u(t_{i+1}, X')|^2].$$

By iteration one can show that with a larger constant  $C$ , we have

$$\mathbb{E}[|\Delta u(t_i, X')|^2] \leq C (\mathbb{E} [|\Delta g(X')|^2] + \Delta I_0^2).$$

Now apply again Lemma IV.5.3 but on the forward equation (IV.5.32) on the interval  $[t_i, t_{i+1}]$ , together with the above inequalities, we get

$$\begin{aligned} \mathbb{E} [\|\Delta X\|_{2,t_{i+1}}^2] &\leq C (\mathbb{E} [\|\Delta X\|_{2,t_i}^2 + |\Delta u(t_{i+1}, X')|^2] + \Delta I_0^2) \\ &\leq C (\mathbb{E} [\|\Delta X\|_{2,t_i}^2 + \Delta g(X')] + \Delta I_0^2). \end{aligned}$$

By iteration one can show that with a larger constant  $C$ , we have

$$\mathbb{E} \left[ \|\Delta X\|_{2,t_i}^2 \right] \leq C(\mathbb{E} [|\Delta x|^2 + \Delta g(X')] + \Delta I_0^2).$$

Injecting the above inequalities into the inequality (IV.5.34), we get

$$\sup_{t \in [t_i, t_{i+1}]} \left\{ \mathbb{E} \left[ \|\Delta X\|_{2,t}^2 + |\Delta Y_t|^2 + \int_t^{t_{i+1}} |\Delta Z_s|^2 ds \right] \right\} \leq C(\mathbb{E} [|\Delta x|^2 + |\Delta g(X')|^2] + \Delta I_0^2).$$

We conclude by summing up both side from  $i = 0$  to  $i = n$ . ■

## IV.6 Proofs

**Proof of Theorem IV.3.2.** We shall follow the steps of the proof of Theorem 2.3 in [110] of Ma, Wu, Zhang (Detao), Zhang (Jianfeng).

(Existence) Let  $0 = t_0 < t_1 < \dots < t_n = T$  be a partition of  $[0, T]$  such that  $\forall i \in \llbracket 1, n \rrbracket, t_{i+1} - t_i > \delta$ . On  $[t_0, t_1]$ , the FBSDE with initial value  $x$  and terminal value  $u(t_1, X)$  has a unique solution  $(X^{t_0, t_1}, Y^{t_0, t_1}, Z^{t_0, t_1})$  that satisfies  $Y_t^{t_0, t_1} = u(t, X)$ . On  $[t_1, t_2]$ , the FBSDE with initial value  $X_{\wedge t_1}^{t_0, t_1}$  and terminal value  $u_{t_2}(X)$  has a unique solution  $(X^{t_1, t_2}, Y^{t_1, t_2}, Z^{t_1, t_2})$  that satisfies again  $Y_t^{t_1, t_2} = u(t, X)$ . The initial condition of  $X^{t_1, t_2}$  is  $X^{t_0, t_1}$ . By patching them together we obtain an forward process  $X^{t_0, t_2}$ , which can be used as initial value for the FBSDE on the interval  $[t_2, t_3]$ . Repeating this procedure forwardly in time  $n$  times, we get a solution on each of the interval of the partition  $0 = t_0 < t_0 < \dots < t_n = T$ .

We notice that the forward process on  $[0, T]$  has been constructed during the above procedure. We only need to prove that the pieces of the backward process can be patched together. Notice that

$$Y_{t_i+}^{t_i, t_{i+1}} = Y_{t_i}^{t_i, t_{i+1}} = u(t_i, X) = Y_{t_i}^{t_{i-1}, t_i}, \quad (\text{IV.6.35})$$

which means the backward process  $Y$  defined on each interval  $[t_i, t_{i+1}]$  by the above procedure is continuous. Moreover, we have  $Y_t = u(t, X)$  and in particular,  $Y_T = g(X)$ . One can check easily that  $(X, Y, Z)$  verifies the FBSDE with initial condition  $X_0 = x$  and terminal condition  $Y_T = g(X)$ .

We can check easily that  $(X, Y, Z)$  is a solution of the FBSDE with initial value  $x$  and terminal value  $u(T, X) = g(X)$ .

(Uniqueness) Let  $(\tilde{X}, \tilde{Y}, \tilde{Z})$  be another solution of the FBSDE with the same initial and terminal condition. By the definition of decoupling field, on the interval  $[t_{n-1}, t_n]$ , we have  $\tilde{Y}_t = u(t, \tilde{X})$ . This implies that  $(\tilde{X}, \tilde{Y}, \tilde{Z})$  satisfies the FBSDE with initial condition  $\tilde{X}_{\wedge t_{n-2}}$  on  $[t_{n-2}, t_{n-1}]$ . Therefore,  $\tilde{Y}_t = u(t, \tilde{X})$  is satisfied on  $[t_{n-2}, t_{n-1}]$ . Repeating this procedure backwardly in time and we get  $\tilde{Y}_t = u(t, \tilde{X})$  for  $t \in [0, T]$ .

On  $[t_0, t_1]$ ,  $(\tilde{X}, \tilde{Y}, \tilde{Z})$  satisfies the FBSDE with initial condition  $x$  and terminal condition  $\tilde{Y}_{t_1} =$

$u(t_1, \tilde{X})$ , by the uniqueness of solution,  $(X_t, Y_t, Z_t) = (\tilde{X}_t, \tilde{Y}_t, \tilde{Z}_t)$  on  $[t_0, t_1]$ . In particular, the FBSDE on  $[t_1, t_2]$  has the same initial condition for  $X$  and  $\tilde{X}$ . Repeating the arguments forwardly in time and we can see that  $(X_t, Y_t, Z_t) = (\tilde{X}_t, \tilde{Y}_t, \tilde{Z}_t)$  on  $[0, T]$ . ■

**Proof of Theorem IV.4.1.** Let  $(y, z) \in \mathbb{H}_T^2 \times \mathbb{H}_T^2$ . Let  $(X, Y, Z)$  be the unique solution of the following decoupled FBSDE :

$$\begin{cases} dX_t = b_t(X, y_t, z_t)dt + \sigma_t(X, y_t, z_t)dW_t & X_0 = x \\ dY_t = -f_t(X, y_t, z_t)dt + Z_t dW_t & Y_T = g(X_{\cdot \wedge T}). \end{cases} \quad \begin{matrix} \text{(IV.6.36)} \\ \text{(IV.6.37)} \end{matrix}$$

We can then define the following mapping  $(y, z) \in \mathbb{H}_T^2 \times \mathbb{H}_T^2 \mapsto (Y, Z) \in \mathbb{H}_T^2 \times \mathbb{H}_T^2$ . Our goal is to show that this mapping is a contraction for some norm that we shall define later. First of all, we notice that if the mapping is indeed a contraction, then the fixed point of the mapping  $(y, z)$  and the corresponding forward process  $X$  defined by the equation (IV.6.36) are a solution of the FBSDE (IV.2.3)-(IV.2.4). Conversely, if  $(X, Y, Z)$  is a solution of the FBSDE (IV.2.3)-(IV.2.4), then  $(Y, Z)$  is a fixed point of the mapping we define above.

Let  $(y, z)$  and  $(y', z')$  be two pairs of progressively measurable processes and let  $(X, Y, Z)$  and  $(X', Y', Z')$  be the corresponding solutions of the above decoupled FBSDE (IV.6.36)-(IV.6.37).

Denote  $\Delta\alpha := \alpha - \alpha'$  for  $\alpha = y, z, X, Y, Z$  and denote

$$\Delta_x \xi_t := \xi_t(X, Y_t, Z_t) - \xi_t(X', Y_t, Z_t)$$

$$\Delta_y \xi_t := \xi_t(X', Y, Z) - \xi_t(X', Y', Z)$$

$$\Delta_z \xi_t := \xi_t(X', Y', Z) - \xi_t(X', Y', Z'),$$

for  $\xi = b, \sigma, f$ . Clearly,

$$\Delta X_t = \int_0^t (\Delta_x b_s + \Delta_y b_s + \Delta_z b_s) ds + \int_0^t (\Delta_x \sigma_s + \Delta_y \sigma_s + \Delta_z \sigma_s) dW_s.$$

By Ito's formula, we get

$$\mathbb{E}[|\Delta X_t|^2] = \mathbb{E}\left[\int_0^t 2\Delta X_s (\Delta_x b_s + \Delta_y b_s + \Delta_z b_s) ds + \int_0^t |\Delta_x \sigma_s + \Delta_y \sigma_s + \Delta_z \sigma_s|^2 ds\right].$$

By Cauchy-Schwarz Inequality and the inequality (IV.2.6), we get

$$2\Delta X_s \cdot \Delta_x b_s \leq 2K_0 |\Delta X_s| \|\Delta X\|_{2,s} \leq K_0 \left(2|X_s|^2 + \int_0^t |\Delta X_s|^2 ds\right);$$

$$2\Delta X_s \cdot \Delta_y b_s \leq 2K_0 |\Delta X_s| |\Delta y_s| \leq K_0 (|\Delta X_s|^2 + |\Delta y_s|^2);$$

$$2\Delta X_s \cdot \Delta_z b_s \leq 2K_0 |\Delta X_s| |\Delta z_s| \leq K_0 \left( \frac{|\Delta X_s|^2}{\epsilon} + \epsilon |\Delta z_s|^2 \right).$$

Combing the above inequalities, we get

$$\int_0^t 2\Delta X_s (\Delta_x b_s + \Delta_y b_s + \Delta_z b_s) ds \leq \int_0^t K_0 ((3+t+\epsilon^{-1})|\Delta X_s|^2 + |\Delta y_s|^2 + \epsilon |\Delta z_s|^2) ds. \quad (\text{IV.6.38})$$

Using Minkowski inequality and arithmetic-geometric inequality, we get

$$\begin{aligned} & |\Delta_x \sigma_s + \Delta_y \sigma_s + \Delta_z \sigma_s|^2 \\ & \leq (K_0 \|\Delta X\|_{2,s} + K_0 |\Delta y_s| + |\nabla_z \sigma|_\infty |\Delta z_s|)^2 \\ & \leq 2K_0^2 \left(1 + \frac{K_0}{\epsilon}\right) (\|\Delta X\|_{2,s}^2 + |\Delta y_s|^2) + (|\nabla_z \sigma|_\infty^2 + K_0 \epsilon) |\Delta z_s|^2 \\ & \leq 2K_0^2 \left(1 + \frac{K_0}{\epsilon}\right) (|\Delta X_s|^2 + \int_0^t |\Delta X_s|^2 ds + |\Delta y_s|^2) + (|\nabla_z \sigma|_\infty^2 + K_0 \epsilon) |\Delta z_s|^2. \end{aligned} \quad (\text{IV.6.39})$$

Combining the inequality (IV.6.38) and (IV.6.39), we get

$$\begin{aligned} & \mathbb{E} \left[ \int_0^t 2\Delta X_s (\Delta_x b_s + \Delta_y b_s + \Delta_z b_s) ds + \int_0^t |\Delta_x \sigma_s + \Delta_y \sigma_s + \Delta_z \sigma_s|^2 ds \right] \\ & \leq \int_0^t C_\epsilon (|\Delta X_s|^2 + |\Delta y_s|^2) + (2K_0 \epsilon + |\nabla_z \sigma|_\infty^2) |\Delta z_s|^2 ds, \end{aligned}$$

where

$$C_\epsilon := 2K_0^2 \left(1 + \frac{K_0}{\epsilon}\right) (1+T) + K_0(3+T+\epsilon^{-1}).$$

By Gronwall Inequality, we get

$$\begin{aligned} \mathbb{E} [|\Delta X_t|^2] & \leq \mathbb{E} \left[ e^{C_\epsilon t} \int_0^t (C_\epsilon |\Delta y_s|^2 + (2K_0 \epsilon + |\nabla_z \sigma|_\infty^2) |\Delta z_s|^2) ds \right] \\ & \leq \mathbb{E} \left[ e^{C_\epsilon T} \int_0^T (C_\epsilon |\Delta y_s|^2 + (2K_0 \epsilon + |\nabla_z \sigma|_\infty^2) |\Delta z_s|^2) ds \right]. \end{aligned}$$

Again by applying the Ito's Formula to the process  $|\Delta Y_t|^2$ , we get

$$\mathbb{E} [|\Delta Y_t|^2 + \int_t^T |\Delta Z_s|^2] = |\Delta Y_T|^2 + \mathbb{E} \left[ \int_t^T 2\Delta Y_s (\Delta_x f_s + \Delta_y f_s + \Delta_z f_s) ds \right]. \quad (\text{IV.6.40})$$

By Cauchy-Schwarz inequality and arithmetic-geometric inequality, we get

$$\begin{aligned} 2\Delta Y_s (\Delta_x f_s + \Delta_y f_s + \Delta_z f_s) & \leq 2K_0 |\Delta Y_s| (\|\Delta X\|_{2,s} + |\Delta y_s| + |\Delta z_s|) \\ & \leq K_0 \left( (2+\epsilon^{-1}) |\Delta Y_s|^2 + \|\Delta X\|_{2,s}^2 + |\Delta y_s|^2 + \epsilon |\Delta z_s|^2 \right). \end{aligned}$$

Combining equation (IV.6.40) and the above inequality, we get

$$\begin{aligned}
 & \mathbb{E}\left[|\Delta Y_t|^2 + \int_t^T |\Delta Z_s|^2\right] \\
 & \leq |\Delta Y_T|^2 + K_0 \mathbb{E}\left[\int_t^T (2 + \epsilon^{-1})|\Delta Y_s|^2 + \epsilon|\Delta z_s|^2 + |\Delta y_s|^2 ds + (T+1) \int_0^T |\Delta X_s|^2 ds\right] \\
 & \leq \mathbb{E}\left[K_1^2 \int_0^T |\Delta X_t|^2 dt + K_1^2 |\Delta X_T|^2\right] + \mathbb{E}\left[\int_t^T \tilde{C}_\epsilon |\Delta Y_s|^2 ds\right. \\
 & \quad \left. + \int_0^T K_0(\epsilon + T(T+1))e^{C_\epsilon T} (2K_0\epsilon + |\nabla_z \sigma|_\infty^2) |\Delta z_s|^2 + (K_0 + T(T+1))K_0 e^{C_\epsilon T} C_\epsilon |\Delta y_s|^2 ds\right] \\
 & \leq \mathbb{E}\left[\int_0^T (T+1)K_1^2 e^{C_\epsilon T} (C_\epsilon |\Delta y_s|^2 + (2K_0\epsilon + |\nabla_z \sigma|_\infty^2) |\Delta z_s|^2) ds + \int_t^T \tilde{C}_\epsilon |\Delta Y_s|^2 ds\right. \\
 & \quad \left. + \int_0^T K_0 e^{C_\epsilon T} (\epsilon + T(T+1))(2K_0\epsilon + |\nabla_z \sigma|_\infty^2) |\Delta z_s|^2 + (K_0 + T(T+1))K_0 e^{C_\epsilon T} C_\epsilon |\Delta y_s|^2 ds\right] \\
 & \leq \mathbb{E}\left[\int_t^T \tilde{C}_\epsilon |\Delta Y_s|^2 + \int_0^T C_y(T, \epsilon) |\Delta y_s|^2 + C_z(T, \epsilon) |\Delta z_s|^2 ds\right]. \tag{IV.6.41}
 \end{aligned}$$

In the last line of the above inequalities, the constants  $\tilde{C}_\epsilon$ ,  $C_y(T, \epsilon)$  and  $C_z(T, \epsilon)$  are given by

$$\tilde{C}_\epsilon := K_0(2 + \epsilon^{-1}), \quad C_y(T, \epsilon) := (T+1)K_1^2 e^{C_\epsilon T} C_\epsilon + K_0 + T(T+1)K_0 e^{C_\epsilon T} C_\epsilon$$

and

$$C_z(T, \epsilon) := (T+1)K_1^2 e^{C_\epsilon T} (2K_0\epsilon + |\nabla_z \sigma|_\infty^2) + K_0(\epsilon + T(T+1))e^{C_\epsilon T} (3K_0\epsilon + |\nabla_z \sigma|_\infty^2).$$

Using Gronwall's inequality on  $Y$ , we get

$$\begin{aligned}
 \mathbb{E}\left[|\Delta Y_t|^2\right] & \leq \mathbb{E}\left[e^{\tilde{C}_\epsilon(T-t)} \int_0^T C_y(T, \epsilon) |\Delta y_s|^2 + C_z(T, \epsilon) |\Delta z_s|^2 ds\right] \\
 & \leq \mathbb{E}\left[e^{\tilde{C}_\epsilon T} \int_0^T C_y(T, \epsilon) |\Delta y_s|^2 + C_z(T, \epsilon) |\Delta z_s|^2 ds\right]. \tag{IV.6.42}
 \end{aligned}$$

Plug the inequality(IV.6.42) into the inequality(IV.6.41), we get, for all  $t \in [0, T]$  :

$$\begin{aligned}
 \mathbb{E}\left[|\Delta Y_t|^2 + \int_t^T |\Delta Z_s|^2\right] & \leq \mathbb{E}\left[(\tilde{C}_\epsilon T e^{\tilde{C}_\epsilon T} + 1) \int_0^T C_y(T, \epsilon) |\Delta y_s|^2 + C_z(T, \epsilon) |\Delta z_s|^2 ds\right] \\
 & \leq \left(T(\tilde{C}_\epsilon T e^{\tilde{C}_\epsilon T} + 1)C_y(T, \epsilon) + (\tilde{C}_\epsilon T e^{\tilde{C}_\epsilon T} + 1)C_z(T, \epsilon)\right) \\
 & \quad \times \sup_{t \in [0, T]} \left\{ \mathbb{E}\left[|\Delta y_t|^2 + \int_t^T |\Delta z_s|^2 ds\right] \right\}.
 \end{aligned}$$

Denote  $\gamma(\epsilon, T) := T(\tilde{C}_\epsilon T e^{\tilde{C}_\epsilon T} + 1)C_y(T, \epsilon) + (\tilde{C}_\epsilon T e^{\tilde{C}_\epsilon T} + 1)C_z(T, \epsilon)$  and

$$\|(y, z)\|_2^2 := \sup_{t \in [0, T]} \left\{ \mathbb{E}\left[|y_t|^2 + \int_t^T |z_s|^2 ds\right] \right\}.$$

For a fixed  $\epsilon > 0$ ,  $\gamma(\epsilon, T)$  converges to  $K_1^2(2K_0\epsilon + |\nabla_z \sigma|_\infty^2) + K_0\epsilon$  while  $T \rightarrow 0$ . Since  $K_1 |\nabla_z \sigma|_\infty < 1$ , we can find a pair of  $\epsilon$  and  $T$  small enough such that  $\gamma(\epsilon, T) < 1$ , in which case, the mapping  $(y, z) \mapsto (Y, Z)$  is a contraction. Denote  $\mathcal{L}$  the space of all  $\mathbb{F}$ -adapted processes  $(Y, Z)$  such that  $\|(Y, Z)\|_2 < +\infty$ . We can show easily that the space  $(\mathcal{L}, \|\cdot\|_2)$  is a Banach space, by the contraction mapping theorem, the mapping  $(y, z) \mapsto (Y, Z)$  has an unique fixed point  $(Y, Z)$ . ■



---

---

# CHAPITRE V

---

## MEAN-FIELD LANGEVIN DYNAMICS AND ENERGY LANDSCAPE OF NEURAL NETWORKS

**Abstract :** Our work is motivated by a desire to study the theoretical underpinning for the convergence of stochastic gradient type algorithms widely used for non-convex learning tasks such as training of neural networks. The key insight is that a certain class of the finite-dimensional non-convex problems becomes convex when lifted to infinite-dimensional space of measures. We leverage this observation and show that the corresponding energy functional defined on the space of probability measures has a unique minimiser which can be characterised by a first-order condition using the notion of linear functional derivative. Next, we study the corresponding gradient flow structure in 2-Wasserstein metric, which we call Mean-Field Langevin Dynamics (MFLD), and show that the flow of marginal laws induced by the gradient flow converges to a stationary distribution, which is exactly the minimiser of the energy functional. We observe that this convergence is exponential under conditions that are satisfied for highly regularised learning tasks. Our proof of convergence to stationary probability measure is novel and it relies on a generalisation of LaSalle's invariance principle combined with HWI inequality. Importantly, we assume neither that interaction potential of MFLD is of convolution type nor that it has any particular symmetric structure. Furthermore, we allow for the general convex objective function, unlike, most papers in the literature that focus on quadratic loss. Finally, we show that the error between finite-dimensional optimisation problem and its infinite-dimensional limit is of order one over the number of parameters.

**Keywords :** Mean-Field Langevin Dynamics, Gradient Flow, Neural Networks

## V.1 Introduction

Neural networks trained with stochastic gradient descent algorithm proved to be extremely successful in number of applications such as computer vision, natural language processing, generative models or reinforcement learning [103]. However, complete mathematical theory that would provide theoretical guarantees for the convergence of machine learning algorithms for non-convex learning tasks has been elusive. On the contrary, empirical experiments demonstrate that classical learning theory [152] may fail to predict the behaviour of modern machine learning algorithms [162]. In fact, it has been observed that the performance of neural networks based algorithms is insensitive to the number of parameters in the hidden layers (provided that this is sufficiently large) and in practice one works with models that have number of parameters larger than the size of the training set [70, 8]. These findings motivate the study of neural networks with large number of parameters which is a subject of this work.

Furthermore while universal representation theorems ensures the existence of the optimal parameters of the network, it is in general not known when such optimal parameters can be efficiently approximated by conventional algorithms, such as stochastic gradient descent. This work aims at revealing the intrinsic connection between the optimality of the network parameters and the dynamic of gradient-descent-type algorithm, using the perspective of the mean-field Langevin equation.

This work builds on the rigorous mathematical framework to study non-convex learning tasks such as training of neural networks developed in Mei, Misiakiewicz and Montanari [117], Chizat and Bach [34], Sirignano and Spiliopoulos [143] as well as Rotskoff and Vanden-Eijnden [138].

We extend some existing results and provide a novel proof technique for mathematical results which provide a theoretical underpinning for the convergence of stochastic gradient type algorithms widely used in practice to train neural networks. We demonstrate how our results apply to a situation when one aims to train one-hidden layer neural network with (noisy) stochastic gradient algorithm.

Let us first briefly recall the classical finite dimensional Langevin equation. Given a *potential* function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  which is Lipschitz continuous and satisfies appropriate growth condition, the overdamped Langevin equation reads

$$dX_t = -\nabla f(X_t)dt + \sigma dW_t, \tag{V.1.1}$$

where  $\sigma$  is a scalar constant and  $W$  is a  $d$ -dimension Brownian motion. One can view this dynamic in two perspectives :

- i) The solution to (V.1.1) is a time-homogeneous Markov diffusion, so under mild condition it admits a unique invariant measure  $m^{\sigma,*}$ , of which the density function must be in the form

$$m^{\sigma,*}(x) = \frac{1}{Z} \exp\left(-\frac{2}{\sigma^2}f(x)\right), \quad \text{for all } x \in \mathbb{R}^d, \quad \text{where } Z := \int_{\mathbb{R}^d} \exp\left(-\frac{2}{\sigma^2}f(x)\right) dx.$$

- ii) The dynamic (V.1.1) can be viewed as the path of a randomised continuous time gradient descent

algorithm.

These two perspectives are unified through the variational form of the invariant measure, namely,  $m^{\sigma,*}$  is the unique minimiser of the free energy function

$$V^\sigma(m) := \int_{\mathbb{R}^d} f(x)m(dx) + \frac{\sigma^2}{2}H(m)$$

over all probability measure  $m$ , where  $H$  is the relative entropy with respect to the Lebesgue measure. The variational perspective has been established in [92] and [93]. Moreover, one may observe that the distribution  $m^{\sigma,*}$  concentrates to the Dirac measure  $\delta_{\arg \min f}$  as  $\sigma \rightarrow 0$  and there is no need to assume that the function  $f$  is convex. This establishes the link between theory of statistical sampling and optimisation and show that Langevin equation plays an important role in the non-convex optimisation. This fact is well-recognized by the communities of numerical optimisation and machine learning [87, 76, 75]

This work aims at generalising the connection between the global minimiser and the invariant measure to the case where the *potential* function is a function defined on a space of probability measures. This is motivated by the following observation on the configuration of neural network. Let us take the example of the network with 1-hidden-layer. While the universal representation theorem, [42, 7] tells us that 1-hidden-layer network can arbitrarily well approximate the continuous function on the compact time interval it does not tell us how to find optimal parameters. One is faced with the following non-convex optimisation problem.

$$\min_{\beta_{n,i} \in \mathbb{R}, \alpha_{n,i} \in \mathbb{R}^{d-1}} \left\{ \int_{\mathbb{R} \times \mathbb{R}^{d-1}} \Phi \left( y - \frac{1}{n} \sum_{i=1}^n \beta_{n,i} \varphi(\alpha_{n,i} \cdot z) \right) \nu(dy, dz) \right\}, \quad (\text{V.1.2})$$

where  $\Phi : \mathbb{R} \rightarrow \mathbb{R}$  is a convex function,  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  is a bounded, continuous, non-constant activation function and  $\nu$  is a measure of compact support representing the data. Let us define the empirical law of the parameters as  $m^n := \frac{1}{n} \sum_{i=1}^n \delta_{\{\beta_{n,i}, \alpha_{n,i}\}}$ . Then

$$\frac{1}{n} \sum_{i=1}^n \beta_{n,i} \varphi(\alpha_{n,i} \cdot z) = \int_{\mathbb{R}^d} \beta \varphi(\alpha \cdot z) m^n(d\beta, d\alpha).$$

To ease notation let us use, for  $x = (\beta, \alpha) \in \mathbb{R}^d$ , the function  $\hat{\varphi}(x, z) := \beta \varphi(\alpha \cdot z)$ , and by  $\mathbb{E}^m$  we denote the expectation of random variable  $X$  under the probability measure  $m$ . Now, instead of (V.1.2), we propose to study the following minimisation problem over the probability measures :

$$\min_m F(m), \quad \text{with} \quad F(m) := \int_{\mathbb{R}^d} \Phi \left( y - \mathbb{E}^m [\hat{\varphi}(X, z)] \right) \nu(dy, dz), \quad (\text{V.1.3})$$

This reformulation is crucial, because the *potential* function  $F$  defined above is convex in the measure space i.e. for any probability measures  $m$  and  $m'$  it holds that

$$F((1 - \alpha)m + \alpha m') \leq (1 - \alpha)F(m) + \alpha F(m') \quad \text{for all } \alpha \in [0, 1].$$

This example demonstrates that a non-convex minimisation problem on a finite-dimensional parameter space becomes a convex minimisation problem when lifted to the infinite dimensional space of probability measures. The key aim of this work is to provide analysis that takes advantage of this observation.

In order to build up the connection between the global minimiser of the convex potential function  $F$  and the upcoming mean-field Langevin equation, as in the classic case, we add the relative entropy  $H$  as a regulariser, but different from the classic case, we use the relative entropy with respect to a Gibbs measure of which the density is proportional to  $e^{-U(x)}$ . A typical choice of the Gibbs measure could be the standard Gaussian distribution. One of our main contributions is to characterise the minimiser of the free energy function

$$V^\sigma := F + \frac{\sigma^2}{2} H$$

using the *linear functional derivative* on the space of probability measures, denoted by  $\frac{\delta}{\delta m}$  (introduced originally in calculus of variations and now used extensively in the theory of mean field games see, e.g. Cardaliaguet et al. [24]). Indeed, we prove the following first order condition :

$$m^* = \arg \min_m V^\sigma(m) \quad \text{if and only if} \quad \frac{\delta F}{\delta m}(m^*, \cdot) + \frac{\sigma^2}{2} \log(m^*) + \frac{\sigma^2}{2} U = \text{constant}.$$

This condition together with the fact that  $m^*$  is a probability measure gives

$$m^*(x) = \frac{1}{Z} \exp \left( -\frac{2}{\sigma^2} \left( \frac{\delta F}{\delta m}(m^*, x) + U(x) \right) \right),$$

where  $Z$  is the normalising constant. We emphasise that throughout  $V$  and hence  $m^*$  depend on the regularisation parameter  $\sigma > 0$ . It is noteworthy that the variational form of the invariant measure of the classic Langevin equation is a particular example of this first order condition. Moreover, given a measure  $m^*$  satisfying the first order condition, it is formally a stationary solution to the nonlinear Fokker–Planck equation :

$$\partial_t m = \nabla \cdot \left( \left( D_m F(m, \cdot) + \frac{\sigma^2}{2} \nabla U \right) m + \frac{\sigma^2}{2} \nabla m \right), \quad (\text{V.1.4})$$

where  $D_m F$  is the *intrinsic derivative* on the probability measure space, defined as  $D_m F(m, x) := \nabla \frac{\delta F}{\delta m}(m, x)$ . Clearly, the particle dynamic corresponding to this Fokker-Planck equation is governed by the *mean field Langevin equation* :

$$dX_t = - \left( D_m F(m_t, X_t) + \frac{\sigma^2}{2} \nabla U(X_t) \right) dt + \sigma dW_t, \quad \text{where } m_t := \text{Law}(X_t). \quad (\text{V.1.5})$$

Therefore, formally, we have already obtained the correspondence between the minimiser of the free energy function and the invariant measure of (V.1.5). In this study, the connection is rigorously proved mainly with a probabilistic argument.

For the particular application to the neural network (V.1.3), it is crucial to observe that the dyna-

mics corresponding to the mean field Langevin dynamics describes exactly the path of the randomised regularized gradient-descent algorithm. More precisely, consider the case where we are given data points  $(y_m, z_m)_{m \in \mathbb{N}}$  which are i.i.d. samples from  $\nu$ . If the loss function  $\Phi$  is simply the square loss then a version of the (randomized, regularized) gradient descent algorithm for the evolution of parameter  $x_k^i$  will simply read as

$$x_{k+1}^i = x_k^i + 2\tau \left( \left( y_k - \frac{1}{N} \sum_{j=1}^N \hat{\varphi}(x_k^j, z_k) \right) \nabla \hat{\varphi}(x_k^i, z_k) - \frac{\sigma^2}{2} \nabla U(x_k^i) \right) + \sigma \sqrt{\tau} \xi_k^i, \quad (\text{V.1.6})$$

with  $\xi_k^i$  independent samples from  $N(0, I_d)$  (for details we refer the reader to Section V.3.2). This evolution is an approximation of (V.1.5) and can be viewed as noisy gradient decent. Indeed, in its original form, the classical stochastic gradient decent (also known as the Robins–Monroe algorithm), is given by (V.1.6) with  $\sigma = 0$ .

### V.1.1 Organisation of the Chapter

The introduction is concluded by Section V.1.2, where we compare the findings to those available in the literature, and by Section V.1.3 recall some basic notions of measure derivatives. All the main results are presented in Section V.2. In Section V.3 we show how the results in Section V.2 apply to in the case of gradient descent training of neural networks. Section V.4 contains all the proofs of the results concerning the free energy function :  $\Gamma$ -convergence when  $\sigma \rightarrow 0$ , particle approximation and the first order condition. In Section V.5 we prove required properties of (V.1.4) and (V.1.5), Section V.6 is used to prove the convergence of the solution to (V.1.4) to an invariant measure which is the minimizer of the free energy function.

### V.1.2 Theoretical Contributions and Literature Review

The study of stationary solutions to nonlocal, diffusive equations (V.1.4) is classical topic with its roots in statistical physics literature and with strong links to Kac’s program in Kinetic theory [120]. We also refer reader to excellent monographs [6] and [3]. In particular, variational approach has been developed in [28, 127, 151] where authors studied dissipation of entropy for granular media equations with the symmetric interaction potential of convolution type (interaction potential corresponds to term  $D_m F$  in (V.1.4)). We also refer a reader to similar results with proofs based on particle approximation of [29, 153, 14], coupling arguments [49] and Khasminskii’s technique [23, 13]. All of the above results impose restrictive condition on interaction potential or/and require it to be sufficiently small. We manage to relax these assumptions allowing for the interaction potential to be arbitrary (but sufficiently regular/bounded) function of measure. Our proof is probabilistic in nature. Using Lasalle’s invariance principle and the HWI inequality from Otto and Villani [128] as the main ingredients, we prove the desired convergence. This approach, to our knowledge, is original, and it clearly justifies the solvability of the randomized/regularized

gradient descent algorithm for neural networks. Finally we clarify how different notions of calculus on the space of probability measures enter our framework. The calculus is critical to work with arbitrary functions of measure. We refer to [26, Chapter 5] for an overview on that topic. The calculus on the measure space enables to derive and quantify the error between finite dimensional optimisation problem and its infinite dimensional limit.

Other results are now available for the mean-field description of non-convex learning problems, see [117, 116, 34, 138, 90, 143]. Let us compare our study to the key results available in the literature. There are essentially three, or, if entropic regularization is included, four key ingredients :

- i) that the finite dimensional optimisation problem is approximated by infinite dimensional problem of minimizing over measures (Theorem V.2.4),
- ii) that the regularized problem approximates the original minimization problem (Proposition V.2.3),
- iii) that on the space of probability measures the minimizers (or, if entropic regularization is included, the unique minimizer) satisfy a first order condition (Proposition V.2.5),
- iv) and finally that on the space of probability measures we have a gradient flow that converges with time to the minimizer (Theorem V.2.11).

Chizat and Bach [34] work without adding entropic regularization which means that their minimization task is convex but not strictly convex. They have results regarding i) and iii). They have a partial result related to iv) in that they prove that if the gradient flow converges to a limit, as  $t \rightarrow \infty$  then objective function also converge to global minimiser. To prove this final convergence result they require the assumption that the activation function is homogenous of either order 1 or order 2 and that, essentially, initial law has full support. The setting used in Chizat and Bach [34] is rather different to that of the results in the present work : in particular we do not need to assume any homogeneity on the activation function and apart mild integrability conditions we do not make assumptions on the initial law. Since we regularize using entropy we obtain convergence of the gradient flow to the minimizer.

Rotskoff and Vanden-Eijnden [138] again work without entropic regularization and have results i). Moreover they provide Central-Limit-Theorem-type fluctuation results. They do show that the output of the network converges to a limit as the time in the gradient flow for the parameter measure goes to infinity. However they do not prove convergence of the parameter measure itself as in iv).

Sirignano and Spiliopoulos in [143] provide detail analysis of i), also studying time-discretisation of continuous time gradient flow. In Section V.3.2, by using links between intrinsic derivative on the space of measure and its finite dimensional projection we provide further insight on the choice of scalings needed to derive non-trivial limit in [143].

The setting of our work is the closest to that of Mei, Misiakiewicz and Montanari [117] in that they use the entropic regularization. For a square loss function  $\Phi$  and a quadratic regularizer  $U$  they prove results on all of i), ii), iii) and iv). In this work we allow a general loss function for all the above results(that this is possible is conjectured in Appendix B of [117]). Due to the special choice of the square loss function, in [117, Lemma 6.10] the authors can compute directly the dynamics of  $F(m_t)$

along the flow of measures defined by (V.1.4). Instead, we obtain the desired dynamics for much more general  $F$  using a pathwise argument based on the Itô calculus (see Theorem V.2.9). The proof of iv) in [117, Lemma 6.12] is based on the Poincaré inequality for the Gaussian distribution and shows that the marginal density converges in  $L^1$ . It is not clear whether this argument can be extended to the case with a general regularizer  $U$ , whereas this work develops a new technique based on LaSalle's invariance principle and the HWI inequality, which allows us to prove the convergence for general  $U$  in the Wasserstein-2 metric, and moreover we observe that in the highly regularized case this convergence is exponential, see Theorem V.2.11.

### V.1.3 Calculus on the Space of Probability Measures

By  $\mathcal{P}(\mathbb{R}^d)$  we denote the space of probability measures on  $\mathbb{R}^d$ , and by  $\mathcal{P}_p(\mathbb{R}^d)$  the subspace of  $\mathcal{P}(\mathbb{R}^d)$  in which the measures have finite  $p$ -moment for  $p \geq 1$ . Note that  $\pi \in \mathcal{P}_p(\mathbb{R}^d \times \mathbb{R}^d)$  is called a coupling of  $\mu$  and  $\nu$  in  $\mathcal{P}_p(\mathbb{R}^d)$ , if for any borel subset  $B$  of  $\mathbb{R}^d$  we have  $\pi(B, \mathbb{R}^d) = \mu(B)$  and  $\pi(\mathbb{R}^d, B) = \nu(B)$ . By  $\mathcal{W}_p$  we denote the Wasserstein- $p$  metric on  $\mathcal{P}_p(\mathbb{R}^d)$ , namely,

$$\mathcal{W}_p(\mu, \nu) := \inf \left\{ \left( \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^p \pi(dx, dy) \right)^{\frac{1}{p}} ; \pi \text{ is a coupling of } \mu \text{ and } \nu \right\} \quad \text{for } \mu, \nu \in \mathcal{P}_p(\mathbb{R}^d).$$

It is convenient to recall that

- i)  $(\mathcal{P}_p(\mathbb{R}^d), \mathcal{W}_p)$  is a Polish space;
- ii)  $\mathcal{W}_p(\mu_n, \mu) \rightarrow 0$  if and only if  $\mu_n$  weakly converge to  $\mu$  and  $\int_{\mathbb{R}^d} |x|^p \mu_n(dx) \rightarrow \int_{\mathbb{R}^d} |x|^p \mu(dx)$ ;
- iii) for  $p' > p$ , the set  $\{\mu \in \mathcal{P}_p(\mathbb{R}^d) : \int_{\mathbb{R}^d} |x|^{p'} \mu(dx) \leq C\}$  is  $\mathcal{W}_p$ -compact.

We say a function  $F : \mathcal{P}(\mathbb{R}^d) \rightarrow \mathbb{R}$  is in  $\mathcal{C}^1$  if there exists a bounded continuous function  $\frac{\delta F}{\delta m} : \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d \rightarrow \mathbb{R}$  such that

$$F(m') - F(m) = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta F}{\delta m}((1 - \lambda)m + \lambda m', x) (m' - m)(dx) d\lambda. \quad (\text{V.1.7})$$

We will refer to  $\frac{\delta F}{\delta m}$  as the linear functional derivative. There is at most one  $\frac{\delta F}{\delta m}$ , up to a constant shift, satisfying (V.1.7). To avoid the ambiguity, we impose

$$\int_{\mathbb{R}^d} \frac{\delta F}{\delta m}(m, x) m(dx) = 0.$$

If  $(m, x) \mapsto \frac{\delta F}{\delta m}(m, x)$  is continuously differentiable in  $x$ , we define its intrinsic derivative  $D_m F : \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  by

$$D_m F(m, x) = \nabla \left( \frac{\delta F}{\delta m}(m, x) \right).$$

In this paper  $\nabla$  always denotes the gradient in the variable  $x \in \mathbb{R}^d$ .

**Example V.1.0.1.** If  $F(m) := \int_{\mathbb{R}^d} \phi(x) m(dx)$  for some bounded continuous function  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ , we have  $\frac{\delta F}{\delta m}(m, x) = \phi(x)$  and  $D_m F(m, x) = \dot{\phi}(x)$ .

It is useful to see what intrinsic measure derivative look like in the special case when we consider empirical measures

$$m^N := \frac{1}{N} \sum_{i=1}^N \delta_{x^i}, \quad \text{where } x^i \in \mathbb{R}^d.$$

Then one can define  $F^N : (\mathbb{R}^d)^N \rightarrow \mathbb{R}$  as  $F^N(x^1, \dots, x^N) = F(m^N)$ . From [30, Proposition 3.1] we know that that if  $F \in \mathcal{C}^1$  then  $F^N \in \mathcal{C}^1$  and for any  $i = 1, \dots, N$  and  $(x^1, \dots, x^N) \in (\mathbb{R}^d)^N$  it holds that

$$\partial_{x^i} F^N(x^1, \dots, x^N) = \frac{1}{N} D_m F(m^N, x^i). \quad (\text{V.1.8})$$

We remark that for notational simplicity in the proofs the constant  $C > 0$  can be different from line to line.

## V.2 Main Results

The objective of this chapter is to study the minimizer(s) of a convex function  $F : \mathcal{P}(\mathbb{R}^d) \rightarrow \mathbb{R}$ .

**Assumption V.2.1.** Assume that  $F \in \mathcal{C}^1$  is convex and bounded from below.

Instead of directly considering the minimization  $\min_m F(m)$ , we propose to first study the regularized version, namely, the minimization of the free energy function :

$$\min_{m \in \mathcal{P}(\mathbb{R}^d)} V^\sigma(m), \quad \text{where } V^\sigma(m) := F(m) + \frac{\sigma^2}{2} H(m), \quad \text{for all } m \in \mathcal{P}(\mathbb{R}^d),$$

where  $H : \mathcal{P}(\mathbb{R}^d) \rightarrow [0, \infty]$  is the relative entropy (Kullback–Leibler divergence) with respect to a given Gibbs measure in  $\mathbb{R}^d$ , namely,

$$H(m) := \int_{\mathbb{R}^d} m(x) \log \left( \frac{m(x)}{g(x)} \right) dx,$$

where

$$g(x) = e^{-U(x)} \quad \text{with } U \text{ s.t. } \int_{\mathbb{R}^d} e^{-U(x)} dx = 1,$$

is the density of the Gibbs measure and the function  $U$  satisfies the following conditions.

**Assumption V.2.2.** The function  $U : \mathbb{R}^d \rightarrow \mathbb{R}$  belongs to  $C^\infty$ . Further,

i) there exist constants  $C_U > 0$  and  $C'_U \in \mathbb{R}$  such that

$$\nabla U(x) \cdot x \geq C_U |x|^2 + C'_U \quad \text{for all } x \in \mathbb{R}^d. \quad (\text{V.2.9})$$



ii)  $\nabla U$  is Lipschitz continuous.

Immediately, we obtain that there exist  $0 \leq C' \leq C$  such that for all  $x \in \mathbb{R}^d$

$$C'|x|^2 - C \leq U(x) \leq C(1 + |x|^2), \quad |\Delta U(x)| \leq C.$$

A typical choice of  $g$  would be the density of the  $d$ -dimensional standard Gaussian distribution. We recall that such relative entropy  $H$  has the properties : it is strictly convex when restricted to measures absolutely continuous with  $g$ , it is weakly lower semi-continuous and its sub-level sets are compact. For more details, we refer the readers to the book [45, Section 1.4]. The original minimization and the regularized one is connected through the following  $\Gamma$ -convergence result.

**Proposition V.2.3.** *Assume that  $F$  is continuous in the topology of weak convergence. Then the sequence of functions  $V^\sigma = F + \frac{\sigma^2}{2}H$   $\Gamma$ -converges to  $F$  when  $\sigma \downarrow 0$ . In particular, given the minimizer  $m^{*,\sigma}$  of  $V^\sigma$ , we have*

$$\limsup_{\sigma \rightarrow 0} F(m^{*,\sigma}) = \inf_{m \in \mathcal{P}_2(\mathbb{R}^d)} F(m).$$

It is a classic property of  $\Gamma$ -convergence that every cluster point of  $\left(\arg\min_m V^\sigma(m)\right)_\sigma$  is a minimizer of  $F$ .

The following theorem shows that we can control the error between the finite and infinite-dimensional optimization problems. It generalises [117, Proposition 2.1] to an arbitrary (smooth) functions of measure. It is an extension of the result from [32, Theorem 2.11].

**Theorem V.2.4.** *We assume that the 2nd order linear functional derivative of  $F$  exists, is jointly continuous in both variables and that there is  $L > 0$  such that for any random variables  $\eta_1, \eta_2$  such that  $\mathbb{E}[|\eta_i|^2] < \infty$ ,  $i = 1, 2$ , it holds that*

$$\mathbb{E} \left[ \sup_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} \left| \frac{\delta F}{\delta m}(\nu, \eta_1) \right| \right] + \mathbb{E} \left[ \sup_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} \left| \frac{\delta^2 F}{\delta m^2}(\nu, \eta_1, \eta_2) \right| \right] \leq L \quad (\text{V.2.10})$$

*If there is an  $m^* \in \mathcal{P}_2(\mathbb{R}^d)$  such that  $F(m^*) = \inf_{m \in \mathcal{P}_2(\mathbb{R}^d)} F(m)$  then with i.i.d  $(X_i^*)_{i=1}^N$  such that  $X_i^* \sim m^*$ ,  $i = 1, \dots, N$  we have that*

$$\left| \inf_{(x_i)_{i=1}^N \subset \mathbb{R}^d} F \left( \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \right) - F(m^*) \right| \leq \frac{2L}{N}.$$

Moreover, when the relative entropy  $H$  is strictly convex, then so is the function  $V$ , and thus the minimizer  $\arg\min_{m \in \mathcal{P}(\mathbb{R}^d)} V(m)$ , if exists, must be unique. It can be characterized by the following first order condition.

**Proposition V.2.5.** *Under Assumption V.2.1 and V.2.2, the function  $V^\sigma$  has a unique minimizer absolutely continuous with respect to Lebesgue measure  $\ell$ , and belonging to  $\mathcal{P}_2(\mathbb{R}^d)$ . Moreover,  $m^* \in \mathcal{P}_2(\mathbb{R}^d) = \operatorname{argmin}_{m \in \mathcal{P}(\mathbb{R}^d)} V^\sigma(m)$  if and only if  $m^*$  is equivalent to Lebesgue measure and*

$$\frac{\delta F}{\delta m}(m^*, \cdot) + \frac{\sigma^2}{2} \log(m^*) + \frac{\sigma^2}{2} U \text{ is a constant, } \ell - a.s., \quad (\text{V.2.11})$$

where we abuse the notation, still denoting by  $m^*$  the density with respect to Lebesgue measure.

Further, we are going to approximate the minimizer of  $V^\sigma$ , using the marginal laws of the solution to the upcoming mean field Langevin equation. Let  $\sigma \in \mathbb{R}_+$  and consider the following McKean–Vlasov SDE :

$$dX_t = - \left( D_m F(m_t, X_t) + \frac{\sigma^2}{2} \nabla U(X_t) \right) dt + \sigma dW_t, \quad (\text{V.2.12})$$

where  $m_t$  is the law of  $X_t$  and  $(W_t)_{t \geq 0}$  is a standard  $d$ -dimensional Brownian motion.

**Remark V.2.6.** *i) Let  $F(m) = \int_{\mathbb{R}^d} f(x)m(dx)$  for some function  $f$  in  $C^1(\mathbb{R}^d, \mathbb{R})$ . We know that  $D_m F(m, x) = \nabla f(x)$ . Hence with this choice of  $F$  and entropy regulariser with respect to the Lebesgue measure, the dynamics (V.2.12) becomes the standard overdamped Langevin equation (V.1.1).*

*ii) If the Gibbs measure is chosen to be a standard Gaussian distribution, the potential of the drift of (V.2.12) becomes  $F(m) + \frac{\sigma^2}{4} \int_{\mathbb{R}^d} |x|^2 m(dx)$ . This shares the same spirit as ridge regression.*

**Assumption V.2.7.** *Assume that the intrinsic derivative  $D_m F : \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  of the function  $F : \mathcal{P}(\mathbb{R}^d) \rightarrow \mathbb{R}$  exists and satisfies the following conditions :*

*i)  $D_m F$  is bounded and Lipschitz continuous, i.e. there exists  $C_F > 0$  such that for all  $x, x' \in \mathbb{R}^d$  and  $m, m' \in \mathcal{P}_2(\mathbb{R}^d)$*

$$|D_m F(m, x) - D_m F(m', x')| \leq C_F (|x - x'| + \mathcal{W}_2(m, m')) \quad (\text{V.2.13})$$

*ii)  $D_m F(m, \cdot) \in C^\infty(\mathbb{R}^d)$  for all  $m \in \mathcal{P}(\mathbb{R}^d)$  ;*

*iii)  $\nabla D_m F : \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d \rightarrow \mathbb{R}^d \times \mathbb{R}^d$  is jointly continuous.*

The well-posedness of the McKean–Vlasov SDE (V.2.12) under Assumption V.2.2 and V.2.7 on the time interval  $[0, t]$ , for any  $t$ , is well known, see e.g. Snitzman [148], so the proof of the following proposition is omitted.

**Proposition V.2.8.** *Under Assumption V.2.2 and V.2.7 the mean field Langevin SDE (V.2.12) has a unique strong solution, if  $m_0 \in \mathcal{P}_2(\mathbb{R}^d)$ . Moreover, the solution is stable with respect to the initial law,*

that is, given  $m_0, m'_0 \in \mathcal{P}_2(\mathbb{R}^d)$ , denoting by  $(m_t)_{t \in \mathbb{R}^+}, (m'_t)_{t \in \mathbb{R}^+}$  the marginal laws of the corresponding solutions to (V.2.12), we have for all  $t > 0$  there is a constant  $C > 0$  such that

$$\mathcal{W}_2(m_t, m'_t) \leq C\mathcal{W}_2(m_0, m'_0).$$

We shall prove the process  $(V^\sigma(m_t))_t$  is decreasing and satisfies the following dynamic.

**Theorem V.2.9.** *Let  $m_0 \in \mathcal{P}_2(\mathbb{R}^d)$ . Under Assumption V.2.2 and V.2.7, we have for any  $t > s > 0$*

$$V^\sigma(m_t) - V^\sigma(m_s) = - \int_s^t \int_{\mathbb{R}^d} \left| D_m F(m_r, x) + \frac{\sigma^2}{2} \frac{\nabla m_r}{m_r}(x) + \frac{\sigma^2}{2} \nabla U(x) \right|^2 m_r(x) dx dr. \quad (\text{V.2.14})$$

**Remark V.2.10.** *In order to prove (V.2.14), we use the generalized Itô calculus as the main tool. Alternative proofs of Theorem V.2.9 can be obtained under the comparable assumptions using the theory of gradient flows, see e.g. the monograph [3]. Also note that our path-wise argument shares the spirit with the recent work [94], which recovers the results of gradient flows for probability measure on the Euclidean space using the Itô calculus but only for linear functional  $F$ .*

Formally, there is a clear connection between the derivative  $\frac{dV^\sigma(m_t)}{dt}$  in (V.2.14) and the first order condition (V.2.11), and it is revealed by the following main theorem.

We call a measure  $m$  an invariant measure of (V.2.12), if  $\text{Law}(X_t) = m$  for all  $t \geq 0$ .

**Theorem V.2.11.** *Let Assumption V.2.1, V.2.2 and V.2.7 hold true and  $m_0 \in \cup_{p>2} \mathcal{P}_p(\mathbb{R}^d)$ . Denote by  $(m_t)_{t \in \mathbb{R}^+}$  the flow of marginal laws of the solution to (V.2.12).*

- *Then, there exists an invariant measure of (V.2.12) equal to  $m^* := \underset{m}{\operatorname{argmin}} V^\sigma(m)$ , and  $\lim_{t \rightarrow \infty} \mathcal{W}_2(m_t, m^*) = 0$ .*
- *Further assume that*

$$|D_m F(m, x) - D_m F(m', x')| \leq C_F |x - x'| + \rho_F \mathcal{W}_1(m, m').$$

*Then, provided that  $\rho_F$  is small enough, we have the contraction :*

$$\mathcal{W}_1(m_t, m'_t) \leq C e^{-\gamma t} \mathcal{W}_1(m_0, m'_0), \quad (\text{V.2.15})$$

*where  $(m'_t)_{t \in \mathbb{R}^+}$  is the flow of marginal laws of the solution to (V.2.12) with the initial law  $m'_0$ .*

**Remark V.2.12.** *As mentioned, the main theoretical contribution of this work is to prove the  $\mathcal{W}_2$ -convergence of the marginal laws of (V.2.12) towards the invariant measure under the mild conditions (Assumption V.2.1, V.2.2 and V.2.7). With the extra condition of the small dependence on the marginal laws, the latter exponential convergence result (V.2.15) was essentially proved in [49] with a reflection coupling argument, so its proof is omitted. Note that only in case that the corresponding optimization*

problem is largely regularized, that is the diffusion constant  $\sigma$  or/and the dissipative constant  $C_U$  in (V.2.9) is comparatively (much) bigger than  $\rho_F$ , the condition for the exponential convergence can hold true.

### V.3 Application to Gradient Descent of Neural Networks

Before proving the main results, we shall first apply them to study the minimization over a neural network. In particular, in Corollary V.3.3 we shall show that the marginal laws of the corresponding mean-field Langevin dynamics converge to the optimal weight of the neural network with 1-hidden layer.

Fix a locally Lipschitz function  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  and for  $l \in \mathbb{N}$  define  $\varphi^l : \mathbb{R}^l \rightarrow \mathbb{R}^l$  as the function given, for  $z = (z_1, \dots, z_l)^\top$  by  $\varphi^l(z) = (\varphi(z_1), \dots, \varphi(z_l))^\top$ . We fix  $L \in \mathbb{N}$  (the number of layers),  $l_k \in \mathbb{N}$ ,  $k = 0, 1, \dots, L-1$  (the size of input to layer  $k$ ) and  $l_L \in \mathbb{N}$  (the size of the network output). A fully connected artificial neural network is then given by  $\Psi = ((\alpha^1, \beta^1), \dots, (\alpha^L, \beta^L)) \in \Pi$ , where, for  $k = 1, \dots, L$ , we have real  $l^k \times l^{k-1}$  matrices  $\alpha^k$  and real  $l^k$ -dimensional vectors  $\beta^k$ . We see that  $\Pi = (\mathbb{R}^{l^1 \times l^0} \times \mathbb{R}^{l^1}) \times (\mathbb{R}^{l^2 \times l^1} \times \mathbb{R}^{l^2}) \times \dots \times (\mathbb{R}^{l^L \times l^{L-1}} \times \mathbb{R}^{l^L})$ . The artificial neural network defines a reconstruction function  $\mathcal{R}\Psi : \mathbb{R}^{l^0} \rightarrow \mathbb{R}^{l^L}$  given recursively, for  $z_0 \in \mathbb{R}^{l^0}$ , by

$$(\mathcal{R}\Psi)(z^0) = \alpha^L z^{L-1} + \beta^L, \quad z^k = \varphi^{l^k}(\alpha^k z^{k-1} + \beta^k), \quad k = 1, \dots, L-1.$$

If for each  $k = 1, \dots, L-1$  we write  $\alpha_i^k, \beta_i^k$  to denote the  $i$ -th row of the matrix  $\alpha^k$  and vector  $\beta^k$  respectively then we can write the reconstruction of the network equivalently as

$$(\mathcal{R}\Psi)(z^0)_i = \alpha_i^L \cdot z^{L-1} + \beta_i^L, \quad (z^k)_i = \varphi(\alpha_i^k \cdot z^{k-1} + \beta_i^k), \quad k = 1, \dots, L-1. \quad (\text{V.3.16})$$

We note that the number of parameters in the network is  $\sum_{i=1}^L (l_{k-1} l_k + l_k)$ .

Given a potential function  $\Phi$  and training data  $(y^j, z^j)_{j=1}^N, (y_j, z_j) \in \mathbb{R}^d$  one approximates the optimal parameters by finding

$$\operatorname{argmin}_{\Psi \in \Pi} \frac{1}{N} \sum_{j=1}^N \Phi(y^j - (\mathcal{R}\Psi)(z^j)). \quad (\text{V.3.17})$$

This is a non-convex minimization problem, so in general hard to solve. Theoretically, the following universal representation theorem ensures that the minimum value should attain 0, provided that  $y = f(z)$  with a continuous function  $f$ .

**Theorem V.3.1** (Universal Representation Theorem). *If an activation function  $\varphi$  is bounded, continuous and non-constant, then for any compact set  $K \subset \mathbb{R}^d$  the set*

$$\left\{ (\mathcal{R}\Psi) : \mathbb{R}^d \rightarrow \mathbb{R} : (\mathcal{R}\Psi) \text{ given by (V.3.16) with } L = 2 \text{ for some } n \in \mathbb{N}, \alpha_j^2, \beta_j^1 \in \mathbb{R}, \alpha_j^1 \in \mathbb{R}^d, j = 1, \dots, n \right\}$$

is dense in  $C(K)$ .

For an elementary proof, we refer the readers to [79, Theorem 2].

### V.3.1 Fully connected 1-hidden layer neural network

Take  $L = 2$ , fix  $d \in \mathbb{N}$  and  $n \in \mathbb{N}$  and consider the following 1-hidden layer neural network for approximating functions from  $\mathbb{R}^d$  to  $\mathbb{R}$ : let  $l_0 = d$ , let  $l_1 = n$ , let  $\beta^2 = 0 \in \mathbb{R}$ ,  $\beta^1 = 0 \in \mathbb{R}^n$ ,  $\alpha^1 \in \mathbb{R}^{n \times d}$ . We will denote, for  $i \in \{1, \dots, l^0\}$ , its  $i$ -th row by  $\alpha_i^1 \in \mathbb{R}^{1 \times d}$ . Let  $\alpha^2 = (\frac{c_1}{n}, \dots, \frac{c_n}{n})^\top$ , where  $c_i \in \mathbb{R}$ . The neural network is  $\Psi^n = \left( (\alpha^1, \beta^1), (\alpha^2, \beta^2) \right)$  (where we emphasise the that the size of the hidden layer is  $n$ ). For  $z \in \mathbb{R}^{l^0}$ , its reconstruction can be written as

$$(\mathcal{R}\Psi^n)(z) = \alpha^2 \varphi^{l^1}(\alpha^1 z) = \frac{1}{n} \sum_{i=1}^n c_i \varphi(\alpha_i^1 \cdot z).$$

The key observation is to note that, due to law of large numbers (and under appropriate technical assumptions)  $\frac{1}{n} \sum_{j=1}^n c_j \varphi(\alpha_j^1 \cdot z) \rightarrow \mathbb{E}^m[B\varphi(A \cdot z)]$  as  $n \rightarrow \infty$ , where  $m$  is the law of the pair of random variables  $(B, A)$  and  $\mathbb{E}^m$  is the expectation under the measure  $m$ . Therefore, another way (indeed a more intrinsic way regarding to the universal representation theorem) to formulate the minimization problem (V.3.17) is :

$$\min_{m \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R})} \tilde{F}(m), \quad \text{where} \quad \tilde{F}(m) := \int_{\mathbb{R}^d} \Phi\left(y - \mathbb{E}^m[B\varphi(A \cdot z)]\right) \nu(dz, dy).$$

For technical reason, we introduce a truncation function  $\ell : \mathbb{R} \rightarrow K$  ( $K$  denotes again some compact set), and consider the truncated version of the minimization :

$$F(m) := \int_{\mathbb{R}^d} \Phi\left(y - \mathbb{E}^m[\ell(B)\varphi(A \cdot z)]\right) \nu(dz, dy).$$

It is crucial to note that in the reformulation the objective function  $F$  becomes a convex function on  $\mathcal{P}(\mathbb{R}^d)$ , provided that  $\Phi$  is convex.

**Assumption V.3.2.** *We apply the following assumptions on the coefficients  $\Phi, \mu, \varphi, \ell$  :*

- i) *the function  $\Phi$  is convex, smooth and  $0 = \Phi(0) = \min_{a \in \mathbb{R}} \Phi(a)$ ;*
- ii) *the data measure  $\mu$  is of compact support;*
- iii) *the truncation function  $\ell \in C_b^\infty(\mathbb{R}^d)$  such that  $\dot{\ell}$  and  $\ddot{\ell}$  are bounded;*
- iv) *the activation function  $\varphi \in C_b^\infty(\mathbb{R}^d)$  such that  $\dot{\varphi}$  and  $\ddot{\varphi}$  are bounded.*

**Corollary V.3.3.** *Under Assumption V.3.2, the function  $F$  satisfies Assumption V.2.1, V.2.7. In particular, with a Gibbs measure of which the function  $U$  satisfies Assumption V.2.2, the corresponding mean field Langevin equation (V.2.12) admits a unique strong solution, given  $m_0 \in \mathcal{P}_2(\mathbb{R}^d)$ . Moreover, the*

flow of marginal laws of the solution,  $(m_t)_{t \in \mathbb{R}^+}$ , satisfies

$$\lim_{t \rightarrow +\infty} \mathcal{W}_2 \left( m_t, \operatorname{argmin}_{m \in \mathcal{P}(\mathbb{R}^d)} V^\sigma(m) \right) = 0.$$

*Démonstration.* Let us define, for  $x = (\beta, \alpha) \in \mathbb{R}^d$ ,  $\beta \in \mathbb{R}$ ,  $\alpha \in \mathbb{R}^{d-1}$  and  $z \in \mathbb{R}^{d-1}$  the function  $\hat{\varphi}(x, z) := \ell(\beta)\varphi(\alpha \cdot z)$ . Then

$$\begin{aligned} \frac{\delta F}{\delta m}(m, x) &= - \int_{\mathbb{R}^d \times \mathbb{R}} \dot{\Phi} \left( y - \mathbb{E}^m[\hat{\varphi}(X, z)] \right) \hat{\varphi}(x, z) \nu(dz, dy) \\ \text{and } D_m F(m, x) &= - \int_{\mathbb{R}^d \times \mathbb{R}} \dot{\Phi} \left( y - \mathbb{E}^m[\hat{\varphi}(X, z)] \right) \nabla \hat{\varphi}(x, z) \nu(dz, dy). \end{aligned}$$

Then it becomes straightforward to verify that  $F$  satisfies both Assumption V.2.1, V.2.7. The rest of the result is direct from Proposition V.2.8 and Theorem V.2.11. ■

### V.3.2 Gradient Descent

Consider independent random variables  $(X_0^i)_{i=1}^N$ ,  $X_0^i \sim m_0$  and independent Brownian motions  $(W^i)_{i=1}^N$ . By approximating the law of the process (V.2.12) by its empirical law we arrive at the following interacting particle system

$$\begin{cases} dX_t^i = - \left( D_m F(m_t^N, X_t^i) + \frac{\sigma^2}{2} \nabla U(X_t^i) \right) dt + \sigma dW_t^i & i = 1, \dots, N, \\ m_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_t^i}. \end{cases} \quad (\text{V.3.18})$$

Note that particles  $(X^i)_{i=1}^N$  are not independent, but their laws are exchangeable. Recall the link between partial derivatives and measure derivative given by (V.1.8) and for any  $(x^1, \dots, x^N) \in (\mathbb{R}^d)^N$  let  $F^N(x^1, \dots, x^N) = F \left( \frac{1}{N} \sum_{i=1}^N \delta_{x^i} \right)$ . Then

$$dX_t^i = - \left( N \partial_{x^i} F^N(X_t^1, \dots, X_t^N) + \frac{\sigma^2}{2} \nabla U(X_t^i) \right) dt + \sigma dW_t^i.$$

Let us define, for  $x = (\beta, \alpha) \in \mathbb{R}^d$ ,  $\beta \in \mathbb{R}$ ,  $\alpha \in \mathbb{R}^{d-1}$  and  $z \in \mathbb{R}^{d-1}$  the function  $\hat{\varphi}(x, z) := \ell(\beta)\varphi(\alpha \cdot z)$ . Then for  $(x^i)_{i=1}^N$  we have

$$F^N(x) = \int_{\mathbb{R}^d} \Phi \left( y - \frac{1}{N} \sum_{j=1}^N \hat{\varphi}(x^j, z) \right) \nu(dz, dy).$$

Hence

$$\partial_{x^i} F^N(x^1, \dots, x^N) = - \frac{1}{N} \int_{\mathbb{R}^d} \dot{\Phi} \left( y - \frac{1}{N} \sum_{j=1}^N \hat{\varphi}(x^j, z) \right) \nabla \hat{\varphi}(x^i, z) \nu(dz, dy),$$

where we denote for all  $z \in \mathbb{R}^{d-1}$

$$\nabla \hat{\varphi}(x^i, z) = \nabla_{(\beta^i, \alpha^i)} [\ell(\beta^i) \varphi(\alpha^i z)] = \begin{pmatrix} \ell(\beta^i) \varphi(\alpha^i z) \\ \ell(\beta^i) \dot{\varphi}(\alpha^i z) \end{pmatrix}.$$

We thus see that (V.3.18) corresponds to

$$dX_t^i = \left( \int_{\mathbb{R}^d} \dot{\Phi} \left( y - \frac{1}{N} \sum_{j=1}^N \hat{\varphi}(X_t^j, z) \right) \nabla \hat{\varphi}(X_t^i, z) \nu(dz, dy) - \frac{\sigma^2}{2} \nabla U(X_t^i) \right) dt + \sigma dW_t^i.$$

This is classical Langevin dynamics (V.1.1) on  $(\mathbb{R}^d)^N$ . One may reasonably expect that a version of Theorem V.2.4 can be proved in this dynamical setup. This has been done for finite time horizon problem in [32]. The extension to the infinite horizon requires uniform in time regularity of the corresponding PDE on Wasserstein space  $(\mathcal{W}_2, \mathcal{P}_2)$  and we leave it for a future research. However rate for uniform propagation of chaos in  $\mathcal{W}_1$  under structural condition on the drift has been proved in [46]. We also remark that for the implementable algorithm one works with time discretisation of (V.3.18) and, at least for the finite time, the error bounds are rather well understood [19, 18, 114, 149, 150].

For a fixed time step  $\tau > 0$  fixing a grid of time points  $t_k = k\tau$ ,  $k = 0, 1, \dots$  we can then write the explicit Euler scheme

$$\begin{aligned} X_{t_{k+1}}^{\tau, i} - X_{t_k}^{\tau, i} \\ = \left( \int_{\mathbb{R}^d} \dot{\Phi} \left( y - \frac{1}{N} \sum_{j=1}^N \hat{\varphi}(X_{t_k}^{\tau, j}, z) \right) \nabla \hat{\varphi}(X_{t_k}^{\tau, i}, z) \nu(dz, dy) - \frac{\sigma^2}{2} \nabla U(X_{t_k}^{\tau, i}) \right) \tau + \sigma (W_{t_{k+1}}^i - W_{t_k}^i). \end{aligned}$$

To relate this to the gradient descent algorithm we consider the case where we are given data points  $(y_m, z_m)_{m \in \mathbb{N}}$  which are i.i.d. samples from  $\nu$ . If the loss function  $\Phi$  is simply the square loss then a version of the (regularized) gradient descent algorithm for the evolution of parameter  $x_k^i$  will simply read as

$$x_{k+1}^i = x_k^i + 2\tau \left( \left( y_k - \frac{1}{N} \sum_{j=1}^N \hat{\varphi}(x_k^j, z_k) \right) \nabla \hat{\varphi}(x_k^i, z_k) - \frac{\sigma^2}{2} \nabla U(x_k^i) \right) + \sigma \sqrt{\tau} \xi_k^i,$$

with  $\xi_k^i$  independent samples from  $N(0, I_d)$ .

## V.4 Free Energy Function

In this section, we study the properties concerning the minimizer of the free energy function  $V^\sigma$ . First, we prove that  $V^\sigma$  is an approximation of  $F$  in the sense of  $\Gamma$ -convergence.

*Proof of Proposition V.2.3.* Let  $(\sigma_n)_{n \in \mathbb{N}}$  be a positive sequence decreasing to 0. On the one hand, since

$F$  is continuous and  $H(m) \geq 0$ , for all  $m_n \rightarrow m$ , we have

$$\liminf_{n \rightarrow +\infty} V^{\sigma_n}(m_n) \geq \lim_{n \rightarrow +\infty} F(m_n) = F(m).$$

On the other hand, given  $m \in \mathcal{P}_2(\mathbb{R}^d)$ , since the function

$$h(x) := x \log(x) \tag{V.4.19}$$

is convex, it follows Jensen's inequality that

$$\int_{\mathbb{R}^d} h(m * f_n) dx \leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} h(f_n(x-y)) m(dy) dx = \int_{\mathbb{R}^d} h(f_n(x)) dx = \int_{\mathbb{R}^d} h(f(x)) dx - d \log(\sigma_n),$$

where  $f$  is the heat kernel and  $f_n(x) = \sigma_n^{-d} f(x/\sigma_n)$ . Besides, we have

$$\int_{\mathbb{R}^d} (m * f_n) \log(g) dx = - \int_{\mathbb{R}^d} m(dy) \int_{\mathbb{R}^d} f_n(x) U(x-y) dx \geq -C \left( 1 + \int_{\mathbb{R}^d} |y|^2 m(dy) \right).$$

The last inequality is due to the quadratic growth of  $U$ . Therefore

$$\limsup_{n \rightarrow +\infty} V^{\sigma_n}(m * f_n) \leq F(m) + \limsup_{n \rightarrow +\infty} \frac{\sigma_n^2}{2} \left\{ \int_{\mathbb{R}^d} h(m * f_n) dx - \int_{\mathbb{R}^d} (m * f_n) \log(g) dx \right\} \leq F(m). \tag{V.4.20}$$

In particular, given a minimizer  $m^{*,\sigma}$  of  $V^\sigma$ , by (V.4.20) we have

$$\limsup_{n \rightarrow \infty} F(m^{*,\sigma_n}) \leq \limsup_{n \rightarrow \infty} V^\sigma(m^{*,\sigma_n}) \leq \limsup_{n \rightarrow +\infty} V^{\sigma_n}(m * f_n) \leq F(m), \quad \text{for all } m \in \mathcal{P}_2(\mathbb{R}^d).$$

■

*Proof of Theorem V.2.4.* Let  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$  be arbitrary. Let  $(X_i)_{i=1}^N$  be i.i.d. with law  $\mu$ . Let  $\mu_N = \frac{1}{N} \sum_{i=1}^N \delta_{X_i}$  and  $m_t^N = \mu + t(\mu_N - \mu)$ ,  $t \in [0, 1]$ . Further let  $(\tilde{X}_i)_{i=1}^N$  be consider i.i.d., independent of  $(X_i)_{i=1}^N$  with law  $\mu$ .

By the definition of linear functional derivatives, we have

$$\begin{aligned} \mathbb{E}[F(\mu_N)] - F(\mu) &= \mathbb{E} \left[ \int_0^1 \int_{\mathbb{R}^d} \frac{\delta F}{\delta m}(m_t^N, v) (\mu_N - \mu)(dv) dt \right] \\ &= \int_0^1 \frac{1}{N} \sum_{i=1}^N \left( \mathbb{E} \left[ \frac{\delta F}{\delta m}(m_t^N, X_1) \right] - \mathbb{E} \left[ \frac{\delta F}{\delta m}(m_t^N, \tilde{X}_1) \right] \right) dt \\ &= \int_0^1 \mathbb{E} \left[ \frac{\delta F}{\delta m}(m_t^N, X_1) - \frac{\delta F}{\delta m}(m_t^N, \tilde{X}_1) \right] dt. \end{aligned} \tag{V.4.21}$$



We introduce the (random) measures

$$\tilde{m}_t^N := m_t^N + \frac{t}{N}(\delta_{\tilde{X}_1} - \delta_{X_1}) \quad \text{and} \quad m_{t,t_1}^N := (\tilde{m}_t^N - m_t^N)t_1 + m_t^N, \quad t, t_1 \in [0, 1],$$

and notice that due to independence of  $(X_i)_{i=1}^N$  and  $(\tilde{X}_i)_{i=1}^N$  we have that

$$\mathbb{E} \left[ \frac{\delta F}{\delta m}(\tilde{m}_t^N, \tilde{X}_1) \right] = \mathbb{E} \left[ \frac{\delta F}{\delta m}(m_t^N, X_1) \right].$$

Therefore,

$$\begin{aligned} \mathbb{E}[F(\mu_N) - F(\mu)] &= \int_0^1 \mathbb{E} \left[ \frac{\delta F}{\delta m}(\tilde{m}_t^N, \tilde{X}_1) - \frac{\delta F}{\delta m}(m_t^N, X_1) \right] dt \\ &= \int_0^1 \mathbb{E} \left[ \int_0^1 \int_{\mathbb{R}^d} \frac{\delta^2 F}{\delta m^2}(m_{t,t_1}^N, \tilde{X}_1, y_1)(\tilde{m}_t^N - m_t^N)(dy_1) dt_1 \right] dt \\ &= \frac{1}{N} \mathbb{E} \left[ \int_0^1 \int_0^1 \int_{\mathbb{R}^d} t \frac{\delta^2 F}{\delta m^2}(m_{t,t_1}^N, \tilde{X}_1, y_1)(\delta_{\tilde{X}_1} - \delta_{X_1})(dy_1) dt_1 dt \right]. \end{aligned} \quad (\text{V.4.22})$$

To conclude, we observe that

$$\begin{aligned} &\mathbb{E} \left[ \left| \int_{\mathbb{R}^d} \frac{\delta^2 F}{\delta m^2}(m_{t,t_1}^N)(\tilde{X}_1, y_1)(\delta_{\tilde{X}_1} - \delta_{X_1})(dy_1) \right| \right] \\ &= \mathbb{E} \left[ \left| \int_{\mathbb{R}^d} \frac{\delta^2 F}{\delta m^2}(m_{t,t_1}^N)(\tilde{X}_1, y_1)\delta_{\tilde{X}_1}(dy_1) - \int_{\mathbb{R}^d} \frac{\delta^2 F}{\delta m^2}(m_{t,t_1}^N)(\tilde{X}_1, y_1)\delta_{X_1}(dy_1) \right| \right] \\ &\leq \mathbb{E} \left[ \sup_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} \left| \frac{\delta^2 F}{\delta m^2}(\nu)(\tilde{X}_1, \tilde{X}_1) \right| + \sup_{\nu \in \mathcal{P}_2(\mathbb{R}^d)} \left| \frac{\delta^2 F}{\delta m^2}(\nu)(\tilde{X}_1, X_1) \right| \right] \leq 2L, \end{aligned}$$

by (V.2.10). We have thus shown that for all  $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ , for all i.i.d.  $(X_i)_{i=1}^N$  with law  $\mu$  and with  $\mu_N = \frac{1}{N} \sum_{i=1}^N \delta_{X_i}$  it holds that

$$|\mathbb{E}[F(\mu_N)] - F(\mu)| \leq \frac{2L}{N}. \quad (\text{V.4.23})$$

From (V.4.23) with i.i.d  $(X_i^*)_{i=1}^N$  such that  $X_i^* \sim m^*$ ,  $i = 1, \dots, N$  we get that

$$\left| \mathbb{E} \left[ F \left( \frac{1}{N} \sum_{i=1}^N \delta_{X_i^*} \right) \right] - F(m^*) \right| \leq \frac{2L}{N}.$$

Let  $(X_i^*)_{i=1}^N$  be i.i.d. such that  $X_i^* \sim m^*$ ,  $i = 1, \dots, N$ . Note that

$$F(m^*) \leq \inf_{(x_i)_{i=1}^N \subset \mathbb{R}^d} F \left( \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \right) \leq \mathbb{E} \left[ F \left( \frac{1}{N} \sum_{i=1}^N \delta_{X_i^*} \right) \right].$$

From this and (V.4.23) we then obtain

$$0 \leq \inf_{(x_i)_{i=1}^N \subset \mathbb{R}^d} F\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i}\right) - F(m^*) \leq \frac{2L}{N}.$$

■

In the rest of the section, we shall discuss the first order condition for the minimizer of the function  $V^\sigma$ . We first show an elementary lemma for convex functions on  $\mathcal{P}(\mathbb{R}^d)$ .

**Lemma V.4.1.** *Under Assumption V.2.1, given  $m, m' \in \mathcal{P}(\mathbb{R}^d)$ , we have*

$$F(m') - F(m) \geq \int_{\mathbb{R}^d} \frac{\delta F}{\delta m}(m, x)(m' - m)(dx). \quad (\text{V.4.24})$$

*Démonstration.* Define  $m^\varepsilon := (1 - \varepsilon)m + \varepsilon m'$ . Since  $F$  is convex, we have

$$\varepsilon(F(m') - F(m)) \geq F(m^\varepsilon) - F(m) = \int_0^\varepsilon \int_{\mathbb{R}^d} \frac{\delta F}{\delta m}(m^s, x)(m' - m)(dx) ds$$

Since  $\frac{\delta F}{\delta m}$  is bounded and continuous, we obtain (V.4.24) by the dominant convergence theorem. ■

*Proof of Proposition V.2.5.*

*Step 1.* We first prove the existence of minimizer. Clearly there exists  $\bar{m} \in \mathcal{P}(\mathbb{R}^d)$  such that  $V^\sigma(\bar{m}) < +\infty$ . Denote

$$\mathcal{S} := \left\{ m : \frac{\sigma^2}{2} H(m) \leq V^\sigma(\bar{m}) - \inf_{m' \in \mathcal{P}(\mathbb{R}^d)} F(m') \right\}.$$

As a sublevel set of the relative entropy  $H$ ,  $\mathcal{S}$  is weakly compact, see e.g. [45, Lemma 1.4.3]. Together with the weak lower semi-continuity of  $V^\sigma$ , the minimum of  $V^\sigma$  on  $\mathcal{S}$  is attained. Notice that for all  $m \notin \mathcal{S}$ , we have  $V^\sigma(m) \geq V^\sigma(\bar{m})$ , so the minimum of  $V^\sigma$  on  $\mathcal{S}$  coincides with the global minimum. Further, since  $V^\sigma$  is strictly convex, the minimizer is unique. Moreover, given  $m^* = \operatorname{argmin}_{m \in \mathcal{P}(\mathbb{R}^d)} V^\sigma(m)$ , we know  $m^* \in \mathcal{S}$ , and thus we have  $H(m^*) < \infty$  as well as  $\mathbb{E}^{m^*}[U(X)] < \infty$ . Therefore,  $m^* \in \mathcal{P}_2(\mathbb{R}^d)$  is absolutely continuous with respect to the Gibbs measure, so also absolutely continuous with respect to the Lebesgue measure.

*Step 2. Sufficient condition :* Let  $m^* \in \mathcal{P}_2(\mathbb{R}^d)$  satisfy (V.2.11), in particular,  $m^*$  is equivalent to the Lebesgue measure. Let  $m \in \mathcal{P}(\mathbb{R}^d)$  such that  $m$  is absolutely continuous with respect to the Lebesgue measure (otherwise  $V^\sigma(m) = +\infty$ ). Let

$$f := \frac{dm}{dm^*}$$

be the Radon-Nikodym derivative. Let  $m^\varepsilon := (1 - \varepsilon)m^* + \varepsilon m = (1 + \varepsilon(f - 1))m^*$  for  $\varepsilon > 0$ . For the simplicity of the notations, denote  $m^\varepsilon(x)$  and  $m^*(x)$  the respective density function of  $m^\varepsilon$  and  $m^*$  with respect to Lebesgue measure. Recall the function  $h$  in (V.4.19) and note that  $h(y) \geq y - 1$  for all  $y \in \mathbb{R}^+$ . Using (V.4.24), we obtain

$$\frac{F(m^\varepsilon) - F(m^*)}{\varepsilon} \geq \frac{1}{\varepsilon} \int_{\mathbb{R}^d} \frac{\delta F}{\delta m}(m^*, \cdot)(m^\varepsilon - m^*) dx = \int_{\mathbb{R}^d} \frac{\delta F}{\delta m}(m^*, \cdot)(f - 1)m^* dx.$$

Moreover

$$\begin{aligned} \frac{\sigma^2}{2\varepsilon} \left( H(m^\varepsilon) - H(m^*) \right) &= \frac{\sigma^2}{2\varepsilon} \int_{\mathbb{R}^d} \left( m^\varepsilon \log \frac{m^\varepsilon}{g} - m^* \log \frac{m^*}{g} \right) dx \\ &= \frac{\sigma^2}{2\varepsilon} \int_{\mathbb{R}^d} (m^\varepsilon - m^*) \log \frac{m^*}{g} dx + \frac{\sigma^2}{2\varepsilon} \int_{\mathbb{R}^d} m^\varepsilon \left( \log \frac{m^\varepsilon}{g} - \log \frac{m^*}{g} \right) dx \\ &= \frac{\sigma^2}{2} \int_{\mathbb{R}^d} (f - 1)m^* \log \frac{m^*}{g} dx + \frac{\sigma^2}{2\varepsilon} \int_{\mathbb{R}^d} m^\varepsilon \log \frac{m^\varepsilon}{m^*} dx \\ &= \frac{\sigma^2}{2} \int_{\mathbb{R}^d} (f - 1)m^* (\log m^* + U) dx + \frac{\sigma^2}{2\varepsilon} \int_{\mathbb{R}^d} h(1 + \varepsilon(f - 1))m^* dx \\ &\geq \frac{\sigma^2}{2} \int_{\mathbb{R}^d} (f - 1)m^* (\log m^* + U) dx + \frac{\sigma^2}{2} \int_{\mathbb{R}^d} (f - 1)m^* dx \\ &= \frac{\sigma^2}{2} \int_{\mathbb{R}^d} (f - 1)m^* (\log m^* + U) dx \end{aligned}$$

since  $\int_{\mathbb{R}^d} (f - 1)m^* dx = \int_{\mathbb{R}^d} (m - m^*) dx = 0$ . Hence

$$\frac{V^\sigma(m^\varepsilon) - V^\sigma(m^*)}{\varepsilon} \geq \int_{\mathbb{R}^d} \left( \frac{\delta F}{\delta m}(m^*, \cdot) + \frac{\sigma^2}{2} \log m^* + \frac{\sigma^2}{2} U \right) (f - 1)m^* dx = 0.$$

**Step 3. Necessary condition :** Let  $m^*$  be the minimizer of  $V^\sigma$ . Let  $m$  a probability measure such that  $H(m) < \infty$ , in particular  $m$  is also absolutely continuous with respect to Lebesgue measure  $\ell$ . As above, denote  $m(x)$  and  $m^*(x)$  the respective density function of  $m$  and  $m^*$  with respect to Lebesgue measure and we have

$$\begin{aligned} &\frac{V^\sigma(m^\varepsilon) - V^\sigma(m^*)}{\varepsilon} \\ &= \frac{1}{\varepsilon} \int_0^\varepsilon \int_{\mathbb{R}^d} \frac{\delta F}{\delta m}(m^s, x)(m(x) - m^*(x)) dx ds \\ &\quad + \frac{\sigma^2}{2\varepsilon} \int_{\mathbb{R}^d} (h(m^\varepsilon(x)) - h(m^*(x)) - \log(g(x))(m^\varepsilon(x) - m^*(x))) dx. \end{aligned}$$

Since  $h$  is convex, we note that for all  $\varepsilon \in (0, 1)$

$$\frac{1}{\varepsilon} (h(m^\varepsilon(x)) - h(m^*(x)) - \log(g(x))(m^\varepsilon(x) - m^*(x))) \leq m(x) \log \left( \frac{m(x)}{g(x)} \right) - m^*(x) \log \left( \frac{m^*(x)}{g(x)} \right).$$

Since  $H(m)$  and  $H(m^*)$  are both finite, the right hand side of the above inequality is integrable. Therefore by Fatou's Lemma we obtain

$$0 \leq \lim_{\varepsilon \rightarrow 0} \frac{V^\sigma(m^\varepsilon) - V^\sigma(m^*)}{\varepsilon} \leq \int_{\mathbb{R}^d} \left( \frac{\delta F}{\delta m}(m^*, x) + \frac{\sigma^2}{2} \log(m^*(x)) + \frac{\sigma^2}{2} U(x) \right) (m(x) - m^*(x)) dx. \quad (\text{V.4.25})$$

Since  $m$  is arbitrary, we first obtain

$$\frac{\delta F}{\delta m}(m^*, \cdot) + \frac{\sigma^2}{2} \log(m^*) + \frac{\sigma^2}{2} U \text{ is a constant, } m^* - a.s.$$

Now suppose that  $m^*$  is not equivalent to Lebesgue measure. There exists a set  $\mathcal{K} \subset \mathbb{R}^d$  such that  $m^*(\mathcal{K}) = 0$  and  $\ell(\mathcal{K}) > 0$ . It follows from (V.4.25) that  $0 \leq C - \int_{\mathcal{K}} \infty dm$ . Since we may choose  $m$  having positive mass on  $\mathcal{K}$ , it is a contradiction. Therefore,  $m^*$  is equivalent to Lebesgue measure and we have

$$\frac{\delta F}{\delta m}(m^*, \cdot) + \frac{\sigma^2}{2} \log(m^*) + \frac{\sigma^2}{2} U \text{ is a constant, } \ell - a.s.$$

■

## V.5 Mean Field Langevin Equations

Recall that

$$b(x, m) := D_m F(m, x) + \frac{\sigma^2}{2} \nabla U(x).$$

Due to Assumption V.2.7 and V.2.2, the function  $b$  is of linear growth.

**Lemma V.5.1.** *Under Assumption V.2.2 and V.2.7, let  $X$  be the strong solution to (V.2.12). If  $m_0 \in \mathcal{P}_p(\mathbb{R}^d)$  for some  $p \geq 2$ , we have*

$$\mathbb{E} \left[ \sup_{t \leq T} |X_t|^p \right] \leq C, \quad \text{for some } C \text{ depending on } p, \sigma, T. \quad (\text{V.5.26})$$

If  $m_0 \in \mathcal{P}_p(\mathbb{R}^d)$  for some  $p \geq 2$ , we have

$$\sup_{t \in \mathbb{R}^+} \mathbb{E} \left[ |X_t|^p \right] \leq C, \quad \text{for some } C \text{ depending on } p, \sigma. \quad (\text{V.5.27})$$

In particular, if  $m_0 \in \cup_{p > 2} \mathcal{P}_p(\mathbb{R}^d)$ , then  $(m_t)_{t \in \mathbb{R}^+}$  belong to a  $\mathcal{W}_2$ -compact subset of  $\mathcal{P}_2(\mathbb{R}^d)$ .

*Démonstration.* Since  $b$  is of linear growth, we have

$$|X_t| \leq |X_0| + \int_0^t C(1 + |X_s|) ds + |\sigma W_t|.$$

Therefore,

$$\sup_{t \leq s} |X_t|^p \leq C \left( |X_0|^p + 1 + \int_0^s \sup_{t \leq r} |X_t|^p dr + \sup_{t \leq s} |\sigma W_t|^p \right).$$

Note that  $\mathbb{E} \left[ \sup_{t \leq s} |\sigma W_t|^p \right] \leq C s^{p/2}$ . Then (V.5.26) follows from the Gronwall inequality.

For the second estimate, we apply the Itô formula and obtain

$$d|X_t|^p = |X_t|^{p-2} \left( -pX_t \cdot b(X_t, m_t) + \frac{p(p-1)}{2} \sigma^2 \right) dt + p\sigma |X_t|^{p-2} X_t \cdot dW_t.$$

Since  $D_m F$  is bounded and  $\nabla U(x) \cdot x \geq C|x|^2 + C'$ , we have

$$\begin{aligned} d|X_t|^p &\leq |X_t|^{p-2} \left( C''|X_t| - \frac{p\sigma^2}{2} (C|X_t|^2 + C') + \frac{p(p-1)}{2} \sigma^2 \right) dt + p\sigma |X_t|^{p-2} X_t \cdot dW_t \\ &\leq |X_t|^{p-2} \left( C - \epsilon |X_t|^2 + \frac{p(p-1)}{2} \sigma^2 \right) dt + p\sigma |X_t|^{p-2} X_t \cdot dW_t, \text{ for some } 0 < \epsilon < \frac{p\sigma^2 C}{2}. \end{aligned}$$

The last inequality is due to the Young inequality. Again by the Itô formula we have

$$d(e^{\epsilon t} |X_t|^p) \leq e^{\epsilon t} \left( |X_t|^{p-2} \left( C + \frac{p(p-1)}{2} \sigma^2 \right) dt + p\sigma |X_t|^{p-2} X_t \cdot dW_t \right) \quad (\text{V.5.28})$$

Further, define the stopping time  $\tau_m := \inf\{t \geq 0 : |X_t| \geq m\}$ . By taking expectation on both sides of (V.5.28), we have

$$\mathbb{E}[e^{\epsilon(\tau_m \wedge t)} |X_{\tau_m \wedge t}|^p] \leq \mathbb{E}[|X_0|^p] + \mathbb{E} \left[ \int_0^{\tau_m \wedge t} e^{\epsilon s} |X_s|^{p-2} \left( C + \frac{p(p-1)}{2} \sigma^2 \right) ds \right]. \quad (\text{V.5.29})$$

In the case  $p = 2$ , it follows from the Fatou lemma and the monotone convergence theorem that

$$\mathbb{E}[|X_t|^2] \leq e^{-\epsilon t} \mathbb{E}[|X_0|^2] + \int_0^t e^{\epsilon(s-t)} (C + \sigma^2) ds \leq C (e^{-\epsilon t} + \epsilon^{-1}(1 - e^{-\epsilon t})),$$

and thus  $\sup_{t \in \mathbb{R}^+} \mathbb{E}[|X_t|^2] < \infty$ . For  $p > 2$ , we again obtain from (V.5.29) that

$$\mathbb{E}[|X_t|^p] \leq e^{-\epsilon t} \mathbb{E}[|X_0|^p] + \int_0^t e^{\epsilon(s-t)} \mathbb{E}[|X_s|^{p-2}] \left( C + \frac{p(p-1)}{2} \sigma^2 \right) ds.$$

Then (V.5.27) follows from induction. ■

**Proposition V.5.2.** *Let Assumption V.2.2 and V.2.7 hold true and assume  $m_0 \in \mathcal{P}_p(\mathbb{R}^d)$  for some  $p \geq 2$ . The marginal law  $m$  of the solution  $X$  to (V.2.12) is a weak solution to Fokker-Planck equation :*

$$\partial_t m = \nabla \cdot \left( b(x, m)m + \frac{\sigma^2}{2} \nabla m \right), \quad (\text{V.5.30})$$

in the sense that for all  $C^\infty$ -function  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$  such that  $\phi, \nabla\phi, \nabla^2\phi$  decay to 0 at infinity, we have

$$\int_{\mathbb{R}^d} \phi(x)(m_t - \tilde{m}_s)(dx) = \int_s^t \int_{\mathbb{R}^d} \left( -\nabla\phi(x)b(x, m_u)m_u(dx) + \frac{\sigma^2}{2}\Delta\phi(x)m_u(dx) \right) du.$$

Moreover, the mapping  $t \mapsto m_t$  is weakly continuous on  $[0, +\infty)$ , the joint density function  $(t, x) \mapsto m(t, x)$  exists and  $m \in C^{1,\infty}((0, \infty) \times \mathbb{R}^d, \mathbb{R})$ . In particular, a stationary solution to the Fokker-Planck equation (V.5.30) is an invariant measure to (V.2.12).

*Démonstration.* By applying the Itô formula on  $\phi(t, X_t)$ , we can verify that  $m$  is a weak solution to (V.5.30). Next, define  $\tilde{b}(x, t) := b(x, m_t)$ . Obviously,  $m$  can be regarded as a weak solution to the linear PDE :

$$\partial_t m = \nabla \cdot \left( \tilde{b}m + \frac{\sigma^2}{2}\nabla m \right). \quad (\text{V.5.31})$$

Then the regularity result follows from a standard argument through  $L^p_{loc}$ -estimate. For details, we refer the readers to the seminal paper [93, p.14-p.15] or the classic book [100, Chapter IV].

Let  $m^*$  be a stationary solution to (V.5.30), and  $X$  be the strong solution to the SDE :

$$dX_t = -b(X_t, m^*)dt + \sigma dW_t.$$

It is easy to verify that given  $\text{Law}(X_0) = m^*$  we have  $\text{Law}(X_t) = m^*$  for all  $t \geq 0$ . Therefore  $X$  is the solution to mean-field Langevin equation (V.2.12) and  $m^*$  is an invariant measure. ■

## V.6 Convergence to the Invariant Measure

Now we are going to show that under mild conditions, the flow of marginal law  $(m_t)_{t \in \mathbb{R}^+}$  converges toward the invariant measure which coincides with the minimizer of  $V^\sigma$ .

**Lemma V.6.1.** *Suppose Assumption V.2.2 and V.2.7 hold true and  $m_0 \in \mathcal{P}_2(\mathbb{R}^d)$ . Let  $m$  be the law of the solution to the mean field Langevin equation (V.2.12). Denote by  $\mathbb{P}_{\sigma, w}$  the scaled Wiener measure<sup>1</sup> with initial distribution  $m_0$ . Then,*

- i) For any  $T > 0$ ,  $\mathbb{P}_{\sigma, w}$  is equivalent to  $m$  on  $\mathcal{F}_T$ , where  $\{\mathcal{F}_t\}$  is the filtration generated by  $X$ , and the relative entropy

$$\mathbb{E}^m \left[ \log \left( \frac{dm}{d\mathbb{P}_{\sigma, w}} \Big|_{\mathcal{F}_T} \right) \right] < \infty. \quad (\text{V.6.32})$$

- ii) For all  $t > 0$ , the marginal law  $m_t$  admits density such that  $m_t > 0$  and  $H(m_t) < \infty$ .

1. Under the scaled Wiener measure  $\mathbb{P}_{\sigma, w}$ , if we denote  $X$  as the canonical process,  $\frac{X}{\sigma}$  is a standard Brownian motion.

*Démonstration.* *i)* We shall prove in the Appendix in Lemma V.7.1 that due to the linear growth in  $x$  of the drift  $b$ ,  $\mathbb{P}_{\sigma,w}$  is equivalent to  $m$ . Also by the linear growth of coefficient, we have

$$\mathbb{E}^m \left[ \log \left( \frac{dm}{d\mathbb{P}_{\sigma,w}} \Big|_{\mathcal{F}_T} \right) \right] = \mathbb{E}^m \left[ \frac{1}{\sigma^2} \int_0^T |b(X_t, m_t)|^2 dt \right] \leq C \mathbb{E}^m \left[ 1 + \sup_{t \leq T} |X_t|^2 \right] < \infty.$$

The last inequality is due to Lemma V.5.1.

*ii)* Since  $\mathbb{P}_{\sigma,w}$  is equivalent to  $m$ , we have  $m_t > 0$ . Denote  $f_{\sigma,t}$  the density function of the marginal law of a standard Brownian motion multiplied by  $\sigma$  with initial distribution  $m_0$ . It follows from the conditional Jensen inequality that for all  $t \in [0, T]$

$$\int_{\mathbb{R}^d} m_t \log \left( \frac{m_t(x)}{f_{\sigma,t}(x)} \right) dx \leq \mathbb{E}^m \left[ \log \left( \frac{dm}{d\mathbb{P}_{\sigma,w}} \Big|_{\mathcal{F}_T} \right) \right] < +\infty. \quad (\text{V.6.33})$$

Further, by the fact  $f_{\sigma,t}(x) \leq \frac{1}{(2\pi t)^{d/2} \sigma}$ , we have

$$\int_{\mathbb{R}^d} m_t(x) \log(f_{\sigma,t}(x)) dx \leq -\frac{d}{2} \log(2\pi t \sigma^2).$$

Finally, note that

$$-\int_{\mathbb{R}^d} m_t(x) \log(g(x)) dx = \int_{\mathbb{R}^d} m_t(x) U(x) dx \leq C \int_{\mathbb{R}^d} m_t(x) |x|^2 dx < \infty.$$

Together with (V.6.33), we have  $H(m_t) < \infty$ . ■

Next, we introduce an interesting result of [58, Theorem 3.10 and Remark 4.13].

**Lemma V.6.2.** *Let  $m$  be a measure equivalent to the scaled Wiener measure  $\mathbb{P}_{\sigma,w}$  such that the relative entropy is finite as in (V.6.32). Then,*

*i) for any  $0 < t < T$  we have  $\int_t^T \int_{\mathbb{R}^d} |\nabla \log(m_s)|^2 m_s dx ds < +\infty$ .*

*ii) given  $t \geq t_0 > 0$  such that the Doléans-Dade exponential  $\mathcal{E}^b(X) := e^{-\int_{t-t_0}^t \frac{b_s}{\sigma^2} dX_s - \int_{t-t_0}^t \frac{1}{2} |\frac{b_s}{\sigma}|^2 ds}$  <sup>2</sup> is conditionally  $\mathbb{L}^2$ -differentiable on the interval  $[t-t_0, t]$  <sup>3</sup>, we have*

$$\nabla \log(m_t(x)) = -\frac{1}{t_0} \mathbb{E} \left[ \int_0^{t_0} \left( 1 + s \nabla b(X_{t-t_0+s}, m_{t-t_0+s}) \right) dW_s^{t-t_0} \Big| X_t = x \right], \quad (\text{V.6.35})$$

2. Again, we slightly abuse the notation, using  $X$  to denote the canonical process of the Wiener space.

3. Denote by  $\mathbb{P}_{\sigma,w}^{t-t_0, x_0}$  the conditional probability of  $\mathbb{P}_{\sigma,w}$  given  $X_{t-t_0} = x_0$ .  $\mathcal{E}^b(X)$  is conditionally  $\mathbb{L}^2$ -differentiable on the interval  $[t-t_0, t]$ , if there exists an absolutely continuous process  $D\mathcal{E}^b := \int_{t-t_0}^t D\mathcal{E}_s^b ds$  with  $D\mathcal{E}_s^b \in \mathbb{L}^2(\mathbb{P}_{\sigma,w}^{t-t_0, x_0})$  for all  $x_0 \in \mathbb{R}^d$  such that for any  $h := \int_{t-t_0}^t \dot{h}_s ds$  with bounded predictable  $\dot{h}$ , we have

$$\lim_{\epsilon \rightarrow 0} \left| \frac{\mathcal{E}^b(X + \epsilon h) - \mathcal{E}^b(X)}{\epsilon} - \langle D\mathcal{E}^b(X), h \rangle \right| = 0, \quad \text{in } \mathbb{L}^2(\mathbb{P}_{\sigma,w}^{t-t_0, x_0}) \text{ for all } x_0 \in \mathbb{R}^d, \quad (\text{V.6.34})$$

where  $\langle D\mathcal{E}^b(X), h \rangle = \int_{t-t_0}^t \dot{h}_s D\mathcal{E}_s^b(X) ds$ .

where  $W_s^{t-t_0} := W_{t-t_0+s} - W_{t-t_0}$  and  $W$  is the Brownian motion in (V.2.12).

We shall prove in the Appendix, Lemma V.7.2, that under Assumption V.2.2 and V.2.7,  $\mathcal{E}^b$  is conditionally  $\mathbb{L}^2$ -differentiable on  $[t - t_0, t]$  for all  $t \geq t_0 > 0$ .

The estimate (i) leads to some other integrability results.

**Lemma V.6.3.** *Suppose Assumption V.2.2 and V.2.7 hold true and  $m_0 \in \mathcal{P}_2(\mathbb{R}^d)$ . We have*

$$\int_t^T \int_{\mathbb{R}^d} |\nabla m_t(x)| dx dt < \infty, \quad \int_t^T \int_{\mathbb{R}^d} |x \cdot \nabla m_t(x)| dx dt < \infty, \quad \text{and} \quad \int_t^T \int_{\mathbb{R}^d} |\Delta m_t(x)| dx dt < \infty.$$

*Démonstration.* By the Young inequality, we have

$$|\nabla m_t| \leq m_t + \left| \frac{\nabla m_t}{m_t} \right|^2 m_t \quad \text{and} \quad |x \cdot \nabla m_t| \leq x^2 m_t + \left| \frac{\nabla m_t}{m_t} \right|^2 m_t.$$

All terms on the right hand sides are integrable, due to Lemma V.6.2, therefore so are  $\nabla m$  and  $x \cdot \nabla m$ . Next, in order to prove the integrability of  $\Delta m$ , we apply Itô's formula :

$$d \log(m_t(X_t)) = \left( \frac{\partial_t m_t(X_t)}{m_t(X_t)} - \frac{\nabla m_t(X_t)}{m_t(X_t)} \cdot b_t(X_t, m_t) - \frac{\sigma^2}{2} \left| \frac{\nabla m_t}{m_t}(X_t) \right|^2 + \frac{\sigma^2}{2} \frac{\Delta m_t(X_t)}{m_t(X_t)} \right) dt + \sigma \frac{\nabla m_t(X_t)}{m_t(X_t)} \cdot dW_t.$$

Together with the Fokker-Planck equation (V.5.30), we have

$$\begin{aligned} & \log(m_t(X_t)) - \log(m_s(X_s)) \\ &= \int_s^t \left( \sigma^2 \frac{\Delta m_r}{m_r}(X_r) - \frac{\sigma^2}{2} \left| \frac{\nabla m_r}{m_r}(X_r) \right|^2 + \nabla \cdot b_r(X_r, m_r) \right) dr + \int_s^t \frac{\nabla m_r(X_r)}{m_r(X_r)} \sigma dW_r. \end{aligned} \quad (\text{V.6.36})$$

By Lemma V.6.2, we have

$$\mathbb{E} \left[ \int_s^t \frac{\nabla m_r(X_r)}{m_r(X_r)} \sigma dW_r \right] = 0.$$

Also recall that  $\nabla \cdot b$  is of linear growth. Taking expectation on both side of (V.6.36), we obtain

$$\mathbb{E} \left[ \int_s^t \sigma^2 \frac{|\Delta m_r|}{m_r}(X_r) dr \right] \leq H(m_t) + H(m_s) + \mathbb{E} \left[ \int_s^t \left( \frac{\sigma^2}{2} \left| \frac{\nabla m_r}{m_r}(X_r) \right|^2 + C(1 + |X_r|) \right) dr \right].$$

By Lemma V.6.1 and Lemma V.6.2, the right hand side is finite. ■

Based on the previous integrability results, the next lemma follows from the integration by part.

**Lemma V.6.4.** *Let  $m_0 \in \mathcal{P}_2(\mathbb{R}^d)$ . Under Assumption V.2.2 and V.2.7 we have for Leb-a.s.  $t$  that*

$$\begin{aligned} & \int_{\mathbb{R}^d} \text{Tr}(\nabla D_m F(m_t, x)) m_t dx = - \int_{\mathbb{R}^d} D_m F(m_t, x) \cdot \nabla m_t dx, \\ & \int_{\mathbb{R}^d} \nabla \cdot \frac{\nabla m_t}{m_t} m_t dx = - \int_{\mathbb{R}^d} \left| \frac{\nabla m_t}{m_t} \right|^2 m_t dx \quad \text{and} \quad \int_{\mathbb{R}^d} \Delta U(x) m_t(x) dx = - \int_{\mathbb{R}^d} \nabla U(x) \cdot \nabla m_t(x) dx. \end{aligned}$$



*Proof of Theorem V.2.9.* By the Itô-type formula given by [26, Theorem 4.14] and Lemma V.6.4, we have

$$\begin{aligned} dF(m_t) &= \int_{\mathbb{R}^d} \left( -|D_m F(m_t, x)|^2 - \frac{\sigma^2}{2} D_m F(m_t, x) \cdot \nabla U(x) + \frac{\sigma^2}{2} \text{Tr}(\nabla D_m F(m_t, x)) \right) m_t dx dt \\ &= \int_{\mathbb{R}^d} \left( -|D_m F(m_t, x)|^2 - \frac{\sigma^2}{2} D_m F(m_t, x) \cdot \left( \nabla U(x) + \frac{\nabla m_t}{m_t} \right) \right) m_t dx dt. \end{aligned} \quad (\text{V.6.37})$$

On the other hand, by (V.6.36), Itô's formula and Lemma V.6.4, we have

$$\begin{aligned} & d\left(\log(m_t(X_t)) + U(X_t)\right) - dM_t \\ &= \left( \sigma^2 \frac{\Delta m_t}{m_t}(X_t) - \frac{\sigma^2}{2} \left| \frac{\nabla m_t}{m_t}(X_t) \right|^2 + \nabla \cdot b_t(X_t, m_t) - \nabla U(X_t) \cdot b(X_t, m_t) + \frac{\sigma^2}{2} \Delta U(X_t) \right) dt, \\ &= \left( \sigma^2 \nabla \cdot \frac{\nabla m_t}{m_t}(X_t) + \frac{\sigma^2}{2} \left| \frac{\nabla m_t}{m_t}(X_t) \right|^2 + \nabla \cdot b_t(X_t, m_t) - \nabla U(X_t) \cdot b(X_t, m_t) + \frac{\sigma^2}{2} \Delta U(X_t) \right) dt. \end{aligned}$$

where  $dM_t = \left( \frac{\nabla m_t}{m_t}(X_t) + X_t \right) \cdot \sigma dW_t$  is a martingale on  $[s, T]$  for any  $0 < s < T$ . By taking expectation on both sides and using Lemma V.6.4, we obtain for  $t > 0$  :

$$\begin{aligned} & dH(m_t) \\ &= \mathbb{E} \left[ -\frac{\sigma^2}{2} \left| \frac{\nabla m_t}{m_t}(X_t) \right|^2 - b_t(X_t, m_t) \cdot \frac{\nabla m_t}{m_t}(X_t) - \nabla U(X_t) \cdot b(X_t, m_t) - \frac{\sigma^2}{2} \nabla U(X_t) \cdot \frac{\nabla m_t}{m_t}(X_t) \right] dt \\ &= \mathbb{E} \left[ -\frac{\sigma^2}{2} \left| \frac{\nabla m_t}{m_t}(X_t) + \nabla U(X_t) \right|^2 - D_m F(X_t, m_t) \cdot \left( \frac{\nabla m_t}{m_t}(X_t) + \nabla U(X_t) \right) \right] dt \\ &= \int_{\mathbb{R}^d} \left( -\frac{\sigma^2}{2} \left| \frac{\nabla m_t}{m_t} + \nabla U(x) \right|^2 - D_m F(m_t, x) \cdot \left( \frac{\nabla m_t}{m_t} + \nabla U(x) \right) \right) m_t(x) dx \end{aligned} \quad (\text{V.6.38})$$

Summing up the equation (V.6.37) and (V.6.38), we obtain (V.2.14). ■

In order to prove there exists an invariant measure of (V.2.12) equal to the minimizer of  $V^\sigma$ , we shall apply Lasalle's invariance principle. Now we simply recall it in our context. Let  $(m_t)_{t \in \mathbb{R}^+}$  be the flow of marginal laws of the solution to (V.2.12), given an initial law  $m_0$ . Define a dynamic system  $S(t)[m_0] := m_t$ . We shall consider the so-called  $w$ -limit set :

$$w(m_0) := \left\{ \mu \in \mathcal{P}_2(\mathbb{R}^d) : \text{there exist } t_n \rightarrow \infty \text{ such that } \mathcal{W}_2\left(S(t_n)[m_0], \mu\right) \rightarrow 0 \right\}$$

**Proposition V.6.5.** *[Invariance Principle] Let Assumption V.2.7 hold true and assume that  $m_0 \in \cup_{p>2} \mathcal{P}_p(\mathbb{R}^d)$ . Then the set  $w(m_0)$  is nonempty, compact and invariant, that is,*

- i) for any  $\mu \in w(m_0)$ , we have  $S(t)[\mu] \in w(m_0)$  for all  $t \in \mathbb{R}^+$ .
- ii) for any  $\mu \in w(m_0)$  and all  $t \in \mathbb{R}^+$ , there exists  $\mu' \in w(m_0)$  such that  $S(t)[\mu'] = \mu$ .

*Démonstration.* Under the upholding assumptions, it follows from Proposition V.2.8 that  $S(t)$  is continuous with respect to the  $\mathcal{W}_2$ -topology. By Lemma V.5.1, we have (V.5.27) with  $p > 2$ , and thus  $(S(t)[m_0])_{t \in \mathbb{R}^+} = (m_t)_{t \in \mathbb{R}^+}$  live in a  $\mathcal{W}_2$ -compact subset of  $\mathcal{P}_2(\mathbb{R}^d)$ . The desired result follows from the invariance principle, see e.g. [74, Theorem 4.3.3]. In order to keep the chapter self-contained, we state the proof as follows.

First, for any  $t \geq 0$ ,  $(m_s)_{s \geq t}$  is relatively compact, hence  $\overline{(m_s)_{s \geq t}}$  is compact. Since the arbitrary intersection of closed sets is closed, the set

$$w(m_0) = \bigcap_{t \geq 0} \overline{(m_s)_{s \geq t}}$$

is compact.

Next, let  $\mu \in w(m_0)$ , by definition we know that there exists a sequence  $(t_N)_{N > 0}$  such that  $S(t_N)[m_0] \rightarrow \mu$ . Let  $t \in \mathbb{R}^+$ , by the continuity of  $S(t) : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathcal{P}_2(\mathbb{R}^d)$ , we have  $S(t + t_N)[m_0] \rightarrow S(t)[\mu]$  and therefore  $S(t)[\mu] \in w(m_0)$ .

Finally, for the second point, let  $t \in \mathbb{R}^+$  and consider the sequence  $(S(t_N - t)[m_0])_N$ . Since  $(m_t)_{t \in \mathbb{R}^+}$  live in a  $\mathcal{W}_2$ -compact subset of  $\mathcal{P}_2(\mathbb{R}^d)$ , there exists a subsequence  $(t_{N'})$  and  $\mu' \in w(m_0)$  such that  $S(t_{N'} - t)[m_0] \rightarrow \mu'$ . Again, by the continuity of  $S(t)$ , we have  $S(t)[\mu'] = \lim_{N' \rightarrow \infty} S(t_{N'} - t + t)m_0 = \mu$ .

■

*Proof of Theorem V.2.11. Step 1.* We first prove that  $m^* \in w(m_0)$ . Since  $w(m_0)$  is compact, there exists  $\tilde{m} \in \operatorname{argmin}_{m \in w(m_0)} V^\sigma(m)$ . By Proposition V.6.5, for  $t > 0$  there exists a probability measure  $\mu \in w(m_0)$  such that  $S(t)[\mu] = \tilde{m}$ . By Theorem V.2.9, for any  $s > 0$  we have

$$V^\sigma(S(t + s)[\mu]) \leq V^\sigma(\tilde{m}).$$

Since  $w(m_0)$  is invariant,  $S(t + s)[\mu] \in w(m_0)$  and thus  $V^\sigma(S(t + s)[\mu]) = V^\sigma(\tilde{m})$ . Again by Theorem V.2.9, we obtain

$$0 = \frac{dV^\sigma(S(t)[\mu])}{dt} = - \int_{\mathbb{R}^d} \left| D_m F(\tilde{m}, x) + \frac{\sigma^2}{2} \frac{\nabla \tilde{m}}{\tilde{m}}(x) + \frac{\sigma^2}{2} \nabla U(x) \right|^2 \tilde{m}(x) dx.$$

Since  $\tilde{m} = S(t)[\mu]$  is equivalent to the Lebesgue measure (Proposition V.6.1), we have

$$D_m F(\tilde{m}, \cdot) + \frac{\sigma^2}{2} \frac{\nabla \tilde{m}}{\tilde{m}} + \frac{\sigma^2}{2} \nabla U = 0. \quad (\text{V.6.39})$$

The probability measure  $\tilde{m}$  is an invariant measure of (V.2.12), because it is a stationary solution to the Fokker-Planck equation (V.5.30). Meanwhile, by Proposition V.2.5 we have  $\tilde{m} = m^*$ . Therefore,  $m^* \in w(m_0)$ .

*Step 2.* Since  $m^* \in w(m_0)$ , there exists a subsequence, denoted by  $(m_{t_n})_{n \in \mathbb{N}}$ , converging to  $m^*$ . We are going to prove that  $V^\sigma(m^*) = \lim_{n \rightarrow \infty} V^\sigma(m_{t_n})$ . It is enough to prove  $\int_{\mathbb{R}^d} m^* \log(m^*) dx = \lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} m_{t_n} \log(m_{t_n}) dx$ . By the lower-semicontinuity of entropy, it is sufficient to prove that

$$\int_{\mathbb{R}^d} m^* \log(m^*) dx \geq \limsup_{n \rightarrow \infty} \int_{\mathbb{R}^d} m_{t_n} \log(m_{t_n}) dx \quad (\text{V.6.40})$$

By (V.6.39), we know that  $-\log m^*$  is semi-convex, so we may apply the HWI inequality in [128, Theorem 3] :

$$\int_{\mathbb{R}^d} m_{t_n} \left( \log(m_{t_n}) - \log(m^*) \right) dx \leq \mathcal{W}_2(m_{t_n}, m^*) \left( \sqrt{I_n} + C \mathcal{W}_2(m_{t_n}, m^*) \right), \quad (\text{V.6.41})$$

where  $I_n$  is the relative Fisher information defined as

$$\begin{aligned} I_n &:= \mathbb{E} \left[ \left| \nabla \log \left( m_{t_n}(X_{t_n}) \right) - \nabla \log \left( m^*(X_{t_n}) \right) \right|^2 \right] \\ &= \mathbb{E} \left[ \left| \nabla \log \left( m_{t_n}(X_{t_n}) \right) + \frac{2}{\sigma^2} D_m F(m^*, X_{t_n}) + \nabla U(X_{t_n}) \right|^2 \right]. \end{aligned} \quad (\text{V.6.42})$$

We are going to show that  $\sup_n I_n < \infty$ . First, since  $D_m F$  is bounded and  $\nabla U$  is of linear growth, by Lemma V.5.1 we have

$$\sup_n \mathbb{E} \left[ \left| \frac{2}{\sigma^2} D_m F(m^*, X_{t_n}) + \nabla U(X_{t_n}) \right|^2 \right] < \infty. \quad (\text{V.6.43})$$

Next, since  $\nabla b$  is bounded, by Lemma V.7.2 and (V.6.35) we have for all  $n$

$$\begin{aligned} \mathbb{E} \left[ \left| \nabla \log \left( m_{t_n}(X_{t_n}) \right) \right|^2 \right] &\leq \inf_{0 < s \leq t_n} \frac{1}{s^2} \int_0^s C(1+r^2) dr \\ &= \inf_{0 < s \leq t_n} C \left( \frac{1}{s} + \frac{s}{3} \right) \leq \frac{2C}{\sqrt{3}}, \quad \text{for } t_n > \sqrt{3}, \end{aligned} \quad (\text{V.6.44})$$

where the constant  $C$  does not depend on  $n$ . Combining (V.6.42), (V.6.43) and (V.6.44) we obtain  $\sup_n I_n < \infty$ . Now the HWI inequality (V.6.41) reads

$$\int_{\mathbb{R}^d} m_{t_n} \left( \log(m_{t_n}) - \log(m^*) \right) dx \leq C \mathcal{W}_2(m_{t_n}, m^*) \left( 1 + \mathcal{W}_2(m_{t_n}, m^*) \right).$$

By letting  $n \rightarrow \infty$ , since  $\mathcal{W}_2(m_{t_n}, m^*) \rightarrow 0$ , we obtain (V.6.40).

*Step 3.* Finally we prove the convergence of the whole sequence  $(m_t)_{t \in \mathbb{R}^+}$  towards  $m^*$ , by showing that the set  $w(m_0)$  is a singleton, namely  $w(m_0) = \{m^*\}$ . Since  $V^\sigma(m_t)$  is non-increasing in  $t$ , there is a constant  $c := \lim_{t \rightarrow \infty} V^\sigma(m_t)$ . Recall that in *Step 2* we proved  $V^\sigma(m^*) = \lim_{n \rightarrow \infty} V^\sigma(m_{t_n})$ , so we obtain  $c = V^\sigma(m^*)$ . On the other hand, for any  $\mu \in w(m_0)$  there is a subsequence  $(m_{t'_n})_{n \in \mathbb{N}}$  converging to  $\mu$

and by the weak lower-semicontinuity of  $V^\sigma$  we have  $V^\sigma(\mu) \leq \liminf_{n \rightarrow \infty} V^\sigma(m_{t'_n}) = c$ . Using the fact that  $m^* = \operatorname{argmin}_{m \in w(m_0)} V^\sigma(m)$ , we have

$$V^\sigma(\mu) = V^\sigma(m^*) = c, \quad \text{for all } \mu \in w(m_0).$$

Finally by the uniqueness of the minimiser of  $V^\sigma$ , we have  $w(m_0) = \{m^*\}$ . ■

## V.7 Appendix

The following result regarding to the change of measure in the Wiener space is classic, see e.g. [9]. For readers' convenience, we provide a transparent proof as follow. Our argument is largely inspired by the one in [10, Lemma 4.1.1].

**Lemma V.7.1.** *Let a function  $(t, x) \mapsto b(t, x)$  be Lipschitz continuous and of linear growth in  $x$ , and a process  $X$  be the strong solution to the SDE :*

$$dX_t = b(t, X_t)dt + \sigma dW_t.$$

Define the following Doléan-Dade exponential for all  $t \in \mathbb{R}^+$

$$\rho_t := \exp\left(\frac{1}{\sigma} \int_0^t b(s, X_s) dW_s - \frac{1}{2\sigma^2} \int_0^t |b(s, X_s)|^2 ds\right). \quad (\text{V.7.45})$$

Then we have  $\mathbb{E}[\rho_t] = 1$  and thus  $\rho$  is a martingale on any finite horizon.

*Démonstration.* First, we shall prove that there exists  $C > 0$  such that for all  $t \in \mathbb{R}^+$ , we have

$$\mathbb{E}[\rho_t |X_t|^2] < C. \quad (\text{V.7.46})$$

By Itô's formula, we have

$$d|X_t|^2 = (2X_t b(t, X_t) + \sigma^2)dt + 2X_t \sigma dW_t,$$

and

$$d(\rho_t |X_t|^2) = \rho_t \left(4X_t b(t, X_t) + \sigma^2\right) dt + \rho_t \left(\frac{1}{\sigma} |X_t|^2 b(t, X_t) + 2X_t \sigma\right) dW_t,$$

and further

$$\begin{aligned} d \frac{\rho_t |X_t|^2}{1 + \varepsilon \rho_t |X_t|^2} &= \frac{\rho_t}{(1 + \varepsilon \rho_t |X_t|^2)^2} \left( \frac{1}{\sigma} |X_t|^2 b(t, X_t) + 2X_t \sigma \right) dW_t \\ &+ \frac{\rho_t}{(1 + \varepsilon \rho_t |X_t|^2)^2} \left( 4X_t b(t, X_t) + \sigma^2 \right) dt \\ &- \frac{\varepsilon \rho_t^2}{(1 + \varepsilon \rho_t |X_t|^2)^3} \left| \frac{1}{\sigma} |X_t|^2 b(t, X_t) + 2X_t \sigma \right|^2 dt. \end{aligned}$$

Note that the integrand of the stochastic integral on the right hand side above is bounded, so the stochastic integral is actually a real martingale. Therefore, by taking the expectation on both sides and using the fact that  $b$  has linear growth in  $x$ , we get

$$\begin{aligned} \frac{d}{dt} \mathbb{E} \left[ \frac{\rho_t |X_t|^2}{1 + \varepsilon \rho_t |X_t|^2} \right] &\leq \mathbb{E} \left[ \frac{\rho_t}{(1 + \varepsilon \rho_t |X_t|^2)^2} \left( 4X_t b(t, X_t) + \sigma^2 \right) \right] \\ &\leq K \mathbb{E} \left[ \frac{\rho_t |X_t|^2}{1 + \varepsilon \rho_t |X_t|^2} + 1 \right]. \end{aligned}$$

By Grönwall inequality, we get

$$\mathbb{E} \left[ \frac{\rho_t |X_t|^2}{1 + \varepsilon \rho_t |X_t|^2} \right] \leq C,$$

for some constant  $C$  which does not depend on  $\varepsilon$ . By Fatou's lemma, we get (V.7.46).

Next, by Itô's formula, we have

$$d \frac{\rho_t}{1 + \varepsilon \rho_t} = \frac{\rho_t b(t, X_t)}{(1 + \varepsilon \rho_t)^2} dW_t - \frac{\varepsilon \rho_t^2 b(t, X_t)^2}{(1 + \varepsilon \rho_t)^3} dt.$$

By (V.7.46), the stochastic integral above is a martingale, so taking the expectation on both sides, we get

$$\mathbb{E} \left[ \frac{\rho_t}{1 + \varepsilon \rho_t} \right] = \frac{1}{1 + \varepsilon} - \int_0^t \mathbb{E} \left[ \frac{\varepsilon \rho_s^2 b(s, X_s)^2}{(1 + \varepsilon \rho_s)^3} \right] ds.$$

Due to the linear growth of  $b$ , the term inside the expectation on the right hand side is bounded by  $C \rho_s (|X_s|^2 + 1)$  for some constant  $C > 0$  independent of  $\varepsilon$ . By the dominated convergence theorem, we get

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[ \frac{\rho_t}{1 + \varepsilon \rho_t} \right] = 1.$$

To conclude, one only needs to note that  $\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[ \frac{\rho_t}{1 + \varepsilon \rho_t} \right] = \mathbb{E}[\rho_t]$ . ■

**Lemma V.7.2.** *Under Assumption V.2.2 and V.2.7, the exponential martingale  $\mathcal{E}(b)$  is conditionally  $\mathbb{L}^2$ -differentiable on  $[t - t_0, t]$ , i.e. the equation (V.6.34) holds true, for all  $t \geq t_0 > 0$ .*

*Démonstration.* Without loss of generality, we may assume  $t = t_0$ . Under the upholding assumptions,

the process  $(b_t)_{t \in [0, t_0]}$  is  $\mathbb{L}^2$ -differentiable. By [126, Lemma 1.3.4], we know that  $\zeta(X) := -\int_0^{t_0} \frac{b_s}{\sigma^2} dX_s - \int_0^{t_0} \frac{1}{2} \left| \frac{b_s}{\sigma} \right|^2 ds$  is  $\mathbb{L}^2$ -differentiable for any  $t_0 > 0$ , namely there exists  $D\zeta$  such that

$$\frac{\zeta(X + \epsilon h) - \zeta(X)}{\epsilon} - \langle D\zeta(X), h \rangle \rightarrow 0, \quad \text{in } \mathbb{L}^2(\mathbb{P}_{\sigma, w}^{0, x}) \text{ for all } x \in \mathbb{R}^d, \text{ as } \epsilon \rightarrow 0.$$

By Proposition 1.3.8 and Proposition 1.3.11 from [126], we may compute  $D\zeta$  explicitly :

$$D\zeta(X) = -\int_0^{t_0} \left( \frac{b_s}{\sigma^2} + \int_s^{t_0} \frac{\nabla b_r}{\sigma^2} (dX_r + b_r dr) \right) ds. \quad (\text{V.7.47})$$

Note that  $\mathcal{E}^b = e^\zeta$ . Therefore, we have

$$\mathcal{E}^b(X + \epsilon h) - \mathcal{E}^b(X) = \int_0^\epsilon \langle \mathcal{E}^b(X + sh) D\zeta(X + sh), h \rangle ds, \quad \mathbb{P}_{\sigma, w}^{0, x}\text{-a.s. for all } x \in \mathbb{R}^d.$$

In order to prove (V.6.34), it is sufficient to prove that for all  $x \in \mathbb{R}^d$

$$\sup_{s \leq 1} \mathbb{E}^{\mathbb{P}_{\sigma, w}^{0, x}} \left[ \left| \langle \mathcal{E}^b(X + sh) D\zeta(X + sh), h \rangle \right|^p \right] < \infty, \quad \text{for some } p > 2.$$

By the form (V.7.47), we have  $\langle D\zeta(X + sh), h \rangle \in \cap_{q > 1} \mathbb{L}^q(\mathbb{P}_{\sigma, w}^{0, x})$ , so it is enough to show

$$\mathbb{E}^{\mathbb{P}_{\sigma, w}^{0, x}} \left[ \left| \mathcal{E}^b(X) \right|^p \right] < \infty, \quad \text{for some } p > 2. \quad (\text{V.7.48})$$

Further, note that

$$\left| \mathcal{E}^b(X) \right|^p = e^{-p \int_0^{t_0} \left( \sigma^{-2} D_m F(m_s, X_s) + \nabla U(X_s) \right) dX_s - \frac{p}{2} \int_0^{t_0} \frac{|b_s|^2}{\sigma^2} ds}.$$

Since  $D_m F$  is bounded, in order to prove (V.7.48), it is enough to show that

$$\mathbb{E}^{\mathbb{P}_{\sigma, w}^{0, x}} \left[ e^{-p \int_0^{t_0} \nabla U(X_s) dX_s} \right] < \infty, \quad \text{for some } p > 2.$$

By Itô formula, we obtain

$$\mathbb{E}^{\mathbb{P}_{\sigma, w}^{0, x}} \left[ e^{-p \int_0^{t_0} \nabla U(X_s) dX_s} \right] = \mathbb{E}^{\mathbb{P}_{\sigma, w}^{0, x}} \left[ e^{-p \left( U(X_{t_0}) - U(x) - \int_0^{t_0} \frac{\sigma^2}{2} \Delta U(X_s) ds \right)} \right] < \infty,$$

where we use the fact that  $U \geq -C$  for some  $C > 0$  and  $\Delta U$  is bounded. ■

---

---

# CHAPITRE VI

---

## MEAN-FIELD LANGEVIN SYSTEM, OPTIMAL CONTROL AND DEEP NEURAL NETWORKS

**Abstract :** In this chapter, we study a regularised relaxed optimal control problem and, in particular, we are concerned with the case where the control variable is of large dimension. We introduce a system of mean-field Langevin equations, the invariant measure of which is shown to be the optimal control of the initial problem under mild conditions. Therefore, this system of processes can be viewed as a continuous-time numerical algorithm for computing the optimal control. As an application, this result endorses the solvability of the stochastic gradient descent algorithm for a wide class of deep neural networks.

**Keywords :** Mean-Field Langevin Dynamics, Gradient Flow, Neural Networks

## VI.1 Introduction

This chapter revisits the classical optimal control problem, that is,

$$\inf_{\alpha} V^0(\alpha), \text{ where } V^0(\alpha) := \int_0^T L(t, X_t^\alpha, \alpha_t) dt + G(X_T^\alpha) \text{ and } \frac{dX_t^\alpha}{dt} = \phi(t, X_t^\alpha, \alpha_t). \quad (\text{VI.1.1})$$

In particular, we aim at providing a feasible algorithm for solving such problem (indeed, its regularized version) when the dimensions of the state  $X$  and of the control  $\alpha$  are both large.

It has been more than half a century since the discovery of Pontryagin's maximum principle [16], which states that in order to be an optimal control to the problem (VI.1.1),  $\alpha^*$  needs to satisfy the forward-backward ODE system :

$$\begin{cases} \alpha_t^* = \operatorname{argmin}_a H(t, X_t^*, a, P_t^*), & \text{where } H(t, x, a, p) := L(t, x, a) + p \cdot \phi(t, x, a), \\ X_t^* = x_0 + \int_0^t \phi(r, X_r^*, \alpha_r^*) dr, \\ P_t^* = \nabla_x G(X_T^*) + \int_0^t \nabla_x H(r, X_r^*, \alpha_r^*, P_r^*) dr. \end{cases} \quad (\text{VI.1.2})$$

It is worth mentioning that this necessary condition becomes sufficient if one imposes convexity condition on the coefficients. To solve the forward-backward system, the most naive way is to follow a fixed-point algorithm, that is, starting with an arbitrary control  $\alpha$ , evaluate the forward equation and then the backward one, and eventually compute a new control  $\tilde{\alpha}$  by solving the optimization problem on the top line. Under some mild conditions, one may show that this mapping  $\alpha \mapsto \tilde{\alpha}$  is a contraction at least on short horizon (i.e. for small  $T$ ), see e.g. [110] for a discussion on a more general setting where the dynamics of  $X$  and  $P$  are allowed to be SDE. However, this algorithm has a major drawback, that is, the optimization on the top line is hard to solve in high dimension (unless in some special cases when the optimizers have analytic forms). That is why after the discovery of Pontryagin's maximum principle, people have studied and widely applied the related gradient descent algorithm, see e.g. [21, 121]. In such iterative algorithm at each step we update the value of the control variable along the direction opposite to that of  $\nabla_a H$ , that is,

$$\alpha^{i+1} := \alpha^i - \eta \nabla_a H(t, X_t^i, a, P_t^i),$$

where  $\eta$  is the learning rate and  $X^i, P^i$  are the forward and backward processes evaluated with  $\alpha^i$ . In order to look into the convergence of such iteration, let us consider the continuous version of this gradient descent algorithm, governed by the following system of ODEs on the infinite horizon :

$$\begin{cases} \frac{d\alpha_t^s}{ds} = -\nabla_a H(t, X_t^s, \alpha_t^s, P_t^s) & \text{on } \{s \geq 0\} \text{ for all } t \in [0, T], \\ X_t^s = x_0 + \int_0^t \phi(r, X_r^s, \alpha_r^s) dr, \\ P_t^s = \nabla_x G(X_T^s) + \int_0^t \nabla_x H(r, X_r^s, \alpha_r^s, P_r^s) dr. \end{cases} \quad (\text{VI.1.3})$$



Curiously, after some careful calculus, one may verify that

$$\frac{dV^0(\alpha^s)}{ds} = - \int_0^T \left| \nabla_a H(r, X_r^s, \alpha_r^s, P_r^s) \right|^2 dr.$$

Therefore  $V^0$  is a natural Lyapunov function for the process  $(\alpha^s)$ , and in order for the equality  $\frac{dV^0(\alpha)}{ds} = 0$  to be true, the control  $\alpha$  must satisfy the forward-backward system (VI.1.2). This analysis (though not completely rigorous) reflects why this algorithm would converge. However, like other gradient-descent type algorithms, it would converge to a local minimizer, since Pontryagin's maximum principle is only a necessary first-order condition. One may attempt to put a convexity condition on the coefficients in order to ensure the local minimizer to be the global one. However, this usually urges  $X$  to be linear in  $\alpha$  (so the function  $\phi$  needs to be linear in  $(x, a)$ ), which largely limits the application of this method.

In order to go beyond the convex case for the optimal control problem, it is natural to recall how the Langevin equation helps to approximate the solution of the non-convex optimization on the real space. Given a function  $F$  not necessarily convex, we know that under some mild conditions the unique invariant measure of the following Langevin equation

$$d\Theta_s = -\dot{F}(\Theta_s)ds + \sigma dW_s \tag{VI.1.4}$$

is the global minimizer of the regularized optimization :

$$\min_{\nu \in \mathcal{P}} \int_{\mathbb{R}^m} F(a)\nu(da) + \frac{\sigma^2}{2} \text{Ent}(\nu), \tag{VI.1.5}$$

where  $W$  is the Brownian motion,  $\mathcal{P}$  is the space of probability measures and the regularizer  $\text{Ent}$  is the relative entropy with respect to the Lebesgue measure, see e.g. [93]. Moreover, the marginal law of the process (VI.1.4) converges to its invariant measure. As analyzed in the recent paper [81], this result is basically due to the fact that the function  $\nu \mapsto \int F(a)\nu(da)$  is convex (indeed linear). In the present work we wish to apply a similar regularization to the optimal control problem. In order to do that we first recall the relaxed formulation of the control problem (VI.1.1). Instead of controlling the process  $\alpha$ , we will control the flow of laws  $(\nu_t)_{t \in [0, T]}$ . Then the controlled process reads

$$X_t = x_0 + \int_0^t \int_{\mathbb{R}^m} \phi(r, X_r, a)\nu_r(da)dr,$$

and we aim at minimizing

$$\inf_{\nu} V(\nu), \quad \text{where} \quad V(\nu) := \int_0^T \int L(t, X_t, a)\nu_t(da)dt + G(X_T).$$

Comparing it to the original control problem (VI.1.1), we obtain that  $\inf_{\nu} V(\nu) \leq \inf_{\alpha} V^0(\alpha)$ . Indeed, due to the classical results in [57, 51], under some mild conditions the values of the minimums of the

two formulations remain the same. Further we add the relative entropy as a regularizer, as in (VI.1.5), and focus on the regularized optimization :

$$\inf_{\nu} V^{\sigma}(\nu), \quad \text{where} \quad V^{\sigma}(\nu) := V(\nu) + \frac{\sigma^2}{2} \int_0^T \text{Ent}(\nu_t) dt. \quad (\text{VI.1.6})$$

We will show that the global minimizer of this regularized control problem is again characterized by the invariant measure of Langevin-type dynamics, however, not a single Langevin equation as in (VI.1.4), but a system of mean-field Langevin equations in the spirit of (VI.1.3), that is,

$$\begin{cases} d\Theta_t^s = -\nabla_a H(t, X_t^s, \Theta_t^s, P_t^s) ds + \sigma dW_s, & \text{for } s \in \mathbb{R}^+, \quad \text{for } t \in [0, T], \quad \text{where} \\ X_t^s = X_0 + \int_0^t \int_{\mathbb{R}^m} \phi(r, X_r^s, a) \nu_r^s(da) dr, & \text{with } \nu_r^s := \text{Law}(\Theta_r^s), \\ P_t^s = \nabla_x G(X_T^s) + \int_t^T \int_{\mathbb{R}^m} \nabla_x H(r, X_r^s, a, P_r^s) \nu_r^s(da) dr, \end{cases} \quad (\text{VI.1.7})$$

The name ‘mean-field’ reflects the fact that the different equations in the system are coupled through (and only through) the marginal laws  $(\nu_t^s)_{t \in [0, T], s \in \mathbb{R}^+}$ . Moreover, we shall show that this characterization holds true not only when  $V$  is convex in  $\nu$  (which is still a quite restrictive case), but also under a set of milder conditions on the coefficients. Also, we prove in both cases that the marginal laws of the system (VI.1.7) converge to its unique invariant measure. In particular, in the latter case we may quantitatively compute the convergence rate.

One concrete motivation of this work is to introduce a new model of deep neural networks, and to provide a methodology of analyzing its efficiency. Our work can be viewed as a natural extension to the recent works [117, 116, 81, 138, 34] in which the authors studied the solvability of the two-layer (i.e. with one hidden layer) neural networks through the gradient flow of probability measures. It was proposed in the recent papers [33, 47, 107] among others, as well as in the course of P.-L. Lions in Collège de France (indeed similar ideas can be dated back to [104, 132], see also a very recent review on this topic [106]), that one may use the continuous-time controlled process as a model to study deep neural networks. One may expect the mathematical model to help understand the following three important properties the deep networks :

- *expressiveness* : why a class of functions can be represented in the form of a given neural network ?
- *trainability* : why the (stochastic) gradient descent algorithm can efficiently optimize the weights ?
- *small overfitting* : how the network representation can be stable with respect to the observed data ?

Using the model of continuous-time controlled process, in some recent papers such as [35, 105], the authors exploited the richness of the vector fields driving the controlled processes to show the expressiveness of the deep networks. The introduction of a new model with relaxed controlled processes is an attempt to address the *trainability* and the *small overfitting* of the deep networks. On the one hand, according to Theorem VI.3.12, under some conditions the mean-field Langevin system (VI.1.7) can be used to find the global minimum of the regularized relaxed control problem (VI.1.6). It can be viewed

as a first step to understand the *trainability* of the deep networks in the relaxed control model. Note that the Langevin dynamics can be viewed as a compromise between the exploitation (the gradient drift) and the exploration (the diffusion), so heuristically can avoid the local minima. This viewpoint is shared in the recent paper [154] which uses the relaxed optimal control to study the reinforcement learning. On the other hand, the entropy regularization is a natural choice to reduce overfitting. In particular, using the new model we provide an insight (see Section VI.4) to the widely applied *dropout* technique to understand its capability of reducing the overfitting.

Finally we would like to cite some other independent research projects sharing a related perspective on the analysis of the training of deep networks, see [88, 108, 12].

The rest of the chapter is organised in the following way. In Section VI.2, we define the relaxed optimal control problem under study and the corresponding system of mean-field Langevin equations. In Section VI.3 we announce the main results of the work, namely, the wellposedness of the system of mean-field Langevin equations and the convergence of the marginal laws of the system towards the optimal control, both in the convex case and in a contraction case. Before giving detailed proofs for the theoretical results, we introduce the application to deep neural networks in Section VI.4. Then in Sections VI.5, VI.6 and VI.7 we provide the proofs of the main results.

## VI.2 Preliminaries

### VI.2.1 Regularized Relaxed Optimal Control

In this chapter we aim to solve the optimal control problems in large dimension (in particular, the control variable is of large dimension). We shall allow the player to apply a mixed strategy, namely, a probability measure  $\nu$  of which the marginal law on the time dimension is the Lebesgue measure, i.e.

$$\nu \in \mathcal{V} := \left\{ \nu \in \mathcal{M}([0, T] \times \mathbb{R}^m) : \nu(dt, da) = \nu_t(da)dt, \text{ for some } \nu_t \in \mathcal{P}(\mathbb{R}^m) \right\},$$

where we denote by  $\mathcal{M}$  the space of measures. The controlled process  $X$  reads :

$$X_t = X_0 + \int_0^t \int_{\mathbb{R}^m} \phi(r, X_r, a, Z) \nu_r(da) dr, \quad \text{for } t \in [0, T],$$

where  $Z$  is an exogenous random variable taking values in a set  $\mathcal{Z}$ . In particular, in the application to the neural networks,  $Z$  would represent the input data. Denote  $\mathcal{G} := \sigma(Z)$  and assume that  $X_0$  is a bounded  $\mathcal{G}$ -measurable random variable. We use the notation  $\mathbb{E}$  as the expectation of random variables on  $\mathcal{G}$ . The relaxed control problem writes :

$$\inf_{\nu} V(\nu), \quad \text{where } V(\nu) := \mathbb{E} \left[ \int_0^T \int_{\mathbb{R}^m} L(t, X_t, a, Z) \nu_t(da) dt + G(X_T, Z) \right]. \quad (\text{VI.2.8})$$

Further in this chapter, instead of addressing the optimal control problem itself, we introduce the following regularized version :

$$\inf_{\nu} V^{\sigma}(\nu), \quad \text{where} \quad V^{\sigma}(\nu) := V(\nu) + \frac{\sigma^2}{2} \int_0^T \text{Ent}(\nu_t) dt, \quad (\text{VI.2.9})$$

where  $\text{Ent}$  is the relative entropy with respect to the Lebesgue measure on  $\mathbb{R}^m$ . It is noteworthy that  $\int_0^T \text{Ent}(\nu_t) dt$  is equal to the relative entropy of  $\nu$  with respect to the Lebesgue measure on  $[0, T] \times \mathbb{R}^m$ .

## VI.2.2 System of Mean-Field Langevin Equations

The following remark establishes a link between the control problem (VI.2.9) and the mean-field Langevin equation.

**Remark VI.2.1.** *Let us consider a simple example of a control problem with the following coefficients :  $X_0 \equiv 0$ ,  $L(a) = \lambda|a|^2$  and  $\phi(t, x, a) = \hat{\phi}(a)$ , that is, we aim to minimize*

$$\inf_{\nu} \mathbb{E} \left[ G \left( \int_0^T \int_{\mathbb{R}^m} \hat{\phi}(a, Z) \nu_t(da) dt \right) \right] + \int_0^T \int_{\mathbb{R}^m} \lambda|a|^2 \nu_t(da) dt + \frac{\sigma^2}{2} \int_0^T \text{Ent}(\nu_t) dt.$$

*Clearly,  $(\nu_t)_{t \in [0, T]}$  are exchangeable, so the optimal control  $\nu^*$  must satisfy  $\nu_0^* = \nu_t^*$  for any  $t \in [0, T]$ . Therefore it is equivalent to minimize*

$$\inf_{\nu_0} \mathbb{E} \left[ G \left( T \int_{\mathbb{R}^m} \hat{\phi}(a, Z) \nu_0(da) \right) \right] + T \int_{\mathbb{R}^m} \lambda|a|^2 \nu_0(da) + \frac{\sigma^2 T}{2} \text{Ent}(\nu_0).$$

*Given a convex function  $G$ , this minimization problem is studied in the recent paper [81], where the authors prove that the marginal laws of the corresponding mean-field Langevin equation converge to the global minimizer. In the present chapter we are going to generalize this result.*

For the general control problem (VI.2.9), we assume that all the coefficients are smooth enough. Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space and  $W$  an  $m$ -dimensional Brownian motion on it. Introduce the following system of mean-field Langevin equations :

$$\left\{ \begin{array}{l} d\Theta_t^s = -\mathbb{E} \left[ \nabla_a H(t, X_t^s, \Theta_t^s, P_t^s, Z) \right] ds + \sigma dW_s, \quad \text{for } s \in \mathbb{R}^+, \quad \text{for } t \in [0, T], \\ \text{where } \begin{array}{l} X_t^s = X_0 + \int_0^t \int_{\mathbb{R}^m} \phi(r, X_r^s, a, Z) \nu_r^s(da) dr, \quad \text{with } \nu_r^s := \text{Law}(\Theta_r^s), \\ P_t^s = \nabla_x G(X_T^s, Z) + \int_t^T \int_{\mathbb{R}^m} \nabla_x H(r, X_r^s, a, P_r^s, Z) \nu_r^s(da) dr, \end{array} \end{array} \right. \quad (\text{VI.2.10})$$

and  $H$  is the Hamiltonian function :

$$H(t, x, a, p, z) := L(t, x, a, z) + p \cdot \phi(t, x, a, z) \quad \text{for } (t, x, a, p) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^d.$$

For readers who are familiar with variational calculus of optimal control (Pontryagin's maximum principle), we note that the process  $(P_t^s)_{t \in [0, T], s \in \mathbb{R}^+}$  has an obvious link to the adjoint process in the maximum principle. This connection will be made clear in the discussion of Section VI.6. We are going to prove that under reasonable assumptions the system of mean-field Langevin equations has a unique solution, and the marginal distribution  $(\nu_t^s)_{t \in [0, T]}$  converges to the global minimizer of the control problem (VI.2.9) as  $s \rightarrow \infty$ .

### VI.2.3 Notation

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space where lives the Brownian motion  $W$ . Denote by  $\mathbb{E}$  the expectation under the probability  $\mathbb{P}$ , or equivalently, the expectation of the randomness produced by the Brownian motion  $W$ . In particular, note the difference between the notations  $\mathbb{E}$  and  $E$ .

In the present work we shall use several different metrics on the measure space. First, recall the  $p$ -Wasserstein distance ( $p \geq 1$ ) on the probability space  $\mathcal{P}(\mathbb{R}^m)$  :

$$\mathcal{W}_p(\mu_0, \nu_0)^p := \inf \left\{ \int_{\mathbb{R}^m} |x - y|^p \pi(dx, dy) : \text{where } \pi \text{ is a coupling of } \mu_0, \nu_0 \in \mathcal{P}(\mathbb{R}^m) \right\}.$$

Further, for  $\mu, \nu \in \mathcal{V}$  we define the metric

$$\overline{\mathcal{W}}_p^T(\mu, \nu) := \left( \int_0^T \mathcal{W}_p(\mu_t, \nu_t)^p dt \right)^{1/p}.$$

In some part of the work, in particular during the discussion of the convex case (see Section VI.3.2 and VI.6), we shall use the following generalized  $p$ -Wasserstein distance on  $\mathcal{V}$  :

$$\mathcal{W}_p^T(\mu, \nu) := T^{1/p} \mathcal{W}_p\left(\frac{\mu}{T}, \frac{\nu}{T}\right) \quad \text{for } \mu, \nu \in \mathcal{V}. \quad (\text{VI.2.11})$$

Comparing the above definitions, clearly we have  $\mathcal{W}_p^T(\mu, \nu) \leq \overline{\mathcal{W}}_p^T(\mu, \nu)$  for any  $\mu, \nu \in \mathcal{V}$ .

In the proofs of the following sections, the constant  $C$  may vary from line to line. Without further specification,  $C$  is always strictly positive.

## VI.3 Main Results

In this section we announce the main results. Their proofs are given in Sections VI.5, VI.6 and VI.7.

Throughout the paper we assume that the Hamiltonian function  $H$  and the terminal cost  $G$  are continuously differentiable in the variables  $(x, a)$ , and the coefficients

$$\phi, \nabla_x G, \nabla_x L, \nabla_x \phi \quad \text{exist and are all bounded.}$$

Therefore  $(X_t^s, P_t^s)$  lives in a compact set  $\mathcal{K}_x \times \mathcal{K}_p$ . From now on we treat  $(t, x, a, p, z) \mapsto H(t, x, a, p, z)$  as a function defined on  $[0, T] \times \mathcal{K}_x \times \mathbb{R}^m \times \mathcal{K}_p \times \mathcal{Z}$ . In particular, whenever we claim that  $H$  satisfies a property (e.g. Lipschitz continuity) globally, it is meant to be true on this set instead of the whole space.

### VI.3.1 Wellposedness of the System of Mean-Field Langevin Equation

**Assumption VI.3.1.** *Assume that the coefficients  $\phi, L, G$  are continuously differentiable in the variables  $(x, a)$  and*

- $\nabla_a \phi, \nabla_a L, \nabla_x \phi, \nabla_x L, \nabla_x G, \phi$  are uniformly Lipschitz continuous in the variables  $(x, a, p)$ ;
- the coefficients  $H, \nabla_a H$  satisfy

$$\sup_{t,z} |H(t, 0, 0, 0, z)| < \infty, \quad \sup_{t,z} |\nabla_a H(t, 0, 0, 0, z)| < \infty. \quad (\text{VI.3.12})$$

Define the space of the continuous measure flows on the horizon  $[0, S]$  :

$$C_p([0, S], \mathcal{V}) := \left\{ \mu = (\mu^s)_{s \in [0, S]} : \mu^s \in \mathcal{V} \text{ and } \lim_{s' \rightarrow s} \overline{W}_p^T(\mu^{s'}, \mu^s) = 0 \text{ for all } s \in [0, S] \right\}.$$

**Theorem VI.3.2.** *Let Assumption VI.3.1 hold true. Given  $(\Theta_t^0)_{t \in [0, T]}$  such that*

$$\int_0^T \mathbb{E}[|\Theta_t^0|^p] dt < \infty, \quad \text{for some } p \geq 1, \quad (\text{VI.3.13})$$

*the system of SDE (VI.2.10) has a unique solution. In particular, the law of the solution  $(\nu^s) \in C_p(\mathbb{R}^+, \mathcal{V})$ .*

One of our main contributions is to observe the decrease of energy along the flow of the solution to the system of mean-field Langevin equations (VI.2.10).

**Assumption VI.3.3.** *We further assume that*

- the coefficients  $\phi, L$  are second-order continuously differentiable in  $a$ ;
- there is  $\epsilon > 0$  such that

$$a \cdot \nabla_a H(t, x, a, p, z) \geq \epsilon |a|^2, \quad \text{for } |a| \text{ big enough}; \quad (\text{VI.3.14})$$

- for fixed  $(t, x, p)$  the mapping  $a \mapsto \mathbb{E}[\nabla_a H(t, x, a, p, Z)]$  belongs to  $C^3$ .

**Theorem VI.3.4** (Gradient flow). *Let Assumptions VI.3.1 and VI.3.3 hold true, and assume that*

$$\int_0^T \mathbb{E}[|\Theta_t^0|^p] dt < \infty \quad \text{for some } p \geq 2. \quad (\text{VI.3.15})$$

Recall the function  $V^\sigma$  defined in (VI.2.9). Let  $(\nu_t^s)$  be the marginal laws of the solution to the system of mean-field Langevin equations (VI.2.10). Then, for each  $s > 0$ ,  $t \in [0, T]$  the law  $\nu_t^s$  admits a density, and for  $s' > s > 0$  we have

$$V^\sigma(\nu^{s'}) - V^\sigma(\nu^s) = - \int_s^{s'} \int_0^T \int_{\mathbb{R}^m} \left| \mathbb{E} \left[ \nabla_a H(t, X_t^r, a, P_t^r, Z) \right] + \frac{\sigma^2}{2} \nabla_a \ln \nu_t^r(a) \right|^2 \nu_t^r(a) da dt d(VI.3.16)$$

## VI.3.2 Convex Case

We first consider the case where the objective function  $V$ , defined in (VI.2.8), is convex in  $\nu$ . More precisely, we assume the following.

**Assumption VI.3.5.** *Let the controlled process  $X$  be linear in  $\nu$ , i.e.*

$$X_t = X_0 + \int_0^t \int_{\mathbb{R}^m} \phi(s, a, Z) \nu_s(da) ds, \quad (VI.3.17)$$

and the dependence on the variables  $(x, a)$  of the function  $L$  is separated, i.e.

$$L(t, x, a, z) = \ell(t, x, z) + c(t, a, z).$$

Further we assume that

- for all  $t \in [0, T]$  the functions  $\ell, G$  are convex in  $x$ ;
- the functions  $\phi, \nabla_x H, \nabla_a H$  are globally Lipschitz continuous in  $t$ ;
- $H$  is continuously differentiable in  $t$  and  $\partial_t H$  is globally Lipschitz continuous in  $(t, a)$ .

**Remark VI.3.6.** *In the present section concerning the convex case we add the regularity assumptions on the coefficients with respect to the variable  $t$ . That is due to the fact that in this part of the paper we will apply the metric  $\mathcal{W}_p^T$  (defined in (VI.2.11)) on the space  $\mathcal{V}$  instead of the usual one  $\overline{\mathcal{W}}_p^T$ .*

Under the above assumptions, it is clear that there exists at least one global minimizer of  $V^\sigma$ . Moreover, the function  $V$  is convex in  $\nu$ , and thus  $V^\sigma$  is strictly convex in  $\nu$  for any  $\sigma > 0$ , so there is one unique global minimizer. By standard variational calculus, the following theorem states a sufficient condition for a control to be the unique global minimizer of the our problem (VI.2.9).

**Theorem VI.3.7** (Sufficient first order condition). *Let Assumption VI.3.5 holds true. If  $\nu^* \in \mathcal{V}$ , equivalent to the Lebesgue measure, satisfies*

$$\mathbb{E} \left[ \nabla_a H(t, X_t^*, \cdot, P_t^*, Z) \right] + \frac{\sigma^2}{2} \nabla_a \ln (\nu_t^*) = 0 \quad (VI.3.18)$$

for Leb-a.s.  $t$ , where  $X^*$  is the controlled process with the control  $\nu^*$  as in (VI.3.17) and  $P^*$  is the following adjoint process

$$P_t^* := \nabla_x G(X_T^*, Z) + \int_t^T \nabla_x \ell(r, X_r^*, Z) dr, \quad (\text{VI.3.19})$$

then  $\nu^*$  is an optimal control of the regularized control problem (VI.2.9).

Combining the sufficient condition above and Theorem VI.3.4, we can prove the following main result in the convex case.

**Theorem VI.3.8.** *Let Assumptions VI.3.1, VI.3.3 and VI.3.5 hold true and let  $(\Theta_t^0)_{t \in [0, T]}$  satisfy (VI.3.15) with  $p > 2$ . Further assume that  $V$  is  $\mathcal{W}_2^T$ -continuous and bounded from below. Denote  $(\nu_t^s)_{t \in [0, T]}^{s \in \mathbb{R}^+}$  the flow of marginal laws of the solution to (VI.2.10). Then there exists an invariant measure of (VI.2.10) equal to  $(\nu_t^*)_{t \in [0, T]} := \underset{\nu}{\operatorname{argmin}} V^\sigma(\nu)$ , and  $(\nu_t^s)_{t \in [0, T]}$  converges to  $(\nu_t^*)_{t \in [0, T]}$ .*

### VI.3.3 Contraction Case

Clearly, the previous convex case has restrictive requirements on the structure of the coefficients. In particular, these requirements cannot all be satisfied in the application to deep neural networks. That drives us to look for another setting in which the system of mean-field Langevin equations leads us to the optimal control of (VI.2.9).

**Proposition VI.3.9** (Necessary first order condition). *Let Assumptions VI.3.1, VI.3.3 hold true and assume that the function  $V$  is bounded from below. Let  $\nu^*$  be an optimal control of  $V^\sigma$  such that  $V^\sigma(\nu^*) < \infty$ . Then  $\nu^*$  is equivalent to the Lebesgue measure and satisfies*

$$\mathbb{E} \left[ \nabla_a H(t, X_t^*, \cdot, P_t^*, Z) \right] + \frac{\sigma^2}{2} \nabla_a \ln(\nu_t^*) = 0, \quad \text{for Leb-a.s. } t \in [0, T], \quad (\text{VI.3.20})$$

where  $(X^*, P^*)$  is the solution of the following ODE

$$\begin{aligned} X_t^* &= X_0 + \int_0^t \int_{\mathbb{R}^m} \phi(r, X_r^*, a, Z) \nu_r^*(da) dr, \\ P_t^* &= \nabla_x G(X_T^*, Z) + \int_t^T \int_{\mathbb{R}^m} \nabla_x H(r, X_r^*, a, P_r^*, Z) \nu_r^*(da) dr. \end{aligned}$$

In particular,  $\nu^*$  is an invariant measure of the system (VI.2.10) satisfying

$$-C(1 + |a|^2) \leq \ln(\nu_t^*(a)) \leq C - C'|a|^2, \quad \text{for some } C \geq C' > 0 \text{ independent of } t, \quad (\text{VI.3.21})$$

and thus  $\nu^*$  has finite  $p$ -moment for all  $p \geq 1$ .

Assume that the control problem (VI.2.9) admits at least one optimal control. The proposition above implies that once we ensure the convergence of the marginal laws of the system (VI.2.10) towards the unique invariant measure, then the limit measure is an (indeed the unique) optimal control of (VI.2.9).



Next we find a sufficient condition for the existence of the unique invariant measure for the system of mean-field Langevin equations. In particular, the convergent rate towards the limit measure is computed explicitly.

**Assumption VI.3.10.** *Assume that there is a continuous function  $\kappa : (0, +\infty) \rightarrow \mathbb{R}$  such that  $\limsup_{r \rightarrow +\infty} \kappa(r) < 0$ ,  $\int_0^1 r\kappa(r)dr < +\infty$  and for any  $(t, x, p, z)$  we have*

$$(a - \tilde{a}) \cdot \left( -\nabla_a H(t, x, a, p, z) + \nabla_a H(t, x, \tilde{a}, p, z) \right) \leq \kappa(|a - \tilde{a}|) |a - \tilde{a}|^2 \quad \text{for all } a, \tilde{a} \in \mathbb{R}^m, a \neq \tilde{a}.$$

**Remark VI.3.11.** *Take the example where the cost function is  $L(t, x, a, z) = \lambda|a|^2$ . Then Assumption VI.3.10 holds true for all  $\lambda > 0$ , provided that  $\nabla_a \phi$  is bounded.*

**Theorem VI.3.12** (Contraction towards the unique invariant measure). *Let Assumptions VI.3.1, VI.3.3 and VI.3.10 hold true. Let  $(\Theta_t^0)_{t \in [0, T]}$ , resp.  $(\tilde{\Theta}_t^0)_{t \in [0, T]}$ , satisfy (VI.3.15) and denote  $\nu^s$ , resp.  $\tilde{\nu}^s$ , the marginal law of the solution of the mean-field Langevin system (VI.2.10). Then there exist constants  $c, \gamma \in (0, \infty)$  such that for any  $s \geq 0$ , we have*

$$\overline{\mathcal{W}}_1^T(\nu^s, \tilde{\nu}^s) \leq e^{\left(\frac{2\gamma T}{\varphi(R_1)} - c\sigma^2\right)s} \frac{2}{\varphi(R_1)} \overline{\mathcal{W}}_1^T(\nu^0, \tilde{\nu}^0), \quad (\text{VI.3.22})$$

where denoting by  $(K_x, K_p)$  the Lipschitz coefficients for the variables  $(x, p)$  of the functions  $\nabla_a H(\cdot, a)$ ,  $\nabla_x H$  and  $\phi$ , by  $K_a$  the Lipschitz coefficient for the variable  $a$  of the functions  $\nabla_x H$  and  $\phi$ , and by  $K_g$  the Lipschitz coefficient of the function  $\nabla_x G$ , we have

$$\gamma = K_x K_a e^{K_x T} + K_p K_a (K_g e^{K_x T} + e^{K_x T} + 1) e^{K_p T},$$

The other coefficients are explicitly given by

$$\varphi(r) = \exp\left(-\frac{1}{2} \int_0^r \frac{u\kappa^+(u)}{\sigma^2} du\right) \quad \text{and} \quad c^{-1} = \int_0^{R_2} \Phi(s)\varphi(s)^{-1} ds,$$

where  $\Phi(r) = \int_0^r \varphi(s) ds$ , and

$$R_1 := \inf\{R \geq 0 : \kappa(r) \leq 0 \text{ for all } r \geq R\}, \quad R_2 := \inf\{R \geq R_1 : \kappa(r)R(R-R_1) \leq -4\sigma^2 \text{ for all } r \geq R\}.$$

In particular, if

$$\frac{2\gamma T}{\varphi(R_1)} < c\sigma^2, \quad (\text{VI.3.23})$$

then there is a unique invariant measure with finite 2-moment.

**Remark VI.3.13.** *The result and the proof of Theorem VI.3.12 reveal the importance of considering*

the relaxed formulation of the control problem instead of the strict one (VI.1.1). As discussed in the introduction, in the setting of the strict formulation, one may let the control  $(\alpha_t)_{t \in [0, T]}$  evolve along the gradient as in (VI.1.3). In this case the limit  $\lim_{s \rightarrow \infty} (a_t^s)_{t \in [0, T]}$ , if exists, depends in general on the initial value  $(\alpha_t^0)_{t \in [0, T]}$ , so it is unlikely to be the optimal control. On the contrary, Theorem VI.3.12 ensures that the gradient flow (VI.2.10) in the relaxed formulation converges to the unique invariant measure independent of the initialization  $\nu^0$ .

**Remark VI.3.14.** Theorem VI.3.12 sheds some light on the exponential ergodicity of the mean-field Langevin system. Its proof is largely based on the reflection coupling studied in [48, 49]. To make the restriction (VI.3.23) more transparent, we consider the case where the running cost is  $L(t, x, a, z) = \frac{1}{2} \lambda |a|^2$ . Then the inequality (VI.3.23) can hold true in one of the following cases :

- $T$  is small,
- $\sigma$  is large,
- $K_a, K_x, K_p$  are small,
- $\lambda$  is large.

In the application to neural networks, one may need to keep the regularizers  $\sigma, \lambda$  not big in order to avoid bias, so using small  $T$  or  $K$  becomes necessary to make (VI.3.23) satisfied. Unfortunately, it may reduce the expressiveness of the network. Therefore, Theorem VI.3.12 only provides a partial answer to the trainability of the deep neural networks.

In order to link the unique invariant measure to the optimal control of (VI.2.9), we need an additional assumption.

**Assumption VI.3.15.** Assume that

- the functional  $V$  is bounded from below,
- $V^\sigma$  has at least an optimal control  $\nu^*$  such that  $V^\sigma(\nu^*) < \infty$ .

In particular, we will see a sufficient condition for the existence of optimal control in Lemma VI.7.1. Finally, combining the results in Proposition VI.3.9 and Theorem VI.3.12, we may conclude :

**Corollary VI.3.16.** Let Assumptions VI.3.1, VI.3.3, VI.3.10 and VI.3.15 hold true. Recall the constants defined in Theorem VI.3.12. If  $\frac{2\gamma T}{\varphi(R_1)} < c\sigma^2$ , then the unique invariant measure of (VI.2.10) is the optimal control of (VI.2.9).

## VI.4 Application to Deep Neural Networks

### VI.4.1 Discretization of mean-field Langevin dynamics and gradient descent

In this subsection we report more explicitly the relation between the mean-field Langevin dynamics and the gradient descent algorithm in practice.

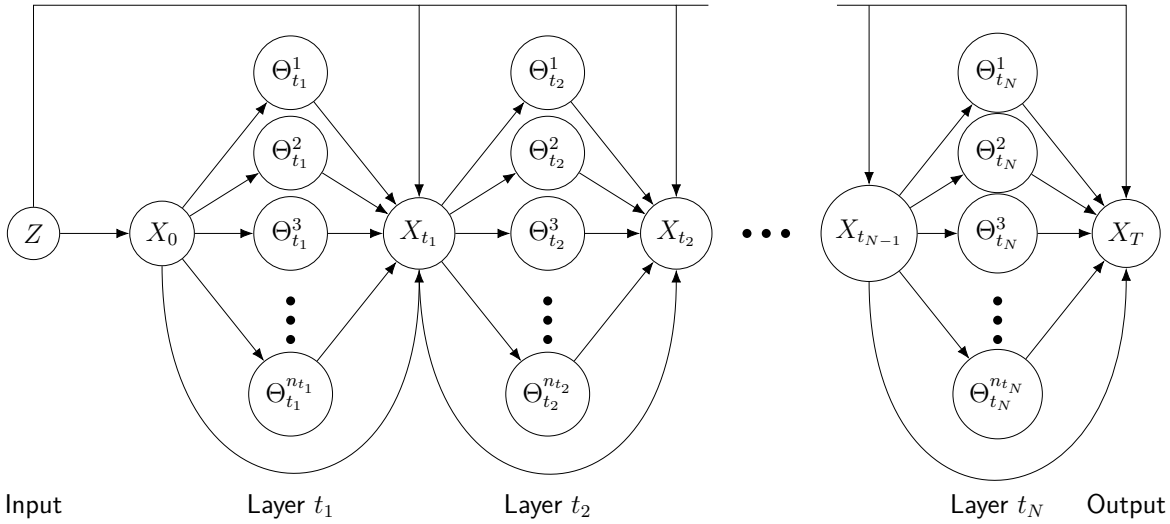
Let the data  $Z$  take values in a compact subset  $\mathcal{Z}$  of  $\mathbb{R}^D$ , and denote by  $Y = f(Z)$  the label of the data. The function  $f$  is unknown, and we want to approximate it using the parametrized function generated by a deep neural network. Recall the controlled dynamics that we use to model the deep neural network :

$$X_T = X_0(Z) + \int_0^T \int_{\mathbb{R}^m} \phi(t, X_t, a, Z) \nu_t(da) dt,$$

where  $\phi$  is the fixed activation function. Once we discretize the equation using the explicit Euler scheme and introduce random variables  $(\Theta_t^j)_{j=1, \dots, n_t}$ , which are independent copies following the law  $\nu_t$ , we find

$$X_{t_{i+1}} \approx X_{t_i} + \frac{\delta t}{n_{t_{i+1}}} \sum_{j=1}^{n_{t_{i+1}}} \phi(t_i, X_{t_i}, \Theta_{t_{i+1}}^j, Z), \quad \text{where } \delta t := t_{i+1} - t_i. \quad (\text{VI.4.24})$$

This discrete dynamics characterizes a type of structure of deep neural networks, and can be equivalently represented by the scheme in Figure VI.1. It appears to be a simplification of Residual Neural Network,



**Figure VI.1:** Neural network corresponding to the relaxed controlled process

see [72].

**Remark VI.4.1.**

- In the previous ODE models on the residual neural networks, the data  $Z$  only enters through the initial value  $X_0$ . Here, in addition, we allow the dependence on the data in the dynamics.
- The most significant feature of this structure is the average pooling (averaging the outputs of the nonlinear activation  $\phi$  as in (VI.4.24)) on each layer, and it is due to the adoption of the relaxed formulation in our model of controlled process.

Given the structure of the neural network, or the scheme of the forward propagation (VI.4.24), it is conventional to optimize the parameters  $(\Theta_{t_i}^j)$  using the stochastic gradient descent algorithm. The

gradients of the parameters are easy to compute, due to the chain rule (or backward propagation) :

$$d\Theta_{t_i}^{j,s} = -\mathbb{E}\left[\nabla_a H(t_i, X_{t_i}^s, \Theta_{t_i}^{j,s}, P_{t_i}^s, Z)\right]ds + \sigma dW_s, \quad (\text{VI.4.25})$$

where  $H(t, x, a, p, z) = \phi(t, x, a, z) \cdot p + \frac{1}{2}\lambda|a|^2$

$$P_{t_{i-1}}^s = P_{t_i}^s - \delta t \sum_{j=1}^{n_{t_{i+1}}} \nabla_x H(t_i, X_{t_i}^s, \Theta_{t_i}^{j,s}, P_{t_i}^s, Z), \quad P_T^s = \nabla_x G(X_T^s, Z),$$

where  $\lambda$  is the coefficient of ridge regularization on each layer, and the horizon  $s \in \mathbb{R}^+$  represents the iterations of gradient descent. In the conventional gradient descent algorithm  $\sigma$  is set to be 0, whereas we add a (small) positive volatility in our model of the regularized optimal control problem. It is important to observe that the above noisy gradient descent dynamics approximates exactly the mean-field Langevin system (VI.2.10), as the number of time steps on the horizon  $t \in [0, T]$  and the number of particles tend to infinity.

**Remark VI.4.2.** Treating the particle system  $(\Theta_{t_i}^j)_{i,j}$  as a high dimensional vector, observe that the dynamics in (VI.4.25) is a Markov diffusion process. Under Assumptions VI.3.1, VI.3.3 and VI.3.10, the marginal laws  $\mu_s := \text{Law}((\Theta_{t_i}^j)_{i,j})$  converge to its invariant measure exponentially quickly, see e.g. [48, Corollary 2]. However, it is important to note that the convergence rate depends on the number of particles. In particular, without conditions such as (VI.3.23), in general the convergence rate would decrease to 0 as the number of particles tends to infinity and it can be viewed as the curse of dimensionality in the over-parametrized optimization. Here we provide a numerical test in training neural works to reflect the theoretical observation. In the toy example, we try to learn the one-dimensional indicator function  $1_{[3,4]}$  on the horizon  $[0, 5]$  using a 20-layer ‘deep’ neural network equipped with the structure in Figure VI.1. With all the other parameters of the networks fixed, we compare the learning errors of the two neural networks : one with 10 neurons per layer, the other with 100 neurons per layer. The result in Figure VI.2 shows that the training errors converge more slowly for the deep network with 100 neurons per layer.

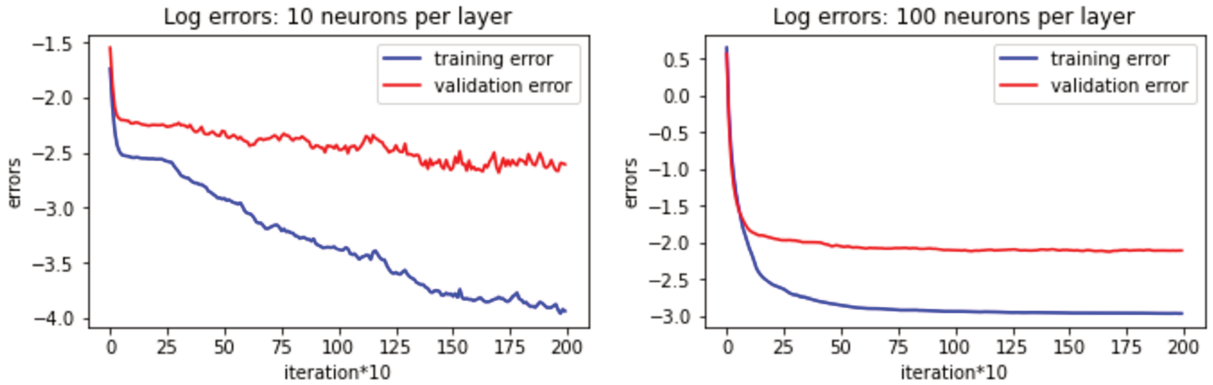
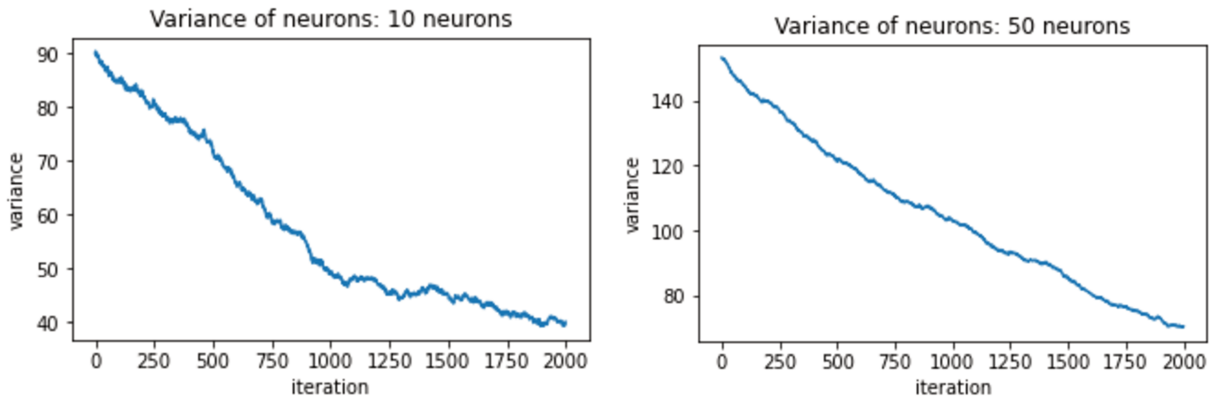


Figure VI.2: Curse of dimensionality in training deep networks

**Remark VI.4.3.** *The curse of dimensionality mentioned above echoes with the common practice, that is, the layers in a deep network are usually set to be narrow. This appears to be in contrast with the fact that in order to use the empirical approximation  $\frac{1}{n_t} \sum_{j=1}^{n_t} \delta_{\Theta_t^j}$  for the law  $\nu_t$ , one needs in general a large number of neurons  $(\Theta_t^j)_j$ . To address this inconsistency heuristically, note that the optimal control  $\nu_t^*$  satisfies the first order equation (VI.3.20), or equivalently,*

$$\nu_t^*(a) = C e^{-\frac{2}{\sigma^2} \mathbb{E} \left[ H(t, X_t^*, a, P_t^*, Z) \right]} = C e^{-\frac{2}{\sigma^2} \mathbb{E} \left[ P_t^* \cdot \phi(t, X_t^*, a) + \lambda |a|^2 \right]}.$$

*In particular, given the value of  $X_t^*$  and  $P_t^*$ , the probability measure  $\nu_t^*$  would concentrate on the support  $\{\operatorname{argmin}_a \mathbb{E} \left[ H(t, X_t^*, a, P_t^*, Z) \right]\}$  as  $\sigma \rightarrow 0$ . Therefore, one might hope to be able to approximate  $\nu_t^*$  using a small number of neurons in the case when  $\sigma$  is small. Further, since during the training  $\nu_t^s$  converges towards  $\nu_t^*$ , the mass of  $\nu_t^s$  should also become concentrated when  $s$  is big enough. While a quantitative study is left for the future research, here we show a numerical test to support the heuristic assertion. As in Remark VI.4.2, we use a 20-layer neural network to learn the toy function  $1_{[3,4]}$ , and set the number of neurons on each layer to be 10 or 50. In both cases, we track the variance of the values of the neurons on the first layer, namely  $(\Theta_{t_1}^i)_i$ , through the training. The result in Figure VI.3 shows*



**Figure VI.3:** Variance of the neurons

*that the variance of the neurons decreases during the training, implying that  $\nu_t$  becomes more and more concentrated and thus can be approximated by fewer and fewer neurons. Based on the observation, one may be encouraged to reduce the number of neurons along the training. Note that such possibility has been explored in [137] for the two-layer neural networks.*

## VI.4.2 A model for studying the dropout

In practice, when training deep networks, at each iteration one can ignore (or ‘drop out’) some randomly chosen neurons. It is well known that the *dropout* technique can help prevent the overfitting. In plain words, dropout is a regularization method that approximates training a large number of neural

networks with different architectures in parallel, see [144]. Here we provide a mathematical perspective for the efficiency of the method, based on the relaxed control model.

Here we focus on the dropout of the hidden neurons (denoted by  $\Theta$  in the figure above), while similar discussion can be made for the visible neurons (denoted by  $X$  above). We need to assume that the activation function satisfies :

$$\phi(t, x, a, z)q = \phi(t, x, qa, z), \quad \text{for all } q \in [0, 1], \quad \text{and} \quad \phi_x(t, x, 0, z) = 0.$$

Introduce the dropout coefficients  $(q_t)_{t \in [0, T]}$  such that  $q_t \in [0, 1]$  for each  $t$ . Then the scaled forward propagation reads

$$\begin{aligned} X_T &= X_0(Z) + \sum_{j=1}^N \int_{\mathbb{R}^m} \phi(t_j, X_{t_j}, q_{t_j} a, Z) \nu_{t_j}(da) (t_{j+1} - t_j) \\ &= X_0(Z) + \sum_{j=1}^N \int_{\mathbb{R}^m} \phi(t_j, X_{t_j}, a, Z) \nu_{t_j}(da) q_{t_j} (t_{j+1} - t_j). \end{aligned}$$

Recall the Hamiltonian function  $H(t, x, a, p, z) = \phi(t, x, a, z) \cdot p + \frac{1}{2} \lambda |a|^2$ . Now for each  $i \in \mathbb{N}$  introduce a random process  $(U_t^{s_i})_{t \in [0, T]}$  independent from  $W$  such that  $\mathbb{P}[U_t^{s_i} = 1] = q_t$  and  $\mathbb{P}[U_t^{s_i} = 0] = 1 - q_t$ . Also assume that  $(U_t^{s_i})_i$  are independent. The random variable  $U_t^s$  models the dropout in practice, namely, if  $U_t^s = 0$  then the corresponding neuron on the  $t$ -th layer will not be trained during the  $s$ -th iteration. Further define the process :

$$\begin{aligned} d\widehat{\Theta}_t^s &= \sum_{i=1}^n \mathbf{1}_{s \in [s_i, s_{i+1})} \left( -U_t^{s_i} \mathbb{E} \left[ \nabla_a \phi(t, \widehat{X}_t^s, \widehat{\Theta}_t^s, Z) \right] \widehat{P}_t^s - \lambda \widehat{\Theta}_t^s \right) ds + \sigma dW_s, \\ \text{where} \quad \widehat{X}_{t_{j+1}}^s &= \widehat{X}_{t_j}^s + \mathbb{E} \left[ \phi(t_j, \widehat{X}_{t_j}^s, \widehat{\Theta}_{t_{j+1}}^s, Z) \right] q_{t_j} (t_{j+1} - t_j) \\ &= \widehat{X}_{t_j}^s + \mathbb{E} \left[ \phi(t_j, \widehat{X}_{t_j}^s, U_{t_j}^s \widehat{\Theta}_{t_{j+1}}^s, Z) \right] (t_{j+1} - t_j), \\ \widehat{P}_{t_{j-1}}^s &= \widehat{P}_{t_j}^s - \mathbb{E} \left[ \nabla_x H(t_j, \widehat{X}_{t_j}^s, \widehat{\Theta}_{t_j}^s, \widehat{P}_{t_j}^s, Z) \right] q_{t_j} (t_j - t_{j-1}) \\ &= \widehat{P}_{t_j}^s - \mathbb{E} \left[ \nabla_x H(t_j, \widehat{X}_{t_j}^s, U_{t_j}^s \widehat{\Theta}_{t_j}^s, \widehat{P}_{t_j}^s, Z) \right] (t_j - t_{j-1}) \quad \widehat{P}_T^s = \nabla_x G(\widehat{X}_T^s, Z). \end{aligned}$$

The dynamics of  $\widehat{\Theta}^s$  can be viewed as the training of the neurons with the dropout involved. Fixing the partition  $(t_j)_j$  on  $[0, T]$ , as the time-grid on the horizon  $s \in \mathbb{R}^+$  becomes infinitely dense, we have the following convergence.

**Lemma VI.4.4.** Fix any  $S > 0$ . Denote by  $\mu^n := \mathbb{P} \circ (\widehat{\Theta}_{t_j}^{\cdot})_{j=1, \dots, N}^{-1} \in \mathcal{P}(C[0, S]^N)$ , and set the time

steps  $s_i \in \{\frac{k}{2^n}S\}$ . Then  $\mu^n$  weakly converges to the law of the following process  $s \in [0, S] \mapsto (\Theta_{t_j}^s)_j$  :

$$d\Theta_t^s = \left( -q_t \mathbb{E} \left[ \nabla_a \phi(t, X_t^s, \Theta_t^s, Z) \right] P_t^s - \lambda \Theta_t^s \right) ds + \sigma dW_s,$$

where

$$X_{t_{j+1}}^s = X_{t_j}^s + \mathbb{E} \left[ \phi(t_j, X_{t_j}^s, \Theta_{t_{j+1}}^s, Z) \right] q_{t_j} (t_{j+1} - t_j),$$

$$P_{t_{j-1}}^s = P_{t_j}^s - \mathbb{E} \left[ \nabla_x H(t_j, X_{t_j}^s, \Theta_{t_j}^s, P_{t_j}^s, Z) \right] q_{t_j} (t_j - t_{j-1}), \quad P_T^s = \nabla_x G(X_T^s, Z).$$

**Remark VI.4.5.** Here we work on the discrete processes on the  $t$ -horizon for technical reasons. Ignoring this difference, the limit process above is related to the control problem :

$$\inf_{\nu} \left( V(\nu) + \frac{\sigma^2}{2q_t} \int_0^T \text{Ent}(\nu_t) dt \right), \quad \text{where} \quad V(\nu) := \mathbb{E} \left[ \int_0^T \int_{\mathbb{R}^m} \frac{\lambda}{2q_t} |a|^2 \nu_t(da) dt + G(X_T, Z) \right],$$

$$dX_t = \int \phi(t, X_t, a, Z) \nu_t(da) dt.$$

Comparing it with the initial control problem, we observe that both regularization coefficients  $\lambda, \sigma$  become bigger. Therefore, we conclude that with the dropout the training is more regularized, so is capable of reducing the overfitting.

*Démonstration.* Denote by  $s \mapsto \tilde{\Theta}^s$  the canonical process on  $C[0, S]^N$  and by  $\mathcal{F}_s$  the canonical filtration. First observe that

$$\sup_n \mathbb{E}^{\mu^n} \left[ |\tilde{\Theta}^s - \tilde{\Theta}^{s'}| \right] \leq C \left( |s - s'| + \sqrt{|s - s'|} \right),$$

and  $(\mu^n)$  lies in the weakly compact. By Itô's formula, we have for any smooth function  $\varphi$  that

$$\begin{aligned} d\varphi(\tilde{\Theta}^s) &= \left( \sum_i 1_{s \in [s_i, s_{i+1})} \nabla \varphi(\tilde{\Theta}^s) \cdot \left( -U_{t_j}^{s_i} \mathbb{E} \left[ \nabla_a \phi(t_j, \widehat{X}_{t_j}^s, \tilde{\Theta}_j^s, Z) \right] \widehat{P}_{t_j}^s - \lambda \tilde{\Theta}_j^s \right) + \frac{1}{2} \sigma^2 \Delta \varphi(\tilde{\Theta}^s) \right) ds + dM_s^n \\ &= \left( \sum_i 1_{s \in [s_i, s_{i+1})} \left( -\nabla \varphi(\tilde{\Theta}^{s_i}) \cdot \left( U_{t_j}^{s_i} \mathbb{E} \left[ \nabla_a \phi(t_j, \widehat{X}_{t_j}^s, \tilde{\Theta}_j^{s_i}, Z) \right] \widehat{P}_{t_j}^s \right)_{j=1, \dots, N} + R_i \cdot (\tilde{\Theta}^s - \tilde{\Theta}^{s_i}) \right) \right. \\ &\quad \left. - \nabla \varphi(\tilde{\Theta}^s) \cdot \lambda \tilde{\Theta}^s + \frac{1}{2} \sigma^2 \Delta \varphi(\tilde{\Theta}^s) \right) ds + dM_s^n, \quad \mu^n\text{-a.s.} \end{aligned}$$

where  $R_i$  is a bounded random variable. Note that  $(U_t^{s_i})_t$  is independent from  $(\tilde{\Theta}_t^{s_i})_t$ . Therefore

$$\begin{aligned} & \mathbb{E}^{\mu^n} [\varphi(\tilde{\Theta}^{s_{i+1}}) | \mathcal{F}_{s_i}] - \varphi(\tilde{\Theta}^{s_i}) \\ &= \mathbb{E}^{\mu^n} \left[ \int_{s_i}^{s_{i+1}} \left( -\nabla \varphi(\tilde{\Theta}^{s_i}) \cdot \left( q_{t_j} \mathbb{E} [\nabla_a \phi(t_j, \widehat{X}_{t_j}^r, \tilde{\Theta}_j^{s_i}, Z)] \widehat{P}_{t_j}^r \right)_{j=1, \dots, N} + R_i \cdot (\tilde{\Theta}^r - \tilde{\Theta}^{s_i}) \right. \right. \\ & \quad \left. \left. - \nabla \varphi(\tilde{\Theta}^r) \cdot \lambda \tilde{\Theta}^r + \frac{1}{2} \sigma^2 \Delta \varphi(\tilde{\Theta}^r) \right) dr \middle| \mathcal{F}_{s_i} \right] \\ &= \mathbb{E}^{\mu^n} \left[ \int_{s_i}^{s_{i+1}} \left( \mathcal{L}^n \varphi(r, \tilde{\Theta}^r) + \tilde{R}_i \cdot (\tilde{\Theta}^r - \tilde{\Theta}^{s_i}) \right) dr \middle| \mathcal{F}_{s_i} \right], \end{aligned}$$

where  $\mathcal{L}^n \varphi(s, a) = -\nabla \varphi(a) \cdot \left( q_{t_j} \mathbb{E} [\nabla_a \phi(t_j, \widehat{X}_{t_j}^s, a, Z)] \widehat{P}_{t_j}^s \right)_{j=1, \dots, N} - \nabla \varphi(a) \cdot \lambda a + \frac{1}{2} \sigma^2 \Delta \varphi(a)$

Let  $\mu^*$  be the limit of a subsequence of  $\mu^n$ , and denote by  $\widehat{X}^*, \widehat{P}^*$  the associated processes and by  $\mathcal{L}^*$  the associated operator. Clearly, we have

$$\widehat{X} \rightarrow \widehat{X}^*, \quad \widehat{P} \rightarrow \widehat{P}^*, \quad \mathbb{E}^{\mu^n} [\tilde{R}_i \cdot (\tilde{\Theta}^r - \tilde{\Theta}^{s_i})] \leq C(|r - s_i| + \sqrt{|r - s_i|}) \rightarrow 0,$$

as the time grid becomes infinitely dense. Therefore we have

$$\begin{aligned} & \mathbb{E}^{\mu^*} \left[ \varphi(\tilde{\Theta}^{s_k}) - \varphi(\tilde{\Theta}^{s_i}) - \int_{s_i}^{s_k} \mathcal{L}^* \varphi(s, \tilde{\Theta}^s) ds \middle| \mathcal{F}_{s_i} \right] \\ &= \lim_{n \rightarrow \infty} \mathbb{E}^{\mu^n} \left[ \varphi(\tilde{\Theta}^{s_k}) - \varphi(\tilde{\Theta}^{s_i}) - \int_{s_i}^{s_k} \mathcal{L}^* \varphi(s, \tilde{\Theta}^s) ds \middle| \mathcal{F}_{s_i} \right] \\ &= \lim_{n \rightarrow \infty} \mathbb{E}^{\mu^n} \left[ \varphi(\tilde{\Theta}^{s_k}) - \varphi(\tilde{\Theta}^{s_i}) - \int_{s_i}^{s_k} \left( \mathcal{L}^n \varphi(r, \tilde{\Theta}^r) + \sum_{i \leq j < k} 1_{s \in [s_j, s_{j+1})} \tilde{R}_j \cdot (\tilde{\Theta}^r - \tilde{\Theta}^{s_j}) \right) ds \middle| \mathcal{F}_{s_i} \right] = 0, \end{aligned}$$

that is,  $\varphi(\tilde{\Theta}^s) - \int_0^s \mathcal{L}^* \varphi(r, \tilde{\Theta}^r) dr$  is a  $\mu^*$ -martingale for all smooth  $\varphi$ . By the uniqueness of the strong solution to the McKean-Vlasov SDE, we can conclude that  $\mu^* = \mathbb{P} \circ (\Theta_{t_j}^i)_{j=1, \dots, N}^{-1}$ .  $\blacksquare$

## VI.5 Wellposedness of the system of mean-field Langevin equations

### VI.5.1 Wellposedness of the System

**Proof of Theorem VI.3.2 :** Let  $S > 0$ . Given any process  $(\mu^s)_{s \in [0, S]} \in C_p([0, S], \mathcal{V})$ , we define for any  $t \in [0, T]$  the process  $(\Theta_t^s)_{s \in [0, S]}$  as the solution of the classical SDE :

$$\left\{ \begin{array}{l} d\Theta_t^s = -\mathbb{E} [\nabla_a H(t, X_t^s(\mu), \Theta_t^s, P_t^s(\mu), Z)] ds + \sigma dW_s, \quad \text{for } s \in [0, S], \\ \text{where } X_t^s(\mu) = X_0 + \int_0^t \int_{\mathbb{R}^m} \phi(r, X_r^s(\mu), a, Z) \mu_r^s(da) dr, \\ P_t^s(\mu) = \nabla_x G(X_T^s, Z) + \int_t^T \int_{\mathbb{R}^m} \nabla_x H(r, X_r^s(\mu), a, P_r^s(\mu), Z) \mu_r^s(da) dr, \quad \text{for } t \in [0, T], \end{array} \right. \quad (\text{VI.5.26})$$



We are going to show that, for  $S$  small enough and independent of  $\Theta_t^0$ , the mapping  $(\mu^s)_{s \in [0, S]} \mapsto (dt \times \text{Law}(\Theta_t^s))_{s \in [0, S]} =: (\nu^s)_{s \in [0, S]}$  is a contraction on the space  $C_p([0, S], \mathcal{V})$  with the following metric :

$$d_p^{T, S}(\nu, \mu) := \sup_{s \leq S} \overline{\mathcal{W}}_p^T(\nu^s, \mu^s),$$

and thus has a fixed point. Then the existence of a unique solution to the system (VI.2.10) follows.

*Step 1.* First we will show the following property for the image of the mapping :  $(\nu^s)_{s \in [0, S]} \in C_p([0, S], \mathcal{V})$ . It suffices to show that

$$\lim_{s' \rightarrow s} \int_0^T \mathbb{E}[|\Theta_t^{s'} - \Theta_t^s|^p] dt = 0 \quad \text{for all } s \in [0, S]. \quad (\text{VI.5.27})$$

Since  $\phi, \nabla_x H, \nabla_x G$  are all bounded, the processes  $(X_t^s), (P_t^s)$  are both uniformly bounded. Further, by the assumption (VI.3.12), the drift terms of the SDEs (VI.5.26) are of linear growth uniformly in  $t$ . Then it follows from the standard estimate of SDE that

$$\mathbb{E}\left[\sup_{s \in [0, S]} |\Theta_t^s|^p\right] \leq C \left(\mathbb{E}[|\Theta_t^0|^p] + 1\right) \quad \text{for all } t \in [0, T],$$

where  $C$  is a constant independent of  $t$ . By the assumption (VI.3.13), we have  $\int_0^T \mathbb{E}\left[\sup_{s' \in [0, S]} |\Theta_t^{s'} - \Theta_t^s|^p\right] dt < \infty$  for all  $s \in [0, S]$ . Then (VI.5.27) follows from the dominated convergence theorem.

*Step 2.* Let  $(\mu^s)_{s \in [0, S]}, (\tilde{\mu}^s)_{s \in [0, S]} \in C_p([0, S], \mathcal{V})$  and denote  $(\Theta_t^s)_{s \in [0, S]}$  and  $(\tilde{\Theta}_t^s)_{s \in [0, S]}$  the corresponding solutions of the SDEs (VI.5.26), respectively. Denote  $\nu_t^s := \text{Law}(\Theta_t^s)$  and  $\tilde{\nu}_t^s := \text{Law}(\tilde{\Theta}_t^s)$ . Denote  $(K_x, K_p, K_a)$  the Lipschitz coefficients in  $(x, p, a)$  of  $\nabla_x H$  and  $\phi$ , and by  $K_g$  the Lipschitz coefficient of  $\nabla_x G$ . We have

$$\begin{aligned} |\delta X_t^s| &:= |X_t^s(\mu) - X_t^s(\tilde{\mu})| \leq \int_0^t (K_x |X_r^s(\mu) - X_r^s(\tilde{\mu})| + K_a \mathcal{W}_1(\mu_r^s, \tilde{\mu}_r^s)) dr, \\ \text{and thus } |\delta X_t^s| &\leq K_a e^{K_x t} \overline{\mathcal{W}}_1^t(\mu^s, \tilde{\mu}^s) \leq K_a e^{K_x T} \overline{\mathcal{W}}_p^T(\mu^s, \tilde{\mu}^s). \end{aligned} \quad (\text{VI.5.28})$$

Similarly, we obtain

$$\begin{aligned} |\delta P_t^s| &\leq |\nabla_x G(X_T^s(\mu)) - \nabla_x G(X_T^s(\tilde{\mu}))| + \int_0^t K_a \mathcal{W}_1(\mu_r^s, \tilde{\mu}_r^s) dr + \int_0^t (K_x |\delta X_r^s| + K_p |\delta P_r^s|) dr \\ &\leq (K_g K_a e^{K_x T} + K_a) \overline{\mathcal{W}}_p^T(\mu^s, \tilde{\mu}^s) + K_x K_a \left(\int_0^t e^{K_x r} dr\right) \overline{\mathcal{W}}_p^T(\mu^s, \tilde{\mu}^s) + \int_0^t K_p |\delta P_r^s| dr \\ &\leq (K_g K_a e^{K_x T} + K_a + K_a e^{K_x T}) \overline{\mathcal{W}}_p^T(\mu^s, \tilde{\mu}^s) + \int_0^t K_p |\delta P_r^s| dr, \end{aligned}$$

$$\text{and thus } |\delta P_t^s| \leq (K_g K_a e^{K_x T} + K_a + K_a e^{K_x T}) e^{K_p T} \overline{\mathcal{W}}_p^T(\mu^s, \tilde{\mu}^s). \quad (\text{VI.5.29})$$

Then define  $\delta\Theta_t^s := \Theta_t^s - \tilde{\Theta}_t^s$ , and we can similarly estimate

$$\begin{aligned} \int_0^T |\delta\Theta_t^s|^p dt &\leq C(p)S^{p-1}K^p e^{C(p)S^{p-1}K^p s} \int_0^T \int_0^s \left( |\delta X_t^r|^p + |\delta P_t^r|^p \right) dr dt \\ &\leq \tilde{C}(p)S^{p-1}K^{p+1}(1+K)e^{C(p)S^{p-1}K^p s + 2KT} \int_0^s \overline{\mathcal{W}}_p^T(\mu^r, \tilde{\mu}^r)^p dr, \end{aligned}$$

where  $C(p), \tilde{C}(p)$  are constants only depending on  $p$ . By taking the expectation on both sides we get

$$\overline{\mathcal{W}}_p^T(\nu^s, \tilde{\nu}^s)^p \leq \tilde{C}(p)S^p K^{p+1}(1+K)e^{C(p)S^p K^p + 2KT} d_p^{T,S}(\mu, \tilde{\mu})^p \quad \text{for } s \in [0, S].$$

Therefore, for  $S$  small enough, the mapping  $(\mu^s)_{s \in [0, S]} \mapsto (\nu^s)_{s \in [0, S]}$  is a contraction.  $\blacksquare$

Next we provide some useful estimates for the solution to the system (VI.2.10).

**Lemma VI.5.1.** *Let Assumption VI.3.1 hold true. Let  $(\Theta_t^0), (\tilde{\Theta}_t^0)$  be two initial values satisfying (VI.3.13), and denote by  $(\nu_t^s), (\tilde{\nu}_t^s)$  the marginal laws of the solutions to the system (VI.2.10), respectively. Then*

$$\overline{\mathcal{W}}_p^T(\nu^s, \tilde{\nu}^s) \leq C \overline{\mathcal{W}}_p^T(\nu^0, \tilde{\nu}^0),$$

for some constant  $C$  possibly depending on  $s$ .

*Démonstration.* This result is a direct consequence of an elementary estimate of SDE. More precisely, for any  $p \geq 1$ , by Grönwall inequality we have

$$\mathbb{E}[|\delta\Theta_t^s|^p] \leq C \left( |\delta\Theta_t^0|^p + \int_0^s [|\delta X_t^r|^p + |\delta P_t^r|^p] dr \right).$$

Using the estimates on  $\delta X_t^r$  and  $\delta P_t^r$  in the proof of Theorem VI.3.2 and Grönwall inequality, we get the following inequality with some constant  $C$  :

$$\overline{\mathcal{W}}_p^T(\nu^s, \tilde{\nu}^s)^p \leq C \int_0^T |\delta\Theta_t^0|^p dt$$

Since the above inequality holds for all couplings  $(\Theta_t^0), (\tilde{\Theta}_t^0)$ , one can choose the coupling such that  $\int_0^T |\delta\Theta_t^0|^p dt = \overline{\mathcal{W}}_p^T(\nu^0, \tilde{\nu}^0)$ .  $\blacksquare$

**Lemma VI.5.2.** *Let Assumptions VI.3.1 and VI.3.3 hold true, and  $(\Theta_t^0)_{t \in [0, T]}$  satisfy (VI.3.15). Then we have*

$$\int_0^T \mathbb{E} \left[ \sup_{s \in [0, S]} |\Theta_t^s|^p \right] dt < \infty \quad \text{for any } S \in \mathbb{R}^+, \quad (\text{VI.5.30})$$

$$\text{as well as } \int_0^T \sup_{s \in \mathbb{R}^+} \mathbb{E} [|\Theta_t^s|^p] dt < \infty. \quad (\text{VI.5.31})$$

*Démonstration.* The result (VI.5.30) follows from the standard SDE estimate, so its proof is omitted.

By the Itô formula, we have

$$d|\Theta_t^s|^p = |\Theta_t^s|^{p-2} \left( -p\Theta_t^s \cdot \mathbb{E}[\nabla_a H(t, X_t^s, \Theta_t^s, P_t^s, Z)] + \frac{\sigma^2}{2} p(p+m-2) \right) ds + \sigma p |\Theta_t^s|^{p-2} \Theta_t^s \cdot dW_s.$$

Now recall the assumptions (VI.3.12) and (VI.3.14) on  $\nabla_a H$ . We obtain :

$$\begin{aligned} d|\Theta_t^s|^p &\leq p|\Theta_t^s|^{p-2} \left( C - \epsilon |\Theta_t^s|^2 1_{\{|\Theta_t^s| \geq M\}} \right) ds + \sigma p |\Theta_t^s|^{p-2} \Theta_t^s \cdot dW_s \quad \text{for some } M > 0, \\ &\leq p|\Theta_t^s|^{p-2} \left( (C + \epsilon M^2) - \epsilon |\Theta_t^s|^2 \right) ds + \sigma p |\Theta_t^s|^{p-2} \Theta_t^s \cdot dW_s, \end{aligned}$$

where  $C$  does not depend on  $t$ . In the case  $p = 2$ , it clearly leads to  $\sup_{s \in \mathbb{R}^+} \mathbb{E}[|\Theta_t^s|^2] \leq C(1 + \mathbb{E}[|\Theta_t^0|^2])$ , due to the Gronwall inequality. Then (VI.5.31) follows. For general  $p > 2$ , the result (VI.5.31) is due to a simple induction.  $\blacksquare$

**Lemma VI.5.3.** *Under the assumptions of Lemma VI.5.2, for each  $t \in [0, T]$  the process  $(X_t^s)_{s \in \mathbb{R}^+}$  in the system (VI.2.10) is Lipschitz continuous in  $s$ .*

*Démonstration.* Let  $s' > s \geq 0$ . Since the function  $\phi$  is Lipschitz continuous in  $(x, a)$ , we have

$$\begin{aligned} |\delta X_t| := |X_t^{s'} - X_t^s| &= \left| \int_0^t \left( \mathbb{E}[\phi(r, X_r^{s'}, \Theta_r^{s'}, Z)] - \mathbb{E}[\phi(r, X_r^s, \Theta_r^s, Z)] \right) dr \right| \\ &\leq \int_0^t \left( C|\delta X_r| + \left| \mathbb{E}[\phi(r, X_r^{s'}, \Theta_r^{s'}, Z) - \phi(r, X_r^s, \Theta_r^s, Z)] \right| \right) dr. \end{aligned}$$

Further, by the Itô formula, we have

$$\begin{aligned} &\left| \mathbb{E}[\phi(r, X_r^{s'}, \Theta_r^{s'}, Z) - \phi(r, X_r^s, \Theta_r^s, Z)] \right| \\ &= \left| \mathbb{E} \left[ \int_s^{s'} \left( -\nabla_a \phi(r, X_r^u, \Theta_r^u, Z) \cdot \mathbb{E}[\nabla_a H(r, X_r^u, \Theta_r^u, P_r^u, Z)] + \frac{\sigma^2}{2} \Delta_{aa} \phi(r, X_r^u, \Theta_r^u, Z) \right) du \right] \right| \\ &\leq C(s' - s) \left( 1 + \sup_{u \in [s, s']} \mathbb{E}[|\Theta_r^u|] \right). \end{aligned}$$

The last inequality is due to the boundedness of  $\nabla_a \phi$ ,  $\Delta_{aa} \phi$  and the uniform linear growth of  $\nabla_a H$  in  $a$ . Finally, it follows from Lemma VI.5.2 and the Gronwall inequality that  $|\delta X_t| \leq C(s' - s)$ .  $\blacksquare$

Given a solution to the system of mean-field Langevin equations (VI.2.10), define

$$b^t(s, a) := -\mathbb{E}[\nabla_a H(t, X_t^s, a, P_t^s, Z)]. \quad (\text{VI.5.32})$$

It is easy to verify that under Assumptions VI.3.1 and VI.3.3, the function  $b^t$  is continuous in  $(s, a)$  and  $C^3$  in  $a$  for all  $t \in [0, T]$ . Due to a classical regularity result in the theory of linear PDEs (see e.g. [93, p.14-15]), we obtain the following result.

**Lemma VI.5.4.** *Let Assumptions VI.3.1 and VI.3.3 hold true. The marginal laws  $(\nu_t^s)$  of the solution to (VI.2.10) are weakly continuous solutions to the Fokker-Planck equations :*

$$\partial_s \nu = \nabla_a \cdot (-b^t \nu + \frac{\sigma^2}{2} \nabla_a \nu) \quad \text{for } t \in [0, T]. \quad (\text{VI.5.33})$$

*In particular, we have that  $(s, a) \mapsto \nu_t^s(a)$  belongs to  $C^{1,2}((0, \infty) \times \mathbb{R}^m)$ .*

## VI.5.2 Gradient Flow

We start by providing an estimate of the value  $\nabla_a \ln(\nu_t^s)$ . First, the following result ensures that  $\ln(\nu_t^s)$  is well defined.

**Lemma VI.5.5.** *Let Assumption VI.3.1 and Assumption VI.3.3 hold true and  $\mathbb{E}[|\Theta_t^0|^2] < \infty$  for some  $t \in [0, T]$ . Denote by  $\mathbb{Q}_t^\sigma$  the scaled Wiener measure<sup>1</sup> with initial distribution  $\nu_t^0$  and by  $(\mathcal{F}_s)_{s \in \mathbb{R}^+}$  the canonical filtration of the Wiener space. Then*

- i) for any finite horizon  $S > 0$ , the law of the solution to (VI.2.10),  $\nu_t := \text{Law}((\Theta_t^s)_{s \in \mathbb{R}^+})$ , is equivalent to  $\mathbb{Q}_t^\sigma$  on  $\mathcal{F}_S$  and the relative entropy*

$$\int \ln \left( \frac{d\nu_t}{d\mathbb{Q}_t^\sigma} \Big|_{\mathcal{F}_S} \right) d\nu_t = \mathbb{E} \left[ \int_0^S |b^t(s, \Theta_t^s)|^2 ds \right] < +\infty. \quad (\text{VI.5.34})$$

- ii) the marginal law  $\nu_t^s$  admits a density such that  $\nu_t^s > 0$  and  $\text{Ent}(\nu_t^s) < +\infty$ .*

The proof of these results is based on the Girsanov theorem and some simple moment estimates. It is similar to the proof of Lemma 6.1 in [81] and thus omitted. Further we have the following regularity result.

**Lemma VI.5.6.** *For  $t \in [0, T]$  and  $(\nu_t^s)_{s \in \mathbb{R}^+}$  the marginal laws of the solution to (VI.2.10), under the same assumptions as in Lemma VI.5.5, we have*

$$\nabla_a \ln(\nu_t^s(a)) = -\frac{1}{s_0} \mathbb{E} \left[ \int_0^{s_0} (1 - r \nabla_a b^t(r, \Theta_t^{s-s_0+r})) dW_r^{s-s_0} \Big| \Theta_t^s = a \right] \quad \text{for } s_0 \in (0, s], \quad (\text{VI.5.35})$$

*where  $W_r^{s-s_0} := W_{s-s_0+r} - W_{s-s_0}$ . In particular, for any  $s > 0$  we have*

$$C := \sup_{r \in [s, \infty)} \int_{\mathbb{R}^m} |\nabla_a \ln(\nu_t^r)|^2 \nu_t^r(a) da < +\infty,$$

*and  $C$  only depends on the Lipschitz constant of  $\nabla_a H$  with respect to  $a$ .*

1. Let  $B$  be the canonical process of the Wiener space and  $\mathbb{Q}$  be the Wiener measure, then the scaled Wiener measure  $\mathbb{Q}^\sigma := \mathbb{Q} \circ (\sigma B)^{-1}$ .

*Démonstration.* The equality (VI.5.35) is shown in Lemma 6.2 in [81]. The proof is based on Lemma 10.2 of the same paper and [58, Theorem 4.7 & Remark 4.13]. Further, we have for all  $s' \geq s$  :

$$\sup_{a \in \mathbb{R}^m} |\nabla_a \ln(\nu_t^{s'}(a))|^2 \leq \inf_{s_0 \in [s, s']} \frac{1}{s_0^2} \mathbb{E} \left[ \int_0^{s_0} |1 - r \nabla_a b^t(r, \Theta_t^{s'-s_0+r})|^2 dr \right].$$

Finally it is enough to note that  $\nabla_a b^t$  is bounded under the assumptions of the present Lemma. ■

**Lemma VI.5.7.** *Let Assumption VI.3.1 and Assumption VI.3.3 hold true and  $\mathbb{E}[|\Theta_t^0|^2] < \infty$ . We have*

$$\begin{aligned} \int_{\mathbb{R}^m} |\nabla_a \nu_t^s(a)| da < +\infty, \quad \int_{\mathbb{R}^m} |a \cdot \nabla_a \nu_t^s(a)| da < +\infty \quad \text{for all } s > 0, \\ \text{and } \int_s^{s'} \int_{\mathbb{R}^m} |\Delta_{aa} \nu_t^r(a)| da dr < +\infty \quad \text{for all } s' > s > 0. \end{aligned}$$

*Démonstration.* By the Young inequality, we have

$$|\nabla_a \nu_t^s(a)| \leq \frac{1}{2} \nu_t^s(a) + \frac{1}{2} \left| \frac{\nabla_a \nu_t^s(a)}{\nu_t^s(a)} \right|^2 \nu_t^s(a) \quad \text{and} \quad |a \cdot \nabla_a \nu_t^s(a)| \leq \frac{1}{2} |a|^2 \nu_t^s(a) + \frac{1}{2} \left| \frac{\nabla_a \nu_t^s(a)}{\nu_t^s(a)} \right|^2 \nu_t^s(a).$$

Since all the terms on the right hand sides are integrable, due to Lemma VI.5.6, therefore so are  $\nabla_a \nu_t^s$  and  $a \cdot \nabla_a \nu_t^s$ . Next, in order to prove the integrability of  $\Delta_{aa} \nu_t^s$ , we apply Itô's formula :

$$d \ln(\nu_t^s(\Theta_t^s)) = \left( \frac{\partial_s \nu_t^s(\Theta_t^s)}{\nu_t^s(\Theta_t^s)} + \frac{\nabla_a \nu_t^s(\Theta_t^s)}{\nu_t^s(\Theta_t^s)} \cdot b^t(s, \Theta_t^s) + \frac{\sigma^2}{2} \Delta_{aa}(\ln(\nu_t^s(\Theta_t^s))) \right) ds + \sigma \frac{\nabla_a \nu_t^s(\Theta_t^s)}{\nu_t^s(\Theta_t^s)} dW_s.$$

Together with the Fokker-Planck equation (VI.5.33), we have

$$d \ln(\nu_t^s(\Theta_t^s)) = \left( \sigma^2 \frac{\Delta_{aa} \nu_t^s(\Theta_t^s)}{\nu_t^s(\Theta_t^s)} - \nabla_a \cdot b^t(s, \Theta_t^s) - \frac{\sigma^2}{2} \frac{|\nabla_a \nu_t^s(\Theta_t^s)|^2}{|\nu_t^s(\Theta_t^s)|^2} \right) ds + \sigma \frac{\nabla_a \nu_t^s(\Theta_t^s)}{\nu_t^s(\Theta_t^s)} dW_s. \quad (\text{VI.5.36})$$

By Lemma VI.5.6, we have  $\mathbb{E} \left[ \int_s^{s'} \frac{\nabla_a \nu_t^r(\Theta_t^r)}{\nu_t^r(\Theta_t^r)} dW_r \right] = 0$ . Also recall that  $\nabla_a \cdot b^t(s, \Theta_t^s)$  is bounded. Taking expectation on both sides of (VI.5.36), we obtain

$$\begin{aligned} \sigma^2 \int_s^{s'} \int_{\mathbb{R}^m} |\Delta_{aa} \nu_t^r(a)| da dr &= \mathbb{E} \left[ \int_s^{s'} \sigma^2 \frac{\Delta_{aa} \nu_t^r(\Theta_t^r)}{\nu_t^r(\Theta_t^r)} dr \right] \\ &\leq \text{Ent}(\nu_t^{s'}) - \text{Ent}(\nu_t^s) + C \mathbb{E} \left[ \int_s^{s'} \left( 1 + \frac{|\nabla_a \nu_t^r(\Theta_t^r)|^2}{\nu_t^r(\Theta_t^r)} \right) ds \right]. \end{aligned}$$

By Lemma VI.5.5 and VI.5.6, the right hand side is finite. ■

Based on the previous integrability results, the next lemma follows from the integration by parts.

**Lemma VI.5.8.** *Under the assumptions of Lemma VI.5.7 we have for  $s > 0$*

$$\begin{aligned} \int_{\mathbb{R}^m} \Delta_{aa} H(t, X_t^s, a, P_t^s, Z) \nu_t^s(a) da &= - \int_{\mathbb{R}^m} \nabla_a H(t, X_t^s, a, P_t^s, Z) \cdot \nabla_a \nu_t^s(a) da \quad \text{for all } s > 0, \\ \int_s^{s'} \int_{\mathbb{R}^m} \Delta_{aa} \left( \ln \nu_t^r(a) \right) \nu_t^r(a) da dr &= - \int_s^{s'} \int_{\mathbb{R}^m} |\nabla_a \ln \nu_t^r(a)|^2 \nu_t^r(a) da dr \quad \text{for all } s' > s > 0. \end{aligned}$$

**Proof of Theorem VI.3.4.** It follows from Lemma VI.5.3 that there exists a bounded process  $(U_t^s)$  such that  $dX_t^s = U_t^s ds$ . On the other hand, note that  $X_t^s = X_0 + \int_0^t \mathbb{E}[\phi(r, X_r^s, \Theta_r^s, Z)] dr$ . By the Itô formula, we get

$$\frac{dX_t^s}{ds} = \int_0^t \int_{\mathbb{R}^m} \left( \nabla_a \phi(r, X_r^s, a, Z) b^r(s, X_r^s) + \nabla_x \phi(r, X_r^s, a, Z) U_r^s + \frac{\sigma^2}{2} \Delta_{aa} \phi(r, X_r^s, a, Z) \right) \nu_r^s(da) dr,$$

where  $b^r$ , defined as in (VI.5.32), is the drift term of the diffusion  $(\Theta_r^s)_{s \in \mathbb{R}^+}$ . In particular, we have

$$\frac{dU_t^s}{dt} = \int_{\mathbb{R}^m} \left( \nabla_a \phi(t, X_t^s, a, Z) b^t(s, X_t^s) + \nabla_x \phi(t, X_t^s, a, Z) U_t^s + \frac{\sigma^2}{2} \Delta_{aa} \phi(t, X_t^s, a, Z) \right) \nu_t^s(da) \quad (\text{VI.5.37})$$

Then note that  $V(\nu^s) = \mathbb{E} \left[ \int_0^T \mathbb{E} [L(t, X_t^s, \Theta_t^s, Z)] dt + G(X_T^s, Z) \right]$ . Again by the Itô formula, we have

$$\begin{aligned} \frac{dV(\nu^s)}{ds} &= \mathbb{E} \left[ \int_0^T \int_{\mathbb{R}^m} \left( \nabla_a L(t, X_t^s, a, Z) \cdot b^t(s, a) + \frac{\sigma^2}{2} \Delta_{aa} L(t, X_t^s, a, Z) \right. \right. \\ &\quad \left. \left. + \nabla_x L(t, X_t^s, a, Z) \cdot U_t^s \right) \nu_t^s(da) dt + \nabla_x G(X_T^s, Z) \cdot U_T^s \right]. \quad (\text{VI.5.38}) \end{aligned}$$

Recall (VI.5.37) and the dynamic of  $(P_t^s)_{t \in [0, T]}$  in (VI.2.10). By integration by parts, we have

$$\begin{aligned} \nabla_x G(X_T^s, Z) \cdot U_T^s &= \int_0^T \int_{\mathbb{R}^m} \left( -U_t^s \cdot \nabla_x H(t, X_t^s, a, P_t^s, Z) + P_t^s \cdot \nabla_x \phi(t, X_t^s, a, Z) U_t^s \right. \\ &\quad \left. + P_t^s \cdot \nabla_a \phi(t, X_t^s, a, Z) b^t(s, a) + P_t^s \cdot \frac{\sigma^2}{2} \Delta_{aa} \phi(t, X_t^s, a, Z) \right) \nu_t^s(da) dt \\ &= \int_0^T \int_{\mathbb{R}^m} \left( -U_t^s \cdot \nabla_x L(t, X_t^s, a, Z) \right. \\ &\quad \left. + P_t^s \cdot \nabla_a \phi(t, X_t^s, a, Z) b^t(s, a) + P_t^s \cdot \frac{\sigma^2}{2} \Delta_{aa} \phi(t, X_t^s, a, Z) \right) \nu_t^s(da) dt. \end{aligned}$$

Together with (VI.5.38), we obtain

$$\begin{aligned} \frac{dV(\nu^s)}{ds} &= \mathbb{E} \left[ \int_0^T \int_{\mathbb{R}^m} \left( b^t(s, a) \cdot \nabla_a H(t, X_t^s, a, P_t^s, Z) + \frac{\sigma^2}{2} \Delta_{aa} H(t, X_t^s, a, P_t^s, Z) \right) \nu_t^s(da) dt \right] \\ &= \int_0^T \int_{\mathbb{R}^m} \left( -|b^t(s, a)|^2 + \mathbb{E} \left[ \frac{\sigma^2}{2} \Delta_{aa} H(t, X_t^s, a, P_t^s, Z) \right] \right) \nu_t^s(da) dt. \end{aligned}$$

Further by Lemma VI.5.8, we have for  $s > 0$

$$\frac{dV(\nu^s)}{ds} = \int_0^T \int_{\mathbb{R}^m} \left( -|b^t(s, a)|^2 + \frac{\sigma^2}{2} b^t(s, a) \cdot \nabla_a \ln \nu_t^s(a) \right) \nu_t^s(da) dt. \quad (\text{VI.5.39})$$

On the other hand, recall formula (VI.5.36). By taking expectation on both sides and applying Lemma VI.5.8, we obtain for any  $s > 0$  :

$$\frac{d\left(\frac{\sigma^2}{2} \int_0^T \text{Ent}(\nu_t^s) dt\right)}{ds} = \int_0^T \int_{\mathbb{R}^m} \left( \frac{\sigma^2}{2} \nabla_a \ln \nu_t^s(a) \cdot b^t(s, a) - \frac{\sigma^4}{4} |\nabla_a \ln \nu_t^s(a)|^2 \right) \nu_t^s(da) dt. \quad (\text{VI.5.40})$$

Summing up (VI.5.39) and (VI.5.40), we finally obtain (VI.3.16). ■

## VI.6 Proof for the Convex Case

### VI.6.1 Sufficient First Order Condition

We are going to apply a standard variational calculus argument in order to derive the sufficient condition for being the optimal control of (VI.2.9).

**Proof of Theorem VI.3.7.** Take a  $\nu \in \mathcal{V}$  such that  $\nu$  is absolutely continuous with respect to Lebesgue measure (otherwise  $\int_0^T \text{Ent}(\nu_t) dt = +\infty$ ), and thus absolutely continuous with respect to the measure  $\nu^*$ . Denote  $X^*$  and  $X$  the controlled processes with  $\nu^*$  and  $\nu$ , respectively, and define  $\delta X := X - X^*$  and  $\delta \nu := \nu - \nu^*$ . By the assumption on convexity of the coefficients, we have

$$\begin{aligned} \delta V &:= \mathbb{E} \left[ \int_0^T \int_{\mathbb{R}^m} \left( L(t, X_t, a, Z) \nu_t(da) - L(t, X_t^*, a, Z) \nu_t^*(da) \right) dt + G(X_T, Z) - G(X_T^*, Z) \right] \\ &\geq \mathbb{E} \left[ \int_0^T \left( \nabla_x \ell(t, X_t^*, Z) \cdot \delta X_t + \int_{\mathbb{R}^m} L(t, X_t^*, a, Z) \delta \nu_t(da) \right) dt + \nabla_x G(X_T^*, Z) \cdot \delta X_T \right] \end{aligned} \quad (\text{VI.6.41})$$

Recall the adjoint process  $P^*$  defined in (VI.3.19). By integration by parts, we have

$$\nabla_x G(X_T^*) \cdot \delta X_T = P_T^* \cdot \delta X_T = \int_0^T \left( \int_{\mathbb{R}^m} P_s^* \cdot \phi(s, a) \delta \nu_s(da) - \nabla_x \ell(s, X_s^*, Z) \cdot \delta X_s \right) ds.$$

Together with (VI.6.41), it leads to

$$\delta V \geq \mathbb{E} \left[ \int_0^T \int_{\mathbb{R}^m} \left( P_t^* \cdot \phi(t, a) + L(t, X_t^*, a, Z) \right) \delta \nu_t(da) dt \right].$$

Further, we are going to compute the difference of the relative entropies. Since  $\nu$  is absolutely continuous with respect to  $\nu^*$ , we may define the Radon-Nikodym derivative  $f_t := \frac{\nu_t}{\nu_t^*}$ . Denote  $h(x) = x \ln(x)$  and

note that  $h(x) \geq x - 1$  for all  $x \in \mathbb{R}^+$ . We have

$$\begin{aligned} \text{Ent}(\nu_t) - \text{Ent}(\nu_t^*) &= \int_{\mathbb{R}^d} (\nu_t \ln \nu_t - \nu_t^* \ln \nu_t^*) dx = \int_{\mathbb{R}^d} (\nu_t - \nu_t^*) \ln \nu_t^* dx + \int_{\mathbb{R}^d} \nu_t (\ln \nu_t - \ln \nu_t^*) dx \\ &= \int_{\mathbb{R}^d} (f_t - 1) \nu_t^* \ln \nu_t^* dx + \int_{\mathbb{R}^d} h(f_t) \nu_t^* dx \geq \int_{\mathbb{R}^d} (f_t - 1) \nu_t^* \ln \nu_t^* dx + \int_{\mathbb{R}^d} (f_t - 1) \nu_t^* dx = \int_{\mathbb{R}^m} \ln(\nu_t^*(a)) \delta \nu_t(da). \end{aligned}$$

The last equality is due to  $\int_{\mathbb{R}^d} (f_t - 1) \nu_t^* dx = \int_{\mathbb{R}^d} (\nu_t - \nu_t^*) dx = 0$ . Finally, by (VI.3.18) we have

$$\begin{aligned} V^\sigma(\nu) - V^\sigma(\nu^*) &\geq \mathbb{E} \left[ \int_0^T \int_{\mathbb{R}^m} \left( P_t^* \cdot \phi(t, a) + L(t, X_t^*, a, Z) + \frac{\sigma^2}{2} \ln(\nu_t^*(a)) \right) \delta \nu_t(da) dt \right] \\ &= \mathbb{E} \left[ \int_0^T \int_{\mathbb{R}^m} \left( H(t, X_t^*, a, P_t^*, Z) + \frac{\sigma^2}{2} \ln(\nu_t^*(a)) \right) \delta \nu_t(da) dt \right] = 0. \end{aligned}$$

■

## VI.6.2 Convergence Towards the Invariant Measure

In order to prove that there exists an invariant measure of (VI.2.10) equal to the minimizer of  $V^\sigma$ , we follow the same strategy as in [81]. For readers' convenience, we shall provide a brief proof. The main ingredients of the proof are LaSalle's invariance principle (see e.g. [74, Theorem 4.3.3]) and the HWI inequality (see [128, Theorem 3]). Let  $(\nu^s)_{s \in \mathbb{R}^+}$  be the flow of marginal laws of the solution of (VI.2.10), given an initial law  $\nu^0$ . Define a dynamic system  $\mathcal{S}(s)[\nu^0] := \nu^s$ . We shall consider the following  $\omega$ -limit set :

$$\omega(\nu^0) := \left\{ \nu \in \mathcal{V} : \text{there exists } s_n \rightarrow +\infty \text{ such that } \mathcal{W}_2^T(\mathcal{S}(s_n)[\nu^0], \nu) \rightarrow 0 \right\}.$$

**Proposition VI.6.1** (Invariance Principle). *Assume that Assumption VI.3.1 and Assumption VI.3.3 hold true and  $\nu^0$  satisfies (VI.3.15) for some  $p > 2$ . Then the set  $\omega(\nu^0)$  is non-empty, compact and invariant, that is*

- i) for any  $\nu \in \omega(\nu^0)$ , we have  $\mathcal{S}(s)[\nu] \in \omega(\nu^0)$  for all  $s \in \mathbb{R}^+$ ;
- ii) for any  $\nu \in \omega(\nu^0)$  and all  $s \in \mathbb{R}^+$ , there exists  $\nu' \in \omega(\nu^0)$  such that  $\mathcal{S}(s)[\nu'] = \nu$ .

*Démonstration.* It is important to note that

- the mapping  $\nu^0 \mapsto \mathcal{S}(s)[\nu^0]$  is  $\mathcal{W}_2^T$ -continuous, due to Lemma VI.5.1;
- the mapping  $s \mapsto \mathcal{S}(s)[\nu^0]$  belongs to  $C_2(\mathbb{R}^+, \mathcal{V})$ , due to Theorem VI.3.2;
- the set  $\{\mathcal{S}(s)[\nu^0], s \in \mathbb{R}^+\}$  belongs to a  $\mathcal{W}_2^T$ -compact set, due to Lemma VI.5.2.

The rest follows the standard argument for LaSalle's invariance principle (see e.g. [74, Theorem 4.3.3] or [81, Proposition 6.5]).

■



**Proof of Theorem VI.3.8.** As in the Step 1 of the proof to [81, Theorem 2.10], using the invariance principle we can prove the existence of a convergent subsequence of the measure flow  $(\nu^{s_n})_{n \in \mathbb{N}}$  such that  $\mathcal{W}_2^T(\nu^{s_n}, \nu^*) \rightarrow 0$ , where  $\nu^* = \arg \min_{\nu} V^\sigma(\nu)$  and satisfies

$$\nu^*(t, a) = C \exp\left(-\frac{2}{\sigma^2} H(t, X_t^*, a, P_t^*, Z)\right).$$

In particular,  $\nu^*$  is log-semiconcave, because one may easily verify that the gradient of the mapping  $(t, a) \mapsto H(t, X_t^*, a, P_t^*, Z)$  is Lipschitz continuous. By the HWI inequality we have

$$\int (\ln \nu^{s_n} - \ln \nu^*) \nu^{s_n}(dt, da) \leq T^{-\frac{1}{2}} \mathcal{W}_2^T(\nu^{s_n}, \nu^*) \left( \sqrt{\mathcal{I}_n/T} + CT^{-\frac{1}{2}} \mathcal{W}_2^T(\nu^{s_n}, \nu^*) \right),$$

where  $\mathcal{I}_n$  is the relative Fisher information defined as

$$\begin{aligned} \mathcal{I}_n &:= \int \left| \nabla_a \ln \nu^{s_n} - \nabla_a \ln \nu^* \right|^2 \nu^{s_n}(dt, da) \\ &= \int \left| \nabla_a \ln \nu^{s_n} + \frac{2}{\sigma^2} \nabla_a H(t, X_t^*, a, P_t^*, Z) \right|^2 \nu^{s_n}(dt, da) \\ &\leq 2 \int \left| \nabla_a \ln \nu^{s_n} \right|^2 \nu^{s_n}(dt, da) + C \left( 1 + \int |a|^2 \nu^{s_n}(dt, da) \right). \end{aligned}$$

Then it follows from Lemma VI.5.6 and VI.5.2 that  $\sup_n \mathcal{I}_n < \infty$ . Together with the fact that  $\mathcal{W}_2^T(\nu^{s_n}, \nu^*) \rightarrow 0$ , we have

$$\limsup_{n \rightarrow \infty} \text{Ent}(\nu^{s_n}) - \text{Ent}(\nu^*) = \limsup_{n \rightarrow \infty} \int (\ln \nu^{s_n} - \ln \nu^*) \nu^{s_n}(dt, da) \leq 0.$$

Since  $\text{Ent}$  is  $\mathcal{W}_2^T$ -lower-semicontinuous, we have  $\lim_{n \rightarrow \infty} \text{Ent}(\nu^{s_n}) = \text{Ent}(\nu^*)$ , and thus  $\lim_{s \rightarrow \infty} V^\sigma(\nu^s) = V^\sigma(\nu^*)$ . We have proved that  $V^\sigma$  is a continuous Lyapunov function along the trajectory of  $(\nu^s)$ . Further we can conclude the proof using the standard argument (see e.g. Step 3 of the proof of Theorem 2.10 [81]). ■

## VI.7 Proofs for the Contraction Case

We first provide a sufficient condition for the regularized control problem (VI.2.9) to have at least one optimal control.

**Lemma VI.7.1.** *Assume that there exists  $\nu \in \mathcal{V}$  such that  $V^\sigma(\nu) < \infty$ , and that there is a function  $U : \mathbb{R}^m \rightarrow \mathbb{R}^+$  such that  $\int_{\mathbb{R}^m} e^{-U(a)} da < \infty$  and  $\bar{V}(\nu) := V(\nu) - \frac{\sigma^2}{2} \int_0^T \int_{\mathbb{R}^m} U(a) \nu_t(da) dt$  is bounded from below and weakly lower-semicontinuous. Then  $\arg \min_{\nu} V^\sigma(\nu) \neq \emptyset$ .*

*Démonstration.* Let  $\bar{\nu} \in \mathcal{V}$  such that  $V^\sigma(\bar{\nu}) < \infty$ . Denote  $C_0 := \inf_{\nu} \bar{V}(\nu)$  and  $\bar{C} := V^\sigma(\bar{\nu}) - C_0$ .

Recall that

$$V^\sigma = \bar{V} + \frac{\sigma^2}{2} \int_0^T \left( \text{Ent}(\nu_t) + \int_{\mathbb{R}^m} U(a) \nu_t(da) \right) dt = \bar{V} + \frac{\sigma^2}{2} I(\nu),$$

where 
$$I(\nu) := \int_0^T \int_{\mathbb{R}^m} \ln \left( \frac{\nu_t(a)}{e^{-U(a)}} \right) \nu_t(da) dt.$$

Note that  $I(\nu)$  is the relative entropy of  $\nu$  with respect to the measure  $dt \times e^{-U(a)} da$ . Therefore,  $I$  is weakly lower-semicontinuous (so is  $V^\sigma$ ) and the sublevel set  $\mathcal{K} := \left\{ \nu \in \mathcal{V} : \frac{\sigma^2}{2} I(\nu) \leq \bar{C} \right\}$  is weakly compact (see e.g. [45, Lemma 1.4.3]). Further note that  $\left\{ \nu \in \mathcal{V} : V^\sigma(\nu) \leq V^\sigma(\bar{\nu}) \right\} \subset \mathcal{K}$ , so  $\inf_{\nu \in \mathcal{V}} V^\sigma(\nu) = \inf_{\nu \in \mathcal{K}} V^\sigma(\nu)$ . Since  $V^\sigma$  is weakly lower-semicontinuous and  $\mathcal{K}$  is weakly compact, there exists a global minimizer in  $\mathcal{K}$ .  $\blacksquare$

Next we prove the necessary condition of being an optimal control.

**Proof of Proposition VI.3.9.** Since  $V^\sigma(\nu^*) < \infty$ , we know that  $\int_0^T \text{Ent}(\nu_t^*) dt < \infty$ . In particular,  $\nu^*$  is absolutely continuous with respect to the Lebesgue measure.

*Step 1.* Let  $\nu \in \mathcal{V}$  be a measure such that  $\int_0^T \text{Ent}(\nu_t) dt < \infty$ , in particular, it is absolutely continuous with respect to the Lebesgue measure. Define  $\nu^\varepsilon := (1 - \varepsilon)\nu^* + \varepsilon\nu$  for  $\varepsilon > 0$ . By standard variational calculus we have

$$\lim_{\varepsilon \rightarrow 0} \frac{V(\nu^\varepsilon) - V(\nu^*)}{\varepsilon} = \int_0^T \int_{\mathbb{R}^m} \mathbb{E} \left[ H(t, X_t^*, a, P_t^*, Z) \right] (\nu(da) - \nu_t^*(da)) dt.$$

Further, define the function  $h(x) := x \ln x$ . We have

$$\frac{1}{\varepsilon} \int_0^T (\text{Ent}(\nu_t^\varepsilon) - \text{Ent}(\nu_t^*)) dt = \frac{1}{\varepsilon} \int_0^T \int_{\mathbb{R}^m} \left( h(\nu_t^\varepsilon(a)) - h(\nu_t^*(a)) \right) da dt$$

Since the function  $h$  is convex, we note that

$$\frac{1}{\varepsilon} \left( h(\nu_t^\varepsilon(a)) - h(\nu_t^*(a)) \right) \leq h(\nu_t(a)) - h(\nu_t^*(a)) \quad \text{for all } \varepsilon \in (0, 1).$$

The right hand side of the inequality above is integrable because both  $\int_0^T \text{Ent}(\nu_t) dt$  and  $\int_0^T \text{Ent}(\nu_t^*) dt$  are finite. Therefore, by Fatou's lemma we obtain

$$\begin{aligned} 0 &\leq \limsup_{\varepsilon \rightarrow 0} \frac{V^\sigma(\nu^\varepsilon) - V^\sigma(\nu^*)}{\varepsilon} \\ &\leq \int_0^T \int_{\mathbb{R}^m} \left( \mathbb{E} \left[ H(t, X_t^*, a, P_t^*, Z) \right] + \frac{\sigma^2}{2} \ln \left( \nu_t^*(a) \right) \right) (\nu_t - \nu_t^*)(da) dt. \end{aligned} \quad (\text{VI.7.42})$$

*Step 2.* We are going to show that for Leb-a.s.  $t$

$$\Xi_t(a) := \mathbb{E}\left[H(t, X_t^*, a, P_t^*, Z)\right] + \frac{\sigma^2}{2} \ln\left(\nu_t^*(a)\right) \text{ is equal to a constant } \nu_t^*\text{-a.s.} \quad (\text{VI.7.43})$$

Define the mean value  $\bar{c}_t := \int_{\mathbb{R}^m} \Xi_t(a) \nu_t^*(da)$  and let  $\epsilon, \epsilon' > 0$ . Consider the measure  $\nu \in \mathcal{V}$  absolutely continuous with respect to  $\nu^*$  such that  $\nu_t = \nu_t^*$  if  $\nu_t^*[\Xi_t \leq \bar{c}_t - \epsilon] < \epsilon'$ , otherwise

$$\frac{d\nu_t}{d\nu_t^*} = \frac{1_{\Xi_t \leq \bar{c}_t - \epsilon}}{\nu_t^*[\Xi_t \leq \bar{c}_t - \epsilon]}.$$

Note that  $\Xi_t \leq \bar{c}_t - \epsilon$ ,  $\nu_t$ -a.s. for  $t$  such that  $\nu_t^*[\Xi_t \leq \bar{c}_t - \epsilon] \geq \epsilon'$ . Then we have

$$\int_0^T \int_{\mathbb{R}^m} \Xi_t(a) (\nu_t(da) - \nu_t^*(da)) dt \leq -\epsilon \int_0^T 1_{\nu_t^*[\Xi_t \leq \bar{c}_t - \epsilon] \geq \epsilon'} dt.$$

Together with (VI.7.42), we conclude  $\nu_t^*[\Xi_t \leq \bar{c}_t - \epsilon] < \epsilon'$  for Leb-a.s.  $t \in [0, T]$ . Since this holds true for arbitrary  $\epsilon', \epsilon > 0$ , we obtain (VI.7.43).

*Step 3.* We are going to show that  $\nu^*$  is equivalent to the Lebesgue measure. First we provide an estimate for the constant  $\bar{c}_t$  above. Since  $\nu_t^*$  is a probability measure, we have

$$\int_{\mathbb{R}^m} \exp\left(\frac{2\left(\bar{c}_t - \mathbb{E}\left[H(t, X_t^*, a, P_t^*, Z)\right]\right)}{\sigma^2}\right) da = 1. \quad (\text{VI.7.44})$$

Moreover, since  $(t, z) \mapsto H(t, 0, 0, 0, z)$  is bounded and  $a \mapsto \nabla_a H(t, x, a, p, z)$  is uniformly Lipschitz continuous, we have

$$\sup_{t, z} \left|H(t, X_t^*, a, P_t^*, z)\right| \leq C(1 + |a|^2). \quad (\text{VI.7.45})$$

On the other hand, following the dissipative assumption (VI.3.14), we may easily prove that there are constants  $C, C' > 0$  such that for all  $(t, a)$

$$H(t, X_t^*, a, P_t^*, z) \geq -C + C'|a|^2. \quad (\text{VI.7.46})$$

Together with (VI.7.44) and (VI.7.45), we prove that  $(\bar{c}_t)_{t \in [0, T]}$  is bounded.

Now suppose that  $\nu^*$  is not equivalent to the Lebesgue measure. Then there is a set  $\mathcal{K} \in [0, T] \times \mathbb{R}^m$  such that  $\nu^*(\mathcal{K}) = 0$  (so  $\ln \nu^* = -\infty$  on  $\mathcal{K}$ ) and  $\text{Leb}[\mathcal{K}] > 0$ . It follows from (VI.7.42) that  $0 \leq C - \int_{\mathcal{K}} \infty d\nu$ . Since we may choose  $\nu$  having positive mass on  $\mathcal{K}$ , it is a contradiction. Therefore  $\nu^*$  must be equivalent to the Lebesgue measure.

*Step 4.* Since  $\nu^*$  is equivalent to the Lebesgue measure, together with (VI.7.43) we obtain (VI.3.20), and thus  $\nu_t^*$  is a stationary solution to the Fokker-Planck equation (VI.5.33) for Leb-a.s.  $t \in [0, T]$ . Therefore  $\nu^*$  is an invariant measure of the mean-field Langevin system (VI.2.10). Finally, the bounds in (VI.3.21) are obtained from the estimates (VI.7.45) and (VI.7.46). ■

In order to prove the main Theorem VI.3.12, the main ingredient is the reflection coupling in Eberle [48]. For the mean-field system, we shall adopt the reflection-synchronous coupling similar to [49].

We fix a parameter  $\epsilon > 0$ . Introduce the Lipschitz functions  $\text{rc} : \mathbb{R}^m \times \mathbb{R}^m \rightarrow [0, 1]$  and  $\text{sc} : \mathbb{R}^m \times \mathbb{R}^m \rightarrow [0, 1]$  satisfying

$$\text{sc}^2(x, y) + \text{rc}^2(x, y) = 1, \quad \text{rc}(x, y) = 1 \text{ for } |x - y| \geq \epsilon, \quad \text{rc}(x, y) = 0 \text{ for } |x - y| \leq \epsilon/2.$$

Let  $(\nu_t^0)_{t \in [0, T]}$  and  $(\tilde{\nu}_t^0)_{t \in [0, T]}$  be two initial measures, and  $(W_s^1), (W_s^2)$  be two independent Brownian motions. For the given  $(\nu_t^0), (\tilde{\nu}_t^0)$  we construct the drift coefficients  $(b^t), (\tilde{b}^t)$  as in (VI.5.32), respectively. Further, for a fixed<sup>2</sup>  $t \in [0, T]$ , define the coupling  $\Sigma_t = (\Theta_t, \tilde{\Theta}_t)$  as the solution to the standard SDE

$$\begin{aligned} d\Theta_t^s &= b^t(s, \Theta_t^s)ds + \text{rc}(\Sigma_t^s)\sigma dW_s^1 + \text{sc}(\Sigma_t^s)\sigma dW_s^2, \\ d\tilde{\Theta}_t^s &= \tilde{b}^t(s, \tilde{\Theta}_t^s)ds + \text{rc}(\Sigma_t^s)(\text{Id} - 2e_s \langle e_s, \cdot \rangle)\sigma dW_s^1 + \text{sc}(\Sigma_t^s)\sigma dW_s^2, \end{aligned}$$

where  $e_s := \frac{\Theta_t^s - \tilde{\Theta}_t^s}{|\Theta_t^s - \tilde{\Theta}_t^s|}$  for  $\Theta_t^s \neq \tilde{\Theta}_t^s$ , otherwise  $e_s := \hat{e}$  some arbitrary fixed unit vector in  $\mathbb{R}^m$ . Next, we construct a concave increasing function  $f$  as in the proof of [49, Theorem 2.3]. Let

$$f(r) := \int_0^r \varphi(s)g(s \wedge R_2)ds, \quad \text{where} \quad g(r) := 1 - \frac{c}{2} \int_0^r \Phi(s)\varphi(s)^{-1}ds,$$

and the function  $\varphi$  and the constant  $R_2$  are defined as in the statement of Theorem VI.3.12. Note that by definition  $\kappa^+(r) = 0$  for any  $r \geq R_1$ , so  $(\varphi(r))_{r \geq R_1}$  is a constant and the function  $f$  is linear on  $[R_2, \infty)$ . Furthermore,  $f$  is twice continuously differentiable on  $(0, R_2) \cup (R_2, +\infty)$ . Clearly for  $r \in (0, R_2)$  we have

$$2\sigma^2 f''(r) = -r\kappa^+(r)f'(r) - c\sigma^2\Phi(r) \leq -r\kappa(r)f'(r) - c\sigma^2 f(r) \quad (\text{VI.7.47})$$

and for  $r \in \mathbb{R}^+$

$$r\varphi(R_1) \leq \Phi(r) \leq 2f(r) \leq 2\Phi(r) \leq 2r. \quad (\text{VI.7.48})$$

Next we prove the following inequality similar to (VI.7.47) on  $(R_2, \infty)$  where  $f'' = 0 : 0 \leq -r\kappa(r)f'(r) - c\sigma^2 f(r)$ . Recall that  $(\varphi(r))_{r \geq R_1}$  is a constant and thus we have  $\Phi(r) = \Phi(R_1) + \varphi(R_1)(r - R_1)$ . Since

2. We are not defining the coupling for the system of SDE's, but for a single SDE with the fixed label  $t$ .

$\Phi(R_1) \geq \varphi(R_1)R_1$ , we have

$$\frac{\Phi(r)}{\Phi(R_2)} = \frac{\Phi(R_1) - \varphi(R_1)R_1 + \varphi(R_1)r}{\Phi(R_1) - \varphi(R_1)R_1 + \varphi(R_1)R_2} \leq \frac{r}{R_2}, \quad \text{for } r \geq R_2. \quad (\text{VI.7.49})$$

Furthermore, it is easy to verify that

$$c^{-1} \geq \int_{R_1}^{R_2} \Phi(s)\varphi^{-1}(s)ds \geq \frac{\varphi^{-1}(R_1)\Phi(R_2)(R_2 - R_1)}{2}. \quad (\text{VI.7.50})$$

Also note that  $g(R_2) = \frac{1}{2}$  due to the definition of  $c$ , and thus  $f'(r) = \frac{\varphi(R_1)}{2}$  for  $r \geq R_2$ . Together with (VI.7.49), (VI.7.50) and the definition of  $R_2$ , we have

$$r\kappa(r)f'(r) \leq -2\sigma^2 \frac{r\varphi(R_1)}{(R_2 - R_1)R_2} \leq -2\sigma^2 \frac{\Phi(r)\varphi(R_1)}{(R_2 - R_1)\Phi(R_2)} \leq -c\sigma^2\Phi(r) \leq -c\sigma^2 f(r).$$

**Proof of Theorem VI.3.12.** *Step 1.* We first use an argument similar to that of the proof of Theorem 2.3 in [49] to obtain some estimates concerning the coupling. Since  $\nu_t^0, \tilde{\nu}_t^0$  have finite 2-moment, we may choose the coupling  $(\Theta_t^0, \tilde{\Theta}_t^0)$  so that

$$\mathcal{W}_1(\nu_t^0, \tilde{\nu}_t^0) = \mathbb{E}[|\Theta_t^0 - \tilde{\Theta}_t^0|] \geq \mathbb{E}\left[f(|\Theta_t^0 - \tilde{\Theta}_t^0|)\right]. \quad (\text{VI.7.51})$$

The last inequality is due to (VI.7.48). On the other hand, for all  $s \geq 0$  we have

$$\mathcal{W}_1(\nu_t^s, \tilde{\nu}_t^s) \leq \mathbb{E}[|\Theta_t^s - \tilde{\Theta}_t^s|] \leq \frac{2}{\varphi(R_1)} \mathbb{E}\left[f(|\Theta_t^s - \tilde{\Theta}_t^s|)\right]. \quad (\text{VI.7.52})$$

Denote  $\delta\Theta_t^s := \Theta_t^s - \tilde{\Theta}_t^s$ . By the definition of the coupling above, we have

$$d\delta\Theta_t^s = \left(b^t(s, \Theta_t^s) - \tilde{b}^t(s, \tilde{\Theta}_t^s)\right)ds + 2\text{rc}(\Sigma_t^s)\sigma d\bar{W}_s,$$

where  $\bar{W}_s := \int_0^s e_r \cdot dW_r^1$  is a one-dimensional Brownian motion. Denote  $r_s := |\delta\Theta_t^s|$  and note that by the definition of  $\text{rc}$  we have  $\text{rc}(\Sigma_t^s) = 0$  whenever  $r_s \leq \epsilon/2$ . Therefore, one may show that

$$dr_s = e_s \cdot \left(b^t(s, \Theta_t^s) - \tilde{b}^t(s, \tilde{\Theta}_t^s)\right)ds + 2\text{rc}(\Sigma_t^s)\sigma d\bar{W}_s.$$

Define  $\mathcal{L}_s^x$  as the right-continuous local time of  $(r_s)$  and  $\mu_f$  as the nonpositive measure representing the second derivative of  $f$ . Then it follows from the Itô-Tanaka formula that

$$\begin{aligned} f(r_s) - f(r_0) &= \int_0^s f'(r_u)e_u \cdot \left(b^t(u, \Theta_t^u) - \tilde{b}^t(u, \tilde{\Theta}_t^u)\right)du + \frac{1}{2} \int_{\mathbb{R}} \mathcal{L}_s^x \mu_f(dx) + M_s, \\ &\leq \int_0^s \left(f'(r_u)e_u \cdot \left(b^t(u, \Theta_t^u) - \tilde{b}^t(u, \tilde{\Theta}_t^u)\right) + 2\text{rc}(\Sigma_t^u)^2 \sigma^2 f''(r_u)\right) du + M_s, \end{aligned}$$

where  $M_s := 2 \int_0^s \text{rc}(\Sigma_t^u) f'(r_u) \sigma d\bar{W}_u$  is a martingale, and the last inequality is due to the concavity of  $f$ . Now it is important to note that under the assumptions of the present Theorem we have

$$e_s \cdot \left( b^t(s, \Theta_t^s) - \tilde{b}^t(s, \tilde{\Theta}_t^s) \right) \leq 1_{\{r_s \geq \epsilon\}} r_s \kappa(r_s) + 1_{\{r_s < \epsilon\}} \gamma \epsilon + \gamma \bar{\mathcal{W}}_1^T(\nu^s, \tilde{\nu}^s),$$

where  $\gamma := K^2(1 + K) \exp(2KT)$  with  $K$  a common Lipschitz coefficient of  $\nabla_a H$ ,  $\nabla_x H$ ,  $\nabla_x G$  and  $\phi$ , which is computed explicitly in the proof of Theorem VI.3.2. Further, since  $f'' \leq 0$  and  $\text{rc}(\Sigma_t^s) = 1$  whenever  $r_s \geq \epsilon$ , we have

$$\begin{aligned} f(r_s) - f(r_0) &\leq \int_0^s \left( 1_{\{r_u \geq \epsilon\}} \left( f'(r_u) r_u \kappa(r_u) + 2\sigma^2 f''(r_u) \right) + 1_{\{r_u < \epsilon\}} \gamma \epsilon + \gamma \bar{\mathcal{W}}_1^T(\nu^u, \tilde{\nu}^u) \right) du + M_s \\ &\leq \int_0^s \left( -1_{\{r_u \geq \epsilon\}} c\sigma^2 f(r_u) + 1_{\{r_u < \epsilon\}} \gamma \epsilon + \gamma \bar{\mathcal{W}}_1^T(\nu^u, \tilde{\nu}^u) \right) du + M_s \\ &= \int_0^s \left( -c\sigma^2 f(r_u) + 1_{\{r_u < \epsilon\}} (\gamma + c\sigma^2) \epsilon + \gamma \bar{\mathcal{W}}_1^T(\nu^u, \tilde{\nu}^u) \right) du + M_s \end{aligned}$$

The last inequality is due to (VI.7.47). It clearly leads to

$$\mathbb{E} \left[ e^{c\sigma^2 s} f(r_s) - f(r_0) \right] \leq \int_0^s e^{c\sigma^2 u} \left( (\gamma + c\sigma^2) \epsilon + \gamma \bar{\mathcal{W}}_1^T(\nu^u, \tilde{\nu}^u) \right) du.$$

Recall (VI.7.51) and (VI.7.52). Together with the estimate above, we obtain

$$\frac{\varphi(R_1)}{2} e^{c\sigma^2 s} \mathcal{W}_1(\nu_t^s, \tilde{\nu}_t^s) - \mathcal{W}_1(\nu_t^0, \tilde{\nu}_t^0) \leq \int_0^s e^{c\sigma^2 u} \left( (\gamma + c\sigma^2) \epsilon + \gamma \bar{\mathcal{W}}_1^T(\nu^u, \tilde{\nu}^u) \right) du. \quad (\text{VI.7.53})$$

*Step 2.* Since for each  $t \in [0, T]$  one may obtain the estimate (VI.7.53) through the previous coupling argument, we have

$$\frac{\varphi(R_1)}{2} e^{c\sigma^2 s} \bar{\mathcal{W}}_1^T(\nu^s, \tilde{\nu}^s) - \bar{\mathcal{W}}_1^T(\nu^0, \tilde{\nu}^0) \leq T \int_0^s e^{c\sigma^2 u} \left( (\gamma + c\sigma^2) \epsilon + \gamma \bar{\mathcal{W}}_1^T(\nu^u, \tilde{\nu}^u) \right) du.$$

By the Gronwall inequality, we have

$$\bar{\mathcal{W}}_1^T(\nu^s, \tilde{\nu}^s) \leq e^{\left( \frac{2\gamma T}{\varphi(R_1)} - c\sigma^2 \right) s} \frac{2}{\varphi(R_1)} \bar{\mathcal{W}}_1^T(\nu^0, \tilde{\nu}^0) + \frac{2T}{\varphi(R_1)} e^{\frac{2\gamma T}{\varphi(R_1)} s} (\gamma + c\sigma^2) \epsilon \int_0^s e^{c\sigma^2 u} du.$$

This holds true for all  $\epsilon > 0$ , so finally we obtain (VI.3.22). ■

## Remerciements

*Ce dernier chapitre marquera la fin de cette thèse qui dura un peu plus de mille cent jours. Ces quelques lignes seront, j'en suis presque sûr, les plus lues de ce manuscrit. Non pas qu'elles soient plus compréhensibles et mieux écrites que le reste, mais le sujet traité intéresse beaucoup plus de personnes.*

*Évidemment, les plus profonds remerciements sont adressés à Nizar Touzi, mon directeur de thèse, que j'ai rencontré pour la première fois dans son cours de contrôle stochastique du master Probabilité et Finance. Le rigueur et la limpidité de son cours me marquèrent et dirigèrent vers une thèse dans ce domaine. Il fut extrêmement généreux de ses idées et de temps, n'hésitant pas à m'emmener en conférences, me mettant dans les meilleures conditions imaginables pour la réalisation de ma thèse.*

*Je remercie également mon confrère et coauteur de plusieurs papiers Zhenjie Ren, maître de conférence à l'Université Paris-Dauphine. J'apprends énormément des nombreuses heures d'études et de discussions durant ces années consacrées au divers sujets variant du Principal-Agent aux jeux à champ-moyen, du Pontryagin aux réseaux de neurones profonds.*

*C'est un grand honneur pour moi d'avoir Johannes Muhle-Karbe et Stéphane Villeneuve comme rapporteurs. Les exposés de Johannes Muhle-Karbe témoignent de la clarté de ses idées et de son intérêt pour la communication. Je remercie également Stéphane Villeneuve pour son excellent cours sur Principal-Agent à la 11<sup>ème</sup> édition de l'école d'été européenne de mathématiques financières ainsi que pour son invitation à présenter un exposé à TSE, l'occasion pour moi de retourner dans la ville qui m'accueillit pour mes classes préparatoires il y a 7 longs années. Un grand merci à Caroline Hillairet, Gilles Pagès, Jean-François Chassagneux et Zhenjie Ren d'avoir accepté de faire partie de mon jury de thèse.*

*Bien entendu ce manuscrit doit beaucoup à mes coauteurs David Šiška, Łukasz Szpruch et Junjian Yang, qui m'aidèrent à condenser mes papiers. Je tiens remercier ensuite ma famille de thèse Romuald Elie, Dylan Possamai, Xiaolu Tan, Gaoyue Guo, Hadrien De March, Heythem Farat, Bowen Sheng, Mehdi Talbi, qui m'accompagnèrent et m'apportèrent un soutien sans faille tout le long de ma thèse. Je remercie également Eduardo Abi Jaber, Ankush Agarwal, Yann Aït-Mokhtar, René Aid, Adrien Barrasso, Cyril Bénézet, Adel Cherchali, Julien Claisse, Giovanni Conforti, Gonçalo dos Reis, Omar El Euch, Pierre Henry-Labordère, Emma Hubert, Yiqing Lin, Gang Liu, Alexandre Richard, Ruoci Sun, Chao Zhou.*

*Dès 2016 je fus accueilli dans ce nid douillet qu'est le CMAP. Je voudrais remercier l'excellente équipe mathématiques financières du laboratoire, en particulier Emmanuel Gobet, Stefano De Marco, Thibaut Mastrolia, Mathieu Rosenbaum. Je ne peux oublier de remercier le personnel administratif toujours très efficace, en particulier Vincent Chapelier, Wilfried Edmond, Nasséra Naar, Alexandra Noiret. Je remercie aussi Thierry Bodineau, Anne de Bouard et Sylvie Méléard. N'oublions pas les doctorants du CMAP Bastien Baldacci, Rémi Besson, Mathilde Boissier, Céline Bonnet, Florian Bourgey, Nicolas Brosse, Juliette Chevallier, Sylvain Ferrand, Antoine Fosset, Antoine Havet, Fedor Goncharov, Paul Jusselin, Frédéric Logé-Munerel, Othmane Mounjid, Kevish Napal, Léa Nicolas, Geneviève Robin, Omar Saadi, Paul Thévenin.*

*Enfin un immense merci aux amis de promo, du master, de passage, à mes anciens professeurs, aux élèves sur qui j'aiguisé ma pédagogie. Je tiens ici à remercier en particulier Élie Casbi, le plus grand pianiste parmi tous les mathématiciens (que je connais), nous sommes nés le même jour, et entrés à l'École Normale Supérieure au même rang ; Sibó Cheng, qui abandonna la carrière de comédien pour les mathématiques grâce à son enthousiasme sans borne pour le filtre de Kalman ; Achille Samaran, le meilleur joueur de poker de France qui arrive en plus à me battre à chaque fois dans AGE OF EMPIRE III sans aucune pitié et Lucie Cheyer, grande inventrice de tartelette aux pêches version EARL GREY et grande spécialiste de l'art culinaire de C. SATIVA, qui arrive à réciter AINSI PARLAIT ZARATHOUSTRA presque sans aucune faute. Je remercie dernièrement ma maman et mon papa avec tout mon cœur, soutien inconditionnel et source de bonheur sans fin.*



---

# BIBLIOGRAPHIE

- [1] AÏD, R., POSSAMAÏ, D., AND TOUZI, N. Optimal electricity demand response contracting with responsiveness incentives. *preprint arXiv :1810.09063* (2018).
- [2] ALASSEUR, C., EKELAND, I., ELIE, R., HERNÁNDEZ SANTIBÁÑEZ, N., AND POSSAMAÏ, D. An adverse selection approach to power pricing. *preprint arXiv :1706.01934* (2017).
- [3] AMBROSIO, L., GIGLI, N., AND SAVARÉ, G. *Gradient flows : in metric spaces and in the space of probability measures*. Springer, 2008.
- [4] ANTONELLI, F. Backward-forward stochastic differential equations. *Ann. Appl. Probab.* 3, 3 (08 1993), 777–793.
- [5] ATHALYE, A., AND CARLINI, N. On the Robustness of the CVPR 2018 White-Box Adversarial Example Defenses. *arXiv e-prints* (Apr 2018), arXiv :1804.03286.
- [6] BAKRY, D., AND ÉMERY, M. Diffusions hypercontractives. In *Séminaire de Probabilités XIX 1983/84*. Springer, 1985, pp. 177–206.
- [7] BARRON, A. R. Universal approximation bounds for superpositions of a sigmoidal function. *IEEE Transactions on Information theory* 39, 3 (1993), 930–945.
- [8] BELKIN, M., RAKHLIN, A., AND TSYBAKOV, A. B. Does data interpolation contradict statistical optimality? *arXiv :1806.09471* (2018).
- [9] BENES, V. Existence of optimal stochastic control laws. *Society for Industrial and Applied Mathematics* (Mai 1970), 446–472.
- [10] BENSOUSSAN, A. *Stochastic Control of Partially Observable Systems*. Cambridge University Press, 1992.
- [11] BIAIS, B., MARIOTTI, T., ROCHET, J.-C., AND VILLENEUVE, S. Large risks, limited liability, and dynamic moral hazard. *Econometrica* 78, 1 (2010), 73–118.
- [12] BO, L., CAPPONI, A., AND LIAO, H. Deep residual learning via large sample mean-field stochastic optimization. *Preprint arXiv :1906.08894* (2019).

- [13] BOGACHEV, V. I., RÖCKNER, M., AND SHAPOSHNIKOV, S. V. Convergence in variation of solutions of nonlinear Fokker–Planck–Kolmogorov equations to stationary measures. *Journal of Functional Analysis* (2019).
- [14] BOLLEY, F., GENTIL, I., AND GUILLIN, A. Uniform convergence to equilibrium for granular media. *Archive for Rational Mechanics and Analysis* 208, 2 (2013), 429–445.
- [15] BOLTON, P., AND DEWATRIPONT, M. *Contract Theory*. MIT Press, 2005.
- [16] BOLTYANSKII, V. G., GAMKRELIDZE, R. V., AND PONTRYAGIN, L. S. The theory of optimal processes. I. The maximum principle. *TRW Space Technology Labs, Los Angeles, California* (1960).
- [17] BORCH, K. Equilibrium in a reinsurance market. *Econometrica* 30, 3 (1962), 424–444.
- [18] BOSSY, M., JOURDAIN, B., ET AL. Rate of convergence of a particle method for the solution of a 1d viscous scalar conservation law in a bounded interval. *The Annals of Probability* 30, 4 (2002), 1797–1832.
- [19] BOSSY, M., AND TALAY, D. A stochastic particle method for the McKean–Vlasov and the Burgers equation. *Mathematics of Computation of the American Mathematical Society* 66, 217 (1997), 157–192.
- [20] BOUCHARD, B. A stochastic target formulation for optimal switching problems in finite horizon. *Stochastics* 81, 2 (2009), 171–197.
- [21] BRYSON, A. E., AND DENHAM, W. F. A steepest ascent method for solving optimum programming problems. *Journal of Applied Mechanics* 29, 2 (1962), 247.
- [22] BUCKDAHN, R., DJEHICHE, B., LI, J., AND PENG, S. Mean-Field Backward Stochastic Differential Equations : A Limit Approach. *The Annals of Probability* 37, 4 (Jul. 2009), 1524–1565.
- [23] BUTKOVSKY, O. On ergodic properties of nonlinear Markov chains and stochastic McKean–Vlasov equations. *Theory of Probability & Its Applications* 58, 4 (2014), 661–674.
- [24] CARDALIAGUET, P., DELARUE, F., LASRY, J.-M., AND LIONS, P.-L. The master equation and the convergence problem in mean field games. *arXiv :1509.02505* (2015).
- [25] CARMONA, R., AND DELARUE, F. *Probabilistic Theory of Mean Field Games with Applications I*. Springer International Publishing, 2018.
- [26] CARMONA, R., AND DELARUE, F. *Probabilistic Theory of Mean Field Games with Applications II*. Springer International Publishing, 2018.
- [27] CARMONA, R., AND LACKER, D. A probabilistic weak formulation of mean field games and applications. *Ann. Appl. Probab.* 25, 3 (2015), 1189–1231.
- [28] CARRILLO, J. A., MCCANN, R. J., AND VILLANI, C. Kinetic equilibration rates for granular media and related equations : entropy dissipation and mass transportation estimates. *Revista Matemática Iberoamericana* 19, 3 (2003), 971–1018.

- [29] CATTIAUX, P., GUILLIN, A., AND MALRIEU, F. Probabilistic approach for granular media equations in the non-uniformly convex case. *Probability Theory and Related Fields* 140, 1-2 (2008), 19–40.
- [30] CHASSAGNEUX, J.-F., CRISAN, D., AND DELARUE, F. A probabilistic approach to classical solutions of the master equation for large population equilibria. *arXiv :1411.3009* (2014).
- [31] CHASSAGNEUX, J.-F., ELIE, R., AND KHARROUBI, I. A note on existence and uniqueness for solutions of multidimensional reflected BSDEs. *Electron. Commun. Probab* 16, 120-128 (2011).
- [32] CHASSAGNEUX, J.-F., SZPRUCH, L., AND TSE, A. Weak quantitative propagation of chaos via differential calculus on the space of measures. *arXiv :1901.02556* (2019).
- [33] CHEN, R. T. Q., RUBANOVA, Y., BETTENCOURT, J., AND DUVENAUD, D. Neural ordinary differential equations. *arXiv :1806.07366* (2019).
- [34] CHIZAT, L., AND BACH, F. On the global convergence of gradient descent for over-parameterized models using optimal transport. In *Advances in neural information processing systems* (2018), pp. 3040–3050.
- [35] CUCHIERO, C., LARSSON, M., AND TEICHMANN, J. Deep neural networks, generic universal interpolation, and controlled ODEs. *arXiv :1908.07838* (2019).
- [36] CVITANIĆ, J., POSSAMAÏ, D., AND TOUZI, N. Dynamic programming approach to principal-agent problems. *Finance and Stochastics* 22, 1 (Jan 2018), 1–37.
- [37] CVITANIC, J., WAN, X., AND YANG, H. Dynamics of contract design with screening. *Management Science* 59 (05 2013), 1229–1244.
- [38] CVITANIĆ, J., WAN, X., AND ZHANG, J. Optimal compensation with hidden action and lump-sum payment in a continuous-time model. *Applied Mathematics and Optimization* 59, 1 (Feb 2009), 99–146.
- [39] CVITANIC, J., AND ZHANG, J. The steepest descent method for forward-backward sdes. *Electron. J. Probab.* 10 (2005), 1468–1495.
- [40] CVITANIĆ, J., AND ZHANG, J. Optimal compensation with adverse selection and dynamic actions. *Math Finan Econ* (2007).
- [41] CVITANIĆ, J., AND ZHANG, J. *Contract theory in continuous-time models*. Springer Finance. Springer, Heidelberg, 2013.
- [42] CYBENKO, G. Approximation by superpositions of a sigmoidal function. *Mathematics of control, signals and systems* 2, 4 (1989), 303–314.
- [43] DELARUE, F. On the existence and uniqueness of solutions to fbsdes in a non-degenerate case. *Stochastic Process* 99, 209-286, *MR1918539* (2002).
- [44] DEMSKI, J. S., AND SAPPINGTON, D. Optimal incentive contracts with multiple agents. *Journal of Economic Theory* 33, 1 (1984), 152–171.

- [45] DUPUIS, P., AND ELLIS, R. S. *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley, 1997.
- [46] DURMUS, A., EBERLE, A., GUILLIN, A., AND ZIMMER, R. An elementary approach to uniform in time propagation of chaos. *arX :1805.11387* (2018).
- [47] E, W., HAN, J., AND LI, Q. A mean-field optimal control formulation of deep learning. *Res Math Sci* 6, 10 (2019).
- [48] EBERLE, A. Reflection couplings and contraction rates for diffusions. *Probability Theory and Related Fields* 166, 3-4 (2016), 851–886.
- [49] EBERLE, A., GUILLIN, A., AND ZIMMER, R. Quantitative Harris-type theorems for diffusions and McKean–Vlasov processes. *Transactions of the American Mathematical Society* 371, 10 (2019), 7135–7173.
- [50] EKELAND, I. Notes on optimal transportation. *Economy Theory* 42 (2010), 437–459.
- [51] EL KAROUÏ, N., NGUYEN, D., AND JEANBLANC-PICQUÉ, M. Compactification methods in the control of degenerate diffusions : existence of an optimal control. *Stochastics* 20, 3 (1987), 169–219.
- [52] EL KAROUÏ, N., PENG, S., AND QUENEZ, M. C. Backward stochastic differential equations in finance. *Math. Finance* 7, 1 (1997), 1–71.
- [53] ELIE, R., AND KHARROUBI, I. Probabilistic representation and approximation for coupled systems of variational inequalities. *Statist. Probab. Lett.* 80, 17-18 (2010), 1388–1396.
- [54] ELIE, R., MASTROLIA, T., AND POSSAMAÏ, D. A tale of a principal and many, many agents. *Mathematics of Operations Research* (to appear).
- [55] ELIE, R., AND POSSAMAÏ, D. Contracting theory with competitive interacting agents. *SIAM J. Control Optim.* 57, 2 (2019), 1157–1188.
- [56] FERNANDES, A., AND PHELAN., C. A Recursive Formulation for Repeated Agency with History Dependence. *Journal of Economic Theory* 91 (2000), 223–247.
- [57] FLEMING, W. H. Generalized solutions in optimal stochastic control. *Differential Games and Control theory II, Proceedings of 2nd Conference, Univ. of Rhode Island, Kingston, RI, 1976, Lect. Notes in Pure and Appl. Math.* 30 (1977), 147–165.
- [58] FÖLLMER, H. Time reversal on Wiener space. In *Stochastic Processes - Mathematics and Physics*, S. A. Alberverio, P. Blanchard, and L. Streit, Eds. Springer, 1986, pp. 119–129.
- [59] FREI, C., AND REIS, G. D. A financial market with interacting investors : does an equilibrium exist? *Mathematics and Financial Economics* 4 (2011), 161–182.
- [60] FROMM, A., AND IMKELLER, P. Existence, Uniqueness and Regularity of Decoupling Fields to Multidimensional Fully Coupled FBSDEs. *ArXiv e-prints* (Oct. 2013).

- [61] FUHRMAN, M., AND MORLAIS, M.-A. Optimal switching problems with an infinite set of modes : an approach by randomization and constrained backward SDEs. *arXiv e-prints* (Nov 2018), arXiv :1811.10886.
- [62] GEIRHOS, R., RUBISCH, P., MICHAELIS, C., BETHGE, M., WICHMANN, F. A., AND BRENDEL, W. ImageNet-trained CNNs are biased towards texture ; increasing shape bias improves accuracy and robustness. *arXiv e-prints* (Nov 2018), arXiv :1811.12231.
- [63] GOODFELLOW, I., BENGIO, Y., AND COURVILLE, A. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- [64] GOODFELLOW, I. J., SHLENS, J., AND SZEGEDY, C. Explaining and Harnessing Adversarial Examples. *arXiv e-prints* (Dec 2014), arXiv :1412.6572.
- [65] GREEN, J., AND STOKEY, N. A comparison of tournaments and contracts. *Journal of Political Economy* 91, 3 (1983), 349–64.
- [66] HAMADENE, S., AND LEPELTIER, J. Backward equations, stochastic control and zero-sum stochastic differential games. *Stochastics and Stochastic Reports* 54, 3-4 (1995), 221–231.
- [67] HAMADÈNE, S., AND ZHANG, J. Switching problem and related system of reflected backward SDEs. *Stochastic Process. Appl.* 120, 4 (2010), 403–426.
- [68] HARRIS, M., KRIEBEL, C. H., AND RAVIV, A. Asymmetric information, incentives and intrafirm resource allocation. *Management Science* 28, 6 (1982), 604–620.
- [69] HARTER, J., AND RICHOUE, A. A stability approach for solving multidimensional quadratic bsdes.
- [70] HASTIE, T., MONTANARI, A., ROSSET, S., AND TIBSHIRANI, R. J. Surprises in high-dimensional ridgeless least squares interpolation. *arXiv :1903.08560* (2019).
- [71] HAUSSMANN, U. G., AND LEPELTIER, J.-P. On the existence of optimal controls. *SIAM J. Control Optim.* 28, 4 (1990), 851–902.
- [72] HE, K., ZHANG, X., REN, S., AND SUN, J. Deep residual learning for image recognition. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2016).
- [73] HELLWIG, M., AND SCHMIDT, K. Discrete-time approximations of the Holmström-Milgrom Brownian-motion model of intertemporal incentive provision. *Econometrica* 70, 6 (2002), 2225–2264.
- [74] HENRY, D. *Geometric Theory of Semilinear Parabolic Equations*. Springer, 1981.
- [75] HOLLEY, R., AND STROOCK, D. Simulated annealing via Sobolev inequalities. *Communications in Mathematical Physics* 115, 4 (1988), 553–569.
- [76] HOLLEY, R. A., KUSUOKA, S., AND STROOCK, D. W. Asymptotics of the spectral gap with applications to the theory of simulated annealing. *Journal of functional analysis* 83, 2 (1989), 333–347.

- [77] HOLMSTROM, B. Moral hazard in teams. *Bell Journal of Economics* 13, 2 (1982), 324–340.
- [78] HOLMSTRÖM, B., AND MILGROM, P. Aggregation and linearity in the provision of intertemporal incentives. *Econometrica* 55, 2 (1987), 303–328.
- [79] HORNIK, K. Approximation capabilities of multilayer feedforward networks. *Neural Networks* 4, 2 (1991), 251–257.
- [80] HU, K. The wellposedness of path-dependent multidimensional forward-backward sde. *Preprint* (2019).
- [81] HU, K., REN, Z., SISKA, D., AND SZPRUCH, L. Mean-field Langevin dynamics and energy landscape of neural networks. *Preprint arXiv :1905.07769* (2019).
- [82] HU, K., REN, Z., AND TOUZI, N. Continuous-time principal-agent problem in degenerate systems, 2019.
- [83] HU, K., REN, Z., AND YANG, J. Principal-agent problem with multiple principals.
- [84] HU, Y., AND PENG, S. Solution of forward-backward stochastic differential equations. *Probability Theory and Related Fields* 103, 2 (Jun 1995), 273–283.
- [85] HU, Y., AND TANG, S. Multi-dimensional BSDE with oblique reflection and optimal switching. *Probability Theory and Related Fields* 147, 1-2 (2010), 89–121.
- [86] HUANG, M., CAINES, P. E., AND MALHAME, R. P. Large-population cost-coupled lqg problems with nonuniform agents : Individual-mass behavior and decentralized  $\varepsilon$ -nash equilibria. *IEEE Transactions on Automatic Control* 52, 9 (Sep. 2007), 1560–1571.
- [87] HWANG, C.-R., ET AL. Laplace’s method revisited : weak convergence of probability measures. *The Annals of Probability* 8, 6 (1980), 1177–1182.
- [88] JABIR, J.-F., SISKA, D., AND SZPRUCH, L. Mean-field neural odes via relaxed optimal control. *Preprint arXiv :1912.05475* (2019).
- [89] JACOD, J. Multivariate Point Processes : Predictable Projection, Radon-Nikodym Derivatives, Representation of Martingales. *Z. Wahrscheinlichkeitstheorie verw* 31, 3 (1975), 235–253.
- [90] JAVANMARD, A., MONDELLI, M., AND MONTANARI, A. Analysis of a two-layer neural network via displacement convexity. *arXiv :1901.01375* (2019).
- [91] JIONGMIN YONG, X. Z. *Stochastic Controls*, vol. 43. Springer-Verlag New York, 1999.
- [92] JORDAN, R., AND KINDERLEHRER, D. An extended variational principle. In *Partial differential equations and applications : collected papers in honor of Carlo Pucci*. CRC, 1996.
- [93] JORDAN, R., KINDERLEHRER, D., AND OTTO, F. The variational formulation of the Fokker–Planck equation. *SIAM J. Math. Anal.* 29, 1 (January 1998), 1–17.
- [94] KARATZAS, I., SCHACHERMAYER, W., AND TSCHIDERER, B. Pathwise Otto calculus. *arXiv :1811.08686* (2019).

- [95] KAZAMAKI, N. *Continuous Exponential Martingales and BMO*. Springer-Verlag Berlin Heidelberg, 1994.
- [96] KEUN KOO, H., SHIM, G., AND SUNG, J. Optimal multi-agent performance measures for team contracts. *Mathematical Finance* 18, 4 (2008), 649–667.
- [97] KOBYLANSKI, M. Backward stochastic differential equations and partial differential equations with quadratic growth. *The Annals of Probability Vol.28, No.2*, 558-602 (2000).
- [98] KRIZHEVSKY, A., SUTSKEVER, I., AND HINTON, G. E. Imagenet classification with deep convolutional neural networks. In *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1* (USA, 2012), NIPS'12, Curran Associates Inc., pp. 1097–1105.
- [99] LACKER, D. Mean field games via controlled martingale problems : existence of Markovian equilibria. *Stochastic Processes and their Applications* 125, 7 (2015), 2856–2894.
- [100] LADYZENSKAJA, O. A., SOLONNIKOV, V. A., AND URAL'CEVA, N. N. *Linear and quasi-linear equations of parabolic type*. Translations of Mathematical Monographs. AMS, 1968.
- [101] LAFFONT, J.-J., AND MARTIMORT, D. The theory of incentives : The principal-agent model / j.j. laffont, d. martimort.
- [102] LASRY, J.-M., AND LIONS, P.-L. Mean field games. *Jpn. J. Math.* 2, 1 (2007), 229–260.
- [103] LECUN, Y., BENGIO, Y., AND HINTON, G. Deep learning. *Nature* 521, 7553 (2015), 436.
- [104] LECUN, Y., TOURESKY, D., HINTON, G., AND SEJNOWSKI, T. A theoretical framework for back-propagation. *Proceedings of the 1988 connectionist models summer school 1* (1988), 21–28.
- [105] LI, Q., LIN, T., AND SHEN, Z. Deep learning via dynamical systems : An approximation perspective. *Preprint arXiv :1912.10382* (2019).
- [106] LIU, G.-H., AND THEODOROU, E. A. Deep learning theory review : An optimal control and dynamical systems perspective. *arXiv : 1908.10920* (2019).
- [107] LIU, H., AND MARKOWICH, P. Selection dynamics for deep neural networks. *arXiv :1905.09076* (2019).
- [108] LU, Y., MA, C., LU, Y., LU, J., AND YING, L. A mean-field analysis of deep resnet and beyond : Towards provable optimization via overparameterization from depth. *Preprint arXiv :2003.05508* (2020).
- [109] MA, J., PROTTER, P., AND YONG, J. Solving forward-backward stochastic differential equations explicitly — a four step scheme. *Probability Theory and Related Fields* 98, 3 (Sep 1994), 339–359.
- [110] MA, J., WU, Z., ZHANG, D., AND ZHANG, J. On well-posedness of forward-backward SDEs—a unified approach. *Ann. Appl. Probab.* 25, 4 (08 2015), 2168–2214.

- 
- [111] MA, J., AND YONG, J. Adapted solution of a degenerate backward spde, with applications. *Stochastic Processes and their Applications* 70, 1 (1997), 59 – 84.
- [112] MA, J., AND YONG, J. On linear, degenerate backward stochastic partial differential equations. *Probability Theory and Related Fields* 113, 2 (Feb 1999), 135–170.
- [113] MA, J., AND YONG, J. *Forward-Backward Stochastic Differential Equations and their Applications*. Springer, Berlin, Heidelberg, 2007.
- [114] MALRIEU, F., ET AL. Convergence to equilibrium for granular media equations and their Euler schemes. *Annals of Applied Probability* 13, 2 (2003), 540–560.
- [115] MASTROLIA, T., AND REN, Z. Principal-Agent problem with common agency without communication. *SIAM Journal of Financial Mathematics* 9, 2 (2018), 775–799.
- [116] MEI, S., MISIAKIEWICZ, T., AND MONTANARI, A. Mean-field theory of two-layers neural networks : dimension-free bounds and kernel limit. *arXiv :1902.06015* (2019).
- [117] MEI, S., MONTANARI, A., AND NGUYEN, P.-M. A mean field view of the landscape of two-layer neural networks. *Proceedings of the National Academy of Sciences* 115, 33 (2018), E7665–E7671.
- [118] MEYER, P. A. *Probability and Potentials*. Blaisdell Publishing Company, 1966.
- [119] MIRRLEES, J. A. The theory of moral hazard and unobservable behaviour : Part i. *The Review of Economic Studies* 66, 1 (1999), 3–21.
- [120] MISCHLER, S., AND MOUHOT, C. Kac’s program in kinetic theory. *Inventiones Mathematicae* 193, 1 (2013), 1–147.
- [121] MITTER, S. K. Successive approximation methods for the solution of optimal control problems. *Automatica* 3, 3-4 (1966), 135–149.
- [122] MOOKHERJEE, D. Optimal incentive schemes with many agents. *Review of Economic Studies* 51, 3 (1984), 433–446.
- [123] MÜLLER, H. The first-best sharing rule in the continuous-time principal-agent problem with exponential utility. *Journal of Economic Theory* 79, 2 (1998), 276–280.
- [124] MÜLLER, H. Asymptotic efficiency in dynamic principal-agent problems. *Journal of Economic Theory* 91, 2 (2000), 292–301.
- [125] NALEBUFF, B., AND STIGLITZ, J. Prices and incentives : Towards a general theory of compensation and competition. *Bell Journal of Economics* 14, 1 (1983), 21–43.
- [126] NUALART, D. *The Malliavin Calculus and Related Topics*. Springer, 2006.
- [127] OTTO, F. The geometry of dissipative evolution equations : the porous medium equation.
- [128] OTTO, F., AND VILLANI, C. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *Journal of Functional Analysis* 173 (2000), 361–400.



- [129] PARDOUX, E., AND PENG, S. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters* 14, 1 (1990), 55–61.
- [130] PARDOUX, E., AND PENG, S. Backward stochastic differential equations and quasilinear parabolic partial differential equations. 200–217.
- [131] PARDOUX, E., AND TANG, S. Forward-backward stochastic differential equations and quasilinear parabolic pdes. *Probability Theory and Related Fields* 114 (01 1999), 123–150.
- [132] PEARLMUTTER, B. A. Gradient calculations for dynamic recurrent neural networks : A survey. *IEEE Transactions on Neural networks* 6, 5 (1995), 1212–1228.
- [133] PENG, S. A general stochastic maximum principle for optimal control problems. *SIAM Journal on Control and Optimization* 28, 4 (1990), 966–979.
- [134] PENG, S., AND WU, Z. Fully coupled forward-backward stochastic differential equations and applications to optimal control. *SIAM Journal on Control and Optimization* 37, 3 (1999), 825–843.
- [135] PHAM, H., VATHANA, L. V., AND ZHOU, X. Y. Optimal Switching over Multiple Regimes. *SIAM J. Control Optim.* 48, 4 (2009), 2217–2253.
- [136] PRAT, J., AND JOVANOVIC, B. Dynamic incentive contracts under parameter uncertainty. *IZA DP No. 5323* (2010).
- [137] ROTSKOFF, G. M., JELASSI, S., BRUNA, J., AND VANDEN-EIJNDEN, E. Global convergence of neuron birth-death dynamics. *Proceedings of the 36th International Conference on Machine Learning, PMLR 97 :5508-5517* (2019).
- [138] ROTSKOFF, G. M., AND VANDEN-EIJNDEN, E. Neural networks as interacting particle systems : Asymptotic convexity of the loss landscape and universal scaling of the approximation error. *arXiv :1805.00915* (2018).
- [139] SANNIKOV, Y. A continuous-time version of the principal-agent problem. *Rev. Econom. Stud.* 75, 3 (2008), 957–984.
- [140] SCHÄTTLER, H., AND SUNG, J. The first-order approach to the continuous-time principal-agent problem with exponential utility. *Journal of Economic Theory* 61, 2 (1993), 331–371.
- [141] SCHÄTTLER, H., AND SUNG, J. On optimal sharing rules in discrete-and continuous-time principal-agent problems with exponential utility. *Journal of Economic Dynamics and Control* 21, 2 (1997), 551–574.
- [142] SILVER, D., HUANG, A., MADDISON, C. J., GUEZ, A., SIFRE, L., VAN DEN DRIESSCHE, G., SCHRITTWIESER, J., ANTONOGLU, I., PANNEERSHELVAM, V., LANCTOT, M., DIELEMAN, S., GREWE, D., NHAM, J., KALCHBRENNER, N., SUTSKEVER, I., LILICRAP, T., LEACH, M., KAVUKCUOGLU, K., GRAEPEL, T., AND HASSABIS, D. Mastering the game of Go with deep neural networks and tree search. *Nature* 529, 7587 (2016), 484–489.

- 
- [143] SIRIGNANO, J., AND SPILIOPOULOS, K. Mean field analysis of neural networks. *arXiv :1805.01053* (2018).
- [144] SRIVASTAVA, N., HINTON, G., KRIZHEVSKY, A., SUTSKEVER, I., AND SALAKHUTDINOV, R. Dropout : A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research* (2014), 1929–1958.
- [145] SRIVASTAVA, N., AND SALAKHUTDINOV, R. R. Multimodal learning with deep boltzmann machines. In *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 2222–2230.
- [146] SUNG, J. Linearity with project selection and controllable diffusion rate in continuous-time principal-agent problems. *The RAND Journal of Economics* 26, 4 (1995), 720–743.
- [147] SUNG, J. Corporate insurance and managerial incentives. *Journal of Economic Dynamics and Control* 74, 2 (1997), 297–332.
- [148] SZNITMAN, A.-S. *Topics in propagation of chaos*. Springer, 1991.
- [149] SZPRUCH, L., TAN, S., AND TSE, A. Iterative particle approximation for McKean–Vlasov sdes with application to multilevel Monte Carlo estimation. *To appear in Annals of Applied Probability* (2019).
- [150] SZPRUCH, L., AND TSE, A. Antithetic multilevel particle system sampling method for McKean–Vlasov SDEs. *arXiv :1903.07063* (2019).
- [151] TUGAUT, J. Convergence to the equilibria for self-stabilizing processes in double-well landscape. *The Annals of Probability* 41, 3A (2013), 1427–1460.
- [152] VAPNIK, V. *The nature of statistical learning theory*. Springer, 2013.
- [153] VERETENNIKOV, A. Y. On ergodic measures for McKean–Vlasov stochastic equations. In *Monte Carlo and Quasi-Monte Carlo Methods 2004*. Springer, 2006, pp. 471–486.
- [154] WANG, H., ZARIPHOPULOU, T., AND ZHOU, X. Exploration versus exploitation in reinforcement learning : A stochastic control approach. *Preprint arXiv :1812.01552* (2018).
- [155] WILLIAMS, N. On dynamic principal-agent problems in continuous time. *University of Wisconsin, Madison* (2009).
- [156] WILLIAMS, N. Persistent Private Information. *University of Wisconsin, Madison* (2009).
- [157] WU, Z., AND YU, Z. Probabilistic interpretation for a system of quasilinear parabolic partial differential equation combined with algebra equations. *Stochastic Processes and their Applications* 124, 12 (2014), 3921 – 3947.
- [158] XING, H., AND ŽITKOVIĆ, G. A class of globally solvable markovian quadratic bsde systems and applications.
- [159] YONG, J. Finding adapted solutions of forward-backward stochastic differential equations : method of continuation. *Probability Theory and Related Fields* 107 (1997), 537–572.

- [160] YONG, J. Forward-backward stochastic differential equations with mixed initial-terminal conditions. *Trans. Amer. Math. Soc.* 362 (2009).
- [161] YU, Z. Equivalent cost functionals and stochastic linear quadratic optimal control problems. *ESAIM : Control, Optimisation and Calculus of Variations* 19, 1 (2013), 78–90.
- [162] ZHANG, C., BENGIO, S., HARDT, M., RECHT, B., AND VINYALS, O. Understanding deep learning requires rethinking generalization. *arXiv :1611.03530* (2016).
- [163] ZHANG, J. The wellposedness of fbsdes. *Discrete Contin. Dyn. Syst. Ser. B* 6 (2006), 927–940.
- [164] ZHANG, J. *Backward Stochastic Differential Equations, From Linear to Fully Nonlinear Theory*. Springer, 2017.
- [165] ZHANG, J. The wellposedness of fbsdes (ii). *arXiv :1708.05785* (2017).

**Titre :** Jeux Différentiels Stochastiques non-Markoviens et Dynamiques de Langevin à Champ-moyen

**Mots clés :** principe du maximum stochastique, problème de Principal-Agent, jeu à champ-moyen, dynamiques de Langevin à champ-moyen, flot du gradient, réseau de neurones

**Résumé :** Cette thèse se compose de deux parties indépendantes et la première regroupant deux problématiques distinctes.

Dans la première partie, nous étudions d'abord le problème de Principal-Agent dans des systèmes dégénérés avec une approche se basant sur le principe du maximum stochastique. Une étude parallèle est consacrée à l'existence, l'unicité et la stabilité de la solution des EDSRs dépendantes de la trajectoire qui apparaissent dans la résolution du problème précédent. Enfin, nous étudions le problème d'aléa moral avec plusieurs Principals. On suppose que l'Agent ne peut travailler que pour un seul Principal à la fois. En utilisant la méthode de randomisation nous caractérisons la fonction valeur de l'Agent et son effort optimal par un processus d'Itô. Nous résolvons ensuite le problème du Principal lorsqu'il y a une infinité de Principals en équilibre à champ-moyen. Nous justifions la formulation à champ-moyen par un argument de propagation de chaos rétrograde.

La motivation de la deuxième partie est de donner un fondement théorique rigoureux pour la convergence

des algorithmes du type descente de gradient utilisés dans la calibration des réseaux de neurones. Dans le cas de réseaux de neurones à une couche, nous montrons que le problème relevant dans l'espace des mesures admet un unique minimiseur caractérisé par une condition du premier ordre. Avec une analyse du comportement à long terme de la dynamique de Langevin à champ-moyen, nous montrons que le flot de la loi marginale induite par la dynamique de Langevin à champ-moyen converge vers la loi stationnaire, qui est le minimiseur du problème. Dans le cas de réseaux de neurones profonds, nous les modélisons à l'aide d'un problème de contrôle optimal en temps continu. Nous donnons d'abord la condition du premier ordre à l'aide du principe de Pontryagin, qui nous aidera ensuite à introduire le système de Langevin à champ-moyen, dont la mesure invariante correspond au minimiseur du problème de contrôle. Enfin, avec la méthode de couplage par réflexion nous montrons que la loi marginale du système de Langevin à champ-moyen converge vers la mesure invariante avec une vitesse exponentielle.

**Title :** Non-Markovian Stochastic Differential Games and Mean-field Langevin Dynamics

**Keywords :** stochastic maximum principles, Principal-Agent problem, mean field games, mean-field Langevin dynamics, gradient flow, neural networks

**Abstract :** Two independent subjects are studied in this thesis, the first of which consists of two distinct problems.

In the first part, we study the Principal-Agent problem in degenerate systems using stochastic maximum principle. A parallel study is devoted to the wellposedness and the stability of path-dependent FBSDEs which appear while solving the aforementioned problem. Lastly we propose an approach to solve the Principal-Agent problem with multiple Principals where the Agent is only allowed to work for one Principal at a time. By using randomization, we show that the value function of the Agent's problem and his optimal control are given by an Itô process and we solve the Principal's problem in the mean-field case. The mean-field formulation is justified using an argument of backward propagation of chaos.

The motivation of the second part is to give a rigorous theoretical underpinning for the convergence of gradient-descent type of algorithms frequently used

in non-convex optimization problems like calibrating a deep neural network. For one-layer neural networks, we show that the corresponding energy function lifted into the measure space has a unique minimiser characterized by some first order condition. By using a probabilistic analysis of the long-time behavior of the mean-field Langevin dynamics, we show that the flow of marginal laws induced by the mean-field Langevin dynamics converges to the stationary distribution, which is exactly the minimiser of the energy function. As for deep neural networks, we first model them as continuous-time optimal control problems. Then, by using Pontryagin maximum principle we find the first order condition and identify the associated mean-field Langevin system, the invariant measure of which solves the control problem. Finally we show the marginal distribution of the mean-field Langevin system converges exponentially to the unique invariant measure by reflection couplings.