



HAL
open science

Mean-field infinite horizon stochastic control and numerical approximation of infinite horizon backward stochastic differential equations

Charu Shardul

► **To cite this version:**

Charu Shardul. Mean-field infinite horizon stochastic control and numerical approximation of infinite horizon backward stochastic differential equations. Mathematics [math]. Université de Bordeaux, 2024. English. NNT: . tel-04627360v1

HAL Id: tel-04627360

<https://theses.hal.science/tel-04627360v1>

Submitted on 27 Jun 2024 (v1), last revised 3 Jul 2024 (v2)

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.

THÈSE PRÉSENTÉE
POUR OBTENIR LE GRADE DE
DOCTEUR DE
L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE MATHÉMATIQUES ET INFORMATIQUE

MATHÉMATIQUES APPLIQUÉES ET CALCUL SCIENTIFIQUE

Par **Charu SHARDUL**

**Contrôle stochastique de type champ moyen en horizon infini et
approximation numérique des équations différentielles stochastiques
rétrogrades en horizon infini**

(Mean-field infinite horizon stochastic control and numerical approximation of infinite
horizon backward stochastic differential equations)

Sous la direction de : **Adrien RICHOU**
Co-directeur : **Emmanuel GOBET**

Soutenue le 14 Mai 2024

Membres du jury :

M. Adrien RICHOU	Maitre de conférences	Université de Bordeaux	Directeur de thèse
M. Emmanuel GOBET	Professeur	École Polytechnique	Co-directeur de thèse
M. Alexandre POPIER	Professeur	Le Mans Université	Rapporteur
M. Mohamed MRAD	Maitre de conférences	Université Sorbonne Paris Nord	Rapporteur
M. Marc ARNAUDON	Professeur	Université de Bordeaux	Examinateur
M. Jean-François CHASSAGNEUX	Professeur	Université Paris Cité	Examinateur
M. Stéphane CRÉPEY	Professeur	Université Paris Cité	Examinateur
Mme. Céline LABART	Maitresse de conférences	Université Savoie Mont-Blanc	Examinatrice

Acknowledgements

Throughout the journey of this thesis, I had the privilege of meeting and working with very talented researchers. I would like to express my deep gratitude towards everyone who contributed to my development from being a student to a researcher and the successful completion of my thesis.

I am very grateful to Adrien Richou for his thorough supervision of my thesis and especially for his kindness and consistent support. His deep mathematical prowess in BSDEs was essential for a lot of the avenues of research we explored together. His depth of understanding and his mathematical rigor have always motivated me to do better and helped me develop a lot of the mathematical aptitude required for our field. I am very fortunate to have him as my supervisor and highly appreciate his numerous efforts to make my PhD a success. I especially thank him for helping me with not just professional but also personal matters. His excellent sense of humor makes our meetings very refreshing.

I would really like to thank Emmanuel Gobet for supervising my thesis along with Adrien. I met him first during my masters thesis and have known him for 5 years now. He has helped me develop the thinking process that is essential for being a good researcher. He is always full of ideas which has been a key motivating factor for me. His proficiency and the vast expanse of knowledge in so many fields continues to astonish me. He always seemed to know some mathematical results which exactly fit any problem I was stuck on.

I also thank him and the Stress Test Chair from Ecole Polytechnique for sponsoring my various visits to meet him in Paris. They also supported me and made it possible for me to attend several conferences and workshops for which I am very grateful. In particular, it enabled me to present my work at the prestigious SIAM conference on financial mathematics and engineering in Philadelphia, USA. Emmanuel has also helped me a lot towards my integration with the scientific community in our field and helped me to establish collaborations with BNP Paribas for future work. I must also thank Marine Saux from CMAP for helping me with several administrative procedures. My visits to CMAP enabled me to make valuable acquaintances with other researchers in the SIMPAS team who work in similar topics as me. In particular, I was fortunate to meet Wanqing, Manon, Michaël, Clara, Jessie, David, Elisa, Célia, Clara, Dorinel, Maxime and others. I'm especially grateful to Maxime Grangereau who guided me to very good references for getting introduced to mean-field stochastic control. His research in the applications of mean-field control to smart electricity grids is very promising and insightful. I must also thank Wanqing whom I met during my masters thesis with Emmanuel and she made the same journey as me. She has been a good friend and helped me with several things when I first arrived in France. I also thank Cyril Bénézet for advising me about the numerical implementation of neural networks for solving BSDEs, which he also did in his research work.

I am also very grateful to IMB and University of Bordeaux for hosting me as a PhD student and contributing to my academic development through various courses and seminars. I also thank the members of the IOP team and fellow doctoral students who shared this journey with me. In

particular, I thank Paul, Jean, Yiye, Gauthier, Théo, Sam, Virginie, Magalie, Bianca, Hui, Issa and others.

I sincerely thank the reviewers for providing a thorough feedback and suggesting several improvements in the manuscript, especially given the short notice. Their valuable feedback has made the quality of this manuscript much better and the arguments much clearer. I also thank them for pointing out the possibilities to allow for the applicability of my work to a broader class of problems in finance.

I wholeheartedly thank all my professors from Center for Excellence in Basic Sciences (CEBS) in India for training me in the essential mathematical basics and helping me develop the mathematical aptitude. I especially thank Prof. Mallikarjuna Rao from IIT Mumbai with whom I worked on BSDEs for the first time and I sincerely thank Prof. Sandeep Juneja from TIFR Mumbai for introducing me to financial mathematics and helping me appreciate the beauty of applying mathematical ideas for difficult problems in finance.

Lastly, I must thank my family and friends for supporting me and especially my mother whose wise words or advice have always given me a strong sense of purpose. The simplicity in her approach towards life and her lessons of kindness have inspired me to become a better person and instilled a clarity of thought which has always helped me in difficult situations. I thank my sister to share my responsibilities towards my family during my prolonged absence.

I thank everyone for believing in me and supporting me through this wonderful journey.

Contents

0	Introduction (en français)	1
0.1	Préliminaires et contexte	1
0.1.1	Contrôle stochastique	1
0.1.2	EDSR et EDSPR	4
0.1.3	Contrôle champs-moyen	7
0.2	Contributions et défis	11
0.2.1	Contrôle stochastique de champ moyen à horizon infini et applications	11
0.2.2	EDSR en horizon infini	17
0.3	Conclusions	21
1	Introduction	23
1.1	Preliminaries and context	23
1.1.1	Stochastic control	23
1.1.2	BSDEs and FBSDEs	25
1.1.3	Mean-field control	29
1.2	Contributions and challenges	32
1.2.1	Mean-field stochastic control in infinite horizon and applications	32
1.2.2	BSDEs in infinite horizon	38
1.3	Conclusions	42
I	Mean-field stochastic control in infinite horizon and applications	45
2	A study of mean-field infinite horizon stochastic control	47
2.1	Introduction	47
2.2	Preliminaries	49
2.2.1	Differentiability on the space of measures	49
2.2.2	The Model	50
2.2.3	A useful estimate	52
2.2.4	Adjoint processes	53
2.3	Necessary conditions for optimality	54
2.4	Sufficient condition	59
2.5	Examples	61
2.5.1	The mean-field infinite horizon LQ case	61
2.5.2	Optimal liquidation problem in infinite horizon	63
2.5.3	Challenges with standard examples	63
3	Long term dynamic portfolio optimization using infinite horizon McKean-Vlasov FBSDEs	65
3.1	Introduction	66

3.2	Model and preliminaries	67
3.2.1	State dynamics and cost functional	67
3.2.2	Results for the mean-field term	71
3.2.3	The control problem and assumption	73
3.3	Conditions for optimality	74
3.3.1	A priori estimate on state process	74
3.3.2	Adjoint processes	76
3.3.3	Necessary and sufficient conditions for optimality	76
3.3.4	Characterisation of optimal control using FBSDE	77
3.4	Numerical Analysis	83
3.4.1	General setting	87
3.4.2	Scheme based on fictitious play	88
3.4.3	Scheme based on dynamic updates	89
3.4.4	Numerical Experiments	90
3.5	Proofs	94
3.5.1	Proof of Theorem 3.3.3	94
3.5.2	Proof of Theorem 3.3.4	97

II BSDEs in infinite horizon 99

4	Numerical approximation of Markovian BSDEs in infinite horizon and elliptic PDEs	101
4.1	Introduction	102
4.1.1	Statement of the problem	104
4.1.2	Our contributions.	104
4.1.3	Notations	105
4.2	Analytical results	105
4.2.1	Model and value function	105
4.2.2	Time randomized Feynman-Kac representation	107
4.2.3	Contraction properties of the fixed point equation	108
4.2.4	Some sufficient assumptions to get $\kappa_\infty < 1$	113
4.2.5	How to check the contraction in L^p	116
4.3	Numerical schemes	118
4.3.1	A first scheme based on a regular grid	118
4.3.2	Some schemes based on neural networks approximation	122
4.3.3	Numerical experiments	125
4.4	Proofs	131
4.4.1	Proof of Proposition 4.2.3	131
4.4.2	Proof of Proposition 4.3.3	135
4.4.3	Proof of Lemma 4.3.1	137

Chapter 0

Introduction (en français)

Contents

0.1	Préliminaires et contexte	1
0.1.1	Contrôle stochastique	1
0.1.2	EDSR et EDSPR	4
0.1.3	Contrôle champs-moyen	7
0.2	Contributions et défis	11
0.2.1	Contrôle stochastique de champ moyen à horizon infini et applications	11
0.2.2	EDSR en horizon infini	17
0.3	Conclusions	21

Les principaux points focaux de cette thèse sont les équations différentielles stochastiques rétrogrades (EDSR) et les équations différentielles stochastiques progressives-rétrogrades de McKean-Vlasov (EDSPR MV) à horizon infini, leurs utilisations et surtout leur lien avec les problèmes de contrôle stochastique à champ moyen en horizon infini. Dans la première partie de la thèse, nous les étudions d'abord dans leur pleine généralité, puis nous abordons un problème d'optimisation de portefeuille à long terme avec un critère d'optimisation à champ moyen et le résolvons numériquement. Dans la deuxième partie, nous concentrons notre attention sur le développement de schémas numériques efficaces pour résoudre les équations différentielles stochastiques rétrogrades (EDSR) à horizon infini, qui ont récemment trouvé des applications dans de nombreux domaines.

L'introduction est structurée comme suit : la première section fournit le contexte plus large de la thèse et offre un aperçu concis du contrôle stochastique et de sa connexion avec les EDSR. Nous présentons également un aperçu des systèmes à champ moyen et du contrôle à champ moyen, ainsi qu'une revue de l'état de l'art. La deuxième section met en évidence les défis auxquels nous avons été confrontés et nos contributions. La troisième et dernière section présente nos conclusions et explore les perspectives futures.

0.1 Préliminaires et contexte

0.1.1 Contrôle stochastique

Contexte

Le contrôle stochastique est une branche spécialisée de la théorie du contrôle qui étend les principes de l'optimisation mathématique aux systèmes dynamiques influencés ou pilotés par

l'incertitude. La contribution initiale la plus remarquable à la théorie du contrôle fut la programmation dynamique de Bellman [16] dans les années 1950, marquant une étape importante en mathématiques appliquées. Kalman [82] établit le lien entre la programmation dynamique et l'équation de Hamilton-Jacobi de la physique classique en 1960. Voir Åström [118] pour plus de perspectives historiques sur le contrôle stochastique. Barles [12] a étudié les solutions de viscosité pour l'équation de Hamilton-Jacobi-Bellman (HJB), posant des bases théoriques pour les problèmes de contrôle stochastique en temps continu.

Applications

Au fil des ans, les méthodologies de contrôle stochastique sont devenues indispensables pour modéliser des problèmes complexes dans de nombreux domaines et plus particulièrement dans les marchés financiers. Le travail fondateur de Merton [94] a jeté les bases de l'application du contrôle stochastique à la sélection optimale de portefeuille, qui est toujours un pilier de la théorie moderne de portefeuille. En étendant ce cadre, Black and Scholes [24] et [110] ont développé le modèle de Black-Scholes-Merton, révolutionnant la valorisation des options et des produits dérivés. L'inclusion de processus de saut dans les modèles de contrôle stochastique a permis de modéliser les événements extrêmes sur les marchés financiers. Merton [95] a introduit le modèle de diffusion à sauts, et Duffie and Pan [55] l'ont étendu pour inclure le risque de défaut, conduisant au développement de modèles sophistiqués de risque de crédit. Fleming and Soner [59] ont fourni des informations sur les stratégies d'investissement optimales en présence de coûts de transaction, abordant un aspect critique du trading réel. Les avancées dans la modélisation de la microstructure du marché, telles que [28, 21, 46, 75, 83], ont trouvé d'autres applications financières du contrôle stochastique. Par ailleurs, il y a eu également de nombreuses applications du contrôle stochastique dans d'autres domaines.

Les méthodologies de contrôle stochastique sont essentielles pour relever les défis posés par la nature dynamique et incertaine du secteur de l'énergie. En finance de l'énergie, ces méthodologies ont souvent été utilisées dans la valorisation et la couverture des produits dérivés de l'énergie. Schwartz and Smith [111] ont introduit des processus de retour à la moyenne pour modéliser les prix de l'énergie, jetant les bases de l'incorporation de techniques de contrôle stochastique. Leur travail a posé les bases pour l'optimisation des stratégies de trading sur les marchés avec des dynamiques de retour à la moyenne. Dixit and Pindyck [53] ont appliqué la théorie des options réelles aux décisions d'investissement dans l'énergie. Les méthodes de contrôle stochastique étaient cruciales pour valoriser et optimiser l'exercice des options réelles dans le contexte de projets énergétiques. Alipour et al. [8] ont abordé le pilotage optimal des systèmes électriques à l'aide du contrôle stochastique. Ils ont fourni des informations sur les stratégies de répartitions optimales en tenant compte des incertitudes sur les prix de l'électricité et la demande. Alhelou et al. [7] se sont concentrés sur l'optimisation du fonctionnement des systèmes énergétiques hybrides renouvelables sous incertitude. Leur travail a mis en évidence l'importance du contrôle stochastique dans la gestion de l'intégration des sources d'énergie renouvelables. Alharbi and Bhattacharya [6] ont fourni une revue complète des approches de contrôle stochastique pour l'optimisation et la dimensionnement de systèmes de stockage d'énergie. Leur travail a abordé le rôle crucial de l'optimisation du stockage d'énergie dans la gestion des systèmes énergétiques.

Aperçu mathématique

Passons maintenant au formalisme mathématique du contrôle stochastique. Soit $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ un espace de probabilité filtré avec une filtration augmentée et continue à droite $(\mathcal{F}_t)_{t \geq 0}$, et considérons l'équation différentielle stochastique suivante (EDS) comme la dynamique du système

contrôlé,

$$dX_t^u = b(X_t^u, u_t) dt + \sum_{j=1}^m \sigma^j(X_t^u, u_t) dW_t^j, \quad X_{t_0} = x_0,$$

où X_t^u est appelée la variable d'état ou le processus d'état du système, prenant des valeurs dans \mathbb{R}^d . $W_t \in \mathbb{R}^m$ est un mouvement brownien adapté à \mathcal{F}_t et $(u_t)_{t \geq 0}$ est un processus progressivement mesurable par rapport à \mathcal{F}_t appelé le contrôle, prenant des valeurs dans un sous-ensemble fermé et convexe non vide $U \subset \mathbb{R}^n$. L'objectif du processus d'optimisation pourrait être de maximiser une certaine récompense ou de minimiser des coûts, qui pourraient être une fonction du processus d'état contrôlé ou du processus de contrôle lui-même, et souvent des deux. Considérons un problème de minimisation de coûts à titre d'exemple. La fonction de coût est généralement donnée comme suit,

$$J^T(x_0, t_0, u) = \mathbb{E} \left[\int_0^T f(X_t^{u, x_0}, u_t) dt + g(X_T^{x_0, u}) \right],$$

où la fonction f est appelée coût d'exécution et g est appelée coût terminal. L'objectif du problème d'optimisation est de trouver $u = (u_t)_{t \geq 0}$ dans un espace noté \mathcal{A} , appelé l'espace des contrôles admissibles (processus prévisibles et bornés pour un horizon fini), qui minimise cette fonction de coût. La fonction $V^T(x_0) = \min_{u \in \mathcal{A}} J^T(x_0, u)$ est appelée fonction valeur et $u^* \in \arg \min_{u \in \mathcal{A}} J^T(x_0, u)$ est appelé un contrôle optimal. On pourrait modéliser des problèmes sans temps terminal en utilisant une fonction de coût avec un facteur d'actualisation pour prendre en compte la valeur décroissante des coûts futurs et un horizon temporel infini, ce qui est principalement notre focus dans la thèse. On écrit la fonction de coût comme suit,

$$J(x_0, u) = \mathbb{E} \left[\int_0^\infty e^{-\rho t} f(X_t^{u, x_0}, u_t) dt \right].$$

La programmation dynamique est l'une des deux principales approches pour résoudre les problèmes de contrôle stochastique. L'équation de Hamilton-Jacobi-Bellman (HJB),

$$-\rho V(x) + \min_{u \in U} \{ \mathcal{L}^u V + f(x, u) \} = 0, \quad \forall x \in \mathbb{R}^d,$$

où \mathcal{L}^u est le générateur du processus d'état. Un contrôle optimal en boucle fermée est alors donné par $u^*(x) = \arg \min_{u \in U} \{ \mathcal{L}^u V + f(x, u) \}$. L'autre approche principale pour résoudre les problèmes de contrôle stochastique est le principe du maximum de Pontryagin stochastique (voir Pontryagin et al. [107]), où l'on définit une fonction appelée hamiltonien, $H(x, u, y, z) = b(x, u) \cdot y + \text{tr}(\sigma^\top(x, u)z) + f(x, u)$, où (y, z) avec $y \in \mathbb{R}^d$ et $z \in \mathbb{R}^{m \times d}$ sont appelées les variables adjointes qui proviennent de la paire de solutions $(Y_t, Z_t)_{t \geq 0}$ de l'EDSR suivante pour un $u \in \mathcal{A}$ donné,

$$dY_t = -\nabla_x H(X_t^u, u_t, Y_t, Z_t) dt + Z_t dW_t.$$

Un contrôle optimal est donné par $u_t^* = \arg \min_{u \in U} H(X_t^u, u, Y_t, Z_t)$ qui est une fonction de la solution de l'EDSR. Avec cette caractérisation du contrôle optimal mise dans la dynamique d'état, on obtient un système couplé d'EDSPRs. Cela forme un lien crucial entre le contrôle stochastique et les EDSRs, qui est exploré en détail dans la thèse. Pour une couverture exhaustive de la connexion entre la théorie du contrôle et les EDSRs, voir Ma et Yong [81]. Pour un excellent fond théorique et les principaux résultats en EDSRs dans le cadre le plus général, voir Pardoux and Răşcanu [100].

0.1.2 EDSR et EDSPR

Contexte

Une EDSR est un type d'EDS qui évolue en temps inverse, où l'on cherche une paire de processus adaptés solutions d'une EDS avec une condition terminale donnée. Les EDSRs ont été introduites dans un cadre linéaire par Bismut [23] et se sont imposées comme un cadre mathématique puissant pour la modélisation et la résolution de problèmes impliquant des processus stochastiques évoluant en temps inverse. La genèse des EDSRs dans le cadre le plus général peut être retracée jusqu'au travail fondateur de Pardoux and Peng [99], marquant un progrès significatif dans le domaine de l'analyse stochastique. De nombreux résultats théoriques utiles ont depuis été obtenus, par exemple, la connexion avec les équations aux dérivées partielles paraboliques semi-linéaires par Buckdahn et al. [33] qui ont étudié les propriétés d'homogénéisation et Pardoux et al. [101], Darling and Pardoux [48] qui ont étudié la connexion entre les équations aux dérivées partielles elliptiques semi-linéaires et les EDSRs avec temps terminal aléatoire, marquant une contribution importante aux EDSRs avec horizon infini. Briand and Hu [30] ont étudié la stabilité des EDSRs avec un temps terminal aléatoire et ont appliqué ces résultats aux équations aux dérivées partielles elliptiques semi-linéaires. Les principales contributions aux EDSRs comprennent l'existence et l'unicité des solutions, les propriétés de régularité et les connexions avec les équations aux dérivées partielles. El Karoui et al. [58] ont fourni des perspectives fondamentales, établissant un lien entre les EDSRs et les équations aux dérivées partielles quasilineaires. Ma and Yong [89] marquent une contribution importante aux EDSPRs. Les travaux de Soner et al. [112] et de Briand and Confortola [29] ont étendu la compréhension de la stabilité et de la robustesse dans les EDSRs. Pour l'existence et l'unicité des EDSRs avec des générateurs ayant une croissance quadratique, voir Briand and Hu [31] et Delbaen, Hu, and Richou [51]. Ces derniers ont prouvé l'existence et l'unicité pour les EDSRs quadratiques avec un générateur convexe et des conditions terminales non bornées admettant tous les moments exponentiels. Chassagneux and Richou [43] ont développé un schéma numérique pour résoudre les EDSRs avec des générateurs ayant une croissance quadratique dans le processus Z et une condition terminale bornée. Ils ont obtenu une erreur de discrétisation temporelle de l'ordre $\frac{1}{2} - \varepsilon$ pour tout $\varepsilon > 0$.

Applications

La polyvalence des EDSRs est évidente dans leurs applications à travers divers domaines. En finance, les EDSRs trouvent leur utilité dans la valorisation des options, la gestion des risques et l'optimisation de portefeuille. Ces applications naturelles furent introduites dans l'article précurseur de El Karoui, Peng, and Quenez [58] (voir également Delong [52] pour les applications financières des EDSRs à sauts). La caractéristique rétrograde des EDSRs est naturelle dans le contexte financier, où les décisions impactent les résultats futurs. Les EDSRs ont joué un rôle crucial dans la gestion des risques et la valorisation des dérivés en finance. Peng [103] ont introduit les équations différentielles stochastiques rétrogrades comme un outil puissant pour résoudre les espérances non linéaires, offrant une polyvalence pour la valorisation des options et l'évaluation des risques. L'application des EDSRs s'est étendue à la modélisation du risque de crédit, comme le démontrent Crépey and Song [47]. L'avènement des EDSPRs a permis une modélisation plus complète des systèmes financiers dynamiques. Balter and Pelsler [11] ont présenté le cadre FBSDE, offrant une approche unifiée de la valorisation et de la couverture dans les marchés incomplets. Les applications comprennent la valorisation des options américaines (voir Barles [13]) et l'étude des stratégies d'investissement et de consommation optimales (voir Chacko and Viceira [41]). Les EDSPRs de champ moyen représentent un développement récent et puissant pour la modélisation de grandes populations d'agents interagissant sur les marchés financiers. Carmona et al. [39] ont introduit les EDSPRs de champ moyen pour la modélisation de l'impact du marché, démontrant leur potentiel pour capturer les effets de rétroaction des

stratégies de trading sur la dynamique du marché. Les EDSRs ont trouvé des applications dans la modélisation et la valorisation des dérivés d'électricité. Benth and Šaltytė-Benth [19] ont utilisé les EDSRs pour modéliser la dynamique des prix spot sur les marchés de l'électricité, capturant l'impact des facteurs d'offre et de demande.

Aperçu mathématique

Explorons maintenant de manière succincte le formalisme mathématique des EDSRs. Une EDSR standard avec un horizon fini est représentée comme suit :

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad \forall t \in [0, T].$$

où (Y_t, Z_t) sont les processus inconnus, ξ est la condition terminale qui est mesurable par rapport à \mathcal{F}_T , la fonction f est appelée générateur, W_t est un mouvement brownien, et t varie dans l'intervalle de temps $[0, T]$. Voici quelques-uns des résultats les plus cruciaux qui ont posé les bases des EDSRs :

Theorem 0.1.1 (Existence and Uniqueness, see El Karoui et al. [58], Pardoux and Peng [99]). *Pour une EDSR donnée avec condition terminale $\xi \in L^2(\mathcal{F}_T)$ et un générateur $f : [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$, si $t \mapsto f(t, x, y, z)$ est progressivement mesurable et s'il existe une constante $K_f > 0$ telle que, pour tous $dt \otimes d\mathbb{P}$ a.s.,*

$$|f(t, y, z) - f(t, \tilde{y}, \tilde{z})| \leq K_f(|y - \tilde{y}| + |z - \tilde{z}|), \quad y, \tilde{y} \in \mathbb{R} \text{ and } z, \tilde{z} \in \mathbb{R}^d,$$

alors il existe une unique solution adaptée (Y, Z) pour cette EDSR.

Dans le cadre de l'horizon infini, nous pouvons écrire l'EDSR comme suit,

$$Y_t = Y_T + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad \forall 0 \leq t \leq T < +\infty.$$

Nous cherchons l'existence et l'unicité dans l'espace suivant,

$$\mathcal{M}^{2,\rho}(\mathbb{R}^d \times \mathbb{R}^{d \times m}) := \left\{ (Y, Z) \in \mathbb{R}^d \times \mathbb{R}^{d \times m} \mid \mathbb{E} \left[\int_0^\infty e^{-\rho t} (|Y_t|^2 + \|Z_t\|^2) dt \right] < \infty \right\}.$$

Des hypothèses plus fortes sont également nécessaires pour l'existence et l'unicité des solutions. En particulier, en plus des hypothèses ci-dessus, nous avons besoin de la "monotonie" du générateur par rapport au processus Y et d'une condition supplémentaire basée sur ρ , la constante de monotonie (notée μ) et la constante de Lipschitz par rapport à Z (notée $K_{f,z}$). Darling and Pardoux [48] et Briand and Hu [30] ont étudié de manière approfondie l'existence et l'unicité des EDSRs avec un temps terminal aléatoire. Cependant, avec peu de changements, on peut établir les mêmes résultats pour le cadre de l'horizon infini.

Theorem 0.1.2 (Existence et unicité pour l'horizon infini, voir [30, 48]). *En plus des hypothèses du théorème 1.1.1, nous devons satisfaire,*

1. *Monotonie par rapport à Y , c'est-à-dire $\forall (y, y') \in \mathbb{R}^d, \forall z \in \mathbb{R}^{d \times m}$ et $\forall s \geq 0$,*

$$\langle y - y', f(s, y, z) - f(s, y', z) \rangle \leq -\mu |y - y'|^2;$$

2. *$\exists \rho \in \mathbb{R}$ tel que $\rho > K_{f,z}^2 - 2\mu$ et*

$$\mathbb{E} \left[\int_0^\infty e^{\rho t} |f(t, 0, 0)|^2 dt \right] \leq K_f.$$

Sous ces hypothèses, une solution unique existe pour l'EDSR à horizon infini dans l'espace $\mathcal{M}^{2,\rho}(\mathbb{R}^d \times \mathbb{R}^{d \times m})$.

La condition sur le facteur d'actualisation est requise pour le cas multidimensionnel mais peut être assouplie pour le cas unidimensionnel, comme l'ont montré Briand and Hu [30]. Ensuite, énonçons quelques estimations de stabilité a priori pour les EDSRs,

Theorem 0.1.3 (Estimations a priori, voir El Karoui et al. [58]). *Les solutions des EDSRs présentent des estimations de stabilité a priori. Plus précisément, pour deux EDSRs avec des conditions terminales et des générateurs différents, notés ξ_1, ξ_2 et f^1, f^2 respectivement, si les conditions du théorème 0.1.1 sont satisfaites, alors l'estimation suivante est vérifiée :*

$$\begin{aligned} & \mathbb{E} \left[\int_0^T e^{\beta t} |Y_t^1 - Y_t^2|^2 dt \right] + \mathbb{E} \left[\int_0^T e^{\beta t} |Z_t^1 - Z_t^2|^2 dt \right] \\ & \leq C \left(\mathbb{E} [|\xi_1 - \xi_2|^2] + \mathbb{E} \left[\int_0^T e^{\beta t} |f^1(t, Y_t^2, Z_t^2) - f^2(t, Y_t^2, Z_t^2)|^2 dt \right] \right), \end{aligned}$$

où (Y_t^1, Z_t^1) et (Y_t^2, Z_t^2) sont les solutions correspondant à (f^1, ξ_1) et (f^2, ξ_2) respectivement et C est une constante positive.

La constante C ci-dessus dépend du temps terminal et l'on peut rencontrer la difficulté de l'explosion de C lorsque $T \rightarrow \infty$. Cependant, de nombreux résultats de stabilité ont été obtenus pour les EDSRs avec un temps terminal aléatoire et infini, par exemple, par Darling and Pardoux [48] et le résultat de stabilité par Briand and Hu [30]. Ensuite, dans le cadre de l'estimation a priori ci-dessus, énonçons le théorème de comparaison qui a été obtenu pour la première fois par Peng [104].

Theorem 0.1.4 (Principe de comparaison, voir El Karoui et al. [58]). *Etant donnés (f^1, ξ^1) et (f^2, ξ^2) et les solutions correspondantes (Y^1, Z^1) et (Y^2, Z^2) qui sont de carré intégrable, supposons que*

- $\xi^1 \geq \xi^2$ \mathbb{P} p.s.
- $f^1(t, Y^2, Z^2) - f^2(t, Y^2, Z^2) \geq 0$ $\mathbb{P} \otimes dt$ p.s.

Alors, nous avons $Y_t^1 \geq Y_t^2$ $\mathbb{P} \otimes dt$ p.s. .

Ensuite, mettons en évidence les différentes contributions et développements importants dans les méthodes numériques de résolution des EDSRs.

Méthodes numériques

Les schémas numériques pour résoudre les EDSRs ont été étudiés par de nombreux chercheurs au fil des ans. Bally and Pagès [10] ont analysé l'erreur numérique pour la méthode de quantification; Bouchard and Touzi [26] ont étudié les taux de convergence des approximations par régression en utilisant les approches par méthode à noyau et par calcul de Malliavin; Del Moral, Rémillard, and Rubenthaler [50] ont présenté un estimateur préservant la convexité et la monotonie pour les options américaines; Gobet and Labart [66] ont proposé un solveur EDSR où une variable de contrôle adaptative est ajoutée en utilisant la connexion entre les EDSRs linéaires et les EDPs paraboliques linéaires ; Guyon and Henry-Labordère [74] ont fourni une approximation basée sur la régression pour la matrice de covariance optimale dans un modèle de volatilité incertaine. Parmi toutes ces différentes approches pour appliquer les méthodes de Monte-Carlo, la plus populaire est vraisemblablement la régression ordinaire par moindres carrés (OLS en abrégé) qui a été développée par Ma, Protter, San Martín, and Torres [91] (voir aussi [45]). Dans ce cadre, sur un horizon temporel fini, on discrétise d'abord le temps puis on simule des trajectoires pour les processus et on utilise une approche de Monte-Carlo. L'estimateur est ensuite calculé

de manière rétrograde dans la grille de temps comme une combinaison linéaire d'un ensemble réduit de fonctions, trouvant des coefficients qui sont optimaux au sens des moindres carrés empiriques.

Gobet, Lemor, and Warin [70] ont d'abord étendu l'OLS pour les EDSRs standards et ont établi les taux de convergence pour un schéma explicite basé sur la programmation dynamique. Gobet and Lemor [67] ont développé des approximations pour les EDSRs avec des sauts et pour les EDSRs réfléchies ; Bender and Steiner [17] ont utilisé une fonction de base martingale pour réduire l'erreur de propagation lors du calcul des attentes conditionnelles ; Bouchard and Warin [27] ont amélioré l'utilisation d'une base locale adaptative où le support de la base de fonction est adapté à la densité des échantillons ; Bernhart, Pham, Tankov, and Warin [20] ont appliqué cette procédure pour approcher la solution pénalisée des EDSRs avec saut contraint et ont établi un taux de convergence logarithmique ; Gobet et ses co-auteurs ont analysé l'erreur de nombreux estimateurs numériques différents pour résoudre des EDSRs : le schéma explicite multi-étapes a été étudié dans Gobet and Turkedjiev [69], l'échantillonnage par stratification avec une parallélisation utilisant le GPU a été considéré pour améliorer les performances de calcul dans Gobet et al. [71], la stratégie de couverture avec un poids de Malliavin a été présentée dans Gobet and Turkedjiev [68].

Au fil des années, les capacités de calcul ont progressé à un point tel qu'elles ont permis aux chercheurs d'étudier des schémas basés sur des méthodes d'apprentissage profond pour résoudre des équations différentielles stochastiques rétrogrades (EDSR) dans des dimensions très élevées (~ 100), exploitant leur puissance pour atténuer le fléau de la dimensionnalité. E, Han, and Jentzen [56] ont été les premiers à appliquer une méthode qu'ils ont appelée "Deep BSDE" où un schéma progressif est considéré et les paramètres du réseau sont optimisés pour minimiser l'erreur quadratique moyenne sur la condition terminale par descente de gradient stochastique ; ensuite, Han and Long [77] ont étendu l'approche des EDSR profondes pour résoudre un système d'équations différentielles stochastiques rétrogrades couplées et ont démontré ses performances pour un système explicitement soluble dans des dimensions très élevées. Récemment, Carmona and Laurière [38] ont présenté une approche similaire pour un problème de contrôle à champ moyen où l'approximation associée d'une EDSPR spécifique découplée de type McKean-Vlasov a été étudiée, voir aussi Pham and Warin [106] pour une suggestion concernant l'architecture du réseau.

0.1.3 Contrôle champs-moyen

Contexte

Le contrôle à champ moyen (MFC pour Mean Fields Control) et les jeux à champ moyen (MFG pour Mean-field games) sont étroitement liés, le premier étant historiquement issu des MFG. Les MFG ont été introduits par Lasry and Lions [85] et Huang et al. [80], qui ont fourni un puissant outil mathématique pour modéliser les interactions stratégiques d'un grand nombre d'agents rationnels. Ces agents, influençant collectivement la dynamique du système, sont confrontés à un compromis entre des objectifs individuels et collectifs. Le formalisme mathématique des jeux à champ moyen implique l'évolution couplée de la distribution des agents et l'optimisation individuelle de leurs fonctions objectives où leurs actions affectent l'ensemble de la population. Cette nature couplée nécessite des techniques numériques avancées pour une approximation précise et une simulation efficace. Des cadres théoriques, tels que la formulation des jeux à champ moyen par Pardoux et al. [102], ont posé les bases pour comprendre le comportement d'équilibre des grandes populations en présence d'interactions à champ moyen. Leur travail donne un aperçu large des jeux à champ moyen et de leurs applications, offrant des perspectives sur les fondements mathématiques de ces modèles. L'étude a servi de pierre angulaire pour les développements théoriques ultérieurs dans le domaine. Pour une couverture plus large et des

perspectives sur les jeux à champ moyen, voir Bensoussan, Frehse, Yam, et al. [18]. Une étude approfondie des MFGs et des MFCs ainsi que des applications peut être trouvée dans les deux volumes de livres écrits par Carmona and Delarue [37], qui ont sans doute fourni à ce jour l'une des meilleures couvertures détaillées du sujet.

A mesure que les systèmes deviennent de plus en plus complexes et interconnectés, les stratégies de contrôle traditionnelles peuvent être insuffisantes pour capturer le comportement émergent des grandes populations. Ainsi, l'accent mis sur les solutions numériques pour les MFGs et les MFCs en utilisant des EDSPRs de McKean-Vlasov est motivé par la nécessité de combler le fossé entre les formulations théoriques et les mises en œuvre réelles. La motivation de cette recherche réside dans le besoin d'une compréhension plus profonde du contrôle stochastique à champ moyen à horizon infini et de ses applications, notamment en optimisation de portefeuille.

Applications

Les avancées en apprentissage par renforcement profond ont trouvé des applications dans les jeux à champ moyen. Ruthotto et al. [109] ont exploré cette utilisation, développant des algorithmes d'apprentissage accéléré pour les équations de Hamilton–Jacobi–Bellman via des réseaux neuronaux. Ce développement théorique met en évidence les synergies entre l'apprentissage par renforcement et le contrôle à champ moyen, ouvrant la voie à des stratégies de prise de décision plus sophistiquées. L'étude a également posé les bases théoriques pour comprendre les implications et les applications plus larges des modèles de champ moyen dans divers contextes scientifiques. Pour une synthèse sur les jeux à champ moyen appliqués à divers problèmes, voir Gomes and Saúde [73]. Cesaroni et al. [40] ont étendu les outils mathématiques aux Laplaceiens fractionnaires, démontrant l'applicabilité étendue des théories de contrôle à champ moyen. Mertens and Sorin [93] ont également fourni des aperçus sur la stabilité des équilibres. Leur étude a posé les bases pour la recherche ultérieure sur l'analyse de stabilité dans le contexte du contrôle stochastique à champ moyen. Bogachev [25] ont présenté une étude complète sur les équations de Fokker–Planck–Kolmogorov, offrant des résultats théoriques saillants sur l'évolution des distributions de probabilité dans les systèmes à champ moyen. Ce travail a des implications importantes dans la compréhension de la dynamique des systèmes interactifs à grande échelle et leurs propriétés statistiques. Bensoussan et al. [18] couvre plusieurs applications principalement d'un point de vue pédagogique. Cardaliaguet and Lehalle [35] ont développé un cadre pour un jeu à champ moyen et contrôles, appliqué au problème classique de liquidation optimale et de concentration des échanges. Barreiro-Gomez and Tembine [14] ont utilisé un problème de contrôle optimal à champ moyen pour un réseau de distribution d'eau à grande échelle avec non seulement la moyenne mais aussi la variance, à la fois de l'état du système et du processus de contrôle.

Connexion avec les EDSPRs et plus d'applications

L'optimisation des stratégies de stockage d'énergie est une application émergente des EDSPRs. De Paola et al. [49] et Li et al. [87] ont proposé un cadre de EDSPRs pour modéliser le fonctionnement optimal des systèmes de stockage d'énergie, en tenant compte des incertitudes dans les prix de l'électricité et dans la génération d'énergie renouvelable. Ce type d'approche a des implications pour améliorer l'efficacité et la viabilité économique des technologies de stockage d'énergie. Les EDSPRs de champ moyen gagnent en popularité dans la modélisation de la dynamique des systèmes d'alimentation interconnectés. Alasseur et al. [4] ont introduit des EDSPRs de champ moyen pour l'analyse de la stabilité du réseau électrique, en tenant compte des interactions entre plusieurs générateurs et charges. Cette approche facilite une compréhension holistique du comportement du réseau dans diverses conditions de fonctionnement. Les collaborations interdisciplinaires entre mathématiciens, économistes et ingénieurs ont conduit à des grandes avancées dans l'application des méthodologies de contrôle stochastique dans le secteur de

l'énergie. Hafiz et al. [76] ont intégré le contrôle stochastique avec des techniques d'apprentissage automatique pour le fonctionnement optimal en temps réel des systèmes de stockage d'énergie, démontrant le potentiel de combiner diverses méthodologies. Acciaio, Backhoff-Veraguas, and Carmona [1] ont étudié un MFG "étendu" en utilisant le principe du maximum et ont fourni des insights sur la connexion avec la théorie du transport optimal et ont résolu un problème de liquidation optimal. Aïd, Basei, and Pham [3] ont utilisé l'approche de McKean-Vlasov pour modéliser la génération d'énergie distribuée. Les travaux plus récents incluent l'utilisation du contrôle stochastique de McKean-Vlasov pour la gestion de l'énergie d'un réseau électrique intelligent par Gobet and Grangereau [64] et dans [65], ils ont développé de nouvelles approches de résolution combinant la méthode de Newton et le principe de Pontryagin pour le contrôle stochastique. Germain et al. [62] ont utilisé des dynamiques de champ moyen avec des contraintes d'état probabilistes et une approche par courbe de niveau pour le problème de stockage d'énergie renouvelable.

Aperçu mathématique

Définissons maintenant clairement le problème général du contrôle à champ moyen avec dépendance à la loi de l'état mais pas du contrôle. Le MFC avec dépendance à la loi de contrôle pourrait être étudié avec la même approche (voir [1] par exemple). Similairement à la formulation du problème de contrôle stochastique, considérons une EDS contrôlée $(X_t^u)_{t \geq 0}$ avec le processus de contrôle $(u_t)_{t \geq 0}$,

$$dX_t^u = b(X_t^u, u_t, \mu_t) dt + \sum_{j=1}^m \sigma^j(X_t^u, u_t, \mu_t) dW_t^j, \quad X_{t_0} = x_0,$$

où μ_t désigne la loi de l'état X_t^u . Avec cette dépendance à la loi, c'est une EDS dite de McKean-Vlasov qui peut être vue comme la limite champ moyen pour un grand nombre de joueurs interagissant dans un MFG (voir [37] pour plus de détails). Lorsque la dépendance à la loi se fait par une fonction de l'espérance de l'état, on parle d'"interaction scalaire", qui est plus facile à traiter que le cas général mais reste applicable à de nombreux problèmes (voir [64, 1] par exemple). Le cadre général nécessite une notion de différentiabilité et une notion appropriée de convexité dans l'espace des mesures. Voir le chapitre 5 de [37] pour une discussion approfondie sur le sujet. Nous utiliserons uniquement les notions de "L-différentiabilité" et de "L-convexité" car elles sont les mieux adaptées aux problèmes dans le cadre probabiliste, ce qui est idéal pour nous car nous utilisons des EDSPRs pour résoudre le MFC. En général, non seulement la dynamique de l'état mais aussi la fonction objective peut dépendre de la loi de l'état:

$$J^T(x_0, u) = \mathbb{E} \left[\int_0^T f(X_t^{u, x_0}, u_t, \mu_t) dt + g(X_T^{x_0, u}, \mu_T) \right],$$

Dans cette thèse, nous utiliserons uniquement le principe du maximum de Pontryagin pour résoudre le MFC. Cependant, la programmation dynamique a été utilisée par certains chercheurs, par exemple par Tembine, Zhu, and Başar [114] et Zhou, Touzi, and Zhang [117]. Lorsque l'on utilise le principe du maximum, le problème de recherche du contrôle optimal dans l'espace infini-dimensionnel des fonctions (espace des contrôles admissibles) peut être converti en la résolution d'un système couplé de EDSPRs de McKean-Vlasov où le contrôle optimal est donné sous la forme d'une fonction des processus adjoints (Y et Z) en utilisant les conditions nécessaires d'optimalité. Un contrôle optimal u_t doit satisfaire,

$$\nabla_u H(t, X_t^u, \mu_t, u, Y_t, Z_t) = 0 \quad dt \otimes d\mathbb{P}p.s.$$

Notons le contrôle optimal $u^*(X_t, Y_t, Z_t)$ simplement par u_t^* où la dépendance des EDSPRs (X, Y, Z) est implicite. De plus, utilisons θ_t comme un raccourci pour un groupe de variables où

la dépendance est claire. Ensuite, le système couplé de EDSRs de McKean-Vlasov est donné par,

$$\begin{aligned} dX_t &= b(X_t, u_t^*, \mu_t) dt + \sum_{j=1}^m \sigma^j(X_t, u_t^*, \mu_t) dW_t^j, \quad X_{t=0} = x_0 \\ dY_t &= \left(Y_t \partial_x b(\theta_t) + \sum_{j=1}^m Z_t^j \partial_x \sigma^j(\theta_t) + \partial_x f(\theta_t) + \partial_\mu f(\theta_t)(X_t) \right) dt - \sum_{j=1}^m Z_t^j dW_t^j, \\ Y_{t=T} &= \xi \in \mathcal{F}_T, \end{aligned}$$

où $\partial_\mu f(\theta)(X_t)$ désigne la L-dérivée par rapport à la loi de l'état, évaluée en X_t . Voir Carmona and Delarue [37] pour une étude détaillée dans le cadre de l'horizon fini. Ensuite, nous discutons des contributions importantes dans la résolution numérique du problème.

Schémas numériques pour les EDSRs de McKean-Vlasov

L'approximation numérique des EDSRs de McKean-Vlasov reste un aspect difficile mais crucial pour comprendre la dynamique des systèmes de champ moyen. De nombreux schémas numériques ont été proposés, chacun adapté pour aborder des complexités spécifiques inhérentes à ces équations. Achdou and Capuzzo-Dolcetta [2] ont posé les bases des méthodes numériques dans les jeux à champ moyen en introduisant des approximations par différences finies. Ce travail fondateur a ouvert des perspectives sur les défis posés par les systèmes de champ moyen et a offert un point de départ pour les recherches ultérieures dans le domaine de l'analyse numérique des modèles de champ moyen. Allant au-delà de l'approche par différences finies, Cardaliaguet and Lehalle [35] ont étendu l'analyse numérique aux jeux à champ moyen avec bruit commun et sauts. Les auteurs ont introduit un schéma numérique robuste capable de gérer les complexités supplémentaires résultant du bruit commun et des processus discontinus. Ce développement a marqué une étape cruciale vers des simulations plus réalistes des systèmes à champ moyen. Les techniques d'apprentissage profond ont été de plus en plus intégrées à l'approximation numérique des modèles à champ moyen. L'approche probabiliste pour les jeux à champ moyen avec des agents principaux et mineurs, présentée par Buckdahn et al. [34], a contribué à la compréhension de la dynamique des agents hétérogènes. Les analyses numériques associées ont fourni des perspectives précieuses sur les défis et les opportunités posés par les caractéristiques diverses des agents dans les systèmes de champ moyen.

Les équations différentielles stochastiques rétrogrades de champ moyen (EDSRMs) ont été au centre des développements numériques récents. Djehiche et al. [54] ont abordé les méthodes numériques pour les EDSRMs, étendant le champ d'application à divers domaines. Ce travail a élargi la boîte à outils numériques disponibles pour les chercheurs travaillant avec des composantes rétrogrades dans les systèmes de champ moyen. Bertucci [22] ont étendu l'analyse numérique aux jeux à champ moyen avec arrêt optimal, introduisant un nouveau cadre pour étudier les problèmes d'arrêt optimal dans un contexte de champ moyen. Ce travail élargit l'applicabilité des schémas numériques pour englober un éventail plus large de scénarios de prise de décision, contribuant à la polyvalence des modèles de champ moyen. Ruthotto et al. [109] ont exploré l'utilisation de l'apprentissage par renforcement profond pour les jeux à champ moyen, développant des algorithmes d'apprentissage accélérés pour les équations de Hamilton–Jacobi–Bellman via des réseaux neuronaux. Ce développement théorique met en évidence les synergies entre l'apprentissage par renforcement et le contrôle à champ moyen, ouvrant la voie à des stratégies de prise de décision plus sophistiquées. Comme mentionné précédemment, avec la méthode dite "Deep BSDE" par [56], de nombreuses utilisations ont été trouvées pour résoudre approximativement les EDSRs de McKean-Vlasov couplés. Han and Long [77] ont développé un algorithme utilisant l'approche du jeu fictif de la théorie des jeux. Quant à eux, Germain, Mikael,

and Warin [63] ont proposé un algorithme avec une mise à jour dynamique de la loi empirique du processus d'état, comparant le taux de convergence de différentes approches. Notre approche utilise ces développements récents dans l'utilisation de réseaux neuronaux pour approcher les EDSRs de McKean-Vlasov afin de résoudre un problème d'optimisation de portefeuille où la dépendance de distribution dans la fonction de coût vient de l'appétit pour le risque de l'agent.

0.2 Contributions et défis

0.2.1 Contrôle stochastique de champ moyen à horizon infini et applications

Problème général

Dans le deuxième chapitre de la thèse, nous étudions d'abord les problèmes de contrôle à champ moyen (MFC) à horizon infini avec la distribution de l'état apparaissant dans la dynamique de l'état et la fonction de coût, mais sans dépendance par rapport à la distribution du contrôle. Nous étudions d'abord les conditions nécessaires d'optimalité qui caractérisent le contrôle en fonction des variables couplées de l'EDSPR, puis les conditions de convexité nécessaires pour prouver l'unicité et les conditions d'optimalité suffisantes. Des résultats analogues ont été bien établis pour les paramètres de horizon fini dans le travail fondateur par Carmona and Delarue [37]. L'approche est basée sur le principe du maximum de Pontryagin. Nous définissons d'abord le problème de contrôle,

$$\begin{aligned} dX_t^u &= b(X_t^u, u_t)dt + \sum_{j=1}^m \sigma^j(X_t^u, u_t)dW_t^j, \quad X_0 = x_0; \\ J(u) &= \mathbb{E} \left[\int_0^\infty e^{-\rho t} (f(X_t^u, u_t, \mu_{X_t^u})) dt \right] \end{aligned}$$

Définissons le Hamiltonien comme suit,

$$H(x, u, \mu, y, z) = y b(x, u) + \sum_{j=1}^m z^j \sigma^j(x, u) + f(x, u, \mu) - \rho y x.$$

Les processus adjoints $(Y_t, Z_t)_{t \geq 0}$ satisfont les équations suivantes,

$$\begin{aligned} dY_t &= \left(Y_t \partial_x b(\theta_t) + \sum_{j=1}^m Z_t^j \partial_x \sigma^j(\theta_t) + \partial_x f(\theta_t) + \partial_\mu f(\theta_t)(X_t^u) - \rho Y_t \right) dt - \sum_{j=1}^m Z_t^j dW_t^j, \\ \lim_{T \rightarrow \infty} \mathbb{E} [e^{-\rho T} \|Y_T\|^2] &= 0 \end{aligned}$$

Ensuite, la condition nécessaire d'optimalité pour le MFC à horizon infini ci-dessus est donnée par,

Theorem 0.2.1 (Conditions nécessaires, voir Théorème (2.3.1)). *Sous certaines hypothèses, lorsque $u \in \mathcal{A}$ est un contrôle optimal et que X est la variable d'état correspondante et que (Y, Z) sont les processus adjoints correspondants comme défini ci-dessus dans (3.3.1), alors nous devons avoir*

$$\partial_u H(X_t^u, u_t, \mu_{X_t^u}, Y_t, Z_t) \cdot (u_t - \nu) \leq 0, \quad dt \otimes d\mathbb{P} \text{ p.p. } \forall \nu \in U.$$

De plus, si le contrôle optimal u appartient à l'intérieur de U $dt \otimes d\mathbb{P}$ p.p., alors

$$\partial_u H(X_t^u, u_t, \mu_{X_t^u}, Y_t, Z_t) = 0.$$

Le principe du maximum est très puissant pour le MFC car l'approche probabiliste basée sur les EDSRs permet une résolution numérique efficace dans des dimensions supérieures, il fournit une

représentation explicite du contrôle optimal en termes de processus adjoints, grâce aux conditions nécessaires ci-dessus. Ensuite, nous énonçons la condition suffisante d'optimalité basée sur la forte L-convexité.

Theorem 0.2.2 (Conditions nécessaires, voir Théorème (2.4.1)). *Sous certaines hypothèses, soit $u \in \mathcal{A}$ un contrôle admissible, X_t^u le processus d'état correspondant et (Y, Z) les processus adjoints correspondants. De plus, supposons que*

$$\mathbb{R}^d \times U \times \mathcal{P}^2(\mathbb{R}^n) \ni (x, u, \mu) \mapsto H(t, x, u, \mu, Y_t, Z_t) \text{ est convexe } dt \otimes d\mathbb{P} \text{ p.p..}$$

Si

$$H(t, X_t^u, u_t, \mu_{X_t^u}, Y_t, Z_t) = \inf_{u \in U} H(t, X_t^u, u, \mu_{X_t^u}, Y_t, Z_t),$$

$dt \otimes d\mathbb{P}$ p.p. alors u est un contrôle optimal.

La condition de convexité ci-dessus doit être comprise comme suit,

$$\begin{aligned} H(t, x', u', \mu', Y_t, Z_t) &\geq H(t, x, u, \mu, Y_t, Z_t) + \partial_x H(t, x, u, \mu, Y_t, Z_t)(x' - x) \\ &\quad + \partial_u H(t, x, u, \mu, Y_t, Z_t)(u' - u) + \tilde{\mathbb{E}} \left[\partial_\mu H(t, x, u, \mu, Y_t, Z_t)(\tilde{X})(\tilde{X}' - \tilde{X}) \right]. \end{aligned}$$

De plus, sous la condition de forte L-convexité,

$$\begin{aligned} H(t, x', u', \mu', Y_t, Z_t) &\geq H(t, x, u, \mu, Y_t, Z_t) + \partial_x H(t, x, u, \mu, Y_t, Z_t)(x' - x) + \alpha |q' - q|^2 \\ &\quad + \partial_u H(t, x, u, \mu, Y_t, Z_t)(u' - u) + \tilde{\mathbb{E}} \left[\partial_\mu H(t, x, u, \mu, Y_t, Z_t)(\tilde{X})(\tilde{X}' - \tilde{X}) \right], \end{aligned}$$

avec $\alpha > 0$, on peut garantir l'unicité du contrôle optimal.

Nous avons également considéré diverses fonctions de dépendance en la distribution qui pourraient être étudiées en plus des interactions scalaires. Nous avons considéré les fonctions suivantes :

1. Distance de Wasserstein au carré entre la distribution de l'état et une distribution cible π^* ,

$$W_2^2(\mu_{X_t^u}, \pi^*) = \inf_{\gamma \in \Pi(\mu_{X_t^u}, \pi^*)} \left\{ \int_{\mathbb{R} \times \mathbb{R}} |x - y|^2 d\gamma(x, y) \right\}.$$

La L-différentiabilité a été démontrée dans un article récent par Alfonsi et Jourdain [5] (voir également Gangbo and Tudorascu [61]) sous la condition qu'il existe un couplage optimal unique entre $\mu_{X_t^u}$ et π^* qui est donné par un plan de transport. Cependant, la L-convexité n'est généralement pas vraie, comme démontré par un contre-exemple dans le chapitre 5, [37]. Cependant, elle est essentielle pour prouver les conditions suffisantes d'optimalité.

2. Divergence de Kullback-Leibler,

$$d_{KL}(\mu|\nu) := \int_{\mathbb{R}^d} \frac{d\mu(x)}{d\nu} \log \left(\frac{d\mu(x)}{d\nu} \right) d\nu(x), \text{ si } \mu \ll \nu \text{ et } +\infty \text{ sinon.}$$

En raison de l'exigence de continuité absolue de la mesure du processus d'état, elle n'est pas directement applicable.

3. Distance d'énergie (voir Rizzo and Székely [108]),

$$D_E(\mu, \nu) = \int_{\mathbb{R}^d} \left| \mathbb{E} \left[e^{i\lambda X} \right] - \mathbb{E} \left[e^{i\lambda Y} \right] \right|^2 w(\lambda) d\lambda.$$

La L-différentiabilité et la L-convexité n'étaient pas claires pour cet exemple.

4. Valeur en risque conditionnelle (CVaR) ou perte attendue (ES),

$$\text{ES}_\alpha(X) = -\frac{1}{\alpha} \int_0^\alpha \text{VaR}_\gamma(X) d\gamma$$

ce qui est difficile à analyser du point de vue de la L-différentiabilité sans connaître la densité du processus d'état a priori, inconnue ici car il s'agit d'un processus contrôlé.

Le manque d'exemples de travail conventionnels a posé un défi majeur dans la thèse lorsqu'il s'agissait d'appliquer l'approche du principe du maximum pour le MFC, nous conduisant à concevoir une dépendance spécifique en la mesure qui ne soit pas des interactions scalaires, et qui satisfasse les conditions de L-différentiabilité et de L-convexité mais aussi soit applicable. Cela nous conduit au chapitre 3 de la thèse.

Problème financier

Le troisième chapitre est dédié à un problème d'optimisation de portefeuille d'actions, à long terme, modélisé comme un problème MFC. Une application pratique de cette approche est également en discussion avec l'équipe de gestion d'actifs de BNP Paribas dans le cadre de la Chaire StressTest (projet de mécénat entre l'Ecole polytechnique, sa fondation, et BNP Paribas). Nous commençons avec un ensemble de N actifs risqués avec des prix $(S^i)_{i=1,\dots,N}$ de type mouvement brownien géométrique (MBG) avec des paramètres b_t^i et $(\sigma_t^{i,j})_{j=1,\dots,m}$ et un actif sans risque S^0 satisfaisant $dS_t^0 = S_t^0 r_t dt$, où r_t est le taux de rendement sans risque. En général, nous pourrions prendre des paramètres stochastiques mais pour les simulations numériques, nous avons choisi des b , σ et r constants. Le processus de richesse résultant pour le portefeuille contrôlé est ensuite obtenu en utilisant la condition d'autofinancement qui exige que l'agent ne retire pas ou ne dépose pas d'argent dans le portefeuille après $t = 0$,

$$dV_t^u = r_t V_t^u dt + \sum_{i=1}^N u_t^i \left((b_t^i - r_t) dt + \sum_{j=1}^m \sigma_t^{i,j} dW_t^j \right).$$

Nous considérons également un portefeuille de référence que nous voulons surpasser. Ce portefeuille, appelé "portefeuille de référence" dorénavant, a des allocations statiques qui ne sont pas modifiées après $t = 0$. Ces allocations sont décrites par $(\gamma^i)_{i=0,\dots,N}$. Le contrôle q_t est la vitesse de transaction des actifs individuels et une partie du processus d'état est $\mathcal{V}_t^q = V_t / \bar{V}_t$ qui est le processus de richesse correspondant au portefeuille contrôlé considéré dans le numéraire du portefeuille de référence. Le point de départ de \mathcal{V}_t^q est donc 1 car nous commençons avec la même richesse pour les deux portefeuilles. Le processus d'état complet est alors décrit par $X_t^q = [\mathcal{V}_t^q \ \delta_t^1 \ \dots \ \delta_t^N]^\top$. On peut montrer que la dynamique d'état est donnée par

$$\begin{aligned} d\mathcal{V}_t^q &= (\alpha_t^\mathcal{V} \mathcal{V}_t^q + \alpha_t^\delta \cdot \delta_t) dt + \sum_{j=1}^m (\beta_t^{\mathcal{V},j} \mathcal{V}_t^q + \beta_t^{\delta,j} \cdot \delta_t) dW_t^j, \\ d\delta_t^i &= q_t^i dt, \quad \delta_{t=0}^i = 0, \quad i \in \{1, 2, \dots, N\}, \end{aligned}$$

donnant une EDS pour \mathcal{V}_t^q dont la dérive et la volatilité sont stochastiques mais bornées et linéaires en \mathcal{V}^q et δ . Les coefficients $\alpha_t^\mathcal{V}$, α_t^δ , $\beta_t^{\mathcal{V},j}$, $\beta_t^{\delta,j}$ donnés dans (3.2.1) sont des fonctions des prix relatifs des actifs $S_t^i = S_t^i / \bar{V}_t$ et d'autres paramètres du modèle. Nous recherchons la vitesse de transaction optimale pour chaque actif qui minimise une fonction coût composée d'un coût quadratique dans le contrôle et d'un terme dépendant de la distribution cible et de la distribution du processus de richesse relatif. Le premier terme nous aide à minimiser le coût et l'impact de nos transactions. La distribution cible non stationnaire est choisie comme la distribution d'un mouvement brownien géométrique avec des paramètres \tilde{b} et $\tilde{\sigma}$, en partant de 1 ce qui impose

des conditions de croissance idéales pour le processus de richesse associé. La dépendance en la distribution pénalise les écarts par rapport à la distribution cible au sens du risque de baisse. La fonction coût est donnée par

$$J^\lambda(q) = \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(g(X_t^q, q_t) + \int_0^\infty \left[\mathbb{E} [k - \mathcal{V}_t^q]_+^\lambda - p_k \right]_+^\lambda d\nu(k) \right) dt \right]$$

où les paramètres p_k proviennent de la distribution cible, λ est un paramètre de lissage pour la fonction $[k - x]_+$ et ν attribue des poids au coût pour chaque niveau de richesse k . Pour le schéma numérique, nous avons pris la fonction g comme étant $\frac{1}{2}|q_t|^2$ mais en général on peut prendre n'importe quelle fonction lisse qui est convexe dans le premier argument et fortement convexe dans le deuxième argument. Nous avons montré dans le Chapitre 3 Proposition 3.2.3 que la fonction ci-dessus avec variable "mesure", désignée par $h^\lambda(\mu_{\mathcal{V}_t^q})$, est L-différentiable et L-convexe. Nous avons également établi qu'elle est Lipschitz-continue par rapport à la mesure du processus de richesse associée (dans un certain sens) (voir Proposition 3.2.4).

Ensuite, le système d'EDSR en horizon infini pour un contrôle admissible donné $(q_t)_{t \geq 0} \in \mathcal{A}$ associé au problème de contrôle est donné par,

$$\begin{aligned} dY_t^1 &= - \left(Y_t^1 \alpha_t^\mathcal{V} + \sum_{j=1}^m Z_t^{j,1} \beta_t^{\mathcal{V},j} + \partial_\mu h(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) - \rho Y_t^1 \right) dt + \sum_{j=1}^m Z_t^{j,1} dW_t^j \\ dY_t^i &= - \left(Y_t^i \alpha_t^{\delta,i} + \sum_{j=1}^m Z_t^{j,i} \beta_t^{\delta,j,i} - \rho Y_t^i \right) dt + \sum_{j=1}^m Z_t^{j,i} dW_t^j, \quad i \in \{2, 3, \dots, N+1\}, \\ \lim_{T \rightarrow \infty} \mathbb{E} [e^{-\rho T} |Y_T^i|^2] &= 0. \end{aligned}$$

Theorem 0.2.3 (Voir Théorème 3.3.2). *Pour un contrôle admissible donné $(q_t)_{t \geq 0}$, sous certaines hypothèses, l'EDSR numéro (3.3.1) admet une solution unique $(Y_t, Z_t)_{t \geq 0}$ dans l'espace $\mathcal{S}_\rho^2(\mathbb{R}^{1 \times (N+1)}) \times L_\rho^2(\mathbb{R}^{m \times (N+1)})$. De plus, il existe $C > 0$ tel que pour deux paires (Y', Z') et (Y, Z) de solutions aux EDSR correspondant aux paires état-contrôle (X', q') et (X, q) respectivement, l'estimation suivante est vérifiée :*

$$\mathbb{E} \left[\int_0^\infty e^{-\rho t} |Y_t' - Y_t|^2 dt \right] + \mathbb{E} \left[\int_0^\infty e^{-\rho t} |Z_t' - Z_t|^2 dt \right] \leq C \mathbb{E} \left[\int_0^\infty e^{-\rho t} |q_t' - q_t|^2 dt \right].$$

Les hypothèses pour l'existence et l'unicité de l'EDSR ci-dessus imposent une borne inférieure pour le facteur d'actualisation du problème de contrôle lorsque nous étudions les conditions d'optimalité.

Theorem 0.2.4 (Conditions nécessaires, voir Théorème 3.3.3). *Supposons que certaines hypothèses sont satisfaites et que $q \in \mathcal{A}$ est optimal. Soit X la variable d'état correspondante et soient (Y, Z) les processus adjoints correspondants tels que définis ci-dessus dans (3.3.1), alors nous devons avoir*

$$\partial_q H(X_t^q, q_t, \mu_{\mathcal{V}_t^q}, Y_t, Z_t) = 0, \quad dt \otimes d\mathbb{P} \text{ p.p.}$$

En utilisant cela, nous pouvons prouver que le contrôle optimal est donné par $q_t^* = -(Y_t^{[2:(N+1)]})^\top$. Pour prouver les conditions suffisantes pour l'optimalité, nous avons utilisé une approche similaire à [37] mais adaptée au cadre d'horizon infini.

Theorem 0.2.5 (Conditions suffisantes, voir 3.3.4). *Supposons que certaines hypothèses sont satisfaites. Soit $q \in \mathcal{A}$ un contrôle admissible, X_t^q le processus d'état correspondant et (Y, Z) les processus adjoints correspondants. Supposons également que le Hamiltonien satisfait :*

$\mathbb{R}^{N+1} \times \mathbb{R}^N \times \mathcal{P}^2(\mathbb{R}) \ni (x, q, \mu) \mapsto H(t, x, q, \mu, Y_t, Z_t)$ est α -fortement convexe $dt \otimes d\mathbb{P}$ p.p., ce qui doit être compris comme suit,

$$\begin{aligned} H(t, x', q', \mu', Y_t, Z_t) &\geq H(t, x, q, \mu, Y_t, Z_t) + \partial_x H(t, x, q, \mu, Y_t, Z_t)(x' - x) \\ &\quad + \partial_q H(t, x, q, \mu, Y_t, Z_t)(q' - q) + \mathbb{E} [\partial_\mu h(\mu)(X)(X' - X)] + \alpha|q' - q|^2 \end{aligned}$$

où $X \sim \mu$ et $X' \sim \mu'$ et $\alpha > 0$. Si

$$H(t, X_t^q, q_t, \mu_{X_t^q}, Y_t, Z_t) = \inf_{u \in \mathbb{R}^N} H(t, X_t^q, u, \mu_{X_t^q}, Y_t, Z_t), \quad dt \otimes d\mathbb{P} \text{ p.p.}$$

alors q est le contrôle optimal unique dans \mathcal{A} .

On peut prouver que le candidat pour le contrôle q^* à partir des conditions nécessaires satisfait également les conditions suffisantes pour l'optimalité. Cela nous donne un système unique d'EDSPR de McKean-Vlasov couplées correspondant à notre problème de contrôle lorsque nous substituons $q_t^* = -(Y_t^{[2:(N+1)]})^\top$ dans la dynamique du processus d'état. L'existence et l'unicité de cette EDSR de McKean-Vlasov sont établies dans le résultat suivant :

Theorem 0.2.6 (Existence et unicité de l'EDSPR, voir Théorème 3.3.6). *Lorsque certaines hypothèses sont satisfaites, l'EDSPR numéro (3.3.4) a une solution unique $(X_t, Y_t, Z_t)_{t \geq 0}$ dans l'espace $L_\rho^2(\mathbb{R}^{N+1}) \times L_\rho^2(\mathbb{R}^{1 \times (N+1)}) \times L_\rho^2(\mathbb{R}^{m \times (N+1)})$.*

Schéma numérique

A des fins numériques, nous considérons une approximation à temps fini de notre méthode pour un grand $T > 0$ et fixons $Y_T = 0$ comme condition terminale et $Y_t = 0$ pour $t > T$. Nous utilisons également une troncature dans le temps et fournissons une certaine estimation de l'erreur d'approximation de troncature dans la Proposition 3.3.7. Pour traiter la dépendance de la distribution, nous utilisons l'approche de [78]. Nous utilisons des itérations de jeu fictif où chaque itération a une dépendance prédéfinie sous la forme d'une fonction du temps. Nous résolvons les EDSR couplées résultantes en considérant une représentation découplée pour (Y, Z) en utilisant des réseaux neuronaux à propagation avant comme approximations de Y_0 et $(Z_{t_i})_{i=0, \dots, N}$ à chaque pas de temps qui se nourrit du pas de temps précédent, créant ainsi un réseau profond de sous-réseaux. Ce type de réseau a également été utilisé par [77] pour les EDSR couplées. L'idée principale du schéma est de résoudre les équations progressives dans le temps et à la fin, de pénaliser les écarts par rapport à la condition terminale et d'entraîner le réseau neuronal pour minimiser ces écarts. Pour plus de détails sur une architecture plus simple du réseau neuronal pour les EDS, voir [56]. Ensuite, nous décrivons précisément le schéma. Les processus de prix des actifs peuvent être écrits explicitement comme suit,

$$\begin{aligned} \mathbf{s}_{t_k}^i &= \mathbf{s}_0^i \exp \left(\left(b^i - \frac{\sum_{j=1}^m (\sigma^{i,j})^2}{2} \right) t_k + \sum_{j=1}^m \sigma^{i,j} W_{t_k}^{a,j} \right), \\ \mathbf{s}_{t_k}^0 &= \mathbf{s}_0^0 e^{rt_k}. \end{aligned}$$

Puis, le processus de prix relatif de l'actif S^i est donné par, $S_{t_k}^i = \frac{\mathbf{s}_{t_k}^i}{\sum_{i=0}^N \gamma^i \mathbf{s}_{t_k}^i}$.

Nous prenons la distribution cible π_t^* pour le processus de richesse relatif \mathcal{V}_t^q au temps t comme la distribution du mouvement brownien géométrique suivant: $dP_t^* = P_t^*(\mu dt + \sigma dW_t)$ avec $P_0^* = 1, \mu = \sigma = 0.05$. Cela définit à la fois le taux de croissance souhaité et la variation autorisée du processus de richesse relatif une fois que nous calculons les paramètres $p_k(t)$. Ceux-ci peuvent

être calculés comme suit : Lorsque \mathcal{V}_t^q atteint la distribution cible, $\left[\mathbb{E}[k - P_t^*]_+^\lambda - p_k(t)\right]_+^\lambda = 0$, ce qui signifie que $p_k(t)$ doit être supérieur ou égal à $\mathbb{E}[k - P_t^*]_+^\lambda$. Nous pouvons prendre

$$p_k(t) = \mathbb{E}[k - P_t^*]_+ = e^{\mu t} \mathbb{E}\left[k e^{-\mu t} - e^{\sigma W_t - \frac{1}{2}\sigma^2 t}\right]_+,$$

qui peut être résolu explicitement en utilisant la formule de Black-Scholes. Avec cela, nous pouvons écrire la L-dérivée comme une fonction de la variable d'état et de l'espérance d'une fonction de l'état,

$$\partial_\mu h(\mu)(\mathcal{V}_t^q) = \int_0^\infty \frac{-1}{1 + e^{\lambda p_k - (\mathbb{E}[\log(1 + e^{\lambda(k - \mathcal{V}_t^q)])])}} \frac{1}{1 + e^{\lambda(\mathcal{V}_t^q - k)}} d\nu(k) \equiv H(\mathcal{V}_t^q, \mathbb{E}[\theta(\mathcal{V}_t^q)]),$$

où $\theta(x) = \log\left(1 + e^{\lambda(k - \mathcal{V}_t^q)}\right)$.

Pour résoudre les EDSR couplées, nous utilisons des champs de découplage en tant que fonctions de l'état et du processus de prix de l'actif associé: $Y_0 = u(X_0, S_0)$; $Z_t = v_t(X_t, S_t, Y_t)$. Le schéma est décrit comme suit (avec p représentant l'itération du jeu fictif) :

1. Initialiser $(\theta^0(t_i))_{i \in \{1, 2, \dots, N\}}$ comme la fonction de prédiction pour $\mathbb{E}[\theta(\mathcal{V}_t)]$. Initialiser les réseaux neuronaux $u^0(x, s)$ et $\{v_{t_i}^0(x, s, y)\}_{i \in \{1, 2, \dots, N+1\}}$ avec des poids aléatoires.
2. Trouver la solution correspondante $(X_{t_i}^0, Y_{t_i}^0, Z_{t_i}^0)_{0 \leq i \leq N}$ par le schéma d'Euler explicite pour les EDSR résultantes, qui est donné comme suit pour toute itération p ,

$$\begin{aligned} X_0^p &= [1, 0, \dots, 0]; \quad Y_0^p = u^p(X_0, S_0); \quad Z_0^p = v_0^p(X_0^p, S_0, Y_0^p); \\ X_{t_{i+1}}^p &= X_{t_i}^p + b(t_i, X_{t_i}^p, Y_{t_i}^p) + \sigma(t_i, X_{t_i}^p) \Delta W_i; \\ Z_{t_i}^p &= v_{t_i}^p(X_{t_i}^p, S_{t_i}, Y_{t_i}^p); \\ Y_{t_{i+1}}^p &= Y_{t_i}^p - f(t_i, X_{t_i}^p, \theta^{p-1}(t_i), Y_{t_i}^p, Z_{t_i}^p) + Z_{t_i}^p \Delta W_i. \end{aligned}$$

3. Entraîner les réseaux neuronaux avec ces données pour minimiser

$$\mathbb{E}[|Y_T^0|^2] \approx \frac{1}{D} \sum_{a=1}^D |Y_{t_N}^{0,a}|^2,$$

puisque la condition terminale de notre problème tronqué est $Y_T = 0$. Cela nous donne les prochaines itérations u^1 et $(v_{t_i}^1)_{1 \leq i \leq N}$ de l'approximation des champs de découplage.

Mettre à jour l'estimation de $\theta^p(t_i)$ avec $\theta^p(t_i) = \frac{1}{D} \sum_{a=1}^D \theta(\mathcal{V}_{t_i}^{p-1})$.

4. Résoudre l'EDSR correspondant en utilisant le schéma d'Euler comme précédemment et itérer de manière similaire.

Nous avons réussi à démontrer que la fonction de répartition tend à se décaler vers la droite lorsque nous augmentons un multiplicateur constant pour le terme champ moyen, augmentant ainsi la richesse moyenne atteinte tout au long de l'horizon temporel et augmentant également la probabilité d'une richesse plus élevée ($\mathcal{V}_t \geq 1$) par rapport au point de départ sans augmenter considérablement la probabilité d'une richesse plus faible ($\mathcal{V}_t \leq 1$). Nous avons également montré que le terme champ moyen, qui peut être vu comme l'écart par rapport à la distribution cible, diminue avec le temps de manière cohérente tout au long de nos tests.

Nous avons rencontré deux défis majeurs dans la partie numérique : 1. L'instabilité numérique du schéma lorsque le paramètre de lissage λ , qui est important pour une meilleure approximation de la L-dérivée, est trop grand; 2. L'instabilité numérique par rapport au coefficient multiplicatif devant le terme de champ moyen.

0.2.2 EDSR en horizon infini

Dans la deuxième partie de la thèse, notre attention se porte sur le développement d'algorithmes efficaces pour approximer numériquement les EDSR markoviennes en horizon infini, ce qui équivaut à résoudre des EDP elliptiques comme mentionné précédemment. L'EDSR est donné par

$$Y_t = Y_T + \int_t^T f(X_s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad \forall 0 \leq t \leq T < +\infty,$$

où le processus solution (Y, Z) prend des valeurs dans un espace approprié et X est la solution de l'EDS suivante :

$$X_t = x + \int_0^t b(X_s) ds + \int_0^t \sigma(X_s) dW_s, \quad \forall t \geq 0.$$

Nous rappelons le résultat général d'existence et d'unicité pour cet EDSR en horizon infini, pour lequel nous renvoyons au Théorème 5.57 de [100] pour le cas multidimensionnel et au Lemme 3.1 de [30] pour le cas unidimensionnel.

Proposition 0.2.1. (Voir Proposition 4.2.1) Supposons que l'Hypothèse 4.2.1 soit vérifiée. On suppose que f est $-\mu$ monotone par rapport à la variable y , est $K_{f,z}$ Lipschitz par rapport à la variable z et a une croissance polynomiale de puissance au plus r par rapport à la variable x .

1. Si $0 < \rho < 2\mu - K_{f,z}^2$ tel que, pour tout $x \in \mathbb{R}^d$,

$$\mathbb{E} \left[\int_0^{+\infty} e^{-\rho s} |X_s^x|^{2r} ds \right] < +\infty,$$

alors, pour tout $x \in \mathbb{R}^d$, l'EDSR (4.2.1) a une solution unique $(Y^x, Z^x) \in \mathcal{S}_{loc}^2 \times \mathcal{M}_{loc}^2$ telle que

$$\lim_{T \rightarrow +\infty} \mathbb{E} [e^{-\rho T} |Y_T^x|^2] = 0.$$

De plus, cette solution satisfait

$$\mathbb{E} \left[\int_0^{+\infty} e^{-\rho s} \|Z_s^x\|^2 ds \right] + \mathbb{E} \left[\sup_{0 \leq s < +\infty} e^{-\rho s} |Y_s^x|^2 \right] \leq C \left(1 + \mathbb{E} \left[\int_0^{+\infty} e^{-\rho s} |X_s^x|^{2r} ds \right] \right),$$

avec C une constante qui ne dépend pas de x .

2. Si $d' = 1$ (c'est-à-dire que Y^x est scalaire) et $r = 0$ alors, pour tout $x \in \mathbb{R}^d$, l'EDSR (4.2.1) a une solution telle que Y^x est un processus continu et borné et (Y^x, Z^x) satisfait

$$\operatorname{ess\,sup}_{\Omega \times \mathbb{R}^+} |Y^x| + \mathbb{E} \left[\int_0^\infty e^{-2\mu s} \|Z_s^x\|^2 ds \right] \leq C,$$

avec C une constante qui ne dépend pas de x . De plus, cette solution est unique dans la classe des processus (Y^x, Z^x) tels que Y^x est continu et borné et $Z^x \in \mathcal{M}_{loc}^2$.

Sous ces hypothèses, nous disposons d'une représentation markovienne $Y_t = u(X_t)$ où u est une fonction continue. En plus de ces hypothèses, nous supposons que b et σ sont continûment différentiables, que σ est bornée, et une condition d'ellipticité forte pour σ , toutes ces conditions assurant la différentiabilité de $(X_t^x)_{t \geq 0}$ par rapport à x . Nous désignons le processus de flot par $(\nabla_x X_t^x)_{t \geq 0}$ et définissons un poids de Malliavin $(U_t^x)_{t \geq 0}$ tel que,

$$U_t^x = \frac{1}{t} \int_0^t \langle \sigma^{-1}(X_s^x) \nabla_x X_s^x, dW_s \rangle \sigma(x) = (U_t^{x,j})_{1 \leq j \leq d},$$

avec lequel nous pouvons établir la représentation suivante de la solution de l'EDSR, qui sera très utile pour concevoir les schémas numériques.

Proposition 0.2.2. (Voir Proposition 4.2.3) Supposons qu'il existe $0 < \rho < 2\mu - K_{f,z}^2 \mathbf{1}_{d'>1}$ et $C > 0$ tels que, pour tout $x \in \mathbb{R}^d$

$$\mathbb{E} \left[\int_0^{+\infty} e^{-\rho s} |X_s^x|^{2r} ds \right] \leq C(1 + |x|^{2r}).$$

Nous considérons deux paramètres $a > \mu - \frac{K_{f,z}^2}{2} \mathbf{1}_{d'>1} > 0$ et $\tilde{a} \geq 2\mu - K_{f,z}^2 \mathbf{1}_{d'>1} > 0$ tels que pour tout ensemble compact $K \subset \mathbb{R}^d$ nous avons

$$\int_0^{+\infty} e^{-\tilde{a}s} \sup_{x \in K} \mathbb{E} [|X_s^x|^{2r}] ds + \sup_{x \in K} \mathbb{E} \left[\int_0^{+\infty} e^{-\tilde{a}s} \|\nabla_x X_s^x\|^2 ds \right] < +\infty.$$

Alors, sous l'hypothèse supplémentaire 4.2.2, il existe une fonction $u \in C^1(\mathbb{R}^d, \mathbb{R}^{d'})$ telle que, pour tout $x \in \mathbb{R}^d$, $Y^x = u(X^x)$, $\nabla_x u(X^x) \sigma(X^x) := \bar{u}(X^x)$ est une version continue de Z^x et nous avons la représentation probabiliste suivante pour (u, \bar{u}) : pour tout $x \in \mathbb{R}^d$,

$$\begin{aligned} u(x) &:= Y_0^x = \mathbb{E} \left[\int_0^\infty e^{-as} (f(X_s^x, u(X_s^x), \bar{u}(X_s^x)) + au(X_s^x)) ds \right] \\ &= \frac{1}{\theta} \mathbb{E} \left[(f(X_E^x, u(X_E^x), \bar{u}(X_E^x)) + au(X_E^x)) e^{-(a-\theta)E} \right] \end{aligned}$$

et

$$\begin{aligned} \bar{u}(x) &:= Z_0^x = \mathbb{E} \left[\int_0^\infty e^{-\tilde{a}s} (f(X_s^x, u(X_s^x), \bar{u}(X_s^x)) + \tilde{a}u(X_s^x)) U_s^x ds \right] \\ &= \frac{\sqrt{\pi}}{\sqrt{\tilde{\theta}}} \mathbb{E} \left[(f(X_{\tilde{E}}^x, u(X_{\tilde{E}}^x), \bar{u}(X_{\tilde{E}}^x)) + \tilde{a}u(X_{\tilde{E}}^x)) \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} U_{\tilde{E}}^x \right] \end{aligned}$$

où $\theta > 0$, $\tilde{\theta} > 0$, $E \sim \mathcal{E}(\theta)$ est indépendant de W et $\tilde{E} \sim \Gamma(1/2, \tilde{\theta})$ est indépendant de W . Enfin, nous avons la croissance suivante:

$$|u(x)| + \|\bar{u}(x)\| \leq C(1 + |x|^r), \quad \forall x \in \mathbb{R}^d.$$

Schémas numériques

Pour des raisons de concision, nous énonçons le schéma avec $X_t^x = W_t^x$, qui est un mouvement brownien commençant à $x \in \mathbb{R}^d$, mais il peut être facilement étendu à une EDS générale satisfaisant les hypothèses ci-dessus. Une manière courante de tester les schémas numériques pour les EDSR avec des solutions analytiques est de commencer avec une fonction donnée $u(x) \in C^2(\mathbb{R}^d)$ avec une croissance polynomiale au plus de degré r (pour être cohérent avec la Proposition 1.2.2) et de construire la fonction f telle que l'EDSR soit satisfaite par $Y_t^x = u(W_t^x)$ et $Z_t^x = \nabla_x u(W_t^x)$. Cela peut être fait en choisissant une fonction appropriée $f_0(y, z)$ qui satisfait les conditions de monotonie et de Lipschitz sur les composantes y et z respectivement. Ensuite, prenons le générateur de l'EDSR comme suit :

$$f(x, y, z) = f_0(y, z) - \frac{1}{2} \Delta u(x) - f_0(u(x), \nabla_x u(x)),$$

où Δ représente l'opérateur Laplacien. Ceci fonctionne car $u(x)$ satisfait l'EDP elliptique, $\frac{1}{2} \Delta u(x) + f(x, u(x), \nabla_x u(x)) = 0$.

1. Schéma basé sur la grille

Le premier schéma est basé sur le choix d'une sous-grille finie $\Pi \subset \delta\mathbb{Z}^d$ dans \mathbb{R}^d , avec $2N + 1$ divisions le long de chaque axe. Nous désignons par \square son enveloppe convexe. Par convention,

nous pouvons supposer que $0 \in \square$. Nous définissons ensuite une interpolation multi-linéaire sur Π avec les fonctions de base suivantes,

$$\forall z \in \Pi, \quad \forall x \in \mathbb{R}^d, \quad \psi_z(x) = \prod_{i=1}^d (1 - |\delta^{-1}(x_i - z_i)|)_+.$$

Pour une fonction $\phi : \Pi \rightarrow \mathbb{R}$, nous définissons $P\phi$ l'interpolation de ϕ sur \mathbb{R}^d comme suit

$$P\phi(x) =: \begin{cases} \sum_{z \in \Pi} \psi_z(x) \phi(z), & x \in \square \\ \sum_{z \in \Pi} \psi_z(\text{Proj}(x, \square)) \phi(z), & x \notin \square. \end{cases}$$

Le poids de Malliavin pour ce cas est donné par $U_t^x = \frac{W_t^x}{t}$. Le schéma est basé sur un argument de point fixe motivé par la représentation dans la Proposition 1.2.2. Le schéma complet est donné comme suit:

1. Avec $\theta < a$ et $\tilde{\theta} < b$, nous simulons M échantillons d'une variable aléatoire exponentielle $E \sim \mathcal{E}(\theta)$ et d'une variable aléatoire avec distribution gamma $\tilde{E} \sim \Gamma(\frac{1}{2}, \tilde{\theta})$ et commençons avec l'estimation initiale $v^0(x) = (u(x), \bar{u}(x)) = (0, 0) \forall x$.
2. Pour un n donné, évaluer $Pv^n = (Pu^n, P\bar{u}^n)$ pour chaque point de la grille z et mettre à jour v^n en Pv^n .
3. Ensuite, nous définissons l'opérateur ϕ et $\bar{\phi}$ comme suit

$$\begin{aligned} \phi(u^n)(x) &= \frac{1}{\theta M} \sum_{j=1}^M \left(f(W_E^{x,j}, u^n(W_E^{x,j}), \bar{u}^n(W_E^{x,j})) + a u^n(W_E^{x,j}) \right) e^{-(a-\theta)E}, \\ \bar{\phi}(u^n)(x) &= \frac{\sqrt{\pi}}{\sqrt{\tilde{\theta}} M} \sum_{j=1}^M \left(f(W_{\tilde{E}}^{x,j}, u^n(W_{\tilde{E}}^{x,j}), \bar{u}^n(W_{\tilde{E}}^{x,j})) + b u^n(W_{\tilde{E}}^{x,j}) \right) \sqrt{\tilde{E}} e^{-(b-\tilde{\theta})\tilde{E}} U_{\tilde{E}}^x \end{aligned}$$

4. Mettre à jour notre estimation pour v comme suit:

$$\begin{aligned} u^{n+1} &= \phi(u^n), \\ \bar{u}^{n+1} &= \bar{\phi}(\bar{u}^n). \end{aligned}$$

Nous avons également mené l'étude suivante sur l'erreur numérique pour ce schéma :

Theorem 0.2.1. Supposons que les hypothèses de la Proposition 4.2.3 sont satisfaites. Nous fixons $r' > r$ où r est défini dans (4.2.9). Supposons également que $\kappa_\infty < 1$ et que $v \in C^2(\mathbb{R}^{d' \times I})$ avec des dérivées secondes bornées. Alors les propriétés suivantes sont vraies.

1. Si σ est constante, $b - \tilde{\theta} - (1+r)K_b > 0$ et $a - \tilde{\theta} - rK_b > 0$, alors

$$\mathbb{E} \left[\sup_{x \in \mathbb{R}^d} \left\| \frac{Pv_M^n(x) - v(x)}{\rho(x)} \right\| \right] = O \left(\delta^2 + \frac{(\ln \tilde{N})^{1/(r+1)}}{\sqrt{M}} + |\tilde{N}\delta|^{r-r'} + \kappa_\infty^n \right).$$

2. Si $f(x, y, z) = f(x, y)$ et $a - \tilde{\theta} - rK_b > 0$, alors

$$\mathbb{E} \left[\sup_{x \in \mathbb{R}^d} \left\| \frac{Pv_M^n(x) - v(x)}{\rho(x)} \right\| \right] = O \left(\delta^2 + \frac{(\ln \tilde{N})^{1/r}}{\sqrt{M}} + |\tilde{N}\delta|^{r-r'} + \kappa_\infty^n \right).$$

3. Si $\beta \geq 2$, $b - (r+1)K_b - \frac{2\beta-1}{2}K_\sigma^2 - \frac{\beta-1}{\beta}\tilde{\theta} > 0$ et $a - rK_b - \frac{\beta-1}{\beta}\theta > 0$, alors

$$\mathbb{E} \left[\sup_{x \in \mathbb{R}^d} \left\| \frac{Pv_M^n(x) - v(x)}{\rho(x)} \right\| \right] = O \left(\delta^2 + \frac{(\tilde{N})^{d/\beta}}{\sqrt{M}} + |\tilde{N}\delta|^{r-r'} + \kappa_\infty^n \right).$$

Le schéma est efficace pour les dimensions faibles mais la complexité de calcul augmente exponentiellement avec d . Pour surmonter la malédiction de la dimension et l'erreur d'approximation au bord, nous proposons un schéma similaire basé sur la contraction mais avec des réseaux de neurones.

2. Schéma de réseau de neurones (RN) avec contraction

Nous utilisons à nouveau la propriété de contraction en considérant une séquence de réseaux de neurones $(\mathcal{U}_{\theta_n, n})_{n \in \mathbb{N}}$, c'est-à-dire une séquence de fonctions $x \mapsto (\mathcal{U}_{\theta_n, n}^1, \mathcal{U}_{\theta_n, n}^2) \in \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d}$ paramétrées par $\theta_n \in \Theta$. Nous fixons $\mathcal{U}_{\theta_0, 0} = 0$. Sachant le réseau de neurones $\mathcal{U}_{\theta_n, n}$, nous entraînons $\mathcal{U}_{\theta_{n+1}, n+1}$ comme suit :

1. Échantillonner $X_0 = X_{t=0} \sim \mu_0$, $E \sim \mathcal{E}(\theta)$ et $\tilde{E} \sim \Gamma(1/2, \tilde{\theta})$.
2. Obtenir des échantillons correspondants du poids de Malliavin $U_E^{X_0}$ et $U_{\tilde{E}}^{X_0}$ aux instants E et \tilde{E} pour le point de départ X_0 . Obtenir également des échantillons du processus d'état $X_E^{X_0}$ et $X_{\tilde{E}}^{X_0}$.
3. Évaluer le réseau de neurones précédent $\mathcal{U}_{\theta_n, n}$ à chaque échantillon de $X_E^{X_0}$ et $X_{\tilde{E}}^{X_0}$ et calculer les échantillons de $\varphi(\mathcal{U}_{\theta_n, n})(X_0)$ donnés par

$$\begin{aligned} \varphi(\mathcal{U}_{\theta_n, n})(X_0) &= \left(\frac{1}{\theta} (f(X_E^{X_0}, \mathcal{U}_{\theta_n, n}^1(X_E^{X_0}), \mathcal{U}_{\theta_n, n}^2(X_E^{X_0})) + a\mathcal{U}_{\theta_n, n}^1(X_E^{X_0})) e^{-(a-\theta)E}, \right. \\ &\quad \left. \sqrt{\frac{\pi}{\tilde{\theta}}} \left(f(X_{\tilde{E}}^{X_0}, \mathcal{U}_{\theta_n, n}^1(X_{\tilde{E}}^{X_0}), \mathcal{U}_{\theta_n, n}^2(X_{\tilde{E}}^{X_0})) + \tilde{a}\mathcal{U}_{\theta_n, n}^1(X_{\tilde{E}}^{X_0}) \right) \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} U_{\tilde{E}}^{X_0} \right). \end{aligned}$$

4. Minimiser la fonction de perte,

$$\mathcal{L}(\mathcal{U}_{\theta, n+1}) := \mathbb{E} \left[\|\mathcal{U}_{\theta, n+1}(X_0) - \varphi(\mathcal{U}_{\theta, n})(X_0)\|^2 \right]$$

et choisir les paramètres θ_{n+1} par

$$\theta_{n+1} \in \arg \min_{\theta \in \Theta} \mathbb{E} \left[\|\mathcal{U}_{\theta, n+1}(X_0) - \varphi(\mathcal{U}_{\theta, n})(X_0)\|^2 \right].$$

Nous avons établi le résultat suivant pour l'erreur d'approximation numérique du schéma :

Proposition 0.2.3. Soit $\varepsilon > 0$ tel que $(1 + \varepsilon)(\kappa_2)^2 < 1$. Alors, l'erreur d'approximation numérique e_n du schéma à la n -ème itération de Picard est bornée par,

$$e_n \leq \frac{(1 + \varepsilon) \max_{k \in \{1, \dots, n\}} r_k}{\varepsilon - \varepsilon(1 + \varepsilon)(\kappa_2)^2} + [(1 + \varepsilon)(\kappa_2)^2]^n e_0,$$

où $r_n := \mathbb{E} \left[\|\mathcal{U}_{\theta_{n+1}, n+1}(X_0) - \varphi(\mathcal{U}_{\theta_n, n})(X_0)\|^2 \right]$ est l'erreur d'approximation du réseau de neurones.

Ce schéma est très efficace même pour les problèmes de grande dimension, mais pour garantir la contraction, nous avons dû imposer des hypothèses supplémentaires. Celles-ci pourraient être assouplies pour le prochain schéma qui ne repose pas sur des arguments de contraction.

3. Schéma de RN sans contraction

Nous ne nous appuyons pas sur la contraction et tentons d'obtenir une approximation directe par réseau de neurones. Nous entraînons le réseau de neurones \mathcal{U}_θ en époques, chacune avec \mathcal{P} étapes de descente de gradient, où le RN est optimisé pour un ensemble donné d'échantillons avant de mettre à jour les échantillons d'entrée dans l'époque suivante. Nous utilisons le même opérateur φ défini pour le schéma précédent. Au début de chaque époque, nous procédons ainsi : échantillonner $X_0 = X_{t=0} \sim \mu_0$; pour chaque échantillon de X_0 , obtenir M échantillons indépendants de $(E, \tilde{E}, X_E^{X_0}, X_{\tilde{E}}^{X_0}, U_E^{X_0}, U_{\tilde{E}}^{X_0})_j$ pour $j = 1, \dots, M$.

Ensuite, dans une époque donnée i avec les M échantillons indépendants

$\left((E, \tilde{E}, X_E^{X_0}, X_{\tilde{E}}^{X_0}, U_E^{X_0}, U_{\tilde{E}}^{X_0})_{j=0, \dots, M} \right)_i$, nous faisons ce qui suit pour les \mathcal{P} prochaines étapes de descente de gradient, $n \in \{1, \dots, \mathcal{P}\}$:

1. Avec les paramètres actuels θ_n^i , nous calculons les M échantillons correspondants $(\varphi^j(\mathcal{U}_{\theta_n^i})(X_0))_{1 \leq j \leq M}$ pour chaque échantillon de X_0 qui est une fonction uniquement de θ_n^i .
2. Les paramètres θ_{n+1}^i sont obtenus en utilisant l'étape d'optimisation avec la fonction de perte

$$\Upsilon_M(\mathcal{U}_\theta) := \mathbb{E} \left[\left\| \mathcal{U}_\theta(X_0) - \frac{1}{M} \sum_{j=1}^M \varphi^j(\mathcal{U}_\theta)(X_0) \right\|^2 \right].$$

Ce schéma surmonte les difficultés des deux schémas précédents, comme nous le démontrons avec le tracé de l'erreur quadratique moyenne et de la constante de Lipschitz de f par rapport au troisième argument. Nos expériences montrent qu'il y a un point au-delà duquel même si les hypothèses sont satisfaites, la contraction échoue et le deuxième schéma aussi, mais le troisième schéma a toujours une erreur raisonnablement petite. Cependant, contrairement aux deux autres schémas, la dimension de l'échantillon est significativement plus élevée et pour l'instant, nous rencontrons des difficultés avec l'analyse numérique du schéma.

0.3 Conclusions

Dans le Chapitre 2, nous étudions le problème de contrôle stochastique de champ moyen à horizon infini où la fonction de coût dépend de la loi du processus d'état. Nous avons prouvé les conditions nécessaires et suffisantes d'optimalité qui exigent la L-différentiabilité et la L-convexité dans l'espace des mesures pour la fonction de coût instantané. Lorsque cette dépendance se fait à travers des interactions scalaires, le problème est plus simple et a été étudié par les chercheurs de manière extensive au fil des ans. Nous avons cherché des exemples de fonctions de mesure autres que les interactions scalaires qui satisfont ces conditions tout en étant pertinentes du point de vue des applications, démontrant ainsi que c'est en effet un problème difficile.

Dans le Chapitre 3, nous commençons par une application en optimisation de portefeuille du problème MFC introduit dans le Chapitre 2. L'objectif est de surperformer une allocation statique dans un portefeuille investi en actions et en actif sans risque en s'appuyant sur une allocation dynamique, en utilisant la vitesse de négociation des actifs comme contrôle avec un critère de minimisation du risque baissier de type champ moyen. Cela a été fait en trouvant un exemple d'une fonction L-différentiable et L-convexe de la loi du processus d'état. Nous avons prouvé les conditions d'optimalité pour le problème de contrôle et établi l'existence et l'unicité de la solution du système couplé correspondant de type EDSPR McKean-Vlasov. Nous avons également développé un schéma numérique basé sur les réseaux neuronaux pour résoudre une version tronquée dans le temps du problème et fourni des bornes exponentielles pour l'erreur

de troncature. Nous avons montré qu'augmenter le multiplicateur du terme de champ moyen inclinait avec succès la distribution de richesse vers la droite, augmentant ainsi la probabilité de richesse relative plus élevée. Néanmoins, le schéma est instable vis-à-vis du paramètre de régularisation pour la L-dérivée (qui a besoin d'être suffisamment grand pour observer plus de variabilité dans les trajectoires) et du coefficient multiplicatif devant le terme de champ-moyen qui nous empêche de prendre des positions agressives.

Dans le Chapitre 4, nous étudions l'approximation numérique des EDSRs à horizon infini. Nous avons développé trois schémas numériques. Le premier schéma est basé sur un argument de point fixe et utilise une approximation par grille spatiale. Le deuxième schéma est également basé sur un argument de point fixe mais ne discrétise pas l'espace et utilise des réseaux neuronaux. Le troisième schéma ne dépend pas de l'argument de point fixe et utilise des réseaux neuronaux comme le deuxième schéma. Nous fournissons également une étude détaillée de l'erreur numérique pour le premier schéma et prouvons des bornes fines sur l'erreur d'approximation, nécessitant des hypothèses supplémentaires pour la contraction. Pour le deuxième schéma, nous avons prouvé la convergence de l'erreur d'approximation vers zéro lorsque la taille du réseau neuronal augmente. Nous n'avons pas prouvé de telles garanties numériques pour le troisième schéma mais nous avons montré qu'il fonctionne mieux que les deux premiers schémas car il ne nécessite pas les hypothèses supplémentaires pour la contraction.

Perspectives Futures

Il existe de nombreuses directions de recherche intéressantes qui ont émergé lors de la réalisation de la thèse. Dans le Chapitre 3, on pourrait trouver une autre fonction de la mesure du processus d'état qui pourrait surmonter le problème de stabilité comme indiqué précédemment et nous pourrions éviter des trajectoires de vitesse de trading trop proches. Cela fournirait une perspective alternative pour modéliser le problème avec le contrôle étant l'allocation d'actifs et la fonction de coût étant une fonction de la loi du contrôle. Il serait également intéressant d'essayer d'autres schémas numériques pour les EDSPRs McKean-Vlasov qui ne nécessitent pas d'itérations de jeu fictif comme cela a été récemment fait par [63].

Dans le Chapitre 4, obtenir des estimations pour l'erreur d'approximation pour le troisième schéma serait très intéressant. De plus, la vérification de $\kappa_p < 1$ en utilisant une version des estimations d'Aronson pour l'horizon infini serait utile.

Chapter 1

Introduction

Contents

1.1 Preliminaries and context	23
1.1.1 Stochastic control	23
1.1.2 BSDEs and FBSDEs	25
1.1.3 Mean-field control	29
1.2 Contributions and challenges	32
1.2.1 Mean-field stochastic control in infinite horizon and applications	32
1.2.2 BSDEs in infinite horizon	38
1.3 Conclusions	42

The main focal points of this thesis are backward stochastic differential equations (BSDEs) and McKean-Vlasov forward-backward stochastic differential equations (MV FBSDEs) in infinite horizon, their uses and most prominently their connection with mean-field stochastic control problems in infinite horizon. In the first part of the thesis, we first study them in full generality and later tackle a long term portfolio optimization problem with a mean-field optimization criterion and solve it numerically. In the second part, we shift our focus on developing efficient numerical schemes for solving backward stochastic differential equations (BSDEs) in infinite horizon, which have found applications in numerous fields recently.

The introduction is structured as follows: The first section provides the broader context of the thesis and provides a concise overview of stochastic control and its connection with BSDEs. We also provide an overview of mean-field systems and mean-field control and review the state of the art. Second section highlights the challenges we faced and our contributions. The third and final section concludes our findings and explores future perspectives.

1.1 Preliminaries and context

1.1.1 Stochastic control

Background

Stochastic control is a specialized branch of control theory which extends the principles of mathematical optimization to dynamical systems influenced or driven by uncertainty. The most prominent initial contribution to control theory was in dynamic programming by Bellman [16] in the 1950s, marking a significant milestone in applied mathematics. Kalman [82] made the connection between dynamic programming and Hamilton-Jacobi equation from classical physics in 1960. See Åström [118] for more historical perspectives on stochastic control. Barles [12]

studied viscosity solutions for the Hamilton-Jacobi-Bellman (HJB) equation, laying theoretical foundations for continuous-time stochastic control problems.

Applications

Over the years, stochastic control methodologies have become indispensable for modeling complex problems from numerous fields and more notably, financial markets. The seminal work of Merton [94] laid the foundation for applying stochastic control to optimal portfolio selection, which is still a cornerstone for the modern portfolio theory. Extending this framework, Black and Scholes [24] and [110] developed the Black-Scholes-Merton model, revolutionizing the pricing of options and derivatives. The inclusion of jump processes in stochastic control models has made way for us to model extreme events in financial markets. Merton [95] introduced the Jump-Diffusion model, and Duffie and Pan [55] extended it to include default risk, leading to the development of sophisticated credit risk models. Fleming and Soner [59] provided insights into optimal investment strategies under transaction costs, addressing a critical aspect of real-world trading. Advances in modeling market microstructure, such as [28, 21, 46, 75, 83] have found further financial applications of stochastic control. Moreover, there have also been numerous applications of stochastic control in other fields.

Stochastic control methodologies are essential for addressing the challenges posed by the dynamic and uncertain nature of the energy sector. In energy finance, these methodologies have often been employed in the pricing and hedging of energy derivatives. Schwartz and Smith [111] introduced mean-reverting processes to model energy prices, providing a foundation for incorporating stochastic control techniques. Their work laid the groundwork for optimizing trading strategies in markets with mean-reverting dynamics. Dixit and Pindyck [53] applied real options theory to energy investment decisions. Stochastic control methods were crucial in valuing and optimizing the exercise of real options in the context of energy projects. Alipour et al. [8] addressed optimal operation of power systems using stochastic control. They provided insights into optimal dispatch strategies considering uncertainties in electricity prices and demand. Alhelou et al. [7] focused on optimizing the operation of hybrid renewable energy systems under uncertainty. Their work highlighted the significance of stochastic control in managing the integration of renewable energy sources. Alharbi and Bhattacharya [6] provided a comprehensive review of stochastic control approaches for optimal operation and sizing of energy storage systems. Their work addressed the crucial role of energy storage optimization in managing energy systems.

Mathematical overview

Next, let us move into the mathematical formalism for stochastic control. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a filtered probability space with augmented and right-continuous filtration $(\mathcal{F}_t)_{t \geq 0}$ and consider the following stochastic differential equation (SDE) as the dynamics of the controlled system,

$$dX_t^u = b(X_t^u, u_t) dt + \sum_{j=1}^m \sigma^j(X_t^u, u_t) dW_t^j, \quad X_{t_0} = x_0,$$

where X_t^u is called the state variable or state process for the system, taking values in \mathbb{R}^d . $W_t \in \mathbb{R}^m$ is a \mathcal{F}_t -adapted Brownian motion and $(u_t)_{t \geq 0}$ is a \mathcal{F}_t -progressively measurable process called the control, which takes values in a non-empty closed and convex subset $U \subset \mathbb{R}^n$. The objective of the optimization process could be to maximize a certain reward or minimize costs, which could be a function of the controlled state process or the control process itself and often both. Let us consider a cost minimization problem as an example. The cost functional is usually

given as follows,

$$J^T(x_0, u) = \mathbb{E} \left[\int_0^T f(X_t^{u, x_0}, u_t) dt + g(X_T^{x_0, u}) \right],$$

where the function f is called the running cost and g is called the terminal cost. The objective of the optimization problem is to find $u = (u_t)_{t \geq 0}$ in a space denoted by \mathcal{A} , called the space of admissible controls (predictable and bounded processes for finite horizon which minimizes this cost functional. The function $V^T(x_0) = \min_{u \in \mathcal{A}} J^T(x_0, u)$ is called the value function and $u^* = \arg \min_{u \in \mathcal{A}} J^T(x_0, u)$ is called an optimal control. One could model problems with no terminal time using a cost functional with a discount factor to account for diminishing value of future costs and an infinite time horizon, which is primarily our focus in the thesis. One writes the cost functional as,

$$J(x_0, u) = \mathbb{E} \left[\int_0^\infty e^{-\rho t} f(X_t^{u, x_0}, u_t) dt \right].$$

Dynamic programming is one of the two main approaches for solving stochastic control problems. The Hamilton-Jacobi-Bellman (HJB) equation,

$$-\rho V(x) + \min_{u \in U} \{ \mathcal{L}^u V + f(x, u) \} = 0, \forall x \in \mathbb{R}^d,$$

where \mathcal{L}^u is the generator of the state process. A closed-loop optimal control is then given by $u^*(x) \in \arg \min_{u \in U} \{ \mathcal{L}^u V + f(x, u) \}$. The other main approach for solving stochastic control problems is the stochastic Pontryagin's maximum principle (see Pontryagin et al. [107]), where we define a function called Hamiltonian, $H(x, u, y, z) = b(x, u) \cdot y + \text{tr}(\sigma^\top(x, u)z) + f(x, u)$, where (y, z) with $y \in \mathbb{R}^d$ and $z \in \mathbb{R}^{m \times d}$ are called the adjoint variables which come from the solution pair $(Y_t, Z_t)_{t \geq 0}$ of the following BSDE for a given $u \in \mathcal{A}$,

$$dY_t = -\nabla_x H(X_t^u, u_t, Y_t, Z_t) dt + Z_t dW_t.$$

An optimal control is given by $u_t^* \in \arg \min_{u \in U} H(X_t^u, u, Y_t, Z_t)$ which is a function of the solution of the BSDE. With this characterization of the optimal control put into the state dynamics, one obtains a coupled system of FBSDEs. This forms a crucial connection between stochastic control and BSDEs, which is explored in detail in the thesis. For an extensive coverage of the connection between control theory and FBSDEs, see Ma and Yong [81]. For an excellent theoretical background and main results in BSDEs in the most general setting, see Pardoux and Râşcanu [100].

1.1.2 BSDEs and FBSDEs

Background

A BSDE is a type of SDE which evolves backwards in time, where one looks for an adapted pair of processes which solve an SDE with a given terminal condition. BSDEs were introduced in a linear setting by Bismut [23] and have emerged as a powerful mathematical framework for modeling and solving problems involving stochastic processes evolving backwards in time. The genesis of BSDEs in the most general setting can be traced back to the seminal work of Pardoux and Peng [99], marking a significant advancement in the field of stochastic analysis. Many useful theoretical results have since been obtained, for example, the connection with semi-linear parabolic PDEs by Buckdahn et al. [33] who studied the homogenization properties and [101, 48] studied the connection between semi-linear elliptic PDEs and BSDEs with random terminal time, marking an important contribution to BSDEs with infinite horizon. Briand and Hu [30] studied the stability of BSDEs with random terminal time and applied his results to semi-linear elliptic PDEs. Major

contributions in BSDEs include the existence and uniqueness of solutions, regularity properties, and connections with partial differential equations. El Karoui et al. [58] provided foundational insights, establishing a connection between BSDEs and quasilinear partial differential equations. Ma and Yong [89] marks an important early contribution to the FBSDEs. The works of Soner et al. [112] and Briand and Confortola [29] further extended the understanding of stability and robustness in BSDEs. For existence and uniqueness of BSDEs with generators having quadratic growth, see Briand and Hu [31] and Delbaen, Hu, and Richou [51]. The latter proved existence and uniqueness for quadratic BSDEs with convex generator and unbounded terminal conditions which admit all exponential moments. Chassagneux and Richou [43] developed numerical scheme for solving BSDEs with generators having quadratic growth in the Z process and bounded terminal condition. They obtained time discretization error of the order $\frac{1}{2} - \varepsilon$ for all $\varepsilon > 0$.

Applications

The versatility of BSDEs is evident in their applications across diverse fields. In finance, BSDEs find utility in option pricing, risk management, and portfolio optimization which were introduced in the seminal work by El Karoui, Peng, and Quenez [58] (see Delong [52] for financial applications of BSDEs with jumps). The backward nature of BSDEs aligns well with financial contexts, where decisions impact future outcomes. BSDEs have played a crucial role in risk management and derivative pricing in finance. Peng [103] reviewed backward stochastic differential equations as a powerful tool for solving non-linear expectations, offering versatility for option pricing and risk assessment. The application of BSDEs has extended to credit risk modeling, as demonstrated by Crépey and Song [47]. The advent of FBSDEs has enabled a more comprehensive modeling of dynamic financial systems. Balter and Pelsser [11] presented the FBSDE framework, offering a unified approach to pricing and hedging in incomplete markets. Applications include the pricing of American options in Barles [13] and the study of optimal investment and consumption strategies in Chacko and Viceira [41]. Mean-Field FBSDEs represent a recent and powerful development for modeling large populations of interacting agents in financial markets. Carmona et al. [39] introduced mean-field FBSDEs for market impact modeling, showcasing their potential in capturing the feedback effects of trading strategies on market dynamics. BSDEs have found applications in modeling and pricing electricity derivatives. Benth and Šaltytė-Benth [19] utilized BSDEs to model spot price dynamics in electricity markets, capturing the impact of supply and demand factors.

Mathematical overview

Let us now delve into a brief mathematical formalism for BSDEs. A standard BSDE in the finite horizon is represented as:

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad \forall t \in [0, T].$$

where (Y_t, Z_t) are the unknown processes, ξ is the terminal condition which is \mathcal{F}_T measurable, f is called the generator function, W_t is a Brownian motion, and t varies in the time interval $[0, T]$. Following are some of the most crucial results which laid the foundations of BSDEs:

Theorem 1.1.1 (Existence and Uniqueness, see El Karoui et al. [58], Pardoux and Peng [99]). *For a given BSDE with terminal condition $\xi \in L^2(\mathcal{F}_T)$ and generator $f : [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$, if $t \mapsto f(t, x, y, z)$ is progressively measurable and the following uniform Lipschitz condition holds: There exist a constant K_f such that, for all $dt \otimes d\mathbb{P}$ a.s.,*

$$|f(t, y, z) - f(t, \tilde{y}, \tilde{z})| \leq K_f(|y - \tilde{y}| + |z - \tilde{z}|), \quad y, \tilde{y} \in \mathbb{R} \text{ and } z, \tilde{z} \in \mathbb{R}^d,$$

then there exists a unique adapted solution (Y, Z) to the BSDE.

In the infinite horizon setting, we can write the BSDE as,

$$Y_t = Y_T + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad \forall 0 \leq t \leq T < +\infty.$$

We look for the existence and uniqueness in the following space,

$$\mathcal{M}^{2,\rho}(\mathbb{R}^d \times \mathbb{R}^{d \times m}) := \left\{ (Y, Z) \in \mathbb{R}^d \times \mathbb{R}^{d \times m} \mid \mathbb{E} \left[\int_0^\infty e^{-\rho t} (|Y_t|^2 + \|Z_t\|^2) dt \right] < \infty \right\}.$$

One also needs stronger assumptions for existence and uniqueness of the solutions. In particular, along with the above assumptions, we need ‘monotonicity’ of the generator with respect to the Y process and an extra condition based on ρ , the monotonicity constant (denoted by μ) and the Lipschitz constant with respect to Z (denoted by $K_{f,z}$). Darling and Pardoux [48] and Briand and Hu [30] have extensively studied the existence and uniqueness for BSDEs with a random terminal time. However, with little change, one can establish the same results for the infinite horizon setting.

Theorem 1.1.2 (Existence and uniqueness infinite horizon, see [30, 48]). *In addition to the assumptions of 1.1.1, we must satisfy,*

1. *Monotonicity with respect to Y , $\exists \mu$ s.t. $\forall (y, y') \in \mathbb{R}^d, \forall z \in \mathbb{R}^{d \times m}$ and $\forall s \geq 0$,*

$$\langle y - y', f(s, y, z) - f(s, y', z) \rangle \leq -\mu |y - y'|^2;$$

2. *$\exists \rho \in \mathbb{R}$ such that $\rho > K_{f,z}^2 - 2\mu$ and*

$$\mathbb{E} \left[\int_0^\infty e^{\rho t} |f(t, 0, 0)|^2 dt \right] \leq K_f.$$

Under these assumptions, a unique solution exists to the above infinite horizon BSDE in the space $\mathcal{M}^{2,\rho}(\mathbb{R}^d \times \mathbb{R}^{d \times m})$.

The condition on the discount factor is required for the multi-dimensional case but it can be relaxed for the one-dimensional setting, as was shown by Briand and Hu [30]. Next, let us state some a priori stability estimates for BSDEs,

Theorem 1.1.3 (A Priori Estimates, see El Karoui et al. [58]). *The solutions of BSDEs exhibit a priori stability estimates. Specifically, for two BSDEs with different terminal conditions and generators, denoted by ξ_1, ξ_2 and f^1, f^2 respectively, if the conditions of Theorem 1.1.1 are satisfied, then the following estimate holds:*

$$\begin{aligned} & \mathbb{E} \left[\int_0^T e^{\beta t} |Y_t^1 - Y_t^2|^2 dt \right] + \mathbb{E} \left[\int_0^T e^{\beta t} |Z_t^1 - Z_t^2|^2 dt \right] \\ & \leq C \left(\mathbb{E} [|\xi_1 - \xi_2|^2] + \mathbb{E} \left[\int_0^T e^{\beta t} |f^1(t, Y_t^2, Z_t^2) - f^2(t, Y_t^2, Z_t^2)|^2 dt \right] \right), \end{aligned}$$

where Y_t^1 and Y_t^2 are the solutions corresponding to (f^1, ξ_1) and (f^2, ξ_2) respectively and C is a positive constant.

The constant C above depends on terminal time and one might face the difficulty of explosion in C when one takes $T \rightarrow \infty$. However, many stability results have been obtained for BSDEs with random and infinite terminal time, for instance, by Darling and Pardoux [48] and the stability result by Briand and Hu [30]. Next, in the setting of the above a priori estimate, we state the comparison theorem which was first obtained by Peng [104].

Theorem 1.1.4 (Comparison Principle, see El Karoui et al. [58]). *Given (f^1, ξ^1) and (f^2, ξ^2) and corresponding solutions (Y^1, Z^1) and (Y^2, Z^2) which are square integrable, suppose that,*

- $\xi^1 \geq \xi^2 \mathbb{P}$ a.s.
- $f^1(t, Y^2, Z^2) - f^2(t, Y^2, Z^2) \geq 0 \mathbb{P} \otimes dt$ a.s.

Then, we have $Y_t^1 \geq Y_t^2 \mathbb{P} \otimes dt$ a.s. .

Next, we highlight the various important contributions and developments in the numerical methods of solving BSDEs.

Numerical methods

Numerical schemes for solving BSDEs have been investigated by many researchers over the years. Bally and Pagès [10] analysed the numerical error for the quantization method; Bouchard and Touzi [26] studied the convergence rates of approximations by regression using the kernel and the Malliavin approaches; Del Moral, Rémillard, and Rubenthaler [50] presented an estimator preserving the convexity and monotonicity for American options; Gobet and Labart [66] proposed a BSDE solver where an adaptive control variate is added making use of the connection between linear BSDEs and linear parabolic PDEs; Guyon and Henry-Labordère [74] provided a regression-based approximation for the optimal co-variance matrix within an uncertain volatility model. Among all above different approaches for applying Monte-Carlo methods, the most popular one is presumably the ordinary linear least-squares regression (OLS in short) which was developed by Ma, Protter, San Martín, and Torres [91] (see also [45]). Under this framework, in a finite time horizon, one first discretizes the time and then simulates paths for the processes and uses Monte-Carlo approach. The estimator is then computed backwards in the time grid as a linear combination of a reduced set of functions, finding coefficients which are optimal in the empirical least-squares sense.

Gobet, Lemor, and Warin [70] first extended OLS for the standard BSDEs and established the convergence rates for a dynamic programming based explicit scheme. Gobet and Lemor [67] developed approximations for BSDEs with jumps and for reflected BSDEs; Bender and Steiner [17] used a martingale basis function to reduce the propagation error when computing the conditional expectations; Bouchard and Warin [27] improvised the use of an adaptive local basis where the support of the function basis is adapted to the density of the samples; Bernhart, Pham, Tankov, and Warin [20] applied this procedure to approximate the penalized solution of BSDEs with constrained jump and they established a logarithmic convergence rate; Gobet and his co-authors have done the error analysis of many different numerical estimators for solving BSDEs: the multi-step explicit scheme was studied in Gobet and Turkedjiev [69], the stratification sampling with parallelization using GPU was considered to improve the computational performance in Gobet et al. [71], the hedging strategy with Malliavin weight was presented in Gobet and Turkedjiev [68].

Over the years, the computational capabilities have advanced to an extent which has allowed researchers to study schemes based on deep learning methods for solving BSDEs in very high dimensions (~ 100), leveraging their power to mitigate the curse of dimensionality. E, Han, and Jentzen [56] were the first to apply a method which they called ‘Deep BSDE’ where a forward scheme is considered and the parameters of network are optimized to minimize the Mean Squared Error (MSE in short) on terminal condition by stochastic gradient descent; then Han and Long [77] extended the deep BSDE approach to solve a system of coupled FBSDE and demonstrated its performance for an explicitly solvable system in very high dimensions and recently, Carmona and Laurière [38] presented a similar approach for a Mean field control problem where the related approximation of a specific decoupled FBSDE of McKean-Vlasov type, see also Pham and Warin [106] for a suggestion for the network architecture.

1.1.3 Mean-field control

Background

Mean-field control (MFC) and Mean-field games (MFGs) are closely related with the former originating from MFGs historically. MFGs were introduced by Lasry and Lions [85] and Huang et al. [80] who provided a powerful mathematical tool for modeling the strategic interactions of a large number of rational agents. These agents, collectively influencing the system dynamics, face a trade-off between individual and collective objectives. The mathematical formalism of mean-field games involves the coupled evolution of the distribution of agents and each optimizing their own objective functions where their actions affect the whole population. This coupled nature necessitates advanced numerical techniques for accurate approximation and efficient simulation. Theoretical frameworks, such as the mean-field game formulation by Pardoux et al. [102], have laid the groundwork for understanding the equilibrium behavior of large populations in the presence of mean-field interactions. Their work provides a comprehensive overview of mean-field games and their applications, offering insights into the mathematical foundations of these models. The study has served as a cornerstone for subsequent theoretical developments in the field. For a broader coverage and perspectives for mean-field games, see Bensoussan, Frehse, Yam, et al. [18]. An extensive study of MFGs and MFCs along with applications can be found in the two volumes of books written by Carmona and Delarue [37], which have arguably provided one of the best detailed coverage of the topic.

As systems become increasingly complex and interconnected, traditional control strategies may fall short in capturing the emergent behavior of large populations. Hence, the focus on numerical solutions for MFGs and MFCs using McKean-Vlasov FBSDEs is motivated by the necessity to bridge the gap between theoretical formulations and real-world implementations. The motivation for this research lies in the need for a deeper understanding of mean-field stochastic control in infinite horizon and its applications, prominently in portfolio optimization.

Applications

Advancements in deep reinforcement learning have found applications in mean-field games. Ruthotto et al. [109] explored the use of deep reinforcement learning for mean-field games, demonstrating accelerated learning algorithms for Hamilton–Jacobi–Bellman equations through neural networks. This theoretical development highlights the synergies between reinforcement learning and mean-field control, paving the way for more sophisticated decision-making strategies. The study also provided a theoretical foundation for understanding the broader implications and applications of mean-field models in diverse scientific contexts. For a survey of mean-field games applied to various problems, see Gomes and Saúde [73]. Cesaroni et al. [40] extended the mathematical tools to fractional Laplacians, showcasing the broad applicability of mean-field control theories. Mertens and Sorin [93] provided insights into the stability of equilibria, contributing to our understanding of the long-term behavior of mean-field systems. The study laid the foundation for subsequent research on stability analysis in the context of mean-field stochastic control. Bogachev [25] presented a comprehensive study on Fokker–Planck–Kolmogorov equations, offering theoretical insights into the evolution of probability distributions in mean-field systems. This work has broad implications for understanding the dynamics of large-scale interacting systems and their statistical properties. Bensoussan et al. [18] covers several applications mostly from a pedagogical point of view. Cardaliaguet and Lehalle [35] developed a framework for a mean-field game of controls, applied to the classical problem of optimal liquidation and trade crowding. Barreiro-Gomez and Tembine [14] used mean-field optimal control problem for a large-scale water distribution network with not only the mean but also the variance of both the system state and control process.

Connection with FBSDEs and more applications

The optimization of energy storage strategies is an emerging application of FBSDEs. De Paola et al. [49] and Li et al. [87] proposed FBSDEs framework to model the optimal operation of energy storage systems, considering uncertainties in electricity prices and renewable energy generation. This kind of approach has implications for improving the efficiency and economic viability of energy storage technologies. Mean-Field FBSDEs are gaining traction in modeling the dynamics of interconnected power systems. Alasseur et al. [4] introduced mean-field FBSDEs for power grid stability analysis, considering the interactions among multiple generators and loads. This approach facilitates a holistic understanding of the grid's behavior under various operating conditions. Interdisciplinary collaborations between mathematicians, economists, and engineers have led to advancements in the application of stochastic control methodologies in the energy sector. Hafiz et al. [76] integrated stochastic control with machine learning techniques for optimal real-time operation of energy storage systems, showcasing the potential of combining diverse methodologies. Acciaio, Backhoff-Veraguas, and Carmona [1] studied an 'extended' MFG using the maximum principle and provided insights into the connection with the theory of optimal transport and solved an optimal liquidation problem. Äid, Basei, and Pham [3] used the McKean-Vlasov approach for modeling distributed energy generation. More recent works include the use of McKean-Vlasov stochastic control for the energy management of a smart grid by Gobet and Grangereau [64] and in [65], they developed a novel approach combining Newton's method and Pontryagin's principle for stochastic control. Germain et al. [62] used mean-field dynamics with probabilistic state constraints and level-set approach for renewable energy storage problem.

Mathematical overview

Let us now clearly define the general problem of mean-field control with dependence on the law of the state but not control. The MFC with dependence on the law of control could be studied with a similar approach but poses additional problems as one cannot convert the problem the problem into a pointwise minimization problem (see [1] for example). Similar to the formulation of stochastic control problem, consider a controlled state SDE $(X_t^u)_{t \geq 0}$ with the control process being $(u_t)_{t \geq 0}$,

$$dX_t^u = b(X_t^u, u_t, \mu_t) dt + \sum_{j=1}^m \sigma^j(X_t^u, u_t, \mu_t) dW_t^j, \quad X_{t_0} = x_0,$$

where μ_t denotes the law of the state X_t^u . With this dependence on the law, it is called a McKean-Vlasov SDE which can be seen as the mean-field limit for a large number of interacting players in a MFG (see [37] for details). When the dependence on the law is through a function of the expectation of the state, it is called 'scalar interaction' which is easier to handle than the general case but still applicable many problems (see [64, 1] for example). The general framework requires a notion of differentiability and an appropriate notion of convexity in the space of measures. See Chapter 5 of [37] for an extensive discussion of the topic. We will be solely using the notion of 'L-differentiability' and 'L-convexity' as they are best suited for problems in the probabilistic setting, which is ideal for us as we use FBSDEs to solve the MFC. In general, not only the state dynamics but also the objective function could be dependent on the law of the state.

$$J^T(x_0, u) = \mathbb{E} \left[\int_0^T f(X_t^{u, x_0}, u_t, \mu_t) dt + g(X_T^{x_0, u}, \mu_T) \right],$$

In this thesis, we will only be using the Pontryagin's maximum principle for solving MFC. However, dynamic programming has been used by some researchers, for example by Tembine, Zhu, and Başar [114] and Zhou, Touzi, and Zhang [117]. When using the maximum principle,

the problem of finding the optimal control in the infinite dimensional space of functions (space of admissible controls) can be converted into solving a coupled system of McKean-Vlasov FBSDEs where the optimal control is given in the feedback form as a function of the adjoint processes (Y and Z) using the necessary optimality conditions. An optimal control u_t must satisfy,

$$\nabla_u H(t, X_t^u, \mu_t, u, Y_t, Z_t) = 0 \quad dt \otimes d\mathbb{P}a.s.$$

Let us denote the optimal control in its feedback form $u^*(X_t, Y_t, Z_t)$ by simply u_t^* where the dependence on the FBSDE (X, Y, Z) is understood. Also, let us use θ_t as a shorthand for a group of variables where the dependence is clear. Then, the coupled system of McKean-Vlasov FBSDE is given by,

$$\begin{aligned} dX_t &= b(X_t, u_t^*, \mu_t) dt + \sum_{j=1}^m \sigma^j(X_t, u_t^*, \mu_t) dW_t^j, \quad X_{t=0} = x_0 \\ dY_t &= \left(Y_t \partial_x b(\theta_t) + \sum_{j=1}^m Z_t^j \partial_x \sigma^j(\theta_t) + \partial_x f(\theta_t) + \partial_\mu f(\theta_t)(X_t) \right) dt - \sum_{j=1}^m Z_t^j dW_t^j, \\ Y_{t=T} &= \xi \in \mathcal{F}_T, \end{aligned}$$

where $\partial_\mu f(\theta)(X_t)$ denotes the L-derivative with respect to the law of the state, evaluated at X_t . See Carmona and Delarue [37] for a detailed study in the finite horizon setting. Next, we discuss the important contributions in the numerical resolution of the problem.

Numerical schemes for McKean-Vlasov FBSDEs

The numerical approximation of McKean-Vlasov FBSDEs remains a challenging yet crucial aspect for understanding the dynamics of mean-field systems. Numerous numerical schemes have been proposed, each tailored to address specific complexities inherent in these equations. Achdou and Capuzzo-Dolcetta [2] laid the foundation for numerical methods in mean-field games by introducing finite-difference approximations. This seminal work provided insights into the challenges posed by mean-field systems and offered a starting point for subsequent research in the field of numerical analysis for mean-field models. Advancing beyond the finite-difference approach, Cardaliaguet and Lehalle [35] extended the numerical analysis to mean-field games with common noise and jumps. The authors introduced a robust numerical scheme capable of handling the additional complexities arising from common noise and discontinuous processes. This development marked a critical step towards more realistic simulations of mean-field systems. Deep learning techniques have been increasingly integrated into the numerical approximation of mean-field models. The probabilistic approach for mean-field games with major and minor agents, presented by Buckdahn et al. [34], contributed to the understanding of heterogeneous agent dynamics. The accompanying numerical analyses provided valuable insights into the challenges and opportunities posed by diverse agent characteristics in mean-field systems.

Mean-field backward stochastic differential equations (MFBSDEs) have been a major focus of recent numerical developments. Djehiche et al. [54] addressed numerical methods for MFBSDEs, extending the scope of applications to various domains. This work broadened the numerical toolkit available for researchers dealing with backward components in mean-field systems. Bertucci [22] extended the numerical analysis to mean-field games of optimal stopping, introducing a novel framework for studying optimal stopping problems within a mean-field context. This work broadens the applicability of numerical schemes to encompass a wider range of decision-making scenarios, contributing to the versatility of mean-field models. Ruthotto et al. [109] explored the use of deep reinforcement learning for mean-field games, demonstrating accelerated learning algorithms for Hamilton–Jacobi–Bellman equations through neural networks. This theoretical development highlights the synergies between reinforcement learning and mean-field

control, paving the way for more sophisticated decision-making strategies. As remarked earlier, with the advent of the Deep BSDE method by [56], many uses were found in solving the coupled McKean-Vlasov FBSDE. Han and Long [77] developed an algorithm utilizing the fictitious play approach from game theory. On the contrary, Germain, Mikael, and Warin [63] proposed an algorithm with dynamic update of the empirical law of the state process, comparing the convergence rate of various approaches. Our approach makes use of these recent developments in the use of neural networks for solving McKean-Vlasov FBSDEs to solve a portfolio optimization problem where the distribution dependence in the cost functional comes from the risk-appetite of the agent.

1.2 Contributions and challenges

1.2.1 Mean-field stochastic control in infinite horizon and applications

General problem

In the second chapter of the thesis, we first study mean-field control problems (MFC) in the infinite horizon with the distribution of the state appearing in the cost functional but there is no dependence on the distribution of the control. We first study the necessary conditions for optimality which characterizes the control as a function of the coupled FBSDE variables and then the convexity conditions required to prove the uniqueness and sufficient optimality conditions. Analogous results have been well established for the finite horizon settings in the seminal work by Carmona and Delarue [37]. The approach is based on Pontryagin's maximum principle. We first define the control problem,

$$\begin{aligned} dX_t^u &= b(X_t^u, u_t)dt + \sum_{j=1}^m \sigma^j(X_t^u, u_t)dW_t^j, \quad X_0 = x_0; \\ J(u) &= \mathbb{E} \left[\int_0^\infty e^{-\rho t} (f(X_t^u, u_t, \mu_{X_t^u})) dt \right] \end{aligned}$$

Let us define the Hamiltonian as,

$$H(x, u, \mu, y, z) = y b(x, u) + \sum_{j=1}^m z^j \sigma^j(x, u) + f(x, u, \mu) - \rho y x.$$

The adjoint processes $(Y_t, Z_t)_{t \geq 0}$ satisfy the following equations,

$$\begin{aligned} dY_t &= \left(Y_t \partial_x b(\theta_t) + \sum_{j=1}^m Z_t^j \partial_x \sigma^j(\theta_t) + \partial_x f(\theta_t) + \partial_\mu f(\theta_t)(X_t^u) - \rho Y_t \right) dt - \sum_{j=1}^m Z_t^j dW_t^j, \\ \lim_{T \rightarrow \infty} \mathbb{E} [e^{-\rho T} \|Y_T\|^2] &= 0 \end{aligned}$$

Then, the necessary condition for optimality for the above infinite horizon MFC is given by,

Theorem 1.2.1 (Necessary conditions, see Theorem (2.3.1)). *Under certain assumptions, when $u \in \mathcal{A}$ is an optimal control and X is the corresponding state variable and (Y, Z) the corresponding adjoint processes as defined above in (3.3.1), then we must have*

$$\partial_u H(\eta_t) \cdot (u_t - \nu) \leq 0, \quad dt \otimes d\mathbb{P} \text{ a.e. } \forall \nu \in U$$

where η_t denotes $(X_t^u, u_t, \mu_{X_t^u}, Y_t, Z_t)$. Furthermore, if the optimal control u belongs to the interior of U $dt \otimes d\mathbb{P}$ a.e., then

$$\partial_u H(\eta_t) = 0.$$

The maximum principle is very powerful for MFC because the probabilistic approach based of BSDEs provides for efficient numerical solvability in higher dimensions and because it provides explicit representation of the optimal control in terms of the adjoint processes, thanks to the above necessary conditions. Next, we state the sufficient condition for optimality based on strong L-convexity.

Theorem 1.2.2 (Necessary conditions, see Theorem (2.4.1)). *Under certain assumptions, let $u \in \mathcal{A}$ be an admissible control, X_t^u the corresponding state process and (Y, Z) the corresponding adjoint processes. Also, assume that*

$$\mathbb{R}^d \times U \times \mathcal{P}^2(\mathbb{R}^n) \ni (x, u, \mu) \mapsto H(t, x, u, \mu, Y_t, Z_t) \text{ is convex } dt \otimes d\mathbb{P} \text{ a.e.}$$

If

$$H(t, X_t^u, u_t, \mu_{X_t^u}, Y_t, Z_t) = \inf_{u \in U} H(t, X_t^u, u, \mu_{X_t^u}, Y_t, Z_t),$$

$dt \otimes d\mathbb{P}$ a.e. then u is an optimal control.

The above convexity condition is to be understood as follows,

$$\begin{aligned} H(t, x', u', \mu', Y_t, Z_t) &\geq H(t, x, u, \mu, Y_t, Z_t) + \partial_x H(t, x, u, \mu, Y_t, Z_t)(x' - x) \\ &\quad + \partial_u H(t, x, u, \mu, Y_t, Z_t)(u' - u) + \tilde{\mathbb{E}} \left[\partial_\mu H(t, x, u, \mu, Y_t, Z_t)(\tilde{X})(\tilde{X}' - \tilde{X}) \right]. \end{aligned}$$

In addition, under the strong L-convexity condition,

$$\begin{aligned} H(t, x', u', \mu', Y_t, Z_t) &\geq H(t, x, u, \mu, Y_t, Z_t) + \partial_x H(t, x, u, \mu, Y_t, Z_t)(x' - x) + \alpha |q' - q|^2 \\ &\quad + \partial_u H(t, x, u, \mu, Y_t, Z_t)(u' - u) + \tilde{\mathbb{E}} \left[\partial_\mu H(t, x, u, \mu, Y_t, Z_t)(\tilde{X})(\tilde{X}' - \tilde{X}) \right], \end{aligned}$$

with $\alpha > 0$, one can ensure the uniqueness of the optimal control. We also considered various distribution dependence functions which could be studied as an example besides scalar interactions which is where we faced most of the challenges in finding the applicability of this approach. We considered the following functions,

1. Squared Wasserstein's distance between distribution of the state and a target distribution π^* ,

$$W_2^2(\mu_{X_t^u}, \pi^*) = \inf_{\gamma \in \Pi(\mu_{X_t^u}, \pi^*)} \left\{ \int_{\mathbb{R} \times \mathbb{R}} |x - y|^2 d\gamma(x, y) \right\}.$$

The L-differentiability was proved in a recent paper by Alfonsi and Jourdain [5] (see also Gangbo and Tudorascu [61]) under the condition that there is a unique optimal coupling between $\mu_{X_t^u}$ and π^* which is given by a transport map. However, the L-convexity does not hold in general, which was shown by a counter-example in Chapter 5, [37]. However, it is essential for proving the sufficient conditions for optimality.

2. Kullback-Leibler divergence,

$$d_{KL}(\mu|\nu) := \int_{\mathbb{R}^d} \frac{d\mu(x)}{d\nu} \log \left(\frac{d\mu(x)}{d\nu} \right) d\nu(x), \text{ if } \mu \ll \nu \text{ and } +\infty \text{ otherwise.}$$

Due to the requirement of absolute continuity of the measure of the state process, it is not directly applicable.

3. Energy distance (see Rizzo and Székely [108]),

$$D_E(\mu, \nu) = \int_{\mathbb{R}^d} \left| \mathbb{E} \left[e^{i\lambda X} \right] - \mathbb{E} \left[e^{i\lambda Y} \right] \right|^2 w(\lambda) d\lambda.$$

The L-differentiability and L-convexity were not clear for this example.

4. Conditional value at risk (CVaR) or expected shortfall (ES),

$$\text{ES}_\alpha(X) = -\frac{1}{\alpha} \int_0^\alpha \text{VaR}_\gamma(X) d\gamma$$

which is difficult to analyze from the point of view of L-differentiability without knowing the density of the state process a priori which is not the case here as it is a controlled process.

The lack of any conventional working examples posed a major challenge in the thesis when it came to applying the maximum principle approach for MFC, leading us to design a measure dependence on our own which was not scalar interactions and not only satisfies the L-differentiability and L-convexity conditions but was also applicable. This leads us to Chapter 3 of the thesis.

Financial problem

The third chapter is dedicated to a long-term stock portfolio optimization problem modeled as an infinite horizon MFC problem. A practical application of this approach is also under discussion with the asset management team of BNP Paribas in the context of the Chaire Stress Test (research program hosted by Fondation de l'Ecole Polytechnique and BNP Paribas). We start with a collection of N risky assets with prices $(S^i)_{i=1,\dots,N}$ satisfying geometric Brownian motion (GBM) with parameters b_t^i and $(\sigma_t^{i,j})_{j=1,\dots,m}$ and a risk-free asset S^0 satisfying $dS_t^0 = S_t^0 r_t dt$, where r_t is the risk-free rate of return. In general, we could take stochastic parameters but for the numerical simulations, we chose constant b , σ and r . The resulting wealth process for the controlled portfolio is then derived using the self-financing condition which requires the agent to not withdraw from or deposit into the portfolio after $t = 0$,

$$dV_t^u = r_t V_t^u dt + \sum_{i=1}^N u_t^i \left((b_t^i - r_t) dt + \sum_{j=1}^m \sigma_t^{i,j} dW_t^j \right).$$

We also consider a benchmark portfolio which we want to outperform. This portfolio, called 'reference portfolio' henceforth, has static allocations which are not changed after $t = 0$. These allocations are denoted by $(\gamma^i)_{i=0,\dots,N}$. The control q_t is the trading speed of the individual assets and part of the state process is $\mathcal{V}_t^q = V_t/\bar{V}_t$ which is the wealth process corresponding to the controlled portfolio considered in the numeraire of the reference portfolio. The starting point of \mathcal{V}_t^q is hence 1 as we start with the same endowment for both portfolios. The full state process is then given by $X_t^q = [\mathcal{V}_t^q \delta_t^1 \cdots \delta_t^N]^\top$. It can be shown that the state dynamics are given as follows,

$$\begin{aligned} d\mathcal{V}_t^q &= (\alpha_t^\mathcal{V} \mathcal{V}_t^q + \alpha_t^\delta \cdot \delta_t) dt + \sum_{j=1}^m (\beta_t^{\mathcal{V},j} \mathcal{V}_t^q + \beta_t^{\delta,j} \cdot \delta_t) dW_t^j, \\ d\delta_t^i &= q_t^i dt, \quad \delta_{t=0}^i = 0, \quad i \in \{1, 2, \dots, N\}, \end{aligned}$$

giving a SDE for \mathcal{V}_t^q whose drift and volatility are stochastic but bounded and linear in \mathcal{V}^q and δ . The coefficients $\alpha_t^\mathcal{V}, \alpha_t^\delta, \beta_t^{\mathcal{V},j}, \beta_t^{\delta,j}$ given in (3.2.1) are functions of the relative asset prices

$S_t^i = \mathbf{S}_t^i / \bar{V}_t$ and other model parameters. We are looking for the optimal trading speed for each asset which minimizes a cost functional consisting of a quadratic cost in control and a term depending on the target distribution and the distribution of the relative wealth process. The former term helps us to minimize the cost and impact of our trades. The non-stationary target distribution is chosen as the distribution of a GBM with parameters \tilde{b} and $\tilde{\sigma}$, starting from 1 which imposes ideal growth conditions for the relative wealth process. The distribution dependence penalizes deviations from the target distribution in the sense of downside risk. The cost functional is given as follows,

$$J^\lambda(q) = \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(g(X_t^q, q_t) + \int_0^\infty \left[\mathbb{E}[k - \mathcal{V}_t^q]_+^\lambda - p_k \right]_+^\lambda d\nu(k) \right) dt \right].$$

where the parameters p_k come from the target distribution, λ is a smoothing parameter for the function $[k - x]_+$ and ν assign weights to the cost for each level of wealth k . For the numerical scheme, we took the function g to be $\frac{1}{2}|q_t|^2$ but in general one can take any smooth function which is convex in first argument and strongly convex in the second argument. We have shown in Chapter 3 Proposition 3.2.3 that the above function of measure, denoted by $h^\lambda(\mu_{\mathcal{V}_t^q})$ is L-differentiable and L-convex. We also established that it is Lipschitz continuous with respect to the measure of the relative wealth process (in a certain sense) in Proposition 3.2.4.

Next, the infinite horizon system of BSDEs for a given admissible control $(q_t)_{t \geq 0} \in \mathcal{A}$ associated to the control problem is given by,

$$\begin{aligned} dY_t^1 &= - \left(Y_t^1 \alpha_t^\mathcal{V} + \sum_{j=1}^m Z_t^{j,1} \beta_t^{\mathcal{V},j} + \partial_\mu h(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) - \rho Y_t^1 \right) dt + \sum_{j=1}^m Z_t^{j,1} dW_t^j \\ dY_t^i &= - \left(Y_t^i \alpha_t^{\delta,i} + \sum_{j=1}^m Z_t^{j,i} \beta_t^{\delta,j,i} - \rho Y_t^i \right) dt + \sum_{j=1}^m Z_t^{j,i} dW_t^j, \quad i \in \{2, 3, \dots, N+1\}, \\ \lim_{T \rightarrow \infty} \mathbb{E} [e^{-\rho T} |Y_T^i|^2] &= 0. \end{aligned}$$

Theorem 1.2.3 (See Theorem 3.3.2). *For a given admissible control $(q_t)_{t \geq 0}$, under certain assumptions, the BSDE (3.3.1) admits a unique solution $(Y_t, Z_t)_{t \geq 0}$ in the space $\mathcal{S}_\rho^2(\mathbb{R}^{1 \times (N+1)}) \times L_\rho^2(\mathbb{R}^{m \times (N+1)})$. Moreover, there exists $C > 0$ such that for two pairs (Y', Z') and (Y, Z) of solutions to BSDEs corresponding to state-control pairs (X', q') and (X, q) respectively, the following estimate holds:*

$$\mathbb{E} \left[\int_0^\infty e^{-\rho t} |Y_t' - Y_t|^2 dt \right] + \mathbb{E} \left[\int_0^\infty e^{-\rho t} |Z_t' - Z_t|^2 dt \right] \leq C \mathbb{E} \left[\int_0^\infty e^{-\rho t} |q_t' - q_t|^2 dt \right].$$

The assumptions for the existence and uniqueness of the above BSDE impose a lower bound for the discount factor of the control problem when we study the optimality conditions.

Theorem 1.2.4 (Necessary conditions, see Theorem 3.3.3). *Suppose that certain assumptions hold and $q \in \mathcal{A}$ is optimal. Let X be the corresponding state variable and let (Y, Z) be the corresponding adjoint processes as defined above in (3.3.1), then we must have*

$$\partial_q H(X_t^q, q_t, \mu_{\mathcal{V}_t^q}, Y_t, Z_t) = 0, \quad dt \otimes d\mathbb{P} \text{ a.e.}$$

Using this, we can prove that the optimal control is given by $q_t^* = -(Y_t^{[2:(N+1)]})^\top$. For proving the sufficient conditions for optimality, we used an approach similar to [37] but adapted for the infinite horizon setting.

Theorem 1.2.5 (Sufficient conditions, see 3.3.4). *Suppose certain assumptions hold. Let $q \in \mathcal{A}$ be an admissible control, X_t^q the corresponding state process and (Y, Z) the corresponding adjoint processes. Also, assume the Hamiltonian satisfies:*

$$\mathbb{R}^{N+1} \times \mathbb{R}^N \times \mathcal{P}^2(\mathbb{R}) \ni (x, q, \mu) \mapsto H(t, x, q, \mu, Y_t, Z_t) \text{ is } \alpha\text{-strongly convex } dt \otimes d\mathbb{P} \text{ a.e.,}$$

which is to be understood as follows,

$$\begin{aligned} H(t, x', q', \mu', Y_t, Z_t) &\geq H(t, x, q, \mu, Y_t, Z_t) + \partial_x H(t, x, q, \mu, Y_t, Z_t)(x' - x) \\ &\quad + \partial_q H(t, x, q, \mu, Y_t, Z_t)(q' - q) + \mathbb{E} [\partial_\mu h(\mu)(X)(X' - X)] + \alpha|q' - q|^2 \end{aligned}$$

where $X \sim \mu$ and $X' \sim \mu'$ and $\alpha > 0$. If

$$H(t, X_t^q, q_t, \mu_{X_t^q}, Y_t, Z_t) = \inf_{u \in \mathbb{R}^N} H(t, X_t^q, u, \mu_{X_t^q}, Y_t, Z_t), \quad dt \otimes d\mathbb{P} \text{ a.e.}$$

then q is the unique optimal control in \mathcal{A} .

It can be proved that the candidate for the control q^* from the necessary conditions satisfies the sufficient conditions for optimality as well. This gives us a unique system of coupled McKean-Vlasov FBSDEs corresponding to our control problem when we plug $q_t^* = -(Y_t^{[2:(N+1)]})^\top$ into the dynamics of the state process. The existence and uniqueness of this McKean-Vlasov FBSDE is established in the following result:

Theorem 1.2.6 (FBSDE existence and uniqueness, see Theorem 3.3.6). *When certain assumptions hold, the FBSDE (3.3.4) has a unique solution $(X_t, Y_t, Z_t)_{t \geq 0}$ in the space $L_\rho^2(\mathbb{R}^{N+1}) \times L_\rho^2(\mathbb{R}^{1 \times (N+1)}) \times L_\rho^2(\mathbb{R}^{m \times (N+1)})$.*

Numerical scheme

For numerical purposes, we consider a finite time approximation of our method for a large $T > 0$ and set $Y_T = 0$ as the terminal condition and $Y_t = 0$ for $t > T$. We also use truncation in time and provide a form of the truncation approximation error in Proposition 3.3.7. In order to deal with the distribution dependence, we use the approach of [78]. We use fictitious play iterations where each iteration has a pre-specified dependence in the form of a function of time. We solve the resulting coupled FBSDE by considering a decoupling field representation for (Y, Z) using feed-forward neural networks as approximations of Y_0 and $(Z_{t_i})_{i=0, \dots, N}$ at each time step which feeds from the previous time step, creating a deep network of sub-networks. This kind of network was also used by [77] for coupled FBSDEs. The main idea of the scheme is to solve the equations forward in time and at the end, penalize deviations from the terminal condition and train the neural network to minimize these deviations. For more details on a simpler architecture of the neural network for BSDEs, see [56]. Next, we precisely describe the scheme. The asset price processes can be explicitly written as,

$$\begin{aligned} \mathbf{S}_{t_k}^i &= \mathbf{S}_0^i \exp \left(\left(b^i - \frac{\sum_{j=1}^m (\sigma^{i,j})^2}{2} \right) t_k + \sum_{j=1}^m \sigma^{i,j} W_{t_k}^{a,j} \right), \\ \mathbf{S}_{t_k}^0 &= \mathbf{S}_0^0 e^{rt_k}. \end{aligned}$$

Then, the relative asset price process S^i is given by, $S_{t_k}^i = \frac{\mathbf{S}_{t_k}^i}{\sum_{i=0}^N \gamma^i \mathbf{S}_{t_k}^i}$.

We take the target distribution π_t^* for the relative wealth process \mathcal{V}_t^q at time t to be the distribution of the following geometric Brownian motion (GBM): $dP_t^* = P_t^*(\mu dt + \sigma dW_t)$ with $P_0^* = 1$, $\mu = \sigma = 0.05$. This sets both the desired growth rate and permitted variation of

the relative wealth process once we calculate the parameters $p_k(t)$. These can be calculated as follows: When \mathcal{V}_t^q attains the target distribution, $\left[\mathbb{E}[k - P_t^*]_+^\lambda - p_k(t)\right]_+^\lambda = 0$, which means that $p_k(t)$ needs to be larger than or equal to $\mathbb{E}[k - P_t^*]_+^\lambda$. We can take

$$p_k(t) = \mathbb{E}[k - P_t^*]_+ = e^{\mu t} \mathbb{E} \left[k e^{-\mu t} - e^{\sigma W_t - \frac{1}{2} \sigma^2 t} \right]_+,$$

which can be solved explicitly using the Black-Scholes formula. With this, we can write the L-derivative as a function of the state variable and expectation of a function of the state,

$$\partial_\mu h(\mu)(\mathcal{V}_t^q) = \int_0^\infty \frac{-1}{1 + e^{\lambda p_k - (\mathbb{E}[\log(1 + e^{\lambda(k - \mathcal{V}_t^q)])])}} \frac{1}{1 + e^{\lambda(\mathcal{V}_t^q - k)}} d\nu(k) \equiv H(\mathcal{V}_t^q, \mathbb{E}[\theta(\mathcal{V}_t^q)]),$$

where $\theta(x) = \log(1 + e^{\lambda(k - \mathcal{V}_t^q)})$.

For solving the coupled FBSDE, we use decoupling fields as functions of the state and relative asset price process: $Y_0 = u(X_0, S_0)$; $Z_t = v_t(X_t, S_t, Y_t)$. The scheme is outlined as follows (with p standing for fictitious play iteration):

1. Initialize $(\theta^0(t_i))_{i \in \{1, 2, \dots, N\}}$ as the guess function for $\mathbb{E}[\theta(\mathcal{V}_t)]$. Initialize the neural networks $u^0(x, s)$ and $\{v_{t_i}^0(x, s, y)\}_{i \in \{1, 2, \dots, N+1\}}$ with random weights.
2. Find the corresponding solution $(X_{t_i}^0, Y_{t_i}^0, Z_{t_i}^0)_{0 \leq i \leq N}$ by the forward Euler scheme for the resulting FBSDE, which is given as follows for any iteration p ,

$$\begin{aligned} X_0^p &= [1, 0, \dots, 0]; \quad Y_0^p = u^p(X_0, S_0); \quad Z_0^p = v_0^p(X_0^p, S_0, Y_0^p); \\ X_{t_{i+1}}^p &= X_{t_i}^p + b(t_i, X_{t_i}^p, Y_{t_i}^p) + \sigma(t_i, X_{t_i}^p) \Delta W_i; \\ Z_{t_i}^p &= v_{t_i}^p(X_{t_i}^p, S_{t_i}, Y_{t_i}^p); \\ Y_{t_{i+1}}^p &= Y_{t_i}^p - f(t_i, X_{t_i}^p, \theta^{p-1}(t_i), Y_{t_i}^p, Z_{t_i}^p) + Z_{t_i}^p \Delta W_i. \end{aligned}$$

3. Train the neural networks with this data to minimize the following,

$$\mathbb{E}[|Y_T^0|^2] \approx \frac{1}{D} \sum_{a=1}^D |Y_{t_N}^{0,a}|^2,$$

since the terminal condition of our truncated problem is $Y_T = 0$. This gives us the next iterates u^1 and $(v_{t_i}^1)_{1 \leq i \leq N}$ of the approximation of decoupling fields. Update the estimate

$$\text{of } \theta^p(t_i) \text{ with } \theta^p(t_i) = \frac{1}{D} \sum_{a=1}^D \theta(\mathcal{V}_{t_i}^{p-1}).$$

4. Solve the corresponding FBSDE using the Euler Scheme as before and iterate similarly.

We have successfully shown that the cumulative distribution function skews towards right as we increase a constant multiplier for the mean-field term, increasing the average wealth attained throughout the time horizon and also increasing the probability of higher wealth ($\mathcal{V}_t \geq 1$) than starting point while not increasing the probability of lower wealth ($\mathcal{V}_t \leq 1$) by much. We have also shown that the mean-field term, which can be seen as the deviation from the target distribution, decreases with time consistently throughout our tests.

We faced two major challenges in the numerics: 1. Numerical instability of the scheme when choosing high values of λ (smoothing parameter), which is essential of better approximation of the L-derivative; 2. Numerical instability with respect to the multiplier in front of the mean-field term. Smaller values of this parameter causes the trajectories of the L-derivative (which is a random variable) to be somewhat close, resulting in the trajectories of the trading speed predictions to be closely packed.

1.2.2 BSDEs in infinite horizon

In the second part of the thesis, our focus is on developing efficient algorithms for numerically approximating Markovian BSDEs in infinite horizon, which is equivalent to solving elliptic PDEs as remarked earlier. The BSDE is given as follows,

$$Y_t = Y_T + \int_t^T f(X_s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad \forall 0 \leq t \leq T < +\infty,$$

where the solution process (Y, Z) takes values in some appropriate space and X is the solution of the following SDE:

$$X_t = x + \int_0^t b(X_s) ds + \int_0^t \sigma(X_s) dW_s, \quad \forall t \geq 0.$$

We recall the general existence and uniqueness result for this infinite horizon BSDE, for which we refer to Theorem 5.57 from [100] for the multidimensional setting and Lemma 3.1 from [30] for the one-dimensional setting.

Proposition 1.2.1. (See Proposition 4.2.1) Let us assume that Assumption 4.2.1 is in force.

1. If there exists $0 < \rho < 2\mu - K_{f,z}^2$ such that, for all $x \in \mathbb{R}^d$,

$$\mathbb{E} \left[\int_0^{+\infty} e^{-\rho s} |X_s^x|^{2r} ds \right] < +\infty,$$

then, for all $x \in \mathbb{R}^d$, BSDE (4.2.1) has a unique solution $(Y^x, Z^x) \in \mathcal{S}_{loc}^2 \times \mathcal{M}_{loc}^2$ such that

$$\lim_{T \rightarrow +\infty} \mathbb{E} [e^{-\rho T} |Y_T^x|^2] = 0.$$

Moreover this solution satisfies

$$\mathbb{E} \left[\int_0^{+\infty} e^{-\rho s} \|Z_s^x\|^2 ds \right] + \mathbb{E} \left[\sup_{0 \leq s < +\infty} e^{-\rho s} |Y_s^x|^2 \right] \leq C \left(1 + \mathbb{E} \left[\int_0^{+\infty} e^{-\rho s} |X_s^x|^{2r} ds \right] \right),$$

with C a constant that does not depend on x .

2. If $d' = 1$ (i.e. Y^x is scalar valued) and $r = 0$ then, for all $x \in \mathbb{R}^d$, BSDE (4.2.1) has a solution such that Y^x is a continuous and bounded process and (Y^x, Z^x) satisfies

$$\operatorname{ess\,sup}_{\Omega \times \mathbb{R}^+} |Y^x| + \mathbb{E} \left[\int_0^{\infty} e^{-2\mu s} \|Z_s^x\|^2 ds \right] \leq C,$$

with C a constant that does not depend on x . Moreover this solution is unique in the class of processes (Y^x, Z^x) such that Y^x is continuous and bounded and $Z^x \in \mathcal{M}_{loc}^2$.

Under these assumptions, we have a Markovian representation $Y_t = u(X_t)$ where u is a continuous function. In addition to these assumptions, we assume continuous differentiability of b and σ , boundedness of σ and a strong ellipticity condition for σ , all of which ensure the differentiability of $(X_t^x)_{t \geq 0}$ with respect to x and we denote the flow process by $(\nabla_x X_t^x)_{t \geq 0}$ and define a Malliavin weight $(U_t^x)_{t \geq 0}$,

$$U_t^x = \frac{1}{t} \int_0^t \langle \sigma^{-1}(X_s^x) \nabla_x X_s^x, dW_s \rangle \sigma(x) = (U_t^{x,j})_{1 \leq j \leq d},$$

using which we can establish the following representation of the solution of BSDE, which will be very helpful for designing the numerical schemes.

Proposition 1.2.2. (See Proposition 4.2.3) Let us assume that there exists $0 < \rho < 2\mu - K_{f,z}^2 \mathbf{1}_{d'>1}$ and $C > 0$ such that, for all $x \in \mathbb{R}^d$

$$\mathbb{E} \left[\int_0^{+\infty} e^{-\rho s} |X_s^x|^{2r} ds \right] \leq C(1 + |x|^{2r}).$$

We consider two parameters $a > \mu - \frac{K_{f,z}^2}{2} \mathbf{1}_{d'>1} > 0$ and $\tilde{a} \geq 2\mu - K_{f,z}^2 \mathbf{1}_{d'>1} > 0$ such that for any compact set $K \subset \mathbb{R}^d$ we have

$$\int_0^{+\infty} e^{-\tilde{a}s} \sup_{x \in K} \mathbb{E} [|X_s^x|^{2r}] ds + \sup_{x \in K} \mathbb{E} \left[\int_0^{+\infty} e^{-\tilde{a}s} \|\nabla_x X_s^x\|^2 ds \right] < +\infty.$$

Then, under extra Assumption 4.2.2, there exists a function $u \in C^1(\mathbb{R}^d, \mathbb{R}^{d'})$ such that, for all $x \in \mathbb{R}^d$, $Y^x = u(X^x)$, $\nabla_x u(X_x) \sigma(X^x) := \bar{u}(X^x)$ is a continuous version of Z^x and we have the following probabilistic representation for (u, \bar{u}) : for all $x \in \mathbb{R}^d$,

$$\begin{aligned} u(x) &:= Y_0^x = \mathbb{E} \left[\int_0^\infty e^{-as} (f(X_s^x, u(X_s^x), \bar{u}(X_s^x)) + au(X_s^x)) ds \right] \\ &= \frac{1}{\theta} \mathbb{E} \left[(f(X_E^x, u(X_E^x), \bar{u}(X_E^x)) + au(X_E^x)) e^{-(a-\theta)E} \right] \end{aligned}$$

and

$$\begin{aligned} \bar{u}(x) &:= Z_0^x = \mathbb{E} \left[\int_0^\infty e^{-\tilde{a}s} (f(X_s^x, u(X_s^x), \bar{u}(X_s^x)) + \tilde{a}u(X_s^x)) U_s^x ds \right] \\ &= \frac{\sqrt{\pi}}{\sqrt{\tilde{\theta}}} \mathbb{E} \left[(f(X_{\tilde{E}}^x, u(X_{\tilde{E}}^x), \bar{u}(X_{\tilde{E}}^x)) + \tilde{a}u(X_{\tilde{E}}^x)) \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} U_{\tilde{E}}^x \right] \end{aligned}$$

where $\theta > 0$, $\tilde{\theta} > 0$, $E \sim \mathcal{E}(\theta)$ is independent of W and $\tilde{E} \sim \Gamma(1/2, \tilde{\theta})$ is independent of W . Lastly, we have the following growth:

$$|u(x)| + \|\bar{u}(x)\| \leq C(1 + |x|^r), \quad \forall x \in \mathbb{R}^d.$$

Numerical Schemes

For brevity of the argument, we state the scheme with $X_t^x = W_t^x$ which is a Brownian motion starting from $x \in \mathbb{R}^d$ but it can be easily extended to a general SDE satisfying above assumptions. A common way to test numerical schemes for BSDEs with analytical solutions is to start with a given function $u(x) \in C^2(\mathbb{R}^d)$ with at most polynomial growth or degree r (to be consistent with Proposition 1.2.2) and construct the function f such that the BSDE is satisfied by $Y_t^x = u(W_t^x)$ and $Z_t^x = \nabla_x u(W_t^x)$. This can be done by choosing a suitable function $f_0(y, z)$ which satisfies the monotonicity and Lipschitz conditions on y and z components respectively. Then, take the generator of the BSDE as follows:

$$f(x, y, z) = f_0(y, z) - \frac{1}{2} \Delta u(x) - f_0(u(x), \nabla_x u(x)),$$

where Δ denotes the Laplacian operator. This works because $u(x)$ satisfies the elliptic PDE, $\frac{1}{2} \Delta u(x) + f(x, u(x), \nabla u(x)) = 0$.

1. Grid based scheme

The first scheme is based on taking a finite sub-grid $\Pi \subset \delta\mathbb{Z}^d$ in \mathbb{R}^d , with $2N + 1$ divisions along each axis. We denote by \square its convex hull of this. By convention, we can assume that $0 \in \square$. We then define a multi-linear interpolation on Π with the following basis functions,

$$\forall z \in \Pi, \quad \forall x \in \mathbb{R}^d, \quad \psi_z(x) = \prod_{i=1}^d (1 - |\delta^{-1}(x_i - z_i)|)_+.$$

For a function $\phi : \Pi \rightarrow \mathbb{R}$, we define $P\phi$ the interpolation of ϕ on \mathbb{R}^d as follows

$$P\phi(x) =: \begin{cases} \sum_{z \in \Pi} \psi_z(x) \phi(z), & x \in \square \\ \sum_{z \in \Pi} \psi_z(\text{Proj}(x, \square)) \phi(z), & x \notin \square. \end{cases}$$

The Malliavin weight for this case is given by $U_t^x = \frac{W_t^x}{t}$. The scheme is based on a fixed point argument motivated from the representation in Proposition 1.2.2. The full scheme is given as follows:

1. With $\theta < a$ and $\tilde{\theta} < b$, we simulate M samples of an exponential r.v. $E \sim \mathcal{E}(\theta)$ and a r.v. with gamma distribution $\tilde{E} \sim \Gamma(\frac{1}{2}, \tilde{\theta})$ and start with the initial guess $v^0(x) = (u(x), \bar{u}(x)) = (0, 0) \forall x$.
2. For a given n , evaluate $Pv^n = (Pu^n, P\bar{u}^n)$ for every grid point z and update v^n to Pv^n .
3. Next, we define the operator ϕ and $\bar{\phi}$ as follows

$$\begin{aligned} \phi(u^n)(x) &= \frac{1}{\theta M} \sum_{j=1}^M \left(f(W_E^{x,j}, u^n(W_E^{x,j}), \bar{u}^n(W_E^{x,j})) + au^n(W_E^{x,j}) \right) e^{-(a-\theta)E} \\ \bar{\phi}(u^n)(x) &= \frac{\sqrt{\pi}}{\sqrt{\tilde{\theta}} M} \sum_{j=1}^M \left(f(W_{\tilde{E}}^{x,j}, u^n(W_{\tilde{E}}^{x,j}), \bar{u}^n(W_{\tilde{E}}^{x,j})) + bu^n(W_{\tilde{E}}^{x,j}) \right) \sqrt{\tilde{E}} e^{-(b-\tilde{\theta})\tilde{E}} U_{\tilde{E}}^x \end{aligned}$$

4. Compute $\|u^n\|_\rho \approx \max_{x \in \Pi} \frac{|u^n(x)|}{\rho(x)}$ and $\|\bar{u}^n\|_\rho \approx \max_{x \in \Pi} \frac{|\bar{u}^n(x)|}{\rho(x)}$.
5. We update our guess for v as follows:

$$\begin{aligned} u^{n+1} &= \phi(u^n), \\ \bar{u}^{n+1} &= \bar{\phi}(\bar{u}^n). \end{aligned}$$

We also have the following study of the numerical error for this scheme:

Theorem 1.2.1. Let us assume that assumptions of Proposition 4.2.3 are fulfilled. We set $r' > r$ where r comes from (4.2.9). We also assume that $\kappa_\infty < 1$ and $v \in C^2(\mathbb{R}^d, \mathbb{R}^d \times \mathbb{R}^{d' \times d})$ with bounded second derivatives.

Then the following holds.

1. If σ is constant, $\tilde{a} - \tilde{\theta} - (1+r)K_b > 0$ and $a - \theta - rK_b > 0$, then

$$\mathbb{E} \left[\sup_{x \in \mathbb{R}^d} \left\| \frac{Pv_M^n(x) - v(x)}{\rho(x)} \right\| \right] = O \left(\delta^2 + \frac{(\ln \tilde{N})^{1/(r+1)}}{\sqrt{M}} + |\tilde{N}\delta|^{r-r'} + \kappa_\infty^n \right).$$

2. If $f(x, y, z) = f(x, y)$ and $a - \theta - rK_b > 0$, then

$$\mathbb{E} \left[\sup_{x \in \mathbb{R}^d} \left\| \frac{Pv_M^n(x) - v(x)}{\rho(x)} \right\| \right] = O \left(\delta^2 + \frac{(\ln \tilde{N})^{1/r}}{\sqrt{M}} + |\tilde{N}\delta|^{r-r'} + \kappa_\infty^n \right).$$

3. If $\beta \geq 2$, $\tilde{a} - (r+1)K_b - \frac{2\beta-1}{2}K_\sigma^2 - \frac{\beta-1}{\beta}\tilde{\theta} > 0$ and $a - rK_b - \frac{\beta-1}{\beta}\theta > 0$, then

$$\mathbb{E} \left[\sup_{x \in \mathbb{R}^d} \left\| \frac{Pv_M^n(x) - v(x)}{\rho(x)} \right\| \right] = O \left(\delta^2 + \frac{(\tilde{N})^{d/\beta}}{\sqrt{M}} + |\tilde{N}\delta|^{r-r'} + \kappa_\infty^n \right).$$

The scheme is efficient for low dimensions but time complexity increases exponentially with d . To overcome the curse of dimensionality and the boundary error, we propose a similar contraction based scheme but with neural networks.

2. NN scheme with contraction

We use once again the contraction property by considering a sequence of neural networks $(\mathcal{U}_{\theta_n, n})_{n \in \mathbb{N}}$, i.e. a sequence of functions $x \mapsto (\mathcal{U}_{\theta_n, n}^1, \mathcal{U}_{\theta_n, n}^2) \in \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d}$ parametrized by $\theta_n \in \Theta$. We set $\mathcal{U}_{\theta_0, 0} = 0$. Knowing the neural network $\mathcal{U}_{\theta_n, n}$, we train $\mathcal{U}_{\theta_{n+1}, n+1}$ as follows:

1. Sample $X_0 = X_{t=0} \sim \mu_0$, $E \sim \mathcal{E}(\theta)$ and $\tilde{E} \sim \Gamma(1/2, \tilde{\theta})$.
2. Get corresponding samples of the Malliavin weight $U_E^{X_0}$ and $U_{\tilde{E}}^{X_0}$ at times E and \tilde{E} for the starting point X_0 . Also, obtain samples of the state process $X_E^{X_0}$ and $X_{\tilde{E}}^{X_0}$.
3. Evaluate the previous neural network $\mathcal{U}_{\theta_n, n}$ at each sample of $X_E^{X_0}$ and $X_{\tilde{E}}^{X_0}$ and compute the samples of $\varphi(\mathcal{U}_{\theta_n, n})(X_0)$ given by,

$$\varphi(\mathcal{U}_{\theta_n, n})(X_0) = \left(\frac{1}{\theta} (f(X_E^{X_0}, \mathcal{U}_{\theta_n, n}^1(X_E^{X_0}), \mathcal{U}_{\theta_n, n}^2(X_E^{X_0})) + a\mathcal{U}_{\theta_n, n}^1(X_E^{X_0})) e^{-(a-\theta)E}, \right. \\ \left. \sqrt{\frac{\pi}{\tilde{\theta}}} \left(f(X_{\tilde{E}}^{X_0}, \mathcal{U}_{\theta_n, n}^1(X_{\tilde{E}}^{X_0}), \mathcal{U}_{\theta_n, n}^2(X_{\tilde{E}}^{X_0})) + \tilde{a}\mathcal{U}_{\theta_n, n}^1(X_{\tilde{E}}^{X_0}) \right) \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} U_{\tilde{E}}^{X_0} \right).$$

4. Minimize the loss function,

$$\mathcal{L}(\mathcal{U}_{\theta, n+1}) := \mathbb{E} [\|\mathcal{U}_{\theta, n+1}(X_0) - \varphi(\mathcal{U}_{\theta_n, n})(X_0)\|^2]$$

and choose the parameters θ_{n+1} ,

$$\theta_{n+1} \in \arg \min_{\theta \in \Theta} \mathbb{E} [\|\mathcal{U}_{\theta, n+1}(X_0) - \varphi(\mathcal{U}_{\theta_n, n})(X_0)\|^2].$$

We also have following result for the numerical approximation error of the scheme:

Proposition 1.2.3. Let $\varepsilon > 0$ be such that $(1+\varepsilon)(\kappa_2)^2 < 1$. Then, the numerical approximation error e_n of the scheme at n -th Picard iteration is bounded by,

$$e_n \leq \frac{(1+\varepsilon) \max_{k \in \{1, \dots, n\}} r_k}{\varepsilon - \varepsilon(1+\varepsilon)(\kappa_2)^2} + [(1+\varepsilon)(\kappa_2)^2]^n e_0,$$

where $r_n := \mathbb{E} [\|\mathcal{U}_{\theta_{n+1}, n+1}(X_0) - \varphi(\mathcal{U}_{\theta_n, n})(X_0)\|^2]$ is the approximation error of the neural network.

This scheme is highly efficient even for high dimensional problems but to ensure contraction, we had to impose additional assumptions. These could be relaxed for the next scheme which does not rely on contraction arguments.

3. NN scheme without contraction

We do not rely on the contraction and try to get a direct neural network approximation. We train the neural network \mathcal{U}_θ in epochs, each with \mathcal{P} gradient descent steps, whereby the NN is optimized for a given set of samples before we update the input samples in the next epoch. We use the same operator φ defined for previous scheme. At the start of each epoch, we do the following: Sample $X_0 = X_{t=0} \sim \mu_0$; For each sample of X_0 , obtain M independent samples of $(E, \tilde{E}, X_E^{X_0}, X_{\tilde{E}}^{X_0}, U_E^{X_0}, U_{\tilde{E}}^{X_0})_j$ for $j = 1, \dots, M$.

Then, within a given epoch i with the M independent samples $\left((E, \tilde{E}, X_E^{X_0}, X_{\tilde{E}}^{X_0}, U_E^{X_0}, U_{\tilde{E}}^{X_0})_{j=0, \dots, M} \right)_i$, we do the following for the next \mathcal{P} gradient descent steps, $n \in \{1, \dots, \mathcal{P}\}$:

1. With the current parameters θ_n^i , we compute the corresponding M samples $(\varphi^j(\mathcal{U}_{\theta_n^i})(X_0))_{1 \leq j \leq M}$ for each sample of X_0 which is a function of only θ_n^i .
2. The parameters θ_{n+1}^i are obtained using the optimization step with the following loss function,

$$\Upsilon_M(\mathcal{U}_\theta) := \mathbb{E} \left[\left\| \mathcal{U}_\theta(X_0) - \frac{1}{M} \sum_{j=1}^M \varphi^j(\mathcal{U}_\theta)(X_0) \right\|^2 \right].$$

This scheme overcomes the difficulties of both the above schemes which we demonstrate with the plot of mean squared error and the Lipschitz constant of f with respect to the third argument. It shows a point beyond which even though the assumptions hold, the contraction fails and so does the second scheme but the third scheme still has reasonably small error. However, unlike the other two schemes, the dimension of the sample is significantly higher and we face difficulty with the numerical analysis of the scheme.

1.3 Conclusions

In Chapter 2, we study the mean-field stochastic control problem in infinite horizon where the cost functional has dependence on the law of the state process. We proved the necessary and sufficient conditions of optimality which required L-differentiability and L-convexity in the measure space for the running cost function. When this dependence is through scalar interactions, the problem is simpler and has been studied by researchers extensively over the years. We looked for examples of functions of measure other than scalar interactions which satisfy these conditions while also being relevant from applications point of view, demonstrating that this is indeed a difficult problem.

In Chapter 3, we start with an application in portfolio optimization of the MFC problem introduced in Chapter 2. The goal is to outperform a static allocation of stocks and a risk-free asset by using dynamic allocation, using the trading speed of the assets as the control with a downside risk minimization criterion of mean-field type. This was done by finding an example of a L-differentiable and L-convex function of the law of the state process. We proved the conditions of optimality for the control problem and established the existence and uniqueness of the solution to the corresponding system of coupled McKean-Vlasov FBSDE. We also developed a numerical scheme based on neural networks for solving a time-truncated version of the problem and provided exponential bounds for the truncation error. We showed that increasing the multiplier of the mean-field term successfully skewed the wealth distribution towards right, increasing the probability of higher relative wealth. However, the scheme is unstable with respect to the smoothing parameter for the L-derivative (which needs to be higher to see more variability in the trajectories) and the multiplier of the mean-field term which inhibits us from taking aggressive positions.

In Chapter 4, we study numerical approximation of BSDEs in infinite horizon. We developed three numerical schemes: The first scheme is based on a fixed point argument and uses grid approximation for the space; The second scheme is also based on fixed point argument but does not discretize space and uses neural networks; The third scheme does not depend on fixed point argument and uses neural networks like the second scheme. We also provide a detailed study of the numerical error for the first scheme and prove tight bounds on the approximation error, requiring additional assumptions for contraction. For the second scheme, we proved the convergence of the approximation error to zero as the size of the neural network increases. We have not proved such numerical guarantees for the third scheme but we have shown that it performs better than the first two schemes as it does not require the additional assumptions for contraction.

Future Perspectives

There are a lot of interesting directions of future research that have emerged while pursuing the thesis. In Chapter 3, one could come up with another function of the measure of the state process which could overcome the problem of stability as pointed out earlier and we could avoid closely packed trading speed trajectories. It would provide an alternative perspective to model the problem with the control being the asset allocation and the cost functional being a function of the law of the control. It would also be interesting to try other numerical schemes for McKean-Vlasov FBSDEs which do not require fictitious play iterations as done recently by [63].

In Chapter 4, obtaining estimates for the approximation error for the third scheme could be very interesting. Also, verification of $\kappa_p < 1$ using a version of Aronson's estimates for infinite horizon would be useful.

Part I

Mean-field stochastic control in infinite horizon and applications

Chapter 2

A study of mean-field infinite horizon stochastic control

Abstract. We study a mean-field stochastic control (MFC) problem in the infinite horizon setting where the running cost function depends on the law of the state process. We use Pontryagin’s stochastic maximum principle and the notion of L-differentiation with respect to the law of the state process to prove a necessary condition for optimality of a control which reduces the problem to point-wise minimization in the control space and gives the form of the optimal control in terms of the state and adjoint processes which satisfy a BSDE. We use the notion of L-convexity to prove a sufficient condition and uniqueness for the optimal control. Using the form of the optimal control in terms of the state and adjoint processes, we write a coupled system of McKean-Vlasov FBSDE. We later provide potential applications and challenges when using some standard mean-field dependencies.

Contents

2.1	Introduction	47
2.2	Preliminaries	49
2.2.1	Differentiability on the space of measures	49
2.2.2	The Model	50
2.2.3	A useful estimate	52
2.2.4	Adjoint processes	53
2.3	Necessary conditions for optimality	54
2.4	Sufficient condition	59
2.5	Examples	61
2.5.1	The mean-field infinite horizon LQ case	61
2.5.2	Optimal liquidation problem in infinite horizon	63
2.5.3	Challenges with standard examples	63

2.1 Introduction

Mean field control (MFC) problems originated from mean field games (MFG) which provide a powerful framework for modeling strategic interactions of a large number of rational agents, originally introduced by Lasry and Lions [85] and Huang et al. [80]. Theoretical frameworks, such as the mean-field game formulation by Pardoux et al. [102], have laid the groundwork

for understanding the equilibrium behavior of large populations in the presence of mean-field interactions. Their work provides a comprehensive overview of mean-field games and their applications, offering insights into the mathematical foundations of these models. The study has served as a cornerstone for subsequent theoretical developments in the field. For a broader coverage and perspectives for mean-field games, see Bensoussan, Frehse, Yam, et al. [18]. MFCs are single agent control problems with the objective function generally depending on the law of the state and/or control process. An extensive study of MFGs and MFCs along with applications can be found in the two volumes of books written by Carmona and Delarue [37], which have arguably provided one of the best detailed coverage of the topic. The optimization of energy storage strategies is an emerging application of FBSDEs. De Paola et al. [49] and Li et al. [87] proposed FBSDEs framework to model the optimal operation of energy storage systems, considering uncertainties in electricity prices and renewable energy generation. This kind of approach has implications for improving the efficiency and economic viability of energy storage technologies. Mean-Field FBSDEs are gaining traction in modeling the dynamics of interconnected power systems. Alasseur et al. [4] introduced mean-field FBSDEs for power grid stability analysis, considering the interactions among multiple generators and loads. This approach facilitates a holistic understanding of the grid's behavior under various operating conditions. Interdisciplinary collaborations between mathematicians, economists, and engineers have led to advancements in the application of stochastic control methodologies in the energy sector. [?] integrated stochastic control with machine learning techniques for optimal real-time operation of energy storage systems, showcasing the potential of combining diverse methodologies. Acciaio, Backhoff-Veraguas, and Carmona [1] studied an 'extended' MFG using the maximum principle and provided insights into the connection with the theory of optimal transport and solved an optimal liquidation problem. Aïd, Basei, and Pham [3] used the McKean-Vlasov approach for modeling distributed energy generation. More recent works include the use of McKean-Vlasov stochastic control for the energy management of a smart grid by Gobet and Grangereau [64] and in [65], they developed a novel approach combining Newton's method and Pontryagin's principle for stochastic control. Germain et al. [62] used mean-field dynamics with probabilistic state constraints and level-set approach for renewable energy storage problem.

Andersson and Djehiche [9] proved sufficient conditions for optimality for the general scalar interactions without the need for L-convexity. Li et al. [88] obtained explicit solutions in terms of Riccati equations for the infinite time horizon in the general mean field LQ setting. Pham and Wei [105] deal with the Dynamic programming approach, bypassing the notions of L-differentiability and L-convexity but do not discuss sufficient conditions for optimality. The literature on different notions of convexity of functions of measure is richer in the field of optimal transport. In optimal transport theory, the notion of displacement convexity is most popular (see Definition 5.77 in [37]). Ohta and Takatsu [97] extensively studied the conditions for displacement convexity of relative entropy. and Proposition 5.79 in [37] proves the equivalence of L-convexity and displacement convexity for continuously L-differentiable functions. The intersection with optimal transport theory and hence the applicability of this notion of convexity is very interesting in itself.

Statement of the Problem. In this chapter, we study an infinite horizon mean-field stochastic control (MFC) problem where the running cost function depends on the law of the state process. The law dependence requires a notion of differentiation and convexity of functions (in the differential form) defined on the space of measures, for which we use the notion of L-differentiability and L-convexity respectively (see Chapter 5 of Carmona and Delarue [37] for more details). We use stochastic Pontryagin maximum principle which leads to necessary optimality conditions imposed on the Hamiltonian of the problem which result into a point-wise minimization problem, giving the explicit form of the control in terms of the state and adjoint processes which satisfy a BSDE. If one includes a dependence on the law of the control in the state dynamics

and/or the cost functional, we call the problem extended MFC (see Acciaio et al. [1] for results on the optimality conditions for extended MFC). However, the dependence on the law of the control results into an additional term in the necessary conditions, which is the L-derivative of the Hamiltonian with respect to the law of control which is in general a random variable which is a function of the control variable. Due to this, one cannot convert the problem into a point-wise minimization problem as illustrated with a counter-example in Remark 4.2 [1].

Outline. In section 2, we introduce some basic notation and discuss the notions of L-differentiability and L-convexity. Also, we describe the problem formally and define the adjoint processes. In section 3, we use the maximum principle to prove a necessary condition for optimality in Theorem 2.3.1. In section 4, we prove a sufficient condition for an admissible control to be optimal and unique under a joint strong L-convexity condition with respect to the control in Theorem 2.4.1. Section 5 discuss two examples with scalar interactions: Infinite horizon mean-field LQ control and optimal liquidation problem in the infinite horizon. Also, we describe the challenges we faced when finding an example outside of scalar interaction, where we tried some standard mean-field functions of measure.

2.2 Preliminaries

2.2.1 Differentiability on the space of measures

We consider a MFC problem with the running cost function depending on the law of the state process. We use Pontryagin's maximum principle, which is convenient in our setting due to the probabilistic notions of L-differentiability and L-convexity on the space of measures (see [37] for more details). Let us consider a space of measures defined on \mathbb{R}^d with finite second moments, denoted by $\mathcal{P}^2(\mathbb{R}^d)$. The squared 2-Wasserstein distance between measure $\mu, \nu \in \mathcal{P}^2(\mathbb{R}^d)$ is defined as,

$$W_2^2(\mu, \nu) = \inf_{\gamma \in \Pi(\mu, \nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\gamma(x, y),$$

where $\Pi(\mu, \nu)$ is the space of all couplings of the measures μ and ν . It can also written in probabilistic terms as follows,

$$W_2^2(\mu, \nu) = \inf_{X \sim \mu, Y \sim \nu} \mathbb{E} [|X - Y|^2],$$

which will be more useful for us in the proofs. Consider a function $f : \mathcal{P}^2(\mathbb{R}^d) \rightarrow \mathbb{R}$ and let $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ be the space of square integrable random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and denote the law of a random variable X by $\mathcal{L}(X)$.

Definition 2.2.1 (Lifting of a function of measure). (See section 5.2. of [37]) A lifting of the function f , denoted by \tilde{f} is defined on the Hilbert space $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ by setting $\tilde{f}(X) = f(\mathcal{L}(X))$.

Example (scalar interactions): Consider $f(\mu) = \int_{\mathbb{R}^d} p(x) d\mu(x) = \langle p, \mu \rangle$. Its lifting function is given by $\tilde{f}(X) = f(\mathcal{L}(X)) = \mathbb{E} [p(X)]$, for any $X \sim \mu$.

Next, we define the notion of L-differentiability on $\mathcal{P}^2(\mathbb{R}^d)$, with the L-derivative denoted by $\partial_\mu f(\mu)(X) \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$,

Definition 2.2.2. [L-differentiability](see Definition 5.2, [37]) The function f above is L-differentiable at $\mu \in \mathcal{P}^2(\mathbb{R}^d)$ if there exists a random variable $X \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ with $X \sim \mu$ such that the

lifted function \tilde{f} is Fréchet differentiable at X with the Fréchet derivative denoted by $D\tilde{f}$, i.e.

$$\lim_{\delta \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d), \|\delta\|_{L^2} \rightarrow 0} \frac{|\tilde{f}(X + \delta) - \tilde{f}(X) - \mathbb{E} [D\tilde{f}(X) \cdot \delta]|}{\|\delta\|_{L^2}} = 0.$$

However, it is important to ensure that this definition is intrinsic, which means that it does not depend on the choice of the random variable $X \sim \mu$. For this we refer to Proposition 5.24 [37]. Next, we define a notion of convexity for L-differentiable functions of measure.

Definition 2.2.3. [L-convexity](Section 5.5 [37]) Suppose the function f is continuously L-differentiable. It is said to be L-convex if for every $\mu, \mu' \in \mathcal{P}^2(\mathbb{R}^d)$, we have

$$f(\mu') - f(\mu) - \mathbb{E} [\partial_\mu f(\mu)(X) \cdot (X' - X)] \geq 0,$$

where $X' \sim \mu'$ and $X \sim \mu$.

Note that this definition of L-convexity is intrinsic because the notion of the L-derivative, on which it is based, is intrinsic.

2.2.2 The Model

In this section we will define a McKean-Vlasov stochastic control problem in the infinite time horizon with the running cost function depending on the law of the state process. We use the approach of [37] and [1] which provide us with similar results in the finite time horizon setting. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a probability space with an augmented and right-continuous filtration $(\mathcal{F}_t)_{t \geq 0}$ and consider the following SDE as the state dynamics of our control problem,

$$dX_t^u = b(X_t^u, u_t) dt + \sum_{j=1}^m \sigma^j(X_t^u, u_t) dW_t^j, \quad X_0 = x_0, \quad (2.2.1)$$

where $X_t \in \mathbb{R}^d$ and $W_t \in \mathbb{R}^m$ is a \mathcal{F} -adapted standard Brownian motion and $(u_t)_{t \geq 0}$ is a \mathcal{F} -progressively measurable process called control taking values in $U = \mathbb{R}^n$. Consider the following cost functional,

$$J(u) := \mathbb{E} \left[\int_0^\infty e^{-\rho t} f(X_t^u, u_t, \mu_{X_t^u}) dt \right],$$

for a given $\rho > 0$ discounting factor and the control u belongs to the space of admissible controls defined as follows:

Definition 2.2.4. We define \mathcal{A} as the set of *admissible controls* $(u_t)_{t \geq 0}$ to have the following properties:

1. $(u_t)_{t \geq 0}$ is \mathcal{F} -progressively measurable,
2. $\mathbb{E} [\int_0^\infty e^{-\rho t} |u_t|^2 dt] < \infty$.

We consider a cost minimization problem in the infinite horizon. An optimal control when it exists is usually denoted by u^* ,

$$u^* = \arg \min_{u \in \mathcal{A}} \mathbb{E} \left[\int_0^\infty e^{-\rho t} f(X_t^u, u_t, \mu_{X_t^u}) dt \right].$$

Definition 2.2.5. For given $k, l \in \mathbb{N}$, $S^2(\mathbb{R}^k)$ (or $S^2(\mathbb{R}^{k \times l})$) is defined as the set of continuous progressively measurable processes $(L_t, t \geq 0)$ with values in \mathbb{R}^k (respectively $\mathbb{R}^{k \times l}$) s.t. $\forall T > 0$,

$$\mathbb{E} \left[\sup_{t \in [0, T]} |L_t|^2 \right]^{\frac{1}{2}} \left(\text{respectively } \mathbb{E} \left[\sup_{t \in [0, T]} \|L_t\|^2 \right]^{\frac{1}{2}} \right) < \infty$$

and define $\Lambda^2(\mathbb{R}^k)$ (or $\Lambda^2(\mathbb{R}^{k \times l})$) as the set of progressively measurable processes $(L_t, t \geq 0)$ with values in \mathbb{R}^k (respectively $\mathbb{R}^{k \times l}$) s.t. $\forall T > 0$,

$$\mathbb{E} \left[\int_0^T |L_t|^2 dt \right]^{\frac{1}{2}} \left(\text{respectively } \mathbb{E} \left[\int_0^T \|L_t\|^2 dt \right]^{\frac{1}{2}} \right) < \infty.$$

Consider the following assumptions on the system dynamics and the running cost.

Assumptions

(H1). The functions $b(x, u)$ and $\sigma(x, u)$ are continuously differentiable with respect to (x, u) . The derivatives $\partial_x b$, $\partial_x \sigma$, $\partial_u b$ and $\partial_u \sigma$ are uniformly bounded, i.e. there exists $K_b > 0$ and $K_\sigma > 0$ s.t. $\forall (x, u) \in \mathbb{R}^d \times \mathbb{R}^n$,

$$\max(|\partial_x b(x, u)|, |\partial_u b(x, u)|) \leq K_b \text{ and } \max \left(\sqrt{\sum_{j=1}^m \|\partial_x \sigma^j(x, u)\|_\infty^2}, \sqrt{\sum_{j=1}^m \|\partial_u \sigma^j(x, u)\|_\infty^2} \right) \leq K_\sigma.$$

(H2). The running cost function $f(x, u, \mu)$ is continuously differentiable with respect to (x, u) and continuously L-differentiable in μ and the derivatives of f have linear growth, i.e. there exists a constant $L > 0$ s.t.

$$\max\{|\partial_x f(x, u, \mu)|, |\partial_u f(x, u, \mu)|\} \leq L(1 + |x| + |u| + \|\mu\|_{\mathcal{M}_2}).$$

where $\|\nu\|_{\mathcal{M}_2} = \sqrt{\mathbb{E}[|Y|^2]}$, $Y \sim \nu$ (note that $\|\mu\|_{\mathcal{M}_2}$ can be seen as $W_2(\mu, \delta_0)$) and the $L^2(\mathbb{R}^d, \mu)$ -norm of the mapping $x' \mapsto \partial_\mu f(x, u, \mu)(x')$ is bounded by $L(1 + |x| + |u| + \|\mu\|_{\mathcal{M}_2})$, i.e.

$$(\mathbb{E} [|\partial_\mu f(x, u, \mu)(X)|^2])^{\frac{1}{2}} \leq L(1 + |x| + |u| + \|\mu\|_{\mathcal{M}_2}), \quad X \sim \mu$$

(H3). The discounting factor ρ is s.t.

$$2K_b + K_\sigma^2 - \rho < 0.$$

(H4). The derivatives of f are Lipschitz continuous in all three arguments, i.e. $\exists L_f$ such that $\forall x, x_1, x_2 \in \mathbb{R}^d$; $u, u_1, u_2 \in \mathbb{R}^n$; $\mu_1, \mu_2 \in \mathcal{P}^2(\mathbb{R}^d)$ and $\mu \in \mathcal{P}^2(\mathbb{R}^d)$, we have

$$\max_{i \in \{x, u\}} |\partial_i f(x_1, u_1, \mu_1) - \partial_i f(x_2, u_2, \mu_2)| \leq L_f (|x_1 - x_2| + |u_1 - u_2| + W_2(\mu_1, \mu_2));$$

$$\mathbb{E} [|\partial_\mu f(x_1, u_1, \mu_1)(X^1) - \partial_\mu f(x_2, u_2, \mu_2)(X^2)|^2]^{\frac{1}{2}} \leq L_f (|x_1 - x_2| + |u_1 - u_2| + W_2(\mu_1, \mu_2))$$

for any $X_1 \sim \mu_1$, $X_2 \sim \mu_2$.

Note. When using the partial L-derivative of a function being evaluated at a random variable, we would use $\tilde{\mathbb{E}}$ to denote expectation over independent copies of the corresponding random variables. For instance, the above inequality would become

$$\begin{aligned} \tilde{\mathbb{E}} [|\partial_\mu f(X^1, u, \mu_1)(\tilde{X}_t^1) - \partial_\mu f(X^2, u, \mu_2)(\tilde{X}_t^2)|^2]^{\frac{1}{2}} &\leq L_f (|X^1 - X^2| + W_2(\mu_1, \mu_2)) \\ &\leq L_f (|X^1 - X^2| + \mathbb{E} [|X^1 - X^2|^2]^{\frac{1}{2}}), \end{aligned}$$

where $X^1 \sim \mu^1$, $X^2 \sim \mu^2$, \tilde{X}^1 and \tilde{X}^2 denote their independent copies. This notion will be later used in the proofs.

2.2.3 A useful estimate

From now onward, we will use C to denote a generic constant whose value is not important for the argument and might change from line to line. We state the following useful estimate on the forward process:

Proposition 2.2.1. *Under H1 and H3, there exists $C > 0$ independent of u and x_0 s.t. $\forall u \in \mathcal{A}$,*

$$\sup_{T \geq 0} \mathbb{E} [e^{-\rho T} |X_T|^2] + \mathbb{E} \left[\int_0^\infty e^{-\rho t} |X_t|^2 dt \right] \leq C \left(1 + |x_0|^2 + \mathbb{E} \left[\int_0^\infty e^{-\rho t} |u_t|^2 dt \right] \right). \quad (2.2.2)$$

Proof.

Use Itô's formula for $e^{-\rho t} |X_t^u|^2$ and Lipschitz continuity of b and σ (implying linear growth of b and σ) from H1 to write

$$\begin{aligned} e^{-\rho T} |X_T|^2 - |x_0|^2 &= \int_0^T e^{-\rho t} \left[-\rho |X_t^u|^2 dt + 2X_t^u \cdot dX_t^u + \sum_{i=1}^N d\langle X^{u,i}, X^{u,i} \rangle_t \right], \\ \mathbb{E} [e^{-\rho T} |X_T|^2] - |x_0|^2 &= \mathbb{E} \left[\int_0^T e^{-\rho t} \left(-\rho |X_t^u|^2 + 2X_t^u \cdot b(X_t^u, u_t) + \sum_{j=1}^m \|\sigma^j(X_t^u, u_t)\|^2 \right) dt \right] \\ &\leq \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(-\rho |X_t^u|^2 + 2|X_t^u|(|b(0,0)| + K_b |X_t^u| + K_b |u_t|) \right. \right. \\ &\quad \left. \left. + \sum_{j=1}^m (\|\sigma^j(0,0)\| + \|\partial_x \sigma^j\|_\infty |X_t^u| + \|\partial_u \sigma^j\|_\infty |u_t|)^2 \right) dt \right] \\ &\leq \mathbb{E} \left[\int_0^T e^{-\rho t} \left((-\rho + 2K_b + K_\sigma^2 + \mu) |X_t|^2 + C_1 + C_2 |u_t|^2 \right) dt \right], \end{aligned}$$

where we use Young's inequality $2|X_t||u_t| \leq \mu |X_t|^2 + \frac{1}{\mu} |u_t|^2$ for arbitrary $\mu > 0$. Owing to assumption H3, one can choose $\mu > 0$ s.t. $-\kappa := (-\rho + 2K_b + K_\sigma^2 + \mu) < 0$. The above rewrites as

$$\mathbb{E} [e^{-\rho T} |X_T|^2] + \kappa \mathbb{E} \left[\int_0^T e^{-\rho t} |X_t|^2 dt \right] \leq |x_0|^2 + \mathbb{E} \left[\int_0^T e^{-\rho t} (C_1 + C_2 |u_t|^2) dt \right].$$

Apply the monotone convergence theorem to complete the proof of the proposition,

$$\sup_{T \geq 0} \mathbb{E} [e^{-\rho T} |X_T|^2] + \mathbb{E} \left[\int_0^\infty e^{-\rho t} |X_t|^2 dt \right] \leq C \left(1 + |x_0|^2 + \mathbb{E} \left[\int_0^\infty e^{-\rho t} |u_t|^2 dt \right] \right).$$

□

The following L^2 estimate for the difference of two state processes can be proved in a similar way:

Corollary 2.2.1.1. *Under the above assumptions, if X and X' are the state processes corresponding to $q, q' \in \mathcal{A}$ then the following L^2 estimate holds:*

$$\sup_{T \geq 0} \mathbb{E} [e^{-\rho T} |X_T' - X_T|^2] + \mathbb{E} \left[\int_0^\infty e^{-\rho t} |X_t' - X_t|^2 dt \right] \leq C \mathbb{E} \left[\int_0^\infty e^{-\rho t} (|q_t' - q_t|^2) dt \right] < \infty.$$

2.2.4 Adjoint processes

From now on, to shorten the notation wherever it causes no confusion, we will use generic placeholder variables like θ and η for arguments of some functions. For example,

$$b(\theta_t) = b(X_t^u, u_t), \quad \sigma(\theta_t) = \sigma(X_t^u, u_t) \quad \text{and} \quad f(\theta_t) = f(X_t^u, u_t, \mu_{X_t^u}),$$

with an understanding that $b(\theta_t)$ and $\sigma(\theta_t)$ do not depend on $\mu_{X_t^u}$. Next, we define the Hamiltonian of the problem as

$$H(\eta) := H(x, u, \mu, y, z) = y b(x, u) + z^\top : \sigma(x, u) + f(x, u, \mu) - \rho y x \quad (2.2.3)$$

where $(y, z) \in \mathbb{R}^{1 \times d} \times \mathbb{R}^{m \times d}$, A^\top denotes the transpose and $A : B$ denotes $\text{Tr}(A^\top B)$. Then, we define the adjoint processes of X to be the pair of processes $(Y, Z) \in S^2(\mathbb{R}^{1 \times d}) \times \Lambda^2(\mathbb{R}^{m \times d})$ which satisfy for every $T > 0$ the following infinite horizon BSDE $\text{dt} \otimes \text{d}\mathbb{P}$ -a.e.,

$$\begin{aligned} & Y_t - Y_T \\ &= \int_t^T (\partial_x H(X_t^u, u_t, \mu_{X_t^u}, Y_t, Z_t) + \tilde{\mathbb{E}}[\partial_\mu H(\tilde{X}_t^u, \tilde{u}_t, \mu_{X_t^u}, \tilde{Y}_t, \tilde{Z}_t)(X_t^u)]) dt - \sum_{j=1}^m \int_t^T Z_t^j dW_t^j \\ &= \int_t^T \left(Y_t \partial_x b(\theta_t) + \sum_{j=1}^m Z_t^j \partial_x \sigma^j(\theta_t) + \partial_x f(\theta_t) + \tilde{\mathbb{E}}[\partial_\mu f(\tilde{\theta}_t)(X_t^u)] - \rho Y_t \right) dt - \sum_{j=1}^m \int_t^T Z_t^j dW_t^j, \end{aligned} \quad (2.2.4)$$

where Z_t^j is the j -th row of Z_t and the first three terms of the generator correspond to the space derivative of the Hamiltonian and the fourth term corresponds to the L-derivative with respect to the law of the state. We have the following existence and uniqueness result for a solution of (2.2.4):

Proposition 2.2.2. *Let $u \in \mathcal{A}$ and the corresponding state process be $(X_t^u)_{t \geq 0}$ and suppose that assumptions H1, H2 and H3 hold. Then, the above BSDE (2.2.4) has a unique solution (Y, Z) in $S^2(\mathbb{R}^{1 \times d}) \times \Gamma^2(\mathbb{R}^{m \times d})$ with $\lim_{t \rightarrow \infty} \mathbb{E}[e^{-\rho t} |Y_t|^2] = 0$.*

Proof. Let us consider $u \in \mathcal{A}$ and the corresponding state process $(X_t^u)_{t \geq 0}$. Some direct computations show that the generator of the BSDE is $(K_b - \rho)$ -monotone with respect to y and K_σ -Lipschitz with respect to z . Then, we just have to apply Theorem 5.57 in Pardoux and Răşcanu [100] which gives a result in a more general setting. In the notation of their assumptions **BSDE** – **H $_\infty$** , we take $\eta = 0$, $p = 2$, $n_p = 1$ and we get $\xi_t = \zeta_t = 0$ for all $t \geq 0$ a.s.. We also take $Q_t = t$, $\alpha_t = 1$, $\mu_t = K_b - \rho$, $l(t) = K_\sigma^2$ and Φ as the generator of BSDE (2.2.4). Using this, the process V_t is then given by

$$V_t = (K_b - \rho)t + \frac{aK_\sigma^2 t}{2} = \frac{1}{2}(2K_b + K_\sigma^2 - \rho)t - \frac{\rho t}{2} + \frac{(a-1)K_\sigma^2 t}{2}, \quad t \geq 0.$$

Since it is true for arbitrary $a > 1$ and assumption H3 implies that $2K_b + K_\sigma^2 - \rho < 0$, we can choose $a > 1$ small enough such that there exists an $\varepsilon > 0$ small enough too satisfying

$$V_t < \frac{-\rho t}{2} - \varepsilon t, \quad t \geq 0.$$

Then, the only non-trivial assumption left to satisfy is

$$\mathbb{E} \left[\int_0^\infty e^{V_t} |\Phi(t, \xi_t, \zeta_t)| dQ_t \right]^p < \infty.$$

But we have, using Hölder's inequality, the linear growth condition from assumption H2, Fubini's theorem, Proposition 2.2.1 and admissibility of control,

$$\mathbb{E} \left[\int_0^\infty e^{V_t} |\Phi(t, 0, 0)| dt \right]^2 \leq \mathbb{E} \left[\int_0^\infty e^{-(\frac{\rho}{2} + \varepsilon)t} |\Phi(t, 0, 0)| dt \right]^2$$

$$\begin{aligned}
 &\leq \mathbb{E} \left[\int_0^\infty e^{-\rho t} |\Phi(t, 0, 0)|^2 dt \right] \mathbb{E} \left[\int_0^\infty e^{-2\epsilon t} dt \right] \\
 &\leq C \mathbb{E} \left[\int_0^\infty e^{-\rho t} (1 + |X_t^u|^2 + |u_t|^2) dt \right] < \infty.
 \end{aligned}$$

□

2.3 Necessary conditions for optimality

Let us now present the necessary conditions that an optimal control must satisfy,

Theorem 2.3.1 (Necessary conditions). *Suppose that the assumptions H1, H2, H3 and H4 are satisfied and $u \in \mathcal{A}$ is optimal. Let X be the corresponding state variable and let (Y, Z) be the corresponding adjoint processes as defined above in (2.2.4), then we must have*

$$\partial_u H(X_t^u, u_t, \mu_{X_t^u}, Y_t, Z_t) = 0, \quad dt \otimes d\mathbb{P} \text{ a.e.}$$

Proof. The proof is split into several steps:

Step 1: $J(u)$ is Gateau differentiable $\forall u \in \mathcal{A}$.

Consider an arbitrary admissible control $v \in \mathcal{A}$. Then $\forall \varepsilon \in [0, 1]$, we can define an admissible control $u^\varepsilon = u + \varepsilon(v - u)$ (thanks to the convexity of $U = \mathbb{R}^n$) and let X^ε denote the solution of the corresponding state equation (2.2.1). Then, consider a formal differentiation of X^ε with respect to ε and denote it by V , defined as the solution of the following SDE,

$$\begin{aligned}
 V_t &= \int_0^t \left(\lim_{\varepsilon \rightarrow 0} \frac{b(\theta_s^\varepsilon) - b(\theta_s)}{\varepsilon} ds + \sum_{j=1}^m \lim_{\varepsilon \rightarrow 0} \frac{\sigma^j(\theta_s^\varepsilon) - \sigma^j(\theta_s)}{\varepsilon} dW_s^j \right) \\
 &= \int_0^t \left(\partial_x b(\theta_s) V_s + \partial_u b(\theta_s)(v_s - u_s) \right) ds \\
 &\quad + \sum_{j=1}^m \int_0^t \left(\partial_x \sigma^j(\theta_s) V_s + \partial_u \sigma^j(\theta_s)(v_s - u_s) \right) dW_s^j.
 \end{aligned} \tag{2.3.1}$$

Now, we want to consider the true Gateaux derivative of X_t^ε w.r.t. ε at $\varepsilon = 0$ and then we would like to prove that it is indeed V_t defined above. For this, define the ε -difference term, $\frac{X_t^\varepsilon - X_t}{\varepsilon}$ as V_t^ε and consider the following

$$V_t^\varepsilon = \frac{X_t^\varepsilon - X_t}{\varepsilon} = \int_0^t \left(\frac{b(\theta_s^\varepsilon) - b(\theta_s)}{\varepsilon} ds + \sum_{j=1}^m \frac{\sigma^j(\theta_s^\varepsilon) - \sigma^j(\theta_s)}{\varepsilon} dW_s^j \right).$$

Using Taylor's formula equality, we get

$$\begin{aligned}
 V_t^\varepsilon &= \int_0^t \left[\left(\int_0^1 \partial_x b(X_s + \lambda(X_s^\varepsilon - X_s), u_s^\varepsilon) \frac{(X_s^\varepsilon - X_s)}{\varepsilon} d\lambda + \int_0^1 \partial_u b(X_s, u_s + \lambda(u_s^\varepsilon - u_s)) \frac{(u_s^\varepsilon - u_s)}{\varepsilon} d\lambda \right) ds \right. \\
 &\quad \left. + \sum_{j=1}^m \left(\int_0^1 \partial_x \sigma^j(X_s + \lambda(X_s^\varepsilon - X_s), u_s^\varepsilon) \frac{(X_s^\varepsilon - X_s)}{\varepsilon} d\lambda + \int_0^1 \partial_u \sigma^j(X_s, u_s + \lambda(u_s^\varepsilon - u_s)) \frac{(u_s^\varepsilon - u_s)}{\varepsilon} d\lambda \right) dW_s^j \right]
 \end{aligned}$$

$$= \int_0^s \left(\left(\dot{b}_x^\varepsilon(s) V_s^\varepsilon + \dot{b}_u^\varepsilon(s) (v_s - u_s) \right) ds + \sum_{j=1}^m \left(\dot{\sigma}_x^{\varepsilon,j}(s) V_s^\varepsilon + \dot{\sigma}_u^{\varepsilon,j}(s) (v_s - u_s) \right) dW_s^j \right)$$

where, in order to simplify, we have used the following notations

$$\begin{aligned} \dot{b}_x^\varepsilon(t) &:= \int_0^1 \partial_x b(X_t + \lambda(X_t^\varepsilon - X_t), u_t^\varepsilon) d\lambda, & \dot{\sigma}_x^{\varepsilon,j}(t) &:= \int_0^1 \partial_x \sigma^j(X_t + \lambda(X_t^\varepsilon - X_t), u_t^\varepsilon) d\lambda, \\ \dot{b}_u^\varepsilon(t) &:= \int_0^1 \partial_u b(X_t, u_t + \lambda(u_t^\varepsilon - u_t)) d\lambda, & \dot{\sigma}_u^{\varepsilon,j}(t) &:= \int_0^1 \partial_u \sigma^j(X_t, u_t + \lambda(u_t^\varepsilon - u_t)) d\lambda. \end{aligned}$$

Denote, without ε , the usual derivative at $\varepsilon = 0$, i.e. $\dot{b}_x(t) = \dot{b}_x^\varepsilon(t)|_{\varepsilon=0} = \partial_x b(\theta_t)$ and so on. It is useful to consider the SDE of residual error process $V_t^\varepsilon - V_t$, which satisfies, owing to (2.3.1),

$$\begin{aligned} V_t^\varepsilon - V_t &= \int_0^s \left(\left(\dot{b}_x^\varepsilon(s) (V_s^\varepsilon - V_s) + (\dot{b}_x^\varepsilon(s) - \dot{b}_x(s)) V_s + (\dot{b}_u^\varepsilon(s) - \dot{b}_u(s)) (v_s - u_s) \right) ds \right. \\ &\quad \left. + \sum_{j=1}^m \left(\dot{\sigma}_x^{\varepsilon,j}(s) (V_s^\varepsilon - V_s) + (\dot{\sigma}_x^{\varepsilon,j}(s) - \dot{\sigma}_x^j(s)) V_s + (\dot{\sigma}_u^{\varepsilon,j}(s) - \dot{\sigma}_u^j(s)) (v_s - u_s) \right) dW_s^j \right). \end{aligned} \quad (2.3.2)$$

Let us assume for a while the following Lemma (whose proof is handled at the end).

Lemma 2.3.2. *Under the assumptions of Theorem 2.3.1, we have*

$$\mathbb{E} \left[\int_0^\infty e^{-\rho t} |V_t|^2 dt \right] + \sup_{\varepsilon \leq 1} \mathbb{E} \left[\int_0^\infty e^{-\rho t} |V_t^\varepsilon|^2 dt \right] < +\infty \text{ and} \quad (2.3.3)$$

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[\int_0^\infty e^{-\rho t} |V_t^\varepsilon - V_t|^2 dt \right] = 0. \quad (2.3.4)$$

Next, let us consider the Gateaux derivative of J at u in the direction $(v - u)$. Let us define the following,

$$\partial J^v(u) := \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(\dot{f}_x(t) V_t + \dot{f}_u(t) (v_t - u_t) + \tilde{\mathbb{E}}[\partial_\mu f(\theta_t)(\tilde{X}_t^u, \tilde{V}_t)] \right) dt \right] \quad (2.3.5)$$

where we set use the short-hand notation $\dot{f}_\ell(t) := \partial_\ell f(\theta_t)$ for $\ell \in \{x, u, \mu\}$. The integrability of ∂J^v will follow from the estimates proved below. We claim that

$$\lim_{\varepsilon \rightarrow 0} \frac{J(u^\varepsilon) - J(u)}{\varepsilon} = \partial J^v(u).$$

Using Taylor formula in an equality form, write

$$\begin{aligned} \frac{J(u^\varepsilon) - J(u)}{\varepsilon} - \partial J^v(u) &= \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(\dot{f}_x^\varepsilon(t) (V_t^\varepsilon - V_t) + (\dot{f}_x^\varepsilon(t) - \dot{f}_x(t)) V_t + \right. \right. \\ &\quad \left. \left. (\dot{f}_u^\varepsilon(t) - \dot{f}_u(t)) (v_t - u_t) + \tilde{\mathbb{E}} \left[(\dot{f}_\mu^\varepsilon(t) - \dot{f}_\mu(t)) \cdot (\tilde{V}_t^\varepsilon - \tilde{V}_t) \right] \right) dt \right] \end{aligned} \quad (2.3.6)$$

where we set

$$\begin{aligned} \dot{f}_x^\varepsilon(t) &:= \int_0^1 \partial_x f(X_t^{\varepsilon,\lambda}, u_t^\varepsilon, \mu_{X_t^\varepsilon}) d\lambda, \\ \dot{f}_u^\varepsilon(t) &:= \int_0^1 \partial_u f(X_t, u_t + \lambda\varepsilon(v_t - u_t), \mu_{X_t^\varepsilon}) d\lambda, \end{aligned}$$

$$\dot{f}_\mu^\varepsilon(t) := \int_0^1 \partial_\mu f(X_t, u_t, \mu_{X_t^\varepsilon, \lambda})(\tilde{X}_t^{\varepsilon, \lambda}) d\lambda,$$

where $X_t^{\varepsilon, \lambda} = X_t + \lambda(X_t^\varepsilon - X_t)$.

i) Let us consider the first term in (2.3.6):

$$\left| \mathbb{E} \left[\int_0^\infty e^{-\rho t} \dot{f}_x^\varepsilon(t) (V_t^\varepsilon - V_t) dt \right] \right| \leq \mathbb{E} \left[\int_0^\infty e^{-\rho t} |\dot{f}_x^\varepsilon(t)|^2 dt \right]^{\frac{1}{2}} \mathbb{E} \left[\int_0^\infty e^{-\rho t} |V_t^\varepsilon - V_t|^2 dt \right]^{\frac{1}{2}} \quad (2.3.7)$$

Since the derivatives of f have linear growth under H2,

$$|\dot{f}_x^\varepsilon(t)|^2 \leq C(1 + |X_t^\varepsilon|^2 + |X_t|^2 + |v_t|^2 + |u_t|^2 + |\mathbb{E}[|X_t^\varepsilon|^2] + \mathbb{E}[|X_t|^2]).$$

Hence, using the uniform a priori estimates of Proposition 2.2.1, we justify that the first factor on the right hand side of (2.3.7) is bounded uniformly in ε (note that the constant C in Proposition 2.2.1 is independent of ε). Then using (2.3.4), we conclude that

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[\int_0^\infty e^{-\rho t} \dot{f}_x^\varepsilon(t) (V_t^\varepsilon - V_t) dt \right] = 0$$

ii) Next, let us consider the second term in (2.3.6). Using the Lipschitz continuity of the derivatives of f under H4, we have

$$\begin{aligned} |\dot{f}_x^\varepsilon(t) - \dot{f}_x(t)| &\leq \int_0^1 \left| \partial_x f(X_t + \lambda(X_t^\varepsilon - X_t), u_t^\varepsilon, \mu_{X_t^\varepsilon}) - \partial_x f(X_t, u_t, \mu_{X_t}) \right| d\lambda \\ &\leq L_f \left(|X_t^\varepsilon - X_t| + \varepsilon |v_t - u_t| + \mathbb{E}[|X_t^\varepsilon - X_t|^2]^{\frac{1}{2}} \right). \end{aligned}$$

Hence, owing to Corollary 2.2.1.1, we easily get that $\mathbb{E} \left[\int_0^\infty e^{-\rho t} |\dot{f}_x^\varepsilon(t) - \dot{f}_x(t)|^2 dt \right] \rightarrow 0$ as $\varepsilon \rightarrow 0$ and we conclude,

$$\left| \mathbb{E} \left[\int_0^\infty e^{-\rho t} (\dot{f}_x^\varepsilon(t) - \dot{f}_x(t)) V_t dt \right] \right| \leq \mathbb{E} \left[\int_0^\infty e^{-\rho t} |\dot{f}_x^\varepsilon(t) - \dot{f}_x(t)|^2 dt \right]^{\frac{1}{2}} \mathbb{E} \left[\int_0^\infty e^{-\rho t} |V_t|^2 dt \right]^{\frac{1}{2}} \rightarrow 0$$

as $\varepsilon \rightarrow 0$.

iii) The third term can be handled in the same way as the first two terms.

iv) For the fourth term, we have the following,

$$\begin{aligned} &\left| \mathbb{E} \left[\int_0^\infty e^{-\rho t} \mathbb{E} \left[(\dot{f}_\mu^\varepsilon(t) - \dot{f}_\mu(t)) \cdot (\tilde{V}_t^\varepsilon - \tilde{V}_t) \right] dt \right] \right| \\ &\leq \left(\mathbb{E} \left[\int_0^\infty e^{-\rho t} |\dot{f}_\mu^\varepsilon(t) - \dot{f}_\mu(t)|^2 dt \right] \right)^{\frac{1}{2}} \left(\mathbb{E} \left[\int_0^\infty e^{-\rho t} |(\tilde{V}_t^\varepsilon - \tilde{V}_t)|^2 dt \right] \right)^{\frac{1}{2}}, \end{aligned}$$

which is similar to the first two terms, except for the Lipschitz requirement for the partial L-derivative of f , which follows from H4,

$$\tilde{\mathbb{E}} \left[|\partial_\mu f(\theta_t)(\tilde{X}_t^{\varepsilon, \lambda}) - \partial_\mu f(\theta_t)(\tilde{X}_t)|^2 \right] \leq L_f^2 \mathbb{E} \left[|X_t^{\varepsilon, \lambda} - X_t|^2 \right].$$

Using Corollary 2.2.1.1 and (2.3.4), the limit of the fourth term is zero as well.

This proves that the functional J is Gateaux differentiable and its Gateaux derivative is ∂J as defined earlier.

Step 2: Transformation of the Gateaux derivative using BSDE. Our aim is to express the formula (2.3.5) in terms of the infinite-horizon BSDE (2.2.4). So, let us use the Itô product rule for the process $e^{-\rho t}Y_tV_t$:

$$\begin{aligned}
 \mathbb{E} \left[e^{-\rho T} Y_T V_T \right] &= \mathbb{E} \left[\int_0^T e^{-\rho t} \left(-\rho Y_t V_t dt + Y_t dV_t + (dY_t) V_t + \sum_{i=1}^d d\langle Y^i, V^i \rangle_t \right) \right] \\
 &= \mathbb{E} \left[\int_0^T e^{-\rho t} \left(-\rho Y_t V_t dt + (Y_t \dot{b}_x(t) V_t + Y_t \dot{b}_u(t) (v_t - u_t)) dt + \left(-Y_t \dot{b}_x(t) V_t \right. \right. \right. \\
 &\quad \left. \left. - \sum_{j=1}^m Z_t^j \dot{\sigma}_x^j(t) V_t - \dot{f}_x(t) V_t - \tilde{\mathbb{E}} \left[\partial_\mu f(\tilde{\theta}_t)(X_t) \right] V_t + \rho Y_t V_t \right) dt + \sum_{j=1}^m \left(Z_t^j \dot{\sigma}_x^j(t) V_t + Z_t^j \dot{\sigma}_u^j(t) (v_t - u_t) \right) dt \right] \\
 &= \mathbb{E} \left[\int_0^T e^{-\rho t} \left(Y_t \dot{b}_u(t) (v_t - u_t) + \sum_{j=1}^m Z_t^j \dot{\sigma}_u^j(t) (v_t - u_t) - \dot{f}_x(t) V_t - \tilde{\mathbb{E}} \left[\partial_\mu f(\tilde{\theta}_t)(X_t) \right] V_t \right) dt \right] \\
 &= \mathbb{E} \left[\int_0^T e^{-\rho t} \left(Y_t \dot{b}_u(t) (v_t - u_t) + \sum_{j=1}^m Z_t^j \dot{\sigma}_u^j(t) (v_t - u_t) - \dot{f}_x(t) V_t - \tilde{\mathbb{E}} \left[\partial_\mu f(\theta_t)(\tilde{X}_t) \tilde{V}_t \right] \right) dt \right],
 \end{aligned}$$

where the expectation of the martingale term is zero and Fubini's theorem is used in the last equality. Next, take the limit as $T \rightarrow +\infty$ in the above identity: on the one hand,

$$\left| \mathbb{E} \left[e^{-\rho T} Y_T V_T \right] \right| \leq \mathbb{E} \left[e^{-\rho T} |Y_T|^2 \right]^{\frac{1}{2}} \mathbb{E} \left[e^{-\rho T} |V_T|^2 \right]^{\frac{1}{2}} \rightarrow 0$$

using the terminal condition (2.2.4) and the a priori estimates (2.3.3). On the other hand, the above integral converges using the L_2 estimates from (2.2.2)-(2.3.3) together with H2. Hence, we have established, for any $u, v \in \mathcal{A}$,

$$\mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(Y_t \dot{b}_u(t) (v_t - u_t) + \sum_{j=1}^m Z_t^j \dot{\sigma}_u^j(t) (v_t - u_t) - \dot{f}_x(t) V_t - \tilde{\mathbb{E}} \left[\partial_\mu f(\theta_t)(\tilde{X}_t) \tilde{V}_t \right] \right) dt \right] = 0,$$

and summing this with (2.3.5), we get

$$\partial J^v(u) = \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(Y_t \dot{b}_u(t) + \sum_{j=1}^m Z_t^j \dot{\sigma}_u^j(t) + \dot{f}_u(t) \right) (v_t - u_t) dt \right].$$

Step 3: necessary condition. Let $u \in \mathcal{A}$ be an optimal control for the above problem.

Then, we must have $\partial J^v(u) = 0$ for all $v \in \mathcal{A}$, that is to say,

$$\partial J^v(u) = \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(Y_t \dot{b}_u(t) + \sum_{j=1}^m Z_t^j \dot{\sigma}_u^j(t) + \dot{f}_u(t) \right) (v_t - u_t) dt \right] = 0, \quad \forall v \in \mathcal{A}.$$

Hence, we obtain the following,

$$\partial J^v(u) = \mathbb{E} \left[\int_0^\infty e^{-\rho t} \partial_u H(\eta_t) (v_t - u_t) dt \right] = 0, \quad \forall v \in \mathcal{A}.$$

Next, for the final step of the theorem, define the control v_t to take value $\nu \in U$ (deterministic) in an arbitrarily chosen progressively measurable set $B \subset [0, \infty) \times \Omega$ with finite measure and equal to u_t elsewhere. Clearly, admissibility of $(u_t)_{t \geq 0}$ implies the admissibility of v_t defined above. Hence, the above condition gives,

$$\mathbb{E} \left[\int_0^\infty \mathbf{1}_B \left(e^{-\rho t} \left(\partial_u H(\eta_t) \right) (\nu - u_t) \right) dt \right] = 0$$

Since this holds for arbitrary B and $\nu \in U$, we must have,

$$\partial_u H(\eta_t) = 0, \quad dt \otimes d\mathbb{P} \text{ a.e.}$$

To prove this, assume contrarily that there exists a $dt \otimes d\mathbb{P}$ -measurable set D with positive measure such that $\partial_u H(\eta_t) \cdot (\nu - u_t) \neq 0$ on D . Then,

$$\mathbb{E} \left[\int_0^\infty \mathbf{1}_D \left(e^{-\rho t} \left(\partial_u H(\eta_t) \right) (\nu - u_t) \right) dt \right] \neq 0$$

which is a contradiction, hence proving the theorem. \square

Proof. [Proof of Lemma 2.3.2] Recall the SDE satisfied by $V_t^\varepsilon - V_t$ from (2.3.2). Let us consider the Itô's formula for $e^{-\rho t} |V_t^\varepsilon - V_t|^2$, for a given $T \geq 0$,

$$\begin{aligned} & e^{-\rho T} |V_T^\varepsilon - V_T|^2 \\ &= \int_0^T e^{-\rho t} \left[\left(-\rho |V_t^\varepsilon - V_t|^2 + 2(V_t^\varepsilon - V_t)^\top \dot{b}_x^\varepsilon(t) (V_t^\varepsilon - V_t) + 2(V_t^\varepsilon - V_t)^\top (\dot{b}_x^\varepsilon(t) - \dot{b}_x(t)) V_t \right. \right. \\ & \quad \left. \left. + 2(V_t^\varepsilon - V_t)^\top (\dot{b}_u^\varepsilon(t) - \dot{b}_u(t)) (v_t - u_t) \right) dt + \sum_{j=1}^m \left| \dot{\sigma}_x^{\varepsilon,j}(t) (V_t^\varepsilon - V_t) + (\dot{\sigma}_x^{\varepsilon,j}(t) - \dot{\sigma}_x^j(t)) V_t \right. \right. \\ & \quad \left. \left. + (\dot{\sigma}_u^{\varepsilon,j}(t) - \dot{\sigma}_u^j(t)) (v_t - u_t) \right|^2 dt \right] + \mathcal{M}_T, \end{aligned}$$

where \mathcal{M}_T is a martingale. Taking expectation on both sides, using Young's inequality for cross-terms and expanding the square after the summation, we get

$$\begin{aligned} & \mathbb{E} [e^{-\rho T} |V_T^\varepsilon - V_T|^2] \\ & \leq \mathbb{E} \left[\int_0^T e^{-\rho t} \left[\left(-\rho + 2K_b + 4\mu \right) |V_t^\varepsilon - V_t|^2 + \frac{1}{\mu} \|\dot{b}_x^\varepsilon(t) - \dot{b}_x(t)\|^2 |V_t|^2 \right. \right. \\ & \quad \left. \left. + \frac{1}{\mu} \|\dot{b}_x^\varepsilon(t) - \dot{b}_x(t)\|^2 |v_t - u_t|^2 + \sum_{j=1}^m \left(\|\dot{\sigma}_x^{\varepsilon,j}(t)\|^2 |V_t^\varepsilon - V_t|^2 + \|\dot{\sigma}_x^{\varepsilon,j}(t) - \dot{\sigma}_x^j(t)\|^2 |V_t|^2 \right. \right. \right. \\ & \quad \left. \left. + \|\dot{\sigma}_u^{\varepsilon,j}(t) - \dot{\sigma}_u^j(t)\|^2 |v_t - u_t|^2 + \frac{1}{\mu} \|\dot{\sigma}_x^{\varepsilon,j}(t)\|^2 \|\dot{\sigma}_x^{\varepsilon,j}(t) - \dot{\sigma}_x^j(t)\|^2 |V_t|^2 \right. \right. \\ & \quad \left. \left. + \frac{1}{\mu} \|\dot{\sigma}_x^{\varepsilon,j}(t)\|^2 \|\dot{\sigma}_x^{\varepsilon,j}(t) - \dot{\sigma}_x^j(t)\|^2 |v_t - u_t|^2 \right. \right. \\ & \quad \left. \left. + \|\dot{\sigma}_u^{\varepsilon,j}(t) - \dot{\sigma}_u^j(t)\|^2 \|\dot{\sigma}_x^{\varepsilon,j}(t) - \dot{\sigma}_x^j(t)\|^2 (|v_t - u_t|^2 + |V_t|^2) \right) dt \right] \\ & \leq \mathbb{E} \left[\int_0^T e^{-\rho t} \left(\left(-\rho + 2K_b + K_\sigma^2 + 4\mu \right) |V_t^\varepsilon - V_t|^2 + h^\varepsilon(t) \right) dt \right] \end{aligned}$$

where we have used the bound on the derivatives of b and σ and $h^\varepsilon(t)$ denotes the terms without $|V_t^\varepsilon - V_t|^2$,

$$h^\varepsilon(t) := \frac{1}{\mu} \|\dot{b}_x^\varepsilon(t) - \dot{b}_x(t)\|^2 |V_t|^2 + \frac{1}{\mu} \|\dot{b}_x^\varepsilon(t) - \dot{b}_x(t)\|^2 |v_t - u_t|^2$$

$$\begin{aligned}
 & + \|\dot{\sigma}_x^{\varepsilon,j}(t) - \dot{\sigma}_x^j(t)\|^2 |V_t|^2 + \|\dot{\sigma}_u^{\varepsilon,j}(t) - \dot{\sigma}_u^j(t)\|^2 |v_t - u_t|^2 \\
 & + \frac{1}{\mu} \|\dot{\sigma}_x^{\varepsilon,j}(t)\|^2 \|\dot{\sigma}_x^{\varepsilon,j}(t) - \dot{\sigma}_x^j(t)\|^2 |V_t|^2 + \frac{1}{\mu} \|\dot{\sigma}_x^{\varepsilon,j}(t)\|^2 \|\dot{\sigma}_x^{\varepsilon,j}(t) - \dot{\sigma}_x^j(t)\|^2 |v_t - u_t|^2 \\
 & + \|\dot{\sigma}_u^{\varepsilon,j}(t) - \dot{\sigma}_u^j(t)\|^2 \|\dot{\sigma}_x^{\varepsilon,j}(t) - \dot{\sigma}_x^j(t)\|^2 (|v_t - u_t|^2 + |V_t|^2)
 \end{aligned}$$

Let us denote $-\rho + 2K_b + K_\sigma^2 + 4\mu$ by $-\kappa$ where $\kappa > 0$ since $\mu > 0$ can be arbitrarily small and assumption H3 holds. Hence, we have $\forall T > 0$,

$$\begin{aligned}
 \mathbb{E}[e^{-\rho T} |V_T^\varepsilon - V_T|^2] + \mathbb{E}\left[\int_0^T e^{-\rho t} \kappa |V_t^\varepsilon - V_t|^2 dt\right] &\leq \mathbb{E}\left[\int_0^\infty e^{-\rho t} h^\varepsilon(t) dt\right] \\
 \mathbb{E}\left[\int_0^\infty e^{-\rho t} \kappa |V_t^\varepsilon - V_t|^2 dt\right] &\leq \mathbb{E}\left[\int_0^\infty e^{-\rho t} h^\varepsilon(t) dt\right]
 \end{aligned}$$

where the last step follows from Monotone convergence theorem. Next, using admissibility of the controls v and u , the bound on the derivatives of b and σ and (2.3.3), we get, $|h^\varepsilon(t)| \leq C(|V_t|^2 + |v_t|^2 + |u_t|^2)$, giving us

$$\mathbb{E}\left[\int_0^\infty e^{-\rho t} h^\varepsilon(t) dt\right] \leq C,$$

where C is a constant independent of ε . Also, using the continuity of the derivatives of b and σ , we get,

$$\lim_{\varepsilon \rightarrow 0} h^\varepsilon(t) = 0, \quad dt \otimes d\mathbb{P} \text{ a.e.}$$

This enables us to use Lebesgue's dominate convergence theorem to conclude that,

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E}\left[\int_0^\infty e^{-\rho t} |V_t^\varepsilon - V_t|^2 dt\right] = 0,$$

proving the lemma. □

Hence, we have a criterion that an optimal control must satisfy, given in terms of a point-wise minimization in the control space. This condition gives us the form of the optimal control in terms of the state and the adjoint processes (Y, Z) .

2.4 Sufficient condition

In this section, we will prove a sufficient condition for an admissible control to be optimal based on the notion of L-convexity. This notion can be further strengthened using strong convexity with respect to the control in order to prove the uniqueness of the optimal control.

Theorem 2.4.1. *Under assumptions H1, H2, H3, H4, let $u \in \mathcal{A}$ be an admissible control, X_t^u the corresponding state process and (Y, Z) the corresponding adjoint processes. Also, assume that*

$$\mathbb{R}^d \times U \times \mathcal{P}^2(\mathbb{R}^n) \ni (x, u, \mu) \mapsto H(t, x, u, \mu, Y_t, Z_t) \text{ is convex } dt \otimes d\mathbb{P} \text{ a.e.}$$

Suppose the following holds,

$$H(X_t^u, u_t, \mu_{X_t^u}, Y_t, Z_t) = \inf_{u \in U} H(X_t^u, u, \mu_{X_t^u}, Y_t, Z_t), \quad dt \otimes d\mathbb{P} \text{ a.e.} \quad (2.4.1)$$

Then u is an optimal control.

The above convexity condition is to be understood as follows,

$$H(x', u', \mu', Y_t, Z_t) \geq H(x, u, \mu, Y_t, Z_t) + \partial_x H(x, u, \mu, Y_t, Z_t)(x' - x)$$

$$+ \partial_u H(x, u, \mu, Y_t, Z_t)(u' - u) + \tilde{\mathbb{E}} \left[\partial_\mu H(x, u, \mu, Y_t, Z_t)(\tilde{X})(\tilde{X}' - \tilde{X}) \right]$$

where $X \sim \mu$ and $X' \sim \mu'$.

Proof. Consider an arbitrary admissible control $(u'_t)_{t \geq 0}$, $(X'_t)_{t \geq 0}$ the corresponding state process and $(\mu'_t)_{t \geq 0}$ the law of the state process. Consider the following, for any $T > 0$

$$\begin{aligned} & J(u') - J(u) \\ &= J_T(u') - J_T(u) + \mathbb{E} \left[\int_0^T e^{-\rho t} (f(X'_t, u'_t, \mu'_t) - f(X_t^u, u_t, \mu_{X_t^u})) dt \right] \\ &= J_T(u') - J_T(u) + \mathbb{E} \left[\int_0^T e^{-\rho t} ([H(X'_t, u'_t, \mu'_t, Y_t, Z_t) - H(X_t^u, u_t, \mu_{X_t^u}, Y_t, Z_t)]) dt \right] \\ &\quad - \mathbb{E} \left[\int_0^T e^{-\rho t} \left(Y_t [b(X'_t, u'_t) - b(X_t^u, u_t)] + \sum_{j=1}^m Z_t^j [\sigma^j(X'_t, u'_t) - \sigma^j(X_t^u, u_t)] - \rho Y_t (X'_t - X_t^u) \right) dt \right], \end{aligned}$$

where $J_T(u) = \mathbb{E} \left[\int_T^\infty e^{-\rho t} f(X_t^u, u_t, \mu_{X_t^u}) dt \right]$. Next, using the adjoint equation and considering the Itô's product formula for $e^{-\rho t} Y_t (X'_t - X_t^u)$, we get,

$$\begin{aligned} d(e^{-\rho t} Y_t (X'_t - X_t^u)) &= e^{-\rho t} \left(-\rho Y_t (X'_t - X_t^u) dt + (dY_t)(X'_t - X_t^u) + Y_t d(X'_t - X_t^u) \right. \\ &\quad \left. + d\langle Y, (X' - X^u) \rangle_t \right) \\ \mathbb{E} [e^{-\rho T} Y_T (X'_T - X_T^u)] &= \mathbb{E} \left[\int_0^T e^{-\rho t} \left(-\rho Y_t (X'_t - X_t^u) - [\partial_x H(\theta_t) + \tilde{\mathbb{E}}[\partial_\mu H(\tilde{\theta}_t)(X_t^u)]] \right) (X'_t - X_t^u) dt \right] \\ &\quad + \mathbb{E} \left[\int_0^T e^{-\rho t} (Y_t [b(\theta'_t) - b(\theta_t)] + Z_t [\sigma(\theta'_t) - \sigma(\theta_t)]) dt \right] \end{aligned}$$

where the expectation of the Itô integrals coming from the first equality is zero, $b(\theta'_t) = b(X'_t, u'_t)$ and so on. Using Fubini's theorem, we can write the second term as,

$$\begin{aligned} \mathbb{E} \left[\int_0^T e^{-\rho t} \left(\tilde{\mathbb{E}}[\partial_\mu H(\tilde{\theta}_t)(X_t^u)](X'_t - X_t^u) \right) dt \right] &= \mathbb{E} \left[\tilde{\mathbb{E}} \left(\int_0^T e^{-\rho t} \left([\partial_\mu H(\theta_t)(\tilde{X}_t^u)](\tilde{X}'_t - \tilde{X}_t^u) \right) dt \right) \right] \\ &= \mathbb{E} \left[\int_0^T e^{-\rho t} \left(\tilde{\mathbb{E}} \left[[\partial_\mu H(\theta_t)(\tilde{X}_t^u)](\tilde{X}'_t - \tilde{X}_t^u) \right] \right) dt \right]. \end{aligned}$$

Next, using (2.4.1),

$$\partial_u H(t, X_t^u, u_t, \mu_{X_t^u}, Y_t, Z_t) = 0,$$

and using the convexity condition, we have,

$$\begin{aligned} & J(u') - J(u) \\ &= J_T(u') - J_T(u) + \mathbb{E} \left[\int_0^T e^{-\rho t} ([H(X'_t, u'_t, \mu'_t, Y_t, Z_t) - H(X_t^u, u_t, \mu_{X_t^u}, Y_t, Z_t)]) dt \right] \\ &\quad - \mathbb{E} [e^{-\rho T} Y_T (X'_T - X_T^u)] - \mathbb{E} \left[\int_0^T e^{-\rho t} \left([\partial_x H(\theta_t)(X'_t - X_t^u) + \tilde{\mathbb{E}}[\partial_\mu H(\tilde{\theta}_t)(X_t^u)](\tilde{X}'_t - \tilde{X}_t^u)] \right) dt \right] \\ &\geq J_T(u') - J_T(u) - \mathbb{E} [e^{-\rho T} Y_T (X'_T - X_T^u)]. \end{aligned}$$

Next, since u and u' are admissible, the integrands of $J(u)$ and $J(u')$ are integrable. Hence,

$$\lim_{T \rightarrow \infty} |J_T(u)| + |J_T(u')| = 0.$$

Next, using Cauchy-Schwartz inequality for expectation, we have,

$$\mathbb{E} [e^{-\rho T} |Y_T(X'_T - X_T^u)|] \leq \mathbb{E} [e^{-\rho T} |Y_T|^2]^{\frac{1}{2}} \mathbb{E} [e^{-\rho T} |X'_T - X_T^u|^2]^{\frac{1}{2}},$$

where the limit of the first factor is zero because of the terminal condition and second factor is finite from the a priori estimate for X (see Proposition 2.2.1).

Hence, we have,

$$J(u') - J(u) \geq 0,$$

implying that u is an optimal control. \square

Next, using the strong convexity condition, we can state the following uniqueness result.

Corollary 2.4.1.1. *Under the above assumptions and the following additional assumption of strong convexity with respect to control,*

$$\begin{aligned} H(t, x', u', \mu', Y_t, Z_t) &\geq H(t, x, u, \mu, Y_t, Z_t) + \partial_x H(t, x, u, \mu, Y_t, Z_t)(x' - x) \\ &\quad + \partial_u H(t, x, u, \mu, Y_t, Z_t)(u' - u) + \tilde{\mathbb{E}} \left[\partial_\mu H(t, x, u, \mu, Y_t, Z_t)(\tilde{X})(\tilde{X}' - \tilde{X}) \right] \\ &\quad + \alpha |u' - u|^2, \end{aligned}$$

we above optimal control is unique $dt \otimes d\mathbb{P}$ a.s.

Proof. The strong convexity condition changes (2.4) as follows,

$$\begin{aligned} &J(u') - J(u) \\ &= J_T(u') - J_T(u) + \mathbb{E} \left[\int_0^T e^{-\rho t} ([H(X'_t, u'_t, \mu'_t, Y_t, Z_t) - H(X_t^u, u_t, \mu_{X_t^u}, Y_t, Z_t)]) dt \right] \\ &\quad - \mathbb{E} [e^{-\rho T} Y_T(X'_T - X_T^u)] - \mathbb{E} \left[\int_0^T e^{-\rho t} \left([\partial_x H(\theta_t) + \tilde{\mathbb{E}}[\partial_\mu H(\tilde{\theta}_t)(X_t^u)]] (X'_t - X_t^u) \right) dt \right] \\ &\geq J_T(u') - J_T(u) - \mathbb{E} [e^{-\rho T} Y_T(X'_T - X_T^u)] + \mathbb{E} \left[\int_0^T e^{-\rho t} (\alpha |u'_t - u_t|^2) dt \right] \end{aligned}$$

The rest of the proof follows and we obtain $J(u') - J(u) \geq \mathbb{E} \left[\int_0^T e^{-\rho t} (\alpha |u'_t - u_t|^2) dt \right]$ implying uniqueness in the $dt \otimes d\mathbb{P}$ sense in \mathcal{A} . \square

Combining this with necessary conditions, we can conclude that if f is strongly convex in control and L-convex with respect to the law of the state process, then the solution of the point-wise minimization problem (2.4.1) is the unique optimal control which can be written in terms of the state and adjoint processes. This enables use to write a coupled system of FBSDEs which are tackled in Chapter 3.

2.5 Examples

2.5.1 The mean-field infinite horizon LQ case

In this sub-section, we study a special case of our problem where the state dynamics are linear in state and control processes and the running cost function is quadratic in the state, control and expectation of the state process. We derive the resulting coupled system of McKean-Vlasov

FBSDE using the necessary conditions which give the explicit form for the optimal control. However, we will not tackle its solvability here. We refer to [88] for this problem in more general setting where they show the equivalence of this problem to solving a system of algebraic Riccati equations and a BSDE. We can write the state dynamics as,

$$dX_t^u = (AX_t^u + Bu_t)dt + (CX_t^u + Du_t)dW_t, \quad X_0^u = x_0. \quad (2.5.1)$$

where $A, C \in \mathbb{R}^{d \times d}$ and $B, D \in \mathbb{R}^{d \times n}$. Let us consider the following cost functional,

$$J(u) = \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(X_t^{u\top} Q X_t^u + u_t^\top R u_t + 2 \left(q(t)^\top X_t^u + p(t)^\top u(t) \right) + \mathbb{E} \left[X_t^{u\top} \right] S \mathbb{E} \left[X_t^u \right] \right) dt \right]$$

where Q and S are symmetric matrices in $\mathbb{R}^{d \times d}$; $q(\cdot)$ and $p(\cdot)$ are progressively measurable and square integrable; R is a symmetric matrix in $\mathbb{R}^{n \times n}$ which is also invertible (this assumption is not important as one could use the pseudo-inverse as done by [88]). We can write the Hamiltonian as done in (2.2.3):

$$\begin{aligned} H(X_t^u, u_t, Y_t, Z_t, \mathbb{E}[X_t^u]) = & Y_t (AX_t^u + Bu_t) + Z_t (CX_t^u + Du_t) + \left(X_t^{u\top} Q X_t^u + u_t^\top R u_t + \right. \\ & \left. 2 \left(q(t)^\top X_t^u + p(t)^\top u_t \right) + \mathbb{E} \left[X_t^{u\top} \right] S \mathbb{E} \left[X_t^u \right] \right) - \rho X_t^u Y_t. \end{aligned}$$

It can be easily verified that our assumptions H1, H2, H3 and H4 hold. We can use Proposition 3.2.1 from Chapter 3 (originally from [36]) to prove that the L-derivative of the term $\mathbb{E} \left[X_t^{u\top} \right] S \mathbb{E} \left[X_t^u \right]$ at X_t^u is given by $2X_t^{u\top} S$. We can write the adjoint process (BSDE) as follows:

$$dY_t = - \left(Y_t (A - \rho \mathbf{I}) + Z_t C + 2X_t^{u\top} Q + 2q(t)^\top + \mathbb{E} \left[2X_t^{u\top} S \right] \right) dt + Z_t dW_t \quad (2.5.2)$$

for all $t \geq 0$ with $\lim_{T \rightarrow \infty} \mathbb{E} \left[e^{-\rho T} |Y_T|^2 \right] = 0$. Note that the process Y_t is written as a row vector taking values in $\mathbb{R}^{1 \times n}$. The existence and uniqueness of the solution of this BSDE for a given admissible control u can be proved using Proposition 2.2.2. The necessary conditions, following Theorem 2.3.1, can be written as,

$$\partial_u H(X_t^u, u, Y_t, Z_t, \mathbb{E}[X_t^u]) = Y_t B + Z_t D + 2u_t^\top R + 2p(t)^\top = 0.$$

Using the invertibility of R , we can simply write the optimal control as a linear function of the processes (Y, Z) , given by

$$u_t = -\frac{1}{2} R^{-1} \left(B^\top Y_t^\top + D^\top Z_t^\top + 2p(t) \right).$$

Using this in the BSDE (2.5.2) and (2.5.1), we get the following FBSDE,

$$\begin{aligned} dX_t^u = & \left(AX_t^u + -\frac{1}{2} B R^{-1} \left(B^\top Y_t^\top + D^\top Z_t^\top + 2p(t) \right) \right) dt + \\ & \left(CX_t^u + -\frac{1}{2} D R^{-1} \left(B^\top Y_t^\top + D^\top Z_t^\top + 2p(t) \right) \right) dW_t, \quad X_0^u = x_0 \\ dY_t = & - \left(Y_t (A - \rho \mathbf{I}) + Z_t C + 2X_t^{u\top} Q + 2q(t)^\top + \mathbb{E} \left[2X_t^{u\top} S \right] \right) dt + Z_t dW_t. \end{aligned}$$

This is an example of a linear and coupled FBSDE which have been well-studied in the literature and there are many approaches for solving them (see Yong [116] for example).

2.5.2 Optimal liquidation problem in infinite horizon

We refer to [1] for the model we consider here, where they consider a finite horizon T by which they want to liquidate their position on a certain asset. When the trades are large enough, one does not get the desired price and there are execution costs incurred which includes market impact. Depending on the size of the trades, there could be either a permanent or temporary impact. We refer to [1] for more references on the subject. There is a trade-off between trading fast when the prices are favorable and trading slowly to reduce the market impact. The price process can be modeled as,

$$dS_t = \lambda \mathbb{E}[\alpha_t] dt + \sigma dW_t, \quad S_0 = s_0,$$

where $\lambda \geq 0$ and $\lambda \mathbb{E}[\alpha_t]$ is the permanent impact due to all agents who are trading this asset. The asset inventory satisfies $dQ_t = \alpha_t dt$ where Q_0 is the initial inventory and we want to deplete it eventually. The wealth process can be written as,

$$dV_t = -\alpha_t(S_t + k\alpha_t)dt, \quad V_0 = 0,$$

where $k\alpha_t$ is the temporary market impact due to each agent. In [1], they consider a finite time horizon and liquidate their position at time T , resulting in the following cost functional.

$$J(\alpha) = \mathbb{E} \left[\int_0^T (\phi Q_t^2 + \alpha_t S_t + k\alpha_t^2) dt - Q_T(S_T - AQ_T) \right],$$

where ϕ is a risk-aversion parameter and $(S_T - AQ_T)$ is the reduced value we obtain for the asset upon forced liquidation at time T . The authors in [1] established explicit solutions for the optimal feedback control and inventory levels for the finite horizon case. Alternatively, one could consider a model with some uncertainty on the terminal horizon. For example, we could have the terminal time as a random variable τ when we are forced to liquidate, which could for instance have exponential distribution $\mathcal{E}(\rho)$ and be independent of the Brownian motion. We want to maximize the wealth at time τ and have a running cost on the size of the inventory. Then, we get the following cost functional,

$$J(\alpha) = \mathbb{E} \left[\int_0^\infty e^{-\rho t} (\phi Q_t^2 + \alpha_t S_t + k\alpha_t^2 - Q_t(S_t - AQ_t)) dt \right],$$

Let us emphasize that the distribution dependence is of scalar interaction type. However, this model is different from our setting due to dependence on the law of the control rather than the state.

2.5.3 Challenges with standard examples

In this subsection, we summarize several examples of possible mean-field dependencies in the cost functional which are not scalar interactions and could serve as a distance from a given target distribution. More precisely, we looked at following examples.

1. Squared 2-Wasserstein's distance between distribution of the state and a target distribution π^* ,

$$W_2^2(\mu_{X_t^u}, \pi^*) = \inf_{\gamma \in \Pi(\mu_{X_t^u}, \pi^*)} \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\gamma(x, y) \right\}.$$

The L-differentiability was proved by Gangbo and Tudorascu [61] and also by Alfonsi and Jourdain [5] by using a purely probabilistic approach, under the condition that there is a unique optimal coupling between $\mu_{X_t^u}$ and π^* which is given by a transport map. However, the L-convexity does not hold in general, see e.g. the counter-example in Chapter 5, [37]. However, it is essential for applying our sufficient conditions for optimality.

2. Kullback-Leibler divergence,

$$d_{KL}(\mu|\nu) := \int_{\mathbb{R}^d} \frac{d\mu(x)}{d\nu} \log \left(\frac{d\mu(x)}{d\nu} \right) d\nu(x), \text{ if } \mu \ll \nu \text{ and } +\infty \text{ otherwise.}$$

The L-differentiability and L-convexity were not clear for this example.

3. Energy distance (see Rizzo and Székely [108]),

$$D_E(\mu, \nu) = \int_{\mathbb{R}^d} \left| \mathbb{E} \left[e^{i\lambda X} \right] - \mathbb{E} \left[e^{i\lambda Y} \right] \right|^2 w(\lambda) d\lambda.$$

The L-differentiability and L-convexity were not clear for this example.

We could not find any classical working examples which posed a major challenge in the thesis when it came to applying the maximum principle approach for our mean-field control problem, leading us to work on designing a measure dependence which was not scalar interactions and not only satisfies the L-differentiability and L-convexity conditions but was also applicable to portfolio optimization. This leads us to Chapter 3 of the thesis.

Chapter 3

Long term dynamic portfolio optimization using infinite horizon McKean-Vlasov FBSDEs

Abstract. We model the problem of long term optimal allocation of a self-financing portfolio of stocks and a risk-free asset as an infinite horizon mean-field stochastic control problem. We aim to outperform a static portfolio by using dynamical allocation, using the trading speed as the control and a term in the running cost functional which is quadratic in control to minimize the market impact and costs associated with the trades. The optimization criterion is based on a mean-field term in the cost functional depending on the distribution of the relative wealth process for the corresponding allocation, which corresponds to a downside risk minimization problem. The model is especially more relevant when the time horizon of the investment is not known a priori and we aim to keep the wealth process close to the target distribution at all times. We study the necessary and sufficient conditions for the optimality of a control and characterize it using the stochastic Pontryagin's maximum principle, proving its equivalence with a coupled system of McKean-Vlasov forward-backward stochastic differential equation (FBSDE) in the infinite horizon. We prove the existence and uniqueness of the solution for this McKean-Vlasov FBSDE using a continuation argument and study two numerical schemes based on a long finite horizon truncation for which we establish bounds on the truncation error. The numerical resolution of the McKean-Vlasov FBSDE involves deep neural networks for which we also provide some numerical experiments.

Contents

3.1	Introduction	66
3.2	Model and preliminaries	67
3.2.1	State dynamics and cost functional	67
3.2.2	Results for the mean-field term	71
3.2.3	The control problem and assumption	73
3.3	Conditions for optimality	74
3.3.1	A priori estimate on state process	74
3.3.2	Adjoint processes	76
3.3.3	Necessary and sufficient conditions for optimality	76
3.3.4	Characterisation of optimal control using FBSDE	77
3.4	Numerical Analysis	83

3.4.1	General setting	87
3.4.2	Scheme based on fictitious play	88
3.4.3	Scheme based on dynamic updates	89
3.4.4	Numerical Experiments	90
3.5	Proofs	94
3.5.1	Proof of Theorem 3.3.3	94
3.5.2	Proof of Theorem 3.3.4	97

3.1 Introduction

In this chapter, we model a long term portfolio optimization problem as a mean-field stochastic control problem with an infinite time horizon. Mean field control (MFC) problems are closely related to mean-field games (MFGs) which were introduced by Lasry and Lions [85] and Huang et al. [80] who provided a powerful mathematical tool for modeling the strategic interactions of a large number of rational agents. Theoretical frameworks, such as the mean-field game formulation by Pardoux et al. [102], have laid the groundwork for understanding the equilibrium behavior of large populations in the presence of mean-field interactions. For a broader coverage and perspectives for mean-field games, see Bensoussan, Frehse, Yam, et al. [18] and an extensive study of MFGs and MFC problems along with applications can be found in the two volumes of books written by Carmona and Delarue [37], which have arguably provided one of the best detailed coverage of the topic.

We use a probabilistic approach based on the stochastic maximum principle to solve this problem, which gives rise to an equivalent coupled system of McKean-Vlasov forward-backward stochastic differential equations (FBSDEs) in the infinite horizon. Methods for solving McKean-Vlasov FBSDEs have gained a lot of traction in the past few years due to numerous applications of MFC and MFGs. Alasseur et al. [4] introduced mean-field FBSDEs for power grid stability analysis, considering the interactions among multiple generators and loads. Hafiz et al. [76] integrated stochastic control with machine learning techniques for optimal real-time operation of energy storage systems, showcasing the potential of combining diverse methodologies. Acciaio, Backhoff-Veraguas, and Carmona [1] studied an ‘extended’ MFG using the maximum principle and provided insights into the connection with the theory of optimal transport and solved an optimal liquidation problem. Aïd, Basei, and Pham [3] used the McKean-Vlasov approach for modeling distributed energy generation. More recent works include the use of McKean-Vlasov stochastic control for the energy management of a smart grid by Gobet and Grangereau [64] and in [65], they developed a novel approach combining Newton’s method and Pontryagin’s principle for stochastic control. Germain et al. [62] used mean-field dynamics with probabilistic state constraints and level-set approach for renewable energy storage problem.

The mean-field term in our problem is not of scalar interactions type which makes the numerical tractability of this problem challenging. However, probabilistic approaches based on FBSDEs have been widely explored by researchers in the last few years. Chassagneux, Crisan, and Delarue [44] used local Picard iterations on small time intervals and used a tree structure to represent the pathwise law of the solution and established bounds on rate of convergence. Alfonsi and Jourdain [5] developed another Picard scheme using a grid structure to represent marginal laws of the solution. However, as the authors remarked, Picard schemes converge only in small time horizons. Several approaches using neural networks (NNs) have become very popular since the seminal work of E, Han, and Jentzen [56] for solving high dimensional parabolic PDEs using BSDEs where they used a deep feed-forward NN using sub-NNs to approximate the Z process on a discrete time grid. Similar approach was used by Han and Long [77] for solving coupled FBSDEs and by Han et al. [78] for general MFC problems and MFGs using a fictitious

play approach. Germain et al. [63] developed several other approaches based on NNs for both the strong and weak formulation of the control problem, of which the most promising was an approach based on dynamic updates of the law at each time step.

Statement of the problem. We consider a long term optimal allocation problem for a portfolio of stocks and a bond in finance where the objective is to attain a target distribution for the wealth process by dynamically controlling the asset allocation at regular time intervals. We model the problem as a continuous time mean-field stochastic control (MFC) problem with infinite horizon. We aim at outperforming a static asset allocation with the same initial endowment and hence consider as state process the ratio of the dynamically controlled wealth process and the static allocation wealth process, which we will henceforth call the relative wealth process. Dynamic allocation of stocks comes at a price as one has to account for the trading costs and market impact for large trades. To take this into account, as a simplification, we use the trading speed of the assets as the control for the problem and a quadratic term in control in the cost functional which serves a two-fold purpose by penalizing high trading speeds and making the cost functional strongly convex with respect to control, helping us obtain existence and uniqueness guarantees for the control problem.

From a financial point of view, it is desirable to have a right-skewed distribution with longer right tails as the target distribution for the relative wealth process. Hence, we take a log-normal target distribution which evolves with time. Precisely, it is given by the distribution of a geometric Brownian motion (GBM) with constant positive drift and volatility. To penalize deviation of the relative wealth process from the target distribution, we add mean-field term to the cost functional which acts as the distance from the target distribution. For this term, we take an approach similar to downside risk minimization, as made precise in the next section.

Contributions. Our main contributions are in modeling a portfolio optimization problem in large time horizon using an infinite horizon mean-field stochastic control with non-time-homogeneous system dynamics using a non-stationary target distribution for a risk-appetite based mean-field criterion and benchmark portfolio out-performance using trading speed as the control. Taking a quadratic term in control in the objective function allows us to penalize high trading speeds, helping to mitigate high transaction costs. In section 2, we state the control problem corresponding to this portfolio optimization problem and prove some results for the mean-field term. In section 3, we prove some necessary and sufficient conditions for the optimality of a control and prove uniqueness under a strong convexity condition with respect to the control. The necessary condition gives the explicit form of the optimal control in terms of the solution of the adjoint process. We also prove the existence and uniqueness of the solution of the coupled system of FBSDE corresponding to this control problem. In section 4, we describe two numerical schemes using neural networks, based on a finite horizon truncation. We provide some numerical experiments and also illustrate that increasing the multiplier of the mean-field term skews the relative wealth distribution towards right while not significantly increasing probability of negative returns.

3.2 Model and preliminaries

3.2.1 State dynamics and cost functional

State dynamics. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a filtration $(\mathcal{F}_t)_{t \geq 0}$ which satisfies the usual conditions and is a natural filtration of a m -dimensional Brownian motion $(W_t)_{t \geq 0}$ with $W_0 = 0$. We assume that the rates of lending and borrowing are the same. Consider a portfolio consisting of N risky assets modeled as stochastic differential equations (SDEs) $(S^i)_{i=1, \dots, N}$ and

a risk-free asset \mathbf{S}^0 which satisfy the following dynamics for all $t \geq 0$,

$$\begin{aligned} d\mathbf{S}_t^i &= \mathbf{S}_t^i \left(b_t^i dt + \sum_{j=1}^m \sigma_t^{i,j} dW_t^j \right), \quad i \in \{1, 2, \dots, N\} \\ d\mathbf{S}_t^0 &= \mathbf{S}_t^0 r_t dt \end{aligned}$$

where b_t^i , $\sigma_t^{i,j}$ and r_t are \mathbb{R} -valued bounded stochastic processes with $|b_t^i| \leq A_b^i$, $|\sigma_t^{i,j}| \leq A_\sigma^{i,j}$ and $0 \leq r_t \leq A_r$. For numerical purposes, we will consider constant b^i , $\sigma^{i,j}$ and r . Denote by the process $(u_t)_{t \geq 0}$ valued in \mathbb{R}^{N+1} the value of wealth in the $N+1$ assets. Consider the value of this portfolio, $(V_t^u)_{t \geq 0}$ valued in \mathbb{R} , given by $V_t^u = \sum_{i=0}^N u_t^i$.

Suppose that the changes in wealth process can only happen due to changes in allocation across the assets and changes in the value of these assets. In particular, the agent does not withdraw money or make further investment after the initial endowment. This restriction is called the ‘*self-financing condition*’ and implies the following dynamics of the wealth process,

$$\begin{aligned} dV_t^u &= \sum_{i=0}^N (u_t^i) \frac{d\mathbf{S}_t^i}{\mathbf{S}_t^i} = u_t^0 r_t dt + \sum_{i=1}^N u_t^i \left(b_t^i dt + \sum_{j=1}^m \sigma_t^{i,j} dW_t^j \right) \\ &= \left(V_t^u - \sum_{i=1}^N u_t^i \right) r_t dt + \sum_{i=1}^N u_t^i \left(b_t^i dt + \sum_{j=1}^m \sigma_t^{i,j} dW_t^j \right) \\ &= r_t V_t^u dt + \sum_{i=1}^N u_t^i \left((b_t^i - r_t) dt + \sum_{j=1}^m \sigma_t^{i,j} dW_t^j \right) \end{aligned}$$

where $V_0^u = v_0 > 0$ is the initial endowment. We would like to control the trading speed of the risky assets, denoted by $(q^i)_{i=1, \dots, N}$ rather than the actual amount held in them. Let us write the vector $[u_t^i]_{i=1}^N$ as $[\delta_t^i \mathbf{S}_t^i]_{i=1}^N$ where δ_t^i is the number of i -th asset at time t and \mathbf{S}_t^i is the price of the i -th asset. So, the wealth process is written as,

$$\begin{aligned} dV_t^q &= r_t V_t^q dt + \sum_{i=1}^N \delta_t^i \mathbf{S}_t^i \left((b_t^i - r_t) dt + \sum_{j=1}^m \sigma_t^{i,j} dW_t^j \right) \\ d\delta_t^i &= q_t^i dt, \end{aligned}$$

where we take $[q_t^i]_{i=1}^N$, the trading speed vector as the control. We do not put any constraints on the sign of V^q . We would like to compare this wealth process with a static investment in the same assets with same initial endowment, i.e. a portfolio $[\gamma^i]_{i=0}^N$ which does not change with time, henceforth called the reference portfolio. This investment strategy gives the wealth process with the following dynamics,

$$d\bar{V}_t = r_t \bar{V}_t dt + \sum_{i=1}^N \gamma^i \mathbf{S}_t^i \left((b_t^i - r_t) dt + \sum_{j=1}^m \sigma_t^{i,j} dW_t^j \right),$$

where $\bar{V}_t > 0$ is guaranteed for all t because the asset allocations are fixed with $\gamma^i \geq 0$. Let us denote $\mathcal{V}_t^q := V_t^q / \bar{V}_t$ and $\mathcal{S}_t^i = \mathbf{S}_t^i / \bar{V}_t$, which are the wealth process and assets prices in the numeraire of the reference portfolio, henceforth called the *relative wealth* and *relative asset price* respectively. Let us assume $\gamma^i > 0$, so that (since $\mathbf{S}_t^i \geq 0$)

$$|\mathcal{S}_t^i| = \left| \frac{\mathbf{S}_t^i}{\bar{V}_t} \right| = \left| \frac{1}{\gamma^i} \frac{\gamma^i \mathbf{S}_t^i}{\sum_{i=0}^N \gamma^i \mathbf{S}_t^i} \right| \leq \frac{1}{\gamma^i}.$$

Using Itô's formula, the dynamics of the relative wealth process are given as follows:

$$\begin{aligned}
 d\mathcal{V}_t^q &= d\left(\frac{V_t^q}{\bar{V}_t}\right) = \frac{dV_t^q}{\bar{V}_t} - \frac{V_t^q d\bar{V}_t}{(\bar{V}_t)^2} + \frac{V_t^q d\langle \bar{V}, \bar{V} \rangle_t}{(\bar{V}_t)^3} - \frac{d\langle V^q, \bar{V} \rangle_t}{(\bar{V}_t)^2} \\
 &= \left[\frac{V_t^q}{\bar{V}_t} r_t dt + \sum_{i=1}^N \frac{S_t^i}{\bar{V}_t} \delta_t^i (b_t^i - r_t) dt + \sum_{j=1}^m \sum_{i=1}^N \frac{S_t^i}{\bar{V}_t} \delta_t^i \sigma_t^{i,j} dW_t^j \right] - \frac{V_t^q}{\bar{V}_t} \left[r_t dt + \sum_{i=1}^N \frac{S_t^i}{\bar{V}_t} \gamma^i (b_t^i - r_t) dt \right. \\
 &\quad \left. + \sum_{j=1}^m \sum_{i=1}^N \frac{S_t^i}{\bar{V}_t} \gamma^i \sigma_t^{i,j} dW_t^j \right] + \frac{V_t^q}{\bar{V}_t} \left[\sum_{j=1}^m \left(\sum_{i=1}^N \frac{S_t^i}{\bar{V}_t} \gamma^i \sigma_t^{i,j} \right)^2 dt \right] - \left[\sum_{j=1}^m \left(\sum_{i=1}^N \frac{S_t^i}{\bar{V}_t} \delta_t^i \sigma_t^{i,j} \right) \left(\sum_{i=1}^N \frac{S_t^i}{\bar{V}_t} \gamma^i \sigma_t^{i,j} \right) dt \right] \\
 &= \left[\sum_{i=1}^N S_t^i (\delta_t^i - \gamma^i \mathcal{V}_t^q) (b_t^i - r_t) - \sum_{j=1}^m \left(\sum_{i=1}^N S_t^i \gamma^i \sigma_t^{i,j} \right) \left(\sum_{i=1}^N S_t^i (\delta_t^i - \mathcal{V}_t^q \gamma^i) \sigma_t^{i,j} \right) \right] dt \\
 &\quad + \sum_{j=1}^m \left[\sum_{i=1}^N S_t^i (\delta_t^i - \gamma^i \mathcal{V}_t^q) \sigma_t^{i,j} \right] dW_t^j. \tag{3.2.1}
 \end{aligned}$$

Let us write the above SDE as follows:

$$d\mathcal{V}_t^q = (\alpha_t^\mathcal{V} \mathcal{V}_t^q + \alpha_t^\delta \cdot \delta_t) dt + \sum_{j=1}^m (\beta_t^{\mathcal{V},j} \mathcal{V}_t^q + \beta_t^{\delta,j} \cdot \delta_t) dW_t^j,$$

giving a SDE whose drift and volatility are stochastic but bounded and linear in \mathcal{V}^q and δ . The coefficients are given in terms of the original parameters as follows:

$$\begin{aligned}
 \alpha_t^\mathcal{V} &= \sum_{j=1}^m \left(\sum_{i=1}^N S_t^i \gamma^i \sigma_t^{i,j} \right)^2 - \sum_{i=1}^N S_t^i \gamma^i (b_t^i - r_t), \\
 \alpha_t^\delta &= \left(S_t^i (b_t^i - r_t) - \sum_{j=1}^m \left(\sum_{k=1}^N S_t^k \gamma^k \sigma_t^{k,j} \right) \left(S_t^i \sigma_t^{i,j} \right) \right)_{i \in \{1, \dots, N\}}, \\
 \beta_t^{\mathcal{V},j} &= - \sum_{i=1}^N S_t^i \gamma^i \sigma_t^{i,j}, \\
 \beta_t^{\delta,j} &= \left(S_t^i \sigma_t^{i,j} \right)_{i \in \{1, \dots, N\}}.
 \end{aligned}$$

Note that the state dynamics depends on δ which is the time integral of the control. So, we consider the state vector for our control problem, $(X_t^q)_{t \geq 0}$ given by the following linear SDE,

$$dX_t^q = d \begin{bmatrix} \mathcal{V}_t^q \\ \delta_t \end{bmatrix} = \begin{bmatrix} \alpha_t^\mathcal{V} \mathcal{V}_t^q + \alpha_t^\delta \delta_t \\ q_t^1 \\ \vdots \\ q_t^N \end{bmatrix} dt + \sum_{j=1}^m \begin{bmatrix} \beta_t^{\mathcal{V},j} \mathcal{V}_t^q + \beta_t^{\delta,j} \delta_t \\ 0 \\ \vdots \\ 0 \end{bmatrix} dW_t^j.$$

Cost functional. The mean-field dependence comes in the form of a downside risk, where we penalize excess losses beyond a certain specified threshold which comes from the target distribution. This threshold is given as the magnitude of loss they are willing to bear for a given value of attained wealth relative to the benchmark. With these considerations, we start with the following cost functional,

$$J^\lambda(q) = \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(g(X_t^q, q_t) + \int_0^\infty \left[\mathbb{E}[k - \mathcal{V}_t^q]_+^\lambda - p_k \right]_+^\lambda d\nu(k) \right) dt \right].$$

where $[k-x]_+$ is the positive part of the function $(k-x)$ and $[k-x]_+^\lambda$ is a smooth approximation of this function and ν is a non-negative measure on \mathbb{R} chosen to ensure integrability with respect to k . For example, one could take ν to be absolutely continuous with respect to Lebesgue measure and hence get a weighted integral over k , however some applications might require use of discrete measures. Note that the cost functional not only depends on the value of the state but also its distribution, making it a McKean-Vlasov control problem. So, we would be interested in finding conditions on g, ρ, p_k and ν in order to ensure the existence of an optimal control and its characterisation as an FBSDE.

In general, we would also want to prove the square integrability of the integral over k , which would give us some conditions on the measure ν and p_k . We get,

$$\begin{aligned} \int_0^\infty \left| \left[\mathbb{E} [k - \mathcal{V}_t^g]_+^\lambda - p_k \right]_+^\lambda \right|^2 d\nu(k) &\leq \int_0^\infty (\mathbb{E} [|\mathcal{V}_t^g|] + |k| + |p_k|)^2 d\nu(k) \\ &\leq C \int_0^\infty (\mathbb{E} [|\mathcal{V}_t^g|^2] + |k|^2 + |p_k|^2) d\nu(k), \end{aligned}$$

where we have used $[x]_+ \leq |x|$, Cauchy-Schwartz inequality for the mixed terms coming from expansion for the last inequality, Jensen's inequality for $(\mathbb{E} [|\mathcal{V}_t^g|])^2 \leq \mathbb{E} [|\mathcal{V}_t^g|^2]$ and C is independent of k and t . See (H0) for detailed assumptions on ν and p_k .

For example, one could take ν to be a discrete measure or an exponential distribution with density $\gamma e^{-\gamma k}$ (for some $\gamma > 0$) provided that p_k has sub-exponential growth. Next, we interpret this cost functional with a financial point of view.

Financial interpretation:

- \mathcal{V} corresponds to the discounted wealth process (relative wealth) for the trading speed g . Having the term $\left[\mathbb{E} [k - \mathcal{V}_t^g]_+^\lambda - p_k \right]_+$ in the cost means that we want to limit the losses below the threshold p_k for a given k . For example, when k is small, one may want to have small p_k so that if the portfolio goes into a certain amount of loss at a small value of wealth, it gets penalised more aggressively and as the wealth grows (penalisation triggered only for higher k), a higher p_k would correspond to a bigger risk appetite.
- We would also like ν to assign less weight at higher values of k . It is not only essential for integrability but also a decreasing weight with k would mean that the losses are less penalised for higher wealth. For example, when the wealth process is at \$10,000, a loss of \$10 would be insignificant and must have smaller penalisation compared to when the wealth process is at \$100.
- For applications other than finance, one may want to include another term $\left[\mathbb{E} [\mathcal{V}_t^g - k]_+ - c_k \right]_+$ when they want to penalise higher values of the state. However, we don't consider this term in our settings because it would mean that we are penalising the profits.

Next, let us consider the following smooth approximation of $[k-x]_+^\lambda$. Note that the derivative of $[k-x]_+$ is -1 on the left side of k , 0 on the right and not defined on k . So we could approximate the derivative by a sigmoid function,

$$\Sigma(x) = \frac{-e^{\lambda(k-x)}}{1 + e^{\lambda(k-x)}}$$

and similarly for $[x-k]_+$. So, we could approximate $[x-k]_+$ and respectively $[k-x]_+$ by the above anti-derivatives,

$$[x-k]_+ \approx [x-k]_+^\lambda := \frac{1}{\lambda} \log \left| 1 + e^{\lambda(x-k)} \right| \tag{3.2.2}$$

$$[k - x]_+ \approx [k - x]_+^\lambda := \frac{1}{\lambda} \log \left| 1 + e^{\lambda(k-x)} \right|$$

and we approach the true value when $\lambda \rightarrow \infty$.

Remark 3.2.1. Alternatively, one could use the following approximation:

$$\begin{aligned} [x - k]_+ &\approx [x - k]_+^\varepsilon := \frac{(x - k + \varepsilon)^2}{4\varepsilon} \quad \forall x \in [k - \varepsilon, k + \varepsilon], [x - k]_+ \text{ otherwise;} \\ [k - x]_+ &\approx [k - x]_+^\varepsilon := \frac{(k - x + \varepsilon)^2}{4\varepsilon} \quad \forall x \in [k - \varepsilon, k + \varepsilon], [k - x]_+ \text{ otherwise,} \end{aligned}$$

where we approach the true value when $\varepsilon \rightarrow 0$.

3.2.2 Results for the mean-field term

When studying the optimality conditions, we would need some regularity conditions for functions of measures in $\mathcal{P}^2(\mathbb{R}^d)$ (space of measures with finite second moment). For this, we recall the notion of Lions differentiability (L-differentiability for short) from [36] (see Definition 2.2.2 from Chapter 2). We also need a notion of convexity of L-differentiable functions for which we recall Definition 2.2.3 from Chapter 2. We have the following useful result on L-differentiability (see Section 5.2.2. Example-1 in [36]):

Proposition 3.2.1. *Let $u(\mu) = \int_{\mathbb{R}^d} h(x) d\mu(x)$ and assume that h is continuously differentiable whose derivative has at most linear growth. Then, the L-derivative of u at X is given by $\partial_\mu u(\mu)(X) = \partial h(X)$ where $X \sim \mu$.*

Next, we state a result from [36] (see Section 5.5.1, Example 1) about L-convex functions:

Proposition 3.2.2. *Given a non-decreasing convex differentiable function $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ and a convex differentiable function $\zeta : \mathbb{R}^d \rightarrow \mathbb{R}$, whose derivative is of at most linear growth, the function*

$$\varphi(\langle \zeta, \mu \rangle) := \varphi \left(\int_{\mathbb{R}^d} \zeta(x) d\mu(x) \right) = \varphi(\mathbb{E}[\zeta(X)])$$

is L-convex where $X \sim \mu$ and the L-derivative at X is given by $\varphi'(\mathbb{E}[\zeta(X)])\partial\zeta(X)$.

Using these results, we can prove the L-differentiability and L-convexity of the measure dependent part of the cost functional, given as follows,

$$h(\mu_{\mathcal{V}_t^q}) := \int_0^\infty \left[\mathbb{E}[k - \mathcal{V}_t^q]_+^\lambda - p_k \right]_+^\lambda d\nu(k),$$

where ν is a probability measure. It will be useful to consider separately the integrand for a given k , $h_k(\mu_{\mathcal{V}_t^q}) := \left[\mathbb{E}[k - \mathcal{V}_t^q]_+^\lambda - p_k \right]_+^\lambda$.

Proposition 3.2.3. *The measure dependent part of the cost function,*

$$h(\mu_{\mathcal{V}_t^q}) := \int_0^\infty h_k(\mu_{\mathcal{V}_t^q}) d\nu(k) = \int_0^\infty \left[\mathbb{E}[k - \mathcal{V}_t^q]_+^\lambda - p_k \right]_+^\lambda d\nu(k),$$

is L-differentiable and L-convex with the L-derivative at \mathcal{V}_t^q given by,

$$\partial_\mu h(\mu)(\mathcal{V}_t^q) = \int_0^\infty \partial_\mu h_k(\mu)(\mathcal{V}_t^q) d\nu(k) = \int_0^\infty \frac{-1}{1 + e^{\lambda p_k - (\mathbb{E}[\log(1 + e^{\lambda(k - \mathcal{V}_t^q)])])}} \frac{1}{1 + e^{\lambda(\mathcal{V}_t^q - k)}} d\nu(k).$$

Proof. Let us use Proposition 3.2.1 and Proposition 3.2.2 and for a given k , take $\varphi(x) = [x - p_k]_+^\lambda$ and $\zeta(x) = [k - x]_+^\lambda$ to establish that for a given k ,

$$h_k(\mu_{\mathcal{V}_t^q}) = \left[\mathbb{E} [k - \mathcal{V}_t^q]_+^\lambda - p_k \right]_+^\lambda = \varphi \left(\mathbb{E} [\zeta(\mathcal{V}_t^q)] \right)$$

is L-convex and the L-derivative at \mathcal{V}_t^q is given by,

$$\partial_\mu h_k(\mu)(\mathcal{V}_t^q) = \frac{-1}{1 + e^{\lambda p_k - (\mathbb{E} [\log(1 + e^{\lambda(k - \mathcal{V}_t^q)])])}} \frac{1}{1 + e^{\lambda(\mathcal{V}_t^q - k)}}.$$

Next, we want to prove that

$$h(\mu_{\mathcal{V}_t^q}) = \int_{\mathbb{R}^+} h_k(\mu_{\mathcal{V}_t^q}) d\nu(k) \text{ is L-differentiable and L-convex.}$$

So, essentially the problem boils down to proving that we can take the L-derivative inside the integral over k . We use dominated convergence theorem to establish this result. Firstly, note that the integrand is bounded by 1, i.e.

$$|\partial_\mu h_k(\mu)(\mathcal{V}_t^q)| \leq 1, \quad \forall k \in [0, \infty], \quad dt \otimes d\mathbb{P} \text{ a.e.}$$

Hence, we have,

$$\int_{\mathbb{R}^+} |\partial_\mu h_k(\mu)(\mathcal{V}_t^q)|^2 d\nu(k) \leq 1.$$

To prove that the L-derivative of $h(\mu)$ exists and is given by (3.2.3), we need to prove the following for $\eta \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$ with $\|\eta\|_{L^2} \leq 1$,

$$\lim_{\|\eta\|_{L^2} \rightarrow 0} \frac{\int_{\mathbb{R}^+} \left(h_k(\mathcal{V}_t^q + \eta) - h_k(\mathcal{V}_t^q) - \mathbb{E} [\partial_\mu h_k(\mu)(\mathcal{V}_t^q) \cdot (\eta)] \right) d\nu(k)}{\|\eta\|_{L^2}} = 0.$$

We would need to take the following limit over η inside the integral and use the result that h_k is L-differentiable. For this we first need to dominate the integrand by something k -integrable uniformly in η ,

$$\frac{|h_k(\mathcal{V}_t^q + \eta) - h_k(\mathcal{V}_t^q) - \mathbb{E} [\partial_\mu h_k(\mu)(\mathcal{V}_t^q) \cdot (\eta)]|}{\|\eta\|_{L^2}} \leq \frac{\left| h_k(\mathcal{V}_t^q + \eta) - h_k(\mathcal{V}_t^q) \right| + \left| \mathbb{E} [\partial_\mu h_k(\mu)(\mathcal{V}_t^q) \cdot (\eta)] \right|}{\|\eta\|_{L^2}}$$

The second term in the numerator is bounded by $\|\eta\|_{L^2}$ after using Cauchy-Schwartz for the scalar product and then the above estimate for the bound on $\partial_\mu h_k$. For the first term, note that the derivative of the function $[k - x]_+^\lambda$ (see equation (3.2.2)) is bounded by 1, hence it is 1-Lipschitz implying that h_k is Lipschitz because,

$$|h_k(\mathcal{V}_t^q + \eta) - h_k(\mathcal{V}_t^q)| \leq |\mathbb{E} [k - \mathcal{V}_t^q - \eta]_+^\lambda - \mathbb{E} [k - \mathcal{V}_t^q]_+^\lambda| \leq \mathbb{E} |\eta| \leq (\mathbb{E} [|\eta|^2])^{1/2} \leq \|\eta\|_{L^2}.$$

Hence, we have a dominating function whose integral is bounded. So, we can take the limit inside the integral over k and as we have established, h_k is L-differentiable, hence the above limit is zero. Therefore, h is L-differentiable. Next, the L-convexity for $h(\mu)$ follows from the L-convexity of $h_k(\mu)$. Take $X \sim \mu$ and $X' \sim \mu'$,

$$\begin{aligned} & h_k(\mu') - h_k(\mu) - \mathbb{E} [\partial_\mu h_k(\mu)(X) \cdot (X' - X)] \geq 0 \tag{3.2.3} \\ \implies & \int_{\mathbb{R}^+} \left(h_k(\mu') - h_k(\mu) - \mathbb{E} [\partial_\mu h_k(\mu)(X) \cdot (X' - X)] \right) d\nu(k) \geq 0 \\ \implies & h(\mu') - h(\mu) - \mathbb{E} [\partial_\mu h(\mu)(X) \cdot (X' - X)] \geq 0, \end{aligned}$$

proving the L-convexity of h . Hence, h is L-differentiable and L-convex. □

Proposition 3.2.4. *The L-derivative of h satisfies the following Lipschitz condition with Lipschitz constant $L_h = 4\lambda^2$,*

$$\mathbb{E} [|\partial_\mu h(\mu')(X') - \partial_\mu h(\mu)(X)|^2] \leq L_h \mathbb{E} [|X' - X|^2]$$

for any $\mu, \mu' \in \mathcal{P}^2(\mathbb{R})$ and $X \sim \mu, X' \sim \mu'$.

Proof. Let us first note that the function,

$$\psi_1(x) = \frac{1}{1 + e^{-\lambda(x-p_k)}} \text{ and } \psi_2(x) = \frac{-1}{1 + e^{-\lambda(k-x)}},$$

are both λ -Lipschitz continuous and both are bounded by 1. Also, $\partial_\mu h_k(\mu)(X) = \psi_1(\mathbb{E}[\zeta_k(X)])\psi_2(X)$, where $\zeta_k(x) = \frac{1}{\lambda} \log(1 + e^{\lambda(k-x)})$ which is 1-Lipschitz. So, we get,

$$\begin{aligned} & |\partial_\mu h_k(\mu')(X') - \partial_\mu h_k(\mu)(X)| \\ &= \left| \psi_1(\mathbb{E}[\zeta_k(X')])\psi_2(X') - \psi_1(\mathbb{E}[\zeta_k(X)])\psi_2(X) \right| \\ &= \left| \psi_1(\mathbb{E}[\zeta_k(X')])\psi_2(X') - \psi_1(\mathbb{E}[\zeta_k(X')])\psi_2(X) + \psi_1(\mathbb{E}[\zeta_k(X')])\psi_2(X) - \psi_1(\mathbb{E}[\zeta_k(X)])\psi_2(X) \right| \\ &\leq \lambda |\psi_1(\mathbb{E}[\zeta_k(X')])| |X' - X| + \lambda \mathbb{E} [|X - X'|] |\psi_2(X)| \\ &\leq \lambda (|X' - X| + \mathbb{E} [|X' - X|]). \end{aligned}$$

Taking square on both sides and then expectation and using Jensen's inequality for the term with the square of expectation, we get,

$$\mathbb{E} [|\partial_\mu h_k(\mu')(X') - \partial_\mu h_k(\mu)(X)|^2] \leq 4\lambda^2 \mathbb{E} [|X' - X|^2].$$

Next, notice that,

$$\begin{aligned} \mathbb{E} [|\partial_\mu h(\mu')(X') - \partial_\mu h(\mu)(X)|^2] &\leq \int_{\mathbb{R}^+} \mathbb{E} [|\partial_\mu h_k(\mu') - \partial_\mu h_k(\mu)|^2] d\nu(k) \\ &\leq 4\lambda^2 \mathbb{E} [|X' - X|^2] \end{aligned}$$

Therefore, it is L_h -Lipschitz continuous with $L_h = 4\lambda^2$. □

3.2.3 The control problem and assumption

We have $[\mathcal{V}_t^q \ \delta_t]^\top$ as the $(N+1)$ -dimensional state process which satisfies the linear SDE from (3.2.1). For convenience of notation, let us re-write the state dynamics as follows,

$$\begin{aligned} dX_t^q &= (A_t^x X_t^q + A_t^q q_t) dt + \sum_{j=1}^m (B_t^{x,j} X_t^q) dW_t^j \\ X_0^q &= [\mathcal{V}_0^q \ 0]^\top \end{aligned}$$

where the matrices A^x and $B^{x,j}$ valued in $\mathbb{R}^{(N+1) \times (N+1)}$ and A^q valued in $\mathbb{R}^{(N+1) \times N}$ are stochastic processes but with bounded norm uniformly in time and are given as follows:

$$A_t^x = \begin{bmatrix} \alpha_t^\mathcal{V} & \alpha_t^\delta & \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}, \quad A_t^q = \begin{bmatrix} 0 & \cdots & 0 \\ 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}, \quad B_t^{x,j} = \begin{bmatrix} \beta_t^{\mathcal{V},j} & \beta_t^{\delta,j} & \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}.$$

The cost functional for the control problem is,

$$J^\lambda(q) = \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(g(X_t^q, q_t) + \int_0^\infty \left[\mathbb{E} [k - \mathcal{V}_t^{q\lambda} - p_k]_+^\lambda d\nu(k) \right) dt \right].$$

Next, we present the assumptions we need on the control space, state parameters and cost-discount factor ρ for finding the necessary and sufficient conditions for optimality for this infinite horizon McKean-Vlasov stochastic control problem. We refer to chapter 2 for the definitions of the set of admissible controls and the spaces $\mathcal{S}_\rho^2(\mathbb{R}^k)$ and $L_\rho^2(\mathbb{R}^k)$.

Next, let us consider the following assumptions on the system dynamics and the running cost:

Assumptions

(H0). For the distribution ν ,

1. $\int_0^\infty d\nu(k) = 1$
2. $\int_0^\infty k^2 d\nu(k) < \infty$
3. $\int_0^\infty |p_k|^2 d\nu(k) < \infty$.

(H1). The stochastic system parameters, b_t^i , r_t and $\sigma_t^{i,j}$ are bounded uniformly in time, i.e.

$$|b_t^i| \leq K_b; |r_t| \leq K_r; |\sigma_t^{i,j}| \leq K_\sigma, dt \otimes d\mathbb{P} \text{ a.e.}$$

(H2). The discount factor ρ satisfies,

$$\rho > \sup_{t \geq 0} \left(3 \|\sigma_t\|_{L^{1,2}}^2 + 2 \sum_{i=1}^N |b_t^i - r_t| \mathbf{1}_{(b_t^i - r_t) < 0} \right), \mathbb{P} \text{ a.s.}$$

where the matrix norm $L^{1,2}$ is defined as, $\left(\sum_{j=1}^m \left(\sum_{i=1}^N |\sigma_t^{i,j}| \right)^2 \right)^{\frac{1}{2}}$. We will denote $\sup_{t \geq 0} \left(3 \|\sigma_t\|_{L^{1,2}}^2 + 2 \sum_{i=1}^N |b_t^i - r_t| \mathbf{1}_{(b_t^i - r_t) < 0} \right)$ by \bar{K} .

(H3). $g(x, q)$ is continuously differentiable with the first derivatives being Lipschitz continuous.

3.3 Conditions for optimality

3.3.1 A priori estimate on state process

In this subsection, we establish L^2 estimates for the state and the corresponding variation process:

Proposition 3.3.1. *When the assumptions (H0), (H1), (H2) and (H3) hold, the state process satisfies the following L^2 estimate: $(X_t^q)_{t \geq 0} \in L_\rho^2(\mathbb{R}^{N+1})$, i.e.*

$$\mathbb{E} \left[\int_0^\infty e^{-\rho t} |X_t^q|^2 dt \right] \leq C \left(|X_0|^2 + \mathbb{E} \left[\int_0^\infty e^{-\rho t} (|q_t|^2) dt \right] \right) < \infty.$$

Proof. Note that $|X_t^q|^2 = |\mathcal{V}_t^q|^2 + |\delta_t|^2$. Next, using the Itô's formula for $e^{-\rho t} |\delta_t|^2$, we get,

$$e^{-\rho T} |\delta_T|^2 = |\delta_0|^2 + \int_0^T e^{-\rho t} (-\rho |\delta_t|^2 + 2\delta_t q_t) dt$$

using Young's inequality, we get, $\forall \gamma > 0$

$$e^{-\rho T} |\delta_T|^2 \leq |\delta_0|^2 + \int_0^T e^{-\rho t} \left(-\rho |\delta_t|^2 + \gamma |\delta_t|^2 + \frac{1}{\gamma} |q_t|^2 \right) dt,$$

taking $0 < \gamma < \rho$ and denoting $\kappa = \rho - \gamma > 0$,

$$\mathbb{E} [e^{-\rho T} |\delta_T|^2] + \mathbb{E} \left[\int_0^T e^{-\rho t} (\kappa |\delta_t|^2) dt \right] \leq |\delta_0|^2 + \frac{1}{\gamma} \mathbb{E} \left[\int_0^T e^{-\rho t} (|q_t|^2) dt \right].$$

Next, doing the same for \mathcal{V}_t^q , we get,

$$\begin{aligned} e^{-\rho T} |\mathcal{V}_T^q|^2 &= |\mathcal{V}_0^q|^2 + \int_0^T e^{-\rho t} \left(-\rho |\mathcal{V}_t^q|^2 + 2\mathcal{V}_t^q (\alpha_t^\mathcal{V} \mathcal{V}_t^q + \alpha_t^\delta \delta_t) + \sum_{j=1}^m \left| \beta_t^{\mathcal{V},j} \mathcal{V}_t^q + \beta_t^{\delta,j} \delta_t \right|^2 \right) dt + \mathcal{M}_T \\ &\leq |\mathcal{V}_0^q|^2 + \left[\int_0^T e^{-\rho t} \left(-\rho |\mathcal{V}_t^q|^2 + 2\alpha_t^\mathcal{V} |\mathcal{V}_t^q|^2 + \gamma |\mathcal{V}_t^q|^2 + \frac{1}{\gamma} |\alpha_t^\delta \delta_t|^2 + \sum_{j=1}^m \left(|\beta_t^{\mathcal{V},j}|^2 |\mathcal{V}_t^q|^2 \right. \right. \right. \\ &\quad \left. \left. \left. + |\beta_t^{\delta,j} \delta_t|^2 + \gamma |\mathcal{V}_t^q|^2 + \frac{1}{\gamma} |\beta_t^{\mathcal{V},j} \beta_t^{\delta,j} \delta_t|^2 \right) dt \right] + \mathcal{M}_T \\ &\leq |\mathcal{V}_0^q|^2 + \left[\int_0^T e^{-\rho t} \left(\left(-\rho + 2\alpha_t^\mathcal{V} + \sum_{j=1}^m |\beta_t^{\mathcal{V},j}|^2 \right) |\mathcal{V}_t^q|^2 + (1+m)\gamma |\mathcal{V}_t^q|^2 + \frac{1}{\gamma} |\alpha_t^\delta \delta_t|^2 \right. \right. \\ &\quad \left. \left. + \sum_{j=1}^m \left(|\beta_t^{\delta,j} \delta_t|^2 + \frac{1}{\gamma} |\beta_t^{\mathcal{V},j} \beta_t^{\delta,j} \delta_t|^2 \right) dt \right] + \mathcal{M}_T, \end{aligned}$$

where we have used Young's inequality for the cross-terms as before and \mathcal{M}_t is a local martingale term. As usual, we can use a localization argument by taking a sequence of stopping time $(\tau_n)_{n \geq 0}$ such that $\mathcal{M}_{\tau_n \wedge T}$ is a martingale and $\tau_n \rightarrow T$ as $n \rightarrow \infty$. Using the assumption (H2),

$$-\rho + 2\alpha_t^\mathcal{V} + \sum_{j=1}^m |\beta_t^{\mathcal{V},j}|^2 \leq -\rho + 3 \|\sigma_t\|_{L^{1,2}}^2 - 2 \sum_{i=1}^N (b_t^i - r_t) \mathbf{1}_{(b_t^i - r_t) < 0} < 0,$$

and as done before, we can take γ arbitrarily small and satisfy $-\xi := -\rho + 2\alpha_t^\mathcal{V} + \sum_{j=1}^m |\beta_t^{\mathcal{V},j}|^2 + (1+m)\gamma < 0$. As a result, we get,

$$\begin{aligned} \mathbb{E} [e^{-\rho T \wedge \tau_n} |\mathcal{V}_T^q|^2] + \xi \mathbb{E} \left[\int_0^{T \wedge \tau_n} |\mathcal{V}_t^q|^2 dt \right] &\leq C \left(|\mathcal{V}_0^q|^2 + \mathbb{E} \left[\int_0^{T \wedge \tau_n} e^{-\rho t} |q_t|^2 dt \right] \right) \\ &\leq C \left(|\mathcal{V}_0^q|^2 + \mathbb{E} \left[\int_0^\infty e^{-\rho t} |q_t|^2 dt \right] \right) \end{aligned}$$

where we have used the above L_ρ^2 estimate on δ . This concludes the proof after taking $n \rightarrow \infty$ and $T \rightarrow \infty$. We can do this because C is independent of T and n . Moreover, it only depends on m and γ which are dependent only on the parameters in our assumptions. \square

The following L^2 estimate for the difference of two state processes will be useful later:

Corollary 3.3.1.1. *Under the above assumptions, if X and X' are the state processes corresponding to $q, q' \in \mathcal{A}$ then the following L^2 estimate holds: $(X_t^q)_{t \geq 0} \in L_\rho^2(\mathbb{R}^{N+1})$, i.e.*

$$\mathbb{E} \left[\int_0^\infty e^{-\rho t} |X_t' - X_t|^2 dt \right] \leq C \mathbb{E} \left[\int_0^\infty e^{-\rho t} (|q_t' - q_t|^2) dt \right] < \infty.$$

Proof. From linearity of the state SDE, $X' - X$ satisfies the same SDE but with the initial condition $X_0' - X_0 = 0$, hence the previous proposition leads to the desired result. \square

3.3.2 Adjoint processes

Let us define the Hamiltonian of the problem as,

$$H(\eta) := H(t, x, q, \mu, y, z) = y (A_t^x x + A_t^q q) + \sum_{j=1}^m z^j B_t^{x,j} x + f(x, q, \mu) - \rho y x,$$

where $(y, z) \in \mathbb{R}^{1 \times (N+1)} \times \mathbb{R}^{m \times (N+1)}$. Let us define a vector $\mathcal{D}_\mu h_t$ valued in $\mathbb{R}^{1 \times (N+1)}$ as,

$$\mathcal{D}_\mu h_t = \left[\partial_\mu h(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) \quad 0 \quad \cdots \quad 0 \right].$$

Then, we define the adjoint processes of X to be the pair of processes $(Y, Z) \in \mathcal{S}^2(\mathbb{R}^{1 \times (N+1)}) \times \Lambda^2(\mathbb{R}^{m \times (N+1)})$ which satisfy for every $T > 0$ the following BSDE $dt \otimes d\mathbb{P}$ -a.e.,

$$\begin{aligned} Y_t - Y_T &= \int_t^T (\partial_x H(X_t^q, q_t, \mu_{\mathcal{V}_t^q}, Y_t, Z_t) + \mathcal{D}_\mu h_t) dt - \sum_{j=1}^m \int_t^T Z_t^j dW_t^j \\ &= \int_t^T \left(Y_t A_t^x + \sum_{j=1}^m Z_t^j B_t^{x,j} + \partial_x g(X_t^q, q_t) + \mathcal{D}_\mu h_t - \rho Y_t \right) dt - \sum_{j=1}^m \int_t^T Z_t^j dW_t^j, \end{aligned} \quad (3.3.1)$$

with $\lim_{T \rightarrow \infty} \mathbb{E} [e^{-\rho T} |Y_T|^2] = 0$ where Z_t^j is the j -th row of Z_t and the first three terms of the generator correspond to the space derivative of the Hamiltonian and the fourth term corresponds to the L-derivative. The *generator* of the BSDE, denoted by $\Phi(t, y, z)$ is given as follows:

$$\Phi(t, y, z) = y(A_t^x - \rho I) + \sum_{j=1}^m z^j B_t^{x,j} + \partial_x g(X_t^q, q_t) + \mathcal{D}_\mu h_t.$$

The following result establishes the existence and uniqueness of the solution of the BSDE for a given admissible control.

Proposition 3.3.2. *For a given admissible control $(q_t)_{t \geq 0}$, under the assumptions (H1), (H2) and (H3), the above BSDE (3.3.1) admits a unique solution $(Y_t, Z_t)_{t \geq 0}$ in the space $\mathcal{S}_\rho^2(\mathbb{R}^{1 \times (N+1)}) \times L_\rho^2(\mathbb{R}^{m \times (N+1)})$. Moreover, there exists $C > 0$ such that for two pairs (Y', Z') and (Y, Z) of solutions to BSDEs corresponding to state-control pairs (X', q') and (X, q) respectively, the following estimate holds:*

$$\mathbb{E} \left[\int_0^\infty e^{-\rho t} |Y'_t - Y_t|^2 dt \right] + \mathbb{E} \left[\int_0^\infty e^{-\rho t} |Z'_t - Z_t|^2 dt \right] \leq C \mathbb{E} \left[\int_0^\infty e^{-\rho t} |q'_t - q_t|^2 dt \right].$$

See Proposition 2.2.2 from Chapter 2 for a proof of this result based on a result from Pardoux and Răşcanu [100]. From now on, to shorten the notation wherever it causes no confusion, we will use the following notation,

$$g(\theta_t) = g(X_t^q, q_t) \quad \text{and} \quad f(\theta_t) = f(X_t^q, q_t, \mu_{\mathcal{V}_t^q}) = g(X_t^q, q_t) + h(\mu_{\mathcal{V}_t^q}),$$

with an understanding that $g(\theta_t)$ does not depend on $\mu_{\mathcal{V}_t^q}$.

3.3.3 Necessary and sufficient conditions for optimality

Let us now present some necessary conditions that an optimal control must satisfy.

Theorem 3.3.3 (Necessary conditions). *Suppose that assumptions (H0), (H1), (H2) and (H3) hold and $q \in \mathcal{A}$ is optimal. Let X be the corresponding state variable and let (Y, Z) be the corresponding adjoint processes as defined above in (3.3.1), then we must have*

$$\partial_q H(X_t^q, q_t, \mu_{\mathcal{V}_t^q}, Y_t, Z_t) = 0, \quad dt \otimes d\mathbb{P} \text{ a.e}$$

The proof of this theorem follows from Theorem 2.3.1 but we prove it specifically for the setting of this chapter in Section 3.5.

Remark 3.3.1. As an example, let us consider $g(x, q) = \frac{\gamma}{2}|q|^2$ for some $\gamma > 0$. The above condition gives the optimal control in terms of the adjoint equation as,

$$q_t = -\frac{1}{\gamma}(Y_t A_t^q)^\top = -\frac{1}{\gamma}(Y_t^{[2:(N+1)]})^\top \in \mathbb{R}^N.$$

Hence, the stochastic maximum principle helps us to convert the problem into a point-wise maximization problem, which gives us the explicit form of the optimal trading speed.

Next, we prove a sufficient condition for optimality for our control problem.

Theorem 3.3.4. *Suppose (H0), (H1), (H2) and (H3) hold. Let $q \in \mathcal{A}$ be an admissible control, X_t^q the corresponding state process and (Y, Z) the corresponding adjoint processes. Also, assume the Hamiltonian satisfies:*

$$\mathbb{R}^{N+1} \times \mathbb{R}^N \times \mathcal{P}^2(\mathbb{R}) \ni (x, q, \mu) \mapsto H(t, x, q, \mu, Y_t, Z_t) \text{ is } \alpha\text{-strongly convex } dt \otimes d\mathbb{P} \text{ a.e.,}$$

which is to be understood as follows,

$$\begin{aligned} H(t, x', q', \mu', Y_t, Z_t) &\geq H(t, x, q, \mu, Y_t, Z_t) + \partial_x H(t, x, q, \mu, Y_t, Z_t)(x' - x) \\ &\quad + \partial_q H(t, x, q, \mu, Y_t, Z_t)(q' - q) + \mathbb{E} [\partial_\mu h(\mu)(X)(X' - X)] + \alpha|q' - q|^2 \end{aligned} \quad (3.3.2)$$

where $X \sim \mu$ and $X' \sim \mu'$ and $\alpha > 0$. If

$$H(t, X_t^q, q_t, \mu_{X_t^q}, Y_t, Z_t) = \inf_{u \in \mathbb{R}^N} H(t, X_t^q, u, \mu_{X_t^q}, Y_t, Z_t), \quad dt \otimes d\mathbb{P} \text{ a.e.} \quad (3.3.3)$$

then q is the unique optimal control in \mathcal{A} .

The proof of this result follows from Theorem 2.4.1 but we prove it specifically for the setting of this chapter in Section 3.5. Let us consider the special case when $g(x, q) = \frac{\gamma}{2}|q|^2$. Then, note that the Hamiltonian satisfies the above strong convexity condition with $\alpha = \frac{\gamma}{2}$. Hence, we have the following result:

Corollary 3.3.4.1. *For $g(x, q) = \frac{\gamma}{2}|q|^2$ with $\gamma > 0$, the Hamiltonian is $\frac{\gamma}{2}$ -strongly convex in the above sense and for a given $q \in \mathcal{A}$, if $\partial_q H(t, X_t^q, q_t, \mu_{X_t^q}, Y_t, Z_t) = 0$ with (Y, Z) being the corresponding solution of the BSDE, then q must be the unique optimal control in \mathcal{A} .*

3.3.4 Characterisation of optimal control using FBSDE

For the rest of the chapter, we will fix $g(x, q) = \frac{1}{2}|q|^2$. From the Remark 3.3.1 obtained from necessary conditions for $g(x, q) = \frac{1}{2}|q|^2$, we can plug the optimal control in terms of Y in place of the q in the both the state dynamics and the adjoint equation to obtain a coupled system of forward-backward SDEs, written as follows,

$$\begin{aligned} dX_t^q &= \left(A_t^x X_t^q - \begin{bmatrix} 0 & Y_t^{[2:(N+1)]} \end{bmatrix}^\top \right) dt + \sum_{j=1}^m (B_t^{x,j} X_t^q) dW_t^j, \quad X_0^q = [\mathcal{V}_0^q \quad 0]^\top, \quad (3.3.4) \\ Y_t - Y_T &= \int_t^T \left(Y_t (A_t^x - \rho I) + \sum_{j=1}^m Z_t^j B_t^{x,j} + \mathcal{D}_\mu h_t \right) dt - \sum_{j=1}^m \int_t^T Z_t^j dW_t^j, \end{aligned}$$

where I denotes the unit matrix in $\mathbb{R}^{(N+1) \times (N+1)}$. Hence, we obtain a linear mean-field FBSDE in infinite horizon.

We use method of continuation for FBSDEs similar to [36] whereby, we will introduce a space of inputs and define a norm on an input as,

$$\mathbb{I} := \left\{ (\mathcal{I}^b, \mathcal{I}^\sigma, \mathcal{I}^f) \mid \mathbb{E} \left[\int_0^\infty e^{-\rho t} |\mathcal{I}_t^{(b,\sigma,f)}|^2 dt \right] < \infty \right\}, \quad \|\mathcal{I}\|_{\mathbb{I}} := \mathbb{E} \left[\int_0^\infty e^{-\rho t} |\mathcal{I}_t|^2 dt \right]^{1/2},$$

where the progressively measurable processes \mathcal{I}^b , $\mathcal{I}^{\sigma,j}$ and \mathcal{I}^f valued in \mathbb{R}^{N+1} , \mathbb{R}^{N+1} and $\mathbb{R}^{1 \times (N+1)}$ are inputs given to the state dynamics (drift and volatility respectively) and the generator the BSDE respectively. We denote the triple $(\mathcal{I}^b, \mathcal{I}^{\sigma,j}, \mathcal{I}^f)$ by \mathcal{I} . The inputs are given to FBSDE in the following manner:

$$\begin{aligned} dX_t^q &= \left(\zeta (A_t^x X_t^q + A_t^q q_t) + \mathcal{I}_t^b \right) dt + \sum_{j=1}^m \left(\zeta B_t^{x,j} X_t^q + \mathcal{I}_t^{\sigma,j} \right) dW_t^j, \quad X_0^q = [\mathcal{V}_0^q \ 0]^\top \quad (3.3.5) \\ Y_t - Y_T &= \int_t^T \left(\zeta \left(Y_t A_t^x + \sum_{j=1}^m Z_t^j B_t^{x,j} + \mathcal{D}_\mu h_t \right) - \rho Y_t + \mathcal{I}_t^f \right) dt - \sum_{j=1}^m \int_t^T Z_t^j dW_t^j, \end{aligned}$$

where the control q satisfies the optimality conditions, letting us write $q_t \equiv \tilde{q}(Y_t) = -Y_t^{[2:N+1]\top}$ and hence causing the coupling between the forward and backward SDEs. The parameter $\zeta \in [0, 1]$ serves as a coupling parameter. The usefulness of introducing these equations with inputs \mathcal{I} is illustrated in the proof of Theorem 3.3.6, where we can go from the solvability of $\Pi(\zeta, \mathcal{I}')$ to the solvability of $\Pi(\zeta + \eta, \mathcal{I})$ for a small η and any \mathcal{I} , given a good choice of inputs \mathcal{I}' , which is the essence of the proof by using continuation argument. For $\zeta = 1$, we have the fully coupled FBSDE (3.3.4) for which we want to establish the existence of a unique solution, whereas, for $\zeta = 0$, we have an uncoupled SDE and BSDE which trivially have unique solution. For sake of convenience, we introduce the following notation:

Definition 3.3.1. (i) A solution (X, Y, Z) for the FBSDE (3.3.5) for a given ζ and $\mathcal{I} \equiv (\mathcal{I}^b, \mathcal{I}^{\sigma,j}, \mathcal{I}^f)$ will be denoted by $\Pi(\zeta, \mathcal{I})$. We define the space of solutions $(X, Y, Z) \in L_\rho^2(\mathbb{R}^{N+1}) \times L_\rho^2(\mathbb{R}^{1 \times (N+1)}) \times L_\rho^2(\mathbb{R}^{m \times (N+1)})$ by $\mathcal{M}_\rho^2(0, \infty)$ and a norm on it as follows:

$$\|\Pi(\zeta, \mathcal{I})\|_{L_\rho^2} := \mathbb{E} \left[\int_0^\infty e^{-\rho t} (|X_t|^2 + |Y_t|^2 + |Z_t|^2) dt \right]^{1/2}$$

(ii) For a given $\zeta \in [0, 1]$, we say that $(\mathcal{E}\mathcal{U}_\zeta)$ holds when the corresponding FBSDE $\Pi(\zeta, \mathcal{I})$ has a unique solution for any input $\mathcal{I} \in \mathbb{I}$.

Before proceeding to the main result, let us remark the definition of the Hamiltonian we will use in the proof is the same as the original Hamiltonian, which for $g(x, q) = \frac{1}{2}|q|^2$ translates to, for any $X \sim \mu$,

$$H(t, x, y, z, q, \mu) = y(A_t^x x + A_t^q q) + \sum_{j=1}^m z^j B_t^{x,j} x + \frac{1}{2}|q|^2 + \partial_\mu h(\mu)(X) - \rho y x,$$

which is $\frac{1}{2}$ -strongly convex and hence has a unique minimizer,

$$\arg \min_{q \in \mathbb{R}^N} H(t, x, y, z, q, \mu) = -(y A_t^q)^\top = -y^{[2:(N+1)]\top} \equiv \tilde{q}(y), \quad (3.3.6)$$

implying that,

$$\partial_q H(t, x, y, z, \tilde{q}(y), \mu) = 0. \quad (3.3.7)$$

The following a priori stability result for the solution of the FBSDE $\Pi(\zeta, \mathcal{I})$ will be useful for the main result:

Lemma 3.3.5 (A priori stability result). *Let $\zeta \in [0, 1]$ such that (\mathcal{EU}_ζ) holds. Then, there exists $C > 0$ independent of ζ such that for any inputs $\mathcal{I}, \mathcal{I}' \in \mathbb{I}$, the corresponding solutions of FBSDEs $\Pi(\zeta, \mathcal{I})$ and $\Pi(\zeta, \mathcal{I}')$ satisfy the following:*

$$\|\Pi(\zeta, \mathcal{I}') - \Pi(\zeta, \mathcal{I})\|_{L^2_p} \leq C \|\mathcal{I} - \mathcal{I}'\|_{\mathbb{I}}$$

Proof. Let $\mathcal{I}, \mathcal{I}' \in \mathbb{I}$ and let $\Pi(\zeta, \mathcal{I}), \Pi(\zeta, \mathcal{I}')$ be corresponding solutions of the FBSDE and we denote the corresponding controls by $q_t \equiv \tilde{q}(Y_t)$ and $q'_t \equiv \tilde{q}(Y'_t)$. Following an approach similar to Theorem (3.3.4) and using Itô's formula, we get for any $T > 0$,

$$\begin{aligned} & \mathbb{E} \left[e^{-\rho T} Y_T (X'_T - X_T) \right] \\ &= \mathbb{E} \left[\int_0^T e^{-\rho t} \left((-\rho Y_t - \zeta [Y_t A_t^x + \sum_{j=1}^m Z_t^j B_t^{x,j} + \mathcal{D}_\mu h_t] + \rho Y_t) (X'_t - X_t^q) \right. \right. \\ & \quad \left. \left. + \zeta Y_t (A_t^x (X'_t - X_t^q) + \zeta A_t^q (q'_t - q_t)) + \zeta \sum_{j=1}^m Z_t^j B_t^{x,j} (X'_t - X_t^q) \right) dt \right] \\ &+ \mathbb{E} \left[\int_0^T e^{-\rho t} \left(Y_t (\mathcal{I}_t^{',b} - \mathcal{I}_t^b) + \sum_{j=1}^m Z_t^j (\mathcal{I}_t^{',\sigma,j} - \mathcal{I}_t^{\sigma,j}) - \mathcal{I}_t^f (X'_t - X_t) \right) dt \right] \\ &= \mathbb{E} \left[\int_0^T e^{-\rho t} \left((-\zeta [\partial_x H(\theta_t) + \rho Y_t + \mathcal{D}_\mu h_t]) (X'_t - X_t^q) \right. \right. \\ & \quad \left. \left. + \zeta Y_t (A_t^x (X'_t - X_t^q) + \zeta A_t^q (q'_t - q_t)) + \zeta \sum_{j=1}^m Z_t^j B_t^{x,j} (X'_t - X_t^q) \right) dt \right] \\ &+ \mathbb{E} \left[\int_0^T e^{-\rho t} \left(Y_t (\mathcal{I}_t^{',b} - \mathcal{I}_t^b) + \sum_{j=1}^m Z_t^j (\mathcal{I}_t^{',\sigma,j} - \mathcal{I}_t^{\sigma,j}) - \mathcal{I}_t^f (X'_t - X_t) \right) dt \right]. \end{aligned}$$

Next, let us consider the difference between the corresponding cost functionals, which can be written as follows:

$$\begin{aligned} & \zeta (J(q') - J(q)) \\ &= \zeta (J_T(q') - J_T(q)) + \mathbb{E} \left[\int_0^T e^{-\rho t} \left(\zeta (H(\theta'_t) - H(\theta_t)) - \zeta (Y_t (A_t^x (X'_t - X_t^q) \right. \right. \\ & \quad \left. \left. + A_t^q (q'_t - q_t)) + \sum_{j=1}^m Z_t^j B_t^{x,j} (X'_t - X_t^q) \right) + \zeta \rho Y_t (X'_t - X_t^q) \right) dt \right] \\ &= \zeta (J_T(q') - J_T(q)) + \mathbb{E} \left[\int_0^T e^{-\rho t} \left(\zeta (H(\theta'_t) - H(\theta_t)) - \zeta [\partial_x H(\theta_t) + \mathcal{D}_\mu h_t + \rho Y_t] (X'_t - X_t^q) \right. \right. \\ & \quad \left. \left. + \zeta \rho Y_t (X'_t - X_t^q) + Y_t (\mathcal{I}_t^{',b} - \mathcal{I}_t^b) + \sum_{j=1}^m Z_t^j (\mathcal{I}_t^{',\sigma,j} - \mathcal{I}_t^{\sigma,j}) - \mathcal{I}_t^f (X'_t - X_t^q) \right) dt \right] \\ &- \mathbb{E} \left[e^{-\rho T} Y_T (X'_T - X_T) \right] \\ &= \zeta (J_T(q') - J_T(q)) + \mathbb{E} \left[\int_0^T e^{-\rho t} \left(\zeta (H(\theta'_t) - H(\theta_t)) - \zeta [\partial_x H(\theta_t) + \mathcal{D}_\mu h_t] (X'_t - X_t^q) \right. \right. \\ & \quad \left. \left. + Y_t (\mathcal{I}_t^{',b} - \mathcal{I}_t^b) + \sum_{j=1}^m Z_t^j (\mathcal{I}_t^{',\sigma,j} - \mathcal{I}_t^{\sigma,j}) - \mathcal{I}_t^f (X'_t - X_t^q) \right) dt \right] - \mathbb{E} \left[e^{-\rho T} Y_T (X'_T - X_T) \right]. \end{aligned}$$

where $\theta'_t = (t, X'_t, Y_t, Z_t, \mu_{\mathcal{V}_t}, q'_t)$ and $\theta_t = (t, X_t, Y_t, Z_t, \mu_{\mathcal{V}_t}, q_t)$. Next, using the $\frac{1}{2}$ -strong convexity of the Hamiltonian (3.3.6) along with (3.3.7) gives,

$$H(\theta'_t) - H(\theta_t) - \partial_x H(\theta_t) (X'_t - X_t^q) - \mathbb{E} [\mathcal{D}_\mu h_t \cdot (X'_t - X_t^q)] \geq \frac{1}{2} |q'_t - q_t|^2,$$

implying the following:

$$\begin{aligned} \zeta(J(q') - J(q)) &\geq \zeta(J_T(q') - J_T(q)) + \mathbb{E} \left[\int_0^T e^{-\rho t} \left(Y_t(\mathcal{I}_t^{',b} - \mathcal{I}_t^b) + \sum_{j=1}^m Z_t^j(\mathcal{I}_t^{',\sigma,j} - \mathcal{I}_t^{\sigma,j}) \right. \right. \\ &\quad \left. \left. - \mathcal{I}_t^f(X'_t - X_t^q) + \zeta \frac{1}{2} |q'_t - q_t|^2 \right) dt \right] - \mathbb{E} [e^{-\rho T} Y_T(X'_T - X_T)]. \end{aligned}$$

Next, we do the same but interchange the roles of q' and q in the above inequality and add it with the above inequality (note that the arguments of strong convexity and criticality of q' still hold because $q'_t = \tilde{q}(Y'_t)$ and \tilde{q} is the minimizer function for H). We obtain the following:

$$\begin{aligned} 0 &\geq \mathbb{E} \left[\int_0^T e^{-\rho t} \left(-(Y'_t - Y_t)(\mathcal{I}_t^{',b} - \mathcal{I}_t^b) - \sum_{j=1}^m (Z_t^{',j} - Z_t^j)(\mathcal{I}_t^{',\sigma,j} - \mathcal{I}_t^{\sigma,j}) \right. \right. \\ &\quad \left. \left. + (\mathcal{I}_t^{',f} - \mathcal{I}_t^f)(X'_t - X_t^q) + \zeta |q'_t - q_t|^2 \right) dt \right] + \mathbb{E} [e^{-\rho T} (Y'_T - Y_T)(X'_T - X_T)], \end{aligned}$$

implying the following,

$$\begin{aligned} \zeta \mathbb{E} \left[\int_0^T e^{-\rho t} |q'_t - q_t|^2 dt \right] &\leq \mathbb{E} \left[\int_0^T e^{-\rho t} \left((Y'_t - Y_t)(\mathcal{I}_t^{',b} - \mathcal{I}_t^b) + \sum_{j=1}^m (Z_t^{',j} - Z_t^j)(\mathcal{I}_t^{',\sigma,j} - \mathcal{I}_t^{\sigma,j}) \right. \right. \\ &\quad \left. \left. - (\mathcal{I}_t^{',f} - \mathcal{I}_t^f)(X'_t - X_t^q) \right) dt \right] - \mathbb{E} [e^{-\rho T} (Y'_T - Y_T)(X'_T - X_T)]. \end{aligned}$$

Using Young's inequality for all the cross terms in the integrand and Cauchy-Schwartz for the last term, we get $\forall \varepsilon > 0$ and $T > 0$,

$$\begin{aligned} \zeta \mathbb{E} \left[\int_0^T e^{-\rho t} |q'_t - q_t|^2 dt \right] &\leq \mathbb{E} \left[\int_0^T e^{-\rho t} \left(\varepsilon |Y'_t - Y_t|^2 + \frac{1}{\varepsilon} |\mathcal{I}_t^{',b} - \mathcal{I}_t^b|^2 + \sum_{j=1}^m \left(\varepsilon |Z_t^{',j} - Z_t^j|^2 \right. \right. \right. \\ &\quad \left. \left. + \frac{1}{\varepsilon} |\mathcal{I}_t^{',\sigma,j} - \mathcal{I}_t^{\sigma,j}|^2 \right) + \varepsilon |X'_t - X_t|^2 + \frac{1}{\varepsilon} |\mathcal{I}_t^{',f} - \mathcal{I}_t^f|^2 \right) dt \right] \\ &\quad + \mathbb{E} [e^{-\rho T} |Y'_T - Y_T|^2]^{1/2} \mathbb{E} [e^{-\rho T} |X'_T - X_T|^2]^{1/2}. \end{aligned}$$

Using the terminal condition for the BSDE and the a priori estimate (proposition (3.3.1)) for X which can be proved for $(X' - X)$ as well, implies that the limit of the last term is zero, giving us the following:

$$\begin{aligned} \zeta \|q'_t - q_t\|_{L^2_\rho}^2 &\leq \varepsilon \|Y'_t - Y_t\|_{L^2_\rho}^2 + \frac{1}{\varepsilon} \|\mathcal{I}_t^{',b} - \mathcal{I}_t^b\|_{\mathbb{H}}^2 + \sum_{j=1}^m \left(\varepsilon \|Z_t^{',j} - Z_t^j\|_{L^2_\rho}^2 + \frac{1}{\varepsilon} \|\mathcal{I}_t^{',\sigma,j} - \mathcal{I}_t^{\sigma,j}\|_{\mathbb{H}}^2 \right) \\ &\quad + \varepsilon \|X'_t - X_t\|_{L^2_\rho}^2 + \frac{1}{\varepsilon} \|\mathcal{I}_t^{',f} - \mathcal{I}_t^f\|_{\mathbb{H}}^2 \\ &= \varepsilon \|\Pi(\zeta, \mathcal{I}') - \Pi(\zeta, \mathcal{I})\|_{L^2_\rho}^2 + \frac{1}{\varepsilon} \|\mathcal{I}' - \mathcal{I}\|_{\mathbb{H}}^2. \end{aligned}$$

Next, using the a priori estimate (Proposition 3.3.1) and Proposition 3.3.2, we have,

$$\|\Pi(\zeta, \mathcal{I}') - \Pi(\zeta, \mathcal{I})\|_{L^2_\rho}^2 \leq C \mathbb{E} \left[\int_0^\infty e^{-\rho t} |q_t - q'_t|^2 dt \right].$$

Note that the constant C is dependent only on the parameters from the assumption. Combining the two inequalities from above, we have the following whenever $\zeta > 0$:

$$\|\Pi(\zeta, \mathcal{I}') - \Pi(\zeta, \mathcal{I})\|_{L^2_\rho}^2 \leq \frac{C\varepsilon}{\zeta} \|\Pi(\zeta, \mathcal{I}') - \Pi(\zeta, \mathcal{I})\|_{L^2_\rho}^2 + \frac{C}{\zeta\varepsilon} \|\mathcal{I}' - \mathcal{I}\|_{\mathbb{H}}^2.$$

Choosing $\varepsilon > 0$ small enough such that $1 - \frac{C\varepsilon}{\zeta} > 0$, we have the proof of the lemma,

$$\|\Pi(\zeta, \mathcal{I}') - \Pi(\zeta, \mathcal{I})\|_{L^2_\rho} \leq C\|\mathcal{I}' - \mathcal{I}\|_{\mathbb{I}},$$

where C depends only on ζ and the parameters from the assumptions. □

Let us now prove the existence and uniqueness of the solution to the FBSDE, which is given in Theorem 6.19. from Carmona and Delarue [36] but adapted here for our specific setting for completeness:

Theorem 3.3.6 (Existence and uniqueness). *When the assumptions (H1), (H2) and (H3) hold, the above FBSDE (3.3.4) has a unique solution $(X_t, Y_t, Z_t)_{t \geq 0}$ in the space $L^2_\rho(\mathbb{R}^{N+1}) \times L^2_\rho(\mathbb{R}^{1 \times (N+1)}) \times L^2_\rho(\mathbb{R}^{m \times (N+1)})$.*

Proof. As remarked earlier, $(\mathcal{EU})_0$ holds. We use an induction argument to prove the following claim:

Claim-1: There exists $\delta > 0$ such that if $(\mathcal{EU})_\zeta$ holds for a given $\zeta \in [0, 1)$, then $(\mathcal{EU})_{\zeta+\eta}$ also holds for any $\eta \in (0, \delta]$ satisfying $\eta + \zeta \leq 1$.

Proof of Claim: Let $\zeta \in [0, 1)$ be such that $(\mathcal{EU})_\zeta$ holds. Then, for any given $\tilde{\mathcal{I}} \in \mathbb{I}$, the corresponding FBSDE has a unique solution $\Pi(\zeta, \tilde{\mathcal{I}})$. Let $\mathcal{I} \in \mathbb{I}$. We construct a mapping $\mathcal{T} : \mathcal{M}_\rho^2(0, \infty) \rightarrow \mathcal{M}_\rho^2(0, \infty)$ as follows:

$(X, Y, Z) \equiv \Theta \mapsto \Theta'$ where Θ' is defined as the unique solution to the FBSDE for the given ζ and inputs \mathcal{I}' (which can be verified to be in \mathbb{I}) defined as follows for $\eta > 0$:

$$\begin{aligned} \mathcal{I}'_t{}^b &= \eta(A_t^x X_t + A_t^q q_t) + \mathcal{I}_t^b, \\ \mathcal{I}'_t{}^{\sigma,j} &= \eta B_t^{x,j} X_t + \mathcal{I}_t^{\sigma,j}, \\ \mathcal{I}'_t{}^f &= \eta \left(Y_t A_t^x + \sum_{j=1}^m Z_t^j B_t^{x,j} + \mathcal{D}_\mu h_t \right) + \mathcal{I}_t^f. \end{aligned}$$

As before, $q_t \equiv \tilde{q}(Y_t)$ where the Y_t comes from Θ (which may not necessarily satisfy a FBSDE). The unique solution $\Pi(\zeta, \mathcal{I}')$ exists from the hypothesis. Next, if Θ was a fixed point of \mathcal{T} , this would imply that the FBSDE corresponding to $\eta + \zeta$ for the given input \mathcal{I} has a unique solution (from the nature of the FBSDE and \mathcal{I}'), i.e. $\Pi(\zeta, \mathcal{I}') = \Pi(\zeta + \eta, \mathcal{I})$. Conversely, if Θ is the solution to FBSDE corresponding to $\zeta + \eta$ and \mathcal{I} , i.e. $\Pi(\zeta, \mathcal{I}') = \Pi(\zeta + \eta, \mathcal{I})$, then Θ would be a fixed point of the mapping \mathcal{T} . Hence, if we prove that \mathcal{T} is a contraction for a small enough η , we would prove the claim since the choice of \mathcal{I} was arbitrary. Let $\Theta^1, \Theta^2 \in \mathcal{M}_\rho^2(0, \infty)$ and Θ'^1, Θ'^2 the corresponding images of \mathcal{T} which from definition, are solutions of FBSDE for their respective inputs. From the previous lemma, we have the following,

$$\|\Theta'^1 - \Theta'^2\|_{L^2_\rho} \leq C\|\mathcal{I}'^1 - \mathcal{I}'^2\|_{\mathbb{I}}.$$

Using the linearity of the SDE, linearity of the BSDE in Y, Z and Lipschitz continuity of the L-derivative, we have

$$\|\mathcal{I}'^1 - \mathcal{I}'^2\|_{\mathbb{I}} \leq C\eta\|\Theta^1 - \Theta^2\|_{L^2_\rho}.$$

Combining the two inequalities, we have,

$$\|\Theta'^1 - \Theta'^2\|_{L^2_\rho} \leq C\eta\|\Theta^1 - \Theta^2\|_{L^2_\rho}.$$

Take $\eta < \frac{1}{C}$ (C from the last inequality) proving the claim that \mathcal{T} is a contraction and hence, must have a fixed point, say $\tilde{\Theta}$ since $\mathcal{M}_\rho^2(0, \infty)$ is a Hilbert space, which, as shown above would be the solution of the FBSDE for $\zeta + \eta$ and \mathcal{I} , i.e. $\tilde{\Theta} = \Pi(\zeta + \eta, \mathcal{I})$. Since the choice of \mathcal{I} was arbitrary, $(\mathcal{E}\mathcal{U})_{\zeta+\eta}$ holds.

Therefore, the proof of the theorem follows from induction. \square

Next, we prove a bound on the error for the solution if one truncates the time horizon to T . For this, we first make a useful remark:

Remark 3.3.2. (Truncated FBSDE) The following FBSDE, with solutions denoted by $\Pi^N(\zeta, \mathcal{I})$, satisfies the above a priori stability (Lemma 3.3.5) and existence and uniqueness results (Theorem 3.3.6) for every $N > 0$:

$$\begin{aligned} dX_t^N &= \zeta b(X_t^N, Y_t^N) dt + \zeta \sum_{j=1}^m \sigma^j(X_t^N) dW_t^j, \quad X_0^q = [\mathcal{V}_0^q \ 0]^\top, \\ Y_t^N - Y_T^N &= \int_t^T \left(\zeta f^N(X_t^N, Y_t^N, Z_t^N, \mathcal{D}_\mu h_t^N) + \mathcal{I}_t^f \right) dt - \sum_{j=1}^m \int_t^T Z_t^{j,N} dW_t^j, \\ \lim_{T \rightarrow \infty} \mathbb{E} [e^{-\rho T} |Y_T^N|^2] &= 0, \end{aligned}$$

where $f^N(x, y, z, \mathcal{D}_\mu h_t^N) = \mathbf{1}_{t \leq N} (f(x, y, z, \mathcal{D}_\mu h_t^N) + \rho y) - \rho y$ with f denoting the generator of the infinite horizon FBSDE, $\Pi(1, \mathcal{I})$ and $\mathcal{D}_\mu h_t^N = [\partial_\mu h(\mu_{\mathcal{V}_t^N})(\mathcal{V}_t^N \ 0 \ \dots \ 0)^\top]$. This is because the the new generator f^N satisfies the same assumptions as the original f and the constant C from the a priori stability (Lemma 3.3.5) depends only on the parameters from the assumptions and on parameter ζ . Moreover, $\Pi^N(\zeta, 0)$ when restricted to the time interval $[0, N]$, satisfies the finite horizon FBSDE,

$$\begin{aligned} dX_t^N &= \zeta b(X_t^N, Y_t^N) dt + \zeta \sum_{j=1}^m \sigma^j(X_t^N) dW_t^j, \quad X_0^q = [\mathcal{V}_0^q \ 0]^\top, \\ Y_t^N &= \int_t^N \zeta f^N(X_t^N, Y_t^N, Z_t^N, \mathcal{D}_\mu h_t^N) dt - \sum_{j=1}^m \int_t^N Z_t^{j,N} dW_t^j, \\ Y_N^N &= 0, \end{aligned}$$

with 0 as the terminal condition. Indeed, we can remark that for all $t \geq N$, $f^N(x, y, z, \mathcal{D}_\mu h_t^N) = -\rho y$ which implies that the FBSDE is decoupled on $[N, \infty)$, i.e. we can solve the BSDE and then plug the solution in the SDE in order to solve it. The BSDE becomes

$$Y_t^N = Y_T^N + \int_t^T -\rho Y_s^N ds - \int_t^T Z_s^N dW_s, \quad N \leq t \leq T,$$

which is simple to solve: $Y_t^N = 0$ and $Z_t^N = 0$ for all $t \geq N$. Moreover, for $t \geq N$, X^N is the solution of

$$dX_t^N = \zeta b(X_t^N, 0) dt + \zeta \sum_{j=1}^m \sigma^j(X_t^N) dW_t^j.$$

Proposition 3.3.7 (Time horizon truncation error). *Let $N \geq 0$ and $\Pi^N = \Pi^N(1, 0)$ denote the truncated solution of the FBSDE under the terminal condition $Y_N = 0$ and let $\Pi^N(1, \mathcal{I})$ denote the solution of the infinite horizon FBSDE with inputs \mathcal{I} , restricted to time N . For a given $\varepsilon > 0$ s.t. $\rho > \rho - \varepsilon > \bar{K}$, we have the following result,*

$$\|\Pi^N - \Pi(1, 0)\|_{L^2}^2 \leq C e^{-\varepsilon N},$$

where $C > 0$ is a constant independent of T and ε .

Proof. The existence and uniqueness result 3.3.6 guarantees that $\Pi^N(1, \mathcal{I})$ exists for all $\mathcal{I} \in \mathbb{I}$. Let us take $\Theta = (X, Y, Z) \equiv \Pi(1, 0)$ which is the solution of the infinite horizon FBSDE. The above remark establishes the existence and uniqueness along with the stability result for $\Pi^N(1, \mathcal{I})$ and that $\Pi^N(1, 0) = \Pi^N$. With $f^N(x, y, z, \mathcal{D}_\mu h_t^N) = \mathbf{1}_{t \leq N} \left(f(x, y, z, \mathcal{D}_\mu h_t^N) + \rho y \right) - \rho y$, let us define the following input,

$$I_t^f = f^N(X_t^N, Y_t^N, Z_t^N, \mathcal{D}_\mu h_t^N) - f(X_t^N, Y_t^N, Z_t^N, \mathcal{D}_\mu h_t^N),$$

where X_t^N, Y_t^N, Z_t^N and $\mathcal{D}_\mu h_t^N$ come from the solution of the finite horizon truncated FBSDE. Then by substitution, it can be verified that $\Pi(1, I^f) = \Pi^N$. This enables us to use the stability result from Lemma 3.3.5 to get the following,

$$\begin{aligned} \|\Pi^N - \Pi(1, 0)\|_{L^2_\rho}^2 &= \|\Pi(1, I^f) - \Pi(1, 0)\|_{L^2_\rho}^2 \leq C \|I^f - 0\|_{\mathbb{I}}^2 \\ &\leq C \mathbb{E} \left[\int_N^\infty e^{-\varepsilon t} e^{-(\rho-\varepsilon)t} |f(\Theta_t)|^2 dt \right] \\ &\leq C e^{-\varepsilon N} \mathbb{E} \left[\int_N^\infty e^{-(\rho-\varepsilon)t} |f(\Theta_t)|^2 dt \right] \\ &\leq C e^{-\varepsilon N} \mathbb{E} \left[\int_N^\infty e^{-(\rho-\varepsilon)t} |Y_t A_t^x + \sum_{j=1}^m Z_t^j B_t^{x,j} + \mathcal{D}_\mu h_t - (\rho - \varepsilon) Y_t + \varepsilon Y_t|^2 dt \right] \end{aligned}$$

We can use the a priori estimates for the BSDE and X_t (for a smaller ρ but it works because $\rho - \varepsilon$ still satisfies the assumption) to obtain a bound on the above expectation which is independent of N and η ,

$$\begin{aligned} \mathbb{E} \left[\int_N^\infty e^{-(\rho-\varepsilon)t} |Y_t A_t^x + \sum_{j=1}^m Z_t^j B_t^{x,j} + \mathcal{D}_\mu h_t - \rho Y_t|^2 dt \right] \\ \leq \mathbb{E} \left[\int_0^\infty e^{-(\rho-\varepsilon)t} |Y_t A_t^x + \sum_{j=1}^m Z_t^j B_t^{x,j} + \mathcal{D}_\mu h_t - \rho Y_t|^2 dt \right] \leq C. \end{aligned}$$

Therefore, we have,

$$\|\Pi^N - \Pi(1, 0)\|_{L^2_\rho}^2 \leq C e^{-\varepsilon N}.$$

□

3.4 Numerical Analysis

In this section, we will tackle the problem of solving a truncated version of the coupled McKean-Vlasov FBSDE and ignore the truncation error. For sake of convenience, let us consider the stock drift, volatility and interest rate to be constants b, σ and r respectively and for the cost functional, we will consider the special case with quadratic cost in control which is of main interest for us, $g(x, q) = \frac{\gamma}{2} |q_t|^2$. This enables us to find the form for the optimal control in terms of the solution of the FBSDE (3.3.4). We will propose a numerical scheme for solving this FBSDE which fully characterises the optimal control. It will be useful to write the SDE and BSDE into different components as we can see that the forward dynamics do not depend on

the first component of the BSDE. So, we essentially obtain the following finite horizon forward-backward system,

$$\begin{aligned} d\mathcal{V}_t^q &= (\alpha_t^\mathcal{V}\mathcal{V}_t^q + \alpha_t^\delta.\delta_t)dt + \sum_{j=1}^m (\beta_t^{\mathcal{V},j}\mathcal{V}_t^q + \beta_t^{\delta,j}.\delta_t)dW_t^j, \quad \mathcal{V}_t^q|_{t=0} = \mathcal{V}_0^q, \\ d\delta_t^i &= -\frac{1}{\gamma}Y_t^{i+1}dt, \quad \delta_t^i|_{t=0} = 0, \quad i \in \{1, 2, \dots, N\}, \\ dY_t^1 &= -\left(Y_t^1\alpha_t^\mathcal{V} + \sum_{j=1}^m Z_t^{j,1}\beta_t^{\mathcal{V},j} + \partial_\mu h(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) - \rho Y_t^1\right)dt + \sum_{j=1}^m Z_t^{j,1}dW_t^j \\ dY_t^i &= -\left(Y_t^1\alpha_t^{\delta,i} + \sum_{j=1}^m Z_t^{j,1}\beta_t^{\delta,j,i} - \rho Y_t^i\right)dt + \sum_{j=1}^m Z_t^{j,i}dW_t^j, \quad i \in \{2, 3, \dots, N+1\}, \\ Y_T^i &= 0, \quad i \in \{1, 2, 3, \dots, N+1\}, \end{aligned}$$

where we use the terminal condition for the truncated time horizon, $Y_T = 0$. In the next proposition, we express Y as a conditional expectation in terms of the L-derivative.

Proposition 3.4.1. *Under the assumptions (H0), (H1), (H2) and (H3), Y^1 and $(Y^i)_{i=2,\dots,N+1}$ have the following representation,*

$$\begin{aligned} Y_s^1 &= e^{\rho s - \int_0^s \alpha_r^\mathcal{V} dr} \mathbb{E} \left[\left(\int_s^T e^{-\rho t + \int_0^t \alpha_r^\mathcal{V} dr} \frac{L_t}{L_s} \partial_\mu h(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) dt \right) \middle| \mathcal{F}_s \right], \\ Y_s^i &= e^{\rho s} \mathbb{E} \left[\int_s^T e^{-\int_0^t \alpha_r^\mathcal{V} dr} \left(\int_t^T \frac{L_u}{L_t} \partial_\mu \hat{h}(\mu_{\mathcal{V}_u})(\mathcal{V}_u) du \right) \alpha_t^{\delta,i} dt \middle| \mathcal{F}_s \right] + \\ &\quad \mathbb{E} \left[\frac{L_T}{L_s} \left(\int_s^T \partial_\mu \hat{h}(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) dt \right) \times \sum_{j=1}^m \left(\int_s^T e^{-\int_0^t \alpha_r^\mathcal{V} dr} \frac{L_s}{L_t} \beta_t^{\delta,j,i} dW_t^j - \right. \right. \\ &\quad \left. \left. \int_s^T e^{-\int_0^t \alpha_r^\mathcal{V} dr} \frac{L_s}{L_t} \beta_t^{\delta,j,i} \beta_t^{\mathcal{V},j} dt \right) \middle| \mathcal{F}_s \right], \quad i = 2, 3, \dots, N+1, \end{aligned}$$

where L_T is the Radon-Nikodym derivative given by,

$$L_T = \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right)_{\mathcal{F}_T} = \exp \left(\int_0^T \sum_{j=1}^m \beta_t^{\mathcal{V},j} dW_t^j - \frac{1}{2} \int_0^T \sum_{j=1}^m |\beta_t^{\mathcal{V},j}|^2 dt \right).$$

Proof. Note that there is no z -dependence in the i -th component of the BSDE for $i \geq 2$, which means we can write the solution as a conditional expectation in terms of the solution of the first component,

$$e^{-\rho s} Y_s^i - e^{-\rho T} \mathbb{E} [Y_T^i | \mathcal{F}_s] = \mathbb{E} \left[\int_s^T e^{-\rho t} \left(Y_t^1 \alpha_t^{\delta,i} + \sum_{j=1}^m Z_t^{j,1} \beta_t^{\delta,j,i} \right) dt \middle| \mathcal{F}_s \right]. \quad (3.4.1)$$

For the first component, we can use a Girsanov transformation to get rid of z -dependence in the generator by introducing a new Brownian motion:

$$\tilde{W}_T^j = W_T^j - \int_0^T \beta_t^{\mathcal{V},j} dt, \quad \forall T \geq 0, \quad j \in \{1, 2, \dots, m\},$$

under the new probability measure \mathbb{Q} . It can be verified that Novikov's condition is satisfied for its Radon-Nikodym derivative. Note that W_t^j was independent of W_t^k for $k \neq j$ in our setting

and the Brownian motion under \mathbb{Q} preserves this property. Next, we can write the first BSDE component under the new measure as,

$$dY_t^1 = -\left(Y_t^1(\alpha_t^\nu - \rho) + \partial_\mu h(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q)\right)dt + \sum_{j=1}^m Z_t^{j,1} d\tilde{W}_t^j, \quad (3.4.2)$$

$$d(e^{-\rho t + \int_0^t \alpha_r^\nu dr} Y_t^1) + e^{-\rho t + \int_0^t \alpha_r^\nu dr} \partial_\mu h(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q)dt = \sum_{j=1}^m e^{-\rho t + \int_0^t \alpha_r^\nu dr} Z_t^{j,1} d\tilde{W}_t^j,$$

enabling us to write it as a conditional expectation,

$$e^{-\rho s + \int_0^s \alpha_r^\nu dr} Y_s^1 - \mathbb{E}^{\mathbb{Q}} \left[e^{-\rho T + \int_0^T \alpha_r^\nu dr} Y_T^1 | \mathcal{F}_s \right] = \mathbb{E}^{\mathbb{Q}} \left[\int_s^T e^{-\rho t + \int_0^t \alpha_r^\nu dr} \partial_\mu h(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) dt \middle| \mathcal{F}_s \right]$$

Denote with a hat the quantity multiplied by $e^{-\rho t + \int_0^t \alpha_r^\nu dr}$. We have

$$\begin{aligned} \hat{Y}_s^1 &= \mathbb{E}^{\mathbb{Q}} \left[\hat{Y}_T^1 + \int_s^T \partial_\mu \hat{h}(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) dt \middle| \mathcal{F}_s \right], \\ &= \mathbb{E}^{\mathbb{P}} \left[\frac{L_T}{L_s} \left(\hat{Y}_T^1 + \int_s^T \partial_\mu \hat{h}(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) dt \right) \middle| \mathcal{F}_s \right] \\ &= \mathbb{E}^{\mathbb{P}} \left[\left(\int_s^T \frac{L_t}{L_s} \partial_\mu \hat{h}(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) dt \right) \middle| \mathcal{F}_s \right], \end{aligned}$$

where we take $Y_T = 0$ and martingale property of L_T in the last equation. The BSDE (3.4.2) can be written as,

$$d(\hat{Y}_t^1) + \partial_\mu \hat{h}(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q)dt = \sum_{j=1}^m \hat{Z}_t^{j,1} d\tilde{W}_t^j.$$

The Itô isometry gives that for a function h in L^2 :

$$\mathbb{E}^{\mathbb{P}} \left[\int_s^T h_t^j Z_t^{j,1} dt \middle| \mathcal{F}_s \right] = \mathbb{E}^{\mathbb{Q}} \left[\frac{L_s}{L_T} \int_s^T h_t^j Z_t^{j,1} dt \middle| \mathcal{F}_s \right] = \mathbb{E}^{\mathbb{Q}} \left[\int_s^T Z_t^{j,1} d\tilde{W}_t^j \times \int_s^T \frac{L_s}{L_T} h_t^j d\tilde{W}_t^j \middle| \mathcal{F}_s \right].$$

Using the fact is $1/L_t$ is \mathcal{F}_t -martingale and the tower property of conditional expectation, we get,

$$\begin{aligned} \mathbb{E}^{\mathbb{P}} \left[\int_s^T h_t^j Z_t^{j,1} dt \middle| \mathcal{F}_s \right] &= \mathbb{E}^{\mathbb{Q}} \left[\frac{L_s}{L_T} \int_s^T h_t^j Z_t^{j,1} dt \middle| \mathcal{F}_s \right] = \int_s^T \mathbb{E}^{\mathbb{Q}} \left[\frac{L_s}{L_t} h_t^j Z_t^{j,1} \middle| \mathcal{F}_s \right] dt \\ &= \mathbb{E}^{\mathbb{Q}} \left[\int_s^T Z_t^{j,1} d\tilde{W}_t^j \times \int_s^T \frac{L_s}{L_t} h_t^j d\tilde{W}_t^j \middle| \mathcal{F}_s \right]. \end{aligned}$$

Using the same approach, one has the following expectation representation of the sum of $Z^{j,1}$,

$$\begin{aligned} &\mathbb{E}^{\mathbb{P}} \left[\int_s^T e^{-\rho t} \sum_{j=1}^m Z_t^{j,1} \beta_t^{\delta,j,i} dt \middle| \mathcal{F}_s \right] \\ &= \mathbb{E}^{\mathbb{Q}} \left[\left(\sum_{j=1}^m \int_s^T e^{-\rho t + \int_0^t \alpha_r^\nu dr} Z_t^{j,1} d\tilde{W}_t^j \right) \times \left(\sum_{j=1}^m \int_s^T e^{-\int_0^t \alpha_r^\nu dr} \frac{L_s}{L_t} \beta_t^{\delta,j,i} d\tilde{W}_t^j \right) \middle| \mathcal{F}_s \right] \end{aligned}$$

$$\begin{aligned}
& \stackrel{(3.4.2)}{=} \mathbb{E}^{\mathbb{Q}} \left[\left(e^{-\rho T + \int_0^T \alpha_r^\nu dr} Y_T^1 - e^{-\rho s + \int_0^s \alpha_r^\nu dr} Y_s^1 + \int_s^T e^{-\rho t + \int_0^t \alpha_r^\nu dr} \partial_\mu h(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) dt \right) \times \right. \\
& \quad \left. \sum_{j=1}^m \int_s^T e^{-\int_0^t \alpha_r^\nu dr} \frac{L_s}{L_t} \beta_t^{\delta,j,i} d\tilde{W}_t^j \middle| \mathcal{F}_s \right] \\
& = \mathbb{E}^{\mathbb{Q}} \left[\left(\hat{Y}_T^1 - \hat{Y}_s^1 + \int_s^T \partial_\mu \hat{h}(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) dt \right) \times \sum_{j=1}^m \int_s^T e^{-\int_0^t \alpha_r^\nu dr} \frac{L_s}{L_t} \beta_t^{\delta,j,i} d\tilde{W}_t^j \middle| \mathcal{F}_s \right] \\
& = \mathbb{E}^{\mathbb{Q}} \left[\left(\hat{Y}_T^1 + \int_s^T \partial_\mu \hat{h}(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) dt \right) \times \sum_{j=1}^m \int_s^T e^{-\int_0^t \alpha_r^\nu dr} \frac{L_s}{L_t} \beta_t^{\delta,j,i} d\tilde{W}_t^j \middle| \mathcal{F}_s \right] - \\
& \quad \hat{Y}_s^1 \mathbb{E}^{\mathbb{Q}} \left[\sum_{j=1}^m \int_s^T e^{-\int_0^t \alpha_r^\nu dr} \frac{L_s}{L_t} \beta_t^{\delta,j,i} d\tilde{W}_t^j \middle| \mathcal{F}_s \right] \\
& = \mathbb{E}^{\mathbb{Q}} \left[\left(\hat{Y}_T^1 + \int_s^T \partial_\mu \hat{h}(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) dt \right) \times \sum_{j=1}^m \int_s^T e^{-\int_0^t \alpha_r^\nu dr} \frac{L_s}{L_t} \beta_t^{\delta,j,i} d\tilde{W}_t^j \middle| \mathcal{F}_s \right] \\
& = \mathbb{E}^{\mathbb{P}} \left[\frac{L_T}{L_s} \times \left(\hat{Y}_T^1 + \int_s^T \partial_\mu \hat{h}(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) dt \right) \times \sum_{j=1}^m \left(\int_s^T e^{-\int_0^t \alpha_r^\nu dr} \frac{L_s}{L_t} \beta_t^{\delta,j,i} dW_t^j - \right. \right. \\
& \quad \left. \left. \int_s^T e^{-\int_0^t \alpha_r^\nu dr} \frac{L_s}{L_t} \beta_t^{\delta,j,i} \beta_t^{\nu,j} dt \right) \middle| \mathcal{F}_s \right] \\
& = \mathbb{E}^{\mathbb{P}} \left[\frac{L_T}{L_s} \times \left(\int_s^T \partial_\mu \hat{h}(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) dt \right) \times \sum_{j=1}^m \left(\int_s^T e^{-\int_0^t \alpha_r^\nu dr} \frac{L_s}{L_t} \beta_t^{\delta,j,i} dW_t^j - \right. \right. \\
& \quad \left. \left. \int_s^T e^{-\int_0^t \alpha_r^\nu dr} \frac{L_s}{L_t} \beta_t^{\delta,j,i} \beta_t^{\nu,j} dt \right) \middle| \mathcal{F}_s \right],
\end{aligned}$$

where we have used in the last equation $Y_T = 0$. Using this expression for the above term, we can express Y_s^i using (3.4.1) as follows,

$$\begin{aligned}
& e^{-\rho s} Y_s^i - e^{-\rho T} \mathbb{E}^{\mathbb{P}} [Y_T^i | \mathcal{F}_s] \\
& = \mathbb{E}^{\mathbb{P}} \left[\int_s^T e^{-\int_0^t \alpha_r^\nu dr} \left(\hat{Y}_t^1 \alpha_t^{\delta,i} + \sum_{j=1}^m \hat{Z}_t^{j,1} \beta_t^{\delta,j,i} \right) dt \middle| \mathcal{F}_s \right] \\
& = \mathbb{E}^{\mathbb{P}} \left[\int_s^T e^{-\int_0^t \alpha_r^\nu dr} \left(\mathbb{E}^{\mathbb{P}} \left[\frac{L_T}{L_t} \left(\hat{Y}_T^1 + \int_t^T \partial_\mu \hat{h}(\mu_{\mathcal{V}_u})(\mathcal{V}_u) du \right) \middle| \mathcal{F}_t \right] \alpha_t^{\delta,i} \right) dt \middle| \mathcal{F}_s \right] + \\
& \quad \mathbb{E}^{\mathbb{P}} \left[\frac{L_T}{L_s} \left(\hat{Y}_T^1 + \int_s^T \partial_\mu \hat{h}(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) dt \right) \times \sum_{j=1}^m \left(\int_s^T e^{-\int_0^t \alpha_r^\nu dr} \frac{L_s}{L_t} \beta_t^{\delta,j,i} dW_t^j \right. \right. \\
& \quad \left. \left. - \int_s^T e^{-\int_0^t \alpha_r^\nu dr} \frac{L_s}{L_t} \beta_t^{\delta,j,i} \beta_t^{\nu,j} dt \right) \middle| \mathcal{F}_s \right] \\
& = \mathbb{E}^{\mathbb{P}} \left[\int_s^T e^{-\int_0^t \alpha_r^\nu dr} \left(\int_t^T \frac{L_u}{L_t} \partial_\mu \hat{h}(\mu_{\mathcal{V}_u})(\mathcal{V}_u) du \right) \alpha_t^{\delta,i} dt \middle| \mathcal{F}_s \right] \\
& \quad + \mathbb{E}^{\mathbb{P}} \left[\frac{L_T}{L_s} \left(\int_s^T \partial_\mu \hat{h}(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) dt \right) \times \sum_{j=1}^m \left(\int_s^T e^{-\int_0^t \alpha_r^\nu dr} \frac{L_s}{L_t} \beta_t^{\delta,j,i} dW_t^j \right. \right.
\end{aligned}$$

$$- \int_s^T e^{-\int_0^t \alpha_r^\nu dr} \frac{L_s}{L_t} \beta_t^{\delta,j,i} \beta_t^{\nu,j} dt \Big| \mathcal{F}_s \Big].$$

In the last inequality, we take $Y_T = 0$. Hence, we get the following expression for Y^i ,

$$\begin{aligned} e^{-\rho s} Y_s^i &= \mathbb{E}^{\mathbb{P}} \left[\int_s^T e^{-\int_0^t \alpha_r^\nu dr} \left(\int_t^T \frac{L_u}{L_t} \partial_\mu \hat{h}(\mu \nu_u)(\nu_u) du \right) \alpha_t^{\delta,i} dt \Big| \mathcal{F}_s \right] \\ &+ \mathbb{E}^{\mathbb{P}} \left[\frac{L_T}{L_s} \left(\int_s^T \partial_\mu \hat{h}(\mu \nu_t^q)(\nu_t^q) dt \right) \times \sum_{j=1}^m \left(\int_s^T e^{-\int_0^t \alpha_r^\nu dr} \frac{L_s}{L_t} \beta_t^{\delta,j,i} dW_t^j \right. \right. \\ &\left. \left. - \int_s^T e^{-\int_0^t \alpha_r^\nu dr} \frac{L_s}{L_t} \beta_t^{\delta,j,i} \beta_t^{\nu,j} dt \right) \Big| \mathcal{F}_s \right]. \end{aligned}$$

□

3.4.1 General setting

Since we took b_t, σ_t, r_t to be constants, the functions $\alpha_t^\nu, \alpha_t^\delta, \beta_t^{\nu,j}$ and $\beta_t^{\delta,j}$ are functions of the relative asset price process S_t , i.e.

$$\alpha_t^\nu \equiv \alpha^\nu(S_t), \quad \alpha_t^\delta \equiv \alpha^\delta(S_t), \quad \beta_t^{\nu,j} \equiv \beta^{\nu,j}(S_t), \quad \beta_t^{\delta,j} \equiv \beta^{\delta,j}(S_t).$$

Also note that we can write the L-derivative as a function of the state variable and expectation of a function of the state,

$$\partial_\mu h(\mu)(\mathcal{V}_t^q) = \int_0^\infty \frac{-1}{1 + e^{\lambda p_k - (\mathbb{E}[\log(1 + e^{\lambda(k - \mathcal{V}_t^q)])])}} \frac{1}{1 + e^{\lambda(\mathcal{V}_t^q - k)}} d\nu(k) \equiv H(\mathcal{V}_t^q, \mathbb{E}[\theta(\mathcal{V}_t^q)]) \quad (3.4.3)$$

where $\theta(x) = \log(1 + e^{\lambda(k - \mathcal{V}_t^q)})$. This motivates us to write the following Markovian representation of Y^1 and $(Y^i)_{i=2, \dots, N+1}$,

$$Y_s^1 = \psi^1(\mathcal{V}_s, S_s, \mathbb{E}[\theta(\mathcal{V}_s)]); \quad Y_s^i = \psi^2(\mathcal{V}_s, S_s, \mathbb{E}[\theta(\mathcal{V}_s)]), \quad i = 2, \dots, N+1,$$

motivated by the notion of decoupling fields (see Ma et al. [92]).

Let us consider a time discretization $\mathbb{T} = \{t_0 = 0, t_1, t_2, \dots, t_{T_{max}} = T\}$. Next, we simulate the Brownian motion with D sample paths $\{\{W_{t_k}^a\}_{k \in \{0,1, \dots, T_{max}\}}\}_{a \in \{0,1, \dots, D\}}$ over the truncated time horizon. We estimate the a -th sample path of the asset price process S_t using the following:

$$\begin{aligned} (\mathbf{S}_{t_k}^a)^i &= \mathbf{S}_0^i \exp \left(\left(b^i - \frac{\sum_{j=1}^m (\sigma^{i,j})^2}{2} \right) t_k + \sum_{j=1}^m \sigma^{i,j} W_{t_k}^{a,j} \right), \\ (\mathbf{S}_{t_k}^a)^0 &= \mathbf{S}_0^0 e^{rt_k}. \end{aligned}$$

Then, the relative asset price process S^i is given by,

$$(S_{t_k}^a)^i = \frac{(\mathbf{S}_{t_k}^a)^i}{\sum_{i=0}^N \gamma^i (\mathbf{S}_{t_k}^a)^i}.$$

The relative wealth process \mathcal{V} is simulated using an Euler scheme for the linear SDE (3.2.1) with stochastic coefficients which are functions of S .

To avoid writing long equations, we will denote the drift and volatility of X_t by $b(t, x, y)$ and $\sigma(t, x)$ respectively where the dependence on y comes by replacing the optimal control q_t by $-\frac{1}{\gamma}(Y_t^{[2:N+1]})^\top$. Also, we use $f(t, x, \theta, y, z)$ to denote the generator of the BSDE since the L-derivative at time t is a function of X_t and $\theta_t := \mathbb{E}[\theta(X_t)]$.

Target distribution and p_k . We take the target distribution π_t^* for the relative wealth process \mathcal{V}_t^q at time t to be the distribution of the following geometric Brownian motion (GBM),

$$dP_t^* = P_t^*(\mu dt + \sigma dW_t), \quad P_0^* = 1,$$

with $\mu = \sigma = 0.05$. This sets both the desired growth rate and permitted variation of the relative wealth process once we calculate the parameters $p_k(t)$. These can be calculated as follows: When \mathcal{V}_t^q attains the target distribution,

$$\left[\mathbb{E} [k - P_t^*]_+^\lambda - p_k(t) \right]_+^\lambda = 0,$$

which means that $p_k(t)$ needs to be larger than or equal to $\mathbb{E} [k - P_t^*]_+^\lambda$. We can take

$$p_k(t) = \mathbb{E} [k - P_t^*]_+ = e^{\mu t} \mathbb{E} \left[k e^{-\mu t} - e^{\sigma W_t - \frac{1}{2} \sigma^2 t} \right]_+.$$

Since $e^{\sigma W_t - \frac{1}{2} \sigma^2 t}$ is a martingale, the above expectation is given by the price of the European put option with risk-free interest rate being zero, the strike price being $k e^{-\mu t}$ and maturity t , which can be calculated using the Black-Scholes formula,

$$p_k(t) = e^{\mu t} (k e^{-\mu t} N(-d_2) - N(-d_1)),$$

where N is the cumulative distribution function of the standard normal distribution and,

$$d_1 = \frac{1}{\sigma \sqrt{t}} \left(\log \frac{1}{k e^{-\mu t}} + \frac{\sigma^2 t}{2} \right); \quad d_2 = d_1 - \sigma \sqrt{t}.$$

For numerical purposes, we will use a discrete measure $\nu = \frac{1}{K} \sum_{i=1}^K \delta_{k_i}$ where the support $\{k_i\}_{i=1, \dots, K}$ is centered at 1 as the relative wealth process starts from 1.

3.4.2 Scheme based on fictitious play

We refer to Han and Long [77] for an approach for solving coupled FBSDEs using neural networks and the article Han et al. [78] for solving McKean Vlasov coupled FBSDEs in the finite horizon. In the latter article, the authors suggest the use of ‘fictitious play’ to find the solution of the McKean Vlasov FBSDE using a fixed point argument where one computes the law dependent part using the previous iteration of the solution and treats the FBSDE as a standard coupled FBSDE. For solving the coupled FBSDE, they use neural networks to approximate the decoupling fields as functions of the state and time, which in our case, are function of time, state and the relative asset price process. Starting from an initial guess, the system evolves forward in time using an Euler scheme and we train the neural networks to minimize the deviation from the terminal condition. See Theorem 1 in [77] for a proof that the resulting simulation error can be bounded by $C \mathbb{E} [|Y_T|^2]$ for some constant $C > 0$. Let the fictitious play iteration be denoted by p . For each p , we use neural networks \mathcal{Y}^p and $(\mathcal{Z}_t^p)_{t \in \{0, t_1, \dots, T_{max}\}}$ to approximate the decoupling fields,

$$Y_0 \approx \mathcal{Y}^p(X_0, S_0); \quad Z_t \approx \mathcal{Z}_t^p(X_t, S_t, Y_t).$$

The numerical scheme is thus outlined as follows:

1. Initialize $(\theta_{t_i}^{p=0})_{i \in \{0, 1, \dots, T_{max}\}}$ as the guess function for $\mathbb{E} [\theta(\mathcal{V}_t)]$.
2. For any p , initialize the neural networks $\mathcal{Y}^p(x, s)$ and $\{\mathcal{Z}_{t_i}^p(x, s, y)\}_{i \in \{0, 1, \dots, T_{max}\}}$ with random weights.

3. Find the corresponding solution $(X_{t_i}^p, Y_{t_i}^p, Z_{t_i}^p)_{0 \leq i \leq T_{max}}$ using the forward Euler scheme for the resulting FBSDE, which is given as follows,

$$\begin{aligned} X_0^p &= [1, 0, \dots, 0]; \quad Y_0^p = \mathcal{Y}^p(X_0, S_0); \quad Z_0^p = \mathcal{Z}_0^p(X_0^p, S_0, Y_0^p); \\ X_{t_{i+1}}^p &= X_{t_i}^p + b(t_i, X_{t_i}^p, Y_{t_i}^p) + \sigma(t_i, X_{t_i}^p) \Delta W_{t_i}; \\ Z_{t_i}^p &= \mathcal{Z}_{t_i}^p(X_{t_i}^p, S_{t_i}, Y_{t_i}^p); \\ Y_{t_{i+1}}^p &= Y_{t_i}^p - f(t_i, X_{t_i}^p, \theta_{t_i}^{p-1}, Y_{t_i}^p, Z_{t_i}^p) + Z_{t_i}^p \Delta W_{t_i}. \end{aligned}$$

4. Train the neural networks with this data to minimize the following,

$$\mathbb{E} [|Y_{T_{max}}^p|^2] \approx \frac{1}{D} \sum_{a=1}^D |Y_{t_{T_{max}}}^{p,a}|^2,$$

since the terminal condition of our truncated problem is $Y_T = 0$. Update the estimates of \mathcal{Y}^p and $(\mathcal{Z}_{t_i}^p)_{1 \leq i \leq T_{max}}$.

5. Update the estimate of $\theta_{t_i}^p$ with $\theta_{t_i}^p = \frac{1}{D} \sum_{a=1}^D \theta(\mathcal{V}_{t_i}^{p-1})$ and solve the corresponding FBSDE using the Euler Scheme as before and iterate.

A natural choice for the initialization of $(\theta_{t_i}^{p=0})_{i \in \{0,1,\dots,T_{max}\}}$ is by computing $\theta(\mathcal{V}_{t=0})$, which is $\log(1 + e^{\lambda(k-\mathcal{V}_0)})$ for all i . The above scheme can be summarized as follows:

Algorithm 1 Feed-forward NN fictitious play

- 1: We start with an initial guess $(\theta_{t_i}^{p=0})_{0 \leq i \leq T_{max}}$ for $\mathbb{E}[\theta(\mathcal{V}_t)]$. At each fictitious play step p , the unknown processes Y_0 and $(Z_t)_{t \in [0, T]}$ are approximated by neural networks (NNs) \mathcal{Y}^p and $(\mathcal{Z}_{t_i}^p)_{0 \leq i \leq T_{max}}$ respectively. We solve the FBSDE forward in time while training the NNs to minimize the expected squared deviation of Y_T from 0.
 - 2: Initialize $(\theta_{t_i}^0)_{0 \leq i \leq T_{max}}$.
 - 3: **for** p from 1 to p_{max} **do**
 - 4: Initialize $X_0^p = [1, 0, \dots, 0]$, $Y_0^p = \mathcal{Y}^p$ and $Z_0^p = \mathcal{Z}_0^p$.
 - 5: Sample D paths of Brownian motion W and stock price S .
 - 6: **for** i from 0 to $T_{max} - 1$ **do**
 - 7: $X_{t_{i+1}}^p \leftarrow X_{t_i}^p + b(t_i, X_{t_i}^p, Y_{t_i}^p) + \sigma(t_i, X_{t_i}^p) \Delta W_{t_i}$
 - 8: $Z_{t_i}^p \leftarrow \mathcal{Z}_{t_i}^p(X_{t_i}^p, S_{t_i}, Y_{t_i}^p)$
 - 9: $Y_{t_{i+1}}^p \leftarrow Y_{t_i}^p - f(t_i, X_{t_i}^p, \theta_{t_i}^{p-1}, Y_{t_i}^p, Z_{t_i}^p) + Z_{t_i}^p \Delta W_{t_i}$.
 - 10: **end for**
 - 11: Train the neural networks to minimize $\frac{1}{D} \sum_{a=1}^D |Y_{t_{T_{max}}}^{p,a}|^2$.
 - 12: Update $\theta_{t_i}^p \leftarrow \frac{1}{D} \sum_{a=1}^D \theta(\mathcal{V}_{t_i}^{p-1,a})$ for all $0 \leq i \leq T_{max}$.
 - 13: **end for**
-

Next, we present a scheme based on dynamic updates of the measure dependent part.

3.4.3 Scheme based on dynamic updates

Germain et al. [62] developed a numerical scheme which does not rely on the above fictitious play iteration. The law dependent part is updated dynamically at each time step using empirical mean for the given batch of particles. The rest of the numerical scheme is same as the previous

one. Also, similar to the previous scheme, one could initialize $(\theta_{t_i}^{p=0})_{i \in \{0,1,\dots,T_{max}\}}$ as $\theta(\mathcal{V}_{t=0}) = \log(1 + e^{\lambda(k-\mathcal{V}_0)})$ for all i but this choice is subjective. The numerical scheme can be summarized as follows:

Algorithm 2 Feed-forward NN dynamic updates

- 1: We start with an initial guess $(\theta_{t_i}^{p=0})_{0 \leq i \leq T_{max}}$ for $\mathbb{E}[\theta(\mathcal{V}_t)]$. The unknown processes Y_0 and $(Z_t)_{t \in [0,T]}$ are approximated by neural networks (NNs) \mathcal{Y} and $(\mathcal{Z}_{t_i})_{0 \leq i \leq T_{max}}$ respectively. At each time step t_i , we update the estimate for the L-derivative using the average of $\theta(\mathcal{V}_{t_{i-1}})$. We solve the FBSDE forward in time while training the NNs to minimize the deviation of Y_T from 0.
 - 2: Initialize $X_0 = [1, 0, \dots, 0]$, $Y_0 = \mathcal{Y}$ and $Z_0 = \mathcal{Z}_0$.
 - 3: Sample D paths of Brownian motion W and stock price S .
 - 4: **for** i from 0 to $T_{max} - 1$ **do**
 - 5: $\theta_{t_i} \leftarrow \frac{1}{D} \sum_{a=1}^D \theta(\mathcal{V}_{t_i}^a)$
 - 6: $X_{t_{i+1}} \leftarrow X_{t_i} + b(t_i, X_{t_i}, Y_{t_i}) + \sigma(t_i, X_{t_i}) \Delta W_{t_i}$
 - 7: $Z_{t_i} \leftarrow \mathcal{Z}_{t_i}(X_{t_i}, S_{t_i}, Y_{t_i})$
 - 8: $Y_{t_{i+1}} \leftarrow Y_{t_i} - f(t_i, X_{t_i}, \theta_{t_i}, Y_{t_i}, Z_{t_i}) + Z_{t_i} \Delta W_{t_i}$.
 - 9: **end for**
 - 10: Train the neural networks to minimize $\frac{1}{D} \sum_{a=1}^D |Y_{t_{T_{max}}}^{p,a}|^2$.
-

3.4.4 Numerical Experiments

In this subsection, we provide some numerical experiments for Algorithm 1 and 2 for the dynamic allocation of one stock and one risky asset. We use the following common setting for both the schemes: terminal time $T = 1$, number of time steps $T_{max} = 52$, number of samples $D = 1024$, support for the L-derivative $\nu_{supp} = \{0.95, 0.96, \dots, 1.05\}$, smoothing parameter $\lambda = 10$, discount factor $\rho = 0.2$, risk-free interest rate $r = 0.026$, drift of risky asset price $b = 0.274$, volatility $\sigma = 0.254$ and starting price $S_{t=0} = 72.14$ (using historical daily adjusted closing price of Apple stock from the year 2010 to 2020). The values of the drift and volatility are yearly since we have taken $T = 1$. The starting value of the wealth process is $V_0 = 1000$ and the starting stock allocation is $\delta_0^1 = 8.317$ units, which means that the starting amount in the risk-free asset is $\delta_0^0 = 400$. We take the static allocation reference portfolio as $\zeta_t^0 = 400$, $\zeta_t^1 = 8.317$, $\forall t \geq 0$, which means that the starting value of the relative wealth process is $\mathcal{V}_0 = 1$. For numerical purposes, we define our system of McKean-Vlasov FBSDE with a constant multiplier ξ for the mean-field which appears in the system as follows,

$$\begin{aligned}
 d\mathcal{V}_t^q &= (\alpha_t^\mathcal{V} \mathcal{V}_t^q + \alpha_t^\delta \cdot \delta_t) dt + \sum_{j=1}^m (\beta_t^{\mathcal{V},j} \mathcal{V}_t^q + \beta_t^{\delta,j} \cdot \delta_t) dW_t^j, \quad \mathcal{V}_t^q \Big|_{t=0} = \mathcal{V}_0^q, \\
 d\delta_t^i &= -\frac{1}{\gamma} Y_t^{i+1} dt, \quad \delta_t^i \Big|_{t=0} = 0, \quad i \in \{1, 2, \dots, N\}, \\
 dY_t^1 &= -\left(Y_t^1 \alpha_t^\mathcal{V} + \sum_{j=1}^m Z_t^{j,1} \beta_t^{\mathcal{V},j} + \xi \partial_\mu h(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) - \rho Y_t^1 \right) dt + \sum_{j=1}^m Z_t^{j,1} dW_t^j \\
 dY_t^i &= -\left(Y_t^1 \alpha_t^{\delta,i} + \sum_{j=1}^m Z_t^{j,1} \beta_t^{\delta,j,i} - \rho Y_t^i \right) dt + \sum_{j=1}^m Z_t^{j,i} dW_t^j, \quad i \in \{2, 3, \dots, N+1\}, \\
 Y_T^i &= 0, \quad i \in \{2, 3, \dots, N+1\}.
 \end{aligned}$$

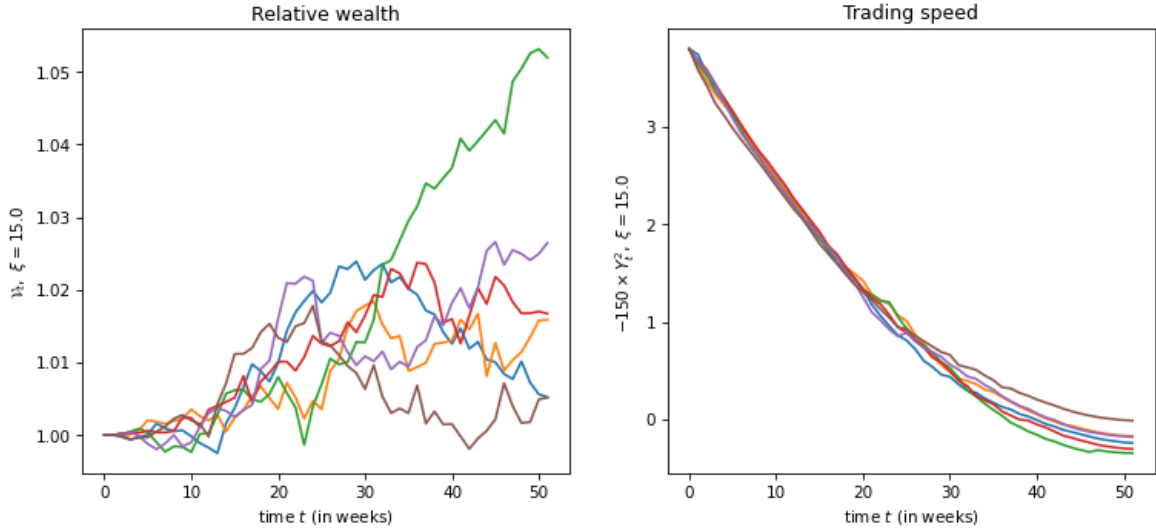


Figure 3.1: Sample trajectories of \mathcal{V}_t and $\frac{1}{\gamma}Y_T^2$ for $\frac{1}{\gamma} = 150$ and $\xi = 15$

The trading speed at time t is given by $-\frac{1}{\gamma}Y_t^{i+1}$. Increasing the parameter γ results in smaller values of the trading speed and the wealth process but result in the terminal values of Y_T^1 and Y_T^2 being closer to zero compared to smaller values of γ . The parameter ξ is very interesting for us to measure the impact of the weight given to the mean-field term which could be seen as the *distance* between the measure of the wealth process and the target distribution (strictly speaking, it is not a distance on $\mathcal{P}^2(\mathbb{R}^d)$). We must remark that there is no explicit solution for this problem to serve as a benchmark.

Fictitious play scheme

The scheme (see Algorithm 1) is based on a fixed point iterative approach where we start with the function $\bar{\theta}_{t_i}^{p=0} = \frac{1}{D} (\log 1 + e^{\lambda(k-1)})$, $\forall i \in \{1, \dots, T_{max}\}$ (with $D = 1024$) as the initial guess for $\mathbb{E}[\theta_i]$ (see (3.4.3)) and solve the resulting coupled FBSDE using a Euler-Maruyama scheme. The resulting solution gives the next iteration $\theta_{t_i}^1$ for all i and so on. For the neural network (NN), we use a sequence of sub-NNs \mathcal{Z}_{t_i} which are fully connected ReLU neural networks with two hidden layers, each with 20 nodes and batch normalization. At time t_i , the input for \mathcal{Z}_{t_i} is $(X_{t_i}, Y_{t_i}, S_{t_i})$. The output is used for next time step to obtain $(X_{t_{i+1}}, Y_{t_{i+1}})$ and at $t_{T_{max}}$, we minimize $\frac{1}{D} (|Y_{T_{max}}^1|^2 + |Y_{T_{max}}^2|^2)$. We use only two fictitious play iterations, each with 4000 gradient descent steps as the solution does not change significantly in further iterations. To ensure a balanced trade-off for a reasonable range of ξ , we have taken $\gamma \approx 0.00667$ ($\frac{1}{\gamma} = 150$ unless specified otherwise).

In Figure 3.1, we plot some sample trajectories of the relative wealth process and the corresponding trajectories of the trading speed for $\xi = 15.0$. We can notice that the trading speed is closely packed and not overly aggressive but terminate close to $Y_T = 0$ consistently. We can change this behavior by using the following loss function instead,

$$\frac{1}{D} (|Y_{T_{max}}^1|^2 + \kappa |Y_{T_{max}}^2|^2),$$

where $\kappa > 0$ is a constant smaller than 1. In Figure 3.2, we plot the relative wealth and trading speed trajectories for $\xi = 15.0$ but $\kappa = 0.4$ and 0.1 to illustrate the trade-off between the wealth returns and terminal condition.

Next, we measure the impact of changing ξ . In Figure 3.3, we plot the *distance* from the target

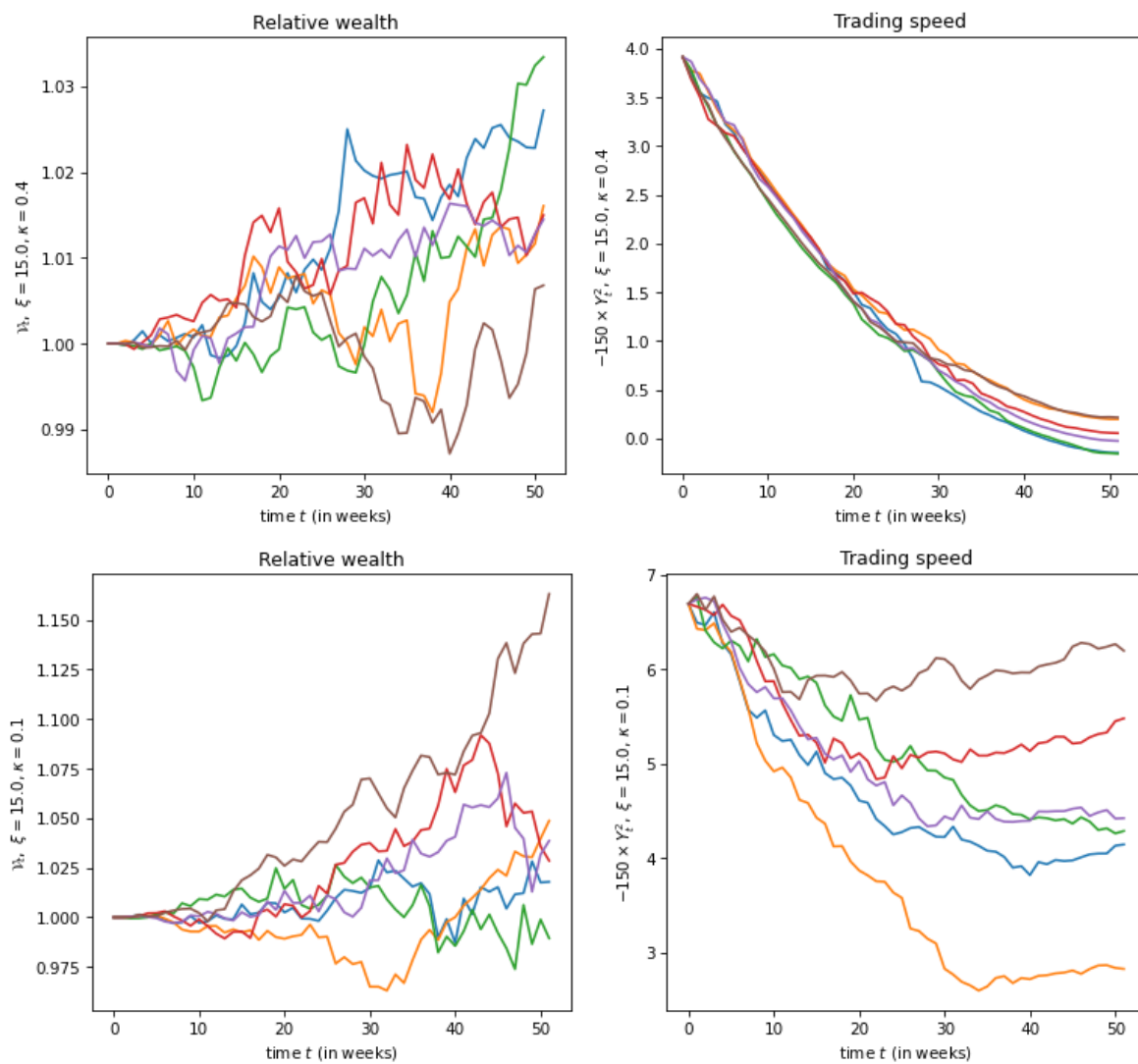


Figure 3.2: Trajectories of \mathcal{V}_t and $\frac{1}{\gamma}Y_T^2$ for $\frac{1}{\gamma} = 150$, $\xi = 15$ and $\kappa = 0.4$ and 0.1 (top and bottom respectively)

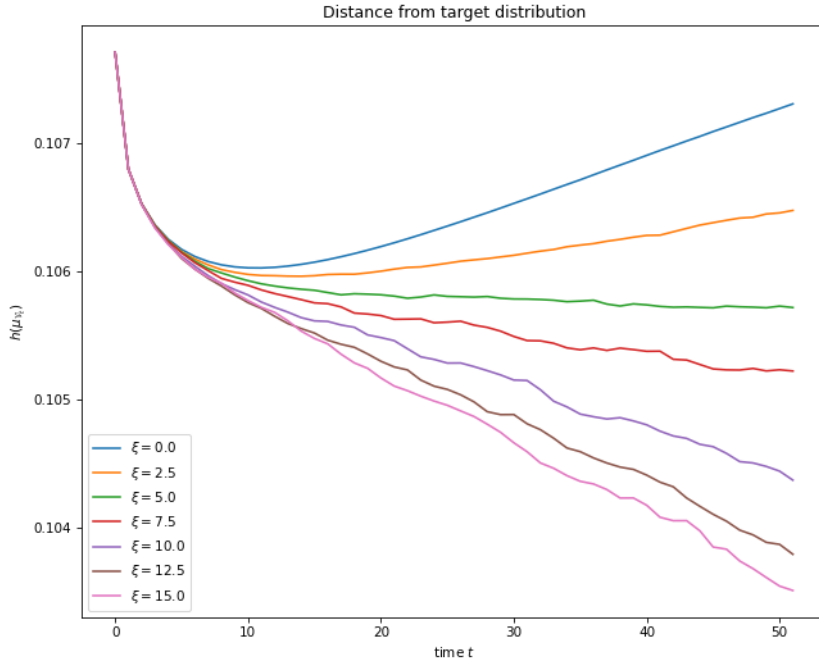


Figure 3.3: Distance between the target distribution and the measure of \mathcal{V}_t for different values of ξ .

distribution for different values of $\xi \in \{0, 2.5, \dots, 15\}$ and we observe that the distance reaches much lower values for higher values of ξ .

In Figure 3.3, we plot the cumulative distribution function (CDF) of the relative wealth process of different values of ξ which illustrates how increasing the multiplier of the mean-field term skews the distribution of wealth towards right while not increasing the probability of negative returns significantly. However, we cannot keep increasing the value of ξ due to the instability when we take $\xi \gtrsim 30$ for instance.

Dynamic update scheme

This scheme (see Algorithm 2) is different from the previous scheme in two key ways: it does not rely on a fixed point approach but rather updates the estimate of $\mathbb{E}[\theta_t]$ at each time step and we use a single neural network for \mathcal{Z}_{t_i} which takes t_i as an additional input. Also, we do not use batch normalization. We use a fully connected NN with two hidden layers with 30 nodes each. We take much smaller values of $\frac{1}{\gamma}$ for this scheme due to the instability for values $\frac{1}{\gamma} \gtrsim 40$. This scheme has comparatively more sensitive to the values ξ . Even for low enough values of $\xi \approx 20$, the terminal values of Y_T^2 are not as close to zero as they were for the first scheme. However, the values of Y_T^1 are very close to zero even for large values of ξ .

The impact of increasing ξ (within the range of stability) has a similar impact on the CDF of the relative wealth process and the *distance* from the target distribution as the first scheme. Increasing ξ has a positive impact on the wealth returns at the cost of the terminal trading speeds being positive.

We noted the following differences in the performance of the two scheme:

1. The second scheme using dynamic updates of θ_{t_i} was more sensitive to the choice of

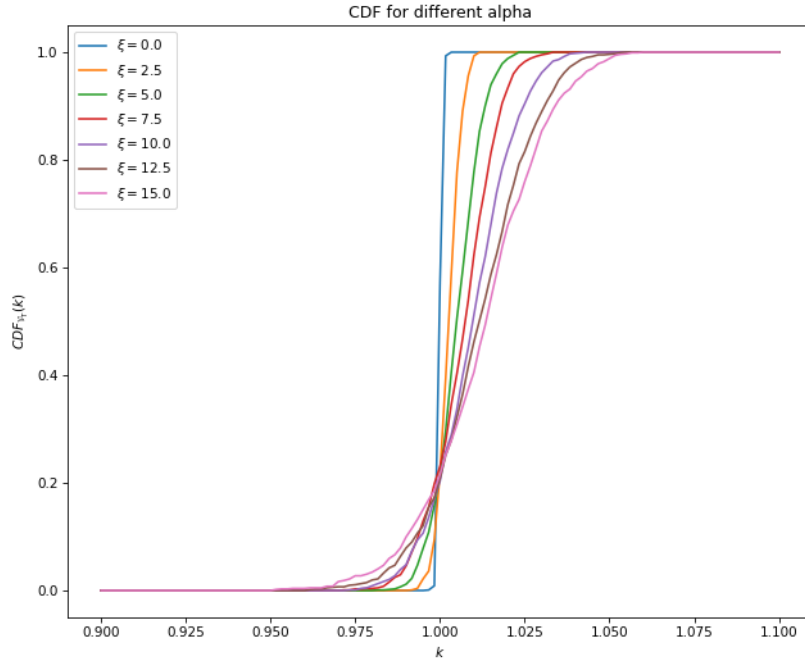


Figure 3.4: CDF of \mathcal{V}_t for different values of ξ .

parameter $\frac{1}{\gamma}$ compared to the scheme with fictitious play iterations. The reason for this is not clear.

2. Within the domain of favorable parameters $\frac{1}{\gamma}$ and ξ , the scheme with dynamic updates was much faster than the scheme based on fictitious play which is natural due to two factors: 1. Using the same NN parametrized by time for Z for the dynamic updates but multiple NNs for the fictitious play scheme; 2. Multiple fictitious play iterations for the first scheme.

3.5 Proofs

3.5.1 Proof of Theorem 3.3.3

Step 1: $J(q)$ is Gateaux differentiable $\forall q \in \mathcal{A}$.

Consider an arbitrary admissible control $p \in \mathcal{A}$. Then $\forall \varepsilon \in [0, 1]$, we can define an admissible control $q^\varepsilon = q + \varepsilon(p - q)$ (thanks to the convexity of \mathcal{A}) and let X^ε denote the solution of the corresponding state equation. The standard procedure is to consider a variational process which is the Gateaux derivative of the state process with respect to ε , but in our case, due to the linearity of the state SDE in state and control, the variational process is just the state process corresponding to the control $(p - q)$, i.e. the variational process D_t satisfies,

$$D_t := \lim_{\varepsilon \rightarrow 0} \frac{X_t^\varepsilon - X_t}{\varepsilon} = \lim_{\varepsilon \rightarrow 0} \frac{\varepsilon X_t^{p-q}}{\varepsilon} = X_t^{p-q}.$$

This can be verified from the linearity in (3.2.1) and the state SDE. Also, note that $\mathcal{V}_t^\varepsilon - \mathcal{V}_t^q = \varepsilon \mathcal{V}_t^{p-q}$.

Next, let us consider the Gateaux derivative of J at q in the direction $(p - q)$: We propose that

the Gateaux derivative is given as follows,

$$\partial J^p(q) := \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(\dot{g}_x(t) \cdot D_t + \dot{g}_q(t)(p_t - q_t) + \mathbb{E} \left[\partial_\mu h(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) \mathcal{V}_t^{p-q} \right] \right) dt \right]. \quad (3.5.1)$$

The integrability of ∂J^p will follow from the estimates proved below. We claim that

$$\lim_{\varepsilon \rightarrow 0} \frac{J(q^\varepsilon) - J(q)}{\varepsilon} = \partial J^p(q).$$

Using Taylor formula in an equality form, we write

$$\begin{aligned} \frac{J(q^\varepsilon) - J(q)}{\varepsilon} - \partial J^p(q) = & \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left((\dot{g}_x^\varepsilon(t) - \dot{g}_x(t)) \cdot D_t + (\dot{g}_q^\varepsilon(t) - \dot{g}_q(t)) \cdot (p_t - q_t) \right. \right. \\ & \left. \left. + \mathbb{E} \left[(\dot{h}_\mu^\varepsilon(t) - \dot{h}_\mu(t)) \mathcal{V}_t^{p-q} \right] \right) dt \right], \end{aligned} \quad (3.5.2)$$

where we set

$$\begin{aligned} \dot{g}_x^\varepsilon(t) &:= \int_0^1 \partial_x g(X_t^q + \lambda(X_t^\varepsilon - X_t^q), q_t^\varepsilon) d\lambda, \\ \dot{g}_q^\varepsilon(t) &:= \int_0^1 \partial_q g(X_t^q, q_t + \lambda\varepsilon(p_t - q_t)) d\lambda, \\ \dot{h}_\mu^\varepsilon(t) &:= \int_0^1 \partial_\mu h(\mu_{\mathcal{V}_t^{\lambda, \varepsilon}})(\mathcal{V}_t^{\lambda, \varepsilon}) d\lambda, \text{ where } \mathcal{V}_t^{\lambda, \varepsilon} := \mathcal{V}_t^q + \lambda(\mathcal{V}_t^\varepsilon - \mathcal{V}_t^q), \end{aligned}$$

and $\dot{g}_x(t)$, $\dot{g}_q(t)$ and $\dot{h}_\mu(t)$ denote their counter-parts with $\varepsilon = 0$.

i) Let us consider the first term in (3.5.2):

$$\left| \mathbb{E} \left[\int_0^\infty e^{-\rho t} (\dot{g}_x^\varepsilon(t) - \dot{g}_x(t)) \cdot D_t dt \right] \right| \leq \left(\mathbb{E} \left[\int_0^\infty e^{-\rho t} |\dot{g}_x^\varepsilon(t) - \dot{g}_x(t)|^2 dt \right] \right)^{\frac{1}{2}} \left(\mathbb{E} \left[\int_0^\infty e^{-\rho t} |D_t|^2 dt \right] \right)^{\frac{1}{2}}.$$

The second factor is finite, thanks to the a priori estimates. For the first factor, we can use the Lipschitz condition on the derivatives of g to obtain,

$$\begin{aligned} |\dot{g}_x^\varepsilon(t) - \dot{g}_x(t)| &\leq \int_0^1 \left| \partial_x g(X_t^q + \lambda(X_t^\varepsilon - X_t^q), q_t^\varepsilon) - \partial_x g(X_t^q, q_t) \right| d\lambda \\ &\leq \varepsilon L_f \left(|X_t^{p-q}| + |p_t - q_t| \right). \end{aligned}$$

Then, using the a priori estimate and the admissibility condition, it follows that,

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[\int_0^\infty e^{-\rho t} |\dot{g}_x^\varepsilon(t) - \dot{g}_x(t)|^2 dt \right] = 0,$$

hence the limit of the first term in (3.5.2) is zero.

ii) Next, the second term can be handled similarly.

iii) For the third term, we have,

$$\left| \mathbb{E} \left[\int_0^\infty e^{-\rho t} \mathbb{E} \left[(\dot{h}_\mu^\varepsilon(t) - \dot{h}_\mu(t)) \cdot \mathcal{V}_t^{p-q} \right] dt \right] \right|$$

$$\leq \left(\mathbb{E} \left[\int_0^\infty e^{-\rho t} |\dot{h}_\mu^\varepsilon(t) - \dot{h}_\mu(t)|^2 dt \right] \right)^{\frac{1}{2}} \left(\mathbb{E} \left[\int_0^\infty e^{-\rho t} |\mathcal{V}_t^{p-q}|^2 dt \right] \right)^{\frac{1}{2}},$$

which is similar to the first term, except for the Lipschitz requirement for the L-derivative of h , which follows from Proposition (3.2.4).

$$\mathbb{E} \left[|\partial_\mu h(\mu_{\mathcal{V}_t^\varepsilon})(\mathcal{V}_t^{\lambda,\varepsilon}) - \partial_\mu h(\mu_{\mathcal{V}_t^q})(cVt)|^2 \right] \leq L_h \mathbb{E} \left[|\mathcal{V}_t^{\lambda,\varepsilon} - \mathcal{V}_t^q|^2 \right],$$

which from the definition of $\mathcal{V}_t^{\lambda,\varepsilon}$, implies,

$$\mathbb{E} \left[|\partial_\mu h(\mu_{\mathcal{V}_t^\varepsilon})(\mathcal{V}_t^{\lambda,\varepsilon}) - \partial_\mu h(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q)|^2 \right] \leq L_h \mathbb{E} \left[|\lambda(\mathcal{V}_t^\varepsilon - \mathcal{V}_t^q)|^2 \right] = L_h \lambda^2 \varepsilon^2 \mathbb{E} \left[|\mathcal{V}_t^{p-q}|^2 \right].$$

From the a priori estimate (3.3.1) and using Fubini's theorem for interchanging expectation and integral, $\mathbb{E} \left[\int_0^\infty e^{-\rho t} (|\mathcal{V}_t^{p-q}|^2) dt \right] < \infty$. Then, using Fubini's theorem for interchanging expectation and integral, (3.5.1) gives,

$$\left| \mathbb{E} \left[\int_0^\infty e^{-\rho t} \mathbb{E} \left[(\dot{h}_\mu^\varepsilon(t) - \dot{h}_\mu(t)) \cdot \mathcal{V}_t^{p-q} \right] dt \right] \right| \leq \varepsilon \sqrt{\frac{L_h}{3}} \left(\mathbb{E} \left[\int_0^\infty e^{-\rho t} |\mathcal{V}_t^{p-q}|^2 dt \right] \right) \leq \varepsilon C,$$

for some constant C . Hence, the limit of the third term is zero as well.

This proves that the functional J is Gateaux differentiable and its Gateaux derivative in the direction p , evaluated at q is $\partial J^p(q)$ as defined earlier in (3.5.1). Since we assumed that q is optimal, the Gateaux derivative should be zero,

$$\partial J^p(q) := \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(\dot{g}_x(t) \cdot D_t + \dot{g}_q(t)(p_t - q_t) + \mathbb{E} \left[\partial_\mu h(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) \mathcal{V}_t^{p-q} \right] \right) dt \right]. \quad (3.5.3)$$

Step 2: Transformation of the Gateaux derivative using BSDE. Our aim is to express the formula (3.5.3) in terms of the infinite-horizon BSDE (3.3.1). So, let us use the Itô product rule for the process $e^{-\rho t} Y_t D_t$:

$$\begin{aligned} & \mathbb{E} \left[e^{-\rho T} Y_T D_T \right] \\ &= \mathbb{E} \left[\int_0^T e^{-\rho t} \left(-\rho Y_t D_t dt + Y_t dD_t + (dY_t) D_t + d\langle Y, D \rangle_t \right) \right] \\ &= \mathbb{E} \left[\int_0^T e^{-\rho t} \left(-\rho Y_t X_t^{p-q} + (Y_t A_t^x X_t^{p-q} + Y_t A_t^q (p_t - q_t)) + \left(\rho Y_t X_t^{p-q} - Y_t A_t^x X_t^{p-q} \right. \right. \right. \\ &\quad \left. \left. - \sum_{j=1}^m Z_t^j B_t^{x,j} X_t^{p-q} - \partial_x g(X_t^q, q_t) X_t^{p-q} - \partial_\mu h(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) \mathcal{V}_t^{p-q} \right) + \sum_{j=1}^m Z_t^j B_t^{x,j} X_t^{p-q} \right) dt \right] \\ &= \mathbb{E} \left[\int_0^T e^{-\rho t} \left(Y_t A_t^q (p_t - q_t) - \partial_x g(X_t^q, q_t) X_t^{p-q} - \partial_\mu h(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) \mathcal{V}_t^{p-q} \right) dt \right] \\ &= \mathbb{E} \left[\int_0^T e^{-\rho t} \left(Y_t A_t^q (p_t - q_t) - \partial_x g(X_t^q, q_t) X_t^{p-q} - \mathbb{E} \left[\partial_\mu h(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) \mathcal{V}_t^{p-q} \right] \right) dt \right], \end{aligned}$$

where the expectation of the martingale term coming from the first step is zero (using standard computations) and we have taken the expectation inside in the last equality. Next, take the limit as $T \rightarrow +\infty$ in the above identity:

$$|\mathbb{E} \left[e^{-\rho T} Y_T X_T^{p-q} \right]| \leq \mathbb{E} \left[e^{-\rho T} |Y_T|^2 \right]^{\frac{1}{2}} \mathbb{E} \left[e^{-\rho T} |X_T^{p-q}|^2 \right]^{\frac{1}{2}} \rightarrow 0,$$

where the limit of the first factor is zero using the terminal condition in (3.3.1) and the second factor has finite limit due to the a priori estimate (3.3.1). Hence, we have established, for any $p, q \in \mathcal{A}$,

$$\mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(Y_t A_t^q (p_t - q_t) - \partial_x g(X_t^q, q_t) X_t^{p-q} - \mathbb{E} \left[\partial_\mu h(\mu_{\mathcal{V}_t^q})(\mathcal{V}_t^q) \mathcal{V}_t^{p-q} \right] \right) dt \right] = 0,$$

and summing this with (3.5.3), we get

$$\partial J^p(q) = \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(Y_t A_t^q + \partial_q g(X_t^q, q_t) \right) (p_t - q_t) dt \right].$$

Step 3: necessary condition. Let $q \in \mathcal{A}$ be an optimal control for the above problem. For a perturbation in any direction p in the control space, the Gateaux derivative should be zero. Using this, we can re-write the $\partial J = 0$ condition in terms of the BSDE as,

$$\partial J^p(q) = \mathbb{E} \left[\int_0^\infty e^{-\rho t} \left(Y_t A_t^q + \partial_q g(X_t^q, q_t) \right) (p_t - q_t) dt \right] = 0, \quad \forall p \in \mathcal{A},$$

which can be written in terms of the Hamiltonian as,

$$\partial J^p(q) = \mathbb{E} \left[\int_0^\infty e^{-\rho t} \partial_q H(\eta_t) (p_t - q_t) dt \right] = 0, \quad \forall p \in \mathcal{A}.$$

Next, for the final step of the theorem, define the control p_t to take value $\nu \in \mathbb{R}^N$ (deterministic) in an arbitrarily chosen progressively measurable set $B \subset [0, \infty) \times \Omega$ and equal to q_t elsewhere. Clearly, admissibility of $(q_t)_{t \geq 0}$ implies the admissibility of p_t defined above. Hence, the above condition gives,

$$\mathbb{E} \left[\int_0^\infty \mathbf{1}_B \left(e^{-\rho t} \left(\partial_q H(\eta_t) \right) (\nu - q_t) \right) dt \right] = 0.$$

Since this holds for arbitrary B , we must have,

$$\partial_q H(\eta_t) (\nu - q_t) = 0, \quad dt \otimes d\mathbb{P} \text{ a.e.}$$

To prove this, assume contrarily that there exists a $dt \otimes d\mathbb{P}$ -measurable set D with positive measure such that $\partial_q H(\eta_t) (\nu - q_t) < 0$ (or > 0) on D . Then,

$$\mathbb{E} \left[\int_0^\infty \mathbf{1}_D e^{-\rho t} \left(\partial_q H(\eta_t) (\nu - q_t) \right) dt \right] < 0 \quad (\text{or } > 0 \text{ respectively})$$

which is a contradiction, hence proving the theorem. □

3.5.2 Proof of Theorem 3.3.4

We have proved the general result in Chapter 2 Theorem 2.4.1 but a proof in our setting is provided here for reader's convenience. Consider an arbitrary admissible control $(q'_t)_{t \geq 0} \in \mathcal{A}$ and X' the corresponding state process and μ' the law of the state process. Consider the following, for any $T > 0$

$$J(q') - J(q)$$

$$\begin{aligned}
&= J_T(q') - J_T(q) + \mathbb{E} \left[\int_0^T e^{-\rho t} \left(f(X'_t, q'_t, \mu'_t) - f(X_t^q, q_t, \mu_{\mathcal{V}_t^q}) \right) dt \right] \\
&= J_T(q') - J_T(q) + \mathbb{E} \left[\int_0^T e^{-\rho t} \left([H(X'_t, q'_t, \mu'_t, Y_t, Z_t) - H(X_t^q, q_t, \mu_{\mathcal{V}_t^q}, Y_t, Z_t)] \right) dt \right] \\
&\quad - \mathbb{E} \left[\int_0^T e^{-\rho t} \left(Y_t (A_t^x(X'_t - X_t^q) + A_t^q(q'_t - q_t)) + \sum_{j=1}^m Z_t^j B_t^{x,j}(X'_t - X_t^q) - \rho Y_t(X'_t - X_t^q) \right) dt \right],
\end{aligned}$$

where $J_T(q) = \mathbb{E} \left[\int_T^\infty e^{-\rho t} f(X_t^q, q_t, \mu_{\mathcal{V}_t^q}) dt \right]$. Next, using the adjoint equation and considering the Itô's product formula for $e^{-\rho t} Y_t(X'_t - X_t^q)$, we get,

$$d(e^{-\rho t} Y_t(X'_t - X_t^q)) = e^{-\rho t} (-\rho Y_t(X'_t - X_t^q) dt + (dY_t)(X'_t - X_t^q) + Y_t d(X'_t - X_t^q) + d\langle Y, (X' - X^q) \rangle_t),$$

$$\begin{aligned}
\mathbb{E} [e^{-\rho T} Y_T(X'_T - X_T^q)] &= \mathbb{E} \left[\int_0^T e^{-\rho t} (-\rho Y_t(X'_t - X_t^q) - [\partial_x H(\theta_t) + \mathcal{D}_\mu h_t](X'_t - X_t^q)) dt \right] \\
&\quad + \mathbb{E} \left[\int_0^T e^{-\rho t} \left(Y_t (A_t^x(X'_t - X_t^q) + A_t^q(q'_t - q_t)) + \sum_{j=1}^m Z_t^j B_t^{x,j}(X'_t - X_t^q) \right) dt \right],
\end{aligned}$$

where the expectation of the martingale terms coming from the first equality is zero. Next, using the minimization condition (3.3.3),

$$\partial_q H(t, X_t^q, q_t, \mu_{\mathcal{V}_t^q}, Y_t, Z_t) = 0,$$

and using (3.3.2), we have,

$$\begin{aligned}
J(q') - J(q) &= J_T(q') - J_T(q) + \mathbb{E} \left[\int_0^T e^{-\rho t} \left([H(X'_t, q'_t, \mu'_t, Y_t, Z_t) - H(X_t^q, q_t, \mu_{\mathcal{V}_t^q}, Y_t, Z_t)] \right) dt \right] \\
&\quad - \mathbb{E} [e^{-\rho T} Y_T(X'_T - X_T^q)] - \mathbb{E} \left[\int_0^T e^{-\rho t} ([\partial_x H(\theta_t) + \mathcal{D}_\mu h_t](X'_t - X_t^q)) dt \right] \\
&\geq J_T(q') - J_T(q) - \mathbb{E} [e^{-\rho T} Y_T(X'_T - X_T^q)] + \mathbb{E} \left[\int_0^T e^{-\rho t} (\alpha |q'_t - q_t|^2) dt \right].
\end{aligned}$$

Next, since q and q' are admissible, the integrands of $J(q)$ and $J(q')$ are integrable. Hence,

$$\lim_{T \rightarrow \infty} |J_T(q)| + |J_T(q')| = 0.$$

Next, using Cauchy-Schwartz inequality for expectation, we have,

$$\mathbb{E} [e^{-\rho T} |Y_T(X'_T - X_T^q)|] \leq \mathbb{E} [e^{-\rho T} |Y_T|^2]^{\frac{1}{2}} \mathbb{E} [e^{-\rho T} |X'_T - X_T^q|^2]^{\frac{1}{2}},$$

where the limit of the first factor is zero because of the terminal condition and second factor is finite from the a priori estimate of X .

Hence, we have,

$$J(q') - J(q) \geq \mathbb{E} \left[\int_0^T e^{-\rho t} (\alpha |q'_t - q_t|^2) dt \right] \geq 0,$$

implying that q is an optimal control and is unique $dt \otimes d\mathbb{P}$ a.s. □

Part II

BSDEs in infinite horizon

Chapter 4

Numerical approximation of Markovian BSDEs in infinite horizon and elliptic PDEs

Abstract. We study backward stochastic differential equations (BSDEs) in the infinite horizon and develop efficient numerical schemes for solving them. We establish a probabilistic representation of the solution of the BSDE using Malliavin derivative and prove results for contraction of a Picard scheme. We develop three numerical schemes, of which the first two are based on a fixed point argument using contraction imposing additional assumptions compared to what is needed for existence and uniqueness of the solution. The first scheme is a space grid based approximation where we establish tight numerical error bounds using a growth truncation argument and it performs well in low dimensions but computational times increase exponentially with dimension. The second scheme uses neural network approximations for which we have proved a convergence result. Using neural networks alleviates the curse of dimensionality, giving good accuracy in very high dimensions. The third scheme also uses neural networks but does not rely of contraction arguments, showcasing good performance even for larger α -Lipschitz outside the domain of contraction.

Contents

4.1	Introduction	102
4.1.1	Statement of the problem	104
4.1.2	Our contributions.	104
4.1.3	Notations	105
4.2	Analytical results	105
4.2.1	Model and value function	105
4.2.2	Time randomized Feynman-Kac representation	107
4.2.3	Contraction properties of the fixed point equation	108
4.2.4	Some sufficient assumptions to get $\kappa_\infty < 1$	113
4.2.5	How to check the contraction in L^p	116
4.3	Numerical schemes	118
4.3.1	A first scheme based on a regular grid	118
4.3.2	Some schemes based on neural networks approximation	122
4.3.3	Numerical experiments	125

4.4	Proofs	131
4.4.1	Proof of Proposition 4.2.3	131
4.4.2	Proof of Proposition 4.3.3	135
4.4.3	Proof of Lemma 4.3.1	137

4.1 Introduction

A BSDE is a type of SDE which evolves backwards in time, where one looks for an adapted pair of processes which solve an SDE with a given terminal condition. BSDEs were introduced in a linear setting by Bismut [23] and have emerged as a powerful mathematical framework for modeling and solving problems involving stochastic processes evolving backward in time. The genesis of BSDEs in the most general setting can be traced back to the seminal work of Pardoux and Peng [99], marking a significant advancement in the field of stochastic analysis. Many useful theoretical results have since been obtained, for example, the connection with semi-linear parabolic PDEs by Buckdahn et al. [33] who studied the homogenization properties and Darling and Pardoux [48] studied the connection between semi-linear elliptic PDEs and BSDEs with random terminal time, marking an important contribution to BSDEs with infinite horizon. Briand and Hu [30] studied the stability of BSDEs with random terminal time and applied his results to semi-linear elliptic PDEs. Major contributions in BSDEs include the existence and uniqueness of solutions, regularity properties, and connections with partial differential equations. El Karoui et al. [57] provided foundational insights, establishing a connection between BSDEs and quasilinear partial differential equations. The works of Soner et al. [112] and Briand and Confortola [29] further extended the understanding of stability and robustness in BSDEs. For existence and uniqueness of BSDEs with generators having quadratic growth, see Briand and Hu [31] and Delbaen, Hu, and Richou [51].

Numerical schemes for solving BSDEs have been investigated by many researchers over the years. Bally and Pagès [10] analysed the numerical error for the quantization method; Bouchard and Touzi [26] studied the convergence rates of approximations by regression using the kernel and the Malliavin approaches; Del Moral, Rémillard, and Rubenthaler [50] presented an estimator preserving the convexity and monotonicity for American options; Gobet and Labart [66] proposed a BSDE solver where an adaptive control variate is added making use of the connection between linear BSDEs and linear parabolic PDEs; Guyon and Henry-Labordère [74] provided a regression-based approximation for the optimal co-variance matrix within an uncertain volatility model. Among all above different approaches for applying Monte-Carlo methods, the most popular one is presumably the ordinary linear least-squares regression (OLS in short) which was developed by Ma, Protter, San Martín, and Torres [91] (see also [45]). Under this framework, in a finite time horizon, one first discretizes the time and then simulates paths for the processes and uses Monte-Carlo approach. The estimator is then computed backwards in the time grid as a linear combination of a reduced set of functions, finding coefficients which are optimal in the empirical least-squares sense.

Gobet, Lemor, and Warin [70] first extended OLS for the standard BSDEs and established the convergence rates for a dynamic programming based explicit scheme. Gobet and Lemor [67] developed approximations for BSDEs with jumps and for reflected BSDEs; Bender and Steiner [17] used a martingale basis function to reduce the propagation error when computing the conditional expectations; Bouchard and Warin [27] improvised the use of an adaptive local basis where the support of the function basis is adapted to the density of the samples; Bernhart, Pham, Tankov, and Warin [20] applied this procedure to approximate the penalized solution of BSDEs with constrained jump and they established a logarithmic convergence rate; Gobet and

his co-authors have done the error analysis of many different numerical estimators for solving BSDEs: the multi-step explicit scheme was studied in Gobet and Turkedjiev [69], the stratification sampling with parallelization using GPU was considered to improve the computational performance in Gobet et al. [71], the hedging strategy with Malliavin weight was presented in Gobet and Turkedjiev [68].

Over the years, the computational capabilities have advanced to an extent which has allowed researchers to study schemes based on deep learning methods for solving BSDEs in very high dimensions (~ 100), leveraging their power to mitigate the curse of dimensionality. E, Han, and Jentzen [56] were the first to apply a method which they called ‘Deep BSDE’ where a forward scheme is considered and the parameters of network are optimized to minimize the Mean Squared Error (MSE in short) on terminal condition by stochastic gradient descent; then Han and Long [77] extended the deep BSDE approach to solve a system of coupled FBSDE and demonstrated its performance for an explicitly solvable system in very high dimensions and recently, Carmona and Laurière [38] presented a similar approach for a Mean field control problem where the related approximation of a specific decoupled FBSDE of McKean-Vlasov type, see also Pham and Warin [106] for a suggestion for the network architecture.

While the literature concerning the simulation of (Markovian) BSDEs in the finite horizon setting is huge, there are few results concerning the infinite horizon setting. In the recent paper [15], authors provide a full-history recursive multilevel Picard approximation scheme for solving semi-linear elliptic PDE. Nevertheless, their framework does not contain possible non linearity with respect to the gradient of the solution: in terms of BSDEs, it means that the generator does not depend on z . A neural network approximation scheme for elliptic PDEs with a boundary is also provided in [96] but without a study of the numerical errors. Let us remark that one possible first attempt to tackle infinite horizon BSDE could be to consider a finite horizon approximation of the problem with a horizon T large enough and use a numerical approximation for the finite horizon BSDE. Nevertheless, numerical error bounds for the approximation of finite horizon BSDE deeply depend on the horizon time T - usually in an exponential way - which could lead to some logarithm speed of convergence.

In this chapter of the thesis, we obtained a fixed point representation for the solution for a BSDE in the infinite time horizon, which we used to design numerical schemes for solving such BSDEs. Due to the Markovian BSDE, we expect the solution (Y_t, Z_t) to be given as $v(X_t) = (u(X_t), \bar{u}(X_t))$ for all $t \geq 0$. We established the following fixed point representation for the solution,

$$\begin{aligned} u(x) := Y_0^x &= \mathbb{E} \left[\int_0^\infty e^{-as} (f(X_s^x, u(X_s^x), \bar{u}(X_s^x)) + au(X_s^x)) ds \right] \\ &= \frac{1}{\theta} \mathbb{E} \left[(f(X_E^x, u(X_E^x), \bar{u}(X_E^x)) + au(X_E^x)) e^{-(a-\theta)E} \right] \end{aligned}$$

and

$$\begin{aligned} \bar{u}(x) := Z_0^x &= \mathbb{E} \left[\int_0^\infty e^{-\tilde{a}s} (f(X_s^x, u(X_s^x), \bar{u}(X_s^x)) + \tilde{a}u(X_s^x)) U_s^x ds \right] \\ &= \frac{\sqrt{\pi}}{\sqrt{\tilde{\theta}}} \mathbb{E} \left[(f(X_{\tilde{E}}^x, u(X_{\tilde{E}}^x), \bar{u}(X_{\tilde{E}}^x)) + \tilde{a}u(X_{\tilde{E}}^x)) \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} U_{\tilde{E}}^x \right] \end{aligned}$$

where $E \sim \mathcal{E}(\theta)$ is an exponential random variable and $\tilde{E} \sim \Gamma(1/2, \tilde{\theta})$ is a gamma distribution. Due to the exponential weights in the integral, using an exponential random variable for time would be a natural choice for both u and \bar{u} . However, due to the Malliavin weight U_s^x in the representation of \bar{u} , there is an explosion of the order \sqrt{s} close to $s = 0$. In order to control it, we use a gamma distribution $\Gamma(1/2, \tilde{\theta})$. In Section 4.2.3, we studied the contraction properties

of this fixed point representation in order to obtain the additional assumptions we need for two of the numerical schemes which rely on the contraction. These assumptions depend on whether the generator for the BSDE depends on Z , as shown in Section 4.2.4. In Section 4.3.1, we provide an explicit framework and obtain tight bounds on the numerical error of the scheme. The numerical error depends on the assumptions we have for the growth of the state SDE X , i.e. we have different error bounds for a Brownian motion from an Orstein-Uhlenbeck process for instance. This scheme is based on grid approximation in space and hence prone to the curse of dimensionality. In Section 4.3.2, we first provide a neural network based scheme relying on contraction which allows us to tackle high dimensional settings. We must remark that the additional assumptions required for the two schemes relying on contraction limit us to problems where the growth of the generator with respect to Z process is not very high. This is the interest of the third scheme which does not rely on contraction.

4.1.1 Statement of the problem

We study the numerical approximation of the Markovian Backward Stochastic Differential Equation (BSDE) in infinite horizon

$$Y_t = Y_T + \int_t^T f(X_s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad \forall 0 \leq t \leq T < +\infty \quad (4.1.1)$$

where the solution process (Y, Z) takes values in some appropriate space and X is the solution of the following SDE:

$$X_t = x + \int_0^t b(X_s) ds + \int_0^t \sigma(X_s) dW_s, \quad 0 \leq t.$$

Some notations and assumptions are stated below.

4.1.2 Our contributions.

In this chapter, we study from a numerical viewpoint a general backward stochastic differential equation (BSDE) (4.1.1) in the infinite horizon setting, which is equivalent to solving the elliptic partial differential equation (PDE) (4.2.3). In Proposition 4.2.3, we establish a representation result for the unique solution of the BSDE (4.2.1) using the general assumptions for existence and uniqueness (see Proposition 4.2.1) where we write (Y, Z) as expectations by introducing a Malliavin weight (4.2.4). Based on this representation result, we propose three numerical schemes, two of which are based on Picard iterations and the third one is a direct scheme. The first two schemes rely on a contraction result given in Proposition 4.2.4, which requires up to impose additional assumptions beyond the ones needed for existence and uniqueness. Results for verifying the contraction are given in Section 4.2.4 and Section 4.2.5.

The first numerical scheme is based on a space grid-based approximation given by (4.3.1) which relies on a Picard scheme. We prove tight bounds on the numerical error given in Theorem 4.3.1. The scheme is highly efficient in low dimensions but suffers from the curse of dimensionality, which motivated us to design the second scheme with neural networks which are known to alleviate such problems. The second scheme is also based on a Picard scheme (see Algorithm 3 for a description of the numerical scheme). We proved a convergence result for the numerical error for the scheme in Proposition 4.3.4. However, for parameters close to and beyond the range for contraction, these two schemes show their shortcomings. The third scheme is based on a direct approach which solves this problem as it does not rely on a Picard iteration approach (see Algorithm 4 for the full scheme). Some numerical experiments are provided in Section 4.3.3.

4.1.3 Notations

In all this work, we consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ which supports a d -dimensional Brownian motion $W = (W^1, \dots, W^d)^\top$. The filtration $(\mathcal{F}_t)_{t \geq 0}$ is the one generated by W augmented by the \mathbb{P} -null sets, so that the filtration satisfies the "usual conditions".

Vector, matrix. $|x|$ denotes the Euclidean norm of $x \in \mathbb{R}^p$; when A is a matrix $\|A\|$ stands for the matrix 2-norm (i.e. subordinated to the Euclidean norm); x^\top denotes the transpose of the vector x and $\text{Tr}(A)$ denotes the trace of A . For a vector $x \in \mathbb{R}^p$ (resp. a matrix $A \in \mathbb{R}^{p \times q}$) and $R \in \mathbb{R}^+$, we denote $[x]_R$ the projection of x (resp. A) on the Euclidean ball $\bar{B}(0, R)$ of \mathbb{R}^p (resp. $\mathbb{R}^{p \times q}$). For $A = (c_1, \dots, c_q) \in \mathbb{R}^{p \times q}$ and $x \in \mathbb{R}^p$, we denote $\langle A, x \rangle$ the row-valued vector $(c_1^\top x, \dots, c_q^\top x)$

Function. $C^0(\mathbb{R}^p, \mathbb{R}^q)$ denotes the set of functions $f : \mathbb{R}^p \rightarrow \mathbb{R}^q$ that are continuous. $C^1(\mathbb{R}^p, \mathbb{R}^q)$ (resp. $C_b^1(\mathbb{R}^p, \mathbb{R}^q)$) denotes the set of functions $f : \mathbb{R}^p \rightarrow \mathbb{R}^q$ that are differentiable with a continuous (resp. continuous bounded) derivative. For a bounded function $f \in C^0(\mathbb{R}^p, \mathbb{R}^q)$ (resp. $C^0(\mathbb{R}^p, \mathbb{R}^{q \times r})$), we denote $|f|_\infty := \sup_{x \in \mathbb{R}^p} |f(x)|$ (resp. $\|f\|_\infty := \sup_{x \in \mathbb{R}^p} \|f(x)\|$).

For a function $f := (f_i)_{1 \leq i \leq p} \in C_b^1(\mathbb{R}^p, \mathbb{R}^q)$, we denote $\nabla_x f$ the function $\mathbb{R}^p \ni x \mapsto (\partial_{x_j} f_i)_{i,j} \in \mathbb{R}^{p \times q}$. In particular, when $p = 1$, $\nabla_x f$ is a row-vector valued function.

Random variables and stochastic processes. For $p \geq 1$, \mathbb{L}_p denotes the set of (scalar or vector-valued) random variables X with finite norm $|X|_p := (\mathbb{E}[|X|^p])^{1/p} < +\infty$. \mathbb{L}_∞ stands for the set of essentially bounded random variables.

\mathcal{S}_T^2 is the set of vector valued adapted continuous processes Y on $[0, T]$ such that

$$\mathbb{E} \left[\sup_{s \in [0, T]} |Y_s|^2 \right] < +\infty.$$

\mathcal{S}_{loc}^2 denotes the set of continuous processes Y on \mathbb{R}^+ such that $(Y_t)_{t \leq T} \in \mathcal{S}_T^2$, for all $T > 0$. \mathcal{M}_T^2 is the set of real matrix-valued predictable processes Z on $[0, T]$ such that

$$\mathbb{E} \left[\int_0^T \|Z_s\|^2 ds \right] < +\infty.$$

\mathcal{M}_{loc}^2 denotes the set of continuous processes Z on \mathbb{R}^+ such that $(Z_t)_{t \leq T} \in \mathcal{M}_T^2$, for all $T > 0$.

Specific distributions. For $\ell > 0$ and $a > 0$, we denote $\Gamma(a, \ell)$ the gamma distribution with density (with respect to the Lebesgue measure)

$$x \mapsto \frac{\ell^a}{\Gamma(a)} x^{a-1} e^{-\ell x} \mathbf{1}_{(0, +\infty)}(x).$$

We recall the scaling property between distributions $\Gamma(a, \ell) \stackrel{d}{=} \ell^{-1} \Gamma(a, 1)$.

We denote by F the cumulative distribution function of the Gaussian distribution $\mathcal{N}(0, 1)$.

4.2 Analytical results

4.2.1 Model and value function

We consider the following Markovian BSDE in infinite horizon

$$Y_t = Y_T + \int_t^T f(X_s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad \forall 0 \leq t \leq T < +\infty \quad (4.2.1)$$

where the solution process (Y, Z) takes values in the space $\mathcal{S}_{loc}^2 \times \mathcal{M}_{loc}^2$ and X is the solution of the following d -dimensional SDE:

$$X_t = x + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s, \quad 0 \leq t. \quad (4.2.2)$$

An existence and uniqueness result for the solution of (4.2.1) will be recall in Proposition (4.2.1). Throughout the paper, we take the following assumptions on f , b and σ .

Assumption 4.2.1. $f(\cdot, y, z)$ is continuous $\forall y \in \mathbb{R}^{d'}, z \in \mathbb{R}^{d' \times d}$ and there exists constants $r \geq 0$, $M_f \geq 0$, $\mu \in \mathbb{R}$, $K_{f,y} \geq 0$, $K_{f,z} \geq 0$, $K_b \geq 0$ and $K_\sigma \geq 0$ such that, $\forall x, x' \in \mathbb{R}^d, y, y' \in \mathbb{R}^{d'}, z, z' \in \mathbb{R}^{d' \times d}$,

1. $|f(x, y, z) - f(x, y', z')| \leq K_{f,y}|y - y'| + K_{f,z}\|z - z'\|$,
2. $|f(x, y, z)| \leq M_f(1 + |x|^r + |y| + \|z\|)$,
3. $\langle f(x, y, z) - f(x, y', z), y - y' \rangle \leq -\mu|y - y'|^2$,
4. $|b(x) - b(x')| \leq K_b|x - x'|$,
5. $\|\sigma(x) - \sigma(x')\| \leq K_\sigma|x - x'|$.

Let us remark that we can assume without restriction that $|\mu| \leq K_{f,y}$.

Since b and σ are Lipschitz, the SDE (4.2.2) has a unique solution for any starting point $x \in \mathbb{R}^d$: whenever necessary to emphasize on the x -dependence of the solution, we shall denote it by X^x and we denote (Y^x, Z^x) the solution of the associated BSDE.

Let us recall some known existence and uniqueness results for this BSDE.

Proposition 4.2.1. Let us assume that Assumption 4.2.1 is in force.

1. If there exists $0 < \rho < 2\mu - K_{f,z}^2$ such that, for all $x \in \mathbb{R}^d$,

$$\mathbb{E} \left[\int_0^{+\infty} e^{-\rho s} |X_s^x|^{2r} ds \right] < +\infty,$$

then, for all $x \in \mathbb{R}^d$, BSDE (4.2.1) has a unique solution $(Y^x, Z^x) \in \mathcal{S}_{loc}^2 \times \mathcal{M}_{loc}^2$ such that

$$\lim_{T \rightarrow +\infty} \mathbb{E} [e^{-\rho T} |Y_T^x|^2] = 0.$$

Moreover this solution satisfies

$$\mathbb{E} \left[\int_0^{+\infty} e^{-\rho s} \|Z_s^x\|^2 ds \right] + \mathbb{E} \left[\sup_{0 \leq s < +\infty} e^{-\rho s} |Y_s^x|^2 \right] \leq C \left(1 + \mathbb{E} \left[\int_0^{+\infty} e^{-\rho s} |X_s^x|^{2r} ds \right] \right),$$

with C a constant that does not depend on x .

2. If $d' = 1$ (i.e. Y^x is scalar valued) and $r = 0$ then, for all $x \in \mathbb{R}^d$, BSDE (4.2.1) has a solution such that Y^x is a continuous and bounded process and (Y^x, Z^x) satisfies

$$\text{ess sup}_{\Omega, t \in \mathbb{R}^+} |Y_t^x| + \mathbb{E} \left[\int_0^\infty e^{-2\mu s} \|Z_s^x\|^2 ds \right] \leq C,$$

with C a constant that does not depend on x . Moreover this solution is unique in the class of processes (Y^x, Z^x) such that Y^x is continuous and bounded and $Z^x \in \mathcal{M}_{loc}^2$.

For proofs, we refer to Theorem 5.57 in [98] for the multidimensional framework and Lemma 3.1 in [30] for the scalar setting.

Under suitable assumptions, the solution of this BSDE has a Markovian representation that gives a probabilistic representation of the following system of elliptic PDEs

$$\mathcal{L}u(x) + f(x, u(x), \nabla_x u(x)\sigma(x)) = 0, \quad \forall x \in \mathbb{R}^d. \quad (4.2.3)$$

where \mathcal{L} is the generator of the semi-group associated to the SDE (4.2.2).

Proposition 4.2.2. Let us assume same assumptions as in Proposition 4.2.1 and the continuity of function f . Then we have the Markovian representation $Y_t = u(X_t)$ where u is a continuous function. Moreover, if the i -th coordinate of f depends only on the i -th row of the matrix z then u is a viscosity solution of (4.2.3).

For the proof in the multidimensional setting we refer to Theorem 5.74 in [98]. For the case $d' = 1$, the proof can be easily adapted: In particular, the continuity follows directly from the stability result given by Theorem 3.4 in [30].

4.2.2 Time randomized Feynman-Kac representation

Now, we want to obtain a probabilistic representation for ∇u in terms of u and ∇u . In order to get it, we use the following extra assumptions on b and σ .

Assumption 4.2.2. We have $\sigma \in C_b^1(\mathbb{R}^d, \mathbb{R}^{d \times d})$ bounded with $\sup_{x \in \mathbb{R}^d} \|\sigma(x)\| = M_\sigma < +\infty$ and $b \in C_b^1(\mathbb{R}^d, \mathbb{R}^d)$. Furthermore, for all x , $\sigma(x)$ is invertible and σ^{-1} is bounded: $\sup_{x \in \mathbb{R}^d} \|\sigma^{-1}(x)\| = M_{\sigma^{-1}} < +\infty$.

Under Assumption 4.2.2, $(X_t^x)_{t \geq 0}$ is differentiable with respect to x , recalling e.g. [84]. We denote $(\nabla_x X_t^x)_{t \geq 0}$ its flow process and we define $(U_t^x)_{t \geq 0}$, a Malliavin weight (row vector valued) given by

$$U_t^x = \frac{1}{t} \int_0^t \langle \sigma^{-1}(X_s^x) \nabla_x X_s^x, dW_s \rangle \sigma(x) = (U_t^{x,j})_{1 \leq j \leq d}. \quad (4.2.4)$$

Proposition 4.2.3. Let us assume that there exists $0 < \rho < 2\mu - K_{f,z}^2 \mathbf{1}_{d' > 1}$ and $C > 0$ such that, for all $x \in \mathbb{R}^d$

$$\mathbb{E} \left[\int_0^{+\infty} e^{-\rho s} |X_s^x|^{2r} ds \right] \leq C(1 + |x|^{2r}). \quad (4.2.5)$$

We consider two parameters $a > \mu - \frac{K_{f,z}^2}{2} \mathbf{1}_{d' > 1} > 0$ and $\tilde{a} \geq 2\mu - K_{f,z}^2 \mathbf{1}_{d' > 1} > 0$ such that for any compact set $K \subset \mathbb{R}^d$ we have

$$\int_0^{+\infty} e^{-\tilde{a}s} \sup_{x \in K} \mathbb{E} [|X_s^x|^{2r}] ds + \sup_{x \in K} \mathbb{E} \left[\int_0^{+\infty} e^{-\tilde{a}s} \|\nabla_x X_s^x\|^2 ds \right] < +\infty. \quad (4.2.6)$$

Then, under extra Assumption 4.2.2, there exists a function $u \in C^1(\mathbb{R}^d, \mathbb{R}^{d'})$ such that, for all $x \in \mathbb{R}^d$, $Y^x = u(X^x)$, $\nabla_x u(X_x)\sigma(X^x) := \bar{u}(X^x)$ is a continuous version of Z^x and we have the following probabilistic representation for (u, \bar{u}) : for all $x \in \mathbb{R}^d$,

$$\begin{aligned} u(x) &:= Y_0^x = \mathbb{E} \left[\int_0^\infty e^{-as} (f(X_s^x, u(X_s^x), \bar{u}(X_s^x)) + au(X_s^x)) ds \right] \\ &= \frac{1}{\theta} \mathbb{E} \left[(f(X_E^x, u(X_E^x), \bar{u}(X_E^x)) + au(X_E^x)) e^{-(a-\theta)E} \right] \end{aligned} \quad (4.2.7)$$

and

$$\begin{aligned}\bar{u}(x) &:= Z_0^x = \mathbb{E} \left[\int_0^\infty e^{-\tilde{a}s} (f(X_s^x, u(X_s^x), \bar{u}(X_s^x)) + \tilde{a}u(X_s^x)) U_s^x ds \right] \\ &= \frac{\sqrt{\pi}}{\sqrt{\tilde{\theta}}} \mathbb{E} \left[(f(X_{\tilde{E}}^x, u(X_{\tilde{E}}^x), \bar{u}(X_{\tilde{E}}^x)) + \tilde{a}u(X_{\tilde{E}}^x)) \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} U_{\tilde{E}}^x \right]\end{aligned}\quad (4.2.8)$$

where $\theta > 0$, $\tilde{\theta} > 0$, $E \sim \mathcal{E}(\theta)$ is independent of W and $\tilde{E} \sim \Gamma(1/2, \tilde{\theta})$ is independent of W . Lastly, we have the following growth:

$$|u(x)| + \|\bar{u}(x)\| \leq C(1 + |x|^r), \quad \forall x \in \mathbb{R}^d. \quad (4.2.9)$$

The proof of Proposition 4.2.3 is postponed to Section 4.4.1.

Remark 4.2.1. We introduce these extra parameters θ and $\tilde{\theta}$ due to the theoretical study of the numerical error where we need sometimes $\theta \neq a$ and/or $\tilde{\theta} \neq b$. We take $\tilde{E} \sim \Gamma(1/2, \tilde{\theta})$ to control the explosion of order $t^{-1/2}$ at 0 of U_t^x .

Remark 4.2.2. By using the Lipschitz assumption on b and σ , and the boundedness on σ , standard estimates on X^x give us: for all $\varepsilon > 0$ and $p \geq 1$,

$$\mathbb{E}[|X_s^x|^{2p}] \leq C_\varepsilon(1 + |x|^{2p})e^{2p(K_b + \varepsilon)s}, \quad \mathbb{E}[\|\nabla_x X_s^x\|^{2p}] \leq Ce^{(2pK_b + p(2p-1)K_\sigma^2)s}, \quad \forall s \geq 0, x \in \mathbb{R}^d.$$

Thus (4.2.6) is satisfied as soon as $\tilde{a} > (2K_b + K_\sigma^2) \vee (2rK_b)$.

4.2.3 Contraction properties of the fixed point equation

We have established that (u, \bar{u}) solves equations (4.2.7) and (4.2.8) that can be seen as some fixed point equation. More precisely we define a map Φ such that, for all measurable functions $w : \mathbb{R}^d \rightarrow \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d}$ denoted by (w^1, w^2) , $\Phi(w)$ is a measurable function from \mathbb{R}^d to $\mathbb{R}^{d'} \times \mathbb{R}^{d' \times d}$ denoted by $(\Phi^1(w), \Phi^2(w))$, given as follows for all $x \in \mathbb{R}^d$,

$$\begin{aligned}\Phi(w)(x) &= \left(\frac{1}{\theta} \mathbb{E} \left[(f(X_E^x, w^1(X_E^x), w^2(X_E^x)) + aw_1(X_E^x)) e^{-(a-\theta)E} \right], \right. \\ &\quad \left. \sqrt{\frac{\pi}{\tilde{\theta}}} \mathbb{E} \left[(f(X_{\tilde{E}}^x, w^1(X_{\tilde{E}}^x), w^2(X_{\tilde{E}}^x)) + \tilde{a}w^1(X_{\tilde{E}}^x)) \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} U_{\tilde{E}}^x \right] \right).\end{aligned}$$

Then, equations (4.2.7) and (4.2.8) correspond to a fixed point of this map, $\Phi(v) = v$. Next, we study the contraction property of Φ . First, we have to define a suitable norm on $C^0(\mathbb{R}^d, \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d})$. Let $\rho : \mathbb{R}^+ \rightarrow [1, +\infty)$ be a positive non decreasing weight function, such that $\lim_{x \rightarrow +\infty} \rho(x) = +\infty$, with a growth at least polynomial of degree r and at most exponential: there exists $C > 0$ such that

$$1 + C^{-1}x^r \leq \rho(x) \leq Ce^{Cx}.$$

Some standard choices will correspond to polynomial or exponential weighting, i.e. $\rho(x) = (1 + x^\alpha)$ or $\rho(x) = \exp(\alpha x)$ for some parameter $\alpha \geq 0$. By a slight abuse of notation, we also define ρ on \mathbb{R}^d by setting $\rho(x) = \rho(|x|)$ for all $x \in \mathbb{R}^d$. The ρ -norm of a function $v : \mathbb{R}^d \rightarrow \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d}$ is defined by

$$\|u\|_\rho = \sup_{x \in \mathbb{R}^d} \frac{\|u(x)\|}{\rho(x)}.$$

Let us also define the following constants, which appear in the upcoming computations,

$$\begin{aligned} c_{\infty,(4.2.10)} &= \sup_{x \in \mathbb{R}^d} \mathbb{E} \left[\frac{e^{-(a-\theta)E} \rho(X_E^x)}{\theta \rho(x)} \right], \\ \tilde{c}_{\infty,(4.2.10)} &= \sup_{x \in \mathbb{R}^d} \mathbb{E} \left[\frac{\sqrt{\pi} \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} |U_{\tilde{E}}^x| \rho(X_{\tilde{E}}^x)}{\sqrt{\tilde{\theta}} \rho(x)} \right]. \end{aligned} \quad (4.2.10)$$

Let us also introduce the following assumption.

Assumption 4.2.3. There exists $\varepsilon > 0$ such that for any compact set $K \subset \mathbb{R}^d$,

$$\sup_{x \in K} \mathbb{E} [\rho(X_E^x)^{1+\varepsilon}] < \infty \text{ and } \sup_{x \in K} \mathbb{E} [\rho(X_{\tilde{E}}^x)^{1+\varepsilon}] < \infty.$$

We establish two preliminary lemmas before the study of the contraction of Φ .

Lemma 4.2.1. Let $h : \mathbb{R}^{d'} \rightarrow \mathbb{R}^{d'}$ a K -Lipschitz function that is $(-\mu)$ -monotone with $0 < \mu \leq K$. Then $h + aI_{d'}$ is $\sqrt{K^2 - 2\mu a + a^2}$ -Lipschitz.

Proof. For any $x, y \in \mathbb{R}^{d'}$, we just have to remark that

$$\begin{aligned} |h(y) + ay - (h(x) + ax)|^2 &= |h(y) - h(x)|^2 + 2a \langle h(y) - h(x), y - x \rangle + a^2 |y - x|^2 \\ &\leq |y - x|^2 (K^2 - 2\mu a + a^2), \end{aligned}$$

and the proof is finished. □

Lemma 4.2.2. Let $a > \theta$, $\tilde{a} > \tilde{\theta}$ and $v = (u, \bar{u}) \in C^0(\mathbb{R}^d, \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d})$ such that $\|v\|_{\rho} < +\infty$. Suppose $c_{\infty,(4.2.10)}, \tilde{c}_{\infty,(4.2.10)} < \infty$. If Assumption 4.2.3 is satisfied, then $\|\Phi(v)\|_{\rho} < +\infty$ and $\Phi(v)(\cdot) \in C^0(\mathbb{R}^d, \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d})$.

Proof. Let $x \in \mathbb{R}^d$. Using the growth assumption on f and the growth conditions on $\rho(x)$, $(1 + |X_E^x|^r \leq C\rho(|X_E^x|))$, we have the following:

$$\begin{aligned} &\frac{1}{\rho(x)} (\|\Phi^1(u, \bar{u})(x)\| + \|\Phi^2(u, \bar{u})(x)\|) \\ &\leq \frac{1}{\rho(x)} \left(\mathbb{E} [C(1 + |X_E^x|^r + |u(X_E^x)| + |\bar{u}(X_E^x)|)] + \mathbb{E} \left[C|U_E^x| \left(1 + |X_E^x|^r + |u(X_E^x)| + |\bar{u}(X_E^x)| \right) \right] \right) \\ &\leq C \left(\frac{1}{\rho(x)} + \mathbb{E} \left[\frac{|X_E^x|^r}{\rho(x)} \right] + \mathbb{E} \left[\frac{|u(X_E^x)|}{\rho(x)} \right] + \mathbb{E} \left[\frac{|\bar{u}(X_E^x)|}{\rho(x)} \right] \right. \\ &\quad \left. + \mathbb{E} \left[\frac{|U_E^x|}{\rho(x)} \right] + \mathbb{E} \left[\frac{|X_E^x|^r |U_E^x|}{\rho(x)} \right] + \mathbb{E} \left[\frac{|u(X_E^x)| |U_E^x|}{\rho(x)} \right] + \mathbb{E} \left[\frac{|\bar{u}(X_E^x)| |U_E^x|}{\rho(x)} \right] \right) \\ &\leq C \left(c_{\infty,(4.2.10)} + \mathbb{E} \left[\frac{|u(X_E^x)| \rho(X_E^x)}{\rho(X_E^x) \rho(x)} \right] + \mathbb{E} \left[\frac{|\bar{u}(X_E^x)| \rho(X_E^x)}{\rho(X_E^x) \rho(x)} \right] \right. \\ &\quad \left. + \tilde{c}_{\infty,(4.2.10)} + \mathbb{E} \left[\frac{|u(X_E^x)| |U_E^x| \rho(X_E^x)}{\rho(X_E^x) \rho(x)} \right] + \mathbb{E} \left[\frac{|\bar{u}(X_E^x)| |U_E^x| \rho(X_E^x)}{\rho(X_E^x) \rho(x)} \right] \right) \\ &\leq C(1 + \|u\|_{\rho} + \|\bar{u}\|_{\rho}) (c_{\infty,(4.2.10)} + \tilde{c}_{\infty,(4.2.10)}) < +\infty. \end{aligned}$$

Hence, $\|\Phi(v)\|_{\rho} < +\infty$ because the above constants are independent of x . Next, let us take a sequence $(x_n)_{n \in \mathbb{N}}$ converging to x in \mathbb{R}^d . We must prove that $\Phi(v)(x^n) \rightarrow \Phi(v)(x)$. Let us

consider the families of random variables $(\xi^n)_{n \in \mathbb{N}}$ and $(\zeta^n)_{n \in \mathbb{N}}$ defined as follows:

$$\begin{aligned}\xi^n &:= \left(f(X_E^{x^n}, u(X_E^{x^n}), \bar{u}(X_E^{x^n})) - f(X_E^x, u(X_E^x), \bar{u}(X_E^x)) + a(u(X_E^{x^n}) - u(X_E^x)) \right) e^{-(a-\theta)E}; \\ \zeta^n &:= \left(f(X_E^{x^n}, u(X_E^{x^n}), \bar{u}(X_E^{x^n})) - f(X_E^x, u(X_E^x), \bar{u}(X_E^x)) + \tilde{a}(u(X_E^{x^n}) - u(X_E^x)) \right) \times \\ &\quad \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} U_E^x.\end{aligned}$$

Using the continuity of f , u and \bar{u} , we can see that $\xi^n \rightarrow 0$ and $\zeta^n \rightarrow 0$ almost surely. Let us consider the following for a given $M > 0$:

$$\mathbb{E}[|\xi^n|] = \mathbb{E}\left[|\xi^n| \mathbf{1}_{|X_E^{x^n}| > M}\right] + \mathbb{E}\left[|\xi^n| \mathbf{1}_{|X_E^{x^n}| \leq M}\right].$$

The second term goes to zero as $n \rightarrow \infty$ using dominated convergence theorem since it is uniformly (in n) bounded by an integrable random variable. Since ρ is an increasing function, we have the following for the first term,

$$\begin{aligned}\mathbb{E}\left[|\xi^n| \mathbf{1}_{|X_E^{x^n}| > M}\right] &\leq \mathbb{E}\left[|\xi^n| \left(\frac{\rho(X_E^{x^n})}{\rho(M)}\right)^\varepsilon\right] \\ &\leq \mathbb{E}\left[C(1 + |X_E^{x^n}|^r + |X_E^x|^r) \left(\frac{\rho(X_E^{x^n})}{\rho(M)}\right)^\varepsilon\right] \\ &\leq \mathbb{E}\left[C(\rho(X_E^{x^n}) + \rho(X_E^x)) \left(\frac{\rho(X_E^{x^n})}{\rho(M)}\right)^\varepsilon\right] \\ &\leq \frac{C}{\rho(M)^\varepsilon} \mathbb{E}[\rho(X_E^{x^n})^{1+\varepsilon} + \rho(X_E^x)^{1+\varepsilon}] \\ &\leq \frac{C}{\rho(M)^\varepsilon},\end{aligned}$$

where the second inequality follows from the growth of f (Assumption 4.2.1), u and \bar{u} (Proposition 4.2.3), the fourth inequality comes from Young inequality and the last inequality follows from the Assumption 4.2.3 because $\{x_n, n \in \mathbb{N}\} \cup \{x\}$ is a compact set. Since the choice of M was arbitrary and ρ is an increasing function,

$$\lim_{n \rightarrow \infty} \mathbb{E}\left[|\xi^n| \mathbf{1}_{|\xi^n| > M}\right] = 0.$$

Hence, we get $\lim_{n \rightarrow +\infty} \mathbb{E}[|\xi^n|] = 0$. By same arguments we can also show that $\lim_{n \rightarrow +\infty} \mathbb{E}[|\zeta^n|] = 0$, proving the continuity of $\Phi(v)$ for a given continuous v . \square

Then we have the following estimate:

Proposition 4.2.4 (L_p -estimate, $1 < p \leq +\infty$).

▷ **The case $p = +\infty$.** Assume that the Assumption 4.2.3 holds, $a > \theta$, $\tilde{a} > \tilde{\theta}$ and $c_{\infty,(4.2.10)}, \tilde{c}_{\infty,(4.2.10)} < +\infty$. Then, for any functions v_1, v_2 in $C^0(\mathbb{R}^d, \mathbb{R}^d \times \mathbb{R}^{d' \times d})$, we have

$$\|\Phi(v_1) - \Phi(v_2)\|_\rho \leq \kappa_\infty \|v_1 - v_2\|_\rho,$$

where

$$\kappa_\infty = \sqrt{c_{\infty,(4.2.10)}^2 \left(\sqrt{K_{f,y}^2 - 2\mu a + a^2 + K_{f,z}}\right)^2 + \tilde{c}_{\infty,(4.2.10)}^2 \left(\sqrt{K_{f,y}^2 - 2\mu \tilde{a} + \tilde{a}^2 + K_{f,z}}\right)^2}.$$

▷ **The case $p \in (1, +\infty)$.** Assume that there exist finite positive constants $c_{p,(4.2.11)}, \tilde{c}_{p,(4.2.11)}, \tilde{c}_{p,(4.2.12)}$ and a probability measure π on \mathbb{R}^d such that for any measurable function v :

$\mathbb{R}^d \rightarrow \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d}$ with a polynomial growth r we have

$$\begin{aligned} \left(\int_{\mathbb{R}^d} \mathbb{E} [\|v(X_E^x)\|^p] \pi(dx) \right)^{1/p} &\leq c_{p,(4.2.11)} \left(\int_{\mathbb{R}^d} \|v(x)\|^p \pi(dx) \right)^{1/p} < +\infty, \\ \left(\int_{\mathbb{R}^d} \mathbb{E} [\|v(X_{\tilde{E}}^x)\|^p] \pi(dx) \right)^{1/p} &\leq \tilde{c}_{p,(4.2.11)} \left(\int_{\mathbb{R}^d} \|v(x)\|^p \pi(dx) \right)^{1/p} < +\infty, \end{aligned} \quad (4.2.11)$$

and

$$\sup_{x \in \mathbb{R}^d} \mathbb{E} \left[\left| \sqrt{\frac{\pi}{\theta}} U_{\tilde{E}}^x \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} \right|^{\frac{p}{p-1}} \right]^{p-1} = \tilde{c}_{p,(4.2.12)}. \quad (4.2.12)$$

Then, for any functions $v_1, v_2 \in C^0(\mathbb{R}^d, \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d})$, we have

$$\left(\int_{\mathbb{R}^d} \|\Phi(v_1(x)) - \Phi(v_2(x))\|^p \pi(dx) \right)^{\frac{1}{p}} \leq \kappa_p \left(\int_{\mathbb{R}^d} \|v_1(x) - v_2(x)\|^p \pi(dx) \right)^{\frac{1}{p}},$$

with

$$\begin{aligned} \kappa_p &= \left(\frac{1}{\theta} \left(\frac{p-1}{ap-\theta} \right)^{p-1} c_{p,(4.2.11)}^p \left(\sqrt{K_{f,y}^2 - 2\mu a + a^2} + K_{f,z} \right)^p \right. \\ &\quad \left. + \tilde{c}_{p,(4.2.12)} \tilde{c}_{p,(4.2.11)}^p \left(\sqrt{K_{f,y}^2 - 2\mu \tilde{a} + \tilde{a}^2} + K_{f,z} \right)^p \right)^{\frac{1}{p}}. \end{aligned}$$

Proof.

▷ Let $p = \infty$. We can assume that the functions v_1, v_2 and $v_1 - v_2$ have a finite ρ -norm, otherwise the above inequality is trivial. From Lemma 4.2.2, $\Phi(v_1)$ and $\Phi(v_2)$ have finite ρ -norms and belong to $C^0(\mathbb{R}^d, \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d})$. We denote $v_1 = (u_1, \bar{u}_1)$ and $v_2 = (u_2, \bar{u}_2)$. Using Lemma 4.2.1, we have for any $x \in \mathbb{R}^d$,

$$\begin{aligned} &\frac{|\Phi^1(v_1)(x) - \Phi^1(v_2)(x)|^2}{\rho(x)^2} \\ &\leq \frac{1}{\theta^2 \rho(x)^2} \mathbb{E} \left[e^{-(a-\theta)E} \left| f(X_E^x, u_1(X_E^x), \bar{u}_1(X_E^x)) - f(X_E^x, u_2(X_E^x), \bar{u}_2(X_E^x)) + a(u_1(X_E^x) - u_2(X_E^x)) \right|^2 \right] \\ &\leq (K_{f,y}^2 - 2\mu a + a^2) \mathbb{E} \left[\frac{|u_1(X_E^x) - u_2(X_E^x)|}{\rho(|X_E^x|)} \frac{\rho(|X_E^x|)}{\rho(x)} \frac{e^{-(a-\theta)E}}{\theta} \right]^2 \\ &\quad + K_{f,z}^2 \mathbb{E} \left[\frac{\|\bar{u}_1(X_E^x) - \bar{u}_2(X_E^x)\|}{\rho(|X_E^x|)} \frac{\rho(|X_E^x|)}{\rho(x)} \frac{e^{-(a-\theta)E}}{\theta} \right]^2 \\ &\leq c_{\infty,(4.2.10)}^2 \left(\sqrt{K_{f,y}^2 - 2\mu a + a^2} \|u_1 - u_2\|_{\rho} + K_{f,z} \|\bar{u}_1 - \bar{u}_2\|_{\rho} \right)^2 \\ &\leq c_{\infty,(4.2.10)}^2 \left(\sqrt{K_{f,y}^2 - 2\mu a + a^2} + K_{f,z} \right)^2 \|v_1 - v_2\|_{\rho}^2. \end{aligned}$$

Following same steps as above for Φ^2 , we obtain,

$$\frac{\|\Phi^2(v_1)(x) - \Phi^2(v_2)(x)\|^2}{\rho(x)^2} \leq \tilde{c}_{\infty,(4.2.10)}^2 \left(\sqrt{K_{f,y}^2 - 2\mu \tilde{a} + \tilde{a}^2} + K_{f,z} \right)^2 \|v_1 - v_2\|_{\rho}^2.$$

Therefore, the advertised result follows for $p = \infty$.

▷ Next, for $p \in (1, \infty)$, we can use Lemma 4.2.1 and Hölder inequality to obtain for the first component,

$$\begin{aligned} & \int_{\mathbb{R}^d} |\Phi^1(v_1)(x) - \Phi^1(v_2)(x)|^p \pi(dx) \\ & \leq \frac{1}{\theta^p} \int_{\mathbb{R}^d} \left(\sqrt{K_{f,y}^2 - 2\mu a + a^2} + K_{f,z} \right)^p \mathbb{E} \left[\|v_1(X_E^x) - v_2(X_E^x)\| e^{-(a-\theta)E} \right]^p \pi(dx) \\ & \leq \mathbb{E} \left[\left(\frac{e^{-(a-\theta)E}}{\theta} \right)^{\frac{p}{p-1}} \right]^{p-1} \int_{\mathbb{R}^d} \left(\sqrt{K_{f,y}^2 - 2\mu a + a^2} + K_{f,z} \right)^p \mathbb{E} [\|v_1(X_E^x) - v_2(X_E^x)\|^p] \pi(dx) \\ & \leq \frac{1}{\theta} \left(\frac{p-1}{ap-\theta} \right)^{p-1} c_{p,(4.2.11)}^p \left(\sqrt{K_{f,y}^2 - 2\mu a + a^2} + K_{f,z} \right)^p \int_{\mathbb{R}^d} \|v_1(x) - v_2(x)\|^p \pi(dx). \end{aligned}$$

Next, similarly for the second component,

$$\begin{aligned} & \int_{\mathbb{R}^d} \|\Phi^2(v_1)(x) - \Phi^2(v_2)(x)\|^p \pi(dx) \\ & \leq \left(\sqrt{\frac{\pi}{\theta}} \right)^p \left(\sqrt{K_{f,y}^2 - 2\mu \tilde{a} + \tilde{a}^2} + K_{f,z} \right)^p \int_{\mathbb{R}^d} \mathbb{E} \left[\|v_1(X_E^x) - v_2(X_E^x)\| \left| U_{\tilde{E}}^x \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} \right| \right]^p \pi(dx) \\ & \leq \left(\sqrt{\frac{\pi}{\theta}} \right)^p \left(\sqrt{K_{f,y}^2 - 2\mu \tilde{a} + \tilde{a}^2} + K_{f,z} \right)^p \times \\ & \quad \int_{\mathbb{R}^d} \mathbb{E} [\|v_1(X_{\tilde{E}}^x) - v_2(X_{\tilde{E}}^x)\|^p] \mathbb{E} \left[|U_{\tilde{E}}^x \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}}|^{\frac{p}{p-1}} \right]^{p-1} \pi(dx) \\ & \leq \left(\sqrt{K_{f,y}^2 - 2\mu \tilde{a} + \tilde{a}^2} + K_{f,z} \right)^p c_{p,(4.2.12)} c_{p,(4.2.11)}^p \int_{\mathbb{R}^d} \mathbb{E} [\|v_1(x) - v_2(x)\|^p] \pi(dx) \end{aligned}$$

where we have used Hölder's inequality. Thus, we obtain the announced result. □

Similarly, when there is no dependance on z in the generator one can prove a special case of this result, stated below.

Proposition 4.2.5. When the generator of the BSDE (4.2.1) does not depend on the Z process, that is $f(x, y, z) = f(x, y)$, conclusions of Proposition 4.2.4 remains true under same assumptions and with following contraction constants:

$$\kappa_\infty = c_{\infty,(4.2.10)} \sqrt{K_{f,y}^2 - 2\mu a + a^2}$$

and for $p \in [1, +\infty)$,

$$\kappa_p = \left(\frac{1}{\theta} \right)^{\frac{1}{p}} \left(\frac{p-1}{ap-\theta} \right)^{\frac{p-1}{p}} c_{p,(4.2.11)} \sqrt{K_{f,y}^2 - 2\mu a + a^2}.$$

Now a straightforward application of the Banach fixed-point theorem yields the following result.

Corollary 4.2.1. Let $p \in [1, +\infty]$. Consider the constants and the probability measure π from Proposition 4.2.4, and assume that

(case $p = +\infty$) : $\kappa_\infty < 1$ and $\|\Phi(0)\|_\rho < +\infty$,

(case $p < +\infty$) : $\kappa_p < 1$ and $\left(\int_{\mathbb{R}^d} \left(\frac{\|\Phi(0)(x)\|}{\rho(x)} \right)^p \pi(dx) \right)^{\frac{1}{p}} < +\infty$.

Then, there is a unique continuous function $v : \mathbb{R}^d \rightarrow \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d}$ with $\sup_{x \in \mathbb{R}^d} \frac{\|v(x)\|}{\rho(x)} < +\infty$

if $p = +\infty$ or $\left(\int_{\mathbb{R}^d} \left(\frac{\|v(x)\|}{\rho(x)} \right)^p \pi(dx) \right)^{\frac{1}{p}} < +\infty$ if $p < +\infty$, satisfying $v = \Phi(v)$, i.e. solving (4.2.7) and (4.2.8). Moreover the solution v satisfies the following bound:

$$\text{(case } p = +\infty \text{)} : \|v\|_\rho \leq \frac{1}{1 - \kappa_\infty} \|\Phi(0)\|_\rho,$$

$$\text{(case } p < +\infty \text{)} : \left(\int_{\mathbb{R}^d} \left(\frac{\|v(x)\|}{\rho(x)} \right)^p \pi(dx) \right)^{\frac{1}{p}} \leq \frac{1}{1 - \kappa_p} \left(\int_{\mathbb{R}^d} \left(\frac{\|\Phi(0)(x)\|}{\rho(x)} \right)^p \pi(dx) \right)^{\frac{1}{p}}.$$

Next, in order to check assumptions of Corollary 4.2.1 we must obtain some bounds on $c_{\infty,(4.2.10)}$, $\tilde{c}_{\infty,(4.2.10)}$, $c_{p,(4.2.11)}$, $\tilde{c}_{p,(4.2.11)}$ and $\tilde{c}_{p,(4.2.12)}$. This is the aim of following subsections. In the remaining we still assume Assumptions 4.2.1 and 4.2.2.

4.2.4 Some sufficient assumptions to get $\kappa_\infty < 1$

The aim of this subsection is to find some sets of assumptions on the generator, the terminal condition and the SDE such that we are allowed to apply Corollary 4.2.1, i.e. Φ is a contraction, recalling that we will use this contraction property to provide a numerical scheme with some convergence results in the next section. These assumptions are often stronger than those requested in Proposition 4.2.3. In particular, they depend on parameters a and \tilde{a} of the exponential and gamma distributions from Proposition 4.2.3 and they give some indications on how to choose optimally these parameters to ensure the contraction property of Φ . We recall that Assumptions 4.2.1 and 4.2.2 are fulfilled. Then, we use following notations:

- b is K_b -Lipschitz and there exist two non-negative constants M_b and L_b such that

$$|b(x)| \leq M_b + L_b|x|, \quad \forall x \in \mathbb{R}^d,$$

- σ is K_σ -Lipschitz and

$$\sup_{x \in \mathbb{R}^d} \|\sigma(x)\| \leq M_\sigma, \quad \sup_{x \in \mathbb{R}^d} \|\sigma^{-1}(x)\| \leq M_{\sigma^{-1}},$$

First, we establish the following elementary result:

Lemma 4.2.3. For all $\varepsilon > 0$ and $p \geq 2$, there exists C_ε that only depends on M_b, L_b, M_σ, p and ε , such that for all t, x , we have

$$\mathbb{E}[|X_t^x|^p] \leq (|x|^p + C_\varepsilon t) e^{(pL_b + \varepsilon)t}.$$

Proof. Applying Itô formula to $|X_t|^p$, we get

$$\begin{aligned} \mathbb{E}[|X_t|^p] &\leq |x|^p + \int_0^t p \mathbb{E}[|X_s|^{p-1} (M_b + L_b|X_s|)] ds + \frac{p(p-1)}{2} \int_0^t M_\sigma^2 \mathbb{E}[|X_s|^{p-2}] ds \\ &\leq |x|^p + C_\varepsilon t + \int_0^t (pL_b + \varepsilon) \mathbb{E}[|X_s|^p] ds \end{aligned}$$

and we just have to use Grönwall lemma to conclude. □

When f only depends on y : $f(x, y, z) = f(x, y)$. We consider (4.2.7) without the component \bar{u} , i.e.

$$u(x) = \frac{1}{\theta} \mathbb{E} \left[(f(X_E^x, u(X_E^x)) + au(X_E^x)) e^{-(a-\theta)E} \right].$$

Due to (4.2.9), it is natural to take a weight function with a growth at least equal to $1 + |x|^r$. Since this weight function will have an impact on the error study of our numerical scheme - to be precise, the error coming from the truncation of the domain, see Remark 4.3.3 - we take here $\rho(x) := 1 + |x|^{r'}$ with $r' \geq r$. Let us remark that the Assumption 4.2.3 holds because of Remark 4.2.2. Also, let us note that this choice implies $\|\Phi(0)\|_\rho < +\infty$ as soon as $c_{\infty,(4.2.10)} < +\infty$.

Proposition 4.2.6. For all $\varepsilon > 0$, $r' > 0$ and $a > r'L_b + \varepsilon$, we have

$$\kappa_\infty \leq \sqrt{K_{f,y}^2 - 2a\mu + a^2} \max \left\{ \frac{1}{a} + \frac{C_\varepsilon}{(a - (r'L_b + \varepsilon))^{1 + \frac{r'}{2} \wedge 1}}, \frac{1}{a - (r'L_b + \varepsilon)} \right\}$$

with C_ε that only depends on M_b, L_b, M_σ, p and ε . If $r' = 0$, we also have

$$\kappa_\infty \leq \frac{\sqrt{K_{f,y}^2 - 2a\mu + a^2}}{a}.$$

In particular, Φ is a contraction as soon as $\mu > r'L_b$ and a is large enough.

Proof. Using the Proposition 4.2.5, we just need a bound on $c_{\infty,(4.2.10)}$. If $r' \geq 2$, then we apply Lemma 4.2.3 to get, for $a > r'L_b + \varepsilon$,

$$\begin{aligned} c_{\infty,(4.2.10)} &\leq \sup_{x \in \mathbb{R}^d} \frac{\int_0^{+\infty} e^{-as} (1 + \mathbb{E}[|X_s^x|^{r'}]) ds}{1 + |x|^{r'}} \\ &\leq \sup_{x \in \mathbb{R}^d} \frac{\int_0^{+\infty} e^{-as} (1 + (C_\varepsilon s + |x|^{r'}) e^{(r'L_b + \varepsilon)s}) ds}{1 + |x|^{r'}} \\ &\leq \sup_{x \in \mathbb{R}^d} \frac{\frac{1}{a} + \frac{C_\varepsilon}{(a - (r'L_b + \varepsilon))^2} + \frac{|x|^{r'}}{a - (r'L_b + \varepsilon)}}{1 + |x|^{r'}} \\ &\leq \max \left\{ \frac{1}{a} + \frac{C_\varepsilon}{(a - (r'L_b + \varepsilon))^2}, \frac{1}{a - (r'L_b + \varepsilon)} \right\}. \end{aligned}$$

When $0 < r' < 2$, we just have to apply Jensen's to get

$$\begin{aligned} c_{\infty,(4.2.10)} &\leq \sup_{x \in \mathbb{R}^d} \frac{\int_0^{+\infty} e^{-as} (1 + \mathbb{E}[|X_s^x|^2]^{r'/2}) ds}{1 + |x|^{r'}} \\ &\leq \sup_{x \in \mathbb{R}^d} \frac{\frac{1}{a} + \int_0^{+\infty} (C_\varepsilon s + |x|^2)^{r'/2} e^{-(a - (r'L_b + \varepsilon))s} ds}{1 + |x|^{r'}} \\ &\leq \sup_{x \in \mathbb{R}^d} \frac{\frac{1}{a} + \frac{1}{a - (r'L_b + \varepsilon)} \left(\int_0^{+\infty} (a - (r'L_b + \varepsilon)) (C_\varepsilon s + |x|^2) e^{-(a - (r'L_b + \varepsilon))s} ds \right)^{r'/2}}{1 + |x|^{r'}} \end{aligned}$$

and we conclude as previously. When $r' = 0$ the result is obvious.

Finally, when $\mu > r'L_b$, we set $0 < \varepsilon < (\mu - r'L_b)/2$. For a large enough

$$\sqrt{K_{f,y}^2 - 2a\mu + a^2} = a - \mu + o(1) \leq a - \mu + \varepsilon,$$

and

$$\kappa_\infty \leq \frac{a - \mu + \varepsilon}{a - (r'L_b + \varepsilon)}$$

which is smaller than 1. □

Remark 4.2.3. • The assumption $\mu > rL_b$ is the condition required in Proposition 4.2.1 to get existence and uniqueness of a solution to the BSDE. If $\mu > rL_b$, then there exists $r' > r$ small enough such that $\mu > r'L_b$. Moreover, if b has a sublinear growth, then we can take L_b as close to 0 as we want and then there is no upper constraint on $r' \geq r$.

- In this special case, we do not need the smoothness of b , σ and the invertibility of σ since there is no Malliavin weight in (4.2.7).

When f depends on y and z . Due to (4.2.9), it is still natural to take the weight function $\rho(x) = 1 + |x|^{r'}$, for $x \in \mathbb{R}^d$ and $r' \geq r$. Once again, the Assumption 4.2.3 holds because of Remark 4.2.2 and this choice implies $\|\Phi(0)\|_\rho < +\infty$ as soon as $c_{\infty,(4.2.10)} < +\infty$.

Proposition 4.2.7. Let us assume that we have following growth estimates on X : there exist some constants $c_{1,(4.2.13)}$, $c_{2,(4.2.13)}$, $c_{3,(4.2.13)}$, $c_{4,(4.2.13)}$, $c_{5,(4.2.13)} \geq 0$ such that, for all $t \geq 0$,

$$\mathbb{E}[|X_t|^{2r'}]^{1/2} \leq (c_{1,(4.2.13)} + c_{2,(4.2.13)}|x|^{r'})e^{c_{3,(4.2.13)}t}, \quad \mathbb{E}[\|\nabla X_t\|^2]^{1/2} \leq c_{4,(4.2.13)}e^{c_{5,(4.2.13)}t}. \quad (4.2.13)$$

Then, for $a > c_{3,(4.2.13)}$ and $\tilde{a} > c_{3,(4.2.13)} + c_{5,(4.2.13)}$ we have following upper-bounds:

$$c_{\infty,(4.2.10)} \leq \max \left\{ \frac{1}{a} + \frac{c_{1,(4.2.13)}}{a - c_{3,(4.2.13)}}, \frac{c_{2,(4.2.13)}}{a - c_{3,(4.2.13)}} \right\}$$

and

$$\begin{aligned} & \tilde{c}_{\infty,(4.2.10)} \\ & \leq M_\sigma M_{\sigma^{-1}} c_{4,(4.2.13)} \sqrt{\pi} \max \left\{ \frac{1}{\sqrt{\tilde{a} - c_{5,(4.2.13)}}} + \frac{c_{1,(4.2.13)}}{\sqrt{\tilde{a} - c_{5,(4.2.13)} - c_{3,(4.2.13)}}}, \frac{c_{2,(4.2.13)}}{\sqrt{\tilde{a} - c_{5,(4.2.13)} - c_{3,(4.2.13)}}} \right\}. \end{aligned}$$

Proof. Let us start to study $c_{\infty,(4.2.10)}$. We have

$$\begin{aligned} \mathbb{E} \left[\frac{e^{-(a-\theta)E}}{\theta} (1 + |X_E^x|^{r'}) \right] &= \int_0^{+\infty} e^{-at} \left(1 + \mathbb{E} \left[|X_t^x|^{r'} \right] \right) dt \leq \int_0^{+\infty} e^{-at} \left(1 + \mathbb{E} \left[|X_t^x|^{2r'} \right]^{1/2} \right) dt \\ &\leq \frac{1}{a} + \frac{c_{1,(4.2.13)}}{a - c_{3,(4.2.13)}} + \frac{c_{2,(4.2.13)}|x|^{r'}}{a - c_{3,(4.2.13)}}. \end{aligned}$$

Thus, we easily obtain that

$$c_{\infty,(4.2.10)} \leq \max \left\{ \frac{1}{a} + \frac{c_{1,(4.2.13)}}{a - c_{3,(4.2.13)}}, \frac{c_{2,(4.2.13)}}{a - c_{3,(4.2.13)}} \right\}.$$

Now, we study $\tilde{c}_{\infty,(4.2.10)}$. We have

$$\begin{aligned} & \mathbb{E} \left[\sqrt{\frac{\pi}{\tilde{\theta}}} \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} |U_{\tilde{E}}^x| (1 + |X_{\tilde{E}}^x|^{r'}) \right] \\ &= \int_0^{+\infty} e^{-\tilde{a}t} \mathbb{E} \left[|U_t^x| (1 + |X_t^x|^{r'}) \right] dt \\ &\leq \int_0^{+\infty} e^{-\tilde{a}t} \mathbb{E} \left[|U_t^x|^2 \right]^{1/2} \left(1 + \mathbb{E} \left[|X_t^x|^{2r'} \right]^{1/2} \right) dt \\ &\leq \int_0^{+\infty} e^{-\tilde{a}t} \frac{M_\sigma M_{\sigma^{-1}}}{\sqrt{t}} \mathbb{E} \left[\frac{1}{t} \int_0^t \|\nabla_x X_s\|^2 ds \right]^{1/2} \left(1 + (c_{1,(4.2.13)} + c_{2,(4.2.13)}|x|^{r'})e^{c_{3,(4.2.13)}t} \right) dt \\ &\leq M_\sigma M_{\sigma^{-1}} c_{4,(4.2.13)} \int_0^{+\infty} e^{-\tilde{a}t} \frac{e^{c_{5,(4.2.13)}t}}{\sqrt{t}} \left(1 + (c_{1,(4.2.13)} + c_{2,(4.2.13)}|x|^{r'})e^{c_{3,(4.2.13)}t} \right) dt \end{aligned}$$

$$\leq M_\sigma M_{\sigma^{-1}} c_{4,(4.2.13)} \sqrt{\pi} \left(\frac{1}{\sqrt{\tilde{a} - c_{5,(4.2.13)}}} + \frac{c_{1,(4.2.13)}}{\sqrt{\tilde{a} - c_{5,(4.2.13)} - c_{3,(4.2.13)}}} + \frac{c_{2,(4.2.13)} |x|^{r'}}{\sqrt{\tilde{a} - c_{5,(4.2.13)} - c_{3,(4.2.13)}}} \right)$$

which gives us the announced result. \square

Remark 4.2.4. Let us take $a = \tilde{a} = K_{f,y}$ and denote $\varepsilon = K_{f,y} - \mu \geq 0$. If $K_{f,y} > c_{3,(4.2.13)} + c_{5,(4.2.13)}$, then

$$\kappa_\infty \leq \sqrt{c_{\infty,(4.2.10)}^2 + \tilde{c}_{\infty,(4.2.10)}^2} \left(\sqrt{2\varepsilon K_{f,y}} + K_{f,z} \right) < 1$$

as soon as ε and $K_{f,z}$ are small enough.

4.2.5 How to check the contraction in L^p

Proposition 4.2.8. Let us assume the growth assumption on $(\nabla X_t)_{t \geq 0}$ given by the right hand side in (4.2.13), that is to say

$$\mathbb{E}[\|\nabla X_t\|^2]^{1/2} \leq c_{4,(4.2.13)} e^{c_{5,(4.2.13)} t}, \quad t \geq 0.$$

If $\tilde{a}p - \tilde{\theta} - c_{5,(4.2.12)}p > 0$, then we have, for $p \geq 2$,

$$\tilde{c}_{p,(4.2.12)} \leq \frac{(\sqrt{\pi} M_\sigma M_{\sigma^{-1}} c_{4,(4.2.12)})^p (p-1)^{\frac{p-1}{2}}}{\sqrt{\tilde{\theta}} (\tilde{a}p - \tilde{\theta} - c_{5,(4.2.12)}p)^{\frac{p-1}{2}}}.$$

Proof.

$$\sup_{x \in \mathbb{R}^d} \mathbb{E} \left[\left| \sqrt{\frac{\pi}{\tilde{\theta}}} U_{\tilde{E}}^x \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} \right|^{\frac{p}{p-1}} \right]^{p-1} \leq \tilde{c}_{p,(4.2.12)}.$$

We have

$$\begin{aligned} \mathbb{E} \left[\left| \sqrt{\frac{\pi}{\tilde{\theta}}} U_{\tilde{E}}^x \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} \right|^{\frac{p}{p-1}} \right] &= \int_0^{+\infty} \mathbb{E} \left[\left| \sqrt{\frac{\pi}{\tilde{\theta}}} U_t^x \sqrt{t} e^{-(\tilde{a}-\tilde{\theta})t} \right|^{\frac{p}{p-1}} \right] \sqrt{\frac{\tilde{\theta}}{\pi t}} e^{-\tilde{\theta}t} dt \\ &\leq \left(\frac{\pi}{\tilde{\theta}} \right)^{\frac{1}{2(p-1)}} \int_0^{+\infty} \mathbb{E} \left[|U_t^x|^{\frac{p}{p-1}} \right] t^{\frac{1}{2(p-1)}} e^{-\frac{\tilde{a}p}{p-1}t + \frac{\tilde{\theta}}{p-1}t} dt \end{aligned}$$

Since $p \geq 2$, then $\frac{p}{p-1} \leq 2$ and Jensen inequality gives us, for all $t > 0$,

$$\begin{aligned} \mathbb{E} \left[|U_t^x|^{\frac{p}{p-1}} \right] &\leq \mathbb{E} \left[|U_t^x|^2 \right]^{\frac{p}{2(p-1)}} \leq \frac{(M_\sigma M_{\sigma^{-1}})^{\frac{p}{p-1}}}{t^{\frac{p}{2(p-1)}}} \left(\frac{1}{t} \int_0^t \mathbb{E} [\|\nabla X_s\|^2] ds \right)^{\frac{p}{2(p-1)}} \\ &\leq \frac{(M_\sigma M_{\sigma^{-1}})^{\frac{p}{p-1}}}{t^{\frac{p}{2(p-1)}}} (c_{4,(4.2.12)} e^{c_{5,(4.2.12)} t})^{\frac{p}{p-1}} \end{aligned}$$

It follows that

$$\begin{aligned} &\mathbb{E} \left[\left| \sqrt{\frac{\pi}{\tilde{\theta}}} U_{\tilde{E}}^x \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} \right|^{\frac{p}{p-1}} \right] \\ &\leq \left(\frac{\pi (M_\sigma M_{\sigma^{-1}} c_{4,(4.2.12)})^{2p}}{\tilde{\theta}} \right)^{\frac{1}{2(p-1)}} \int_0^{+\infty} \frac{1}{\sqrt{t}} e^{-\frac{\tilde{a}p}{p-1}t + \frac{\tilde{\theta}}{p-1}t + \frac{c_{5,(4.2.12)}^p}{p-1}t} dt \end{aligned}$$

$$= \left(\frac{\pi(M_\sigma M_{\sigma^{-1}} c_{4,(4.2.12)})^{2p}}{\tilde{\theta}} \right)^{\frac{1}{2(p-1)}} \frac{\sqrt{\pi(p-1)}}{\sqrt{\tilde{a}p - \tilde{\theta} - c_{5,(4.2.12)}p}}$$

and finally we obtain

$$\sup_{x \in \mathbb{R}^d} \mathbb{E} \left[\left| \sqrt{\frac{\pi}{\tilde{\theta}}} U_E^x \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} \right|^{\frac{p}{p-1}} \right]^{p-1} \leq \frac{(\sqrt{\pi} M_\sigma M_{\sigma^{-1}} c_{4,(4.2.12)})^p (p-1)^{\frac{p-1}{2}}}{\sqrt{\tilde{\theta}} (\tilde{a}p - \tilde{\theta} - c_{5,(4.2.12)}p)^{\frac{p-1}{2}}}$$

which gives us the announced result. \square

Remark 4.2.5. If $p < 2$, it is also possible to get an upper-bound for $\tilde{c}_{p,(4.2.12)}$ by using same computations if we have a deterministic upper-bound for ∇X_t which is the case when the SDE has an additive noise.

It remains to prove the existence of $c_{p,(4.2.11)}$ and $\tilde{c}_{p,(4.2.11)}$. The following proposition treat the special Brownian case.

Proposition 4.2.9. Let us assume that $X = W$ and let us consider $\pi(dx) = \frac{c}{(1+|x|)^{pr+d+1}} dx$ where c is such that π is a density measure on \mathbb{R}^d . Then we can take

$$c_{p,(4.2.11)}^p = \int_0^{+\infty} (t \vee 1)^{\frac{pr+d+1}{2}} \theta e^{-\theta t} dt \int_{\mathbb{R}^d} (1+|y|)^{pr+d+1} \frac{1}{(2\pi)^{d/2}} e^{-\frac{|y|^2}{2}} dy < +\infty$$

and

$$\tilde{c}_{p,(4.2.11)}^p = \int_0^{+\infty} \frac{(t \vee 1)^{\frac{pr+d+1}{2}} \sqrt{\tilde{\theta}}}{\sqrt{t}} \frac{\sqrt{\tilde{\theta}}}{\sqrt{\pi}} e^{-\tilde{\theta}t} dt \int_{\mathbb{R}^d} (1+|y|)^{pr+d+1} \frac{1}{(2\pi)^{d/2}} e^{-\frac{|y|^2}{2}} dy < +\infty.$$

Proof. Firstly, we remark that

$$\int_{\mathbb{R}^d} \|v(x)\|^p \pi(dx) \leq C \int_{\mathbb{R}^d} \frac{(1+|x|)^{pr}}{(1+|x|)^{pr+d+1}} dx \leq C \int_{\mathbb{R}^d} \frac{1}{(1+|x|)^{d+1}} dx < +\infty.$$

Now we compute

$$\begin{aligned} \int_{\mathbb{R}^d} \mathbb{E} [\|v(X_E^x)\|^p] \pi(dx) &= \int_0^{+\infty} \int_{\mathbb{R}^{2d}} \|v(x + \sqrt{t}y)\|^p \frac{1}{(2\pi)^{d/2}} e^{-\frac{|y|^2}{2}} \frac{c}{(1+|x|)^{pr+d+1}} \theta e^{-\theta t} dx dy dt \\ &= \int_0^{+\infty} \int_{\mathbb{R}^{2d}} \|v(z)\|^p \frac{1}{(2\pi)^{d/2}} e^{-\frac{|y|^2}{2}} \frac{c}{(1+|z - \sqrt{t}y|)^{pr+d+1}} \theta e^{-\theta t} dz dy dt. \end{aligned}$$

Since we have

$$\frac{(1+|z|)^{pr+d+1}}{(1+|z - \sqrt{t}y|)^{pr+d+1}} \leq \left(\frac{1+|z - \sqrt{t}y| + \sqrt{t}|y|}{1+|z - \sqrt{t}y|} \right)^{pr+d+1} \leq (1 + \sqrt{t}|y|)^{pr+d+1},$$

then

$$\begin{aligned} &\int_{\mathbb{R}^d} \mathbb{E} [\|v(X_E^x)\|^p] \pi(dx) \\ &\leq \int_{\mathbb{R}^d} \|v(z)\|^p \pi(dz) \int_0^{+\infty} \int_{\mathbb{R}^d} (1 + \sqrt{t}|y|)^{pr+d+1} \frac{1}{(2\pi)^{d/2}} e^{-\frac{|y|^2}{2}} \theta e^{-\theta t} dy dt \\ &\leq \int_{\mathbb{R}^d} \|v(z)\|^p \pi(dz) \int_0^{+\infty} (t \vee 1)^{\frac{pr+d+1}{2}} \theta e^{-\theta t} dt \int_{\mathbb{R}^d} (1+|y|)^{pr+d+1} \frac{1}{(2\pi)^{d/2}} e^{-\frac{|y|^2}{2}} dy \end{aligned}$$

which gives us the first announced result. The second estimate follows same computations. \square

Remark 4.2.6. When X is not the Brownian motion, it would be possible to use some Aronson estimates (in the uniformly elliptic case) in order to get the same kind of bound for $c_{p,(4.2.11)}$ and $\tilde{c}_{p,(4.2.11)}$. In particular, it could be proved using the same approach as in the proof of Theorem 1 in Gobet, López-Salas, and Vázquez [72].

4.3 Numerical schemes

The aim of this section is to define some schemes in order to provide a numerical approximation of the function $v = (u, \bar{u})$ solution of (4.2.7) and (4.2.8). The first proposed scheme relies on the contraction property given by Proposition 4.2.4 and a space discretization through a regular grid. A full study of this scheme error is obtained, see Theorem 4.3.1. Then, we also provide a second and a third scheme based on neural networks approximation.

4.3.1 A first scheme based on a regular grid

We denote Π a non empty finite subgrid of $\delta\mathbb{Z}^d$, N its cardinality, $\delta > 0$ its mesh size and \square its convex hull (in \mathbb{R}^d). Without loss of generality we can assume that $0 \in \square$. In order to define a multilinear interpolation procedure on Π , we consider the following basis functions:

$$\forall z \in \Pi, \quad \forall x \in \mathbb{R}^d, \quad \psi_z(x) := \prod_{i=1}^d (1 - |\delta^{-1}(x_i - z_i)|)_+,$$

where $(\cdot)_+$ denotes the positive part function. For a function $\phi : \Pi \rightarrow \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d}$, we define $P\phi$ the interpolation of ϕ on \mathbb{R}^d as follows

$$P\phi(x) =: \begin{cases} \sum_{z \in \Pi} \psi_z(x) \phi(z), & x \in \square \\ \sum_{z \in \Pi} \psi_z(\text{Proj}(x, \square)) \phi(z), & x \notin \square. \end{cases}$$

By a small abuse of notation we also consider this interpolation operator for functions $\phi : \mathbb{R}^d \rightarrow \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d}$ by defining $P\phi := P\phi|_{\Pi}$. By the same definition, P can act also on vector-valued and matrix-valued functions.

This interpolation operator satisfies following standard properties.

Proposition 4.3.1. Let us consider $\phi, \tilde{\phi} : \mathbb{R}^d \rightarrow \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d}$. Then we have

1.

$$\|P\phi - P\tilde{\phi}\|_{\infty} \leq \|\phi - \tilde{\phi}\|_{\infty}.$$

2. If $\phi \in C^2(\mathbb{R}^d, \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d})$, then there exists $C \geq 0$ only depending on $|\nabla^2 \phi|_{\infty, \square}$ such that

$$\|P\phi - \phi\|_{\infty, \square} \leq C\delta^2.$$

3.

$$\|P\phi(x)\| \leq P\rho(x) \sup_{z \in \Pi} \left\| \frac{\phi(z)}{\rho(z)} \right\|, \quad \forall x \in \mathbb{R}^d.$$

4.

$$\sup_{x \in \mathbb{R}^d} \left| \frac{P\rho(x)}{\rho(x)} \right| \leq \sup_{x \in \square, y \in \square, |y-x| \leq \sqrt{d}\delta} \frac{\rho(y)}{\rho(x)}.$$

Proof. The two first points are standard. For the third point, we have, for all $x \in \mathbb{R}^d$,

$$\begin{aligned} \|P\phi(x)\| &= \left\| \sum_{z \in \Pi} \psi_z(\text{Proj}(x, \square)) \phi(z) \right\| \leq \sum_{z \in \Pi} \psi_z(\text{Proj}(x, \square)) \rho(z) \left\| \frac{\phi(z)}{\rho(z)} \right\| \\ &\leq \sup_{z \in \Pi} \left\| \frac{\phi(z)}{\rho(z)} \right\| \sum_{z' \in \Pi} \psi_{z'}(\text{Proj}(x, \square)) \rho(z') = P\rho(x) \sup_{z \in \Pi} \left\| \frac{\phi(z)}{\rho(z)} \right\|. \end{aligned}$$

So, it just remains to prove the fourth point. Since $0 \in \square$ and ρ is an increasing function with respect to $|x|$, then $\frac{P\rho(x)}{\rho(x)} \leq 1$ for all $x \notin \square$. When $x \in \square$,

$$|P\rho(x)| \leq \sup_{z \in \Pi, \sup_{1 \leq j \leq d} |x_j - z_j| < \delta} \rho(z) \leq \sup_{y \in \square, |y-x| < \sqrt{d}\delta} \rho(y)$$

which proves the result. \square

We also define a growth truncation operator T_{R, ρ_r} by

$$T_{R, \rho_r} : \phi \mapsto \left[\frac{\phi}{\rho_r} \right]_R \rho_r$$

for any function $\phi : \mathbb{R}^d \rightarrow \mathbb{R}^d \times \mathbb{R}^{d' \times d}$ or $\phi : \Pi \rightarrow \mathbb{R}^d \times \mathbb{R}^{d' \times d}$, where $[\cdot]_R$ denotes the projection onto the Euclidean ball $\bar{B}(0, R)$ of $\mathbb{R}^d \times \mathbb{R}^{d' \times d}$ and $\rho_r(x) = 1 + |x|^r$.

Definition of the scheme.

We are now able to define our scheme in the next definition. The idea is to use the Picard iteration, to use a spatial approximation of functions onto Π and to approximate expectations by empirical means. The sample size of empirical means will be denoted by M .

Definition 4.3.1. We construct a sequence of random functions $v_M^n : \Omega \times \Pi \rightarrow \mathbb{R}^d \times \mathbb{R}^{d' \times d}$, $n \in \mathbb{N}$ such that $v_M^0 = 0$ and, for all $n \in \mathbb{N}$, $z \in \Pi$,

$$v_M^{n+1}(z) = T_{B, \rho_r} \left(\frac{1}{M} \sum_{j=1}^M R_{n,j}^{(\cdot)}(Pv_M^n) \right) (z), \quad (4.3.1)$$

where $B \geq \|v\|_{\rho_r}$ and for any $\phi = (w^1, w^2) : \mathbb{R}^d \rightarrow \mathbb{R}^d \times \mathbb{R}^{d' \times d}$ and $x \in \Pi$, $(R_{n,j}^x(\phi))_{n \in \mathbb{N}, j \in \mathbb{N}^*}$ are i.i.d. random variables with the same distribution as

$$R^x(\phi) := \left(\frac{1}{\theta} (f(X_E^x, w^1(X_E^x), w^2(X_E^x)) + aw_1(X_E^x)) e^{-(a-\theta)E}, \right. \\ \left. \sqrt{\frac{\pi}{\tilde{\theta}}} \left(f(X_{\tilde{E}}^x, w^1(X_{\tilde{E}}^x), w^2(X_{\tilde{E}}^x)) + \tilde{a}w^1(X_{\tilde{E}}^x) \right) \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} U_{\tilde{E}}^x \right),$$

recalling that $\theta \in (0, a)$, $\tilde{\theta} \in (0, \tilde{a})$, $E \sim \mathcal{E}(\theta)$, $\tilde{E} \sim \Gamma(1/2, \tilde{\theta})$ are independent and independent of W .

Remark 4.3.1. At first sight, there is a weak dependence of the scheme on $v = (u, \bar{u})$, the unknown solution of (4.2.7) and (4.2.8), through the condition $B \geq \|v\|_{\rho_r}$. Nevertheless, Corollary 4.2.1 gives an upper bound on the ρ_r -norm of v . Another possibility to set B is to track a tight upper bound for the constant C in inequality (4.2.9) from Proposition 4.2.3. However, it seems that the truncation procedure is not necessary in practice since we observe the convergence in our numerical experiments without applying the truncation: see Section 4.3.3.

Remark 4.3.2. The scheme given by Definition 4.3.1 is fully implementable and we are able to give a complete study of the numerical error below: see Theorem 4.3.3. On the other hand, it is well known that a grid spacial approximation has a major drawback: The size of the grid exponentially increases when dimension d linearly increases and thus, it is not possible in practice to get a numerical scheme that works as soon as d is too large. Obviously, it is also possible to replace the grid spacial approximation by an other dimensional robust spacial approximation as a neural network for example, even if, in this case, the theoretical study of the numerical error would be more complicated. This is the purpose of Section 4.3.2.

Theoretical study of the scheme

The analysis will be done on $L_{\infty, \rho}$. In order to treat the statistical error we consider Orlicz norms. We denote $\Psi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ an Orlicz function, that is a continuous non-decreasing function, vanishing in zero and with $\lim_{x \rightarrow +\infty} \Psi(x) = +\infty$ and we define the Ψ -Orlicz norm of a real random vector Y by

$$|Y|_{\Psi} := \inf \left\{ c > 0, \mathbb{E} \left[\Psi \left(\frac{|Y|}{c} \right) \right] \leq 1 \right\}.$$

We easily generalize previous definition to matrix-valued random variables. We also assume that Ψ is convex¹, which implies in particular that $|\cdot|_{\Psi}$ is a norm, and an increasing function in order to insure that Ψ^{-1} is a concave function defined on \mathbb{R}^+ .

In practice we will use following convex and increasing Orlicz functions.

- For $\beta \in (0, 1]$, $\Psi_{\beta}^{LT} := e^{x^{\beta}} - 1$. We remark that $|Y|_{\Psi_{\beta}^{LT}} < +\infty$ implies that there exists $\varepsilon > 0$ such that $\mathbb{E}[e^{\varepsilon Y^{\beta}}] < +\infty$ which means that Y is light-tailed.
- For $\beta > 1$, $\Psi_{\beta}^{HT} := e^{(\ln(1+x))^{\beta}} - 1$. Observe that when $|Y|_{\Psi_{\beta}^{LT}} < +\infty$, Y has finite polynomial moment of order p for any $p > 0$ but Y^{α} may not have exponential moment for any $\alpha > 0$. So, this Orlicz function is well suited for heavy-tailed random variables that are not fat-tailed. In particular the case $\beta = 2$ fits well to log-normal tailed random variables.
- For $\beta > 1$, $\Psi_{\beta}^{FT} := x^{\beta}$. We consider these Orlicz functions when Y has some polynomial moments of finite order. This is the case for fat-tailed random variables.

All these Orlicz functions satisfy some Talagrand inequality and maximal inequality.

Proposition 4.3.2. Ψ_{β}^{LT} ($\beta \in (0, 1]$), Ψ_{β}^{HT} ($\beta > 1$) and Ψ_{β}^{FT} ($\beta > 1$) satisfy:

1. **[Talagrand inequality]** There exists a universal constant C_{Ψ} such that, for all sequence $(Y_k)_{1 \leq k \leq K}$ of independent, mean zero, random variables satisfying $|Y_k|_{\Psi} < +\infty$ for all $0 \leq k \leq K$, we have

$$\left| \sum_{k=1}^K Y_k \right|_{\Psi} \leq C_{\Psi} \left(\mathbb{E} \left[\left| \sum_{k=1}^K Y_k \right|^2 \right] + \left| \max_{1 \leq k \leq K} |Y_k| \right|_{\Psi} \right). \quad (4.3.2)$$

2. **[Maximal inequality]** There exists a universal constant C_{Ψ} such that, for all sequence $(Y_k)_{1 \leq k \leq K}$ of random variables satisfying $|Y_k|_{\Psi} < +\infty$ for all $0 \leq k \leq K$, we have

$$\left| \max_{1 \leq k \leq K} |Y_k| \right|_{\Psi} \leq C_{\Psi} \Psi^{-1}(K) \max_{1 \leq k \leq K} |Y_k|_{\Psi}. \quad (4.3.3)$$

Proof. Talagrand inequality for Ψ_{β}^{LT} and Ψ_{β}^{FT} comes from Theorem 3 and Theorem 1 in [113]. Talagrand inequality for Ψ_{β}^{HT} is given by Theorem 2.1 in [42]. Maximal inequality is provided by Lemma 2.2.2 in [115] (see also Theorem 2.2 in [42] for Ψ_{β}^{HT}). \square

Proposition 4.3.3. Let us assume following hypothesis:

1. Assumptions 4.2.1 and 4.2.2 are fulfilled.
2. $\kappa_{\infty} < 1$, recalling that κ_{∞} is defined in Proposition 4.2.4. Proposition 4.2.4 gives us that the fixed point equation $\Phi(v) = v$ has a unique solution $v \in C^0(\mathbb{R}^d, \mathbb{R}^{d'} \times \mathbb{R}^{d \times d'})$ such that $\|v(x)\| \leq C(1 + |x|^r)$.

¹convex Orlicz functions are also referred to “Young functions” or “N-functions” in the literature.

3. $v \in C^2(\mathbb{R}^d, \mathbb{R}^{d'} \times \mathbb{R}^{d \times d'})$ with bounded second derivatives.
4. There exists a convex and increasing Orlicz function Ψ satisfying Talagrand inequality (4.3.2), Maximal inequality (4.3.3) and there exists $C > 0$, $\delta > 0$ and $z \in \Pi$, such that, for all functions $\phi : \mathbb{R}^d \rightarrow \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d}$ continuous with a growth $\|\phi(x)\| \leq B(1 + |x|^r)$

$$\left| \frac{R^z(P\phi)}{\rho(z)} - \mathbb{E} \left[\frac{R^z(P\phi)}{\rho(z)} \right] \right|_{\Psi} + \mathbb{E} \left[\left\| \frac{R^z(P\phi)}{\rho(z)} - \mathbb{E} \left[\frac{R^z(P\phi)}{\rho(z)} \right] \right\|^2 \right] \leq C. \quad (4.3.4)$$

Then there exists a constant $C > 0$ that does not depend on M , n and δ such that

$$\begin{aligned} & \mathbb{E} \left[\sup_{x \in \mathbb{R}^d} \left\| \frac{Pv_M^n(x) - v(x)}{\rho(x)} \right\| \right] \\ & \leq C \left(\sup_{x \in \mathbb{R}^d} \left| \frac{P\rho(x)}{\rho(x)} \right| \frac{\Psi^{-1}(N)}{\sqrt{M}} \left(1 + \frac{\Psi^{-1}(M)}{\sqrt{M}} \right) + \delta^2 + \sup_{x \in \mathbb{R}^d \setminus \square} \left(\frac{1 + |x|^r}{\rho(x)} \right) + \kappa_{\infty}^n \sup_{x \in \mathbb{R}^d} \left\| \frac{v(x)}{\rho(x)} \right\| \right). \end{aligned}$$

The proof of Proposition 4.3.3 is postponed to Section 4.4.2.

Remark 4.3.3. The upper bound obtained for the numerical error in Proposition 4.3.3 can be easily analyzed:

1. The first term is the statistical error coming from the approximation of the expectation by an empirical mean. When we consider Orlicz functions Ψ_{β}^{LT} ($\beta \in (0, 1]$), Ψ_{β}^{HT} ($\beta > 1$) and Ψ_{β}^{FT} ($\beta \geq 2$), then

$$\begin{aligned} & \frac{(\Psi_{\beta}^{LT})^{-1}(N)}{\sqrt{M}} \left(1 + \frac{(\Psi_{\beta}^{LT})^{-1}(M)}{\sqrt{M}} \right) = O \left(\frac{(\ln(N))^{1/\beta}}{\sqrt{M}} \right) \\ & \frac{(\Psi_{\beta}^{HT})^{-1}(N)}{\sqrt{M}} \left(1 + \frac{(\Psi_{\beta}^{HT})^{-1}(M)}{\sqrt{M}} \right) = O \left(\frac{e^{(\ln(N))^{1/\beta}}}{\sqrt{M}} \right) \\ & \frac{(\Psi_{\beta}^{FT})^{-1}(N)}{\sqrt{M}} \left(1 + \frac{(\Psi_{\beta}^{FT})^{-1}(M)}{\sqrt{M}} \right) = O \left(\frac{N^{1/\beta}}{\sqrt{M}} \right). \end{aligned}$$

2. The second term is related to the space discretization by a discrete grid $\delta\mathbb{Z}^d$.
3. The third term is a truncation error. In order to get a good control on it we must consider a weight function ρ with a bigger growth than the solution v .
4. The last term comes from the Picard procedure using the contraction property.

We specify now the error given by Proposition 4.3.3 for several set of assumptions on b , σ and f . We also assume that our grid Π is centered in 0, and is given by

$$\left\{ (i_1\delta, \dots, i_d\delta) \mid i_k \in \{-\tilde{N}, \dots, \tilde{N}\}, k \in \{1, \dots, d\} \right\}$$

for a given $\tilde{N} \in \mathbb{N}$. It implies that $N = (2\tilde{N} + 1)^d$. Finally, we choose $\rho(x) = 1 + |x|^{r'}$ with $r' > r$.

Lemma 4.3.1. Let us assume that assumptions of Proposition 4.2.3 are fulfilled. We set $r' > r$ where r comes from (4.2.9).

1. If σ is constant, $\tilde{a} - \tilde{\theta} - (1 + r)K_b > 0$ and $a - \theta - rK_b > 0$, then (4.3.4) is fulfilled for $\Psi_{1/(r+1)}^{LT}$.

2. If $f(x, y, z) = f(x, y)$ and $a - \theta - rK_b > 0$, then (4.3.4) is fulfilled for $\Psi_{1/(r\vee 1)}^{LT}$.
3. If $\beta \geq 2$, $\tilde{a} - (r+1)K_b - \frac{2\beta-1}{2}K_\sigma^2 - \frac{\beta-1}{\beta}\tilde{\theta} > 0$ and $a - rK_b - \frac{\beta-1}{\beta}\theta > 0$, then (4.3.4) is fulfilled for Ψ_β^{FT} .

The proof of Lemma 4.3.1 is postponed to Section 4.4.3.

Theorem 4.3.1. Let us assume that assumptions of Proposition 4.2.3 are fulfilled. We set $r' > r$ where r comes from (4.2.9). We also assume that $\kappa_\infty < 1$ and $v \in C^2(\mathbb{R}^d, \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d'})$ with bounded second derivatives.

Then the following holds.

1. If σ is constant, $\tilde{a} - \tilde{\theta} - (1+r)K_b > 0$ and $a - \theta - rK_b > 0$, then

$$\mathbb{E} \left[\sup_{x \in \mathbb{R}^d} \left\| \frac{Pv_M^n(x) - v(x)}{\rho(x)} \right\| \right] = O \left(\delta^2 + \frac{(\ln \tilde{N})^{1/(r+1)}}{\sqrt{M}} + |\tilde{N}\delta|^{r-r'} + \kappa_\infty^n \right).$$

2. If $f(x, y, z) = f(x, y)$ and $a - \theta - rK_b > 0$, then

$$\mathbb{E} \left[\sup_{x \in \mathbb{R}^d} \left\| \frac{Pv_M^n(x) - v(x)}{\rho(x)} \right\| \right] = O \left(\delta^2 + \frac{(\ln \tilde{N})^{1/r}}{\sqrt{M}} + |\tilde{N}\delta|^{r-r'} + \kappa_\infty^n \right).$$

3. If $\beta \geq 2$, $\tilde{a} - (r+1)K_b - \frac{2\beta-1}{2}K_\sigma^2 - \frac{\beta-1}{\beta}\tilde{\theta} > 0$ and $a - rK_b - \frac{\beta-1}{\beta}\theta > 0$, then

$$\mathbb{E} \left[\sup_{x \in \mathbb{R}^d} \left\| \frac{Pv_M^n(x) - v(x)}{\rho(x)} \right\| \right] = O \left(\delta^2 + \frac{(\tilde{N})^{d/\beta}}{\sqrt{M}} + |\tilde{N}\delta|^{r-r'} + \kappa_\infty^n \right).$$

Proof.

Thanks to Lemma 4.3.1 and Proposition 4.3.2 we are allowed to apply Proposition 4.3.3 and we just have to study all the terms in the upper-bound. First of all, we have

$$\sup_{x \in \mathbb{R}^d} \left| \frac{P\rho(x)}{\rho(x)} \right| \leq \sup_{x \in \mathbb{R}^d} \frac{1 + (\sqrt{d}\delta + |x|)^{r'}}{1 + |x|^{r'}} = 1 + (\sqrt{d}\delta)^{r'} = O(1),$$

and

$$\sup_{x \in \mathbb{R}^d \setminus \square} \left(\frac{1 + |x|^r}{\rho(x)} \right) \leq \sup_{x \in \mathbb{R}^d \setminus \square} \left(\frac{1 + |x|^r}{1 + |x|^{r'}} \right) \leq \sup_{|x| \geq \tilde{N}\delta} \frac{1 + |x|^r}{1 + |x|^{r'}} = O(|\tilde{N}\delta|^{r-r'}).$$

Moreover, by using Remark 4.3.3 we obtain the result. □

4.3.2 Some schemes based on neural networks approximation

Contraction based scheme using neural networks approximation

We use once again the contraction property by considering a sequence of neural networks (NN for short) $(\mathcal{U}_{\zeta_n, n})_{n \in \mathbb{N}}$, i.e. a sequence of functions $x \mapsto (\mathcal{U}_{\zeta_n, n}^1, \mathcal{U}_{\zeta_n, n}^2) \in \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d}$ parametrized by $\zeta_n \in \Theta_n$. We set $\mathcal{U}_{\zeta_0, 0} = 0$. Knowing the neural network $\mathcal{U}_{\zeta_n, n}$, we train $\mathcal{U}_{\zeta_{n+1}, n+1}$ as follows:

1. Sample $X_0 = X_{t=0} \sim \mu_0$, $E \sim \mathcal{E}(\theta)$ and $\tilde{E} \sim \Gamma(1/2, \tilde{\theta})$.
2. Get corresponding samples of the Malliavin weight $U_E^{X_0}$, $U_{\tilde{E}}^{X_0}$ and state processes $X_E^{X_0}$, $X_{\tilde{E}}^{X_0}$ at times E and \tilde{E} for the starting point X_0 .

3. Evaluate the previous neural network $\mathcal{U}_{\zeta_n, n}$ at each sample of $X_E^{X_0}$ and $X_{\tilde{E}}^{X_0}$ and compute the samples of $\varphi(\mathcal{U}_{\zeta_n, n})(X_0)$ given by,

$$\varphi(\mathcal{U}_{\zeta_n, n})(X_0) = \left(\frac{1}{\theta} (f(X_E^{X_0}, \mathcal{U}_{\zeta_n, n}^1(X_E^{X_0}), \mathcal{U}_{\zeta_n, n}^2(X_E^{X_0})) + a\mathcal{U}_{\zeta_n, n}^1(X_E^{X_0})) e^{-(a-\theta)E}, \right. \\ \left. \sqrt{\frac{\pi}{\tilde{\theta}}} \left(f(X_{\tilde{E}}^{X_0}, \mathcal{U}_{\zeta_n, n}^1(X_{\tilde{E}}^{X_0}), \mathcal{U}_{\zeta_n, n}^2(X_{\tilde{E}}^{X_0})) + \tilde{a}\mathcal{U}_{\zeta_n, n}^1(X_{\tilde{E}}^{X_0}) \right) \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} U_{\tilde{E}}^{X_0} \right).$$

4. Minimize the loss function,

$$\mathcal{L}(\mathcal{U}_{\zeta, n+1}) := \mathbb{E} [\|\mathcal{U}_{\zeta, n+1}(X_0) - \varphi(\mathcal{U}_{\zeta_n, n})(X_0)\|^2], \quad \zeta \in \Theta_{n+1}$$

and choose the parameters ζ_{n+1} ,

$$\zeta_{n+1} \in \arg \min_{\zeta \in \Theta_{n+1}} \mathbb{E} [\|\mathcal{U}_{\zeta, n+1}(X_0) - \varphi(\mathcal{U}_{\zeta_n, n})(X_0)\|^2].$$

The scheme can be summarized as follows.

Algorithm 3 Contraction based scheme using NN

We start with an initial guess for the neural networks for $(\mathcal{U}_{\theta_0, 0}^1, \mathcal{U}_{\theta_0, 0}^2) = (0, 0)$. At a given step $n + 1$ of the Picard iteration, we sample a random starting point from the distribution μ_0 and compute $\varphi(\mathcal{U}_{\zeta_n, n})$ and we minimize the loss function $\mathcal{L}(\mathcal{U}_{\theta, n+1})$.

- 1: Initialize $(\mathcal{U}_{\theta_0, 0}^1, \mathcal{U}_{\theta_0, 0}^2) \leftarrow (0, 0)$.
- 2: **for** n from 1 to N **do** ▷ Picard iterations
- 3: Obtain M samples of $X_{t=0} \sim \mu_0$, $E \sim \mathcal{E}(\theta)$ and $\tilde{E} \sim \Gamma(\frac{1}{2}, \tilde{\theta})$.
- 4: Obtain corresponding M samples of $U_E^{X_0}$, $U_{\tilde{E}}^{X_0}$, $X_E^{X_0}$ and $X_{\tilde{E}}^{X_0}$.
- 5: Define

$$\varphi^1(\mathcal{U}_{\zeta_n, n})(X_0) \leftarrow \frac{1}{\theta} (f(X_E^{X_0}, \mathcal{U}_{\zeta_n, n}^1(X_E^{X_0}), \mathcal{U}_{\zeta_n, n}^2(X_E^{X_0})) + a\mathcal{U}_{\zeta_n, n}^1(X_E^{X_0})) e^{-(a-\theta)E}.$$

- 6: Define

$$\varphi^2(\mathcal{U}_{\zeta_n, n})(X_0) \leftarrow \sqrt{\frac{\pi}{\tilde{\theta}}} \left(f(X_{\tilde{E}}^{X_0}, \mathcal{U}_{\zeta_n, n}^1(X_{\tilde{E}}^{X_0}), \mathcal{U}_{\zeta_n, n}^2(X_{\tilde{E}}^{X_0})) + \tilde{a}\mathcal{U}_{\zeta_n, n}^1(X_{\tilde{E}}^{X_0}) \right) \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} U_{\tilde{E}}^{X_0}.$$

- 7: Set $\varphi(\mathcal{U}_{\zeta_n, n})(X_0) \leftarrow (\varphi^1(\mathcal{U}_{\zeta_n, n})(X_0), \varphi^2(\mathcal{U}_{\zeta_n, n})(X_0))$.
- 8: Train the neural network $\mathcal{U}_{\zeta, n+1}$ to minimize

$$\mathcal{L}(\mathcal{U}_{\zeta, n+1}) = \frac{1}{M} \sum_{k=1}^M \|\mathcal{U}_{\zeta, n+1}(X_0^k) - \varphi(\mathcal{U}_{\zeta_n, n})(X_0^k)\|^2, \quad \zeta \in \Theta_{n+1}$$

using ADAM optimizer.

- 9: Update $\zeta_{n+1} \leftarrow \arg \min_{\zeta \in \Theta_{n+1}} \mathcal{L}(\mathcal{U}_{\zeta, n+1})$.

10: **end for**

Now we want to study the error

$$e_n := \int_{\mathbb{R}^d} \|\mathcal{U}_{\zeta_n, n}(x) - v(x)\|^2 \mu_0(dx), \quad n \in \mathbb{N}.$$

Let us introduce $r_n := \mathbb{E} [\|\mathcal{U}_{\zeta_n, n}(X_0) - \varphi(\mathcal{U}_{\zeta_{n-1}, n-1})(X_0)\|^2]$, for $n \in \mathbb{N}$, which is the accuracy of the neural network approximation at the n -th Picard iteration. We have the following result for the convergence of this scheme to the solution of the BSDE (4.2.1).

Proposition 4.3.4. We assume assumptions of Proposition 4.2.4 for $p = 2$ and $\pi = \mu_0$. Let $\varepsilon > 0$ be such that $(1 + \varepsilon)(\kappa_2)^2 < 1$ where κ_2 is defined in Proposition 4.2.4. We also assume that we can solve exactly at each Picard step the optimization problem. Then, the numerical approximation error of the scheme at n -th Picard iteration is bounded by

$$e_n \leq \frac{(1 + \varepsilon) \sup_{1 \leq k \leq n} r_k}{\varepsilon - \varepsilon(1 + \varepsilon)(\kappa_2)^2} + [(1 + \varepsilon)(\kappa_2)^2]^n e_0.$$

Proof.

We denote by $\mathcal{N}_\Theta := \{\mathcal{U}_\zeta, \zeta \in \Theta\}$ the set of functions $\mathbb{R}^d \rightarrow \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d}$ that can be mimicked by a neural network with parameters in Θ .

We assume that $\kappa_2 < 1$. Then Corollary 4.2.1 gives us the existence and uniqueness of a solution v to (4.2.7) and (4.2.8). We also assume that, $\zeta \in \Theta_n$ is numerically computable, i.e. we ignore numerical optimisation errors. In particular, $\mathcal{U}_{\zeta_n, n}$ is deterministic for all $n \in \mathbb{N}$. For given parameters $\zeta \in \Theta_{n+1}$ and $\mathcal{U}_\zeta \in \mathcal{N}_{\Theta_{n+1}}$, we have

$$\begin{aligned} \mathbb{E}[\|\mathcal{U}_\zeta(X_0) - \varphi(\mathcal{U}_{\zeta_n, n})(X_0)\|^2] &= \int_{\mathbb{R}^d} \|\mathcal{U}_\zeta(x) - \mathbb{E}[\varphi(\mathcal{U}_{\zeta_n, n})(x)]\|^2 \mu_0(dx) \\ &\quad + \int_{\mathbb{R}^d} \mathbb{E}[\|\mathbb{E}[\varphi(\mathcal{U}_{\zeta_n, n})(x)] - \varphi(\mathcal{U}_{\zeta_n, n})(x)\|^2] \mu_0(dx). \end{aligned}$$

Thus, we obtain

$$\mathcal{U}_{\zeta_{n+1}, n+1} = \text{Proj}(\mathbb{E}[\varphi(\mathcal{U}_{\zeta_n, n})(x)], \mathcal{N}_{\Theta_{n+1}}).$$

Note that for any $\varepsilon > 0$,

$$\begin{aligned} e_{n+1} &\leq \left(1 + \frac{1}{\varepsilon}\right) \mathbb{E}[\|\mathcal{U}_{\zeta_{n+1}, n+1}(X_0) - \varphi(\mathcal{U}_{\zeta_n, n})(X_0)\|^2] + (1 + \varepsilon)(\kappa_2)^2 \mathbb{E}[\|\varphi(\mathcal{U}_{\zeta_n, n})(X_0) - v(X_0)\|^2] \\ &\leq \left(1 + \frac{1}{\varepsilon}\right) r_{n+1} + (1 + \varepsilon)(\kappa_2)^2 e_n. \end{aligned}$$

Let $\varepsilon > 0$ such that $(1 + \varepsilon)(\kappa_2)^2 < 1$. Then, we have

$$e_n \leq \frac{(1 + \varepsilon) \max_{1 \leq k \leq n} r_k}{\varepsilon - \varepsilon(1 + \varepsilon)(\kappa_2)^2} + [(1 + \varepsilon)(\kappa_2)^2]^n e_0.$$

□

This convergence result is clearly theoretical since it is well known that there is no theoretical guaranty for the numerical optimization procedure to converge to a global minimum and even if we would have convergence, we do not take into account the numerical error for the approximation procedure. Moreover, it is known that theoretically, we can find at each step a Θ_n large enough such that r_n is as small as we want, see for instance Theorem 1 of Hornik [79] (see also Leshno et al. [86]). Nevertheless, we do not know in practice how to specify Θ_n , i.e. which architecture of NN we have to consider, in order to be sure that r_n is smaller than a given upper bound. Finally, we must remark that this scheme suffers from the same drawback as the grid scheme since it requires some contraction in order to have the Picard procedure that converges. The aim of the next NN scheme is to overcome this limitation.

Direct scheme using neural networks approximation

Here we propose a scheme using a neural network approximation that do not rely on a contraction property. We train a neural network \mathcal{U}_ζ , for $\zeta \in \Theta$, in epochs, each with \mathcal{P} gradient descent

steps, whereby the NN is optimized for a given set of samples before we update the input samples in the next epoch.

The proposed scheme is defined as follows.

Algorithm 4 Direct scheme using NN

- 1: Initialisation of ζ_0^0
 - 2: **for** i from 1 to I **do** ▷ Iterating over epochs
 - 3: Obtain M_x samples of $X_{t=0} \sim \mu_0$.
 - 4: For each sample of X_0 , obtain M samples of $(E, \tilde{E}, X_E^{X_0}, X_{\tilde{E}}^{X_0}, U_E^{X_0}, U_{\tilde{E}}^{X_0})$ where

$$E \sim \mathcal{E}(\theta) \text{ and } \tilde{E} \sim \Gamma(0.5, \tilde{\theta}).$$
 - 5: Set $\zeta_{n=0}^i \leftarrow \zeta_{n=\mathcal{P}}^{i-1}$. ▷ Using the NN from the end of the previous epoch
 - 6: **for** n from 0 to $\mathcal{P} - 1$ **do** ▷ Gradient descent steps
 - 7: Define for each $k = 1, \dots, M_x$ and $j = 1, 2, \dots, M$,

$$\varphi^{1,j}(\mathcal{U}_{\zeta_n^i})(X_0^k) \leftarrow \frac{1}{\theta} (f(X_{E^j}^{X_0^k}, \mathcal{U}_{\zeta_n^i}^1(X_{E^j}^{X_0^k}), \mathcal{U}_{\zeta_n^i}^2(X_{E^j}^{X_0^k})) + a\mathcal{U}_{\zeta_n^i}^1(X_{E^j}^{X_0^k})) e^{-(a-\theta)E^j}.$$
 - 8: Define for each $k = 1, \dots, M_x$ and $j = 1, 2, \dots, M$,

$$\varphi^{2,j}(\mathcal{U}_{\zeta_n^i})(X_0^k) \leftarrow \sqrt{\frac{\pi}{\tilde{\theta}}} \left(f(X_{\tilde{E}^j}^{X_0^k}, \mathcal{U}_{\zeta_n^i}^1(X_{\tilde{E}^j}^{X_0^k}), \mathcal{U}_{\zeta_n^i}^2(X_{\tilde{E}^j}^{X_0^k})) + \tilde{a}\mathcal{U}_{\zeta_n^i}^1(X_{\tilde{E}^j}^{X_0^k}) \right) \sqrt{\tilde{E}^j} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}^j} U_{\tilde{E}^j}^{X_0^k,j}.$$
 - 9: Set $\varphi^j(\mathcal{U}_{\zeta_n^i})(X_0^k) \leftarrow (\varphi^{1,j}(\mathcal{U}_{\zeta_n^i})(X_0^k), \varphi^{2,j}(\mathcal{U}_{\zeta_n^i})(X_0^k)) \forall j = 1, \dots, M, \forall k = 1, \dots, M_x$.
 - 10: Given ζ_n^i , we get ζ_{n+1}^{i+1} using one ADAM optimization step for the loss function,

$$\Upsilon_{M, M_x}(\mathcal{U}_{\zeta}) := \frac{1}{M_x} \sum_{k=1}^{M_x} \left\| \mathcal{U}_{\zeta}(X_0^k) - \frac{1}{M} \sum_{j=1}^M \varphi^j(\mathcal{U}_{\theta})(X_0^k) \right\|^2.$$
 - 11: **end for**
 - 12: **end for**
-

The theoretical study of this third scheme is not tackled here and is left for some future works.

4.3.3 Numerical experiments

In this subsection, we provide some numerical experiments for the three schemes described above. For the process X_t^x , we take the Brownian motion W_t^x starting from $x \in \mathbb{R}^d$, i.e. $b(x) = 0$ and $\sigma(x) = \mathbb{I}_d$ (identity matrix of dimension d) in (3.2.1). We design examples with analytical solutions in the following way: We start with a given function $u(x) \in C^2(\mathbb{R}^{d'})$ and construct the function f such that the BSDE is satisfied by $Y_t^x = u(W_t^x)$ and $Z_t^x = \nabla_x u(W_t^x)$. This can be done by choosing a suitable function $f_0(x, y, z)$ which satisfies Assumption 4.2.1, in particular, with μ as the y-monotonicity constant and $K_{f,z}$ as the z-Lipschitz continuity constant. Then, take the generator of the BSDE as follows:

$$f(x, y, z) = f_0(x, y, z) - \frac{1}{2} \Delta u(x) - f_0(x, u(x), \nabla_x u(x))$$

where $\Delta u := \nabla \cdot \nabla u$ so that we satisfy the PDE (4.2.3) which is equivalent to our BSDE. We use the following example with $d' = 1$ for illustration:

$$u(x) = \frac{1}{d} \sum_{i=1}^d \tan^{-1}(x_i)$$

$$\begin{aligned}\bar{u}(x) &= \frac{1}{d} \left[\frac{1}{1+x_i^2} \right]_{i \in \{1, \dots, d\}} \\ f_0(x, y, z) &= -cy + \cos(y + |x|) + K_z \sin(|z|) \\ f(x, y, z) &= f_0(x, y, z) + \frac{1}{d} \sum_{i=1}^d \frac{x_i}{(1+x_i^2)^2} - f_0(x, u(x), \bar{u}(x)).\end{aligned}$$

The Malliavin weight for this setting can be calculated as per (4.2.4), which gives $U_t^x = \frac{W_t^x}{t}$.

First Scheme

We provide some numerical experiments for the grid-based contraction scheme Definition 4.3.1. For the numerical purposes, we are not using the truncation operator T_{B, ρ_r} to keep things simpler. Also, we have taken the following values for the parameters: $c = 2$, $K_z = 0.1$, $a = 2$, $\tilde{a} = 2$, $\theta = 1.5$ and $\tilde{\theta} = 1.5$ unless specified otherwise. We define the following estimates of error for any function v ,

$$\begin{aligned}\mathcal{E}_{\infty, n}^{d, r}(v) &:= \left\{ \|v_M^n(x) - v(x)\| \mid x = (i_1\delta, \dots, i_d\delta), i_k \in \{-(\tilde{N} - r), \dots, (\tilde{N} - r)\}, k \in \{1, \dots, d\} \right\} \\ \mathfrak{E}_{\infty, n}^{d, r}(v) &:= \sup \mathcal{E}_{\infty, n}^{d, r},\end{aligned}$$

where n denotes the Picard iteration, d is dimension of the Brownian motion and we take $\delta = 0.3$. Figure 4.1 shows the plot of Picard iterations for the sample size $M = 4000$ with $\tilde{N} = 20$ and $r = 10$ corresponding to the grid $\{-6, -5.7, \dots, 5.7, 6\}$ for possible starting points of X_0^x and $\{-3, -2.7, \dots, 2.7, 3\}$ for plotting the solution. Since we use a grid approximation for this scheme, it is natural to expect the choice of the grid to affect the errors especially as we get closer to the boundary. This is illustrated in Figure 4.2 where we plot the errors for two different cases as a box plot: first, when the grid of starting points for X_0^x is $\{-3, -2.7, \dots, 2.7, 3\}$ and second, when it is $\{-6, -5.7, \dots, 5.7, 6\}$. By standard convention, the inter-quantile range (IQR) for the box plot is between $Q_1 = q(0.25)$ to $Q_3 = q(0.75)$ quantiles and the whiskers are at $Q_1 - 1.5IQR$ and $Q_3 + 1.5IQR$. We denote the deviation of the solution at the n -th Picard iteration from the analytical solution as,

$$\begin{aligned}\Delta u_{\tilde{N}, r}^n(x) &:= |u^n(x) - u(x)|, \\ \Delta \bar{u}_{\tilde{N}, r}^n(x) &:= |\bar{u}^n(x) - \bar{u}(x)|.\end{aligned}$$

For the 5-th Picard iteration, the average deviation from the analytical solution in the former case was,

$$\begin{aligned}\frac{1}{2\tilde{N} + 1} \sum_{i=0}^{2\tilde{N}+1} \Delta u_{10, 0}^5(x_i) &\approx 1.62 \times 10^{-2}, \\ \frac{1}{2\tilde{N} + 1} \sum_{i=0}^{2\tilde{N}+1} \Delta \bar{u}_{10, 0}^5(x_i) &\approx 2.41 \times 10^{-2},\end{aligned}$$

and the maximum error was,

$$\begin{aligned}\mathfrak{E}_{\infty, 5}^{10, 0}(u) &\approx 4.88 \times 10^{-2} \quad \text{at } x = 3.0, \\ \mathfrak{E}_{\infty, 5}^{10, 0}(\bar{u}) &\approx 4.49 \times 10^{-2} \quad \text{at } x = 2.4.\end{aligned}$$

For the latter case, the errors were,

$$\frac{1}{2\tilde{N} + 1} \sum_{i=0}^{2\tilde{N}+1} \Delta u_{20, 10}^5(x_i) \approx 9.39 \times 10^{-3},$$

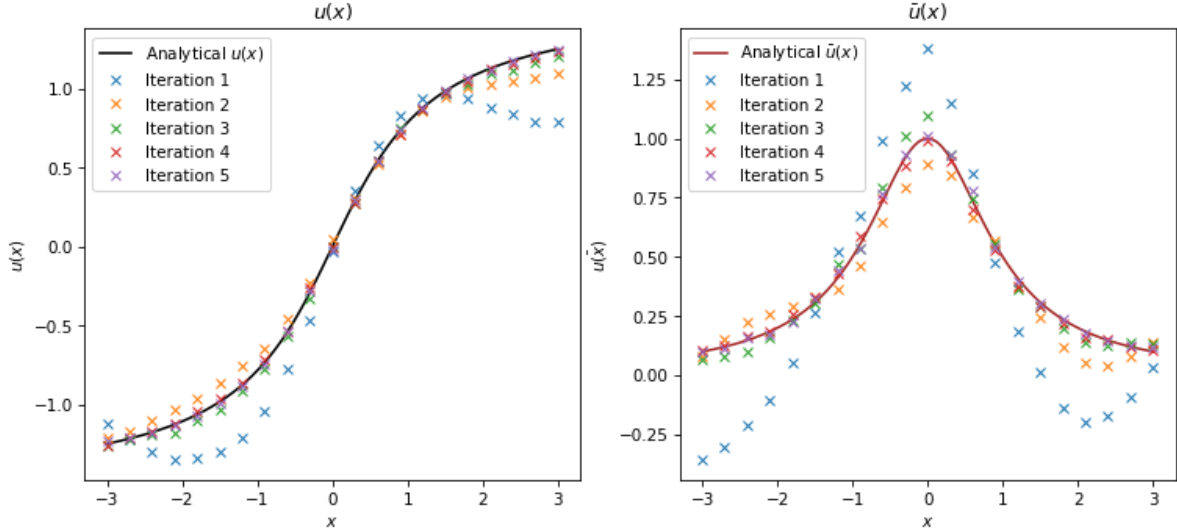


Figure 4.1: Picard iterations for $\tilde{N} = 20$ and $r = 10$ (training grid $\{-6, -5.7, \dots, 5.7, 6\}$).

$$\begin{aligned} \frac{1}{2\tilde{N} + 1} \sum_{i=0}^{2\tilde{N}+1} \Delta \bar{u}_{20,10}^5(x_i) &\approx 1.06 \times 10^{-2}, \\ \mathfrak{E}_{\infty,5}^{10,0}(u) &\approx 2.19 \times 10^{-2} \quad \text{at } x = 0.9, \\ \mathfrak{E}_{\infty,5}^{10,0}(\bar{u}) &\approx 3.31 \times 10^{-2} \quad \text{at } x = -0.9. \end{aligned}$$

Hence, we observe that when the grid of the starting points for X_t^x is larger, we get smaller errors overall but especially closer to the boundary as illustrated with Figure 4.2.

However, this grid-based contraction scheme suffers from the curse of dimensionality as the time complexity for a grid of size $2\tilde{N} + 1$ in each dimension is of the order $O((2\tilde{N} + 1)^d)$ where d is the dimension of the process X (Brownian motion in this case). This motivated us to consider a scheme based on neural networks as described below.

Second scheme

The second scheme also utilizes contraction but relies on neural networks for the approximation procedure (see Algorithm 3). We have used the following values for the parameters: $c = 2$, $K_z = 0.1$, $a = 2$, $\tilde{a} = 2$, $\theta = 1.5$ and $\hat{\theta} = 1.5$ unless specified otherwise. For the distribution of the starting point X_0 , we use a normal distribution $\mathcal{N}(0, \sigma)$ with $\sigma = 2.0$ since we are obtaining the plots within the domain $[-3, 3]$ for each dimension and $\sigma = 2$ gives us a good balance of coverage of the domain and ensures that enough samples are outside this domain so that the errors close to the boundary are small. We use a fully connected neural network with two hidden layers, each with $20 + d$ nodes and we use ReLU activation function. The initial learning rate is 5×10^{-4} which decays by a factor of 0.9 after 1000 steps. In Figure 4.3, we provide a plot of the solution for the first five Picard iterations for the one-dimensional case, i.e. $d = 1$ with the sample size $M = 30000$.

In Figure 4.4, we plot the odd numbered Picard iterations from $n = 1$ to $n = 7$ for the solution corresponding to $d = 2$ and have used a sample size of $M = 40000$ at each iteration. Compared to the one-dimensional case, we need a larger sample size and more Picard iterations for comparable errors which is to be expected. To study the effect of d on the errors, we used $M = 40000$ and obtained $M_{err} = 1000$ test samples with starting points X_0 sampled from $\mathcal{N}(0, \sigma_{err})$ and plotted

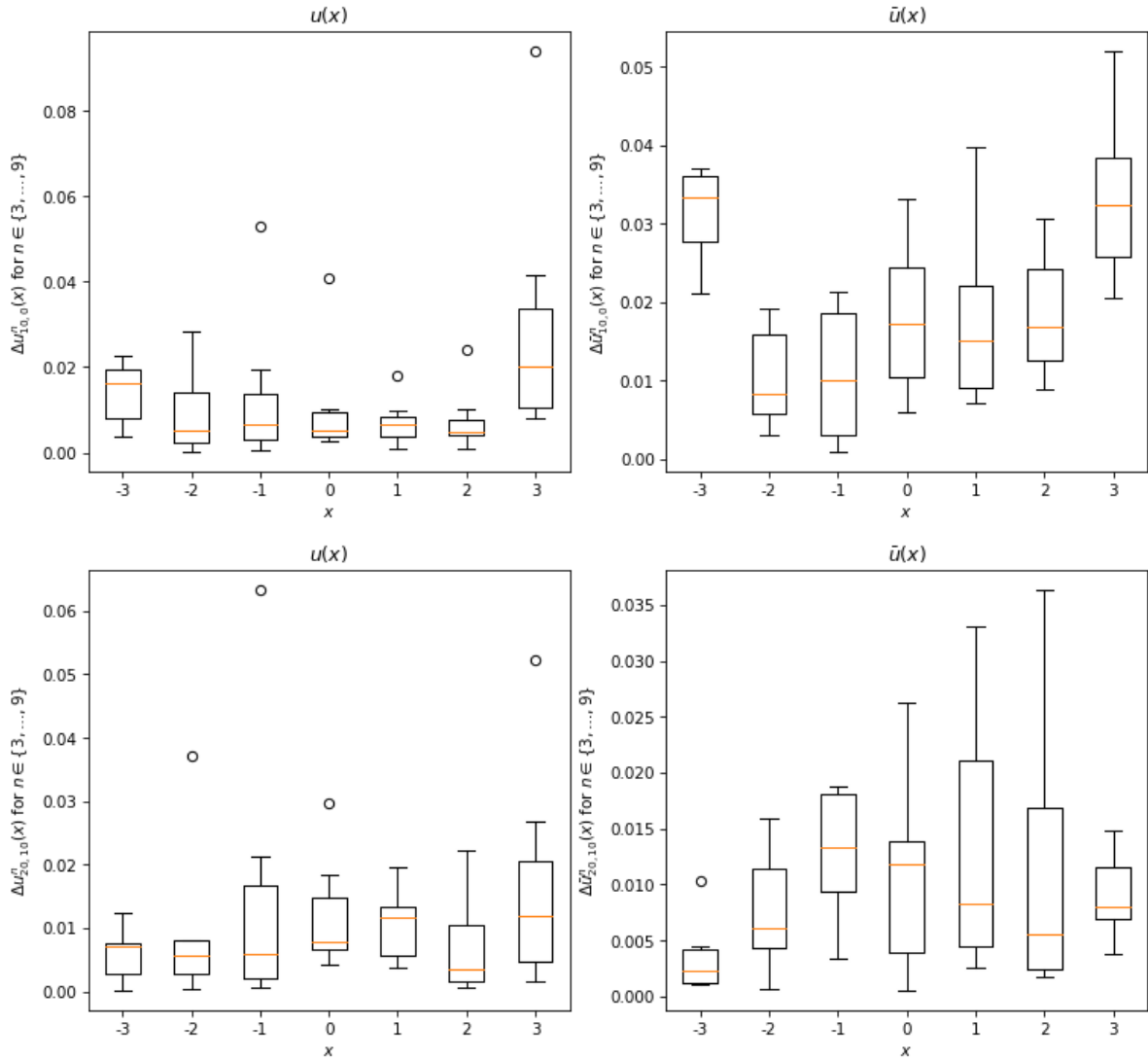


Figure 4.2: Comparison of $\Delta u_{N,r}^n$ and $\Delta \tilde{u}_{N,r}^n$ for $\tilde{N} = 10, r = 0$ (top) and $\tilde{N} = 20, r = 10$ (bottom) for picard iterations $n = 3, \dots, 9$.

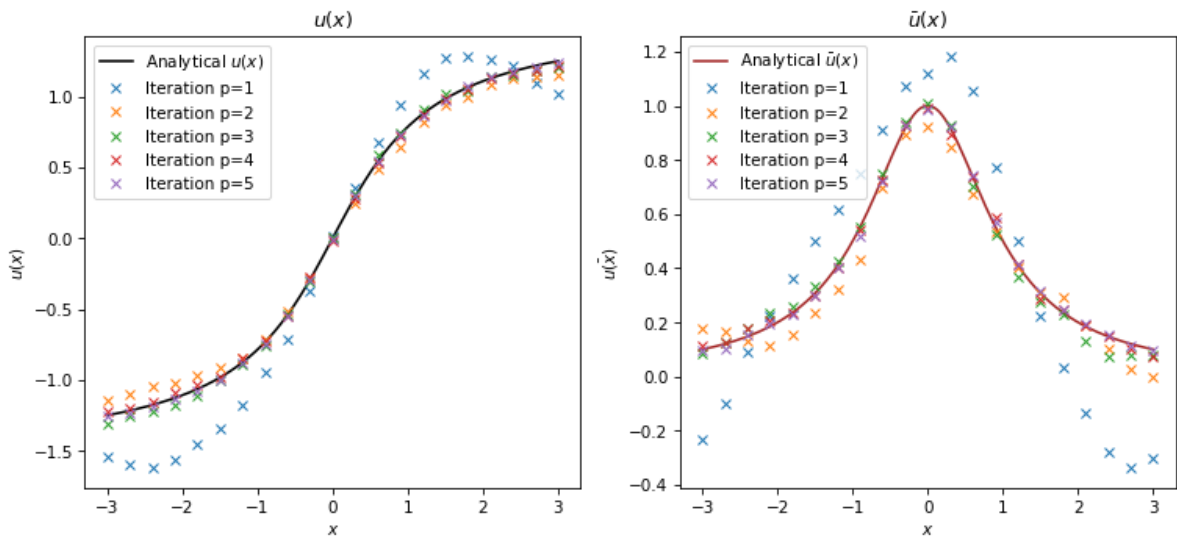


Figure 4.3: Picard iterations for Algorithm 3 in one dimension.

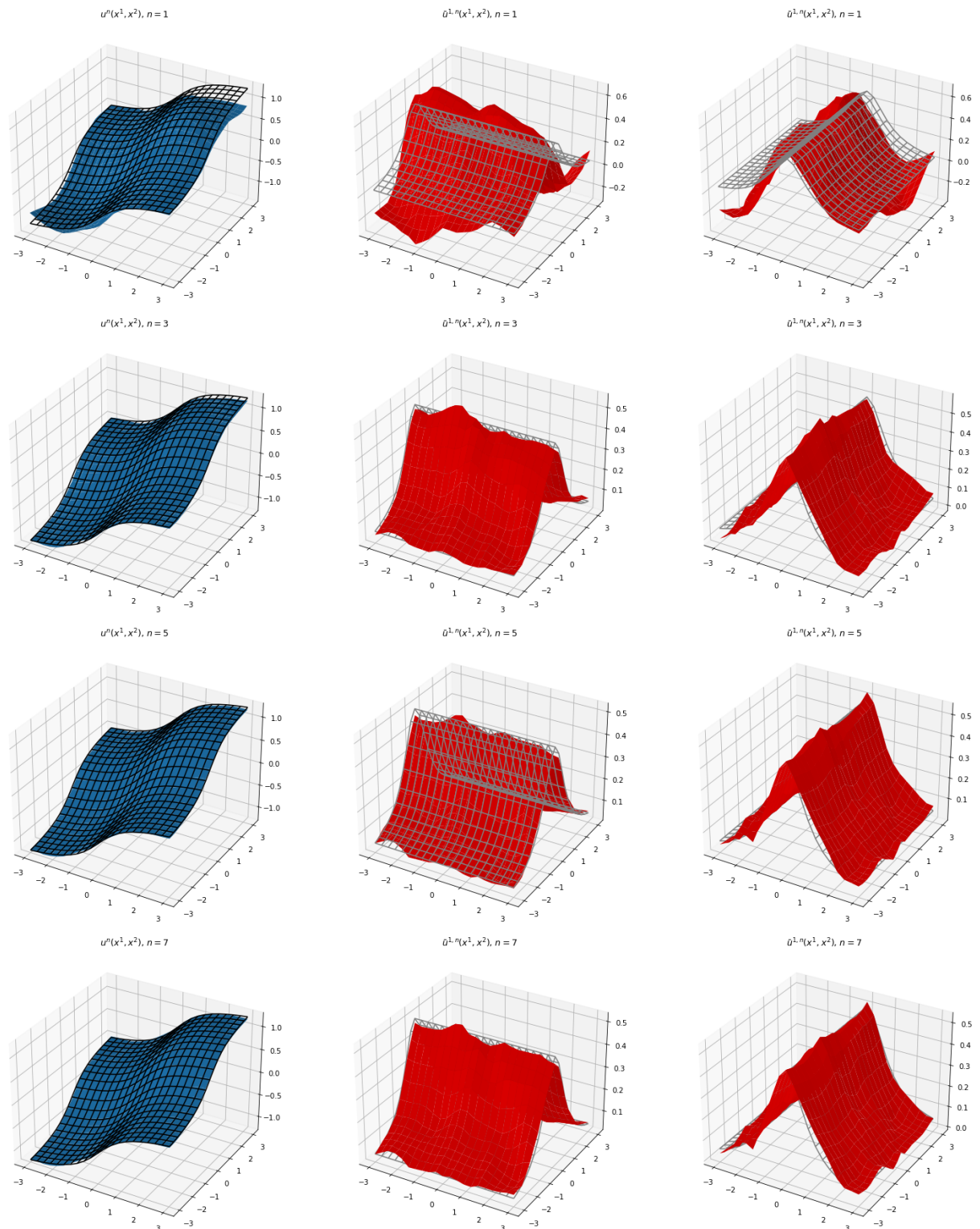


Figure 4.4: Picard iterations (Algorithm 3) for u^n , $\bar{u}^{1,n}$ and $\bar{u}^{2,n}$ for $d = 2$ and $n = 1, 3, 5$ and 7 (from top to bottom). The wire-frame plots are used for the analytical solution.

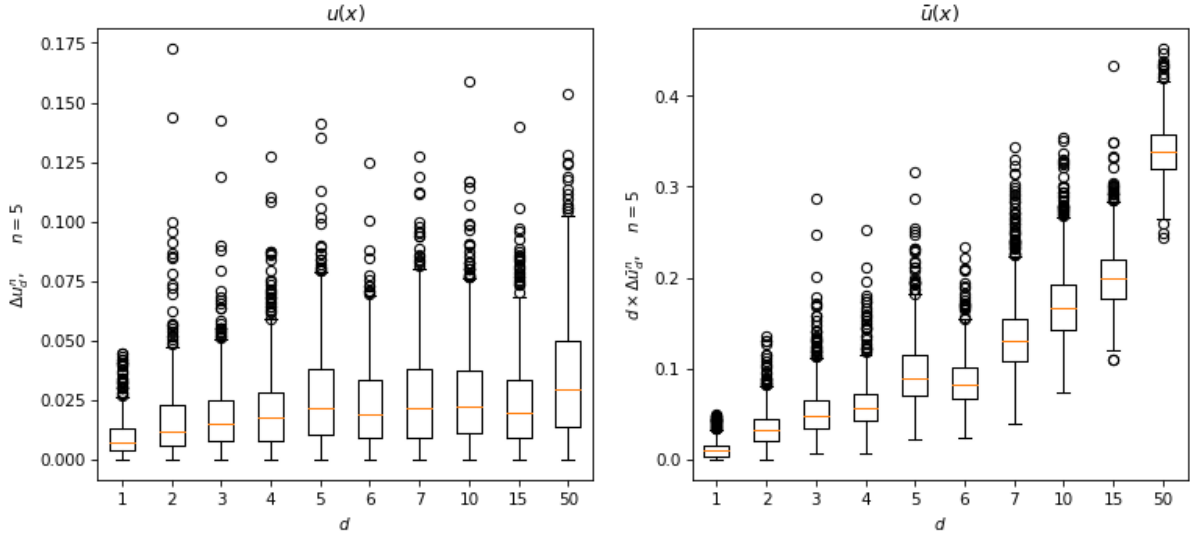


Figure 4.5: Boxplot of errors Δu_d^n and $\Delta \bar{u}_d^n$ for $n = 5$ and different values of d between 1 and 50.

the errors for different values of d , as illustrated in Figure 4.5. We took $\sigma_{err} = 2.0$ which is the same as the one used for training. If we use a smaller σ_{err} for testing, we observe smaller errors. This effect is similar to the truncation of boundary illustrated for the previous scheme. Let us first define the errors as,

$$\begin{aligned} \Delta u_d^n &:= \left\{ |u^n(x_i) - u(x_i)| \mid i \in \{1, \dots, M_{err}\} \right\} \\ \Delta \bar{u}_d^n &:= \left\{ \sqrt{\frac{1}{d} \sum_{k=1}^d (\bar{u}^{n,k}(x_i) - \bar{u}^k(x_i))^2} \mid i \in \{1, \dots, M_{err}\}} \right\} \end{aligned} \quad (4.3.5)$$

We observe that for a fixed number of samples and Picard iterations, the error in u does not grow significantly even for $d = 50$ but for $\{\bar{u}^k\}_{k \in \{1, \dots, d\}}$, the value of the function along each dimension decreases by a factor of $\frac{1}{d}$. Hence, to get a more accurate comparison of the errors, we plot the errors Δu_d^n and $d \times \Delta \bar{u}_d^n$ in Figure 4.5. The running times are almost comparable for each d , thanks to the neural networks' ability to overcome the curse of dimensionality. However, the errors $d \times \Delta \bar{u}_d^n$ grow significantly, suggesting that we need larger sample size for better coverage of the higher dimensional domain when it comes to the computation of the gradient \bar{u} .

Next, we investigate the effect of K_z on the errors. See Remark 4.2.4 to see how increasing K_z can deny the contraction of the scheme, which is we we have taken a small value of $K_z = 0.1$ so far. Theoretically speaking, the existence and uniqueness of the solution is guaranteed for K_z strictly smaller than $\sqrt{2\mu}$ (see Proposition 4.2.1), which is $\sqrt{2}$ in our case. In Figure 4.6, we plot the natural log of the errors for different values of K_z to investigate the breaking point of the scheme for values of K_z close to the theoretical limit, where we have taken $d = 1$ and $M = 30000$. We observe that from $K_z = 1.6$ and onwards, the errors increase significantly. However, we do not observe an explosion even though the existence and uniqueness are not guaranteed.

Third scheme

The third scheme is based on a direct approach not relying on contraction (see Algorithm 4). We use the same neural network architecture as the previous scheme and use an initial learning rate 5×10^{-4} which decays by a factor of 0.6 after 500 steps. We use $M_x = 512$ and for each

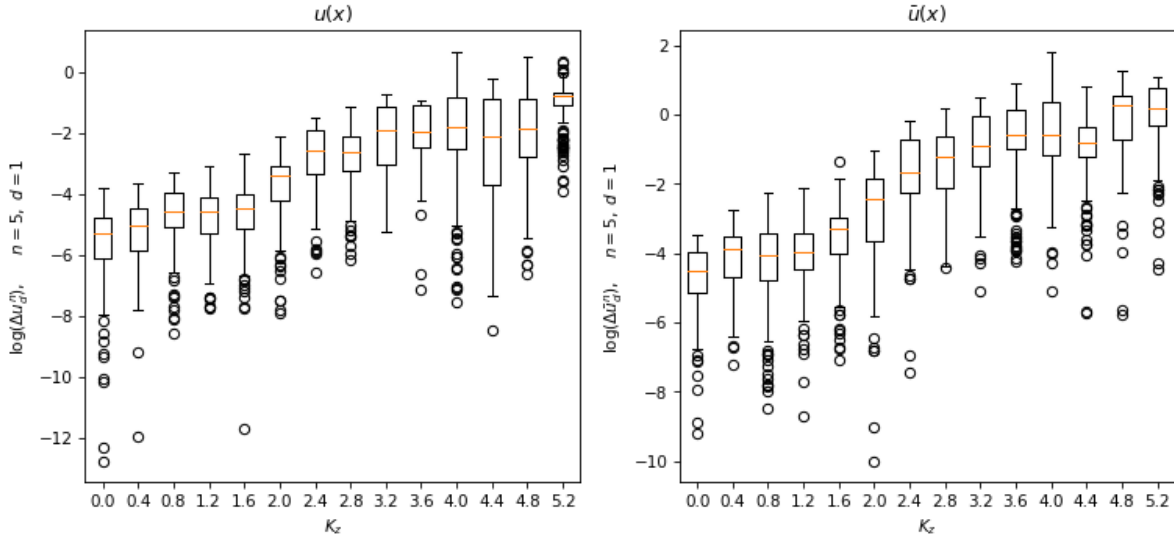


Figure 4.6: Boxplot of errors $\log(\Delta u_d^n)$ and $\log(\Delta \bar{u}_d^n)$ for $n = 5$ and different values of $K_z \in \{0, 0.4, \dots, 5.2\}$ for the second scheme.

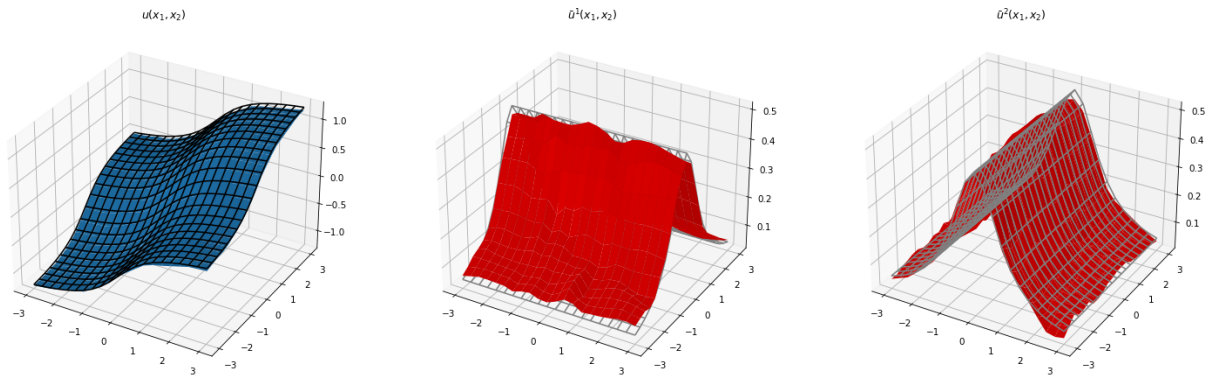


Figure 4.7: Solution $u(x)$ and $\bar{u}(x)$ for $d = 2$ for the third scheme (Algorithm 4).

sample of x , we use $M = 3000$. The values of the system parameters are: $c = 2$, $K_z = 0.1$, $a = 2$, $b = 2$, $\theta = 1.5$ and $\hat{\theta} = 1.5$ unless specified otherwise. In Figure 4.7, we plot the solution for $d = 2$. Next, using a similar approach as the previous scheme, we define the errors Δu_d and $\Delta \bar{u}_d$ similar to (4.3.5) but without n and plot in Figure 4.8 the natural log of the errors for different values of K_z between 0 and 6 which is well beyond the theoretical guarantee for existence and uniqueness. We observe that the errors, unlike the second scheme, do not tend to increase significantly with K_z which is to be expected as there is no contraction required. However, it is surprising that the errors remain low well beyond the theoretical limit for K_z which is $\sqrt{2}$.

4.4 Proofs

4.4.1 Proof of Proposition 4.2.3

1. Let us start by proving a probabilistic representation for u . Applying Itô formula to $e^{-at}Y_t^x$ we get, for all $0 \leq T$,

$$u(x) = Y_0^x = \mathbb{E} \left[e^{-aT} Y_T^x + \int_0^T e^{-as} (f(X_s^x, Y_s^x, Z_s^x) + aY_s^x) ds \right]. \quad (4.4.1)$$

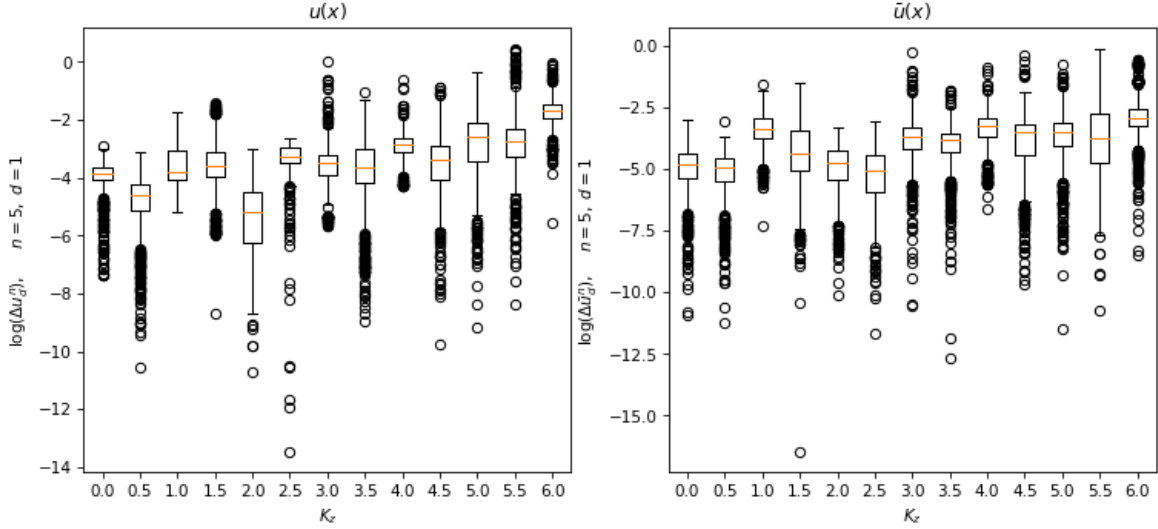


Figure 4.8: Boxplot of errors $\log(\Delta u_d)$ and $\log(\Delta \bar{u}_d)$ for different values of $K_z \in \{0, 0.5, \dots, 6.0\}$ for the third scheme.

By using Proposition 4.2.1 and Young's inequality, we have

$$\begin{aligned} & \mathbb{E} \left[\int_0^{+\infty} e^{-as} |f(X_s^x, Y_s^x, Z_s^x) + aY_s^x| ds \right] \\ & \leq C \mathbb{E} \left[\int_0^{+\infty} e^{-(2a-\lambda)s} + e^{-\lambda s} (|Y_s^x|^2 + \|Z_s^x\|^2 + |X_s^x|^{2r}) ds \right] < +\infty, \end{aligned}$$

where we have set $\lambda := 2\mu - K_{f,z}^2 \mathbf{1}_{d' > 1}$. The Lebesgue theorem gives us that the integral term in (4.4.1) tends to

$$\mathbb{E} \left[\int_0^{+\infty} e^{-as} (f(X_s^x, Y_s^x, Z_s^x) + aY_s^x) ds \right]$$

when $T \rightarrow +\infty$. Moreover, by Proposition 4.2.1 we get

$$\lim_{T \rightarrow +\infty} \mathbb{E} [e^{-aT} |Y_T^x|^2] \leq \lim_{T \rightarrow +\infty} \left(e^{-(2a-\lambda)T} + \mathbb{E} [e^{-\lambda T} |Y_T^x|^2] \right) = 0.$$

Thus, we have

$$u(x) = \mathbb{E} \left[\int_0^{\infty} e^{-as} (f(X_s^x, u(X_s^x), Z_s^x) + au(X_s^x)) ds \right]$$

and the growth on u given by (4.2.9) follows from Proposition 4.2.1 and (4.2.5).

2.a. Now let us assume that $f \in C_b^1(\mathbb{R}^d \times \mathbb{R}^{d'} \times \mathbb{R}^{d' \times d}, \mathbb{R}^{d'})$ and let us consider the following finite horizon BSDE

$$Y_t^{N,x} = \int_t^N f(X_s^x, Y_s^{N,x}, Z_s^{N,x}) ds - \int_t^N Z_s^{N,x} dW_s, \quad 0 \leq t \leq N,$$

with solution $(Y^{N,x}, Z^{N,x}) \in \mathcal{S}_N^2 \times \mathcal{M}_N^2$ defined on $[0, N]$ for any $x \in \mathbb{R}^d$ and $N > 0$. We can set $Y_t^{N,x} := 0$ and $Z_t^{N,x} := 0$ for all $t > N$. Let us remark that for any $\tilde{a} > 0$ we also have

$$Y_t^{N,x} = \int_t^N e^{-\tilde{a}(s-t)} (f(X_s^x, Y_s^{N,x}, Z_s^{N,x}) + \tilde{a}Y_s^{N,x}) ds - \int_t^N e^{-\tilde{a}(s-t)} Z_s^{N,x} dW_s, \quad 0 \leq t \leq N.$$

By applying Theorem 4.2 in [90], we have the representation

$$Y_t^{N,x} = u_N(t, X_t^x), \quad \forall t \in [0, N],$$

with $u_N(t, \cdot) \in C^1(\mathbb{R}^d, \mathbb{R}^{d'})$ for all $t \in [0, N)$ and there exists a continuous version of $Z^{N,x}$ given by

$$Z_t^{N,x} = \nabla_x u_N(t, X_t^x) \sigma(X_t^x) := \bar{u}_N(t, X_t^x), \quad t \in [0, N),$$

and

$$\begin{aligned} \bar{u}_N(t, X_t^x) = & \mathbb{E} \left[\int_t^N e^{-\bar{a}(s-t)} (f(X_s^x, Y_s^{N,x}, Z_s^{N,x}) + \bar{a}Y_s^{N,x}) \right. \\ & \left. \times \frac{1}{s-t} \int_t^s \langle \sigma^{-1}(X_r^x) \nabla_x X_r^x (\nabla X_t^x)^{-1}, dW_r \rangle \sigma(X_t^x) ds \middle| \mathcal{F}_t \right]. \end{aligned}$$

Finally, from proofs of Theorem 5.57 in [98] and Lemma 3.1 in [30] we also have a uniform growth estimate for u_N : there exists a constant $C > 0$ such that

$$|u_N(t, x)| \leq C(1 + |x|^r), \quad \forall t \geq 0, x \in \mathbb{R}^d, N > 0.$$

Now we send N to infinity. In order to do it we denote, for all $x \in \mathbb{R}^d$,

$$\bar{u}(x) = \mathbb{E} \left[\int_0^{+\infty} e^{-\bar{a}s} (f(X_s^x, Y_s^x, Z_s^x) + \bar{a}Y_s^x) U_s^x ds \right],$$

and we firstly show that \bar{u} is well defined. Recalling the growth of u (resp. u_N), we can use the Bismut-Elworthy representation for Z^x (resp. $Z^{N,x}$) on the time interval $[0, 1]$, see e.g. Corollary 3.11 of [60], to get that

$$\|Z_s^x\| + \|Z_s^{N,x}\| \leq C(1 + |X_s^x|^r), \quad \forall (s, x) \in [0, 1] \times \mathbb{R}^d, N \geq 2.$$

Then, we obtain

$$\begin{aligned} & \mathbb{E} \left[\int_0^{+\infty} e^{-\bar{a}s} \| (f(X_s^x, Y_s^x, Z_s^x) + \bar{a}Y_s^x) U_s^x \| ds \right] \\ & \leq C \mathbb{E} \left[\left(1 + \sup_{s \in [0,1]} |X_s^x|^r \right) \int_0^1 |U_s^x| ds \right] \\ & \quad + C \mathbb{E} \left[\int_1^{+\infty} e^{-\bar{a}s} (1 + |X_s^x|^r + \|Z_s^x\|) \left| \int_0^s \langle \sigma^{-1}(X_u^x) \nabla_x X_u^x, dW_u \rangle \right| ds \right] \\ & \leq C(1 + |x|^r) \left(\int_0^1 \frac{ds}{s^{1/2}} \right)^{1/2} \left(\int_0^1 \frac{1}{s^{3/2}} \int_0^s \mathbb{E} [\|\sigma^{-1}(X_u^x) \nabla_x X_u^x\|^2] duds \right)^{1/2} \\ & \quad + C \mathbb{E} \left[\int_0^{+\infty} e^{-\bar{a}s} (1 + |X_s^x|^{2r} + \|Z_s^x\|^2) ds \right]^{1/2} \mathbb{E} \left[\int_0^{+\infty} e^{-\bar{a}s} \|\nabla_x X_s^x\|^2 ds \right]^{1/2} \\ & \leq C(1 + |x|^r) \end{aligned}$$

by applying estimates on Y^x and Z^x given by Proposition 4.2.1, the growth of u , Cauchy-Schwarz inequality, Fubini theorem, (4.2.6), (4.2.5) and standard estimates on X^x and $\nabla_x X^x$.

Now, we estimate the error between $\bar{u}_N(0, \cdot)$ and \bar{u} . Thanks to Assumption 4.2.1 and by using same calculations as previously, we have, for all $x \in \mathbb{R}^d$, $N \geq 2$ and $\varepsilon \in (0, 1]$,

$$\|\nabla_x u_N(0, x) - \bar{u}(x) \sigma^{-1}(x)\|$$

$$\begin{aligned}
&\leq C \|\sigma^{-1}(x)\| \mathbb{E} \left[\int_0^{+\infty} e^{-\tilde{a}s} (|Y_s^{N,x} - Y_s^x| + \|Z_s^{N,x} - Z_s^{t,x}\|) |U_s^x| ds \right] \\
&\quad + C \|\sigma^{-1}(x)\| \mathbb{E} \left[\int_N^{+\infty} e^{-\tilde{a}s} |f(X_s^x, 0, 0)| |U_s^x| ds \right] \\
&\leq C \mathbb{E} \left[\sup_{s \in [0,1]} (|Y_s^x| + |Y_s^{N,x}| + \|Z_s^x\| + \|Z_s^{N,x}\|) \int_0^\varepsilon |U_s^x| ds \right] \\
&\quad + C \mathbb{E} \left[\int_\varepsilon^{+\infty} e^{-\tilde{a}s} |U_s^x|^2 ds \right]^{1/2} \mathbb{E} \left[\int_0^{+\infty} e^{-\tilde{a}s} (|Y_s^{N,x} - Y_s^x|^2 + \|Z_s^{N,x} - Z_s^x\|^2) ds \right]^{1/2} \\
&\quad + C \mathbb{E} \left[\int_1^{+\infty} e^{-\tilde{a}s} \|\nabla_x X_s^x\|^2 ds \right]^{1/2} \mathbb{E} \left[\int_N^{+\infty} e^{-\tilde{a}s} (1 + |X_s^x|^{2r}) ds \right]^{1/2} \\
&\leq C(1 + |x|^r) \sqrt{\varepsilon} \\
&\quad + C_\varepsilon \mathbb{E} \left[\int_0^{+\infty} e^{-\tilde{a}s} \|\nabla_x X_s^x\|^2 ds \right]^{1/2} \mathbb{E} \left[\int_0^{+\infty} e^{-\tilde{a}s} (|Y_s^{N,x} - Y_s^x|^2 + \|Z_s^{N,x} - Z_s^x\|^2) ds \right]^{1/2} \\
&\quad + C \mathbb{E} \left[\int_0^{+\infty} e^{-\tilde{a}s} \|\nabla_x X_s^x\|^2 ds \right]^{1/2} \mathbb{E} \left[\int_N^{+\infty} e^{-\tilde{a}s} (1 + |X_s^x|^{2r}) ds \right]^{1/2}
\end{aligned}$$

and (4.2.6) gives us, for any compact set $K \subset \mathbb{R}^d$,

$$\begin{aligned}
&\sup_{x \in K} \|\nabla_x u_N(0, x) - \bar{u}(x) \sigma^{-1}(x)\| \\
&\leq C_K \sqrt{\varepsilon} + C_K \left(\int_N^{+\infty} e^{-\tilde{a}s} (1 + \sup_{x \in K} \mathbb{E}[|X_s^x|^{2r}]) ds \right)^{1/2} \\
&\quad + C_{K,\varepsilon} \sup_{x \in K} \mathbb{E} \left[\int_0^{+\infty} e^{-\tilde{a}s} (|Y_s^{N,x} - Y_s^x|^2 + \|Z_s^{N,x} - Z_s^x\|^2) ds \right]^{1/2}.
\end{aligned}$$

By using (5.24) in [98] and by checking carefully the proof of Lemma 3.1 in [30], we can easily obtain

$$\lim_{N \rightarrow +\infty} \sup_{x \in K} \mathbb{E} \left[\int_0^{+\infty} e^{-\tilde{a}s} (|Y_s^{N,x} - Y_s^x|^2 + \|Z_s^{N,x} - Z_s^x\|^2) ds \right]^{1/2} = 0,$$

recalling that $\tilde{a} \geq 2\mu - K_{f,z}^2 \mathbf{1}_{d>1}$. Then, we get

$$\limsup_{N \rightarrow +\infty} \sup_{x \in K} \|\nabla_x u_N(0, x) - \bar{u}(x) \sigma^{-1}(x)\| \leq C_K \sqrt{\varepsilon}, \quad \forall \varepsilon \in (0, 1],$$

which implies that

$$\lim_{N \rightarrow +\infty} \sup_{x \in K} \|\nabla_x u_N(0, x) - \bar{u}(x) \sigma^{-1}(x)\| = 0.$$

We also have from proofs of Theorem 5.57 in [98] and Lemma 3.1 in [30] that

$$\lim_{N \rightarrow +\infty} u_N(0, x) = u(x), \quad \text{for all } x \in \mathbb{R}^d.$$

Finally, we can use a theorem of interchange between limit and differentiation to get that $(u_N(0, \cdot), \nabla_x u_N(0, \cdot))$ tends to $(u, \bar{u} \sigma^{-1})$ uniformly on all compacts, $u \in C^1(\mathbb{R}^d, \mathbb{R}^d)$ and $\nabla_x u = \bar{u} \sigma^{-1}$. By standard arguments, see e.g. Corollary 4.1 in [58], $\bar{u}(X^x)$ is also a continuous version of Z^x .

2.b. All calculations done in part 2.a. and estimates obtained do not depend on the derivative of f . Then, we can consider a smooth simple approximation of f by a sequence $(f_n)_{n \in \mathbb{N}}$ and use the same strategy of proof as in 2.a. to tackle the unsmooth setting by sending n to infinity.

4.4.2 Proof of Proposition 4.3.3

Let us denote, for $n \geq 0$,

$$e_{\infty, n+1} := \mathbb{E} \left[\sup_{x \in \mathbb{R}^d} \left\| \frac{Pv_M^{n+1}(x) - v(x)}{\rho(x)} \right\| \right].$$

We have

$$e_{\infty, n+1} \leq \mathcal{E}_{\infty, 1} + \mathcal{E}_{\infty, 2} + \mathcal{E}_{\infty, 3}$$

with

$$\begin{aligned} \mathcal{E}_{\infty, 1} &:= \mathbb{E} \left[\sup_{x \in \mathbb{R}^d} \left\| \frac{Pv_M^{n+1}(x) - P(T_{B, \rho_r}(\mathbb{E}_{v_M^n}[R^{(\cdot)}(Pv_M^n)]))(x)}{\rho(x)} \right\| \right] \\ \mathcal{E}_{\infty, 2} &:= \mathbb{E} \left[\sup_{x \in \mathbb{R}^d} \left\| \frac{Pv(x) - P(T_{B, \rho_r}(\mathbb{E}_{v_M^n}[R^{(\cdot)}(Pv_M^n)]))(x)}{\rho(x)} \right\| \right] \\ \mathcal{E}_{\infty, 3} &:= \sup_{x \in \mathbb{R}^d} \left\| \frac{Pv(x) - v(x)}{\rho(x)} \right\|. \end{aligned}$$

Error $\mathcal{E}_{\infty, 1}$: Recalling that P is linear, $[\cdot]_B$ is 1-Lipschitz and applying inequality 3 in Proposition 4.3.1, we get

$$\begin{aligned} \mathcal{E}_{\infty, 1} &\leq \mathbb{E} \left[\sup_{z \in \Pi} \frac{\rho_r(z)}{\rho(z)} \left\| \left[\mathbb{E}_{v_M^n} \left[\frac{R^z(Pv_M^n)}{\rho_r(z)} \right] \right]_B - \left[\frac{1}{M} \sum_{j=1}^M \frac{R_{n,j}^z(Pv_M^n)}{\rho_r(z)} \right]_B \right\| \right] \sup_{x \in \mathbb{R}^d} \left| \frac{P\rho(x)}{\rho(x)} \right| \\ &\leq \sup_{x \in \mathbb{R}^d} \left| \frac{P\rho(x)}{\rho(x)} \right| \mathbb{E} \left[\sup_{z \in \Pi} \|H_z\| \right], \end{aligned} \quad (4.4.2)$$

where

$$H_z := \frac{1}{M} \sum_{j=1}^M \frac{R_{n,j}^z(Pv_M^n)}{\rho_r(z)} - \mathbb{E}_{v_M^n} \left[\frac{R^z(Pv_M^n)}{\rho_r(z)} \right].$$

Since Ψ is a convex and increasing function, Ψ^{-1} is concave and Jensen inequality gives us

$$\mathbb{E} \left[\sup_{z \in \Pi} \|H_z\| \right] \leq \left| \sup_{z \in \Pi} \|H_z\| \right|_{\Psi} \Psi^{-1} \left(\mathbb{E} \left[\Psi \left(\frac{\sup_{z \in \Pi} \|H_z\|}{\left| \sup_{z \in \Pi} \|H_z\| \right|_{\Psi}} \right) \right] \right) \leq \Psi^{-1}(1) \left| \sup_{z \in \Pi} \|H_z\| \right|_{\Psi}.$$

Then we upper bound the right hand side of the previous inequality by applying Maximal inequality (4.3.3) to get

$$\mathbb{E} \left[\sup_{z \in \Pi} \|H_z\| \right] \leq C \Psi^{-1}(N) \sup_{z \in \Pi} |H_z|_{\Psi}. \quad (4.4.3)$$

Since, conditionally to v_M^n , H_z is a sum of centered independent r.v., we can apply a conditional version of Talagrand inequality (4.3.2):

$$|H_z|_{\Psi, v_M^n} \leq C \left(\mathbb{E} [\|H_z\|] + \left| \sup_{1 \leq j \leq M} \left\| \frac{R_j^z(Pv_M^n)}{\rho_r(z)} - \mathbb{E}_{v_M^n} \left[\frac{R^z(Pv_M^n)}{\rho_r(z)} \right] \right\| \right) \right)_{\Psi, v_M^n}$$

where the notation $\|\cdot\|_{\Psi, v_M^n}$ means that the Orlicz norm is computed conditionally to v_M^n . It gives us, using (4.3.4) and Maximal inequality (4.3.3),

$$\begin{aligned} |H_z|_{\Psi, v_M^n} &\leq C \frac{1}{\sqrt{M}} \mathbb{E}_{v_M^n} \left[\left\| \frac{R^z(Pv_M^n)}{\rho(z)} - \mathbb{E}_{v_M^n} \left[\frac{R^z(Pv_M^n)}{\rho(z)} \right] \right\|^2 \right]^{1/2} \\ &\quad + C \frac{\Psi^{-1}(M)}{M} \left| \frac{R^z(Pv_M^n)}{\rho(z)} - \mathbb{E}_{v_M^n} \left[\frac{R^z(Pv_M^n)}{\rho(z)} \right] \right|_{\Psi, v_M^n} \\ &\leq C(M^{-1/2} + M^{-1}\Psi^{-1}(M)). \end{aligned}$$

Since the previous bound is deterministic, we just have to put it in (4.4.3) and (4.4.2) to get

$$\mathcal{E}_{\infty,1} \leq C \sup_{x \in \mathbb{R}^d} \left| \frac{P\rho(x)}{\rho(x)} \right| \frac{\Psi^{-1}(N)}{\sqrt{M}} \left(1 + \frac{\Psi^{-1}(M)}{\sqrt{M}} \right).$$

Error $\mathcal{E}_{\infty,2}$: Recalling Proposition 4.2.4, we have that v is the unique solution of the fixed point equation $\Phi(v) = v$. Moreover, $B \geq \|v\|_{\rho_r}$ implies that $T_{B,\rho_r}(\Phi(v)) = v$. Then we get, by using the linearity of P , the first inequality in Proposition 4.3.1, the fact that $[\cdot]_B$ is 1-Lipschitz and Proposition 4.2.4,

$$\begin{aligned} &\mathbb{E} \left[\sup_{x \in \mathbb{R}^d} \left\| \frac{Pv(x) - P(T_{B,\rho_r}(\mathbb{E}_{v_M^n}[R^{(\cdot)}(Pv_M^n)]))(x)}{\rho(x)} \right\| \right] \\ &= \mathbb{E} \left[\sup_{x \in \mathbb{R}^d} \left\| \frac{P(T_{B,\rho_r}\Phi(v(\cdot)))(x) - P(T_{B,\rho_r}\Phi(Pv_M^n(\cdot)))(x)}{\rho(x)} \right\| \right] \\ &\leq \mathbb{E} \left[\sup_{x \in \mathbb{R}^d} \left\| \frac{\Phi(v(\cdot))(x) - \Phi(Pv_M^n(\cdot))(x)}{\rho(x)} \right\| \right] \\ &\leq \kappa_{\infty} \mathbb{E} \left[\sup_{x \in \mathbb{R}^d} \left\| \frac{v(x) - Pv_M^n(x)}{\rho(x)} \right\| \right] = \kappa_{\infty} e_{\infty,n}. \end{aligned}$$

Error $\mathcal{E}_{\infty,3}$: We have assumed that v is C^2 . Then, by using the second inequality in Proposition 4.3.1 and the growth of v , we have

$$\begin{aligned} \sup_{x \in \mathbb{R}^d} \left\| \frac{Pv(x) - v(x)}{\rho(x)} \right\| &\leq \sup_{x \in \square} \left\| \frac{Pv(x) - v(x)}{\rho(x)} \right\| + \sup_{x \in \mathbb{R}^d \setminus \square} \left\| \frac{Pv(x) - v(x)}{\rho(x)} \right\| \\ &\leq C\delta^2 + C \sup_{x \in \mathbb{R}^d \setminus \square} \left(\frac{1 + |x|^r}{\rho(x)} \right). \end{aligned}$$

where C does not depend on Π since $\nabla^2 v$ is assumed to be bounded on \mathbb{R}^d .

Error $e_{\infty,n}$: Now we just have to collect previous estimates to get

$$e_{\infty,n+1} \leq C \sup_{x \in \mathbb{R}^d} \left| \frac{P\rho(x)}{\rho(x)} \right| \frac{\Psi^{-1}(N)}{\sqrt{M}} \left(1 + \frac{\Psi^{-1}(M)}{\sqrt{M}} \right) + \kappa_{\infty} e_{\infty,n} + C\delta^2 + C \sup_{x \in \mathbb{R}^d \setminus \square} \left(\frac{1 + |x|^r}{\rho(x)} \right)$$

which gives us

$$e_{\infty,n} \leq C \left(\sum_{k=0}^{n-1} \kappa_{\infty}^k \right) \left(\sup_{x \in \mathbb{R}^d} \left| \frac{P\rho(x)}{\rho(x)} \right| \frac{\Psi^{-1}(N)}{\sqrt{M}} \left(1 + \frac{\Psi^{-1}(M)}{\sqrt{M}} \right) + \delta^2 + \sup_{x \in \mathbb{R}^d \setminus \square} \left(\frac{1 + |x|^r}{\rho(x)} \right) \right) + \kappa_{\infty}^n e_{\infty,0}.$$

4.4.3 Proof of Lemma 4.3.1

1. We have for all $c > 0$, by using the growth of ϕ , and Young inequality

$$\begin{aligned} \mathbb{E} \left[e^{c \left\| \frac{R^z(P\phi)}{\rho(z)} \right\|^{\frac{1}{r+1}}} \right] &\leq \mathbb{E} \left[e^{cC \left(\frac{|U_{\tilde{E}}^z| \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} (1+|X_{\tilde{E}}^z|^r) + e^{-(a-\theta)E} (1+|X_{\tilde{E}}^z|^r)}{\rho(z)} \right)^{\frac{1}{r+1}}} \right] \\ &\leq \mathbb{E} \left[e^{cC \left(\frac{|U_{\tilde{E}}^z| \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} (1+|X_{\tilde{E}}^z|^r)}{\rho(z)} \right)^{\frac{1}{r+1}}} e^{cC \left(\frac{|e^{-(a-\theta)E} (1+|X_{\tilde{E}}^z|^r)}{\rho(z)} \right)^{\frac{1}{r+1}}} \right] \\ &\leq \mathbb{E} \left[e^{cC \left(\frac{|U_{\tilde{E}}^z| \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} (1+|X_{\tilde{E}}^z|^r)}{\rho(z)} \right)^{\frac{1}{r+1}}} \right] + \mathbb{E} \left[e^{cC \left(\frac{e^{-(a-\theta)E} (1+|X_{\tilde{E}}^z|^r)}{\rho(z)} \right)^{\frac{1}{r+1}}} \right] \end{aligned}$$

Young inequality gives us, if $r > 0$,

$$\left(\frac{|U_{\tilde{E}}^z| \sqrt{\tilde{E}} e^{-(\tilde{a}-\tilde{\theta})\tilde{E}} (1+|X_{\tilde{E}}^z|^r)}{\rho(z)} \right)^{\frac{1}{r+1}} \leq e^{-\frac{(\tilde{a}-\tilde{\theta})\tilde{E}}{r+1}} \tilde{E} \left(\frac{|U_{\tilde{E}}^z| \sqrt{\tilde{E}}}{r+1} + \frac{r}{r+1} \left(\frac{(1+|X_{\tilde{E}}^z|^r)}{\rho(z)} \right)^{\frac{1}{r}} \right)$$

and the previous upper-bound stays true without the second term when $r = 0$. Then, Hölder's inequality gives us,

$$\begin{aligned} \mathbb{E} \left[e^{c \left\| \frac{R^z(P\phi)}{\rho(z)} \right\|^{\frac{1}{r+1}}} \right] &\leq e^{cC} \mathbb{E} \left[e^{cC |U_{\tilde{E}}^z| \sqrt{\tilde{E}} e^{-\frac{(\tilde{a}-\tilde{\theta})\tilde{E}}{1+r}}} \right] \mathbb{E} \left[e^{cC \mathbf{1}_{r>0} \frac{|X_{\tilde{E}}^z|}{(1+|z|^{r'})^{1/r}} e^{-\frac{(\tilde{a}-\tilde{\theta})\tilde{E}}{1+r}}} \right] \\ &\quad + e^{cC} \mathbb{E} \left[e^{cC \frac{|X_{\tilde{E}}^z|^{r/(r+1)}}{(1+|z|^{r'})^{1/(r+1)}} e^{-\frac{(a-\theta)E}{r+1}}} \right], \end{aligned}$$

where, as usual, C may change from one term to another but it does not depend on n , M , δ and z . Let us upper bound three previous terms (only two if $r = 0$).

$$\mathbb{E} \left[e^{cC |U_{\tilde{E}}^z| \sqrt{\tilde{E}} e^{-\frac{(\tilde{a}-\tilde{\theta})\tilde{E}}{1+r}}} \right] \leq \int_0^{+\infty} \frac{\sqrt{\tilde{\theta}}}{\sqrt{\pi} \sqrt{t}} \mathbb{E} \left[e^{cC \frac{1}{\sqrt{t}} \left| \int_0^t \langle \sigma^{-1} \nabla_x X_s^z, dW_s \rangle \right| M_\sigma e^{-\frac{(\tilde{a}-\tilde{\theta})\tilde{E}}{1+r} t}} \right] e^{-\tilde{\theta} t} dt.$$

Since $\|\sigma^{-1} \nabla_x X_s^z\| \leq M_{\sigma^{-1}} C e^{K_b t}$ for $0 \leq s \leq t$, it follows from the Dambis-Dubins-Schwarz representation of the continuous martingale $t \mapsto \int_0^t \langle \sigma^{-1} \nabla_x X_s^z, dW_s \rangle$ that

$$\begin{aligned} \mathbb{E} \left[e^{cC \frac{1}{\sqrt{t}} \left| \int_0^t \langle \sigma^{-1} \nabla_x X_s^z, dW_s \rangle \right| M_\sigma e^{-\frac{(\tilde{a}-\tilde{\theta})\tilde{E}}{1+r} t}} \right] &\leq \mathbb{E} \left[\sup_{0 \leq s \leq M_{\sigma^{-1}}^2 C e^{2K_b t}} e^{cC \frac{1}{\sqrt{t}} |W_s| e^{-\frac{(\tilde{a}-\tilde{\theta})\tilde{E}}{1+r} t}} \right] \\ &\leq \mathbb{E} \left[\sup_{0 \leq s \leq 1} e^{cC |W_s| e^{-\frac{(\tilde{a}-\tilde{\theta}-(1+r)K_b)\tilde{E}}{1+r} t}} \right] \\ &\leq \mathbb{E} \left[e^{cC \sup_{0 \leq s \leq 1} W_s} e^{cC \sup_{0 \leq s \leq 1} (-W_s)} \right] \\ &\leq \mathbb{E} \left[e^{cC |W_1|} \right] \leq e^{(c+c^2)C} \end{aligned}$$

using assumption $\tilde{a} - \tilde{\theta} - (1+r)K_b > 0$ and the fact that $\sup_{0 \leq s \leq 1} W_s$ and $|W_1|$ have the same distribution. Thus we get that

$$\mathbb{E} \left[e^{cC |U_{\tilde{E}}^z| \sqrt{\tilde{E}} e^{-\frac{(\tilde{a}-\tilde{\theta})\tilde{E}}{1+r}}} \right] \leq e^{(c+c^2)C}.$$

By mimicking computations of page 563 in [32] and using again previous arguments, we have

$$\begin{aligned} \mathbb{E} \left[e^{cC \frac{|X_{\tilde{E}}^z|}{(1+|z|^{r'})^{1/r}} e^{-\frac{(\tilde{a}-\tilde{\theta})}{1+r} \tilde{E}}} \right] &\leq \int_0^{+\infty} \frac{\sqrt{\tilde{\theta}}}{\sqrt{\pi}\sqrt{t}} \mathbb{E} \left[e^{cC \frac{|z|+t+|\int_0^t \sigma dW_s|}{(1+|z|^{r'})^{1/r}} e^{K_b t} e^{-\frac{(\tilde{a}-\tilde{\theta})}{1+r} t}} \right] e^{-\tilde{\theta}t} dt \\ &\leq \int_0^{+\infty} \frac{\sqrt{\tilde{\theta}}}{\sqrt{\pi}\sqrt{t}} \mathbb{E} \left[\sup_{0 \leq s \leq M_\sigma^2} e^{cC(1+t+\sqrt{t}|W_s|)} e^{-\frac{(\tilde{a}-\tilde{\theta}-(1+r)K_b)t}{1+r}} \right] e^{-\tilde{\theta}t} dt \\ &\leq \int_0^{+\infty} \frac{\sqrt{\tilde{\theta}}}{\sqrt{\pi}\sqrt{t}} \mathbb{E} \left[\sup_{0 \leq s \leq 1} e^{cC(1+|W_s|)} \right] e^{-\tilde{\theta}t} dt \leq e^{(c+c^2)C}. \end{aligned}$$

By the same token, recalling that $a - \tilde{\theta} - rK_b > 0$, we also have

$$\mathbb{E} \left[e^{cC \frac{|X_{\tilde{E}}^z|^{r/(r+1)}}{(1+|z|^{r'})^{1/(r+1)}} e^{-\frac{(a-\tilde{\theta})}{r+1} E}} \right] \leq e^{(c+c^2)C}.$$

By collecting previous estimates, we finally get

$$\mathbb{E} \left[e^{c \left\| \frac{R^z(\phi)}{\rho(z)} \right\|^{\frac{1}{r+1}}} \right] \leq e^{(c+c^2)C}. \quad (4.4.4)$$

Moreover, we also have

$$\begin{aligned} \left\| \mathbb{E} \left[\left(\frac{R^z(P\phi)}{\rho(z)} \right) \right] \right\| &\leq \mathbb{E} \left[\left\| \frac{R^z(P\phi)}{\rho(z)} \right\|^2 \right]^{1/2} \\ &\leq C \left(\int_0^{+\infty} \frac{1}{\sqrt{t}} \mathbb{E} \left[\frac{t|U_t^z|^2(1+|X_t^z|^{2r})}{\rho(z)^2} \right] e^{-(2\tilde{a}-\tilde{\theta})t} dt \right)^{1/2} \\ &\quad + C \left(\int_0^{+\infty} \mathbb{E} \left[\frac{1+|X_t^z|^{2r}}{\rho(z)^2} \right] e^{-(2a-\theta)t} dt \right)^{1/2} \\ &\leq C \left(\int_0^{+\infty} \frac{1}{\sqrt{t}} \mathbb{E} [t^2|U_t^z|^4]^{1/2} \mathbb{E} \left[\frac{1+|X_t^z|^{4r}}{\rho(z)^4} \right]^{1/2} e^{-(2\tilde{a}-\tilde{\theta})t} dt \right)^{1/2} \\ &\quad + C \left(\int_0^{+\infty} \mathbb{E} \left[\frac{1+|X_t^z|^{2r}}{\rho(z)^2} \right] e^{-(2a-\theta)t} dt \right)^{1/2}. \end{aligned}$$

Then, using Remark 4.2.2, we get, for all $\varepsilon > 0$,

$$\begin{aligned} \mathbb{E} \left[\left\| \frac{R^z(P\phi)}{\rho(z)} \right\|^2 \right]^{1/2} &\leq C_\varepsilon \left(\int_0^{+\infty} \frac{1}{\sqrt{t}} e^{(2K_b+3K_\sigma^2)t} e^{(2rK_b+\varepsilon)t} e^{-(2\tilde{a}-\tilde{\theta})t} dt \right)^{1/2} \\ &\quad + C_\varepsilon \left(\int_0^{+\infty} e^{(2rK_b+\varepsilon)t} e^{-(2a-\theta)t} dt \right)^{1/2}. \end{aligned}$$

This upper bound is finite for ε small enough, recalling that $\tilde{a}-\tilde{\theta}-(1+r)K_b > 0$, $a-\tilde{\theta}-rK_b > 0$ and $K_\sigma = 0$ since σ is constant. In particular, it implies the second part of (4.3.4). Finally, by using previous bound and (4.4.4), we also have

$$\mathbb{E} \left[\Psi_{1/(r+1)}^{LT} \left(c \frac{R^z(P\phi)}{\rho(z)} - c \mathbb{E} \left[\frac{R^z(P\phi)}{\rho(z)} \right] \right) \right] \leq e^{(c+c^2)C} - 1.$$

Since $c \mapsto e^{(c+c^2)C} - 1$ is a continuous increasing function from \mathbb{R}^+ to \mathbb{R}^+ , we can conclude that the first part of (4.3.4) also holds true.

2. When σ is not constant, the proof of point 1. becomes false only due to the fact that $\nabla_x X^z$ has not a deterministic upper-bound, which hinder the treatment of the term

$$\mathbb{E} \left[e^{cC|U_{\tilde{E}}^z| \sqrt{\tilde{E}} e^{-\frac{(\tilde{a}-\tilde{\theta})}{1+r}\tilde{E}}} \right].$$

Nevertheless, this problem disappears when $f(x, y, z) = f(x, y)$. Then the proof of point 2. follows the same proof as the proof of point 1. remarking that it is enough to take the power $1/(r \vee 1)$ instead of $1/(r + 1)$.

3. As in the proof of point 1. we have

$$\begin{aligned} \left\| \mathbb{E} \left[\left(\frac{R^z(P\phi)}{\rho(z)} \right) \right] \right\| &\leq \mathbb{E} \left[\left\| \frac{R^z(P\phi)}{\rho(z)} \right\|^2 \right]^{1/2} \leq \mathbb{E} \left[\left\| \frac{R^z(P\phi)}{\rho(z)} \right\|^\beta \right]^{1/\beta} \\ &\leq C \left(\int_0^{+\infty} \frac{1}{\sqrt{t}} \mathbb{E} \left[\frac{t^{\beta/2} |U_t^z|^\beta (1 + |X_t^z|^{\beta r})}{\rho(z)^\beta} \right] e^{-(\beta\tilde{a} - (\beta-1)\tilde{\theta})t} dt \right)^{1/\beta} \\ &\quad + C \left(\int_0^{+\infty} \mathbb{E} \left[\frac{1 + |X_t^z|^{\beta r}}{\rho(z)^\beta} \right] e^{-(\beta a - (\beta-1)\theta)t} dt \right)^{1/\beta} \\ &\leq C \left(\int_0^{+\infty} \frac{1}{\sqrt{t}} \mathbb{E} \left[t^\beta |U_t^z|^{2\beta} \right]^{1/2} \mathbb{E} \left[\frac{1 + |X_t^z|^{2\beta r}}{\rho(z)^{2\beta}} \right]^{1/2} e^{-(\beta\tilde{a} - (\beta-1)\tilde{\theta})t} dt \right)^{1/\beta} \\ &\quad + C \left(\int_0^{+\infty} \mathbb{E} \left[\frac{1 + |X_t^z|^{\beta r}}{\rho(z)^\beta} \right] e^{-(\beta a - (\beta-1)\theta)t} dt \right)^{1/\beta}. \end{aligned}$$

Thanks to Burkholder-Davis-Gundy and Jensen inequality, we have

$$\mathbb{E} \left[t^\beta |U_t^z|^{2\beta} \right] \leq C \mathbb{E} \left[\left(\frac{1}{t} \int_0^t \|\nabla_x X_s\|^2 ds \right)^\beta \right] \leq C \frac{1}{t} \int_0^t \mathbb{E} \left[\|\nabla_x X_s\|^{2\beta} \right] ds.$$

Then, using Remark 4.2.2, we get, for all $\varepsilon > 0$,

$$\begin{aligned} \mathbb{E} \left[\left\| \frac{R^z(P\phi)}{\rho(z)} \right\|^\beta \right]^{1/\beta} &\leq C_\varepsilon \left(\int_0^{+\infty} \frac{1}{\sqrt{t}} e^{(\beta K_b + \frac{\beta(2\beta-1)}{2} K_\sigma^2)t} e^{(\beta r K_b + \varepsilon)t} e^{-(\beta\tilde{a} - (\beta-1)\tilde{\theta})t} dt \right)^{1/\beta} \\ &\quad + C_\varepsilon \left(\int_0^{+\infty} e^{(\beta r K_b + \varepsilon)t} e^{-(\beta a - (\beta-1)\theta)t} dt \right)^{1/\beta}. \end{aligned}$$

This upper bound is finite for ε small enough, recalling that $\tilde{a} - (r+1)K_b - \frac{2\beta-1}{2}K_\sigma^2 - \frac{\beta-1}{\beta}\tilde{\theta} > 0$ and $a - rK_b - \frac{\beta-1}{\beta}\theta > 0$. In particular, it implies (4.3.4).

Bibliography

- [1] Beatrice Acciaio, Julio Backhoff-Veraguas, and René Carmona. Extended mean field control problems: stochastic maximum principle and transport perspective. *SIAM journal on Control and Optimization*, 57(6):3666–3693, 2019.
- [2] Yves Achdou and Italo Capuzzo-Dolcetta. Mean field games: numerical methods. *SIAM Journal on Numerical Analysis*, 48(3):1136–1162, 2010.
- [3] René Aïd, Matteo Basei, and Huyên Pham. A McKean–Vlasov approach to distributed electricity generation development. *Mathematical Methods of Operations Research*, 91: 269–310, 2020.
- [4] Clémence Alasseur, Imen Ben Taher, and Anis Matoussi. An extended mean field game for storage in smart grids. *Journal of Optimization Theory and Applications*, 184:644–670, 2020.
- [5] Alfonsi and Jourdain. Squared quadratic Wasserstein distance: optimal couplings and Lions differentiability. *ESAIM: PS*, 24:703–717, 2020. doi: 10.1051/ps/2020013. URL <https://doi.org/10.1051/ps/2020013>.
- [6] Hisham Alharbi and Kankar Bhattacharya. Stochastic optimal planning of battery energy storage systems for isolated microgrids. *IEEE Transactions on Sustainable Energy*, 9(1): 211–227, 2018. doi: 10.1109/TSTE.2017.2724514.
- [7] Hassan Haes Alhelou, Mohamad-Esmail Hamedani-Golshan, Reza Zamani, Ehsan Heydarian-Forushani, and Pierluigi Siano. Challenges and opportunities of load frequency control in conventional, modern and future smart power systems: A comprehensive review. *Energies*, 11(10):2497, 2018.
- [8] Manijeh Alipour, Kazem Zare, and Heresh Seyedi. A multi-follower bilevel stochastic programming approach for energy management of combined heat and power microgrids. *Energy*, 149(C):135–146, 2018. doi: 10.1016/j.energy.2018.02. URL <https://ideas.repec.org/a/eee/energy/v149y2018icp135-146.html>.
- [9] Daniel Andersson and Boualem Djehiche. A maximum principle for SDEs of mean-field type. *Applied Mathematics & Optimization*, 63(3):341–356, 2011.
- [10] Vincent Bally and Gilles Pagès. A quantization algorithm for solving multidimensional discrete-time optimal stopping problems. *Bernoulli*, 9(6):1003–1049, 2003.
- [11] Anne G. Balter and Antoon Pelsser. Pricing and hedging in incomplete markets with model uncertainty. *European Journal of Operational Research*, 282(3):911–925, 2020. ISSN 0377-2217. doi: <https://doi.org/10.1016/j.ejor.2019.09.054>. URL <https://www.sciencedirect.com/science/article/pii/S0377221719308112>.
- [12] Guy Barles. *Solutions de Viscosité des Équations de Hamilton-Jacobi*. Mathématiques & Applications. Springer-Verlag, 1994.

- [13] Guy Barles. *Convergence of Numerical Schemes for Degenerate Parabolic Equations Arising in Finance Theory*, page 1–21. Publications of the Newton Institute. Cambridge University Press, 1997.
- [14] Julian Barreiro-Gomez and Hamidou Tembine. Mean-field-type model predictive control: An application to water distribution networks. *IEEE Access*, 7:135332–135339, 2019.
- [15] Christian Beck, Lukas Gonon, and Arnulf Jentzen. Overcoming the curse of dimensionality in the numerical approximation of high-dimensional semilinear elliptic partial differential equations, 2020.
- [16] Richard Bellman. *Dynamic Programming*. Princeton University Press, 1957.
- [17] Christian Bender and Jessica Steiner. Least-Squares Monte Carlo for Backward SDEs. *Annals of Applied Probability*, 15:2172–2202, 2005.
- [18] Alain Bensoussan, Jens Frehse, Phillip Yam, et al. *Mean field games and mean field type control theory*, volume 101. Springer, 2013.
- [19] Fred Espen Benth and Juratė Šaltytė-Benth. Stochastic modelling of temperature variations with a view towards weather derivatives. *Applied Mathematical Finance*, 12(1): 53–85, 2005.
- [20] Marie Bernhart, Huyên Pham, Peter Tankov, and Xavier Warin. Swing options valuation: A BSDE with constrained jumps approach. In *Numerical Methods in Finance*, pages 379–400. Springer, 2012.
- [21] Dimitris Bertsimas and Andrew Lo. Optimal Control of Execution Costs. *Journal of Financial Markets*, 1(1):1–50, 1998.
- [22] Charles Bertucci. Optimal stopping in mean field games, an obstacle problem approach. *Journal de Mathématiques Pures et Appliquées*, 120:165–194, 2018. ISSN 0021-7824. doi: <https://doi.org/10.1016/j.matpur.2017.09.016>. URL <https://www.sciencedirect.com/science/article/pii/S0021782417301411>.
- [23] Jean-Michel Bismut. Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, 44:384–404, 1973.
- [24] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.
- [25] Vladimir I Bogachev. Fokker–Planck–Kolmogorov equations. *Mathematical Surveys and Monographs*, 207, 2017.
- [26] Bruno Bouchard and Nizar Touzi. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Processes and their Applications*, 111(2):175–206, 2004.
- [27] Bruno Bouchard and Xavier Warin. Monte-Carlo valuation of American options: facts and new algorithms to improve existing methods. In R. Carmona, P. Del Moral, P. Hu, and N. Oudjane, editors, *Numerical Methods in Finance*, volume 12 of *Springer Proceedings in Mathematics*, pages 215–255. Springer, 2012.
- [28] Jean-Philippe Bouchaud, Marc Mezard, and Marc Potters. Statistical properties of stock order books: Empirical results and models. *Quantitative Finance*, 2(4):251–256, 2002.
- [29] Philippe Briand and Fulvio Confortola. Stability of backward stochastic differential equations with jumps and applications. *The Annals of Probability*, 36(3):973–986, 2008.

- [30] Philippe Briand and Ying Hu. Stability of BSDEs with random terminal time and homogenization of semilinear elliptic PDEs. *Journal of Functional Analysis*, 155(2):455–494, 1998.
- [31] Philippe Briand and Ying Hu. BSDE with quadratic growth and unbounded terminal value. *Probab. Theory Relat. Fields*, 136(4):604–618, 2006. ISSN 0178-8051. doi: 10.1007/s00440-006-0497-0.
- [32] Philippe Briand and Ying Hu. Quadratic BSDEs with convex generators and unbounded terminal conditions. *Probab. Theory Related Fields*, 141(3-4):543–567, 2008. ISSN 0178-8051.
- [33] Rainer Buckdahn, Yaozhong Hu, and Shige Peng. Probabilistic Approach to Homogenizations of Systems of Semilinear Parabolic PDE’s. Preprint, 1996.
- [34] Rainer Buckdahn, Juan Li, and Shige Peng. Nonlinear stochastic differential games involving a major player and a large number of collectively acting minor agents. *SIAM Journal on Control and Optimization*, 52(1):451–492, 2014. doi: 10.1137/130933174. URL <https://doi.org/10.1137/130933174>.
- [35] Pierre Cardaliaguet and Charles-Albert Lehalle. Mean field game of controls and an application to trade crowding. *Mathematics and Financial Economics*, 12:335–363, 2018.
- [36] René Carmona and François Delarue. *Probabilistic Theory of Mean Field Games with Applications I: Mean Field FBSDEs, Control, and Games*. Springer International Publishing, 2018.
- [37] René Carmona and François Delarue. *Probabilistic theory of mean field games with applications I-II*. Springer, 2018.
- [38] René Carmona and Mathieu Laurière. Convergence analysis of machine learning algorithms for the numerical solution of mean field control and games: Ii—the finite horizon case. *The Annals of Applied Probability*, 32(6):4065–4105, 2022.
- [39] René Carmona, François Delarue, and Daniel Lacker. Mean-field forward-backward stochastic differential equations. *The Annals of Probability*, 46(2):1115–1201, 2018.
- [40] Annalisa Cesaroni, Marco Cirant, Serena Dipierro, Matteo Novaga, and Enrico Valdinoci. On stationary fractional mean field games. *Journal de Mathématiques Pures et Appliquées*, 122:1–22, 2019. ISSN 0021-7824. doi: <https://doi.org/10.1016/j.matpur.2017.10.013>. URL <https://www.sciencedirect.com/science/article/pii/S0021782417301642>.
- [41] George Chacko and Luis M Viceira. Dynamic consumption and portfolio choice with stochastic volatility in incomplete markets. *The Review of Financial Studies*, 18(4):1369–1402, 2005.
- [42] Linda Chamakh, Emmanule Gobet, and Wenjun Liu. Orlicz norms and concentration inequalities for β -heavy tailed random variables. hal-03175697.
- [43] Jean-François Chassagneux and Adrien Richou. Numerical simulation of quadratic BSDEs. 2016.
- [44] Jean-François Chassagneux, Dan Crisan, and François Delarue. Numerical method for FBSDEs OF McKean–Vlasov type. *The Annals of Applied Probability*, 29(3):pp. 1640–1684, 2019. ISSN 10505164, 21688737. URL <https://www.jstor.org/stable/26729310>.

- [45] Emmanuelle Clément, Damien Lamberton, and Philip Protter. An analysis of a least squares regression method for American option pricing. *Finance and Stochastics*, 6(4):449, 2002.
- [46] Rama Cont and Arseniy Kukanov. Optimal order placement in limit order markets. *Management Science*, 60(11):2732–2747, 2014.
- [47] Stéphane Crépey and Shiqi Song. BSDEs of counterparty risk. *Stochastic Processes and their Applications*, 125(8):3023–3052, 2015. ISSN 0304-4149. doi: <https://doi.org/10.1016/j.spa.2015.02.010>. URL <https://www.sciencedirect.com/science/article/pii/S0304414915000605>.
- [48] Richard WR Darling and Etienne Pardoux. Backwards SDE with random terminal time and applications to semilinear elliptic PDE. *The Annals of Probability*, 25(3):1135–1159, 1997.
- [49] Antonio De Paola, David Angeli, and Goran Strbac. Distributed control of micro-storage devices with mean field games. *IEEE Transactions on Smart Grid*, 7(2):1119–1127, 2015.
- [50] Pierre Del Moral, Bruno Rémillard, and Sylvain Rubenthaler. Monte carlo approximations of american options that preserve monotonicity and convexity. In *Numerical Methods in Finance: Bordeaux, June 2010*, pages 115–143. Springer, 2012.
- [51] Freddy Delbaen, Ying Hu, and Adrien Richou. On the uniqueness of solutions to quadratic BSDEs with convex generators and unbounded terminal conditions. In *Annales de l’IHP Probabilités et statistiques*, volume 47, pages 559–574, 2011.
- [52] Łukasz Delong. *Backward stochastic differential equations with jumps and their actuarial and financial applications*. Springer, 2013.
- [53] Avinash K Dixit and Robert S Pindyck. *Investment under uncertainty*. Princeton University Press, 1994.
- [54] Boualem Djehiche, Romuald Elie, and Said Hamadène. Mean-field reflected backward stochastic differential equations. *arXiv preprint arXiv:1911.06079*, 2019.
- [55] Darrell Duffie and Jun Pan. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, 2000.
- [56] Weinan E, Jiequn Han, and Arnulf Jentzen. Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. *Communications in Mathematics and Statistics*, 5(4):349–380, 2017.
- [57] Nicole El Karoui, Shige Peng, and Marie Christine Quenez. A general result of existence and uniqueness of backward stochastic differential equations. *Probability Theory and Related Fields*, 97(4):553–564, 1993.
- [58] Nicole El Karoui, Shige Peng, and Marie C. Quenez. Backward stochastic differential equations in finance. *Math. Finance*, 7(1):1–71, 1997. ISSN 0960-1627.
- [59] Wendell H Fleming and Halil Mete Soner. Optimal investment strategies for constrained investors. *SIAM Journal on Control and Optimization*, 35(6):1780–1811, 1997.
- [60] Marco Fuhrman and Gianmario Tessitore. The Bismut-Elworthy formula for backward SDEs and applications to nonlinear Kolmogorov equations and control in infinite dimensional spaces. *Stoch. Stoch. Rep.*, 74(1-2):429–464, 2002. ISSN 1045-1129. doi: [10.1080/10451120290024856](https://doi.org/10.1080/10451120290024856). URL <http://dx.doi.org/10.1080/10451120290024856>.

- [61] Wilfrid Gangbo and Adrian Tudorascu. On differentiability in the Wasserstein space and well-posedness for Hamilton–Jacobi equations. *Journal de Mathématiques Pures et Appliquées*, 125:119–174, 2019.
- [62] Maximilien Germain, Huyên Pham, and Xavier Warin. A level-set approach to the control of state-constrained McKean-Vlasov equations: application to renewable energy storage and portfolio selection. *arXiv preprint arXiv:2112.11059*, 2021.
- [63] Maximilien Germain, Joseph Mikael, and Xavier Warin. Numerical resolution of McKean-Vlasov FBSDEs using neural networks. *Methodology and Computing in Applied Probability*, 24(4):2557–2586, 2022.
- [64] Emmanuel Gobet and Maxime Grangereau. Extended McKean-Vlasov optimal stochastic control applied to smart grid management. 2021.
- [65] Emmanuel Gobet and Maxime Grangereau. Newton method for stochastic control problems. *SIAM Journal on Control and Optimization*, 60(5):2996–3025, 2022.
- [66] Emmanuel Gobet and Céline Labart. Solving BSDE with adaptive control variate. *SIAM Journal on Numerical Analysis*, 48(1):257–277, 2010.
- [67] Emmanuel Gobet and Jean-Philippe Lemor. Numerical simulation of BSDEs using empirical regression methods: theory and practice. *arXiv preprint arXiv:0806.4447*, 2008.
- [68] Emmanuel Gobet and Plamen Turkedjiev. Approximation of backward stochastic differential equations using Malliavin weights and least-squares regression. *Bernoulli*, 22(1):530–562, 2016.
- [69] Emmanuel Gobet and Plamen Turkedjiev. Linear regression MDP scheme for discrete backward stochastic differential equations under general conditions. *Mathematics of Computation*, 85(299):1359–1391, 2016.
- [70] Emmanuel Gobet, Jean-Philippe Lemor, and Xavier Warin. A regression-based Monte Carlo method to solve backward stochastic differential equations. 2005.
- [71] Emmanuel Gobet, Javier Lopez-Salas, Plamen Turkedjiev, and Carlos Vasquez. Stratified regression Monte-Carlo scheme for semilinear PDEs and BSDEs with large scale parallelization on GPUs. *SIAM Journal of Scientific Computing*, 38(6):C652–C677, 2016.
- [72] Emmanuel Gobet, José Germán López-Salas, and Carlos Vázquez. Quasi-Regression Monte-Carlo scheme for semi-linear PDEs and BSDEs with large scale parallelization on GPUs. *Archives of Computational Methods in Engineering*, 27(3):889–921, April 2019. doi: 10.1007/s11831-019-09335-x. URL <https://hal.science/hal-01904457>.
- [73] Diogo A Gomes and João Saúde. Mean field games models—a brief survey. *Dynamic Games and Applications*, 4:110–154, 2014.
- [74] Julien Guyon and Pierre Henry-Labordère. The uncertain volatility model: a Monte Carlo approach. *The Journal of Computational Finance*, 14(3):37, 2011.
- [75] O. Guéant, C.A. Lehalle, and J.F. Tapia. Optimal portfolio liquidation with limit orders. *Market Microstructure and Liquidity*, 1(01):1350001, 2013.
- [76] Faeza Hafiz, MA Awal, Anderson Rodrigo de Queiroz, and Iqbal Husain. Real-time stochastic optimization of energy storage management using deep learning-based forecasts for residential pv applications. *IEEE Transactions on Industry Applications*, 56(3):2216–2226, 2020.

- [77] Jiequn Han and Jihao Long. Convergence of the deep BSDE method for coupled FBSDEs. *Probability, Uncertainty and Quantitative Risk*, 5(1):1–33, 2020.
- [78] Jiequn Han, Ruimeng Hu, and Jihao Long. Learning high-dimensional McKean-Vlasov forward-backward stochastic differential equations with general distribution dependence. *arXiv preprint arXiv:2204.11924*, 2022.
- [79] Kurt Hornik. Approximation capabilities of multilayer feedforward networks. *Neural networks*, 4(2):251–257, 1991.
- [80] Minyi Huang, Peter E Caines, and Roland P Malhame. Large population stochastic dynamic games: closed-loop McKean–Vlasov systems and the Nash certainty equivalence principle. *Communications in Information & Systems*, 6(3):221–252, 2006.
- [81] Jiongmin Yong Jin Ma. *Method of Optimal Control*, pages 51–79. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007. ISBN 978-3-540-48831-6. doi: 10.1007/978-3-540-48831-6_3. URL https://doi.org/10.1007/978-3-540-48831-6_3.
- [82] Rudolf E. Kalman. The theory of optimal control and the calculus of variations. *Bulletin of the American Mathematical Society*, 66(4):276–281, 1960.
- [83] Idris Kharroubi and Huy en Pham. Optimal portfolio liquidation with execution cost and risk. *Mathematics of Operations Research*, 39(4):933–961, 2014.
- [84] Nicolai V. Krylov. *Controlled diffusion processes*, volume 14 of *Applications of Mathematics*. Springer-Verlag, New York-Berlin, 1980. ISBN 0-387-90461-1. Translated from the Russian by A. B. Aries.
- [85] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese journal of mathematics*, 2(1):229–260, 2007.
- [86] Moshe Leshno, Vladimir Ya Lin, Allan Pinkus, and Shimon Schocken. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. *Neural networks*, 6(6):861–867, 1993.
- [87] Feng Li, Roland P Malham e, and Jerome Le Ny. Mean field game based control of dispersed energy storage devices with constrained inputs. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 4861–4866. IEEE, 2016.
- [88] Xun Li, Jingtao Shi, and Jiongmin Yong. Mean-field linear-quadratic stochastic differential games in an infinite horizon. *arXiv preprint arXiv:2007.06130*, 2020.
- [89] Jin Ma and Jiongmin Yong. *Forward-backward stochastic differential equations and their applications*. Number 1702. Springer Science & Business Media, 1999.
- [90] Jin Ma and Jianfeng Zhang. Representation theorems for backward stochastic differential equations. *Ann. Appl. Probab.*, 12(4):1390–1418, 2002.
- [91] Jin Ma, Philip Protter, Juan San Mart ın, and Soledad Torres. Numerical method for backward stochastic differential equations. *Annals of Applied Probability*, 12(1):302–316, 2002.
- [92] Jin Ma, Zhen Wu, Detao Zhang, and Jianfeng Zhang. On well-posedness of forward–backward SDEs—A unified approach. *The Annals of Applied Probability*, 25(4), August 2015. ISSN 1050-5164. doi: 10.1214/14-aap1046. URL <http://dx.doi.org/10.1214/14-AAP1046>.

- [93] Jean-François Mertens and Sylvain Sorin. *Game-theoretic methods in general equilibrium analysis*, volume 77. Springer Science & Business Media, 2013.
- [94] Robert C Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics*, pages 247–257, 1969.
- [95] Robert C Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1-2):125–144, 1976.
- [96] N. Nüsken and L. Richter. Interpolating between BSDEs and PINNs – deep learning for elliptic and parabolic boundary value problems, 2021. URL <https://arxiv.org/abs/2112.03749>.
- [97] Shin-Ichi Ohta and Asuka Takatsu. Displacement convexity of generalized relative entropies. ii. *arXiv preprint arXiv:1112.5554*, 2011.
- [98] E. Pardoux and A. Răşcanu. *Stochastic Differential Equations, Backward SDEs, Partial Differential Equations*, volume 69 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, 2014.
- [99] Etienne Pardoux and Shige Peng. Adapted solutions of backward stochastic differential equations. *Systems & Control Letters*, 14(1):55–61, 1990.
- [100] Etienne Pardoux and Aurel Răşcanu. *Stochastic differential equations, Backward SDEs, Partial differential equations*, volume 69. Springer, 2014.
- [101] Etienne Pardoux, Frédéric Pradeilles, and Zusheng Rao. Probabilistic interpretation of a system of semi-linear parabolic partial differential equations. *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*, 33(4):467–490, 1997. ISSN 0246-0203. doi: [https://doi.org/10.1016/S0246-0203\(97\)80101-X](https://doi.org/10.1016/S0246-0203(97)80101-X). URL <https://www.sciencedirect.com/science/article/pii/S024602039780101X>.
- [102] Étienne Pardoux, Jean-Michel Lasry, and Pierre-Louis Lions. Mean field games and applications. *Paris-Princeton Lectures on Mathematical Finance 2013*, pages 200–266, 2014.
- [103] Shige Peng. *Backward Stochastic Differential Equation, Nonlinear Expectation and Their Applications*, pages 393–432. doi: 10.1142/9789814324359_0019. URL https://www.worldscientific.com/doi/abs/10.1142/9789814324359_0019.
- [104] Shige Peng. Stochastic Hamilton-Jacobi-Bellman Equations. *Journal Name*, Volume:Page Range, 1992.
- [105] Huyên Pham and Xiaoli Wei. Bellman equation and viscosity solutions for mean-field stochastic control problem. *ESAIM: Control, Optimisation and Calculus of Variations*, 24(1):437–461, 2018.
- [106] Huyên Pham and Xavier Warin. Mean-field neural networks: learning mappings on Wasserstein space. *arXiv preprint arXiv:2210.15179*, 2022.
- [107] Lev S Pontryagin, Vladimir G Boltyanskii, Ruvim V Gamkrelidze, and Evgenii F Mishchenko. *Mathematical theory of optimal processes*. Interscience, New York, 1962.
- [108] Maria L Rizzo and Gábor J Székely. Energy distance. *wiley interdisciplinary reviews: Computational statistics*, 8(1):27–38, 2016.

- [109] Lars Ruthotto, Stanley J Osher, Wuchen Li, Levon Nurbekyan, and Samy Wu Fung. A machine learning framework for solving high-dimensional mean field game and mean field control problems. *Proceedings of the National Academy of Sciences*, 117(17):9183–9193, 2020.
- [110] Myron Scholes and Fischer Black. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.
- [111] Eduardo S Schwartz and James E Smith. Short-term variations and long-term dynamics in commodity prices. *Management Science*, 46(7):893–911, 2000.
- [112] Halil Mete Soner, Nizar Touzi, and Jianfeng Zhang. Stochastic maximum principle and potential games. *SIAM Journal on Control and Optimization*, 46(2):388–404, 2007.
- [113] M. Talagrand. Isoperimetry and integrability of the sum of independent Banach-space valued random variables. *Ann. Probab.*, 17(4):1546–1570, 1989. ISSN 0091-1798. URL [http://links.jstor.org/sici?sici=0091-1798\(198910\)17:4<1546:IAIOTS>2.0.CO;2-T&origin=MSN](http://links.jstor.org/sici?sici=0091-1798(198910)17:4<1546:IAIOTS>2.0.CO;2-T&origin=MSN).
- [114] Hamidou Tembine, Quanyan Zhu, and Tamer Başar. Risk-sensitive mean-field games. *IEEE Transactions on Automatic Control*, 59(4):835–850, 2013.
- [115] Aad W. van der Vaart and Jon A. Wellner. *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer Series in Statistics. Springer-Verlag, New York, 1996. ISBN 0-387-94640-3. doi: 10.1007/978-1-4757-2545-2. URL <https://doi.org/10.1007/978-1-4757-2545-2>.
- [116] Jiongmin Yong. Linear forward-backward stochastic differential equations with random coefficients. *Probability theory and related fields*, 135(1):53–83, 2006.
- [117] Jianjun Zhou, Nizar Touzi, and Jianfeng Zhang. Viscosity Solutions for HJB Equations on the Process Space: Application to Mean Field Control with Common Noise. *arXiv preprint arXiv:2401.04920*, 2024.
- [118] Karl J. Åström. *Introduction to Stochastic Control Theory*. Academic Press, 1970.

Contrôle stochastique de type champ moyen en horizon infini et approximation numérique des équations différentielles stochastiques rétrogrades en horizon infini

Résumé :

Dans la première partie de la thèse, nous étudions un problème de contrôle stochastique de type champ moyen en horizon infini où la fonction de coût dépend de la loi du processus d'état. Nous prouvons les conditions nécessaires et suffisantes d'optimalité qui exigent la L-différentiabilité et la L-convexité dans l'espace des mesures pour la fonction de coût instantané. Ensuite, nous commençons par une application en optimisation de portefeuille du problème de contrôle de type champ-moyen en horizon infini. L'objectif est de surperformer une allocation statique dans un portefeuille investi en actions et en actif sans risque en s'appuyant sur une allocation dynamique, en utilisant la vitesse de traitement des actifs comme contrôle avec un critère de minimisation du risque baissier de type champ-moyen. Nous avons prouvé les conditions d'optimalité pour le problème de contrôle et établi l'existence et l'unicité de la solution du système couplé d'équations différentielles stochastiques progressives-rétrogrades de type McKean-Vlasov. Nous avons également développé un schéma numérique basé sur les réseaux de neurones pour résoudre une version tronquée dans le temps du problème et fourni des bornes exponentielles pour l'erreur de troncature. Les expériences numériques suggèrent qu'augmenter le multiplicateur du terme de champ moyen incline avec succès la distribution de richesse vers la droite, augmentant ainsi la probabilité d'obtenir une richesse relative plus élevée.

Dans une seconde partie, nous étudions l'approximation numérique des équations différentielles stochastiques rétrogrades en horizon infini. Nous avons développé trois schémas numériques. Le premier schéma est basé sur une itération de Picard et utilise une approximation spatiale par grille. Le deuxième schéma est également basé sur un schéma de Picard mais utilise des réseaux de neurones. Le troisième schéma n'est pas basé sur une itération de Picard et repose sur des réseaux de neurone. Nous fournissons également une étude détaillée de l'erreur numérique pour le premier schéma et prouvons des bornes fines sur l'erreur d'approximation, nécessitant des hypothèses supplémentaires pour la contraction. Pour le deuxième schéma, nous prouvons la convergence de l'erreur d'approximation vers zéro lorsque la taille du réseau de neurones augmente. Les expérimentations numériques suggèrent que le troisième schéma est plus performant que les deux premiers lorsque la contraction n'est plus satisfaite.

Mots-clés : contrôle stochastique, contrôle de type champ moyen, équations différentielles stochastiques rétrogrades, optimisation de portefeuille

Mean-field infinite horizon stochastic control and numerical approximation of infinite horizon backward stochastic differential equations

Abstract:

In the first part of this thesis, we study a mean-field stochastic control problem in infinite horizon where the cost functional has dependence on the law of the state process. We prove the necessary and sufficient conditions of optimality which requires L-differentiability and L-convexity in the measure space for the running cost function. Then, we start with an application in portfolio optimization of this mean-field control problem in the infinite horizon. The goal is to outperform a static allocation of stocks and a risk-free asset by using dynamic allocation, using the trading speed of the assets as the control with a downside risk minimization criterion of mean-field type. We prove the conditions of optimality for the control problem and establish the existence and uniqueness of the solution to the corresponding system of coupled McKean-Vlasov forward-backward stochastic differential equations. We also develop a numerical scheme based on neural networks for solving a time-truncated version of the problem and provide exponential bounds for the truncation error. Numerical experiments suggest that increasing the multiplier of the mean-field term successfully skewed the wealth distribution towards right, increasing the probability of higher relative wealth.

In a second part, we study numerical approximation of backward stochastic differential equations in infinite horizon. We develop three numerical schemes: The first scheme is based on a Picard procedure and uses grid approximation for the space; the second one is also based on a Picard procedure and uses neural networks; the third scheme does not rely on a Picard procedure and uses neural networks like the second one. We also provide a detailed study of the numerical error for the first scheme and prove tight bounds on the approximation error, requiring additional assumptions for contraction. For the second scheme, we proved the convergence of the approximation error to zero as the size of the neural network increases. Numerical experiments also suggest that the third scheme performs better than the first two schemes when the contraction is not fulfilled.

Keywords: stochastic control, mean-field control, backward stochastic differential equations, portfolio optimization

Unité de recherche :

Institut de Mathématiques de Bordeaux, UMR 5251,
351 Cours de la Libération, 33405 Talence cedex, France.