

**MEAN FIELD GAME THEORY: CONSENSUS,
LEADER-FOLLOWER AND MAJOR-MINOR AGENT
SYSTEMS**

Mojtaba Nourian

Department of Electrical and Computer Engineering
McGill University, Montréal

14 September 2012

A thesis submitted to McGill University
in partial fulfilment of the requirements of the degree of
Doctor of Philosophy

© MOJTABA NOURIAN, September 2012

ABSTRACT

This thesis focuses on Mean Field Game (MFG) theory with applications to consensus, flocking, leader-follower and major-minor agent systems. The MFG methodology addresses a class of dynamic games with a large number of minor agents in which each agent interacts with the average or so-called mean field effect of other agents via couplings in their individual dynamics and cost functions. A minor agent is an agent which, asymptotically as the population size goes to infinity, has a negligible influence on the overall system while the overall populations effect on it is significant. The thesis is presented in three main parts.

The first part consists of applications of the MFG methodology to large population consensus and flocking behaviour. In these formulations each agent seeks to minimize its individual quadratic discounted or long time average (i.e., ergodic) cost functions involving the mean of the states of all other agents. The resulting MFG control strategies steer each agent's state toward the initial state population mean, and by applying these decentralized strategies, the system reaches mean-consensus asymptotically as time and population size go to infinity.

The second part is concerned with the extension of the mean field linear-quadratic-Gaussian (MF LQG) framework so as to model the collective system dynamics which include large population of leaders and followers, and an unknown (to the followers) reference trajectory for the leaders. The cost of each leader is based on a trade-off between moving toward the reference trajectory and staying near leaders' own centroid. On the other hand, followers react by tracking a convex combination of their

ABSTRACT

own centroid and the centroid of the leaders. The MF LQG equations characterizing the Nash equilibrium for infinite population systems are derived, and under appropriate conditions, they have a unique solution leading to decentralized control laws. The computation of the followers mean field control laws requires knowledge of the complete reference trajectory which is in general not known to the followers but is estimated by a likelihood ratio based adaptation scheme based on noisy observations taken by the followers on a random sample of leaders.

The final part focuses on large population dynamic games with nonlinear stochastic dynamical systems involving agents of the following mixed types: (i) a major agent, and (ii) a large population of minor agents. The major and minor agents are coupled via both: (i) their individual nonlinear stochastic dynamics of controlled McKean-Vlasov type, and (ii) their individual finite time horizon nonlinear cost functions. A distinct feature of MFG problems with mixed agents is that even asymptotically (as the population size approaches infinity) the noise process of the major agent causes random fluctuation of the mean field behaviour of the minor agents. To deal with this, a stochastic mean field system is formulated in contrast to the deterministic mean field system employed in standard MFG problems.

RÉSUMÉ

Cette thèse se concentre sur la théorie des jeux à population importante (en Anglais, Mean Field Games (MFG)) avec des applications aux systèmes de consensus, flocking, chef-suiveur et aux systèmes d'agents majeure-mineure. La méthodologie MFG aborde une classe de jeux dynamiques avec un grand nombre d'agents mineures dans laquelle chaque agent interagit avec l'effet du champ moyen des autres agents par l'intermédiaire d'accouplements dans leurs dynamiques individuelles et des fonctions de coût. Un agent mineur est un agent qui a une influence négligeable sur l'ensemble du système, mais sur lequel la population globale a un effet significatif. Cette thèse est présentée en trois parties principales.

La première partie développe des applications de la méthodologie MFG au consensus d'une population importante et le comportement de flocking. Dans ces formulations, chaque agent cherche à minimiser ses coûts quadratiques individuels, soit escomptés, soit moyennés en temps (c'est-à-dire ergodique), impliquant la moyenne des états de tous les autres agents. Les stratégies résultant de contrôle MFG orientent l'état de chaque agent vers la moyenne de la population initiale, et en appliquant ces stratégies décentralisés, le système atteint un consensus moyen asymptotiquement en temps et en population.

La deuxième partie s'intéresse à l'extension du cadre des jeux à population importante linéaire-quadratique-Gaussienne (MF LQG) pour modéliser la dynamique du système collective qui comprennent une grande population de chefs et de suiveurs, et une trajectoire de référence pour les chefs qui est inconnue aux suiveurs. Le coût

RÉSUMÉ

de chaque chef est basé sur un compromis entre le déplacement vers la trajectoire de référence et de rester près du centre de gravité propre des chefs. D'autre part, les suiveurs réagissent en faisant le suivi d'une combinaison convexe de leur centre de gravité propre et celui des chefs. Les équations MF LQG qui caractérisent l'équilibre de Nash pour les systèmes de population infinie sont dérivées, et, étant donné des conditions appropriées, ils ont des solutions uniques qui mènent aux lois de contrôle décentralisées. Les calculs des lois de contrôle MFG des suiveurs nécessitent la connaissance complète de la trajectoire de référence qui n'est pas généralement connue aux suiveurs, mais qui est estimée par un rapport de vraisemblance, basé sur des observations bruitées d'un échantillon aléatoire des chefs.

La dernière partie se concentre sur les jeux dynamiques des populations importantes avec des systèmes dynamiques stochastiques non-linéaires impliquant des agents mixtes suivants: (i) un agent majeur, et (ii) une grande population d'agents mineurs. Les agents majeurs et mineurs sont couplés par ces deux: (i) leurs propres dynamiques stochastiques non-linéaires et contrôlées de type McKean-Vlasov, et (ii) leurs fonctions de coûts individuelles non-linéaires à horizon de temps fini. Une caractéristique distincte des problèmes MFG avec des agents mixtes est que, même asymptotiquement (lorsque la taille de la population tend vers l'infini), le processus de bruit de l'agent majeur provoque une fluctuation aléatoire du comportement du champ moyen des agents mineurs. Pour faire face à cela, un système stochastique à champ moyen est introduite comme extension du système déterministe de champ moyen des problèmes de MFG standard.

CLAIMS OF ORIGINALITY

The main contributions of the thesis are as follows:

Chapter 2

- Formulation of consensus models as: (i) linear-quadratic-Gaussian (LQG) dynamic games, (ii) decentralized LQG optimal control, and (iii) centralized linear-quadratic-regulator (LQR) optimal control problems.
- Derivation of the corresponding Nash and social mean field (MF) (or Nash and social certainty equivalence) systems, and proof of existence and uniqueness of their solutions.
- Establishment of some equivalence relationships between the Nash and social MF best response control laws, and the centralized LQR control laws.
- Proof of consensus behaviour for systems with Nash and social MF best response control laws, and centralized LQR control laws.
- Investigating the connectivity role of the localized mean field cost-coupling weight matrix in reaching consensus for heterogeneous sub-populations systems in the LQG dynamic game formulation.

Chapter 3

- Formulation of a mean field game (MFG) consensus problem where the initial states for all the agents are not necessarily distributed according to a Gaussian distribution.

CLAIMS OF ORIGINALITY

- Derivation of the corresponding MF system which consists of two coupled deterministic equations: (i) a backward in time Hamilton-Jacobi-Bellman (HJB), and (ii) a forward in time Fokker-Planck-Kolmogorov (FPK), which are also coupled to a (spatially averaged) cost coupling function.
- Explicit expression of the stationary solution for the MF system, and analysis of the linear stability and the nonlinear stability of the stationary system.
- Formulation of evolution (i.e., forward in time) MF (EMF) equations for systems with Long Time Average (LTA) or ergodic cost functions, and analysis of the linear stability and the nonlinear stability of the stationary system.

Chapter 4

- Formulation of a MFG flocking problem with nonlinear cost-couplings which are inspired by the Cucker-Smale flocking algorithm.
- Derivation of the corresponding MF system which consists of two coupled deterministic equations: (i) a nonlinear backward in time HJB, and (ii) a nonlinear forward in time FPK, which are also coupled to a (spatially averaged) cost coupling function.
- Explicit expression of the stationary solution for the MF system, and analysis of the linear stability and the nonlinear stability of the stationary system.

Chapter 5

- Formulation of a LQG dynamic game based model of collective dynamics which include leaders, followers and an unknown (to the followers) reference trajectory for the leaders.
- Derivation of two coupled MF LQG systems for the leaders and the followers, and proof of existence and uniqueness of their solutions.

- Development of a likelihood ratio based adaptation scheme for the adaptive followers (to identify reference trajectory of the leaders), and the proof of its convergence.
- Proof of the ϵ -Nash equilibrium property of the (estimation based) adaptive MF control laws for the followers and the MF control laws for the leaders.

Chapter 6

- Formulation of a large population dynamic game with nonlinear stochastic dynamical systems involving agents of the following mixed types: (i) a major agent, and (ii) a large population of minor agents.
- Establishment of a mean field convergence theorem for the major and minor MFG problem.
- Derivation of a major-minor stochastic mean field (MM SMF) system by using the theory of backward stochastic differential equations (BSDE). The MM SMF system consists of: (i) two stochastic Hamilton-Jacobi-Bellman (SHJB) equations, and (ii) two stochastic McKean-Vlasov (SMV) equations or stochastic Fokker-Planck-Kolmogorov (SFPK) equations.
- Proof of existence and uniqueness of solution to the MM SMF system by a Banach fixed point argument with random coefficients in the Wasserstein space of stochastic probability measures.
- Retrieval of the MM SMF LQG equations.
- Proof of the ϵ -Nash equilibrium (with respect the full information admissible control sets) of the SMF best response control processes for the overall MM SMF system.

N.B. Almost all of the work above appears in articles which have been published or are currently under review and revision for publication; see pages ix-xii.

Contribution of Co-authors

The work reported in this thesis has been conducted almost entirely by the doctoral candidate. To be more specific, the contribution of each co-author is as follows:

- The doctoral candidate conducted the research reported in Chapter 2 to Chapter 6 and wrote the corresponding manuscripts.
- Professor Peter E. Caines provided advice and comments on the research reported in Chapter 2 to Chapter 6, and helped in editing the corresponding manuscripts. His contributions amounted to 10% in Chapters 2 and 5, 15% in Chapters 3 and 4, and 25% in Chapter 6.
- Professor Roland P. Malhamé provided advice and comments on the research reported in Chapter 2 to Chapter 5 and helped in editing the corresponding manuscripts. His contributions amounted to 10% in Chapters 2 and 5, and 15% in Chapters 3 and 4.
- Professor Minyi Huang provided advice and comments on the research reported in Chapters 2 and Chapter 5, and helped in editing the corresponding manuscripts. His contributions amounted to 10% of those papers.

CURRICULUM VITAE

Mojtaba Nourian was born on March 30, 1983 in Gonabad, Khorasan, Iran. He received dual B.Sc. degrees in applied mathematics and in electrical engineering and the M.Sc. degree in applied mathematics from Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran, in 2005 and 2007, respectively.

List of Publications

Book Chapter

- [135] M. Nourian, P. E. Caines, R. P. Malhamé and M. Huang, “Leader-follower Cucker-Smale type flocking synthesized via mean field stochastic control theory,” In *Brain, Body and Machines*, (Proceedings of CIM 25th Anniversary Symposium, McGill University), J. Angeles et. al. Eds., Chapter 9, pp. 283-298, Springer: Berlin, Germany, 2010.

Journal Papers (Published and to Appear)

- [138] M. Nourian, P. E. Caines, R. P. Malhamé and M. Huang, “Nash, social and centralized solutions to consensus problems via mean field control theory,” To appear as a regular paper in *IEEE Transactions on Automatic Control*, vol. 58, no. 3, Mar. 2013.

- [137] M. Nourian, P. E. Caines, R. P. Malhamé and M. Huang, “Mean field LQG control in leader-follower stochastic multi-agent systems: Likelihood ratio based adaptation,” To appear as a regular paper in *IEEE Transactions on Automatic Control*, vol. 57, no. 11, Nov. 2012.
- [140] M. Nourian, R. P. Malhamé, M. Huang and P. E. Caines, “Mean field (NCE) formulation of estimation based leader-follower collective dynamics,” *International Journal of Robotics and Automation*, vol. 26, no. 1, pp. 120-129, 2011.

Journal Papers (Under Revision)

- [126] M. Nourian and P. E. Caines, “ ϵ -Nash mean field game theory for nonlinear stochastic dynamical systems with major and minor agents,” Under revision in *SIAM Journal on Control and Optimization*, submitted: Aug. 2012 (available in ArXive: <http://arxiv.org/abs/1209.5684>).
- [129] M. Nourian, P. E. Caines and R. P. Malhamé, “A mean field game synthesis of initial mean consensus problems: A continuum approach for non-Gaussian behaviour,” Under revision in *IEEE Transactions on Automatic Control* (submitted Nov. 201, revised Sep. 2012).

Papers in Conference Proceedings (Published and to Appear)

- [128] M. Nourian and P. E. Caines, “ ϵ -Nash mean field theory for nonlinear stochastic games with mixed agents,” To be presented at the *51th IEEE Conference on Decision and Control (CDC)*, Maui, HI, Dec. 2012.
- [131] M. Nourian, P. E. Caines and R. P. Malhamé, “An evolution mean field equation system for initial mean consensus behaviour: A stability analysis,” Proc. of the *50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, Orlando, FL, Dec. 2011, pp. 5029-5034.

- [133] M. Nourian, P. E. Caines and R. P. Malhamé, “A solution to the initial mean consensus problem via a continuum based mean field control approach,” Proc. of the *50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, Orlando, FL, Dec. 2011, pp. 5708-5713.
- [132] M. Nourian, P. E. Caines and R. P. Malhamé, “Mean field analysis of controlled Cucker-Smale type flocking: Linear analysis and perturbation equations,” Proc. of the *18th IFAC World Congress*, Milano, Italy, Aug. 2011, pp. 4471-4476.
- [139] M. Nourian, R. P. Malhamé, M. Huang and P. E. Caines, “Optimality of adaptation based mean field (NCE) control laws in leader-follower stochastic dynamic games,” Proc. of the *49th IEEE Conference on Decision and Control (CDC) Conference*, Atlanta, GA, Dec. 2010, pp. 2270-2275.
- [130] M. Nourian, P. E. Caines and R. P. Malhamé, “Synthesis of Cucker-Smale type flocking via mean field stochastic control theory: Nash equilibria,” Proc. of the *48th Allerton Conference on Communication, Control and Computing*, Monticello, IL, Sep. 2010, pp. 814-819.
- [136] M. Nourian, P. E. Caines, R. P. Malhamé and M. Huang, “A solution to the consensus problem via stochastic mean field control,” In the *2nd IFAC Workshop on Distributed Estimation and Control of Networked Systems (NecSys)*, Annecy, France, Sep. 2010, pp. 323-328.
- [134] M. Nourian, P. E. Caines, R. P. Malhamé and M. Huang, “Derivation of consensus algorithm dynamics from mean field stochastic control NCE equations,” In the *1st IFAC Workshop on Distributed Estimation and Control of Networked Systems (NecSys)*, Venice, Italy, Sep. 2009, pp. 13-18.

Talks at Meetings with Abstract Volumes

- [127] M. Nourian and P. E. Caines, “ ϵ -Nash mean field theory for nonlinear stochastic dynamic games with major-minor agents,” Presented at the *20th International Symposium on Mathematical Theory of Networks and Systems (MTNS)*, Melbourne, Australia, Jul. 2012.
- M. Nourian, P. E. Caines and R. P. Malhamé, “A continuum mean field stochastic control based consensus problem,” Presented at the *SIAM Conference on Control and Its Applications*, Baltimore, MA, Abstract Volume pp. 34, Jul. 2011.
- M. Nourian, P. E. Caines and R. P. Malhamé, “An evolution mean field equation system of the consensus problem,” Presented at the *International Conference of Optimization Days*, Montreal, QC, Abstract Volume pp. 25, May 2011.
- M. Nourian, R. P. Malhamé, M. Huang and P. E. Caines, “Adaptive mean field (NCE) methodology in leader-follower stochastic dynamic games,” Presented at the *4th Biannual Meeting on Systems and Control Theory*, Queen’s University, Kingston, ON, May 2010.
- M. Nourian, P. E. Caines, R. P. Malhamé and M. Huang, “Derivation of consensus algorithm dynamics from mean field stochastic control equations,” Presented at the *International Conference of Optimization Days*, Montreal, QC, Abstract Volume pp. 52, May 2009.

ACKNOWLEDGEMENTS

I am indebted to a great many people who have helped me along my path and without whom this thesis would have never been possible.

First, I would like to express my utmost gratitude to my supervisor Professor Peter E. Caines who was abundantly helpful and offered invaluable encouragement, support and guidance. He is a knowledgeable teacher, a superb scholar and a great advisor who provided a lot of original inspiration and technical guidance for this thesis. In particular, I would like to thank him for directing me towards the fascinating and challenging area of Mean Field Games (MFG) or Nash Certainty Equivalence (NCE). Like all Peter's students I will miss his famous calls during the weekdays and the weekends. I also thank Peter's family for their kindness.

I am greatly indebted to Professor Roland P. Malhamé of GERAD and École Polytechnique de Montréal and Professor Minyi Huang of Carleton University for their encouragement and support. They provided me with many helpful advice and comments which undoubtedly improved the quality of my work. I thank Roland for his easy to understand explanations of technical details; I thank Minyi for always sharing his original works with me.

I consider myself quite fortunate indeed to work with great scholars like Peter, Roland and Minyi who are the pioneers of the MFG (NCE) theory. I must thank them for their efforts arranging our "Friday GERAD Sessions" and "Mean Field Games Days" which are educational, intellectually stimulating and very productive. I also wish to thank the CIM director, Professor Benoit Boulet.

ACKNOWLEDGEMENTS

I am grateful to my thesis examination committee members, Professors Tamer Başar (UIUC), Serdar Yüksel (Queen's), Aditya Mahajan and Michael Rabbat, for the very helpful comments and suggestions.

I thank my lovely friend Sajedeh Mahdavi for sharing with me the fun life of Montreal and for always making me smile. My special thanks also go to her and Peter for translating the abstract of the thesis into French. Many thanks go to my colleagues Farzin Taringoo, Arman C. Kizilkale, Arthur Lazarte, Vahid Raissi Dehkourdi, Rahi Modirnia, Huan Xu, Dmitry Gromov and Peng Jia.

I am grateful for the financial support throughout my Ph.D. studies provided by NSERC discovery research grant and GERAD doctoral student's aid complement. I also thank the organizers and sponsors of the Mean Field Games Workshop (May 2011, Rome, Italy), SADCO Summer School on Optimal Control (September 2011, London, UK) and From Mean Field Control to Weak KAM Dynamics Workshop (May 2012, Coventry, UK) for their support.

On a special note, I am very grateful to Professors Mehdi Dehghan, Mohammad Bagher Menhaj and Ali D. Mohammadi, and to Ms. Tahere Seyedena for their invaluable support during my study at Amirkabir University of Technology.

Special thanks to my old friends Mohamadreza Hosseini, Sajjad Lakzian and Amin Sakzad for their help and encouragement in starting my Ph.D. at McGill. I also thank Mohammad Reza Navabi for introducing me to the interesting area of Systems and Control Theory.

Last but not least, I would like to thank my parents for all their endless love, support and patience. Thanks to them for always encouraging and inspiring me to pursue my education. I also thank my amazing brothers, Morteza and Ali Akbar, and my sweet sisters, Mojgan and Mahnaz, and their adorable children for all the love. I am truly blessed to have such a wonderful family.

Mojtaba Nourian

McGill University, Montréal, Québec

September 2012

TABLE OF CONTENTS

ABSTRACT	i
RÉSUMÉ	iii
CLAIMS OF ORIGINALITY	v
CURRICULUM VITAE	ix
ACKNOWLEDGEMENTS	xiii
LIST OF FIGURES	xxi
CHAPTER 1. Introduction	1
1.1. Structure of The Thesis	5
CHAPTER 2. Nash, Social and Centralized Solutions to Mean Field Consensus Problems	15
2.1. Introduction	16
2.2. Problem Formulations and Terminology	20
2.2.1. Stochastic Decentralized Dynamic Game Consensus	21
2.2.2. Stochastic Decentralized Social Optimal Consensus	22
2.2.3. Deterministic Centralized Social Optimal Consensus	23
2.3. Applications	24
2.4. Preliminary Optimal Control of a Single Agent	25
2.5. Stochastic Mean Field Dynamic Game Consensus Model	26
2.5.1. The Stability and Performance Analysis of Mean Field Control Laws	31

TABLE OF CONTENTS

2.5.2.	ϵ -Nash Equilibrium Property of Mean Field Control Laws	32
2.5.3.	Extension to the Case of Long Time Average (LTA) Cost Functions	34
2.6.	Stochastic Mean Field Social Optimum Consensus Model	35
2.6.1.	Extension to the Case of Long Time Average (LTA) Cost Functions	39
2.7.	Social Optimal LQR Consensus Model	40
2.7.1.	Deterministic Mean Field Algorithm with Feedback	42
2.8.	Game Consensus Model with Random Cost-Couplings	44
2.9.	Game Consensus Model with Localized Cost-Couplings	47
2.10.	Numerical Examples	52
2.11.	Chapter Summary	57
2.12.	Appendix	58
CHAPTER 3. A Mean Field Game Synthesis of Initial Mean Consensus		
	Problems: A Continuum Approach for Non-Gaussian Behaviour	61
3.1.	Introduction	62
3.2.	Dynamic Game Consensus Model with Non-Gaussian Initial States . .	64
3.3.	A Continuum Mean Field Game Approach	65
3.3.1.	Mean Field Approximation	65
3.3.2.	The Mean Field System	67
3.4.	Analysis of the Mean Field System	69
3.4.1.	Gaussian Stationary Solution	69
3.4.2.	Stability Analysis of the Linearized Mean Field System	70
3.4.3.	Stability Analysis of the Nonlinear Mean Field System	73
3.5.	Mean-Consensus	74
3.6.	ϵ -Nash Equilibrium Property of MF Control Laws	75
3.7.	Dynamic Game Consensus Model with LTA (Ergodic) Costs	76
3.7.1.	The Evolution Mean Field System	79
3.7.2.	Stability Analysis of the Evolution Mean Field System	80
3.8.	Numerical Example	83
3.9.	Chapter Summary	83

3.10. Appendix	84
CHAPTER 4. Synthesis of Mean Field Controlled Cucker-Smale Type Flocking:	
A Maxwellian Distribution	89
4.1. Introduction	90
4.2. The Uncontrolled Cucker-Smale Model	92
4.3. The Controlled Flocking Model	93
4.4. A Mean Field Game Approach	95
4.4.1. Mean Field Approximation	96
4.4.2. Preliminary Optimal Control of a Single Agent	96
4.4.3. The Mean Field System	98
4.5. Maxwellian Stationary Solution	101
4.6. ϵ -Nash Equilibrium Property	102
4.7. Chapter Summary	104
CHAPTER 5. Mean Field LQG Control in Leader-Follower Stochastic Multi-	
Agent Systems: Likelihood Ratio Based Adaptation	105
5.1. Introduction	106
5.2. Problem Formulation, Terminology and some Applications	108
5.2.1. Leaders Stochastic LQG Dynamic Game Model	109
5.2.2. Followers Stochastic LQG Dynamic Game Model	110
5.2.3. Followers Observation Processes	112
5.2.4. Applications	112
5.3. Mean Field LQG Stochastic Control Theory	114
5.3.1. Preliminary LQG Optimal Control of a Single Agent	114
5.3.2. The MF LQG Systems	115
5.3.3. Analysis of MF Systems for Leaders and Followers	119
5.4. Estimation Procedure for the Adaptive Followers	122
5.4.1. The Likelihood Function	122
5.4.2. The Likelihood Ratio	124

TABLE OF CONTENTS

5.4.3. Main Estimation Theorem 125

5.5. The Stability Analysis of the Mean Field Control Laws 127

5.5.1. The Follower’s Adaptive Mean Field Algorithm 128

5.5.2. Stability Analysis 129

5.6. ϵ -Nash Equilibrium Property of the Mean Field Based Control Laws . 129

5.7. Numerical Example 135

5.8. Chapter Summary 138

5.9. Appendix 140

CHAPTER 6. Mean Field Game Theory for Nonlinear Stochastic Dynamical

Systems with Major and Minor Agents 147

6.1. Introduction 148

6.1.1. Notation and Terminology 151

6.2. Problem Formulation 152

6.2.1. Assumptions 154

6.3. Major and Minor Agents Mean Field Convergence Theorem 155

6.4. A Preliminary Nonlinear Stochastic Optimal Control Problem 161

6.5. The Major and Minor Agents’ Stochastic Mean Field System 165

6.5.1. Stochastic Optimal Control Problem of the Major Agent 166

6.5.2. Stochastic Optimal Control Problem of the Generic Minor Agent . 168

6.5.3. The Mean Field Game Consistency Condition 172

6.6. Existence and Uniqueness of Solutions to the MM SMF System 173

6.6.1. Analysis of the Major Agent’s SMF System 175

6.6.2. Analysis of the Minor Agents’ SMF System 180

6.6.3. Analysis of the Joint Major and Minor Agents’ SMF System . . . 186

6.7. Applications 197

6.7.1. The MM SMF LQG System 197

6.7.2. Synchronization of Coupled Nonlinear Oscillators Game 201

6.8. ϵ -Nash Equilibrium Property of the SMF Control Laws 203

6.9. Chapter Summary 210

TABLE OF CONTENTS

CHAPTER 7. Conclusion and Future Research 211

REFERENCES 213

LIST OF FIGURES

2.1 Example 2.1: (A) Trajectories of agents' states when $N = 500$, (B) Histogram of the system at time $t = 20$	52
2.2 Example 2.1: (A) The convergence of the population mean; (B) Curve of ϵ_1^N with respect to N	53
2.3 Example 2.2: Trajectories of agents' states in (A) deterministic MF social, and (B) centralized LQR cases.	54
2.4 Example 2.3: (A) Trajectories of agents' states, (B) Histogram of the system at time $t = 20$	55
2.5 Example 2.4: (A) Trajectories of agents' states, (B) Histogram of the system at time $t = 20$	56
2.6 Example 2.5: (A) Trajectories of agents' states, (B) Histogram of the system at time $t = 20$	57
3.1 (A) The contour lines of population density functions, and (B) trajectories of agents' states when $N = 500$	83
5.1 Four possible reference trajectories of the leaders (reference one is the true reference trajectory) corresponding to parameters $\delta_1, \delta_2, \delta_3, \delta_4$ in the set Δ , respectively, and the initial states of leaders and followers.	136

LIST OF FIGURES

5.2 $(1/t) \log$ of likelihood ratios $((1/t) \log(H_t^m(\delta_i)/H_t^m(\delta_j)), \delta_i, \delta_j \in \Delta)$ such that in each figure there is a fixed parameter of set Δ in the numerator of the ratio and the parameter in the denominator changes in the set Δ 137

5.3 State trajectories of leaders (red), non-adaptive followers (blue), and the adaptive follower (black). 138

CHAPTER 1

Introduction

This thesis investigates control and optimization of large-scale stochastic dynamical systems involving multiple agents. Multi-agent control and coordination problems arise in a wide range of areas including telecommunications [9], wireless sensor networks [41], power control in wireless communication systems [7], renewable energy systems [94], vehicle formations [58, 147, 156], competing or cooperating mobile robots [35, 108], flocking and swarming [46, 110, 163], micro-economics and finance [100, 112]. The complexity in the form of uncertainty and complicated interaction and communication among agents of such large population stochastic dynamical systems make centralized control infeasible. Therefore an important issue is the development of decentralized solutions so that each individual agent may implement a strategy based on its local information together with statistical information on the population of agents.

An important class of large-scale stochastic dynamical systems is that of dynamic games with a large number of minor agents in which each agent interacts with the average or so-called mean field effect of other agents via couplings in their individual dynamics and cost functions. A minor agent has a negligible influence on the overall system while the overall population of minor agents has a significant effect on any such agent.

It is worth pointing out that the very early interest in games with a large number of agents is in the book of Von Neumann and Morgenstern [170]: “When the number of participants becomes really great, some hope emerges that the influence of every particular participant will become negligible, and that the above difficulties may recede and a more conventional theory becomes possible.” Among the many papers on such large population game theoretic models we first mention [13] which presents a general equilibrium featuring a continuum of agents. Since then there has accumulated a vast literature on such game models (see [38, 89] and the references therein).

Large population models with game theory features arise in fields such as wireless network resource allocation [7, 8, 74], renewable energy [118], biology [150, 167], advertising competition [57], public health [16, 24], economics [100] and sociology [26]. The reader is referred to [15] for the theory and applications of noncooperative dynamic (differential) games with finite number of agents.

Since 2003, for large population stochastic dynamic games with mean field (MF) couplings, the Mean Field Games (MFG) (or Nash certainty equivalence (NCE)) theory has been developed as a decentralized methodology in a series of papers by Huang together with Caines and Malhamé, see [76, 79] for the mean field linear-quadratic-Gaussian (MF LQG) framework, and [31, 78, 85] for a general formulation of nonlinear McKean-Vlasov type MF stochastic control problems.

For dynamic games with mean field interactions in cost functions a closely related approach has been independently developed by Lasry and Lions [101–103] (see also [36, 66]) where the term mean field games was initially used. For models of many firm industry dynamics, Weintraub et. al. proposed the notion of oblivious equilibrium (OE) to approximate a Markov perfect equilibrium (MPE) of a dynamic game with a large number of agents [174, 175]. In an OE, each individual agent is oblivious to the state of the overall system and makes its decision based only on its own state variable together with a consistently defined mean field. The OE for large population stochastic games with unbounded cost functions are analyzed in [5].

The central idea of the MFG methodology is to establish the existence of an equilibrium relationship between the individual strategies and the mass effect (i.e., the overall effect of the population on a given agent) as the population size goes to infinity [79]. Specifically, in the equilibrium: (i) the individual strategy of each agent is a best response to the infinite population mass effect in the sense of a so-called ϵ -Nash equilibrium, and (ii) the set of strategies collectively replicates the mass effect, this being a dynamical game theoretic fixed point property. The defining property of the MFG equilibrium with individual strategies $\{u_i^o : 1 \leq i \leq N\}$ requires that for any given $\epsilon > 0$, there exists $N(\epsilon)$ such that for any population size $N(\epsilon) \leq N$, when any agent j , $1 \leq j \leq N$, distinct from i employs u_j^o , then agent i can benefit at most ϵ by unilaterally deviating from his strategy u_i^o , and this holds for all $1 \leq i \leq N$.

The MFG feedback strategies display the possibly counterintuitive nature of MFG control which is that in the infinite population limit, except for some statistical information on the parameter distribution and the initial mean state distribution of the population of agents, no observations of other agents' states are necessary to achieve Nash equilibrium behaviour and this property persists with negligible incremental cost for sufficiently large finite populations.

In [106], the MF LQG framework is extended to systems of agents with Long Time Average (LTA) (i.e., ergodic) cost functions such that the set of control laws possesses an almost sure (a.s.) asymptotic Nash equilibrium property, while in [82], the interaction consistency based approach was applied to models where the agents are cooperative and seek socially optimal decisions, and asymptotic decentralized social optimum strategies are obtained. The stochastic adaptive control of MF LQG models with LTA cost functions is studied in [90, 92] where the agents estimate their own dynamical parameters, and the population's dynamical and cost function distribution parameters. The MF LQG framework is extended to the case of agents with localized interactions in their cost functions [80, 83] (see [104] for some numerical simulations). Kolokoltsov et. al. [97] extended the MFG theory to general nonlinear Markov systems as formulated in the monograph [95].

In contrast to [76, 79] with only minor agents, Huang [75] extended the idea of mean field to LQG dynamic games with agents of the following mixed types: (i) a major agent, and (ii) a large population of minor agents. The major agent has a significant influence on minor agents while each minor agent has a negligible impact on other agents in a large population system (see [75, 124]).

Other research in this area include mean field Markov decision team problem with discrete state and action spaces [165], discrete time and finite state space MFG [61], numerical solutions of MFG systems [3, 4], linear-quadratic (LQ) MFG [14], risk-sensitive MFG [166], long time average analysis of MFG [37], MF LQG control with egoistic and altruistic agents [81, 91], MF LQG models with random parameters [171, 172], and two player zero-sum games with binary MF interactions [96]. The reader is referred to the survey paper [28] for some works on MFG theory up to 2011.

In a different framework, a stochastic maximum principle for control problems of mean field type is studied in [10] where the state process is governed by a stochastic differential equation (SDE) in which the coefficients depend on the law of the SDE. In the model of [10], the control action of each agent has significant impact on the mean field, in contrast, an individual agent in the system has little impact on the mean field. The reader is referred to [29, 30] for the analysis of forward–backward stochastic differential equations (FBSDEs) of mean field type and their related partial differential equations (PDEs).

The MFG methodology has been applied to wireless power control [76, 164], coupled nonlinear oscillators subject to random disturbances [177], particle filtering [176], crowd dynamics [49], large population electric vehicles [116] and some models in economics [6, 65, 175].

This thesis is presented in three main parts. The first part consists of applications of the MFG methodology to large population consensus and flocking behaviour. The second part is focused on the extension of the MF LQG framework so as to model the collective system dynamics which include large population of leaders and followers, and an unknown (to the followers) reference trajectory for the leaders. The final

part extends the major and minor MF LQG theory to MFG models with nonlinear stochastic dynamical systems of controlled McKean-Vlasov type.

1.1. Structure of The Thesis

The thesis is organized as follows:

Chapter 2. Nash, Social and Centralized Solutions to Mean Field Consensus Problems. A consensus process is a process for dynamically achieving (by continuous state feedback) an agreement between the members of a group of agents on some common state property such as position, velocity or information. The formulation of consensus systems is one of the important issues in the area of multi-agent control and coordination.

Among many papers on the consensus problems in the systems and control area we first mention here [88,120,142,152,168] and comprehensive surveys [60,141,154] of works up to 2011. Consensus algorithms with noisy measurement or random network connectivity have been addressed in [84,86,87,107,148] among others. The key element of all of these consensus algorithms, which we shall refer to as standard consensus (SC) algorithms, is the use of local feedback by local communication (subject to the network topology) between agents to reach an agreement.

In the standard consensus literature the overall population's initial state contributes to the steady-state (equilibrium) behaviour of the system. This is mainly due to situations in many practical applications where the goal is reaching agreement on some value based on the system's initial state (see for example [60,154] and the references therein).

However, the connectivity of the network structure needed for the above SC models (even for the less demanding "frequently connected" hypotheses) may not hold. Moreover, the SC algorithms require communication with other agents in the system and for large N this leads to high communication and computational complexity.

In this chapter the initial mean consensus behaviour of a set of agents is synthesized from the fundamental optimization principles of (i) stochastic dynamic games, and (ii) optimal control rather than to analyze the behaviour resulting from ad-hoc feedback laws.

In the stochastic dynamic game formulation each agent seeks to minimize its individual quadratic discounted or LTA cost function involving the mean of the states of all other agents [134, 136, 138]. The limiting infinite population MF system is derived and its unique solution is explicitly computed [136, 138]. The resulting MF control strategies steer each agent's state toward the initial state population mean, and by applying these decentralized strategies, the system reaches mean-consensus asymptotically as time and population size go to infinity. Furthermore, these control laws possess an ϵ_N -Nash equilibrium property where ϵ_N goes to zero as the population size N goes to infinity [136, 138]. The analysis is extended to the case of random mean field couplings [138].

In the social cooperative formulation the basic objective is to minimize a social cost as the sum of the individual cost functions containing mean field couplings [82]. In this formulation it is shown that for any individual agent the decentralized mean field social (MF Social) control strategy is the same as the mean field Nash (MF Nash) equilibrium strategy [138]. Hence,

$$\text{MF-Nash Controls } U_{\text{Nash}}^\infty = \text{MF-Social Controls } U_{\text{Soc}}^\infty.$$

On the other hand, the solution to the centralized linear-quadratic-regulator (LQR) optimal control formulation yields the Standard Consensus (SC) algorithm whenever the graph representing the corresponding topology of the network is Completely Connected (CC) [138]. Hence,

$$\text{Cen. LQR Controls } U_{\text{Cen}}^N = \text{SC-CC Controls } U_{\text{SC}}^N.$$

Moreover, a system with centralized control laws reaches consensus on the initial state distribution mean as time and population size N go to infinity [138]. Hence,

asymptotically in time,

$$\begin{aligned} \text{MF-Nash Controls } U_{\text{Nash}}^\infty &= \text{MF-Social Controls } U_{\text{Soc}}^\infty \\ &= \text{Cen. LQR Controls } U_{\text{Cen}}^\infty = \text{SC-CC Controls } U_{\text{SC}}^\infty. \end{aligned}$$

It is important to note that in the MF consensus models, similar to the SC algorithms, the overall population's initial state contributes to the steady-state (equilibrium) behaviour of the system such that the (time) expectation of the system's steady-state solution is the overall population's initial state distribution mean. This is in contrast to situations for many MF solutions, such as the ones in [79, 82], where any initial data information is destroyed as the processes evolve since the overall population's initial state does not affect the steady-state behaviour of the system. This is because of the nature of the cost-couplings between individual agents and the external mass of agents.

In the MF consensus models: (i) each agent has a priori information on the initial state distribution mean of the overall population, (ii) the system of agents achieves mean-consensus without requiring communication with other agents. Whereas in the SC algorithms: (i) agents need no a priori information on the initial state distribution of the overall population but require local communication with other agents, (ii) consensus can be achieved if the union of the interaction graphs for the system is connected frequently enough as the system evolves.

The SC algorithms require communication with other agents in the system and for large N this leads to high communication and computational complexity. On the other hand, the decentralized MF control laws do not require even local communication but need a priori information on the mean of the system's initial state distribution. In the context of centralized LQ models with finite populations, a trade-off between the use of a priori statistical information on a system's initial distribution and communication among its agents has been formulated and analyzed in [157].

The extension of the uniform weight cost-coupling MF consensus model to the case of agents with non-uniform mean field cost-couplings which corresponds to a

heterogeneous system with homogeneous sub-populations. is also investigated (see [80, 83] for the localized MF LQG models). In the localized model with connected topology specified by the cost-coupling weight matrix, the unique stationary equilibrium yields consensus in the weighted average of initial states (which depends on the left eigenvector of the weight matrix corresponding to the unique eigenvalue 1). Let the system cost-coupling weight matrix correspond to an adjacency matrix of a graph with more than one connected component, then each associated sub-population can only converge to the initial distribution mean of its connected component. Correspondingly, in the deterministic SC problem, one of the key hypotheses which is used to establish the convergence to consensus is that the system graph is connected.

Chapter 3. A Continuum Approach to Mean Field Game Consensus Problems: A Non-Gaussian Behaviour. This work presents a continuum approach to a non-Gaussian initial mean consensus problem via MF nonlinear stochastic control theory developed in [78, 85, 177]. In this problem formulation: (i) each agent has simple stochastic dynamics with inputs directly controlling its state's rate of change, and (ii) each agent seeks to minimize by continuous state feedback its individual discounted or LTA cost functions involving the mean of the states of all other agents.

Unlike [136, 138], the initial states for all the agents of the model are not necessarily assumed to be distributed according to a Gaussian distribution, and so the MF LQG framework of [136, 138] cannot be employed. Consequently, for the infinite population limit a general continuum (i.e., PDE) formulation is required.

The resulting continuum based MF system of the consensus model consists of two coupled deterministic equations: (i) a nonlinear (backward in time) Hamilton-Jacobi-Bellman (HJB), and (ii) a nonlinear (forward in time) Fokker-Planck-Kolmogorov (FPK), which are also coupled to a (spatially averaged) cost coupling function approximating the aggregate effect of the agents in the infinite population limit.

The stationary solutions of the MF system is explicitly given and its linear stability (base on the approach of [64]) and nonlinear stability are analyzed [129]. In a finite population system (analogous to the MF LQG framework): (i) the resulting decentralized MF control strategies possess an ϵ_N -Nash equilibrium property where ϵ_N goes to zero as the population size N approaches infinity, and (ii) these MF control strategies steer each individual's state toward the initial state population mean which is reached asymptotically as time goes to infinity. Hence, the system with decentralized MF control strategies reaches mean-consensus on the initial state population mean asymptotically as time and population size go to infinity [129, 133].

In the case of agents with LTA cost functions the solution of the HJB equation is the relative value function which represents perturbations around the steady-state optimal cost rate with respect to an asymptotically stationary process. It turns out that this HJB equation in the MF system of equations has a larger class of stable perturbed solutions in forward time than in backward time [64]. Therefore, an Evolution (i.e., forward in time) Mean Field (EMF) system of consensus model is studied where the initial states for all the agents are not necessarily assumed to be distributed according to a Gaussian distribution [131].

The EMF system consists of two coupled (forward in time) deterministic PDEs which are also coupled to the cost coupling function. The forward in time mean field process has previously appeared in the study of MFG models in [4, 64].

Chapter 4. Synthesis of Mean Field Controlled Cucker-Smale Type Flocking: A Maxwellian Distribution. Collective motion such as the flocking of birds, schooling of fish and swarming of bacteria is one of the most widespread phenomenon in nature. The study of collective motion in nature is of interest not only to model and analyze these widespread phenomena but also because ideas from these behaviours can be used by engineers to develop efficient algorithms for a wide range of applications.

A group of agents has a flocking behaviour if: (i) agents' velocities reach consensus on a common value (e.g., mean of initial velocities), and (ii) the distance between agents remains bounded.

There are two main classes of models for the flocking behaviour: (i) individual based models in the form of coupled Ordinary (Stochastic) Differential Equations (O(S)DEs) (see for instance [46, 169]) where in these algorithms a key element is the use of local feedback involving local communication (subject to the network topology) between agents so as to reach an agreement, and (ii) continuum models in the form of Partial (or integro-partial) Differential Equations (PDEs) to model the collective motion in the case of systems with large (infinite) populations (see [40, 68, 167] among many other papers). The continuum models can be derived from the individual based models in the large population limit by use of the kinetic theory of gases, hydrodynamic and mean field theory (see for instance [40]).

Cucker and Smale formulated an interesting individual based flocking model for a group of agents [46]. This model is motivated by the collective motion of a group of birds such that each bird updates its velocity as a weighted velocities of all the other birds. The weights in this model are functions of the relative distance of the birds such that as the mutual distance between two birds increases the influence of their velocities on each other decreases.

This work is concerned with the synthesis of a controlled flocking model via MFG theory [130, 132]. In this problem formulation the state of each agent consists of both its position and its controlled velocity such that: (i) all agents have similar stochastic dynamics, and (ii) each agent seeks to minimize by continuous state feedback its individual discounted cost functions involving a nonlinear (relative distance based) weighted mean of the velocity states of all other agents. The cost functions are based on the normalized Cucker-Smale (CS) flocking algorithm in its original uncontrolled formulation.

For this dynamic game problem, the MF system of equations which consists of coupled deterministic HJB and FPK equations is derived approximating the stochastic system of agents as the population size goes to infinity. Subject to the existence of a unique solution to the MF system of equations: (i) the stationary solution of the MF system of equations is a Maxwellian distribution function, (ii) the set of MF control laws for the system possesses an ϵ_N -Nash equilibrium property where ϵ_N goes to zero as the population size N approaches infinity. Hence, this model may be regarded as a controlled game theoretic formulation of a flocking behaviour in which each agent, instead of responding to an ad-hoc algorithm, obtains its control law from a game theoretic Nash equilibrium [130, 132].

Chapter 5. Mean Field LQG Control in Leader-Follower Stochastic Multi-Agent Systems: Likelihood Ratio Based Adaptation. In this work a LQG dynamic game based model of collective dynamics is developed which include leaders, followers and a reference trajectory for the leaders [137]. There are many applications of this model in flocking [63], formation control [109], economics and finance [2], and social opinion models with a large number of leaders (e.g., important members of a party) and followers [53] (see [137]).

The cost of each leader is based on a trade-off between moving toward a certain reference trajectory which is unknown to the followers and staying near their own centroid. On the other hand, followers react by tracking a convex combination of their own centroid and the centroid of the leaders. The MF system characterizing the Nash equilibrium for infinite population systems are derived, and under appropriate conditions, they have a unique solution leading to decentralized control laws. Furthermore, for large but finite population systems, such controls are shown to correspond to so-called ϵ -Nash equilibria.

The computation of the followers' control laws requires knowledge of the complete reference trajectory which is in general not known to the followers but is estimated by a likelihood ratio based adaptation scheme based on noisy observations taken by the

followers on a random sample of leaders. Under appropriate identifiability conditions, it is established that this identification scheme is able to select the exact reference trajectory model within a finite class of candidates in a finite deterministic time almost surely as the number of samples goes to infinity. As a result, the (estimation based) adaptive MF control laws of the followers together with the MF control laws of the leaders give rise to a dynamic stochastic Nash equilibrium for the overall leader-follower system [137].

It is worth pointing out that a non-adaptive but general model with weighted couplings in the leaders and followers' cost functions (which depended on the locality parameters of the agents) is developed in [140] which also presents the main adaptation result of the uniform cost coupling model in the case that the followers "only" track the centroid of the leaders. Subsequently, in [139] the optimality property of the (tracking like) adaptive followers' MF control laws is studied. A complete analysis of a more general (and realistic) scenario where the followers are tracking a convex combination of their own centroid and the centroid of the leaders is presented in [137] where an ϵ -Nash equilibrium is achieved for the adaptive followers' MF control laws.

The leader-follower model of [137] is extended to the case of non-adaptive agents with nonlinear Cucker-Smale type cost coupling functions in [135].

Chapter 6. Mean Field Game Theory for Nonlinear Stochastic Dynamical Systems with Major and Minor Agents. This work is concerned with a large population stochastic dynamic game involving nonlinear stochastic dynamical systems involving agents of the following mixed types: (i) a major agent, and (ii) a large population of minor agents [126, 128]. The major and minor agents are coupled via both: (i) their individual nonlinear stochastic dynamics, and (ii) their individual finite time horizon nonlinear cost functions. This model extends the MF LQG model for major and minor agents [75, 124] to the case of nonlinear stochastic dynamic games formulation of controlled McKean-Vlasov type [85].

Applications of the major and minor formulation may be found in charging control of plug-in electric vehicles [117, 178], social opinion models [53] with a finite number of leaders, and power markets involving large consumers and large utilities together with many domestic consumers represented by smart meter agents and possibly large numbers of renewable energy based generators [93].

A distinctive feature of the mixed agent MFG problem is that even asymptotically (as the population size N approaches infinity) the noise process of the major agent causes random fluctuation of the mean field behaviour of the minor agents [75, 124]. To deal with this, the overall asymptotic ($N \rightarrow \infty$) mean field game problem is decomposed into: (i) two non-standard Stochastic Optimal Control Problems (SOCs) with random coefficient processes which yield forward adapted stochastic best response control processes determined from the solution of (backward in time) stochastic Hamilton-Jacobi-Bellman (SHJB) equations, and (ii) two stochastic (coefficient) McKean-Vlasov (SMV) equations which characterize the state of the major agent and the measure determining the mean field behaviour of the minor agents. (i) and (ii) are coupled in the following way: the forward adapted stochastic best response control processes in (i) involve the state of the major agent and the distribution measure corresponding to the mean field behaviour of the minor agents in (ii) where these in turn depend upon the best response control processes themselves.

Existence and uniqueness of the solution to the Stochastic Mean Field (SMF) system (SHJB and SMV equations) is established by a fixed point argument in the Wasserstein space of random probability measures [126]. In the case that minor agents are coupled to the major agent only through their cost functions, the ϵ_N -Nash equilibrium property of the SMF best response control possess is shown for a finite N population system where $\epsilon_N = O(1/\sqrt{N})$ [126].

As a particular but important case, the results of Nguyen and Huang [124] for MM SMF LQG systems with homogeneous population are retrieved. In addition, the results of this chapter are illustrated with a major and minor agent version of a game model of the synchronization of coupled nonlinear oscillators [126].

Chapter 7. Conclusion and Future Research. This chapter contains some concluding remarks and some future research directions.

CHAPTER 2

Nash, Social and Centralized Solutions to Mean Field Consensus Problems

The purpose of this chapter is to synthesize initial mean consensus behaviour of a set of agents from the fundamental optimization principles of (i) stochastic dynamic games, and (ii) optimal control. In the stochastic dynamic game model each agent seeks to minimize its individual quadratic discounted cost function involving the mean of the states of all other agents. In this formulation we derive the limiting infinite population mean field system and explicitly compute its unique solution. The resulting Mean Field (MF) control strategies steer each individual's state toward the initial state population mean which is reached asymptotically as time goes to infinity, thus achieving mean consensus. Furthermore, these control laws possess an ϵ_N -Nash equilibrium property where ϵ_N goes to zero as the population size N goes to infinity. Furthermore, the analysis is extended to the cases of: (i) random mean field cost-couplings, and (ii) agents with non-uniform mean field cost-couplings which corresponds to a heterogeneous system with homogeneous sub-populations.

In the social cooperative formulation the basic objective is to minimize a social cost as the sum of the individual cost functions containing mean field coupling. In this formulation we show that for any individual agent the decentralized mean field social (MF Social) control strategy is the same as the mean field Nash (MF Nash)

equilibrium strategy. Hence,

$$\text{MF-Nash Controls } U_{\text{Nash}}^\infty = \text{MF-Social Controls } U_{\text{Soc}}^\infty.$$

On the other hand, the solution to the centralized LQR optimal control formulation yields the Standard Consensus (SC) algorithm whenever the graph representing the corresponding topology of the network is Completely Connected (CC). Hence,

$$\text{Cent. LQR Controls } U_{\text{Cent}}^N = \text{SC-CC Controls } U_{\text{SC}}^N.$$

Moreover, a system with centralized control laws reaches consensus on the initial state distribution mean as time and population size N go to infinity. Hence, asymptotically in time,

$$\begin{aligned} \text{MF-Nash Controls } U_{\text{Nash}}^N &= \text{MF-Social Controls } U_{\text{Soc}}^N \\ &= \text{Cent. LQR Controls } U_{\text{Cent}}^\infty = \text{SC-CC Controls } U_{\text{SC}}^\infty. \end{aligned}$$

Finally, the analysis is extended to Long Time Average (LTA) (i.e., ergodic) cost functions case.

2.1. Introduction

A *consensus process* is the process of dynamically reaching an agreement between the agents of a group on some common state properties such as position or velocity. The formulation of consensus systems is one of the important issues in the area of multi-agent control and coordination, and has been an active area of research in the systems and control community over the past decade.

Among the many papers on the consensus problems in the systems and control area we first mention here [88, 142, 152, 168] and comprehensive surveys [60, 141, 154] of works up to 2011. Consensus algorithms with noisy measurement or random network connectivity have been addressed in [84, 86, 87, 107, 148] among others. For consensus algorithms with noisy measurements, Huang et. al. took a stochastic approximation approach with a decreasing step size in [86, 87]. The key element

of all of these consensus algorithms, which we shall refer to as *standard consensus* (SC) algorithms, is the use of local feedback by local communication (subject to the network topology) between agents to reach an agreement.

In the standard consensus literature the “overall population’s initial state” contributes to the steady-state (equilibrium) behaviour of the system. This is mainly due to situations in many practical applications where the goal is reaching agreement on some value based on the system’s initial state (see for example [60, 141, 154] and the references therein).

Some optimality issues in consensus problems have been addressed in the literature. The authors in [34] studied the optimal Laplacian matrix by using a linear-quadratic-regulator (LQR) optimization approach with respect to the weights of the network topology in continuous and discrete time while in [158] the authors designed a semi-decentralized optimal control strategy for the standard consensus algorithms by minimizing the individual cost of each agent. On the other hand, a game theoretic interpretation of locally optimal nonlinear consensus algorithms as mechanism design problems is proposed in [17] by imposing individual objective functions.

However, the connectivity of the network structure needed for the above SC models (even for the less demanding “frequently connected” hypotheses) may not hold. Moreover, the SC algorithms require communication with other agents in the system and for large N this leads to high communication and computational complexity.

In this chapter we develop an optimization approach to the study of “initial mean” consensus problems. Our aim is to synthesize from the theory of optimal control the consensus behaviour of a set of agents rather than to analyze the behaviour resulting from ad-hoc feedback laws. This chapter includes the following three approaches to the synthesis of initial mean consensus behaviour: (i) dynamic games, (ii) decentralized optimal control, and (iii) centralized LQR optimal control theory. In all of these problem formulations each agent in the system has simple stochastic or deterministic dynamics with inputs directly controlling the rate of change of the agents’ states. In

the stochastic dynamic game formulation, each agent seeks to minimize its individual quadratic discounted or Long Time Average (LTA) cost functions involving the mean of the states of all other agents, whereas in the social cooperative formulation the basic objective is to minimize a social cost as the sum of these individual cost functions.

The main contributions of this chapter are as follows:

- (i) In the stochastic dynamic game formulation we derive the limiting infinite population Mean Field (MF) system and explicitly compute its unique solution. The resulting MF control strategies steer each individual's state toward the initial state population mean which is reached asymptotically as time goes to infinity, thus achieving mean-consensus. Furthermore, these control laws possess an ϵ_N -Nash equilibrium property where $\epsilon_N \rightarrow 0$ as the population size N goes to infinity. Furthermore, the analysis is extended to the cases of: (i) random mean field cost-couplings, and (ii) agents with non-uniform mean field cost-couplings which corresponds to a heterogeneous system with homogeneous sub-populations.
- (ii) In the social stochastic formulation we show that in the infinite population case the resulting MF system is the same as the MF game system. Hence, the resulting MF social control strategy is the same as the MF Nash strategy, and so

$$\text{MF-Nash Controls } U_{\text{Nash}}^\infty = \text{MF-Social Controls } U_{\text{Soc}}^\infty.$$

- (iii) We show that the solution to the centralized LQR optimal control formulation yields the Standard Consensus (SC) algorithm whenever the graph representing the corresponding topology of the network is Completely Connected (CC). Hence,

$$\text{Cent. LQR Controls } U_{\text{Cent}}^N = \text{SC-CC Controls } U_{\text{SC}}^N.$$

- (iv) In the MF set-up each agent has a priori information on the initial state distribution mean of the overall population; relaxing this a priori information in the

deterministic case gives rise to the centralized feedback of other agent states. We show that a system with this observation feedback algorithm reaches consensus on the initial state distribution mean as time and population size N go to infinity. Hence, asymptotically in time,

$$\begin{aligned} \text{MF-Nash Controls } U_{\text{Nash}}^\infty &= \text{MF-Social Controls } U_{\text{Soc}}^\infty \\ &= \text{Cent. LQR Controls } U_{\text{Cent}}^\infty = \text{SC-CC Controls } U_{\text{SC}}^\infty. \end{aligned}$$

- (v) Finally, the analysis is extended to Long Time Average (LTA) (i.e., ergodic) cost functions case.

In SC algorithms the topology connectivity of the system dynamics is important whereas in MF consensus models the “a priori” information (on the initial state distribution of the overall population) plays a critical role. More precisely, in the MF consensus model considered in this chapter: (i) each agent has a priori information on the initial state distribution mean of the overall population, and (ii) the system of agents achieves mean-consensus without requiring communication with other agents. Whereas in SC Algorithms: (i) agents need no a priori information on the initial state distribution of the overall population but require local communication with other agents, (ii) consensus can be achieved if the union of the interaction graphs for the system is connected frequently enough as the system evolves (see for example [149, 153]). Moreover, the MF consensus approach allows one to compute the transient cost of moving towards consensus.

In the context of centralized LQ models with finite populations, a trade-off between the use of a priori statistical information on a system’s initial distribution and communication among its agents has been formulated and analyzed in [157].

In the case of dynamic game consensus formulation of localized mean field cost-coupling weight matrix with connected topology, the unique stationary equilibrium yields consensus in the weighted average of initial conditions which depends on the left eigenvector of the weight matrix corresponding to the unique eigenvalue one. Let the system cost-coupling weight matrix correspond to an adjacency matrix of a graph

with more than one connected component, then each associated sub-population can only converge to the initial distribution mean of its connected component.

In the MF consensus model of this chapter the (time) expectation of the system's steady-state solution is the overall population's initial state distribution mean (which we define as *initial mean consensus*). This is in contrast to situations for many MF solutions, such as the ones in [79, 82], where any initial data information is destroyed as the processes evolve since the overall population's initial state does not affect the steady-state behaviour of the system. This is because of the nature of the cost-couplings between individual agents and the external mass of agents.

The organization of the chapter is as follows. Section 2.2 is dedicated to the problem formulation and terminology. Some applications of the models are presented in Section 2.3. Section 2.4 gives preliminary results on linear optimal tracking. The stochastic MF dynamic game consensus models are synthesized and analyzed in Section 2.5. Section 2.6 presents the stochastic MF social optimal consensus models. The social optimal LQR consensus models are presented in Section 2.7. The stochastic MF dynamic game consensus model with (i) a random cost-coupling weight matrix, and (ii) nonuniform localized mean field cost-couplings are presented in Sections 2.8 and 2.9, respectively. 2.10 presents sample numerical simulations of the models. Concluding remarks are stated in Section 2.11.

2.2. Problem Formulations and Terminology

The following notation will be used in this chapter. We use the integer valued subscript as the label for an individual agent of the population. In addition, overbar denotes the expected value of a random variable, i.e., $\bar{z}(t) := Ez(t)$. The integer N is reserved to denote the population size of the system. We use the superscripts N and ∞ for a process (such as state, control, etc.) to indicate the dependence on the population size N or the case of a system with infinite population, respectively. $f(N) = O(g(N))$ means that there are positive constants k and M such that $0 \leq f(N) \leq kg(N)$ for all $N \geq M$.

2.2.1. Stochastic Decentralized Dynamic Game Consensus. Consider a system of N agents. The dynamics of the i^{th} agent is given by a controlled stochastic differential equation

$$dz_i(t) = u_i(t)dt + \sigma dw_i(t), \quad t \geq 0, \quad (2.1)$$

where $z_i(\cdot)$, $u_i(\cdot) \in \mathbb{R}$ are the state and control input of agent i , respectively; σ is a non-negative scalar; and $\{w_i : 1 \leq i \leq N\}$ denotes a sequence of independent standard scalar Wiener processes on some *filtered probability space* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ where \mathcal{F}_t is defined as the σ -field $\sigma(z_i(0), w_i(\tau) : 1 \leq i \leq N, \tau < t)$. We assume that the initial states $\{z_i(0) : 1 \leq i \leq N\}$ are measurable on \mathcal{F}_0 , independent, and independent of $\{w_i : 1 \leq i \leq N\}$. Denote the state and the control of the overall system, respectively, as $z \equiv (z_1, \dots, z_N)^T$ and $u \equiv (u_1, \dots, u_N)^T$. Let u_{-i} be defined as $(u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_N)$.

In this problem formulation the agents are individually incentive driven and non-cooperative such that each agent seeks to minimize its individual quadratic cost function containing a mean field coupling to the states of all other agents. More precisely, the objective of each individual agent i , $1 \leq i \leq N$, is to minimize its discounted cost function given by

$$J_i^N(u_i, u_{-i}) := E \int_0^\infty e^{-\rho t} \left((z_i(t) - \frac{1}{N-1} \sum_{j=1, j \neq i}^N z_j(t))^2 + r u_i^2(t) \right) dt, \quad (2.2)$$

where $r > 0$, $\rho > 0$ is the discount factor, and $z_{-i}^N(\cdot) := \frac{1}{N-1} \sum_{j=1, j \neq i}^N z_j(\cdot)$ is called the *mean field term*. To indicate the dependence of J_i on $u_i(\cdot)$, $u_{-i}(\cdot)$ and the population size N , we write it as $J_i^N(u_i, u_{-i})$. For minimization of J_i^N , the *admissible control* set is taken as

$$\mathcal{U}_i := \left\{ u_i(\cdot) : u_i(t) \text{ is adapted to sigma-field } \sigma(z_j(s) : s \leq t, 1 \leq j \leq N) \right\}.$$

It is important to note that $\text{span}\{\mathbf{1}_N\}$ is an unobservable subspace for the system, where $\mathbf{1}_N$ is the N -dimensional vector of all ones

In this problem we are interested in a competitive solution (u_1, \dots, u_N) for the system of N agents with decentralized information within the mean field modelling. In this scenario, each agent i , $1 \leq i \leq N$, with dynamics (2.1) is associated with the cost J_i^N defined in (2.2), and the objective is to obtain a set of ϵ -Nash strategies such that each control u_i is only a function of time t and local information z_i .

The extensions of this model to the cases of: (i) Long Time Average (LTA) cost functions (see [136]), (ii) random mean field cost-couplings, and (iii) nonuniform localized mean field cost-couplings are presented in Sections 2.5.3, 2.8 and 2.9, respectively.

REMARK 2.1. *The results of this chapter can be extended to the case of agents with cost functions (see Section 2.3):*

$$J_i^N(u_i, u_{-i}) := E \int_0^\infty e^{-\rho t} \left(\frac{1}{N-1} \sum_{j=1, j \neq i}^N (z_i(t) - z_j(t))^2 + r u_i^2(t) \right) dt, \quad 1 \leq i \leq N. \quad (2.3)$$

2.2.2. Stochastic Decentralized Social Optimal Consensus. In this problem formulation, within the mean field modelling, we study the case that the agents seek social optimal decisions and are cooperative. More precisely, the objective of the agents in the system is to minimize a *social cost* defined as

$$J_{\text{soc}}^N(u) := \sum_{i=1}^N J_i^N(u), \quad (2.4)$$

where J_i^N is the individual cost for agent i , $1 \leq i \leq N$, defined in (2.2).

In this case, it is important to note that each individual agent should take into account both reducing its own cost and the social impact of such reductions on the sum of the costs of all other agents. The *social admissible control* is given by

$$\mathcal{U}_{\text{soc}} := \left\{ u(\cdot) : u_i(t, \omega) \text{ is adapted to } \mathcal{F}_t, \forall i \right\},$$

where $\omega \in \Omega$ explicitly indicates the dependence of u_i on the sample. Each $u = (u_1, \dots, u_N) \in \mathcal{U}_{\text{soc}}$ may be viewed as a function of the initial state of the system

$z(0) := (z_1(0), \dots, z_N(0))$ and Brownian motions without being related to the system's state process $z(t)$. The benefit of considering \mathcal{U}_{soc} is that one may find the impact of a generic agent's control on the social cost of the system by fixing the controls of all other agents and perturbing that agent's strategy (see [82]).

In this problem we are interested in a cooperative solution $u = (u_1, \dots, u_N) \in \mathcal{U}_{\text{soc}}$ for the system of N agents with dynamics (2.1) and decentralized information to attain the minimum of the social cost J_{soc}^N defined in (2.4). In this decentralized information pattern, the control of the i^{th} agent u_i is only a function of time t and local information z_i .

We discuss the extension of this model to the case of LTA cost functions in Section 2.6.1.

2.2.3. Deterministic Centralized Social Optimal Consensus. In this problem formulation for the system of N agents with deterministic dynamics:

$$dz_i(t) = u_i(t)dt, \quad 1 \leq i \leq N, \quad t \geq 0, \quad (2.5)$$

we are interested in a social solution $u(\cdot) = (u_1(\cdot), \dots, u_N(\cdot))$ with centralized information to attain the minimum of the social cost:

$$J_{\text{soc,det}}^N(u) := \sum_{i=1}^N \int_0^\infty e^{-\rho t} \left((z_i(t) - \frac{1}{N-1} \sum_{j=1, j \neq i}^N z_j(t))^2 + r u_i^2(t) \right) dt. \quad (2.6)$$

In this centralized information pattern, the feedback control of the i^{th} agent u_i is a function of time t and the global information of the system (z_1, \dots, z_N) . See Remark 2.8 for the extension of this model to the case of agents with LTA cost functions [136].

The reason that we do not consider the stochastic case of the centralized social cost formulation is that the noise causes a steady drift of the agents' states during the centralized feedback iterations which eliminates the possibility of convergent (agreement) group behaviour (see Example 1 in [86]). However, note that this case remains implicitly stochastic since the initial states are viewed as drawn from a common probabilistic distribution.

2.3. Applications

The mean field consensus formulations are motivated by many social, economic, and engineering models. Here, we provide an example of a large population mean field consensus model in the synchronization of nonlinear coupled oscillators; the reader is referred to Section 5.2.4 of Chapter 5 for an economic (finance) example (see [137]).

In [177] Yin et. al. formulate a nonlinear dynamic games model of synchronization of coupled oscillators. Consider a population of N oscillators with dynamics,

$$d\theta_i(t) = (\omega_i + u_i(t))dt + \sigma dw_i(t), \quad 1 \leq i \leq N, \quad t \geq 0, \quad (2.7)$$

where $\theta_i(t) \in [0, 2\pi]$ is the phase of the i^{th} oscillator at time t , $u_i(\cdot)$ is the control input, σ is a non-negative scalar, and $\{\omega_i : 1 \leq i \leq N\}$ denotes a sequence of independent standard scalar Wiener processes. It is assumed that the initial states $\{\theta_i(0)\}$ are chosen independently according to the uniform distribution on $[0, 2\pi]$. It is assumed that at time $t = 0$, the N scalars $\{\omega_i\}$ are chosen independently according to a fixed distribution with density g by Assumption (A1) in [177]. For a homogeneous population g is a Dirac delta function (e.g., $g(\omega) = \delta(\omega)$). The objective of the i^{th} oscillator is to minimize its own cost criterion (see [177]).

$$\eta_i^N(u_i, u_{-i}) := \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T (c(\theta_i; \theta_{-i}) + \frac{1}{2}ru_i^2(t))dt, \quad (2.8)$$

where $\theta_{-i} = (\theta_j)_{j \neq i}$, r is a positive scalar, and c is the cost function:

$$c(\theta_i; \theta_{-i}) := \frac{1}{2N} \sum_{j \neq i}^N \sin^2 \left(\frac{\theta_i - \theta_j}{2} \right).$$

The linearization of the cost function c around the aligned state $\theta_1 = \dots = \theta_N$ gives the LTA version (i.e., the discount factor ρ is zero) of the cost functions (2.3). Hence, the linearized cost version of a homogeneous population (e.g., zero natural frequency or $g(\omega) = \delta(\omega)$) gives us the stochastic dynamic game consensus model with LTA cost functions (see Subsection 2.5.3).

There are many other similar applications of the models considered in this chapter in flocking [46], crowd flow dynamics [49], and social opinion models with a very large number of agents [111]. It is important to note that in all of these examples the “initial state” of the agents contributes to the stationary equilibrium behaviour of the system.

2.4. Preliminary Optimal Control of a Single Agent

Consider a single agent called agent i with linear stochastic dynamics

$$dz_i(t) = u_i(t)dt + \sigma dw_i(t), \quad t \geq 0, \quad (2.9)$$

where $z_i(\cdot) \in \mathbb{R}$ is the state; $u_i(\cdot) \in \mathbb{R}$ is the control input; w_i denotes a standard scalar Wiener process; σ is a non-negative scalar; and $z_i(0)$ is given. The initial state $z_i(0)$ is independent of the process w_i . The objective of this agent is to minimize its discounted cost function given by

$$J(u_i) := E \int_0^\infty e^{-\rho t} \left((z_i(t) - \phi(t))^2 + ru_i^2(t) \right) dt, \quad (2.10)$$

where $\phi(\cdot)$ is a known bounded and continuous function, and r and ρ are positive scalars. For minimization of $J(u_i)$, the *admissible control* set is taken as

$$\mathcal{U}_i := \left\{ u_i(\cdot) : u_i(t) \text{ is adapted to sigma-field } \sigma(z_i(0), w_i(s) : s \leq t), \right. \\ \left. E \int_0^\infty e^{-\rho t} (z_i^2 + u_i^2) dt < \infty \right\}.$$

The set \mathcal{U}_i is nonempty due to controllability of (2.9).

THEOREM 2.1. (*Special case of Propositions 3.1-3.3 in [79]*) *For the optimal control problem (2.9)-(2.10):*

(a) *the algebraic Riccati equation $p^2 + r\rho p - r = 0$ has a unique positive solution*

$$p = (-r\rho + \sqrt{(r\rho)^2 + 4r})/2,$$

(b) the differential equation

$$\frac{ds(t)}{dt} = \left(\rho + \frac{p}{r}\right)s(t) + \phi(t), \quad t \geq 0,$$

has a unique bounded solution: $s(t) = -\int_t^\infty e^{(\rho+p/r)(t-\tau)}\phi(\tau)d\tau$ for $t \geq 0$,

(c) the unique optimal control law $u_i^o := \arg \inf_{u_i \in \mathcal{U}_i} J(u_i)$ is given by

$$u_i^o(\cdot) = -(1/r)(pz_i(\cdot) + s(\cdot)),$$

(d) the optimal cost value is given by

$$J(u_i^o) \equiv \inf_{u_i \in \mathcal{U}_i} J(u_i) = pEz_i^2(0) + 2s(0)Ez_i(0) + q(0),$$

where q is the unique bounded solution of the equation:

$$\frac{dq(t)}{dt} = \rho q + \frac{1}{r}s_i^2 - (\phi(t))^2 - \sigma^2 p.$$

□

REMARK 2.2. The LTA cost version of Theorem 2.1 for LQG optimal tracking problems may be found in [106].

2.5. Stochastic Mean Field Dynamic Game Consensus Model

In this section we consider the stochastic decentralized dynamic game consensus model (2.1)-(2.2) (see Section 2.2.1).

Let the *empirical distribution function* associated with N agents be defined by

$$F_N(x) := \frac{1}{N} \sum_{i=1}^N 1_{\{\bar{z}_i(0) < x\}},$$

where $1_{\{\bar{z}_i(0) < x\}} = 1$ if $\bar{z}_i(0) < x$, and $1_{\{\bar{z}_i(0) < x\}} = 0$ otherwise. We enunciate the assumption:

(A2.1) We assume that (i) the initial states $\{z_i(0) : 1 \leq i \leq N\}$ are independent, and there exists a constant k independent of N such that $\sup_{1 \leq i \leq N} E|z_i(0)|^2 \leq k < \infty$, and (ii) $\{F_N : N \geq 1\}$ converges weakly to a probability distribution F , *i.e.*, for any

bounded and continuous function $\phi(x)$ on \mathbb{R} ,

$$\lim_{N \rightarrow \infty} \int_{\mathbb{R}} \phi(x) dF_N(x) = \int_{\mathbb{R}} \phi(x) dF(x).$$

□

REMARK 2.3. *It is important to note that if the sequence $\{z_i(0) : 1 \leq i \leq N\}$ is generated by independent randomized observations on the Gaussian distribution F , then **(A2.1)**-(ii) holds with probability one by the Strong Law of Large Numbers or the Glivenko-Cantelli theorem [42].*

We take a representative agent and let the expected value of its initial state be denoted by θ which takes a value from a fixed compact set Θ independent of N (this follows from **(A2.1)**). The state process of this agent may be denoted by z_θ and we denote its mean trajectory by $\bar{z}_\theta = Ez_\theta$.

For the design of decentralized control, our idea is to consider the population limit and approximate $z_{-\theta}^N(\cdot) := \frac{1}{N-1} \sum_{j=1}^N z_{\theta_j} 1_{\{\theta_j \neq \theta\}}$ in the finite population model by a deterministic function $\phi^\infty(\cdot)$. Let p be given as in Theorem 2.1-(a). For the infinite population with parameter distribution $F(\theta)$, we construct the following *MF game* (or *Nash Certainty Equivalence* (NCE)) system of equations (see [79]):

$$\frac{ds(t)}{dt} = \left(\rho + \frac{p}{r}\right)s(t) + \phi^\infty(t), \quad (2.11)$$

$$\frac{d\bar{z}_\theta(t)}{dt} = -\frac{1}{r}(p\bar{z}_\theta(t) + s(t)), \quad \bar{z}_\theta(0) = \theta \text{ given}, \quad (2.12)$$

$$\phi^\infty(\cdot) = \int_{\Theta} \bar{z}_\theta(\cdot) dF(\theta), \quad (2.13)$$

where $\phi^\infty(t)$ denotes the average state of the agents in the population limit (i.e., $N \rightarrow \infty$). In particular, $\phi^\infty(0) = \int_{\Theta} \bar{z}_\theta(0) dF(\theta)$ which is the mean value of the overall population's initial state.

Equation system (2.11)-(2.13) prescribes a mass function $\phi^\infty(\cdot)$ characterized by the property that it is reproduced as in (2.13) as the average of all agents' states in the continuum of the agents whenever each individual agent optimally tracks the

same mass $\phi^\infty(\cdot)$ by application of the MF control law

$$u_\theta^\infty(\cdot) = -\frac{1}{r}(pz_\theta(\cdot) + s(\cdot)). \quad (2.14)$$

More precisely, (2.11) is the mass offset optimal tracking equation; (2.12) is obtained by taking expectation of the closed-loop dynamics of the generic agent θ using the control law $u_\theta^\infty(\cdot)$.

By integrating (2.12) with respect to the measure dF , the system of equations (2.11)-(2.13) yields the system

$$\frac{ds(t)}{dt} = \left(\rho + \frac{p}{r}\right)s(t) + \phi^\infty(t), \quad (2.15)$$

$$\frac{d\phi^\infty(t)}{dt} = -\frac{p}{r}\phi^\infty(t) - \frac{1}{r}s(t), \quad \phi^\infty(0) = \int_{\Theta} \bar{z}_\theta(0)dF(\theta), \quad (2.16)$$

for the couple $(s(\cdot), \phi^\infty(\cdot))$ where $\phi^\infty = \int_{\Theta} \bar{z}_\theta dF(\theta)$.

THEOREM 2.2. *The system (2.15)-(2.16) has a unique bounded solution:*

$$(s(t), \phi^\infty(t)) = (-p\phi^\infty(0), \phi^\infty(0)), \quad t \geq 0. \quad (2.17)$$

Proof. We may write the system (2.15)-(2.16) as

$$\frac{d}{dt} \begin{pmatrix} s(t) \\ \phi^\infty(t) \end{pmatrix} = A \begin{pmatrix} s(t) \\ \phi^\infty(t) \end{pmatrix}, \quad (2.18)$$

where

$$A = \begin{pmatrix} \rho + p/r & 1 \\ -1/r & -p/r \end{pmatrix}.$$

A is a singular matrix and may be brought to the diagonal form via $J = P^{-1}AP$ where

$$J = \begin{pmatrix} 0 & 0 \\ 0 & \rho \end{pmatrix}, \quad P = \begin{pmatrix} -p & 1 \\ 1 & -p/r \end{pmatrix}.$$

Therefore, we may write the solution of (2.18) as

$$\begin{aligned} \begin{pmatrix} s(t) \\ \phi^\infty(t) \end{pmatrix} &= e^{At} \begin{pmatrix} s(0) \\ \phi^\infty(0) \end{pmatrix} = P e^{Jt} P^{-1} \begin{pmatrix} s(0) \\ \phi^\infty(0) \end{pmatrix} \\ &= \frac{1}{\rho p} \begin{pmatrix} -(p^2/r)s(0) - p\phi^\infty(0) + e^{\rho t} (s(0) + p\phi^\infty(0)) \\ (p/r)s(0) + \phi^\infty(0) - e^{\rho t} (p/r) (s(0) + p\phi^\infty(0)) \end{pmatrix}, \quad t \geq 0. \end{aligned}$$

Hence, by using the solution of Riccati equation in Theorem 2.1-(a), the unique bounded solution of the system (2.15)-(2.16) is given in (2.17). \square

REMARK 2.4. (2.17) gives the unique solution of the MF system (2.15)-(2.16) irrespective of how large the discount factor ρ is. This is not the case of the general MF linear-quadratic-Gaussian (MF LQG) model with discounted cost functions considered in [79] (see Proposition 4.2 in [79]). Moreover, the approach of Theorem 2.2 to the existence and uniqueness analysis of the MF game model is in general different from fixed point arguments developed in [79, 82, 106] because a related operator (e.g., (4.11) in [79]) is no longer a contraction mapping.

REMARK 2.5. The results of this section can be extended to the case of system of agents with general uniform linear dynamics,

$$dz_i(t) = (az_i(t) + bu_i(t))dt + \sigma dw_i(t), \quad 1 \leq i \leq N, \quad t \geq 0,$$

and discounted cost functions (2.10) where $b \neq 0$ and a is equal to the discount factor $\rho > 0$. This is the only case that the solution method of Theorem 2.2 still holds, i.e., the mass function $\phi^\infty(t)$, $t \geq 0$, remains constant and is equal to $\phi^\infty(0)$.

In a system with finite population N , the infinite population control laws in (2.14) yield the following MF control laws:

$$u_i^o(\cdot) := \frac{-1}{r} (pz_i(\cdot) + s(\cdot)) = \frac{-p}{r} (z_i(\cdot) - \phi^\infty(0)), \quad 1 \leq i \leq N, \quad (2.19)$$

where the infinite population mass effect $\phi^\infty(\cdot)$ is equal to $\phi^\infty(0)$ by (2.17).

It is important to note that the MF control law $u_i^o(\cdot)$ in (2.19) is the optimal tracking control input for the i^{th} agent with dynamics (2.1) and cost function (2.2) where the mean field term $z_{-i}^N(\cdot) = \frac{1}{N-1} \sum_{j=1, j \neq i}^N z_j(\cdot)$ is approximated by the infinite population mass function $\phi^\infty(\cdot) \equiv \phi^\infty(0)$; in other words, $u_i^o(\cdot) \in \mathcal{U}_i$ is the unique optimal control of the i^{th} agent, $1 \leq i \leq N$, with dynamics (2.1) and cost function:

$$J_i^\infty(u_i, \phi^\infty(0)) := E \int_0^\infty e^{-\rho t} \left((z_i(t) - \phi^\infty(0))^2 + r u_i^2(t) \right) dt. \quad (2.20)$$

DEFINITION 2.1. *Mean-consensus is said to be achieved asymptotically for a group of N agents if $\lim_{t \rightarrow \infty} |\bar{z}_i(t) - \bar{z}_j(t)| = 0$ for any i and j , $1 \leq i \neq j \leq N$. \square*

By the MF control laws $u_i^o(\cdot)$, $1 \leq i \leq N$, the state of any agent i , $z_i(\cdot)$, follows an *Ornstein-Uhlenbeck* process of negative feedback around $\phi^\infty(0)$:

$$dz_i(t) = d(z_i(t) - \phi^\infty(0)) = \frac{-p}{r} (z_i(t) - \phi^\infty(0)) dt + \sigma dw_i(t),$$

which has the following solution:

$$z_i^o(t) = \phi^\infty(0) + e^{-(p/r)t} (z_i(0) - \phi^\infty(0)) + \sigma \int_0^t e^{-(p/r)(t-\tau)} dw_i(\tau), \quad t \geq 0. \quad (2.21)$$

We use this solution in the proof of the following theorem.

THEOREM 2.3. *By use of the MF control laws (2.19) in the dynamic game model, (2.1)-(2.2), a mean-consensus is reached asymptotically as time goes to infinity with individual asymptotic variance $\frac{\sigma^2 r}{2p}$.*

Proof. By (2.21) we get $\lim_{t \rightarrow \infty} \bar{z}_i^o(t) = \phi^\infty(0)$, $1 \leq i \leq N$, which shows that all the agents reach mean-consensus asymptotically as time goes to infinity. The mean-consensus value is $\phi^\infty(0)$. For the asymptotic individual variance, by the Itô isometry we get

$$\lim_{t \rightarrow \infty} E(z_i^o(t) - \phi^\infty(0))^2 = \sigma^2 \lim_{t \rightarrow \infty} \int_0^t e^{-\frac{2p(t-\tau)}{r}} d\tau = \frac{\sigma^2 r}{2p},$$

for any i , $1 \leq i \leq N$. \square

2.5.1. The Stability and Performance Analysis of Mean Field Control Laws.

THEOREM 2.4. (*Stability of the MF control laws*) Assume **(A2.1)** holds. Then

$$\sup_{N \geq 1} \max_{1 \leq i \leq N} E \int_0^\infty e^{-\rho t} \left((z_i^o(s))^2 + (u_i^o(s))^2 \right) ds < \infty. \quad (2.22)$$

Proof. See the appendix. \square

In the following theorems the infinite population mass effect, $\phi^\infty(\cdot)$, approximation to the finite population closed-loop centroid of flock of agents is justified.

THEOREM 2.5. (*Convergence in mean-square*) Under **(A2.1)** we have

$$\lim_{N \rightarrow \infty} E \left(\frac{1}{N} \sum_{i=1}^N z_i^o(t) - \phi^\infty(0) \right)^2 = 0, \quad \forall t \geq 0. \quad (2.23)$$

Proof. By (2.21) for any fixed $t \geq 0$ we have

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N z_i^o(t) &= \phi^\infty(0) + e^{-(p/r)t} \left(\frac{1}{N} \sum_{i=1}^N z_i(0) - \phi^\infty(0) \right) \\ &\quad + \frac{\sigma}{N} \sum_{i=1}^N \int_0^t e^{-(p/r)(t-\tau)} dw_i(\tau). \end{aligned} \quad (2.24)$$

By the independence of initial states and Wiener processes, and the Itô isometry we get

$$\begin{aligned} E \left(\frac{1}{N} \sum_{i=1}^N z_i^o(t) - \phi^\infty(0) \right)^2 &= e^{-2(p/r)t} \left(\frac{1}{N} \sum_{i=1}^N \bar{z}_i(0) - \phi^\infty(0) \right)^2 \\ &\quad + \frac{\sigma^2 r}{2pN} (1 - e^{-2(p/r)t}). \end{aligned} \quad (2.25)$$

Hence, (2.25) and **(A2.1)** yield (2.23). \square

THEOREM 2.6. *Assume (A2.1) holds. Then*

$$a) \quad \lim_{N \rightarrow \infty} E \int_0^T \left(\frac{1}{N} \sum_{i=1}^N z_i^o(t) - \phi^\infty(0) \right)^2 dt = 0, \quad 0 < T < \infty, \quad (2.26)$$

$$b) \quad \lim_{N \rightarrow \infty} E \int_0^\infty e^{-\rho t} \left(\frac{1}{N} \sum_{i=1}^N z_i^o(t) - \phi^\infty(0) \right)^2 dt = 0. \quad (2.27)$$

Proof. See the appendix. □

THEOREM 2.7. *Assume (A2.1) holds. Then*

$$\lim_{N \rightarrow \infty} \max_{1 \leq i \leq N} |J_i^N(u_i^o, u_{-i}^o) - J_i^\infty(u_i^o, \phi^\infty)| = 0, \quad (2.28)$$

where

$$J_i^\infty(u_i^o, \phi^\infty) = p(\bar{z}_i(0) - \phi^\infty(0))^2 + \frac{\sigma^2 p}{\rho}, \quad 1 \leq i \leq N.$$

Proof. By the Cauchy-Schwarz inequality and (2.27) we obtain (2.28). But, by (2.17) and Theorem 2.1-(d) we get the value of $J_i^\infty(u_i^o, \phi^\infty)$. □

2.5.2. ϵ -Nash Equilibrium Property of Mean Field Control Laws.

DEFINITION 2.2. [79] *Given $\epsilon > 0$, the set of controls $\{u_i^o \in \mathcal{U}_i : 1 \leq i \leq N\}$ for N agents generates an ϵ -Nash equilibrium with respect to the costs $\{J_i^N : 1 \leq i \leq N\}$, if*

$$J_i^N(u_i^o, u_{-i}^o) - \epsilon \leq \inf_{u_i \in \mathcal{U}_i} J_i^N(u_i, u_{-i}^o) \leq J_i^N(u_i^o, u_{-i}^o), \quad 1 \leq i \leq N.$$

□

For a generic agent $1 \leq i \leq N$ denote

$$(\epsilon_i^N)^2 := E \int_0^\infty e^{-\rho t} \left(\frac{1}{N-1} \sum_{j=1, j \neq i}^N z_j^o(t) - \phi^\infty(0) \right)^2 dt, \quad 1 \leq i \leq N, \quad (2.29)$$

where $z_i^o(\cdot)$ is the closed-loop solution of the i^{th} agent's dynamics (2.21).

LEMMA 2.1. *Assume (A2.1) holds. Then $\lim_{N \rightarrow \infty} \epsilon_i^N = 0$ for any $1 \leq i \leq N$.*

Proof. See the appendix. □

THEOREM 2.8. *Assume (A2.1) holds. Let $\epsilon_N := O(\max_{1 \leq i \leq N} \epsilon_i^N)$, then the set of MF control laws $\{u_i^o \in \mathcal{U}_i : 1 \leq i \leq N\}$ in (2.19) generates an ϵ_N -Nash equilibrium, i.e., for any fixed i , $1 \leq i \leq N$, we have*

$$J_i^N(u_i^o, u_{-i}^o) - \epsilon_N \leq \inf_{u_i \in \mathcal{U}_i} J_i^N(u_i, u_{-i}^o) \leq J_i^N(u_i^o, u_{-i}^o) \quad (2.30)$$

where $\lim_{N \rightarrow \infty} \epsilon_N = 0$.

Proof. The right inequality is trivial. To establish the left inequality, we see that for a fixed i , $1 \leq i \leq N$, we have

$$\begin{aligned} & J_i^N(u_i^o, u_{-i}^o) \\ &= E \int_0^\infty e^{-\rho t} \left((z_i^o(t) - \frac{1}{N-1} \sum_{j \neq i}^N z_j^o(t))^2 + r(u_i^o(t))^2 \right) dt \quad (2.31) \\ &\leq E \int_0^\infty e^{-\rho t} \left((z_i^o(t) - \phi^\infty(0))^2 + r(u_i^o(t))^2 \right) dt \\ &\quad + E \int_0^\infty e^{-\rho t} \left(\frac{1}{N-1} \sum_{j \neq i}^N z_j^o(t) - \phi^\infty(0) \right)^2 dt \\ &\quad + 2E \int_0^\infty e^{-\rho t} \left(z_i^o(t) - \phi^\infty(0) \right) \left(\phi^\infty(0) - \frac{1}{N-1} \sum_{j \neq i}^N z_j^o(t) \right) dt \\ &=: J_i^\infty(u_i^o, \phi^\infty(0)) + I_1^N + I_2^N \quad (2.32) \end{aligned}$$

where J_i^∞ is defined in (2.20). We have $I_1^N = (\epsilon_i^N)^2$ where ϵ_i^N is defined in (2.29). By the Cauchy-Schwarz inequality we have $|I_2^N| \leq 2\sqrt{k}\epsilon_i^N$ where

$$k = \max_{1 \leq i \leq N} E \int_0^\infty e^{-\rho t} \left(z_i^o(t) - \phi^\infty(0) \right)^2 ds < \infty$$

is independent of N by (2.22). Therefore,

$$I_1^N + I_2^N = O(\epsilon_i^N). \quad (2.33)$$

But, $u_i^o(\cdot)$ is the optimal tracking control with respect to the cost $J_i^\infty(u_i, \phi^\infty(0))$ (i.e., $u_i^o(\cdot) = \arg \inf_{u_i \in \mathcal{U}_i} J_i^\infty(u_i, \phi^\infty(0))$), by the construction of the MF system (2.11)-(2.13). A similar argument yields (see Theorem 5.6 in [79])

$$\begin{aligned} J_i^\infty(u_i^o, \phi^\infty(0)) &\leq \inf_{u_i \in \mathcal{U}_i} J_i^\infty(u_i, \phi^\infty(0)) \\ &\leq \inf_{u_i \in \mathcal{U}_i} J_i^N(u_i, u_{-i}^o) + O(\epsilon_i^N). \end{aligned} \quad (2.34)$$

Hence, (2.32), (2.33) and (2.34) yield

$$\begin{aligned} J_i^N(u_i^o, u_{-i}^o) &\leq J_i^\infty(u_i^o, \phi^\infty(0)) + O(\epsilon_i^N) \\ &\leq \inf_{u_i \in \mathcal{U}_i} J_i^N(u_i, u_{-i}^o) + \epsilon_N \end{aligned}$$

where $\epsilon_N := O(\max_{1 \leq i \leq N} \epsilon_i^N)$. Lemma 2.1 implies that $\lim_{N \rightarrow \infty} \epsilon_N = 0$. \square

2.5.3. Extension to the Case of Long Time Average (LTA) Cost Functions. Assume that in a system with population size N and individual dynamics (2.1), the objective of the i^{th} individual agent is to minimize the *Long Time Average* (LTA) cost function (see our work [136]):

$$J_{\text{LTA},i}^N(u_i, u_{-i}) := \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left((z_i(t) - \frac{1}{N-1} \sum_{j=1, j \neq i}^N z_j(t))^2 + r u_i^2(t) \right) dt. \quad (2.35)$$

Then the *MF system* of the dynamic game problem (2.1)-(2.35) is given by [136]:

$$\frac{ds(t)}{dt} = \frac{1}{\sqrt{r}} s(t) + \phi^\infty(t), \quad (2.36)$$

$$\frac{d\phi^\infty(t)}{dt} = -\frac{1}{\sqrt{r}} \phi^\infty(t) - \frac{1}{r} s(t), \quad \phi^\infty(0) = \int_{\Theta} \bar{z}_\theta(0) dF(\theta), \quad (2.37)$$

and \sqrt{r} is the positive solution of the algebraic Riccati equation $p^2/r - 1 = 0$.

Similarly to Theorem 2.2 one can show that the system (2.36)-(2.37) has a unique bounded solution:

$$(s(t), \phi^\infty(t)) = (-\sqrt{r} \phi^\infty(0), \phi^\infty(0)), \quad t \geq 0.$$

The set of MF control laws in the discounted model (2.1)-(2.35) is:

$$u_i^o(\cdot) := \frac{-1}{\sqrt{r}}(z_i(\cdot) - \phi^\infty(0)), \quad 1 \leq i \leq N, \quad (2.38)$$

which yields the closed loop solution:

$$z_i^o(t) = \phi^\infty(0) + e^{-\frac{t}{\sqrt{r}}}(z_i(0) - \phi^\infty(0)) + \sigma \int_0^t e^{-\frac{(t-\tau)}{\sqrt{r}}} dw_i(\tau), \quad t \geq 0, \quad (2.39)$$

for the i^{th} agent.

It is clear that the system reaches mean-consensus in $\phi^\infty(0)$ asymptotically as time goes to infinity (similar to Theorem 2.3) and moreover these control strategies possess an ϵ_N -Nash equilibrium property almost surely (a.s.) where ϵ_N goes to zero a.s. as the population size N approaches infinity (similar to Theorem 2.8) [136].

REMARK 2.6. *Since the LTA cost criterion is popular in applications where transients are fast and negligible (one is choosing essentially from the attainable “steady states” solutions), the resulting optimal control is not unique and the transient behaviour has no effect on this cost criterion. It is important to note that the optimal control could be made unique by avoiding time averaging in the cost, and instead resorting to the concept of overtaking optimality (see [27]).*

2.6. Stochastic Mean Field Social Optimum Consensus Model

In this section we take the *Social Certainty Equivalence* (SCE) approach introduced in [82] to design a set of decentralized strategies for minimizing the social cost J_{soc}^N defined in (2.4) as the population size N approaches infinity. In this case it is necessary for each individual agent not only to reduce its own cost but also to consider the impact of such reductions on the sum of the costs of all other agents.

To apply a mean field approximation to this problem we take the following steps: First, we quantify the impact of a control perturbation of a generic agent into the social cost change of the system; Second, in a system with finite population size N we determine a mean field coupling structure in the cost function of a generic agent as the effect of the mass of all other agents; Finally, in the infinite population limit

we approximate the mean field coupling by a deterministic function called the infinite population mass effect and derive the set of mean field equations to characterize that mass effect.

Let the social optimal control minimizing J_{soc}^N be denoted by $\hat{u} := (\hat{u}_1, \dots, \hat{u}_N) \in \mathcal{U}_{\text{soc}}$. Let $\hat{z}_i(\cdot)$ be the corresponding closed-loop solution of the dynamics of the i^{th} individual agent (2.1) by applying $\hat{u}_i(\cdot)$. Denote $\hat{z}_{-i}^N(\cdot) := \frac{1}{N-1} \sum_{j=1, j \neq i}^N \hat{z}_j(\cdot)$ and $\hat{u}_{-i} := (\hat{u}_1, \dots, \hat{u}_{i-1}, \hat{u}_{i+1}, \dots, \hat{u}_N)$. Let the *individual based social admissible control* set for the i^{th} agent be

$$\mathcal{U}_{\text{soc},i} := \{u_i(\cdot) : u_i(t, \omega) \text{ is adapted to } \mathcal{F}_t\},$$

where $\omega \in \Omega$ explicitly indicates the dependence of u_i on the sample.

LEMMA 2.2. **[82]** *Let $z(0)$ be given and $\hat{u} \in \mathcal{U}_{\text{soc}}$ attain the minimum of J_{soc}^N in the admissible control set \mathcal{U}_{soc} . Then $\hat{u}_i \in \mathcal{U}_{\text{soc},i}$ is the unique optimal control of the control problem:*

$$\begin{aligned} dz_i(t) &= u_i(t)dt + \sigma dw_i(t), & t \geq 0, \\ \inf_{u_i \in \mathcal{U}_{\text{soc},i}} J_{\text{soc}}^N(u_i, \hat{u}_{-i}) &\equiv \inf_{u_i \in \mathcal{U}_{\text{soc},i}} \sum_{k=1}^N J_k^N(u_i, \hat{u}_{-i}), \end{aligned} \quad (2.40)$$

i. e., $\hat{u}_i = \arg \inf_{u_i \in \mathcal{U}_{\text{soc},i}} J_{\text{soc}}^N(u_i, \hat{u}_{-i})$. □

We note that in the lemma above since J_{soc}^N is convex in (z, u) and $r > 0$, the existence and uniqueness of an optimal control \hat{u} holds (see [82]).

LEMMA 2.3. *The optimal control $\hat{u}_i \in \mathcal{U}_{\text{soc},i}$ in the preceding lemma is the solution of the control problem:*

$$\arg \inf_{u_i \in \mathcal{U}_{\text{soc},i}} J_{\text{soc},i}^N(u_i) := \arg \inf_{u_i \in \mathcal{U}_{\text{soc},i}} E \int_0^\infty e^{-\rho t} \left(\left(1 + \frac{1}{N-1}\right) (z_i^2 - 2z_i \hat{z}_{-i}^N) + r u_i^2 \right) dt. \quad (2.41)$$

Proof. Here we take the approach of Lemma 5 in [82]. Since $\hat{u} \in \mathcal{U}_{\text{soc}}$, the process \hat{u}_{-i} has been specified in advance. Hence, \hat{u}_{-i} and subsequently $\hat{z}_{-i}^N(\cdot) =$

$\frac{1}{N-1} \sum_{j=1, j \neq i}^N \hat{z}_j(\cdot)$ do not change with $\hat{u}_i \in \mathcal{U}_{\text{soc}, i}$. We may rewrite the cost $J_{\text{soc}}^N(u_i, \hat{u}_{-i})$ in (2.40) as follows. First, we have

$$J_i^N(u_i, \hat{u}_{-i}) = E \int_0^\infty e^{-\rho t} (d_i^i(t) + e_i^i(t)) dt, \quad (2.42)$$

where

$$d_i^i := z_i^2 - 2z_i \hat{z}_{-i}^N + r u_i^2, \quad e_i^i := (\hat{z}_{-i}^N)^2.$$

Second, for $1 \leq j \neq i \leq N$, we have

$$J_j^N(u_i, \hat{u}_{-i}) = E \int_0^\infty e^{-\rho t} (d_j^i(t) + e_j^i(t)) dt, \quad (2.43)$$

where

$$d_j^i := \left(\frac{1}{N-1} \right)^2 z_i^2 - \frac{2z_i}{N-1} \left(\hat{z}_j - \frac{1}{N-1} \sum_{k=1, k \neq i, j}^N \hat{z}_k \right),$$

$$e_j^i := \left(\hat{z}_j - \frac{1}{N-1} \sum_{k=1, k \neq i, j}^N \hat{z}_k \right)^2 + r \hat{u}_j^2.$$

Then using (2.42) and (2.43), and noting that $e_k^i(\cdot)$, $1 \leq k \leq N$, do not change with u_i , the optimal control problem (2.40) may be shown to be equivalent to

$$\begin{aligned} \arg \inf_{u_i \in \mathcal{U}_{\text{soc}, i}} J_{\text{soc}}^N(u_i, \hat{u}_{-i}) &\equiv \arg \inf_{u_i \in \mathcal{U}_{\text{soc}, i}} \sum_{k=1}^N J_k^N(u_i, \hat{u}_{-i}) \\ &= \arg \inf_{u_i \in \mathcal{U}_{\text{soc}, i}} E \int_0^\infty e^{-\rho t} \left(\sum_{k=1}^N d_k^i(t) \right) dt, \end{aligned}$$

which is (2.41), where we observe that all functions corresponding to the arg-inf operations above exist. \square

The cost (2.41) identifies a mean field coupling structure in that all other agents' effect on the i^{th} generic agent appears in the form of $\hat{z}_{-i}^N(\cdot)$ which does not change with $u_i(\cdot)$. Now we may approximate $\hat{z}_{-i}^N(\cdot) = \frac{1}{N-1} \sum_{j=1, j \neq i}^N \hat{z}_j(\cdot)$ in (2.41) by a deterministic bounded and continuous function $\hat{\phi}^\infty(\cdot)$ as the population size N approaches infinity.

The following optimal control problem may be seen as an infinite population approximation to the problem of Lemma 2.3:

$$\arg \inf_{u_i \in \mathcal{U}_{\text{soc},i}} J_{\text{soc},i}^{\infty}(u_i) := \arg \inf_{u_i \in \mathcal{U}_{\text{soc},i}} E \int_0^{\infty} e^{-\rho t} \left(z_i^2(t) - 2z_i(t)\hat{\phi}^{\infty}(t) + ru_i^2(t) \right) dt, \quad (2.44)$$

where $\hat{\phi}^{\infty}(\cdot)$ is a bounded and continuous function.

Applying the SCE approach developed in [82], we formulate the following *MF social system*:

$$\frac{ds(t)}{dt} = \left(\rho + \frac{p}{r} \right) s(t) + \hat{\phi}^{\infty}(t), \quad (2.45)$$

$$\frac{d\bar{z}_{\theta}(t)}{dt} = -\frac{1}{r} (p\bar{z}_{\theta}(t) + s(t)), \quad \bar{z}_{\theta}(0) = \theta \text{ given}, \quad (2.46)$$

$$\hat{\phi}^{\infty}(\cdot) = \int_{\Theta} \bar{z}_{\theta}(\cdot) dF(\theta), \quad (2.47)$$

in the continuum of the agents. Equation (2.45) is the mass offset optimal tracking equation; (2.46) is the closed-loop dynamics of the generic agent θ with the mean field optimal control law

$$\hat{u}_{\theta}^{\infty}(\cdot) = -\frac{1}{r} (pz_{\theta}(\cdot) + s(\cdot)), \quad (2.48)$$

and given $z_{\theta}(0)$; and in (2.47) $\hat{\phi}^{\infty}(\cdot) \equiv \lim_{N \rightarrow \infty} (1/N) \sum_{i=1}^N \hat{z}_i(\cdot)$ acts as an approximation to $\hat{z}_{-\theta}^N(\cdot)$.

REMARK 2.7. *The equation system (33)-(35) in [82] with parameters $A = 0$, $B = 1$, $R = r$, $Q = 1$, $\Pi = p$, $\Gamma = 1$ and $\eta = 0$ also gives the MF social system (2.11)-(2.13).*

But the MF social system, (2.45)-(2.47), is the same as the MF game system of equation, (2.11)-(2.13). This shows that the impact of each individual agent on the cost of other agents becomes negligible as the number of agents goes to infinity. The unique solution of equation (2.45) is $s(\cdot) = -p\hat{\phi}^{\infty}(0)$ and $\hat{\phi}^{\infty}(\cdot)$ in (2.47) equals $\hat{\phi}^{\infty}(0) = \phi^{\infty}(0)$ based on Theorem 2.2.

In a system with finite population N the infinite population control laws in (2.48) yield the following *MF social control laws* for each individual agent i :

$$\hat{u}_i(\cdot) := \frac{-p}{r}(z_i(\cdot) - \phi^\infty(0)), \quad 1 \leq i \leq N, \quad (2.49)$$

which is the same as the MF Nash control law (2.19). Hence,

$$\text{MF-Nash Controls } U_{\text{Nash}}^\infty = \text{MF-Social Controls } U_{\text{Soc}}^\infty.$$

Applying the MF social control laws (2.49) yields the following solution:

$$\begin{aligned} \hat{z}_i(t) &= \phi^\infty(0) + e^{-(p/r)t}(z_i(0) - \phi^\infty(0)) \\ &\quad + \sigma \int_0^t e^{-(p/r)(t-\tau)} dw_i(\tau), \quad t \geq 0, \quad 1 \leq i \leq N. \end{aligned} \quad (2.50)$$

But, these solutions are the same as (2.21). Hence, a system of agents with the social control laws (2.49) reaches mean-consensus asymptotically as time goes to infinity by Theorem 2.3.

The infinite population mass effect approximation, $\hat{\phi}^\infty(\cdot)$, to the finite population closed-loop centroid of flock of agents is justified in Theorems 2.5 and 2.6 and the Lemma below.

LEMMA 2.4. *Assume (A2.1) holds. Then*

$$|J_{\text{soc},i}^N(\hat{u}_i) - J_{\text{soc},i}^\infty(\hat{u}_i)| = O(\epsilon_i^N), \quad 1 \leq i \leq N, \quad (2.51)$$

where $J_{\text{soc},i}^N$ and $J_{\text{soc},i}^\infty$ are respectively given in (2.41) and (2.44), and ϵ_i^N is defined in (2.29).

Proof. By the Cauchy-Schwarz inequality, (2.22) and (2.27) we obtain the desired estimate where $\lim_{N \rightarrow \infty} \epsilon_i^N = 0$ for any $1 \leq i \leq N$ (see Lemma 2.1). \square

2.6.1. Extension to the Case of Long Time Average (LTA) Cost Functions. In the case of LTA cost functions, the objective of the agents with dynamics

(2.1) is to minimize

$$J_{\text{Ita,soc}}^N(u) := \sum_{i=1}^N J_{\text{Ita},i}^N(u), \quad (2.52)$$

where $J_{\text{Ita},i}^N$ is given in (2.35). Following arguments exactly parallel to those used in the derivation of Lemma 2.3 and equation system (2.45)-(2.47), one can show that the MF social system for the optimal control problem (2.1)-(2.52) is the equation system (2.36)-(2.37) for the LTA game problem. Therefore, the set of corresponding MF social control laws is the same as the set of MF Nash control laws $\{u_i^o(\cdot), 1 \leq i \leq N\}$ given in (2.38).

2.7. Social Optimal LQR Consensus Model

In this section we consider the social optimal control problem (2.5)-(2.6) with centralized information (see Section 2.2.3). In a system with N agents we may write the social cost function (2.6) as

$$\begin{aligned} J_{\text{soc,det}}^N(u) &\equiv \int_0^\infty e^{-\rho t} \sum_{i=1}^N \left(\left(z_i - \frac{1}{N-1} \sum_{j \neq i}^N z_j \right)^2 + r u_i^2 \right) dt \\ &= \int_0^\infty e^{-\rho t} \left((L_N z)^T L_N z + u^T R_N u \right) dt, \end{aligned} \quad (2.53)$$

where $z = (z_1, \dots, z_N)$, $u = (u_1, \dots, u_N)$, $R_N = \text{diag}(r, \dots, r)$ and $L_N = (l_{ij})_{N \times N}$ where $l_{ii} = 1$ and $l_{ij} = -1/(N-1)$ for all $j \neq i$, $1 \leq i, j \leq N$. L_N is a time-invariant Laplacian matrix corresponding to a Completely Connected (CC) (i.e., clique) graph which is symmetric and positive semi-definite with a single zero eigenvalue (see [141]). Since $L_N \mathbf{1}_N = 0$, $\text{span}\{\mathbf{1}_N\}$ is an unobservable subspace for the system (2.5)-(2.6).

The vector form of the system dynamics (2.5) is

$$dz(t) = u(t)dt, \quad t \geq 0. \quad (2.54)$$

Let Π_N be the positive semi-definite solution of the discounted (matrix) algebraic Riccati equation:

$$-\rho\Pi_N - \Pi_N R_N^{-1} \Pi_N + L_N^T L_N = 0. \quad (2.55)$$

Then, the unique optimal control law for the system with dynamics (2.54) and centralized LQR cost function (2.53) is given by

$$u^*(t) = -R_N^{-1} \Pi_N z(t), \quad t \geq 0, \quad (2.56)$$

which gives the centralized dynamics:

$$dz^*(t) = -R_N^{-1} \Pi_N z^*(t) dt, \quad t \geq 0. \quad (2.57)$$

LEMMA 2.5. *The solution of the Riccati equation (2.55) is a positive semi-definite matrix with a single zero eigenvalue.*

Proof. Since $L_N \mathbf{1}_N = 0$, by multiplying the right hand side of (2.55) by $\mathbf{1}_N$ we get $0 \leq \rho \Pi_N \mathbf{1}_N = -\Pi_N R_N^{-1} \Pi_N \mathbf{1}_N \leq 0$. This results in $\Pi_N \mathbf{1}_N = 0$ which shows that zero is an eigenvalue of Π_N . By using an orthogonal transformation Ψ such that $\Psi^T L_N^T L_N \Psi = \text{Diag}(\lambda_i)$ from (2.55) we obtain

$$-\rho \Psi^T \Pi_N \Psi - \Psi^T \Pi_N \Psi R_N^{-1} \Psi^T \Pi_N \Psi + \Psi^T L_N^T L_N \Psi = 0.$$

We require the entry of $\Psi^T \Pi_N \Psi$ at the first row and the first column to be zero, corresponding to the unobservable state in the new coordinate system. Then we may find a unique $\Psi^T \Pi_N \Psi \geq 0$ of rank $N - 1$, and subsequently find $\Pi_N \geq 0$ to (2.55). Therefore, the matrix $\hat{\Pi}_N$ is a Laplacian matrix with a single zero eigenvalue. \square

THEOREM 2.9. *The system (2.57) reaches consensus in $(1/N) \sum_{i=1}^N z_i(0)$ asymptotically as time goes to infinity.*

Proof. By Theorem 1 in [141] the system reaches consensus asymptotically as time goes to infinity. Since the network is undirected the agents reach *average consensus* (AC), i.e., consensus value is $(1/N) \sum_{i=1}^N z_i(0)$ (see [141]). \square

The above theorem implies that

$$\text{Cent. LQR Controls } U_{\text{Cent}}^N = \text{SC-CC Controls } U_{\text{SC}}^N.$$

It is important to note that the non-zero minimal discounted social cost with MF controls is in general different from the non-zero minimal discounted social cost with centralized control laws since the feedback control laws are in general different.

REMARK 2.8. The results of this section can be extended to the case of agents with LTA cost functions (see [136]). Furthermore, in the deterministic formulation for a finite population system it can be shown that (i) the LTA cost of each individual at the MF Nash equilibrium, (ii) the minimal LTA social cost with decentralized MF strategies and (iii) the minimal LTA social cost with centralized information are equal to zero. Hence, this MF consensus problem formulation provides a class of dynamic games where social efficiency is achieved for a finite population.

2.7.1. Deterministic Mean Field Algorithm with Feedback. In a deterministic system (i.e., $\sigma = 0$) with population $N \geq 2$ let the state of the i^{th} agent be

$$z_i(t) = \phi^\infty(0) + e^{-(p/r)t} (z_i(0) - \phi^\infty(0)), \quad t \geq 0, \quad 1 \leq i \leq N, \quad (2.58)$$

by the game (2.19) or MF social control laws (2.49). This is the solution of the dynamics

$$dz_i(t) = -\frac{p}{r} (z_i(t) - \phi^\infty(0)) dt, \quad t \geq 0, \quad 1 \leq i \leq N,$$

where $z_i(0) = \bar{z}_i(0)$ and $\phi^\infty(0)$ are given to any agent i , $1 \leq i \leq N$.

Alternatively, we assume that the agents have no a priori information on $\phi^\infty(0)$ and need to estimate it based on the centralized information dynamics (2.57) by observing the states of other agents over time. More precisely, we assume that the

i^{th} agent has the following feedback dynamics:

$$dx_i^N(t) = -\frac{p}{r} \left(x_i^N(t) - \frac{1}{N-1} \sum_{j=1, j \neq i}^N x_j^N(t) \right) dt, \quad t \geq 0, \quad 1 \leq i \leq N,$$

where $x_i(0) = z_i(0)$ is given.

THEOREM 2.10. *Assume (A2.1) holds. Then*

$$a) \quad \lim_{N \rightarrow \infty} \lim_{t \rightarrow \infty} |z_i(t) - x_i^N(t)| = 0, \quad (2.59)$$

$$b) \quad \lim_{N \rightarrow \infty} \lim_{t \rightarrow \infty} x_i^N(t) = \phi^\infty(0). \quad (2.60)$$

Proof. a) For the i^{th} agent, $1 \leq i \leq N$, let $y_i^N(\cdot) = z_i(\cdot) - x_i^N(\cdot)$ where $y_i^N(0) = 0$.

Then

$$\begin{aligned} dy_i^N(t) &= dz_i(t) - dx_i^N(t) \\ &= -\frac{p}{r} \left((z_i(t) - \phi^\infty(0)) - (x_i^N(t) - \frac{1}{N-1} \sum_{j=1, j \neq i}^N x_j^N(t)) \right) dt \\ &= -\frac{p}{r} \left(y_i^N(t) - \phi^\infty(0) + \frac{1}{N-1} \sum_{j=1, j \neq i}^N (z_j(t) - y_j^N(t)) \right) dt, \quad t \geq 0, \end{aligned}$$

which by (2.58) results in

$$dy_i^N(t) = -\frac{p}{r} \left(y_i^N(t) - \frac{1}{N-1} \sum_{j=1, j \neq i}^N y_j^N(t) + \frac{e^{-(p/r)t}}{N-1} \sum_{j=1, j \neq i}^N (z_j(0) - \phi^\infty(0)) \right) dt,$$

which we may write in vector form as

$$dy^N(t) = -\frac{p}{r} L_N y^N(t) dt + \frac{p}{r} e^{-(p/r)t} I_N \delta^N dt, \quad t \geq 0, \quad (2.61)$$

where I_N is the identity matrix, Laplacian matrix L_N is defined in (2.53), $y^N := (y_1^N, \dots, y_N^N)$, $\delta^N = (\delta_1^N, \dots, \delta_N^N)$ and $\delta_i^N := \frac{1}{N-1} \sum_{j=1, j \neq i}^N (z_j(0) - \phi^\infty(0))$ for every

$1 \leq i \leq N$. The solution of (2.61) is

$$\begin{aligned} y^N(t) &= -\frac{p}{r} \int_0^t e^{-(p/r)L_N(t-\tau)} e^{-(p/r)I_N\tau} \delta^N d\tau \\ &= \left(e^{-(p/r)I_N t} - e^{-(p/r)L_N t} \right) A_N^{-1} \delta^N, \quad t \geq 0, \end{aligned}$$

where $A_N := I_N - L_N = \frac{1}{N-1}(\mathbf{1}_{N \times N} - I_N)$ is a doubly stochastic matrix with inverse $A_N^{-1} = \mathbf{1}_{N \times N} - (N-1)I_N$ (note that $\mathbf{1}_{N \times N}$ is the $N \times N$ matrix of ones). Since $\lim_{t \rightarrow \infty} \exp(- (p/r)L_N t) = \frac{1}{N} \mathbf{1}_{N \times N}$ (see [141]) we have

$$\lim_{t \rightarrow \infty} \|y^N(t)\| \leq \sup_{1 \leq i \leq N} |\delta_i^N|,$$

where $\lim_{N \rightarrow \infty} |\delta_i^N| = 0$ for any fixed $1 \leq i \leq N$ by (A2.1) and hence we obtain (2.59).

b) (2.60) follows directly from (2.58) and (2.59). \square

The preceding Theorem implies that in the deterministic problem formulation

$$\begin{aligned} \text{MF-Nash Controls } U_{\text{Nash}}^\infty &= \text{MF-Social Controls } U_{\text{Soc}}^\infty \\ &= \text{Cent. LQR Controls } U_{\text{Cent}}^\infty = \text{SC-CC Controls } U_{\text{SC}}^\infty, \end{aligned}$$

in the limit as time goes to infinity.

REMARK 2.9. *In general it is essential that the a priori information incorporates the population initial mean $\phi^\infty(0)$. If the agents have no a priori information on $\phi^\infty(0)$, then they could rely on an initial finite phase to reach a (necessarily approximate) consensus. Past that phase, they could revert to relying on the obtained approximate $\phi^\infty(0)$.*

2.8. Game Consensus Model with Random Cost-Couplings

Consider a system of N agents with stochastic dynamics (2.1). We introduce a random undirected weight matrix $\Omega_N = [\omega_{ij}^{(N)}]$ with all zeros on the main diagonal,

and off diagonal elements

$$\omega_{ij}^{(N)} = \begin{cases} \frac{1}{q(N-1)} & \text{with probability } q, \\ 0 & \text{with probability } 1 - q, \end{cases} \quad (2.62)$$

for $1 \leq i \neq j \leq N$, where $q \in (0, 1)$. The existence of an undirected link between agent i and agent $j \neq i$ is determined randomly, independently of other links and all the random processes and initial states (see the centralized LQ consensus model of [157] with randomly switching communication graphs).

Since the weights $\omega_{ij}^{(N)}$ are independent random variables, the *expected value* of the weight matrix, $E\Omega_N = \bar{\Omega}_N$, may be defined entry wise as

$$\bar{\omega}_{ij}^{(N)} = \begin{cases} \frac{1}{N-1} & i \neq j, \\ 0 & i = j, \end{cases}$$

for $1 \leq i, j \leq N$ (see [148]). Therefore, the expected weight matrix $\bar{\Omega}_N$ corresponds to a Completely Connected (CC) topology.

Let $[t_k, t_{k+1})$, $k = 1, 2, \dots$, be an infinite sequence of time intervals where $t_1 = 0$, $t_{k+1} = t_k + \tau$, where $0 < \tau < \infty$. We assume that the weight matrix $\Omega(t)$ is constrained to change randomly (based on (2.62)) at the transition instants t_k , $k = 1, 2, \dots$. The weights of the random matrix $\Omega(\cdot)$ at each interval are completely independent of their values on all other previous and future intervals and hence $\Omega(\cdot)$ is an *independent and identically distributed* (i.i.d.) random matrix.

The objective of the i^{th} individual agent is to minimize the discounted cost function:

$$J_{\text{rand},i}^N(u_i, u_{-i}) := E \int_0^\infty e^{-\rho t} \left((z_i(t) - \sum_{j=1}^N \omega_{ij}^{(N)}(t) z_j(t))^2 + r u_i^2(t) \right) dt, \quad (2.63)$$

where r and ρ are positive integers. In (2.63) the term $\sum_{j=1}^N \omega_{ij}^{(N)}(\cdot) z_j(\cdot)$ is called the *random mean field* coupling.

By using the expected weight matrix $\bar{\Omega}_N$ which gives the deterministic weights $1/(N-1)$ of a completely connected topology, the MF game system for the dynamic

game problem (2.1)-(2.63) is the same as equation system (2.11)-(2.13). Therefore, the set of MF control laws is (2.19) which results in the solution (2.21):

$$z_i^o(t) = \phi^\infty(0) + e^{-(p/r)t}(z_i(0) - \phi^\infty(0)) + \sigma \int_0^t e^{-(p/r)(t-\tau)} dw_i(\tau), \quad t \geq 0.$$

for the i^{th} agent ($1 \leq i \leq N$) where $\phi^\infty(0) = \int_{\Theta} \bar{z}_\theta(0) dF(\theta)$ is the mean value of the initial states. Hence, by Theorem 2.3 the system reaches mean-consensus asymptotically as time goes to infinity.

Denote

$$(\eta_i^N)^2 := E \int_0^\infty e^{-\rho t} \left(\sum_{j=1}^N (\omega_{ij}^{(N)}(t) - \bar{\omega}_{ij}^{(N)}) z_j^o(t) \right)^2 dt, \quad 1 \leq i \leq N. \quad (2.64)$$

THEOREM 2.11. *Assume (A2.1) holds. Then the set of MF control laws $\{u_i^o \in \mathcal{U}_i : 1 \leq i \leq N\}$ in (2.19) for the dynamic game problem (2.1)-(2.63) satisfies*

$$J_{\text{rand},i}^N(u_i^o, u_{-i}^o) - O(\epsilon_i^N) - O(\eta_i^N) \leq \inf_{u_i \in \mathcal{U}_i} J_i^N(u_i, u_{-i}^o), \quad 1 \leq i \leq N,$$

where $\lim_{N \rightarrow \infty} \epsilon_i^N = 0$ and $\eta_i^N = O\left(\sqrt{1-q}/\sqrt{Nq}\right)$.

Proof. For a fixed i , $1 \leq i \leq N$, we have

$$\begin{aligned} J_{\text{rand},i}^N(u_i^o, u_{-i}^o) &= E \int_0^\infty e^{-\rho t} \left((z_i^o(t) - \sum_{j=1}^N \omega_{ij}^{(N)}(t) z_j^o(t))^2 + r(u_i^o(t))^2 \right) dt \\ &= E \int_0^\infty e^{-\rho t} \left(\left(z_i^o(t) - \sum_{j=1}^N \bar{\omega}_{ij}^{(N)} z_j^o(t) + \sum_{j=1}^N (\bar{\omega}_{ij}^{(N)} - \omega_{ij}^{(N)}(t)) z_j^o(t) \right)^2 + r(u_i^o(t))^2 \right) dt \\ &\leq E \int_0^\infty e^{-\rho t} \left((z_i^o(t) - \frac{1}{N-1} \sum_{j \neq i}^N z_j^o(t))^2 + r(u_i^o(t))^2 \right) dt + O(\eta_i^N), \\ &= J_i^N(u_i^o, u_{-i}^o) + O(\eta_i^N), \end{aligned} \quad (2.65)$$

by the Cauchy-Schwarz inequality and (2.22), where $J_i^N(u_i^o, u_{-i}^o)$ is given in (2.31) (and hence is bounded by (2.32)) and η_i^N is defined in (2.64). The inequality (2.65)

together with (2.30) in the statement of theorem 2.8 yields in

$$J_{\text{rand},i}^N(u^o) \leq J_i^N(u_i^o, u_{-i}^o) + O(\eta_i^N) \leq \inf_{u_i \in \mathcal{U}_i} J_i^N(u_i, u_{-i}^o) + O(\epsilon_i^N) + O(\eta_i^N),$$

where ϵ_i^N is defined in (2.29) and $\lim_{N \rightarrow \infty} \epsilon_i^N = 0$ by Lemma 2.1.

Let $[\Omega_N(\cdot)]_i$ be the i^{th} row of the random matrix $\Omega_N(\cdot)$. Then we have

$$(\eta_i^N)^2 = E \int_0^\infty e^{-\rho t} \left(([\Omega_N(t)]_i - [\bar{\Omega}_N]_i) z^o(t) \right)^2 dt.$$

By the mutual independency of random weights in the matrices $\Omega_N(\cdot)$ (see (2.62)) and the fact that for fixed $1 \leq i \neq j \leq N$,

$$E(\omega_{ij}^{(N)} - \bar{\omega}_{ij}^{(N)})^2 = \frac{1-q}{q(N-1)^2},$$

we get

$$E \left(([\Omega_N(t)]_i - [\bar{\Omega}_N]_i) \left([\Omega_N(t)]_i - [\bar{\Omega}_N]_i \right)^T \right) = \frac{N(1-q)}{q(N-1)^2}.$$

Hence, by the independence of the i.i.d. random matrix $\Omega_N(\cdot)$ and the process $z^o(\cdot)$, we get

$$(\eta_i^N)^2 \leq \frac{kN(1-q)}{q(N-1)^2},$$

where

$$k := \sup_{N \geq 1} \max_{1 \leq i \leq N} E \int_0^\infty e^{-\rho t} (z_i^o(s))^2 ds < \infty,$$

is independent of N by (2.22). □

2.9. Game Consensus Model with Localized Cost-Couplings

In this model there is a finite number of groups (types) of homogeneous agents within a heterogeneous system where agents in each group assign nonuniform weights to agents in other groups in their cost functions. This is to take into account the possibility of locally related interactions (with possible spatial interpretation) between groups of agents within the population (see [80, 83]).

Consider a system of N agents with stochastic dynamics (2.1). The finite set $\Theta := \{\theta_1, \dots, \theta_K\}$ of distinct elements is used to model K groups (types) within the population [80, 83]. The i^{th} agent, $1 \leq i < \infty$, in the system is assigned with a type parameter l_i taking values from the finite set Θ which indicates the group that this agent belongs to. Let the intra-group coupling weight matrix be

$$\Omega := (\omega_{\theta_i \theta_j})_{K \times K} \quad (2.66)$$

where $\omega_{\theta_i \theta_j} \geq 0$ for any $\theta_i, \theta_j \in \Theta$ and $\sum_{j=1}^K \omega_{\theta_i \theta_j} \neq 0$ for each $\theta_i \in \Theta$. Denote

$$\sum_{i=1}^N 1_{\{l_i = \theta_k\}} = N_k, \quad 1 \leq k \leq K,$$

then we define the weight coefficients $\omega_{l_i l_j}^N$, $1 \leq i, j \leq N$, between agents as

$$\omega_{l_i l_j}^{(N)} = \begin{cases} 1/N_k & \text{for } l_i, l_j \in \theta_k, \\ \omega_{\theta_i \theta_j} / N_{k'} & \text{for } l_i \in \theta_k, l_j \in \theta_{k'}. \end{cases}$$

Let $\pi_k^N = N_k/N$ then $\pi^N := (\pi_1^N, \dots, \pi_K^N)$ is a probability vector which gives the empirical distribution of the system of agents with type parameters l_1, \dots, l_N . In a large population system, a natural way to model the sequence of type parameters l_1, \dots, l_N is to view it as being truncated from an infinite sequence $\{l_i, i \geq 1\}$ which exhibits certain statistical properties introduced in the following assumption (see [75]).

(A2.2) There exists a probability vector π such that

$$\lim_{N \rightarrow \infty} \pi^N = \pi := (\pi_1, \dots, \pi_K)$$

where $\min_{1 \leq k \leq K} \pi_k > 0$ (the probability vector π shows the relative frequency of each of the K groups).

For each $1 \leq i \leq N$, let

$$\phi_i^N(t) := \left(\sum_{j=1}^N \omega_{l_i l_j}^{(N)} z_j(t) \right) / \left(\sum_{j=1}^N \omega_{l_i l_j}^{(N)} \right), \quad t \geq 0, \quad (2.67)$$

be the normalized weighted mean corresponding to the i^{th} agent with type parameter l_i . The objective of the i^{th} agent ($1 \leq i \leq N$) is to minimize its locality dependent discounted cost function given by

$$J_{\text{loc},i}^N(u_i, u_{-i}) := E \int_0^\infty e^{-\rho t} \left((z_i(t) - \phi_i^N(t))^2 + r u_i^2(t) \right) dt, \quad (2.68)$$

where r and ρ are positive integers.

It is important to note that the model (2.1)-(2.68) is the multi-class extension of the dynamic game consensus model (2.1)-(2.2) with only one group (type) of agents.

Let the empirical distribution functions associated with N agents in K groups be defined by

$$F_k^N(x) := \frac{1}{N_k} \sum_{i=1}^N 1_{\{l_i=\theta_k: \bar{z}_i(0) < x\}}, \quad 1 \leq k \leq K$$

where $1_{\{l_i=\theta_k: \bar{z}_i(0) < x\}} = 1$ if agent i is of type k (i.e., $l_i = \theta_k$) and $\bar{z}_i(0) < x$, and $1_{\{\bar{z}_i(0) < x\}} = 0$ otherwise. We enunciate the following assumption which is the multi-class version of **(A2.1)**:

(A2.3) We assume that (i) the initial states $\{z_i(0) : 1 \leq i \leq N\}$ are independent, and there exists a constant C independent of N such that $\sup_{1 \leq i \leq N} E|z_i(0)|^2 \leq C < \infty$, and (ii) for each $1 \leq k \leq K$, $\{F_k^N : N \geq 1\}$ converges weakly to a Gaussian probability distribution F_k . \square

We now take a representative agent of type $\theta \in \Theta \equiv \{\theta_1, \dots, \theta_K\}$. The state process of this agent may be denoted by z_θ .

Let p be given as in Theorem 2.1-(a). For the infinite population we construct the following MF game (NCE) system (see [80, 83]):

$$\frac{ds_\theta(t)}{dt} = \left(\rho + \frac{p}{r} \right) s_\theta(t) + \phi_\theta^\infty(t), \quad \theta \in \Theta \quad (2.69)$$

$$\frac{d\bar{z}_\theta(t)}{dt} = -\frac{p}{r} \bar{z}_\theta(t) - \frac{1}{r} s_\theta(t), \quad \bar{z}_\theta(0) \quad (2.70)$$

$$\phi_\theta^\infty(\cdot) = \left(\sum_{\theta' \in \Theta} \pi_{\theta'} \omega_{\theta\theta'} \bar{z}_{\theta'}(\cdot) \right) / \left(\sum_{\theta' \in \Theta} \pi_{\theta'} \omega_{\theta\theta'} \right) \quad (2.71)$$

where $\bar{z}_\theta(0) := \int_{\mathbb{R}} \bar{z}(0) dF_k(\bar{z}(0))$ is the mean value of the agents' initial states of type $\theta \in \Theta$. System (2.69)-(2.71) is constructed such that the representative agent of type $\theta \in \Theta$ carries out optimal tracking of the local mass function $\phi_\theta^\infty(\cdot)$ which, in turn, depends on normalized locality related coupling as expressed in (2.71) in the continuum of agents.

We note that in the construction of individual strategies:

$$u_\theta^\infty(t) = -\frac{1}{r}(pz_\theta(t) + s_\theta(t)), \quad t \geq 0, \quad \theta \in \Theta. \quad (2.72)$$

each agent needs to know the weight matrix Ω , the probability vector π , and the distribution functions F_k , $1 \leq k \leq K$. But the agent is not required to know specific information on a particular neighbor, such as its type or its initial state.

Let $\bar{z} := (\bar{z}_{\theta_1}, \dots, \bar{z}_{\theta_K})^T$ and $s := (s_{\theta_1}, \dots, s_{\theta_K})^T$, then the system (2.69)-(2.71) may be written in vector form as:

$$\frac{ds(t)}{dt} = \frac{1}{p}s(t) + W\bar{z}(t), \quad (2.73)$$

$$\frac{d\bar{z}(t)}{dt} = -\frac{p}{r}\bar{z}(t) - \frac{1}{r}s(t), \quad \bar{z}(0) \text{ given}, \quad (2.74)$$

where $1/p = \rho + p/r$ by the Riccati equation (see Theorem 2.1-(a)), and $(W)_{ij} := (\pi_j \omega_{\theta_i, \theta_j}) / (\sum_{k=1}^K \pi_k \omega_{\theta_i, \theta_k})$ for $1 \leq i, j \leq K$. Matrix W is a row-stochastic matrix since all its row sums are 1.

We write the steady-state equations of the system (2.73)-(2.74) as

$$(1/p)s_\infty + W\bar{z}_\infty = 0, \quad (p/r)\bar{z}_\infty + (1/r)s_\infty = 0, \quad (2.75)$$

where the subscript ∞ indicated the steady-state solution.

DEFINITION 2.3. *A stochastic matrix A is irreducible if its corresponding digraph is strongly connected [119].*

THEOREM 2.12. *If W is irreducible then the unique stationary solution of the system (2.73)-(2.74) is*

$$(s_\infty, \bar{z}_\infty) = \left(-p \frac{\gamma^T \bar{z}(0)}{\gamma^T \mathbf{1}_K} \mathbf{1}_K, \frac{\gamma^T \bar{z}(0)}{\gamma^T \mathbf{1}_K} \mathbf{1}_K \right), \quad (2.76)$$

where γ^T is the unique left-hand Perron vector for W . Hence, agents reach mean-consensus on $\frac{\gamma^T \bar{z}(0)}{\gamma^T \mathbf{1}_K} \mathbf{1}_K$.

Proof: The algebraic equations in (2.75) give $W \bar{z}_\infty = \bar{z}_\infty$ which indicates that \bar{z}_∞ is the right eigenvector of matrix W corresponding to eigenvalue 1. Hence, we have $\bar{z}_\infty = \alpha \mathbf{1}_N$ for a nonzero constant α .

Since matrix W is irreducible there exists a unique Perron vector γ^T such that $\gamma^T W = \gamma^T$ (see [119]). Multiplying equations (2.73) and (2.74) by γ^T yields

$$\frac{d(\gamma^T s(t))}{dt} = \frac{1}{p} \gamma^T s(t) + \gamma^T \bar{z}(t), \quad (2.77)$$

$$\frac{d(\gamma^T \bar{z}(t))}{dt} = -\frac{p}{r} \gamma^T \bar{z}(t) - \frac{1}{r} \gamma^T s(t), \quad \gamma^T \bar{z}(0). \quad (2.78)$$

Now by Theorem 2.2 the unique bounded solution of system (2.77)-(2.78) is given by

$$(\gamma^T s(t), \gamma^T \bar{z}(t)) = (-p \gamma^T \bar{z}(0), \gamma^T \bar{z}(0)), \quad t \geq 0.$$

This together with $\bar{z}_\infty = \alpha \mathbf{1}_N$ gives

$$\gamma^T \bar{z}_\infty = \alpha \gamma^T \mathbf{1}_N = \gamma^T \bar{z}(0)$$

which determines α uniquely as $(\gamma^T \bar{z}(0)) / (\gamma^T \mathbf{1}_N)$. Hence we have $\bar{z}_\infty = \frac{\gamma^T \bar{z}(0)}{\gamma^T \mathbf{1}_K} \mathbf{1}_K$ and by (2.75) $s_\infty = -p \frac{\gamma^T \bar{z}(0)}{\gamma^T \mathbf{1}_K} \mathbf{1}_K$. \square

REMARK 2.10. *The existence and uniqueness of transient solution ($0 \leq t < \infty$) to the system (2.73)-(2.74) can be shown by a contraction mapping argument similar to Theorem 2 in [80].*

The proof of the following theorem is similar to the one of Theorem 7 in [83] with some modifications.

THEOREM 2.13. *Assume (A1.2) and (A1.3) hold. Then the set of MF control laws for the finite population system $\{u_i^o \in \mathcal{U}_i : 1 \leq i \leq N\}$ given in (2.72) generates an ϵ_N -Nash equilibrium such that*

$$J_{loc,i}^N(u_i^o, u_{-i}^o) - O(\epsilon^N) \leq \inf_{u_i \in \mathcal{U}_i} J_{loc,i}^N(u_i, u_{-i}^o) \leq J_{loc,i}^N(u_i^o, u_{-i}^o), \quad 1 \leq i \leq N,$$

where $\lim_{N \rightarrow \infty} \epsilon^N = 0$. □

2.10. Numerical Examples

EXAMPLE 2.1. *Consider a system of 500 agents with the MF game consensus closed-loop dynamics (2.21) where $r = 10$, $\rho = 0.2$ and $\sigma = 0.05$. The initial states of the agents are taken independently from a standard normal distribution, i.e., a Gaussian distribution with mean zero and variance one. The state trajectories of the agents are shown in Fig. 2.1A. Fig. 2.1B illustrates the histogram of the system at the final time $t = 20$. As shown in Fig. 2.1B the agents reach mean-consensus in the population's initial mean $\phi^\infty(0) = 0$. Fig. 2.2A illustrates*

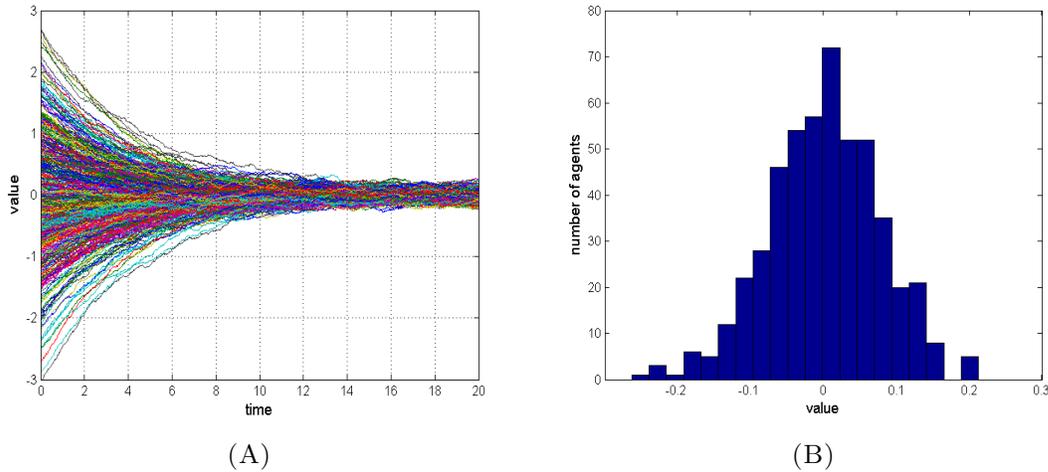


FIGURE 2.1. Example 2.1: (A) Trajectories of agents' states when $N = 500$, (B) Histogram of the system at time $t = 20$.

$$E\left(\frac{1}{N} \sum_{i=1}^N z_i^o(t) - \phi^\infty(0)\right)^2, \quad 0 \leq t \leq 20,$$

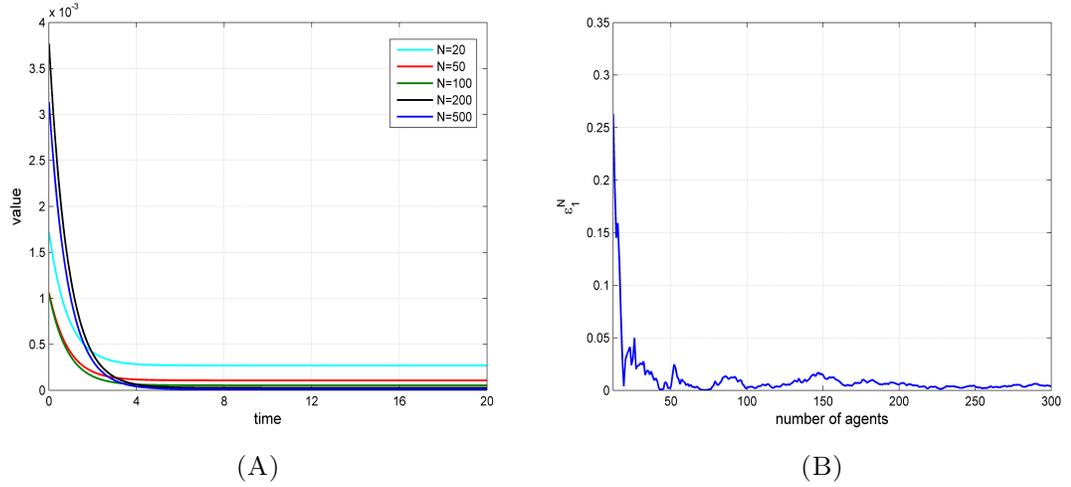


FIGURE 2.2. Example 2.1: (A) The convergence of the population mean; (B) Curve of ϵ_1^N with respect to N .

for population sizes $N = 20, 50, 100, 200$ and 500. As shown in Fig. 2.2A and based on Theorem 2.5, this value goes to zero as the population size N approaches infinity. The curve of ϵ_1^N with respect to N defined in (2.29) is shown in Fig. 2.2B which approaches zero as N goes to infinity.

EXAMPLE 2.2. Consider a system of 500 agents where $r = 10$ and $\rho = 0.2$. In Fig. 2.3 the state trajectories of the deterministic system (i.e., $\sigma = 0$) are shown. The agents in Fig. 2.3A apply the deterministic MF social control laws (2.49) and their closed-loop dynamics are given in (2.50). The agents in Fig. 2.3B apply the centralized control law (2.56). As it is shown in Fig. 2.3 the agents reach consensus asymptotically as time goes to infinity in both scenarios, even though the transient solutions of these strategies will in general be different. The social LQR cost value J_{soc}^N is equal to $1.1084e+003$, and the social cost of the MF control laws is equal to $1.1093e+003$. For the first agent J_1^∞ defined in (2.20) is equal to 4.2676; the first agent's individual costs with MF control (2.19) and centralized control (2.56) are 4.2743 and 4.2742, respectively.

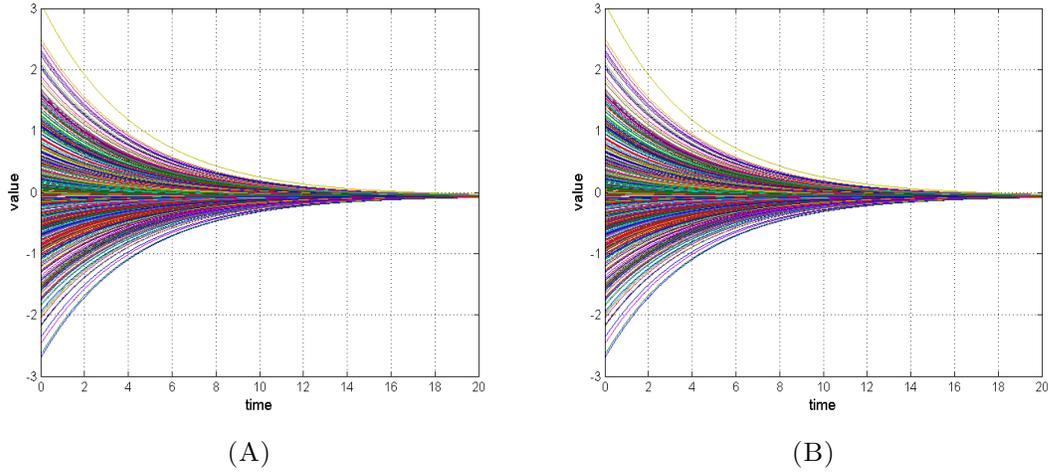


FIGURE 2.3. Example 2.2: Trajectories of agents' states in (A) deterministic MF social, and (B) centralized LQR cases.

EXAMPLE 2.3. Consider a system of 500 agents with localized cost-coupling dynamic game model (2.1)-(2.68) where $r = 10$, $\rho = 0.2$ and $\sigma = 0.1$. We assume 5 group of 100 agents within the population. Let the intra-group coupling weight matrix and its corresponding unique left-hand Perron vector be

$$\Omega = \begin{pmatrix} 0 & 0.2 & 0.4 & 0.3 & 0.1 \\ 0.1 & 0 & 0.5 & 0.3 & 0.1 \\ 0.2 & 0.2 & 0 & 0.1 & 0.5 \\ 0.15 & 0.15 & 0.2 & 0 & 0.5 \\ 0.3 & 0.2 & 0.4 & 0.1 & 0 \end{pmatrix}, \quad \gamma^T = \begin{pmatrix} 0.363 \\ 0.347 \\ 0.597 \\ 0.326 \\ 0.532 \end{pmatrix}^T.$$

Since the probability vector π is $(0.2, 0.2, 0.2, 0.2, 0.2)$ we have $W = \Omega$ which is a row-stochastic matrix. The initial states of the agents of each group are taking independently from Gaussian distributions with variance 0.5 and means -9, -5, 1, 6 and 10 (i.e., $\bar{z}(0) = [-9, -5, 1, 6, 10]^T$). By Theorem 2.12 the agents reach mean-consensus in $\alpha = (\gamma^T \bar{z}(0)) / (\gamma^T \mathbf{1}_N) = 1.3281$. The associated MF game (NCE) system is numerically solved with a time step size of 0.01. The decentralized control law is applied, and Fig. 2.4A shows the individual trajectories on the time interval $[0, 20]$. Fig. 2.4B

illustrates the histogram of the system at the final time $t = 20$. As shown in Fig. 2.4B the agents reach mean-consensus in $\alpha = 1.3281$.

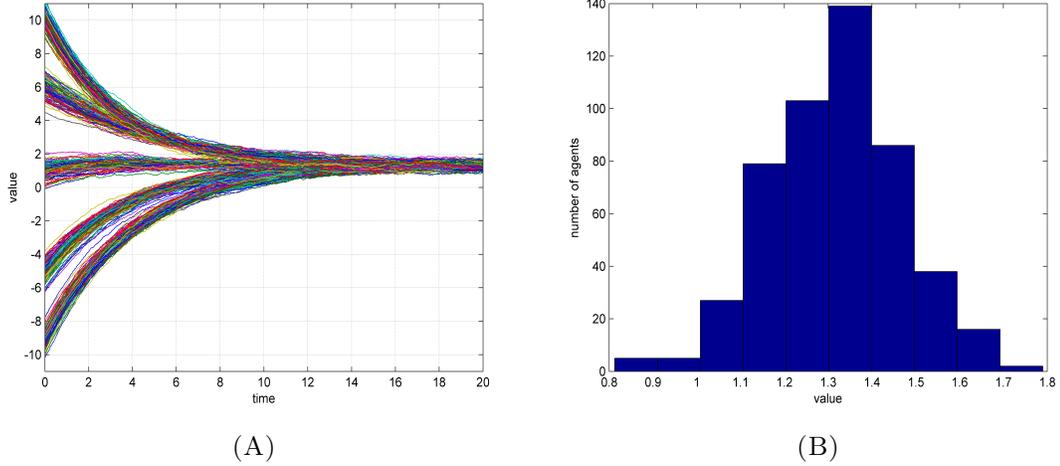


FIGURE 2.4. Example 2.3: (A) Trajectories of agents' states, (B) Histogram of the system at time $t = 20$.

EXAMPLE 2.4. We take the values of Example 2.3. But, the intra-group coupling weight matrix and its corresponding unique left-hand Perron vector are

$$\Omega = \begin{pmatrix} 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0.5 & 0.5 & 0 & 0 & 0 \end{pmatrix}, \quad \gamma^T = \begin{pmatrix} 0.105 \\ 0.315 \\ 0.157 \\ 0.210 \\ 0.210 \end{pmatrix}^T.$$

The non-negative stochastic matrix Ω is irreducible. Fig. 2.5A shows the individual trajectories on the time interval $[0, 20]$. Fig. 2.5B illustrates the histogram of the system at the final time $t = 20$. As shown in Fig. 2.5B the agents reach mean-consensus in $\alpha = 0.999$.

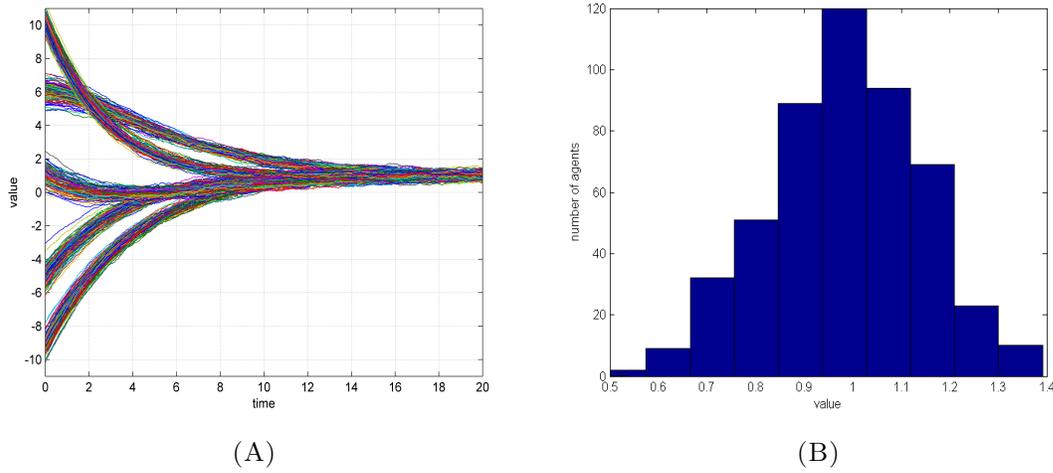


FIGURE 2.5. Example 2.4: (A) Trajectories of agents' states, (B) Histogram of the system at time $t = 20$.

EXAMPLE 2.5. We take the values of Example 2.3. But, the intra-group coupling weight matrix is

$$\Omega = \begin{pmatrix} 0 & 0.6 & 0.4 & 0 & 0 \\ 0.7 & 0 & 0.3 & 0 & 0 \\ 0.6 & 0.4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Ω correspond to an adjacency matrix of a graph where γ_1^T and γ_2^T are left-hand eigenvectors of the eigenvalue one:

$$\gamma_1^T = \begin{pmatrix} 0.677 \\ 0.584 \\ 0.446 \\ 0 \\ 0 \end{pmatrix}^T, \quad \gamma_2^T = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}^T.$$

As shown in Figs. 2.6A and 2.6B each associated sub-population converges to the initial distribution mean of its connected component.

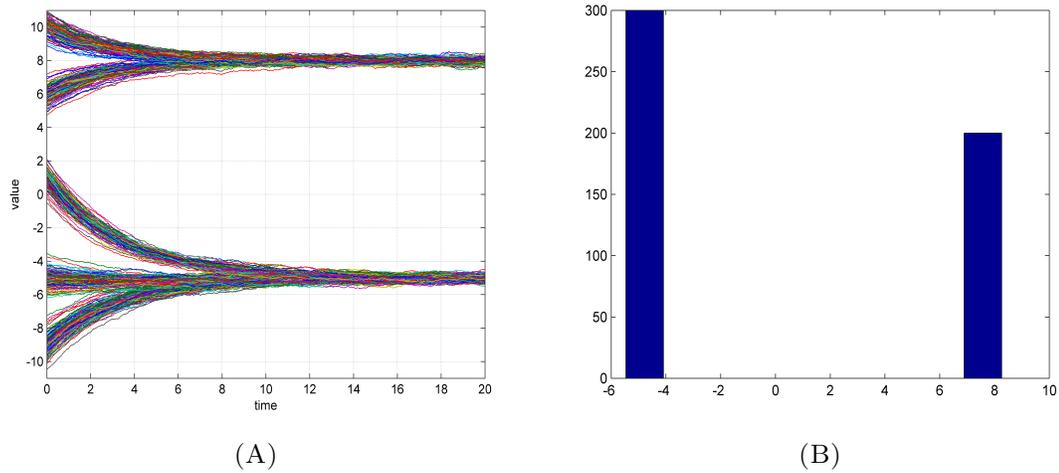


FIGURE 2.6. Example 2.5: (A) Trajectories of agents' states, (B) Histogram of the system at time $t = 20$.

2.11. Chapter Summary

In the Mean Field (MF) consensus model: (i) each agent has a priori information on the initial state distribution mean of the overall population, (ii) the system of agents achieves mean-consensus without requiring communication with other agents. Whereas in the Standard Consensus (SC) algorithms: (i) agents need no a priori information on the initial state distribution of the overall population but require local communication with other agents, (ii) consensus can be achieved if the union of the interaction graphs for the system is connected frequently enough as the system evolves.

The SC algorithms require communication with other agents in the system and for large N this leads to high communication and computational complexity. On the other hand, the decentralized MF control laws do not require even local communication but need a priori information on the mean of the system's initial state distribution.

The uniform weight cost-coupling MF game model is extended to the case of agents with non-uniform mean field cost-couplings which corresponds to a heterogeneous system with homogeneous sub-populations. In the localized model with connected topology specified by the cost-coupling weight matrix, the unique stationary equilibrium yields consensus in the weighted average of initial states (which depends on the left eigenvector of the weight matrix corresponding to the unique eigenvalue 1). Let the system cost-coupling weight matrix correspond to an adjacency matrix of a graph with more than one connected component, then each associated sub-population can only converge to the initial distribution mean of its connected component. Correspondingly, in the deterministic SC problem, one of the key hypotheses which is used to establish the convergence to consensus is that the system graph is connected.

2.12. Appendix

Proof of Theorem 2.4: For a generic agent i , $1 \leq i \leq N$, we have the closed-loop solution (2.21). But,

$$E \int_0^\infty e^{-\rho t} (\phi^\infty(0))^2 dt = (\phi^\infty(0))^2 / \rho, \quad (2.79)$$

$$E \int_0^\infty e^{-\rho t} (e^{-(p/r)t} (z_i(0) - \phi^\infty(0)))^2 dt = (\bar{z}_i(0) - \phi^\infty(0))^2 / (2p/r + \rho), \quad (2.80)$$

and by the Itô isometry

$$E \int_0^\infty e^{-\rho t} \left(\int_0^t e^{-(p/r)(t-\tau)} dw_i(\tau) \right)^2 dt = \frac{r}{2p} (1/\rho - 1/(2p/r + \rho)). \quad (2.81)$$

Now by (2.21) and (2.79)-(2.81) we get

$$E \int_0^\infty e^{-\rho t} (z_i^o(t))^2 dt := k < \infty, \quad (2.82)$$

since $\sup_{\bar{z}_i(0) \in \Theta} |\bar{z}_i(0)| < \infty$ by the compactness of the set Θ . This together with (2.19) results in

$$E \int_0^\infty e^{-\rho t} (u_i^o(t))^2 dt \leq \frac{p}{r} (k + \phi^\infty(0))^2 < \infty. \quad (2.83)$$

Since k is independent of i and N , by (2.82) and (2.83) we obtain (2.22). \square

Proof of Theorem 2.6: a) By (2.25) and since the set of expected values of initial states Θ is a compact set, there exists a finite k independent of N and t such that

$$E\left(\frac{1}{N}\sum_{i=1}^N z_i^o(t) - \phi^\infty(0)\right)^2 < k.$$

Hence, Theorem 2.5 and the *Lebesgue Dominated Convergence theorem* (see [42]) imply (2.26).

b) See the proof of Lemma 2.1 below. \square

Proof of Lemma 2.1: For a fixed $1 \leq i \leq N$ by (2.21) we get

$$\begin{aligned} (\epsilon_i^N)^2 = E \int_0^\infty e^{-\rho t} & \left(e^{-(p/r)t} \left(\frac{1}{N-1} \sum_{j=1, j \neq i}^N z_j(0) - \phi^\infty(0) \right) \right. \\ & \left. + \frac{\sigma}{N-1} \sum_{j=1, j \neq i}^N \int_0^t e^{-(p/r)(t-\tau)} dw_j(\tau) \right)^2 dt. \end{aligned}$$

Then by the independence of initial states and Wiener processes, and the Itô isometry we have

$$\begin{aligned} (\epsilon_i^N)^2 = \frac{1}{\rho + 2(p/r)} & \left(\frac{1}{N-1} \sum_{j=1, j \neq i}^N \bar{z}_j(0) - \phi^\infty(0) \right)^2 \\ & + \frac{\sigma^2 r}{2p(N-1)} \left(\frac{1}{\rho} - \frac{1}{\rho + 2(p/r)} \right). \end{aligned}$$

Hence, $\lim_{N \rightarrow \infty} \epsilon_i^N = 0$ by (A2.1). \square

CHAPTER 3

A Mean Field Game Synthesis of Initial Mean Consensus Problems: A Continuum Approach for Non-Gaussian Behaviour

This chapter presents a continuum approach to a non-Gaussian initial mean consensus problem studied in previous chapter. For this dynamic game problem, a set of coupled deterministic (Hamilton-Jacobi-Bellman (HJB) and Fokker-Planck-Kolmogorov (FPK)) equations is derived approximating the stochastic system of agents as the population size goes to infinity. In a finite population system (analogous to the mean field linear-quadratic-Gaussian (MF LQG) framework in previous chapter): (i) the resulting decentralized mean field (MF) control strategies possess an ϵ_N -Nash equilibrium property where ϵ_N goes to zero as the population size N approaches infinity, and (ii) these MF control strategies steer each individual's state toward the initial state population mean which is reached asymptotically as time goes to infinity. Hence, the system with decentralized MF control strategies reaches mean-consensus on the initial state population mean asymptotically as time and population size go to infinity.

In the case of agents with Long Time Average (LTA) (i.e., ergodic) cost functions the solution of the HJB equation is the relative value function which represents perturbations around the steady-state optimal cost rate with respect to an asymptotically stationary process. It turns out that this HJB equation in the MF system

has a larger class of stable perturbed solutions in forward time than in backward time. Therefore, an Evolution (i.e., forward in time) Mean Field (EMF) system of consensus model is studied. The EMF system consists of two coupled (forward in time) deterministic PDEs which are also coupled to the cost coupling function.

3.1. Introduction

There are two main classes of models relevant to the study of consensus behaviour: (i) Individual based (Lagrangian) models in the form of coupled Ordinary (Stochastic) Differential Equations (O(S)DEs) (see for example [169]). A key element of many individual based algorithms is the use of local feedback involving local communication (subject to the network topology) between agents so as to reach an agreement. (ii) Continuum based (Eulerian) models in the form of Partial Differential Equations (PDEs) in large population systems (see [40], among many other papers).

In Chapter 2 we synthesized consensus behaviour as a dynamic game problem via mean field linear-quadratic-Gaussian (MF LQG) control theory. In this chapter we develop another approach to the study of consensus problems. Unlike the dynamic game consensus formulation of previous chapter (see Section 2.2.1) with Gaussian initial states (see (A2.1) in Chapter 2), the initial states for all the agents of the model in this chapter are not necessarily assumed to be distributed according to a Gaussian distribution, and so the MF LQG framework of Chapter 2 cannot be employed. Consequently, for the infinite population limit a general continuum (i.e., PDE) formulation is required.

The resulting continuum based mean field (MF) system of the DGCM consists of two coupled deterministic equations: (i) a nonlinear (backward in time) Hamilton-Jacobi-Bellman (HJB), and (ii) a nonlinear (forward in time) Fokker-Planck-Kolmogorov (FPK), which are also coupled to a (spatially averaged) cost coupling function approximating the aggregate effect of the agents in the infinite population limit. We present the stationary solutions, linear stability and the nonlinear stability analyses of the continuum MF system. Analogous to the MF LQG framework, we show

(i) the ϵ_N -Nash equilibrium property of the resulting MF control laws, and (ii) the mean-consensus behaviour of the system by applying these MF control laws. It is also important to note that in the non-Gaussian consensus problem the stationary solution of the system is itself Gaussian.

In the case of agents with Long Time Average (LTA) (i.e., ergodic) cost functions the solution of the HJB equation is the relative value function which represents perturbations around the steady-state optimal cost rate with respect to an asymptotically stationary process. It turns out that this HJB equation in the MF system has a larger class of stable perturbed solutions in forward time than in backward time [64]. Therefore, an Evolution (i.e., forward in time) Mean Field (EMF) system of consensus model is studied where the initial states for all the agents are not necessarily assumed to be distributed according to a Gaussian distribution. The EMF system consists of two coupled (forward in time) deterministic PDEs which are also coupled to the cost coupling function. The forward in time mean field process has previously appeared in the study of MFG models in [4, 64].

The following notation will be used in this chapter. We use the integer valued subscript as the label for an individual agent of the population. The integer N is reserved to denote the population size of the system. We use the superscripts N for a process to indicate the dependence on the population size. The symbols ∂_t and ∂_z respectively denote the partial derivative with respect to variables t and z , and ∂_{zz}^2 denotes the second derivative with respect to z .

The chapter is organized as follows. Section 3.2 is dedicated to the problem formulation and the applications of the model. The continuum based MF control approach to the consensus model is presented in Section 3.3. In Section 3.4 we study the stationary solution and the stability analysis of the MF system. The mean-consensus and ϵ -Nash equilibrium properties of the MF control laws are respectively established in Sections 3.5 and 3.6. Section 3.7 studies the MF and EMF systems for agents with LTA cost functions. A sample numerical simulation of the model is

presented illustrating the results in Section 3.8. Concluding remarks are stated in Section 3.9.

3.2. Dynamic Game Consensus Model with Non-Gaussian Initial States

We recall the dynamic game consensus model (2.1)-(2.2) of previous chapter (see Section 2.2.1). Consider a system of N agents. The dynamics of the i^{th} agent is given by a controlled SDE:

$$dz_i(t) = u_i(t)dt + \sigma dw_i(t), \quad t \geq 0, \quad 1 \leq i \leq N, \quad (3.1)$$

where consistent with the standard consensus models $z_i(\cdot)$ and $u_i(\cdot)$ are the scalar state and control input of agent i ; σ is a non-negative scalar; and $\{w_i : 1 \leq i \leq N\}$ denotes a sequence of independent standard scalar Wiener processes on some filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ where \mathcal{F}_t is defined as the σ -field $\sigma(z_i(0), w_i(\tau) : 1 \leq i \leq N, \tau < t)$. We assume that the initial states $\{z_i(0) : 1 \leq i \leq N\}$ are independent, and independent of $\{w_i : 1 \leq i \leq N\}$.

It is important to note that unlike the model in Chapter 2 (see **A2.1**), the initial states for all the agents are not necessarily assumed to be distributed according to a Gaussian distribution.

In this problem formulation each agent seeks to minimize its individual cost function involving the mean of the states of all other agents. Let the cost-coupling function be defined homogeneously throughout the population by $c^N(z_i; z_{-i}) := (z_i - \frac{1}{N-1} \sum_{j \neq i}^N z_j)^2$ for a generic agent i . Then, the objective of each individual agent i is to minimize its discounted cost function given by

$$J_i^N(u_i, u_{-i}) := E \int_0^\infty e^{-\rho t} (c^N(z_i; z_{-i}) + r u_i^2(t)) dt, \quad 1 \leq i \leq N, \quad (3.2)$$

where $r > 0$ is the control penalty and $\rho > 0$ is the discount factor. To indicate the dependence of J_i on u_i , u_{-i} and the population size N , we write it as $J_i^N(u_i, u_{-i})$.

For minimization of J_i^N , the admissible control set is taken as

$$\mathcal{U}_i := \{u_i(\cdot) : u_i(t) \text{ is adapted to sigma-field } \sigma(z_j(s) : s \leq t, 1 \leq j \leq N)\}.$$

The reader is referred to Section 3.7 for the case of agents with Long Time Average (LTA) (i.e., ergodic) cost functions (2.35) (see Chapter 2, Section 2.5.3).

REMARK 3.1. *The results of this chapter can easily be extended to the case of agents with cost-coupling functions of the form $c^N(z_i; z_{-i}) := \frac{1}{N-1} \sum_{j \neq i}^N (z_i - z_j)^2$ (see Remark 2.1 and Section 2.3).*

The Dynamic Game Consensus Model (DGCM) (3.1)-(3.2) is motivated by many social, economic, and engineering applications. The reader is referred to Chapter 2, Section 2.3 for a synchronization of coupled oscillators example, and to Chapter 5, Section 5.2.4 for an economic (finance) example.

3.3. A Continuum Mean Field Game Approach

We take the following steps to the DGCM (3.1)-(3.2) based on the nonlinear MF control approach developed in [78, 85, 177]:

- (i) *The infinite population limit:* In this step a Nash equilibrium is characterized by an “equilibrium relationship” between the individual strategies and the mass effect (i.e., the overall effect of the population on a given agent). This equilibrium relationship is described by the so-called MF system.
- (ii) *ϵ_N -Nash equilibrium for the finite N model:* The distributed continuum based MF control laws (derived from the MF system in Step 1) establish an ϵ_N -Nash equilibrium for the finite N population DGCM where ϵ_N goes to zero asymptotically as N approaches infinity.

3.3.1. Mean Field Approximation. In a large N population system, by the Law of Large Numbers (LLN) we approximate the cost-coupling function for a “generic” agent i in (3.2), $c^N(z_i(\cdot), z_{-i}(\cdot))$, by a deterministic function $c(z, \cdot)$ which only depends on $z = z_i$. Replacing the function $c^N(z_i, z_{-i})$ with the deterministic

function $c(z_i, \cdot)$ in the i^{th} agent's cost function (3.2) reduces the DGCM (3.1)-(3.2) to a set of N independent optimal control problems.

3.3.1.1. *Hamilton-Jacobi-Bellman (HJB) Equation.* We now consider a “single agent” Optimal Control Problem (OCP):

$$dz(t) = u(t)dt + \sigma dw(t), \quad t \geq 0, \quad (3.3)$$

$$\inf_{u \in \mathcal{U}} J(u) := \inf_{u \in \mathcal{U}} E \int_0^\infty e^{-\rho t} (c(z, t) + ru^2(t)) dt, \quad (3.4)$$

where $z(\cdot), u(\cdot) \in \mathbb{R}$ are the state and control input, respectively; $w(\cdot)$ denotes a standard scalar Wiener process; $c(z, \cdot)$ is a positive function; and \mathcal{U} is the corresponding admissible control set of the generic agent. An admissible control $u^\circ(\cdot) \in \mathcal{U}$ is called optimal if $J(u^\circ) = \inf_{u \in \mathcal{U}} J(u)$.

For $x \in \mathbb{R}$ and $0 \leq t < \infty$ we define the value function $v(\cdot, \cdot)$ for the OCP (3.3)-(3.4) by

$$v(x, t) := \inf_{(u(s))_{s \geq t} \in \mathcal{U}} E \left[\int_t^\infty e^{-\rho(s-t)} (c(z(s), s) + r(u(s))^2) ds \mid z(t) = x \right].$$

By employing a standard dynamic programming argument and using Itô's formula we get the following result (see [59]).

THEOREM 3.1. (*HJB for the OCP*) *Assume that the function $c(z, t)$ is Lipschitz continuous with respect to z and uniformly continuous with respect to t , and assume the value function $v(z, t)$ for the OCP (3.3)-(3.4) is a C^1 function of variable t and C^2 function of variable z , then $v(z, t)$ solves the (backward in time) Hamilton-Jacobi-Bellman (HJB) equation*

$$\partial_t v(z, t) + H(\partial_z v(z, t)) + \frac{\sigma^2}{2} \partial_{zz}^2 v(z, t) + c(z, t) = \rho v(z, t), \quad z \in \mathbb{R}, \quad t \geq 0, \quad (3.5)$$

with boundary condition $\lim_{t \rightarrow \infty} e^{-\rho t} v(z(t), t) = 0$, where the Hamiltonian $H(\cdot)$ is defined as $H(p) \equiv \min_{u \in \mathcal{U}} H(p, u) := \min_{u \in \mathcal{U}} \{up + ru^2\}$ for p in \mathbb{R} . \square

The solution of the OCP (3.3)-(3.4) is

$$u^o(z, t) := \arg \min_{u \in \mathcal{U}} H(\partial_z v(z, t), u) = -\frac{1}{2r} \partial_z v(z, t).$$

Substituting $u^o(z, t)$ into (3.5) yields the (backward in time) HJB equation ($z \in \mathbb{R}$, $t \geq 0$):=

$$\partial_t v(z, t) - \frac{1}{4r} (\partial_z v(z, t))^2 + \frac{\sigma^2}{2} \partial_{zz}^2 v(z, t) + c(z, t) = \rho v(z, t), \quad (3.6)$$

with boundary condition $\lim_{t \rightarrow \infty} e^{-\rho t} v(z(t), t) = 0$.

3.3.1.2. Fokker-Planck-Kolmogorov (FPK) Equation. Under the state feedback optimal control law $u^o(z, t) = -\frac{1}{2r} \partial_z v(z, t) \in C^1$, the evolution of the density $f(z, \cdot)$ of the generic agent (3.3) satisfies the (forward in time) Fokker-Planck-Kolmogorov (FPK) equation

$$\partial_t f(z, t) - \frac{1}{2r} \partial_z \left((\partial_z v(z, t)) f(z, t) \right) = \frac{\sigma^2}{2} \partial_{zz}^2 f(z, t), \quad z \in \mathbb{R}, t \geq 0, \quad (3.7)$$

with initial condition $f(z, 0) \geq 0$. We note that $v(z, \cdot)$ in (3.7) is the solution of the HJB equation (3.6). Let us assume the boundary condition $\lim_{|z| \rightarrow \infty} f(z, t) = 0$ for all $t \geq 0$.

3.3.1.3. Cost-Coupling (CC) Function. For a generic agent i , the Law of Large Numbers (LLN) suggests the approximation of the Cost-Coupling (CC) function $c^N(z_i, z_{-i}^o)$ for a large N population system by

$$\bar{c}(z_i, t) = \left(z_i - \int_{\mathbb{R}} z f(z, t) dz \right)^2 = \left(\int_{\mathbb{R}} (z_i - z) f(z, t) dz \right)^2, \quad z_i \in \mathbb{R}, t \geq 0. \quad (3.8)$$

where $f(z, \cdot)$ is the solution of the FPK equation (3.7).

3.3.2. The Mean Field System. Let

$$f_N(x, 0) := \frac{1}{N} \sum_{i=1}^N \delta(x - z_i(0)),$$

be the *empirical probability density* associated with N agents where δ is the Dirac delta function. We enunciate the following assumption (see **(A2.1)**):

(A3.1) We assume that (i) the initial states $\{z_i(0) : 1 \leq i \leq N\}$ are independent, and there exists a constant k independent of N such that $\sup_{1 \leq i \leq N} E|z_i(0)|^2 \leq k < \infty$, and (ii) $\{f_N(x, 0) : N \geq 1\}$ converges weakly to f_0 , i.e., for any bounded continuous function $\phi(x)$ on \mathbb{R} , we have $\lim_{N \rightarrow \infty} \int \phi(x) f_N(x, 0) dx = \int \phi(x) f_0(x) dx$. \square

REMARK 3.2. *It is important to note that if the sequence $\{z_i(0) : 1 \leq i \leq N\}$ is generated by independent random observations on the density function f_0 , then (A3.1)-(ii) holds with probability one by the Strong Law of Large Numbers or the Glivenko-Cantelli theorem [42].*

We now aim to construct the equilibrium relationship (between the individual strategies and the mass effect) in the stochastic MF control theory. For nonlinear MF stochastic control problems, a general formulation using equations of the McKean-Vlasov type is given in [85]. However, for the synchronization of coupled oscillators formulated as a game problem a compact system of coupled MF equations is given in [177] within a nonlinear SDE problem formulation.

The key idea of the MF control methodology is to prescribe a spatially averaged mass function $\bar{c}(z, \cdot)$ characterized by the property that it is reproduced as the average of all agents' states in the continuum of agents whenever each individual agent optimally tracks the same mass function $\bar{c}(z, \cdot)$.

Applying the nonlinear MF stochastic control approach (developed in [78, 85, 177]) to our DGCM (3.1)-(3.2) in the infinite population limit (or (3.3)-(3.4) for a generic agent) yields the nonlinear *continuum based MF system* ($z \in \mathbb{R}$, $t \geq 0$):

$$[\text{MF-HJB}] \quad \partial_t v(z, t) = \frac{1}{4r} (\partial_z v(z, t))^2 - \bar{c}(z, t) + \rho v(z, t) - \frac{\sigma^2}{2} \partial_{zz}^2 v(z, t), \quad (3.9)$$

$$[\text{MF-FPK}] \quad \partial_t f(z, t) = \frac{1}{2r} \partial_z \left((\partial_z v(z, t)) f(z, t) \right) + \frac{\sigma^2}{2} \partial_{zz}^2 f(z, t), \quad (3.10)$$

$$[\text{MF-CC}] \quad \bar{c}(z, t) = \left(\int_{\mathbb{R}} (z - z') f(z', t) dz' \right)^2, \quad (3.11)$$

where $f(z, 0) = f_0(z)$ is the given initial population density, necessarily $\int_{\mathbb{R}} f(z, t) dz = 1$ for any $t \geq 0$, and it is assumed that

$$\lim_{|z| \rightarrow \infty} f(z, t) = 0, \quad \lim_{t \rightarrow \infty} e^{-\rho t} v(z(t), t) = 0.$$

We refer the reader to the corresponding MF system (2.11)-(2.13) of Chapter 2 for the case of Gaussian initial states.

The system of equations (3.9)-(3.11) consists of: (i) the nonlinear (backward in time) *MF-HJB* equation (3.6) which describes the HJB equation of a generic agent's discounted optimal problem (3.3)-(3.4) with cost coupling $\bar{c}(z, \cdot)$, (ii) the nonlinear (forward in time) *MF-FPK* equation (3.7) which describes the evolution of the population density with the best response state feedback control law

$$u^o(z, t) := -\frac{1}{2r} \partial_z v(z, t), \quad z \in \mathbb{R}, t \geq 0, \quad (3.12)$$

and (iii) the spatially averaged *MF-CC* function presented in (3.8).

3.4. Analysis of the Mean Field System

3.4.1. Gaussian Stationary Solution. The MF system (3.9)-(3.11) in stationary form is

$$\frac{1}{4r} (\partial_z v_\infty(z))^2 - \frac{\sigma^2}{2} \partial_{zz}^2 v_\infty(z) = \bar{c}_\infty(z) - \rho v_\infty(z), \quad z \in \mathbb{R}, \quad (3.13)$$

$$\frac{1}{2r} \partial_z \left((\partial_z v_\infty(z)) f_\infty(z) \right) = -\frac{\sigma^2}{2} \partial_{zz}^2 f_\infty(z), \quad z \in \mathbb{R}, \quad (3.14)$$

$$\bar{c}_\infty(z) = \left(\int_{\mathbb{R}} (z - z') f_\infty(z') dz' \right)^2, \quad z \in \mathbb{R}, \quad (3.15)$$

where the density $f_\infty(z)$ satisfies $\int_{\mathbb{R}} f_\infty(z) dz = 1$, and $\lim_{t \rightarrow \infty} e^{-\rho t} v_\infty(z(t)) = 0$.

The stationary equation system (3.13)-(3.15) is related to (3.9)-(3.11) by the fact that the steady-state population density of the system, $f_\infty(z) := \lim_{t \rightarrow \infty} f(z, t)$, gives a time independent cost-coupling $\bar{c}_\infty(z)$ in (3.11) which yields a time independent solution $v_\infty(z)$ to the MF-HJB equation (3.9). Furthermore, $f_\infty(z)$ and $v_\infty(z)$ solve the stationary MF-FPK equation (3.14).

THEOREM 3.2. *For any arbitrary $\mu \in \mathbb{R}$, there exists the following solution to the system (3.13)-(3.15):*

$$v_\infty(z) = \gamma(z - \mu)^2 + \eta, \text{ where } \gamma := \frac{-r\rho + \sqrt{(r\rho)^2 + 4r}}{2} > 0, \quad \eta := \frac{\sigma^2\gamma}{\rho}, \quad (3.16)$$

$$f_\infty(z) = \frac{1}{\sqrt{2\pi s^2}} \exp\left(-\frac{(z - \mu)^2}{2s^2}\right), \quad \bar{c}_\infty(z) = (z - \mu)^2, \text{ where } s^2 := \frac{\sigma^2 r}{2\gamma}. \quad (3.17)$$

Proof. The assertion of the theorem is straightforward to verify by substituting (3.16)-(3.17) into the system (3.13)-(3.15). \square

Let us note that the steady-state solution of the system $f_\infty(z)$ is a Gaussian density function. We further note that in the class of stable solutions to the MF system (3.9)-(3.11), μ in Theorem 3.2 is uniquely determined as the initial state population mean $\int_{\mathbb{R}} z f(z, 0) dz$ (see Section 3.5).

3.4.2. Stability Analysis of the Linearized Mean Field System.

In this subsection by taking the approach of [64] we study the stability of the Gaussian steady-state density function $f_\infty(z)$ based on small perturbations of the form $f_\epsilon(z, 0) = f_\infty(z)(1 + \epsilon\tilde{f}(z, 0))$ on $f_\infty(z)$ such that $f_\epsilon(z, 0)$ is a probability density. We let the subsequent additive and multiplicative perturbations of the solution to the MF system (3.9)-(3.11) be of the forms

$$v_\epsilon(z, t) = v_\infty(z) + \epsilon\tilde{v}(z, t), \quad f_\epsilon(z, t) = f_\infty(z)(1 + \epsilon\tilde{f}(z, t)), \quad (3.18)$$

$$\bar{c}_\epsilon(z, t) = \bar{c}_\infty(z) + \epsilon\tilde{c}(z, t), \quad (3.19)$$

for $z \in \mathbb{R}$ and $t \geq 0$, where $v_\infty(z)$, $f_\infty(z)$ and $\bar{c}_\infty(z)$ are given in (3.16)-(3.17). Since $f_\epsilon(z, \cdot)$ satisfies the MF-FPK equation (3.10), $f_\epsilon(z, t)$ for any time $t > 0$ is necessarily a probability density, i.e., $f_\epsilon(z, t) = f_\infty(z)(1 + \epsilon\tilde{f}(z, t)) \geq 0$ and $\int_{\mathbb{R}} f_\epsilon(z, t) dz = 1$.

REMARK 3.3. *The reason why we take the relative perturbation form of the density function $f_\epsilon(z, \cdot)$ in (3.18) is that it permits us to employ the Hermite series expansion for the resulting linearized equation system.*

PROPOSITION 3.1. *The linearization of the nonlinear MF system (3.9)-(3.11) around the stationary solution (3.16)-(3.17) takes the form ($z \in \mathbb{R}$, $t \geq 0$):*

$$\partial_t \tilde{v}(z, t) = \frac{\gamma}{r}(z - \mu) \partial_z \tilde{v}(z, t) - \frac{\sigma^2}{2} \partial_{zz}^2 \tilde{v}(z, t) + \rho \tilde{v}(z, t) - \tilde{c}(z, t), \quad (3.20)$$

$$\begin{aligned} \partial_t \tilde{f}(z, t) = & -\left(\frac{\gamma}{r}(z - \mu) \partial_z \tilde{f}(z, t) - \frac{\sigma^2}{2} \partial_{zz}^2 \tilde{f}(z, t)\right) \\ & - \frac{1}{\sigma^2 r} \left(\frac{\gamma}{r}(z - \mu) \partial_z \tilde{v}(z, t) - \frac{\sigma^2}{2} \partial_{zz}^2 \tilde{v}(z, t)\right), \end{aligned} \quad (3.21)$$

$$\tilde{c}(z, t) = -2(z - \mu) \left(\int_{\mathbb{R}} z \tilde{f}(z, t) f_{\infty}(z) dz \right), \quad (3.22)$$

with given $\tilde{f}(z, 0)$, and boundary conditions

$$\lim_{|z| \rightarrow \infty} \tilde{f}(z, t) = 0, \quad \lim_{t \rightarrow \infty} e^{-\rho t} \tilde{v}(z(t), t) = 0.$$

Proof. See the appendix. □

For the analysis of the linearized equation system (3.20)-(3.22) we introduce the Hermite polynomials associated to the Hilbert space $L^2(f_{\infty}(z)dz)$. In this space we have the inner product $\langle g, h \rangle := \int_{\mathbb{R}} g(z)h(z)f_{\infty}(z)dz$ and the norm is given by $\|g\|_{L^2} := \langle g, g \rangle^{1/2}$.

DEFINITION 3.1. *For given μ and $s^2 = \frac{\sigma^2 r}{2\gamma}$ in (3.16)-(3.17) we define the normalized Hermite polynomials of the space $L^2(f_{\infty}(z)dz)$ as (see [64] or Chapter 22 in [1])*

$$H_n(z) := \frac{(-1)^n s^n}{\sqrt{n!}} \exp\left(\frac{(z - \mu)^2}{2s^2}\right) \frac{d^n}{dz^n} \exp\left(-\frac{(z - \mu)^2}{2s^2}\right), \quad z \in \mathbb{R}, \quad n \in \mathbb{N}_0.$$

In particular, we have $H_0(z) = 1$ and $H_1(z) = \frac{z - \mu}{s}$, $z \in \mathbb{R}$. □

PROPOSITION 3.2. *We have the following:*

- (a) *The countable family of Hermite polynomials $\{H_n(z) : n \in \mathbb{N}_0\}$ forms an orthogonal basis of the Hilbert space $L^2(f_{\infty}(z)dz)$ such that $\langle H_m, H_n \rangle = \delta(n, m)$ where δ is the Kronecker delta.*

(b) The Hermite polynomials $\{H_n(z) : n \in \mathbb{N}_0\}$ are eigenfunctions of the operator

$$\mathcal{L}g(z) := \frac{\gamma}{r}(z - \mu)\partial_z g(z) - \frac{\sigma^2}{2}\partial_{zz}^2 g(z), \quad z \in \mathbb{R}, \quad (3.23)$$

such that $\mathcal{L}H_n(z) = n\frac{\gamma}{r}H_n(z)$ for any $n \in \mathbb{N}_0$.

Proof. See the appendix. \square

By using the operator \mathcal{L} defined in (3.23) we can rewrite the equations (3.20)-(3.21) as

$$\partial_t \tilde{v}(z, t) = \mathcal{L}\tilde{v}(z, t) + \rho\tilde{v}(z, t) - \tilde{c}(z, t), \quad z \in \mathbb{R}, \quad t \geq 0, \quad (3.24)$$

$$\partial_t \tilde{f}(z, t) = -\frac{1}{\sigma^2 r}\mathcal{L}\tilde{v}(z, t) - \mathcal{L}\tilde{f}(z, t), \quad z \in \mathbb{R}, \quad t \geq 0, \quad (3.25)$$

where $\tilde{c}(z, \cdot)$ is given in (3.22), and the boundary conditions on \tilde{v} and \tilde{f} are those in Proposition 3.1.

DEFINITION 3.2. A steady-state solution $f_\infty(z)$ of the nonlinear MF system (3.9)-(3.11) is linearly asymptotically stable with respect to a set of initial perturbations

$$S = \left\{ f_\epsilon(z, 0) = f_\infty(z)(1 + \epsilon\tilde{f}(z, 0)) : f_\epsilon(z, 0) \geq 0, \int_{\mathbb{R}} f_\epsilon(z, 0) dz = 1, \right. \\ \left. \tilde{f}(z, 0) \in L^2(f_\infty(z) dz) \right\}$$

if there exists a solution $\tilde{f}(z, \cdot) \in L^2(f_\infty(z) dz)$ to the linearized equation system (3.20)-(3.22) such that $\lim_{t \rightarrow \infty} \|\tilde{f}(z, t)\|_{L^2} = 0$. \square

We define the set of ϵ perturbed initial density functions

$$S_{\text{per}}^{\text{Lin}}(\epsilon) := \left\{ f_\epsilon(z, 0) = f_\infty(z)(1 + \epsilon\tilde{f}(z, 0)) : \tilde{f}(z, 0) = \sum_{n=2}^{\infty} k_n(0)H_n(z) \in L^2(f_\infty(z) dz), \right. \\ \left. \epsilon\tilde{f}(z, 0) \geq -1 \right\}.$$

THEOREM 3.3. Let $f_\infty(z)$ be the steady-state solution of the nonlinear MF system (3.9)-(3.11). Then, $f_\infty(z)$ is linearly asymptotically stable with respect to the set of ϵ perturbed initial density functions $S_{\text{per}}^{\text{Lin}}(\epsilon)$, moreover, in this case the ϵ perturbed

solutions (3.18)-(3.19) take the forms $v_\epsilon(z, t) = v_\infty(z)$, $\bar{c}_\epsilon(z, t) = \bar{c}_\infty(z)$ and

$$f_\epsilon(z, t) = f_\infty(z) \left(1 + \epsilon \sum_{n=2}^{\infty} \exp\left(-n \frac{\gamma}{r} t\right) k_n(0) H_n(z) \right), \quad (3.26)$$

where $z \in \mathbb{R}$, $t \geq 0$.

Proof. See the appendix. □

3.4.3. Stability Analysis of the Nonlinear Mean Field System. We now present an infinite dimensional convex set of initial perturbations on the steady-state density $f_\infty(z)$ which yields time-varying solutions to the nonlinear MF system (3.9)-(3.11).

DEFINITION 3.3. *A steady-state solution $f_\infty(z)$ of the nonlinear equation system (3.9)-(3.11) is asymptotically stable with respect to a set of initial perturbations*

$$S = \left\{ f(z, 0) = f_\infty(z) (1 + \tilde{f}(z, 0)) : f(z, 0) \geq 0, \int_{\mathbb{R}} f(z, 0) dz = 1, \right. \\ \left. \tilde{f}(z, 0) \in L^2(f_\infty(z) dz) \right\}$$

if there exists a solution $f(z, \cdot) \in L^2(f_\infty(z) dz)$ to the nonlinear MF system (3.9)-(3.11) such that $\lim_{t \rightarrow \infty} \|f(z, t) - f_\infty(z)\|_{L^2} = 0$. □

We define the set of initial density functions

$$S_{\text{per}}^{\text{NL}} := \left\{ f(z, 0) = f_\infty(z) (1 + \tilde{f}(z, 0)) : \tilde{f}(z, 0) = \sum_{n=2}^{\infty} k_n(0) H_n(z) \in L^2(f_\infty(z) dz), \right. \\ \left. \tilde{f}(z, 0) \geq -1 \right\}.$$

We observe that $S_{\text{per}}^{\text{NL}} = S_{\text{per}}^{\text{Lin}}(1)$.

THEOREM 3.4. *If the initial density function $f(z, 0)$ is in the set $S_{\text{per}}^{\text{NL}}$, then the solution to the nonlinear MF system (3.9)-(3.11) takes the forms $v(z, t) = v_\infty(z)$,*

$\bar{c}(z, t) = \bar{c}_\infty(z)$ and

$$f(z, t) = f_\infty(z) \left(1 + \sum_{n=2}^{\infty} \exp\left(-n\frac{\gamma}{r}t\right) k_n(0) H_n(z) \right), \quad (3.27)$$

where $z \in \mathbb{R}$, $t \geq 0$. Moreover, the steady-state Gaussian solution $f_\infty(z)$ is asymptotically stable.

Proof. See the appendix. □

REMARK 3.4. The set S_{per}^{NL} is non-empty since $f(z, 0) = f_\infty(z)(1 + H_2(z)) \in S_{per}^{NL}$ where $H_2(z) \equiv (1/\sqrt{2}) \left(\frac{(z-\mu)^2}{s^2} - 1 \right)$. In general, for each function $h(z) \geq -1$ in the space $L^2(f_\infty(z)dz)$ which satisfies the mass and the mean preservation conditions (i.e., $\int_{\mathbb{R}} h(z)f_\infty(z)dz = \int_{\mathbb{R}} zh(z)f_\infty(z)dz = 0$) the density function $g(z) = f_\infty(z)(1 + h(z))$ is a member of the set S_{per}^{NL} .

3.5. Mean-Consensus

We recall the following definition from Chapter 2 (see Definition 2.1).

DEFINITION 3.4. Mean-consensus is said to be achieved asymptotically for a group of N agents if $\lim_{t \rightarrow \infty} |E(z_i(t) - z_j(t))| = 0$ for any i and j , $1 \leq i \neq j \leq N$. If the mean-consensus value is the initial state population mean of the system then the initial mean-consensus is said to be achieved. □

In the class of stable solutions to the MF-FPK equations given in (3.26) and (3.27), $\mu \equiv \int_{\mathbb{R}} zf_\infty(z)dz$ is uniquely determined as the initial state population mean $\int_{\mathbb{R}} zf(z, 0)dz$. This is due the fact that $\int_{\mathbb{R}} zH_n(z)f_\infty(z)dz = \langle z, H_n(z) \rangle = \langle \mu H_0(z) + sH_1(z), H_n(z) \rangle = 0$ for $n \geq 2$, by the orthogonality property of the Hermite polynomials (see Proposition 3.2-(a)). Moreover, using the MF continuum based control law (3.12) for a finite N population system (3.1)-(3.2) yields the individual control

$$u_i^o(t) = -\frac{1}{2r} \partial_z v(z, t) \Big|_{z=z_i(t)} = -\frac{\gamma}{r} (z_i(t) - \mu), \quad t \geq 0 \quad 1 \leq i \leq N, \quad (3.28)$$

where $\mu = \int_{\mathbb{R}} z f(z, 0) dz$ is the initial state population mean. Applying the control laws (3.28) to the agents' dynamics (3.1) yields the closed-loop solutions:

$$z_i^o(t) = \mu + e^{-\frac{\gamma}{r}t} (z_i(0) - \mu) + \sigma \int_0^t e^{-\frac{\gamma}{r}(t-\tau)} dw_i(\tau), \quad t \geq 0 \quad 1 \leq i \leq N. \quad (3.29)$$

We now have the following theorem which is the same as Theorem 2.3.

THEOREM 3.5. *By applying the continuum based MF control laws (3.28) in a finite population DGCM (3.1)-(3.2), an initial mean-consensus is reached asymptotically (as time goes to infinity) with individual asymptotic variance $s^2 = \frac{\sigma^2 r}{2\gamma}$.* \square

We note that mean-consensus can be called “weak mean-square consensus” where the agents reach mean-consensus asymptotically (as time goes to infinity) with finite asymptotic variance for each individual agent.

The reader is referred to Section 2.5.1 for the stability and performance analyses of the MF control laws. Moreover, the infinite population mass effect approximation to the finite population closed-loop centroid of flock of agents is justified in Theorems 2.5 and 2.6 of Chapter 2.

3.6. ϵ -Nash Equilibrium Property of MF Control Laws

We recall the following definition from Chapter 2 (see Definition 2.2).

DEFINITION 3.5. [79] *Given $\epsilon > 0$, the set of controls $\{u_i^o \in \mathcal{U}_i : 1 \leq i \leq N\}$ for N agents generates an ϵ -Nash equilibrium with respect to the costs $\{J_i^N : 1 \leq i \leq N\}$, if $J_i^N(u_i^o, u_{-i}^o) - \epsilon \leq \inf_{u_i \in \mathcal{U}_i} J_i^N(u_i, u_{-i}^o) \leq J_i^N(u_i^o, u_{-i}^o)$, for any $1 \leq i \leq N$.* \square

For a generic agent $1 \leq i \leq N$ denote (see 2.29)

$$\epsilon_N^2 := \max_{1 \leq i \leq N} E \int_0^\infty e^{-\rho t} \left(\frac{1}{N-1} \sum_{j=1, j \neq i}^N z_j^o(t) - \mu \right)^2 dt,$$

where $z_i^o(\cdot)$ is the closed-loop solution of the i^{th} agent's dynamics given in (3.29). Due to the fact that in the class of stable solutions (3.26) and (3.27) to the MF-FPK

equation (3.10) we have $\int_{\mathbb{R}} z f(z, t) dz = \int_{\mathbb{R}} z f_0(z) dz = \mu$ for all $t > 0$, the proof of the following theorem is similar to Theorem 2.8 in Chapter 2.

THEOREM 3.6. *Assume (A3.1) holds. Then the set of MF control laws for the finite population system $\{u_i^o \in \mathcal{U}_i : 1 \leq i \leq N\}$ given in (3.28) generates an ϵ_N -Nash equilibrium such that*

$$J_i^N(u_i^o, u_{-i}^o) - \epsilon_N \leq \inf_{u_i \in \mathcal{U}_i} J_i^N(u_i, u_{-i}^o) \leq J_i^N(u_i^o, u_{-i}^o), \quad 1 \leq i \leq N,$$

where $\lim_{N \rightarrow \infty} \epsilon_N = 0$. □

3.7. Dynamic Game Consensus Model with LTA (Ergodic) Costs

Assume that in a system with population size N and individual dynamics (3.1), the objective of the i^{th} individual agent is to almost surely (a.s.) minimize the Long Time Average (LTA) (i.e., ergodic) cost function (see [133]):

$$J_{\text{LTA},i}^N(u_i, u_{-i}) := \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left((z_i(t) - \frac{1}{N-1} \sum_{j=1, j \neq i}^N z_j(t))^2 + r u_i^2(t) \right) dt, \quad (3.30)$$

where r is a positive integer. The admissible control set of the i^{th} agent is [106]

$$\mathcal{U}_i := \left\{ u_i(\cdot) : u_i(t) \text{ is adapted to the sigma-field } \mathcal{F}_t, |z_i(T)|^2 = o(\sqrt{T}), \int_0^T (z_i(t))^2 dt = O(T), \text{ a.s.} \right\}.$$

First, we consider a “single agent” LTA optimal control problem (OCP):

$$dz(t) = u(t)dt + \sigma dw(t), \quad t \geq 0, \quad (3.31)$$

$$\inf_{u \in \mathcal{U}} J(u) := \inf_{u \in \mathcal{U}} \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T (c(z, t) + r u^2(t)) dt, \quad (3.32)$$

where $z(\cdot), u(\cdot) \in \mathbb{R}$ are the state and control input, respectively; $w(\cdot)$ denotes a standard scalar Wiener process; $c(z, \cdot)$ is a positive function; and \mathcal{U} is the corresponding admissible control set of the generic agent.

An admissible control $u^\circ(\cdot) \in \mathcal{U}$ is called a.s. optimal if there exists a constant ρ° such that

$$J(u^\circ) = \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(c(z^\circ(t), t) + r(u^\circ(t))^2 \right) dt = \rho^\circ, \text{ a.s.},$$

where $z^\circ(\cdot)$ is the solution of (3.31) under $u^\circ(\cdot)$, and for any other admissible control $u(\cdot) \in \mathcal{U}$, we have a.s. $J(u) \geq \rho^\circ$.

For $x \in \mathbb{R}$ and $0 \leq t < \infty$ we define the *relative value function* $v(\cdot, \cdot)$ for the OCP (3.31)-(3.32) by (see [22])

$$v(x, t) := \inf_{u \in \mathcal{U}} E \left[\int_t^\infty \left(c(z(s), s) + r(u(s))^2 - \rho^\circ \right) ds \mid z(t) = x \right]. \quad (3.33)$$

We have the following result (see [11, 22]):

THEOREM 3.7. (*HJB for the LTA OCP*) *Assume that the function $c(z, t)$ is Lipschitz continuous with respect to z and uniformly continuous with respect to t , and assume the value function $v(z, t)$ for the OCP (3.31)-(3.32) is a C^1 function of variable t and C^2 function of variable z , then $v(z, t)$ solves the (backward in time) Hamilton-Jacobi-Bellman (HJB) equation*

$$\partial_t v(z, t) + H(\partial_z v(z, t)) + \frac{\sigma^2}{2} \partial_{zz}^2 v(z, t) + c(z, t) = \rho^\circ, \quad (3.34)$$

with boundary condition $\lim_{t \rightarrow \infty} v(z(t), t)/t = 0$, where the Hamiltonian $H(\cdot)$ is defined as $H(p) \equiv \min_{u \in \mathcal{U}} H(p, u) := \min_{u \in \mathcal{U}} \{up + ru^2\}$ for p in \mathbb{R} . \square

The solution of the LTA OCP (3.31)-(3.31) is

$$u^\circ(t) := \arg \min_{u \in \mathcal{U}} H(\partial_z v(z, t), u) = -\frac{1}{2r} \partial_z v(z, t), \quad t \geq 0.$$

Substituting $u^\circ(\cdot)$ into the HJB equation (3.34) yields the (backward in time) ergodic HJB equation:

$$\partial_t v(z, t) - \frac{1}{4r} (\partial_z v(z, t))^2 + \frac{\sigma^2}{2} \partial_{zz}^2 v(z, t) + c(z, t) = \rho^\circ,$$

with boundary condition $\lim_{t \rightarrow \infty} v(z(t), t)/t = 0$.

Following argument exactly parallel to those used in derivation system of (3.9)-(3.11) in Section 3.3 (see [133]), we obtain the continuum based MF system of the LTA dynamic game problem, (3.1) and (3.30):

$$[\text{MF-HJB}] \quad \partial_t v(z, t) = \frac{1}{4r} (\partial_z v(z, t))^2 - \bar{c}(z, t) + \rho^o - \frac{\sigma^2}{2} \partial_{zz}^2 v(z, t), \quad (3.35)$$

$$[\text{MF-FPK}] \quad \partial_t f(z, t) = \frac{1}{2r} \partial_z \left((\partial_z v(z, t)) f(z, t) \right) + \frac{\sigma^2}{2} \partial_{zz}^2 f(z, t), \quad (3.36)$$

$$[\text{MF-CC}] \quad \bar{c}(z, t) = \left(\int_{\mathbb{R}} (z - z') f(z', t) dz' \right)^2. \quad (3.37)$$

where $v(\cdot, \cdot)$ is the relative value function, ρ^o is the best response optimal cost, $f(z, 0) = f_0(z)$ is the given initial population density, necessarily $\int_{\mathbb{R}} f(z, t) dz = 1$ for any $t \geq 0$, and it is assumed that $\lim_{|z| \rightarrow \infty} f(z, t) = 0$ and $\lim_{t \rightarrow \infty} v(z(t), t)/t = 0$.

In the stationary setting, the MF system (3.35)-(3.37) takes the form:

$$\frac{1}{4r} (\partial_z v_{\infty}(z))^2 - \frac{\sigma^2}{2} \partial_{zz}^2 v_{\infty}(z) = \bar{c}_{\infty}(z) - \rho^o, \quad (3.38)$$

$$\frac{1}{2r} \partial_z \left((\partial_z v_{\infty}(z)) f_{\infty}(z) \right) = -\frac{\sigma^2}{2} \partial_{zz}^2 f_{\infty}(z), \quad (3.39)$$

$$\bar{c}_{\infty}(z) = \left(\int_{\mathbb{R}} (z - z') f_{\infty}(z') dz' \right)^2, \quad (3.40)$$

where the density $f_{\infty}(z)$ satisfies $\int_{\mathbb{R}} f_{\infty}(z) dz = 1$, and $\lim_{t \rightarrow \infty} v_{\infty}(z(t))/t = 0$.

The assertion of the following theorem is straightforward to verify.

THEOREM 3.8. *For any arbitrary $\mu \in \mathbb{R}$, there exists the following solution of the stationary system (3.38)-(3.40):*

$$v_{\infty}(z) = \sqrt{r}(z - \mu)^2, \quad \rho^o = \sigma^2 \sqrt{r}, \quad (3.41)$$

$$f_{\infty}(z) = \frac{1}{\sqrt{2\pi s^2}} \exp\left(-\frac{(z - \mu)^2}{2s^2}\right), \quad \bar{c}_{\infty}(z) = (z - \mu)^2, \quad (3.42)$$

where $s^2 := \frac{\sigma^2 \sqrt{r}}{2}$, and $v_{\infty}(z)$ is defined up to a constant. \square

By following arguments exactly parallel to those used in Section 3.4.2, we can show the linear stability and nonlinear stability of the MF system (3.35)-(3.37) with respect to the sets $S_{\text{per}}^{\text{Lin}}(\epsilon)$ and $S_{\text{per}}^{\text{NL}}$, respectively.

The MF continuum based control law for a finite N population system, (3.1) and (3.30), yields the individual control

$$u_i^o(t) = -\frac{1}{2r} \partial_z v(z, t) \Big|_{z=z_i(t)} = -\frac{1}{\sqrt{r}} (z_i(t) - \mu), \quad t \geq 0 \quad 1 \leq i \leq N, \quad (3.43)$$

where $\mu = \int_{\mathbb{R}} z f(z, 0) dz$ is the initial state population mean. The mean-consensus and the ϵ -Nash property of the resulting MF control laws follow from similar arguments in Sections 3.5 and 3.6, respectively (see [133]).

3.7.1. The Evolution Mean Field System. The relative value function $v(\cdot, \cdot)$ defined in (3.33) represents perturbations around the steady-state optimal cost rate. It turns out that the corresponding HJB equation (3.35) in the MF system has a larger class of stable perturbed solutions in forward time than in backward time (see [64]).

In this section we introduce an Evolution (i.e., forward in time) Mean Field (EMF) system (based on [64]) to exhibit a forward in time process which asymptotically (as time goes to infinity) converges to the stationary equilibrium solution (3.41)-(3.42) (where $\mu = \int_{\mathbb{R}} z f(z, 0) dz$ is the initial state population mean) from any initial in a infinitesimal neighbourhood of this equilibrium.

The EMF system is given by

$$\partial_t v(z, t) = -\frac{1}{4r} (\partial_z v(z, t))^2 + \bar{c}(z, t) - \rho^o + \frac{\sigma^2}{2} \partial_{zz}^2 v(z, t), \quad (3.44)$$

$$\partial_t f(z, t) = \frac{1}{2r} \partial_z \left((\partial_z v(z, t)) f(z, t) \right) + \frac{\sigma^2}{2} \partial_{zz}^2 f(z, t), \quad (3.45)$$

$$\bar{c}(z, t) = \left(\int_{\mathbb{R}} (z - z') f(z', t) dz' \right)^2, \quad (3.46)$$

for $t \geq 0$, where $f(z, 0)$ and $v(z, 0) \equiv v_\infty(z)$, necessarily $\int_{\mathbb{R}} f(z, t) dz = 1$ for any $t \geq 0$, and it is assumed that $\lim_{|z| \rightarrow \infty} f(z, t) = 0$ and $\lim_{t \rightarrow \infty} v(z(t), t)/t = 0$.

In the EMF system (3.44)-(3.46) the equations (3.45)-(3.46) are the same as (3.36)-(3.37) but the backward in time MF-HJB equation (3.35) is replaced by a forward in time equation (3.44). It is important to note that the stationary solution

of the EMF system (3.44)-(3.46) is the same as that of the MF system (3.35)-(3.37) (see Theorem 3.8).

3.7.2. Stability Analysis of the Evolution Mean Field System. First, by taking a similar approach to the one of Section 3.4.2 (see [64]) we study the stability of the Gaussian steady-state density function $f_\infty(z)$ based on small perturbations of the forms: (i) $f_\epsilon(z, 0) = f_\infty(z)(1 + \epsilon\tilde{f}(z, 0))$ on $f_\infty(z)$ such that $f_\epsilon(z, 0)$ is a probability density, and (ii) $v_\epsilon(z, 0) = v_\infty(z) + \epsilon\tilde{v}(z, 0)$. We let the additive and multiplicative perturbations of the solution to the EMF system (3.35)-(3.37) be of the forms

$$v_\epsilon(z, t) = v_\infty(z) + \epsilon\tilde{v}(z, t), \quad f_\epsilon(z, t) = f_\infty(z)(1 + \epsilon\tilde{f}(z, t)), \quad (3.47)$$

$$\bar{c}_\epsilon(z, t) = \bar{c}_\infty(z) + \epsilon\tilde{c}(z, t), \quad (3.48)$$

for $z \in \mathbb{R}$ and $t \geq 0$, where $v_\infty(z)$, $f_\infty(z)$ and $\bar{c}_\infty(z)$ are given in (3.41)-(3.42). Since $f_\epsilon(z, \cdot)$ satisfies the MF-FPK equation (3.36), $f_\epsilon(z, t)$ for any time $t > 0$ is necessarily a probability density, i.e., $f_\epsilon(z, t) = f_\infty(z)(1 + \epsilon\tilde{f}(z, t)) \geq 0$ and $\int_{\mathbb{R}} f_\epsilon(z, t) dz = 1$.

The proof of the following theorem is similar to the proof of Proposition 3.1.

PROPOSITION 3.3. *The linearization of the EMF system (3.44)-(3.46) around the stationary equilibrium solution (3.41)-(3.42) takes the form ($z \in \mathbb{R}$, $t \geq 0$):*

$$\partial_t \tilde{v}(z, t) = -\frac{(z - \mu)}{\sqrt{r}} \partial_z \tilde{v}(z, t) + \frac{\sigma^2}{2} \partial_{zz}^2 \tilde{v}(z, t) + \tilde{c}(z, t), \quad (3.49)$$

$$\begin{aligned} \partial_t \tilde{f}(z, t) = & -\frac{(z - \mu)}{\sqrt{r}} \partial_z \tilde{f}(z, t) + \frac{\sigma^2}{2} \partial_{zz}^2 \tilde{f}(z, t) \\ & - \frac{1}{\sigma^2 r} \left(\frac{(z - \mu)}{\sqrt{r}} \partial_z \tilde{v}(z, t) - \frac{\sigma^2}{2} \partial_{zz}^2 \tilde{v}(z, t) \right), \end{aligned} \quad (3.50)$$

$$\tilde{c}(z, t) = -2(z - \mu) \left(\int_{\mathbb{R}} z \tilde{f}(z, t) f_\infty(z) dz \right), \quad (3.51)$$

with given $\tilde{f}(z, 0)$ and $\tilde{v}(z, 0)$, and boundary conditions $\lim_{|z| \rightarrow \infty} \tilde{f}(z, t) = 0$ and $\lim_{t \rightarrow \infty} \tilde{v}(z(t), t)/t = 0$.

□

In this section we let the Hermite polynomials $\{H_n(z) : n \in \mathbb{N}_0\}$ be given in Definition 3.1 with $s^2 \equiv \frac{\sigma^2 \sqrt{r}}{2}$. These polynomials are eigenfunctions of the operator

$$\mathcal{L}_1 g(z) := \frac{1}{\sqrt{r}}(z - \mu) \partial_z g(z) - \frac{\sigma^2}{2} \partial_{zz}^2 g(z), \quad z \in \mathbb{R}, \quad (3.52)$$

such that $\mathcal{L}_1 H_n(z) = \frac{n}{\sqrt{r}} H_n(z)$ for any $n \in \mathbb{N}_+$.

By using the operator \mathcal{L}_1 defined in (3.52) we can rewrite the equation system (3.49)-(3.50) as

$$\partial_t \tilde{v}(z, t) = -\mathcal{L}_1 \tilde{v}(z, t) + \tilde{c}(z, t), \quad (3.53)$$

$$\partial_t \tilde{f}(z, t) = -\frac{1}{\sigma^2 r} \mathcal{L}_1 \tilde{v}(z, t) - \mathcal{L}_1 \tilde{f}(z, t), \quad (3.54)$$

where $\tilde{c}(z, \cdot)$ is given in (3.51), and the initial and boundary conditions are those in Proposition 3.3.

We recall that

$$S_{\text{per}}^{\text{Lin}}(\epsilon) = \left\{ f_\epsilon(z, 0) = f_\infty(z)(1 + \epsilon \tilde{f}(z, 0)) : \tilde{f}(z, 0) = \sum_{n=2}^{\infty} k_n(0) H_n(z) \in L^2(f_\infty(z) dz), \right. \\ \left. \epsilon \tilde{f}(z, 0) \geq -1 \right\},$$

and $S_{\text{per}}^{\text{NL}} = S_{\text{per}}^{\text{Lin}}(1)$. We also let

$$\hat{S}_{\text{per}}^{\text{Lin}}(\epsilon) := \left\{ v_\epsilon(z, 0) = v_\infty(z) + \epsilon \tilde{v}(z, 0) : \tilde{v}(z, 0) = \sum_{n=2}^{\infty} l_n(0) H_n(z) \in L^2(f_\infty(z) dz) \right\},$$

and $\hat{S}_{\text{per}}^{\text{NL}} := \hat{S}_{\text{per}}^{\text{Lin}}(1)$.

THEOREM 3.9. *Let $f_\infty(z)$ be the steady-state solution of the nonlinear EMF system (3.44)-(3.46). Then, $f_\infty(z)$ is linearly asymptotically stable with respect to the set of ϵ perturbed initial density functions $S_{\text{per}}^{\text{Lin}}(\epsilon)$ and the set of ϵ perturbed functions*

$\hat{S}_{per}^{Lin}(\epsilon)$, moreover, in this case the ϵ perturbed solutions (3.47)-(3.48) take the forms:

$$\begin{aligned} v_\epsilon(z, t) &= v_\infty(z) + \epsilon \sum_{n=2}^{\infty} l_n(0) \exp\left(-\frac{nt}{\sqrt{r}}\right) H_n(z), \\ f_\epsilon(z, t) &= f_\infty(z) \left(1 + \epsilon \sum_{n=2}^{\infty} \left(k_n(0) - \frac{nt}{\sigma^2 r \sqrt{r}} l_n(0)\right) \exp\left(-\frac{nt}{\sqrt{r}}\right) H_n(z)\right), \\ \bar{c}_\epsilon(z, t) &= \bar{c}_\infty(z), \end{aligned}$$

where $z \in \mathbb{R}$, $t \geq 0$.

Proof. The proof is similar to Theorem 3.3 (see Theorem 10 in [131]). \square

THEOREM 3.10. *If the initial density function $f(z, 0)$ is in the set S_{per}^{NL} and $v(z, 0)$ is in the set \hat{S}_{per}^{NL} , then the solution to the EMF system (3.44)-(3.46) takes the forms*

$$\begin{aligned} v_\epsilon(z, t) &= v_\infty(z) + \sum_{n=2}^{\infty} l_n(0) \exp\left(-\frac{nt}{\sqrt{r}}\right) H_n(z), \\ f_\epsilon(z, t) &= f_\infty(z) \left(1 + \sum_{n=2}^{\infty} \left(k_n(0) - \frac{nt}{\sigma^2 r \sqrt{r}} l_n(0)\right) \exp\left(-\frac{nt}{\sqrt{r}}\right) H_n(z)\right), \\ \bar{c}_\epsilon(z, t) &= \bar{c}_\infty(z), \end{aligned}$$

where $z \in \mathbb{R}$, $t \geq 0$. Moreover, the steady-state Gaussian solution $f_\infty(z)$ is asymptotically stable.

Proof. The proof is similar to Theorem 3.4. \square

We note that $v(z, \cdot)$ in Theorem 3.10 yields the control law:

$$\hat{u}_i(t) := -\frac{1}{2r} \partial_z v(z, t) \Big|_{z=z_i(t)} = -\frac{1}{\sqrt{r}} (z_i(t) - \mu) + \sum_{n=2}^{\infty} n l_n(0) \exp\left(-\frac{nt}{\sqrt{r}}\right) H_{n-1}(z_i(t)),$$

where we use the fact that $\partial_z H_n(z) = n H_{n-1}(z)$. Hence, the resulting control law of the (forward in time) EMF system (3.44)-(3.46): (i) gives the same asymptotic steady-state solution and performance as in the (backward/forward) MF system (3.35)-(3.37), and (ii) has a larger class of stable perturbed solutions than the control law (3.43) in the transient state.

3.8. Numerical Example

Consider a system (3.1)-(3.2) of 500 agents with $\sigma = 0.05$, $r = 3$ and $\rho = 0.2$. The initial states of the agents are taking independently from a uniform distribution with support $[-3, 3]$. Fig. 3.1A shows the contour lines of the evolution of the population density function $f(z, t)$ as given in (3.27). The state trajectories (3.29) of all the agents of the system are shown in Fig. 3.1B. As shown in Figs. 3.1A and 3.1B the agents reach mean consensus in $\mu = 0$ asymptotically (as time goes to infinity) with individual asymptotic variance $s^2 = \frac{\sigma^2 r}{2\gamma}$.

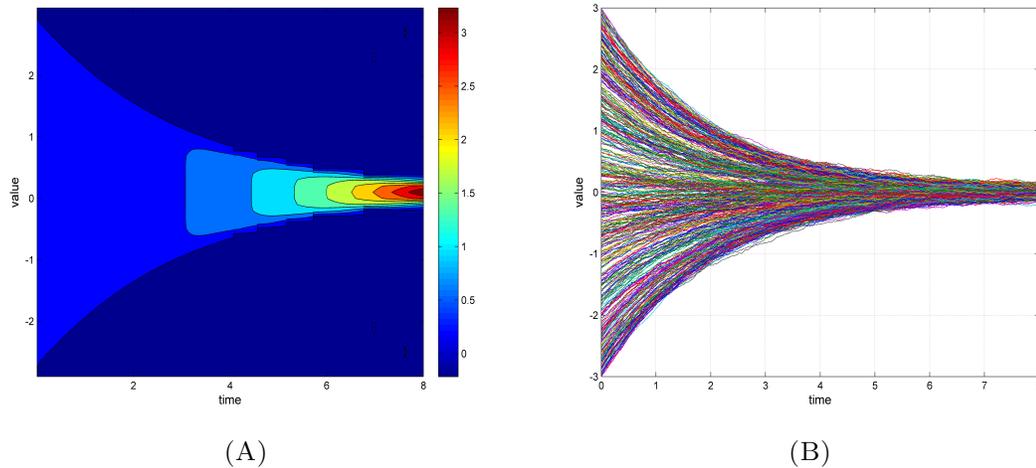


FIGURE 3.1. (A) The contour lines of population density functions, and (B) trajectories of agents' states when $N = 500$.

3.9. Chapter Summary

This chapter presents a synthesis of consensus behaviour as a stochastic dynamic game problem. In this problem formulation each agent in the system: (i) has a simple stochastic dynamics with inputs directly controlling its state's rate of change, and (ii) seeks to minimize its individual cost function containing a mean field coupling to the states of all other agents. We take a continuum approach to this problem via mean field stochastic control theory. Based on this methodology we synthesize a set of mean field decentralized ϵ_N -Nash equilibrium strategies for a system with population

size N . The resulting Mean Field (MF) control strategies steer each individual's state toward the initial state population mean which is reached asymptotically as time goes to infinity, thus achieving mean consensus. As the population size of this finite N population system goes to infinity ϵ_N goes to zero and the set of mean field control strategies becomes an exact Nash equilibrium.

In the case of agents with Long Time Average (LTA) (i.e., ergodic) cost functions, an Evolution (i.e., forward in time) Mean Field (EMF) system of consensus model is studied. The EMF system consists of two coupled (forward in time) deterministic PDEs which are also coupled to the cost coupling function.

3.10. Appendix

Proof of Proposition 3.1: Here, we follow similar arguments to the ones of Proposition 8 in [64]. By substituting $v_\epsilon(z, \cdot)$ and $\bar{c}_\epsilon(z, \cdot)$ from (3.18) into the MF-HJB equation (3.9) we get

$$\begin{aligned} \epsilon \partial_t \tilde{v}(z, t) &= \frac{1}{4r} (\partial_z v_\infty(z))^2 - \bar{c}_\infty(z) + \rho v_\infty(z) - \frac{\sigma^2}{2} \partial_{zz}^2 v_\infty(z) \\ &+ \epsilon \left(\frac{\gamma}{r} (z - \mu) \partial_z \tilde{v}(z, t) - \tilde{c}(z, t) + \rho \tilde{v}(z, t) - \frac{\sigma^2}{2} \partial_{zz}^2 \tilde{v}(z, t) \right) + O(\epsilon^2). \end{aligned}$$

where we use $\partial_z v_\infty(z) = 2\gamma(z - \mu)$ by (3.16). By (3.13) we have

$$\frac{1}{4r} (\partial_z v_\infty(z))^2 - \bar{c}_\infty(z) + \rho v_\infty(z) - \frac{\sigma^2}{2} \partial_{zz}^2 v_\infty(z) = 0,$$

and hence the first order terms in ϵ in the above equation yield (3.20).

By substituting $f_\epsilon(z, \cdot)$ and $v_\epsilon(z, \cdot)$ from (3.18) into the MF-FPK equation (3.21) we get

$$\begin{aligned} \epsilon \partial_t \tilde{f}(z, t) f_\infty(z) &= \frac{1}{2r} \partial_z ((\partial_z v_\infty(z)) f_\infty(z)) \\ &+ \frac{\sigma^2}{2} \partial_{zz}^2 f_\infty(z) + \epsilon \frac{\sigma^2}{2} \partial_{zz}^2 (\tilde{f}(z, t) f_\infty(z)) \\ &+ \epsilon \frac{1}{2r} \left(\partial_z (2\gamma(z - \mu) \tilde{f}(z, t) f_\infty(z)) + (\partial_z \tilde{v}(z, t)) f_\infty(z) \right) + O(\epsilon^2). \end{aligned} \quad (3.55)$$

But we have $\frac{1}{2r}\partial_z((\partial_z v_\infty(z))f_\infty(z)) + \frac{\sigma^2}{2}\partial_{zz}^2 f_\infty(z) = 0$, by (3.14). Since $\partial_z f_\infty(z) = -(\frac{z-\mu}{s^2})f_\infty(z)$ and $\partial_{zz}^2 f_\infty(z) = (\frac{(z-\mu)^2}{s^4} - \frac{1}{s^2})f_\infty(z)$, the first order terms in ϵ of (3.55) yield (3.21).

Finally, by substituting $f_\epsilon(z, \cdot)$ into (3.11) we get

$$\begin{aligned} \left(\int_{\mathbb{R}}(z-z')f_\epsilon(z',t)dz'\right)^2 &= ((z-\mu) + \epsilon \int_{\mathbb{R}}(z-z')\tilde{f}(z',t)f_\infty(z')dz')^2 \\ &= (z-\mu)^2 + 2\epsilon(z-\mu)\left(\int_{\mathbb{R}}(z-z')\tilde{f}(z',t)f_\infty(z')dz'\right) + O(\epsilon^2). \end{aligned}$$

Since $f_\epsilon(z, \cdot)$ in (3.18) is a probability density we have $\int_{\mathbb{R}}\tilde{f}(z,t)f_\infty(z)dz = 0$ for all $t \geq 0$. Hence, the first order terms in ϵ in above equation yields (3.22) by using the definition of $\bar{c}_\epsilon(z, \cdot)$ in (3.18). \square

Proof of Proposition 3.2: (a) See Chapter 22 in [1].

(b) The Hermite polynomials $H_n(z)$, $n \in \mathbb{N}_0$, have the generating function

$$\phi(z, h) = \sum_{n=0}^{\infty} H_n(z) \frac{h^n}{n!} = \exp\left((z-\mu)h - \frac{s^2 h^2}{2}\right).$$

The generating function ϕ satisfies the PDE

$$\frac{\gamma}{r}(z-\mu)\partial_z\phi(z, h) - \frac{\sigma^2}{2}\partial_{zz}^2\phi(z, h) = h\frac{\gamma}{r}\partial_h\phi(z, h),$$

and hence $H_n(z)$, $n \in \mathbb{N}_0$, is the solution of the Hermite differential equation

$$\frac{\gamma}{r}(z-\mu)\partial_z g(z) - \frac{\sigma^2}{2}\partial_{zz}^2 g(z) = n\frac{\gamma}{r}g(z),$$

or Hermite eigenvalue problem $\mathcal{L}g(z) = n\frac{\gamma}{r}g(z)$. \square

Proof of Theorem 3.3: Let

$$\tilde{f}(z, t) = \sum_{n=0}^{\infty} k_n(t)H_n(z), \quad \tilde{v}(z, t) = \sum_{n=0}^{\infty} l_n(t)H_n(z), \quad z \in \mathbb{R}, t \geq 0.$$

Note that at any time $t \geq 0$ the perturbed cost $\tilde{c}(z, t) \in \text{span}(H_1(z))$ (see (3.56) below). We now consider the equation system (3.24)-(3.25) in the Hermite coordinates. For $n = 0$ we get $\partial_t l_0(t) = \rho l_0(t)$ from (3.24). But, since $\rho > 0$ the only function in

$L^2(f_\infty(z)dz)$ that satisfies this equation is the zero function. Next, since $f_\epsilon(z, \cdot)$ in (3.18) is a probability density we have $k_0(t) = 0$, $t \geq 0$, which also satisfies the equation (3.25) for the Hermite coordinate $n = 0$. For $n \geq 2$ we can rewrite the equation system (3.24)-(3.25) in the Hermite coordinates as the ODE (by Proposition 3.2-(b))

$$\frac{\partial}{\partial t} \begin{pmatrix} l_n(t) \\ k_n(t) \end{pmatrix} = \begin{pmatrix} n\frac{\gamma}{r} + \rho & 0 \\ -n\frac{\gamma}{\sigma^2 r^2} & -n\frac{\gamma}{r} \end{pmatrix} \begin{pmatrix} l_n(t) \\ k_n(t) \end{pmatrix}, \quad t \geq 0, \quad n \geq 2.$$

But, since $n\frac{\gamma}{r} + \rho > 0$ the only function in $L^2(f_\infty(z)dz)$ that satisfies $\partial_t l_n(t) = (n\frac{\gamma}{r} + \rho)l_n(t)$ is the zero function. Hence, the above equation gives us $k_n(t) = \exp(-n\frac{\gamma}{r}t)k_n(0)$, $t \geq 0$, for $n \geq 2$.

But, from (3.22) we have

$$\begin{aligned} \tilde{c}(z, t) &= -2sH_1(z) < \mu H_0(z) + sH_1(z), \sum_{n=1}^{\infty} k_n(\cdot)H_n(z) > \\ &= -2s^2k_1(t)H_1(z), \quad t \geq 0, \end{aligned} \quad (3.56)$$

by the orthogonality property of the Hermite polynomials (see Proposition 3.2-(a)) and the fact that $z = \mu H_0(z) + sH_1(z)$. Then the Hermite coordinates of the equation system (3.24)-(3.25) for $n = 1$ satisfy the ODE (by Proposition 3.2-(b))

$$\frac{\partial}{\partial t} \begin{pmatrix} l_1(t) \\ k_1(t) \end{pmatrix} = \begin{pmatrix} \frac{\gamma}{r} + \rho & 2s^2 \\ -\frac{\gamma}{\sigma^2 r^2} & -\frac{\gamma}{r} \end{pmatrix} \begin{pmatrix} l_1(t) \\ k_1(t) \end{pmatrix} =: A \begin{pmatrix} l_1(t) \\ k_1(t) \end{pmatrix}, \quad t \geq 0, \quad (3.57)$$

where γ and s^2 are respectively given in (3.16) and (3.17). A is a singular matrix and may be brought to the diagonal form via $J = P^{-1}AP$ where

$$J = \begin{pmatrix} 0 & 0 \\ 0 & \rho \end{pmatrix}, \quad P = \begin{pmatrix} -\sigma^2 r & 1 \\ 1 & -\frac{\gamma^2}{\sigma^2 r^2} \end{pmatrix}.$$

We may write the solution of (3.57):

$$\begin{aligned} \begin{pmatrix} l_1(t) \\ k_1(t) \end{pmatrix} &= e^{At} \begin{pmatrix} l_1(0) \\ k_1(0) \end{pmatrix} = P e^{Jt} P^{-1} \begin{pmatrix} l_1(0) \\ k_1(0) \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\rho} \left(-\frac{\gamma}{r} l_1(0) - \frac{\sigma^2 r}{\gamma} k_1(0) \right) + \frac{1}{\rho \gamma} e^{\rho t} (l_1(0) + \sigma^2 r k_1(0)) \\ \frac{1}{\rho} \left(\frac{\gamma}{\sigma^2 r^2} l_1(0) + \frac{1}{\gamma} k_1(0) \right) - \frac{1}{\rho r} e^{\rho t} \left(\frac{\gamma}{\sigma^2 r} l_1(0) + \gamma k_1(0) \right) \end{pmatrix}, \end{aligned}$$

for $t \geq 0$. Hence, its unique bounded solution in L^2 is given by $(l_1(t), k_1(t)) = (l_1(0), k_1(0))$ for all $t \geq 0$, where $l_1(0) = -\sigma^2 r k_1(0)$ (note that γ is the positive solution of the algebraic equation $\gamma^2 + r\rho\gamma - r = 0$, as given in (3.16)).

Therefore, we have the following bounded and C^∞ solution to the equation system (3.20)-(3.22):

$$\begin{aligned} \tilde{v}(z, t) &= -\sigma^2 r k_1(0) H_1(z), & \tilde{c}(z, t) &= -2s^2 k_1(0) H_1(z), \\ \tilde{f}(z, t) &= k_1(0) H_1(z) + \sum_{n=2}^{\infty} k_n(t) H_n(z) \\ &\equiv k_1(0) H_1(z) + \sum_{n=2}^{\infty} \exp\left(-n \frac{\gamma}{r} t\right) k_n(0) H_n(z), \end{aligned}$$

in $\mathbb{R} \times [0, \infty)$. By the Cauchy-Schwarz inequality and the integral test for convergence, it can be shown that for any fixed t , $(k_n(t))_n$ is in the space l^1 (i.e., the space of sequences whose series is absolutely convergent).

Now for a fixed time t , by Parseval's theorem we get

$$\begin{aligned} \|\tilde{f}(z, t)\|_{L^2} &= \left(k_1^2(0) + \sum_{n=2}^{\infty} k_n^2(t) \right)^{1/2} \\ &\equiv \left(k_1^2(0) + \sum_{n=2}^{\infty} \exp\left(-2n \frac{\gamma}{r} t\right) k_n^2(0) \right)^{1/2}, \quad z \in \mathbb{R}, t \geq 0. \end{aligned}$$

Since $\lim_{t \rightarrow \infty} k_n(t) = 0$ for $n \geq 2$, the Lebesgue Dominated Convergence theorem implies that $\lim_{t \rightarrow \infty} \|\tilde{f}(z, t)\|_{L^2} = k_1(0)$ which is zero if $k_1(0) = 0$ (let us note that the assumption $k_1(0) = 0$ is satisfied if $\tilde{f}(z, 0)$ is an even function in the space $L^2(f_\infty(z) dz)$). Hence, the steady-state solution $f_\infty(z)$ is linearly asymptotically stable

with respect to the set of initial perturbations $S_{\text{per}}^{\text{Lin}}(\epsilon)$ based on Definition 3.2. Note that the restriction $k_1(0) = 0$ is enforced in $S_{\text{per}}^{\text{Lin}}(\epsilon)$ by assuming the representation $\tilde{f}(z, 0) = \sum_{n=2}^{\infty} k_n(0)H_n(z)$. \square

Proof of Theorem 3.4: We apply a fixed-point argument. For any fixed time $t \geq 0$, let $f(z, t)$ be of the form

$$f(z, t) = f_{\infty}(z) \left(1 + \sum_{n=2}^{\infty} k_n(t) H_n(z) \right) \in S_{\text{per}}^{\text{NL}}, \quad z \in \mathbb{R}. \quad (3.58)$$

Then from the MF-CC equation (3.11) we get $\bar{c}(z, t) = (z - \mu)^2 \equiv c_{\infty}(z)$, $t \geq 0$, since $\int_{\mathbb{R}} z f(z, \cdot) dz = \mu$ by the orthogonality property of the Hermite polynomials (see Proposition 3.2-(a)). This $\bar{c}(z, \cdot)$ gives a well-defined solution to the MF-HJB equation (3.9) as $v(z, t) = \gamma(z - \mu)^2 + \eta \equiv v_{\infty}(z)$, $t \geq 0$, where γ and η are given in (3.16). Applying this $v(z, \cdot)$ into the MF-FPK equation (3.10) gives us

$$\partial_t f(z, t) = \frac{\gamma}{r} \partial_z \left((z - \mu) f(z, t) \right) + \frac{\sigma^2}{2} \partial_{zz}^2 f(z, t),$$

where $f(z, 0) = f_{\infty}(z) \left(1 + \sum_{n=2}^{\infty} k_n(0) H_n(z) \right) \in S_{\text{per}}^{\text{NL}}$ is given. By using the operator \mathcal{L} defined in Proposition 3.2-(b), it can be shown that the solution to the above equation is (3.27) which is indeed in the form of (3.58). We can also show, similar to the proof of Theorem 3.3-(b), that $\lim_{t \rightarrow \infty} \left\| \frac{f(z, t) - f_{\infty}(z)}{f_{\infty}(z)} \right\|_{L^2} = 0$. Hence, the steady-state solution $f_{\infty}(z)$ is asymptotically stable with respect to the set of initial perturbations $S_{\text{per}}^{\text{NL}}$ based on Definition 3.3. \square

CHAPTER 4

Synthesis of Mean Field Controlled Cucker-Smale Type Flocking: A Maxwellian Distribution

This chapter is concerned with the synthesis of a Mean Field (MF) flocking model. In this problem formulation the state of each agent consists of both its position and its controlled velocity such that: (i) all agents have similar stochastic dynamics, and (ii) each agent seeks to minimize by continuous state feedback its individual discounted cost function involving a nonlinear (relative distance based) weighted mean of the velocity states of all other agents. The cost functions are based on the normalized Cucker-Smale (CS) flocking algorithm in its original uncontrolled formulation. For this dynamic game problem, the MF system consisting of coupled deterministic Hamilton-Jacobi-Bellman (HJB) and Fokker-Planck-Kolmogorov (FPK) equations and an infinite population cost-coupling is derived approximating the stochastic system of agents as the population size goes to infinity. Subject to the existence of a unique solution to the MF system: (i) the stationary solution of the MF system is a Maxwellian distribution function, (ii) the set of MF control laws for the system possesses an ϵ_N -Nash equilibrium property where ϵ_N goes to zero as the population size N approaches infinity.

4.1. Introduction

Collective motion such as the flocking of birds, schooling of fish and swarming of bacteria is one of the most widespread phenomenon in nature. Scientists from different disciplines have studied such emergent behaviour for the past fifty years to understand the general mechanisms of cooperative phenomena and their potential applications (see [25, 47, 105, 155] and the references therein).

The study of collective motion in nature is of interest not only to model and analyze these widespread phenomena but also because ideas from these behaviours can be used by engineers to develop efficient algorithms for a wide range of applications (see [108, 147] among many other papers).

A group of agents has a flocking behaviour if: (i) agents' velocities converge to a common value (e.g., mean of initial velocities), i.e., consensus in velocity, and (ii) the distance between agents remains bounded.

There are two main classes of models for the flocking behaviour: (i) individual based models in the form of coupled Ordinary (Stochastic) Differential Equations (O(S)DEs) (see for instance [46, 169]) where in these algorithms a key element is the use of local feedback involving local communication (subject to the network topology) between agents so as to reach an agreement, and (ii) continuum models in the form of Partial (or integro-partial) Differential Equations (PDEs) to model the collective motion in the case of systems with large populations (see [40, 50, 68, 69, 167] among many others). The continuum models can be derived from the individual based models in the large population limit by use of the kinetic theory of gases, hydrodynamic and mean field theory (see for instance [39, 40, 68]).

Two fundamental individual based models, (i) Cucker-Smale flocking model, (ii) self-propelling, friction and attraction-repulsion model, and their corresponding continuum formulations can be found in the comprehensive survey paper [40].

In [46] Cucker and Smale (CS) formulated an interesting individual based flocking model for a group of agents. This model is motivated by the collective motion of a group of birds such that each bird updates its velocity as a weighted velocities of all

the other birds. The weights in this model are functions of the relative distance of the birds such that as the mutual distance between two birds increases the influence of their velocities on each other decreases. Recently, the CS model with normalized (relative distance based) communication rates is studied in [121].

Several extensions of the CS model are addressed to such problems as hierarchical leadership [160], stochastic and noisy environment [45, 67], collision avoidance [44, 144] and space vehicle control [147], among others.

This chapter studies the synthesis of a Mean Field (MF) flocking model. In this problem formulation the state of each agent consists of both its position and its controlled velocity such that: (i) all agents have similar stochastic dynamics, and (ii) each agent seeks to minimize by continuous state feedback its individual discounted cost function involving a nonlinear (relative distance based) weighted mean of the velocity states of all other agents. The cost functions are based on the normalized CS flocking algorithm in its original uncontrolled formulation. For this dynamic game problem, the MF system consisting of coupled deterministic Hamilton-Jacobi-Bellman (HJB) and Fokker-Planck-Kolmogorov (FPK) equations and an infinite population cost-coupling is derived approximating the stochastic system of agents as the population size goes to infinity. Subject to the existence of a unique solution to the MF system: (i) the stationary solution of the MF system is a Maxwellian distribution function, (ii) the set of MF control laws for the system possesses an ϵ_N -Nash equilibrium property where ϵ_N goes to zero as the population size N approaches infinity.

Hence, the model of this chapter may be regarded as a controlled game theoretic formulation of a flocking model in which each agent, instead of responding to an ad-hoc algorithm, obtains its control law from a game theoretic Nash equilibrium.

The following notation will be used throughout the chapter. We use the integer valued subscript as the label for an individual agent of the population. The integer N is reserved to denote the population size of the system. We use the superscripts N for a process to indicate the dependence on the population size. Let \mathbb{R}^n denote the n -dimensional real Euclidean space with the standard Euclidean norm $|\cdot|$. The

transpose of a vector (or matrix) x is denoted by x^T . $\text{tr}(A)$ denotes the trace of a square matrix A . The gradient vector and matrix of second order partial derivatives of f with respect to variable x are denoted by $D_x f$ and $D_{xx}^2 f$, respectively. The symbol ∂_t denotes the partial derivative with respect to variables t .

The chapter is organized as follows. Some background of fundamental uncontrolled CS flocking algorithm and its corresponding continuum formulation is presented in Section 4.2. Section 4.3 is dedicated to the problem formulation. The MF control approach to the flocking problem is presented in Section 4.4. Section 4.5 presents the Maxwellian stationary solution of the MF system. The ϵ -Nash equilibrium properties of the MF control laws is established in Section 4.6. Concluding remarks are stated in Section 4.7.

4.2. The Uncontrolled Cucker-Smale Model

The fundamental uncontrolled CS model [46] for a system of population N is given by the nonlinear system of ODEs:

$$\begin{cases} dx_i(t) = v_i(t)dt, \\ dv_i(t) = \frac{1}{N} \sum_{j=1}^N a(\|x_i(t) - x_j(t)\|)(v_j(t) - v_i(t))dt, \end{cases} \quad 1 \leq i \leq N, \quad t > 0,$$

where $x_i \in \mathbb{R}^n$ and $v_i \in \mathbb{R}^n$ are, respectively, position and velocity vectors of the i^{th} agent, $1 \leq i \leq N$, with initial states $x_i(0), v_i(0)$, and the communication rates are given by

$$a(\|x_i(t) - x_j(t)\|) := \frac{1}{(1 + \|x_i(t) - x_j(t)\|^2)^\beta}, \quad (4.1)$$

for some fixed $\beta \geq 0$ (see [121] for a CS model with normalized communication rates).

It is shown in [46] that the agents' velocities converge to a common value (the average of initial velocities) regardless of the initial configurations when $\beta < 1/2$ and also the distance between agents remain fixed and bounded but not necessarily the same. This result was improved in [68] in the case of $\beta = 1/2$.

The corresponding continuum model of the individual uncontrolled CS algorithm is the advection equation [39]

$$\frac{\partial f}{\partial t}(x, v, t) + v^T D_x f(x, v, t) = D_v^T (\xi(f)(x, v, t) f(x, v, t)), \quad f(x, v, 0),$$

where $f(x, v, t)$ is the density function of particles positioned at $(x, t) \in \mathbb{R}^n \times \mathbb{R}_+$ with velocity $v \in \mathbb{R}^n$, and the velocity field $\xi(f)(x, v, t)$ is defined as

$$\xi(f)(x, v, t) = \int_{\mathbb{R}^{2n}} a(\|x - x'\|)(v - v') f(x', v', t) dx' dv'.$$

See the comprehensive survey paper [40] (and the references therein) for the derivation of the continuum CS flocking model from the individual based CS algorithm in large populations via kinetic theory.

4.3. The Controlled Flocking Model

Consider a system of N agents. The dynamics of the agents are given by controlled Stochastic Differential Equations (SDEs):

$$\begin{cases} dx_i(t) = v_i(t) dt, \\ \\ dv_i(t) = u_i(t) dt + C dw_i(t), \end{cases} \quad 1 \leq i \leq N \quad (4.2)$$

where $x_i(\cdot) \in \mathbb{R}^n$ is the position, $v_i(\cdot) \in \mathbb{R}^n$ is the velocity, $u_i(\cdot) \in \mathbb{R}^n$ is the control input, and $\{w_i : 1 \leq i \leq N\}$ denotes a set of N independent p -dimensional standard Wiener processes. The set of initial data $\{(x_i(0), v_i(0)) : 1 \leq i \leq N\}$ are assumed to be independent and also independent of $\{w_i : 1 \leq i \leq N\}$ with finite second moments. The noise intensity matrix C is in $\mathbb{R}^{n \times p}$.

The admissible control set of the i^{th} agent is taken as

$$\mathcal{U}_i := \{u_i(\cdot) : u_i(t) \text{ is adapted to sigma-field } \sigma((x_j(s), v_j(s)) : s \leq t, 1 \leq j \leq N)\}.$$

Let the nonlinear cost-coupling function be defined homogeneously throughout the population by

$$\phi^N((x_i, v_i); (x, v)_{-i}) := \left\| \frac{1}{\sum_{j=1}^N a(\|x_i - x_j\|)} \sum_{j=1}^N a(\|x_i - x_j\|)(v_j - v_i) \right\|^2, \quad (4.3)$$

for a generic agent i , where $(x, v)_{-i} := ((x, v)_1, \dots, (x, v)_{i-1}, (x, v)_{i+1}, \dots, (x, v)_N)$, and the weight function

$$a(\|x_i - x_j\|) := \frac{1}{(1 + \|x_i - x_j\|^2)^\beta}, \quad (4.4)$$

with $\beta > 0$, is based on (4.1). We note that the cost-coupling function of each agent involves normalized (relative distance based) weighted mean of the velocity states of all other agents.

The objective of the i^{th} individual agent, $1 \leq i \leq N$, is to minimize (over the admissible control set \mathcal{U}_i) its discounted cost function

$$J_i^N(u_i, u_{-i}) := E \int_0^\infty e^{-\rho t} \left(\phi^N((x_i, v_i); (x, v)_{-i}) + u_i^T R u_i \right) dt, \quad (4.5)$$

where the constant $\rho > 0$ is the discounted factor, and R is a symmetric positive definite matrix in $\mathbb{R}^{n \times n}$. To indicate the dependence of J_i on u_i , $u_{-i} := (u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_N)$ and the population size N , we write it as $J_i^N(u_i, u_{-i})$. It is important to note that the generic agent i is coupled to all other agents via the nonlinear cost-coupling function $\phi^N((x_i, v_i); (x, v)_{-i})$.

The model (4.2)-(4.5) may be regarded as a controlled game theoretic formulation of a normalized CS flocking model (see (2.2) in [121] for a normalized uncontrolled CS model). We note that the mean field game consensus model (3.1)-(3.2) studied in previous chapter is the scalar version of the flocking model (4.2)-(4.5) with $\beta = 0$.

REMARK 4.1. *The results of this chapter can easily be extended to the case of agents with cost-coupling functions of the form*

$$\phi^N((x_i, v_i); (x, v)_{-i}) := \frac{1}{\sum_{j=1}^N a(\|x_i - x_j\|)} \sum_{j=1}^N a(\|x_i - x_j\|) \|v_j - v_i\|^2, \quad (4.6)$$

or agents with Long Time Average (LTA) (i.e., ergodic) cost functions (see [130, 132]).

For each i , let $z_i := [x_i, v_i]^T$ and rewrite (4.2) as

$$dz_i(t) = (Fz_i(t) + Gu_i(t))dt + Ddw_i(t), \quad 1 \leq i \leq N, \quad (4.7)$$

where

$$F = \begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix}, \quad G = \begin{pmatrix} 0 \\ I \end{pmatrix}, \quad D = \begin{pmatrix} 0 \\ C \end{pmatrix}.$$

The discounted cost function (4.5) may be rewritten as

$$J_i^N(u_i, u_{-i}) := E \int_0^\infty e^{-\rho t} (\phi^N(z_i; z_{-i}) + u_i^T R u_i) dt, \quad (4.8)$$

where

$$\phi^N(z_i; z_{-i}) := \left\| \frac{1}{\sum_{j=1}^N a(\|x_i - x_j\|)} \sum_{j=1}^N a(\|x_i - x_j\|) (v_j - v_i) \right\|^2, \quad (4.9)$$

where $z_{-i} := (z_1, \dots, z_{i-1}, z_{i+1}, \dots, z_N)$.

4.4. A Mean Field Game Approach

Similar to Chapter 3 we take the following steps to the dynamic game flocking model (4.2)-(4.5) based on the nonlinear MF control approach developed in [78, 85, 177]:

- (i) *The infinite population limit:* In this step a Nash equilibrium is characterized by an “equilibrium relationship” between the individual strategies and the mass effect (i.e., the overall effect of the population on a given agent). This equilibrium relationship is described by the so-called MF system.
- (ii) ϵ_N -Nash equilibrium for the finite N model: The distributed continuum based MF control laws (derived from the MF system in Step 1) establish an ϵ_N -Nash equilibrium for the finite N population dynamic game flocking

model (4.2)-(4.5) where ϵ_N goes to zero asymptotically as N approaches infinity.

4.4.1. Mean Field Approximation. In a large N population system, the mean field approach suggests that the cost-coupling function for a “generic” agent i in (4.9), $\phi^N(z_i; z_{-i})$, be approximated by a deterministic function $\phi(z, \cdot)$ which only depends on $z = z_i$.

Replacing the function $\phi^N(z_i; z_{-i})$ with the deterministic function $\phi(z, \cdot)$ in the i^{th} agent’s cost function (4.8) reduces the dynamic game flocking model (4.7)-(4.8) to a set of N independent optimal control problems.

4.4.2. Preliminary Optimal Control of a Single Agent. We now consider a “single agent” Optimal Control Problem (OCP):

$$dz(t) = (Fz(t) + Gu(t))dt + Ddw(t), \quad t \geq 0, \quad (4.10)$$

$$\inf_{u \in \mathcal{U}} J(u) := \inf_{u \in \mathcal{U}} E \int_0^\infty e^{-\rho t} (\phi(z(t), t) + u_i^T(t)Ru_i(t))dt, \quad (4.11)$$

where $z(\cdot) \in \mathbb{R}^{2n}$, $u(\cdot) \in \mathbb{R}^n$ are the state and control input, respectively; $z(0)$ is given; $w(\cdot)$ denotes a p -dimensional standard Wiener processes; $\phi(z, \cdot)$ is a positive function, and \mathcal{U} is the corresponding admissible control set of the generic agent. An admissible control $u^o(\cdot) \in \mathcal{U}$ is called optimal if $J(u^o) = \inf_{u \in \mathcal{U}} J(u)$.

For $x \in \mathbb{R}^{2n}$ and $0 \leq t < \infty$ we define the value function $h(\cdot, \cdot)$ for the OCP (4.10)-(4.11) by

$$h(x, t) := \inf_{(u(s))_{s \geq t} \in \mathcal{U}} E \left[\int_t^\infty e^{-\rho(s-t)} (\phi(z(s), s) + u_i^T(s)Ru_i(s)) ds \mid z(t) = x \right].$$

By employing a standard dynamic programming argument and using Itô’s formula we get the following result (see [59]).

THEOREM 4.1. (*HJB for the OCP*) Assume that the function $\phi(z, t)$ is Lipschitz continuous with respect to z and uniformly continuous with respect to t , and assume the value function $h(z, t)$ for the OCP (4.10)-(4.11) is a C^1 function of variable t and

C^2 function of variable z , then $h(z, t)$ solves the (backward in time) Hamilton-Jacobi-Bellman (HJB) equation

$$\partial_t h(z, t) + H(z, D_z h) + \frac{1}{2} \text{tr}(DD^T D_{zz}^2 h(z, t)) + \phi(z, t) = \rho h(z, t), \quad z \in \mathbb{R}^{2n}, \quad t \geq 0, \quad (4.12)$$

with boundary condition $\lim_{t \rightarrow \infty} e^{-\rho t} h(z(t), t) = 0$, where the Hamiltonian $H(\cdot)$ is defined as

$$H(z, p) \equiv \min_{u \in \mathcal{U}} H(z, p, u) := \min_{u \in \mathcal{U}} \{(Fz + Gu)^T p + u^T R u\},$$

for p in \mathbb{R}^{2n} . □

The solution of the OCP (4.10)-(4.11) is

$$u^o(z, t) := \arg \min_{u \in \mathcal{U}} H(z, D_z h(z, t), u) = -\frac{1}{2} R^{-1} G^T D_z h(z, t). \quad (4.13)$$

Substituting $u^o(z, t)$ into (4.12) yields the (backward in time) HJB equation ($z \in \mathbb{R}^{2n}$, $t \geq 0$)

$$\begin{aligned} \partial_t h(z, t) + \left(Fz - \frac{1}{4} GR^{-1} G^T D_z h(z, t) \right)^T D_z h(z, t) \\ + \frac{1}{2} \text{tr}(DD^T D_{zz}^2 h(z, t)) + \phi(z, t) = \rho h(z, t), \end{aligned} \quad (4.14)$$

with boundary condition $\lim_{t \rightarrow \infty} e^{-\rho t} h(z(t), t) = 0$.

4.4.2.1. *Fokker-Planck-Kolmogorov (FPK) Equation.* Under the state feedback optimal control law $u^o(z, t)$ given in (4.13), the evolution of the density $f(z, \cdot)$ of the generic agent (4.10) satisfies the (forward in time) Fokker-Planck-Kolmogorov (FPK) equation ($z \in \mathbb{R}^{2n}$, $t \geq 0$)

$$\partial_t f(z, t) + D_z^T \left(\left(Fz - \frac{1}{2} GR^{-1} G^T D_z h(z, t) \right) f(z, t) \right) = \frac{1}{2} \text{tr}(DD^T D_{zz}^2 f(z, t)), \quad (4.15)$$

with initial condition $f(z, 0) \geq 0$. We note that $h(z, \cdot)$ in (4.15) is the solution of the HJB equation (4.14). Let us assume the boundary condition $\lim_{\|z\| \rightarrow \infty} f(z, t) = 0$ for all $t \geq 0$.

4.4.2.2. *Nonlinear Cost-Coupling (CC) Function.* For a generic agent i , the Law of Large Numbers (LLN) suggests the approximation of the Cost-Coupling (CC) function $\phi^N(z_i, z_{-i}^o)$ in (4.9) (where z_{-i}^o is the state of all agents $\{j : 1 \leq j \leq N\}$ distinct from agent i which evolve according to the SDEs (4.7) with optimal control laws $u_i^o(\cdot) := u^o(z, \cdot)|_{z=z_i}$) for a large N population system by

$$\bar{\phi}(z_i, t) \equiv \bar{\phi}(x_i, v_i, t) = \left\| \frac{\int_{\mathbb{R}^{2n}} a(\|x_i - x'\|)(v' - v_i)f(x', v', t)dx'dv'}{\int_{\mathbb{R}^{2n}} a(\|x_i - x'\|)f(x', v', t)dx'dv'} \right\|^2, \quad (4.16)$$

where $f(x, v, \cdot) \equiv f(z, \cdot)$ is the solution of the equation (4.15), and the function $a(\|x - x'\|)$ is defined in (4.4).

4.4.3. The Mean Field System.

Let

$$f_N(x, v, 0) := \frac{1}{N} \sum_{i=1}^N \delta(x - x_i(0))\delta(v - v_i(0)),$$

be the initial empirical density function associated with N agents where δ is the Dirac delta. We enunciate the following assumption:

(A4.1) We assume that: (i) the sequence of initial conditions $\{(x_i(0), v_i(0)) : 1 \leq i \leq N\}$ has a compactly supported probability density $f(x, v, 0) \equiv f_0(x, v)$ such that $\int_{\mathcal{A}} f_0(x, v)dx dv = 1$ where \mathcal{A} is a compact interval containing all $(x_i(0), v_i(0))$, $1 \leq i \leq N$, and (ii) $\{f_N(x, v, 0) : N \geq 1\}$ converges weakly to $f_0(x, v)$ almost surely, i.e., for any $\psi(x, v) \in C_b^{2n}$ (the space of bounded continuous functions on \mathbb{R}^{2n}),

$$\lim_{N \rightarrow \infty} \int \psi(x, v)f_N(x, v, 0)dx dv = \int \psi(x, v)f_0(x, v)dx dv, \quad (a.s.).$$

□

REMARK 4.2. *If the sequence $\{z_i(0) \equiv [x_i(0), v_i(0)]^T : 1 \leq i \leq N\}$ is generated by independent random observations on the density function $f_0(x, v)$, then (A4.1)-(ii) holds with probability one by the Strong Law of Large Numbers or the Glivenko-Cantelli theorem [42].*

We now aim to construct the equilibrium relationship (between the individual strategies and the mass effect) in the stochastic MF control theory. For nonlinear MF stochastic control problems, a general formulation using equations of the McKean-Vlasov type is given in [85]. However, for the synchronization of coupled oscillators formulated as a game problem a compact system of coupled MF equations is given in [177] within a nonlinear SDE problem formulation.

The key idea of the MF control methodology is to prescribe a spatially averaged mass function $\bar{\phi}(z, \cdot)$ characterized by the property that it is reproduced as the average of all agents' states in the continuum of agents whenever each individual agent optimally tracks the same mass function $\bar{\phi}(z, \cdot)$.

Applying the nonlinear MF stochastic control approach (developed in [78, 85, 177]) to our dynamic game flocking model (4.7)-(4.8) in the infinite population limit (or (4.10)-(4.11) for a generic agent) yields the nonlinear *continuum based MF system* ($z \in \mathbb{R}^{2n}$, $t \geq 0$):

$$\begin{aligned} \text{[MF-HJB]} \quad \partial_t h(z, t) + \left(Fz - \frac{1}{4}GR^{-1}G^T D_z h(z, t) \right)^T D_z h(z, t) \\ + \frac{1}{2} \text{tr}(DD^T D_{zz}^2 h(z, t)) + \bar{\phi}(z, t) = \rho h(z, t), \end{aligned} \quad (4.17)$$

$$\begin{aligned} \text{[MF-FPK]} \quad \partial_t f(z, t) + D_z^T \left(\left(Fz - \frac{1}{2}GR^{-1}G^T D_z h(z, t) \right) f(z, t) \right) \\ = \frac{1}{2} \text{tr}(DD^T D_{zz}^2 f(z, t)), \end{aligned} \quad (4.18)$$

$$\text{[MF-CC]} \quad \bar{\phi}(z, t) \equiv \bar{\phi}(x, v, t) = \left\| \frac{\int_{\mathbb{R}^{2n}} a(\|x - x'\|)(v' - v) f(x', v', t) dx' dv'}{\int_{\mathbb{R}^{2n}} a(\|x - x'\|) f(x', v', t) dx' dv'} \right\|^2, \quad (4.19)$$

where $f(z, 0) \equiv f_0(x, v)$ is the given initial population density, necessarily we have $\int_{\mathbb{R}^{2n}} f(z, t) dz = 1$ for any $t \geq 0$, and it is assumed that $\lim_{\|z\| \rightarrow \infty} f(z, t) = 0$ and $\lim_{t \rightarrow \infty} e^{-\rho t} h(z(t), t) = 0$.

The system of equations (4.17)-(4.19) consists of: (i) the nonlinear (backward in time) *MF-HJB* equation (4.14) which describes the HJB equation of a generic agent's discounted optimal problem (4.10)-(4.11) with cost coupling $\bar{\phi}(z, \cdot)$, (ii) the

nonlinear (forward in time) *MF-FPK* equation (4.15) which describes the evolution of the population density with the best response state feedback control law

$$u^o(z, t) := -\frac{1}{2}R^{-1}G^T D_z h(z, t), \quad z \in \mathbb{R}^{2n}, \quad t \geq 0, \quad (4.20)$$

and (iii) the spatially averaged *MF-CC* function presented in (4.16).

The MF system (4.17)-(4.19) with respect to the position and velocity variables takes the following form ($x, v \in \mathbb{R}^n, t \geq 0$)

$$\begin{aligned} \text{[MF-HJB]} \quad \partial_t h(x, v, t) + v^T D_x h(x, v, t) - \frac{1}{4}R^{-1} \|D_v h(x, v, t)\|^2 \\ + \frac{1}{2}\text{tr}(CC^T D_{vv}^2 h(x, v, t)) + \bar{\phi}(x, v, t) = \rho h(x, v, t), \end{aligned} \quad (4.21)$$

$$\begin{aligned} \text{[MF-FPK]} \quad \partial_t f(x, v, t) + v D_x^T f(x, v, t) - \frac{1}{2}R^{-1} D_v^T \left((D_v h(x, v, t)) f(x, v, t) \right) \\ = \frac{1}{2}\text{tr}(CC^T D_{vv}^2 f(x, v, t)), \end{aligned} \quad (4.22)$$

$$\text{[MF-CC]} \quad \bar{\phi}(x, v, t) = \left\| \frac{\int_{\mathbb{R}^{2n}} a(\|x - x'\|)(v' - v) f(x', v', t) dx' dv'}{\int_{\mathbb{R}^{2n}} a(\|x - x'\|) f(x', v', t) dx' dv'} \right\|^2, \quad (4.23)$$

where $f(x, v, 0) \equiv f_0(x, v)$ is the given initial population density, necessarily we have $\int_{\mathbb{R}^{2n}} f(x, v, t) dx dv = 1$ for any $t \geq 0$, and it is assumed that

$$\lim_{\|x\| \text{ or } \|v\| \rightarrow \infty} f(x, v, t) = 0, \quad \lim_{t \rightarrow \infty} e^{-\rho t} h(x(t), v(t), t) = 0.$$

4.5. Maxwellian Stationary Solution

The MF system (4.21)-(4.23) in stationary form is

$$\begin{aligned} v^T D_x h_\infty(x, v) - \frac{1}{4} R^{-1} \|D_v h_\infty(x, v)\|^2 \\ + \frac{1}{2} \text{tr}(CC^T D_{vv}^2 h_\infty(x, v)) + \bar{\phi}_\infty(x, v) = \rho h_\infty(x, v), \end{aligned} \quad (4.24)$$

$$\begin{aligned} v D_x^T f_\infty(x, v) - \frac{1}{2} R^{-1} D_v^T \left((D_v h_\infty(x, v)) f_\infty(x, v) \right) \\ = \frac{1}{2} \text{tr}(CC^T D_{vv}^2 f_\infty(x, v)), \end{aligned} \quad (4.25)$$

$$\bar{\phi}_\infty(x, v) = \left\| \frac{\int_{\mathbb{R}^{2n}} a(\|x - x'\|) (v' - v) f_\infty(x', v') dx' dv'}{\int_{\mathbb{R}^{2n}} a(\|x - x'\|) f_\infty(x', v') dx' dv'} \right\|^2, \quad (4.26)$$

where the density $f_\infty(x, v)$ satisfies $\int_{\mathbb{R}^{2n}} f_\infty(x, v) dx dv = 1$, and

$$\lim_{t \rightarrow \infty} e^{-\rho t} v_\infty(x(t), v(t)) = 0.$$

The stationary equation system (4.24)-(4.26) is related to (4.21)-(4.23) by the fact that the steady-state population density of the system, $f_\infty(x, v) := \lim_{t \rightarrow \infty} f(x, v, t)$, gives a time independent cost-coupling $\bar{\phi}_\infty(x, v)$ in (4.19) which yields a time independent solution $h_\infty(x, v)$ to the MF-HJB equation (4.21). Furthermore, $f_\infty(x, v)$ and $h_\infty(x, v)$ solve the stationary MF-FPK equation (4.25).

Let $\Pi > 0$ be the unique solution of the deterministic algebraic Riccati equation

$$- \Pi R^{-1} \Pi + I = \rho \Pi.$$

THEOREM 4.2. *Assume that the weight function $a(\|x\|)$ is integrable, i.e.,*

$$\int_{\mathbb{R}^n} a(\|x\|) dx < \infty,$$

and $\Sigma = CC^T > 0$. Then for any arbitrary $\mu \in \mathbb{R}^n$, there exists the following solution to the system (4.24)-(4.26):

$$\begin{aligned} h_\infty(v) &= (v - \mu)^T \Pi (v - \mu) + \eta, & \eta &= \text{tr}(\Sigma)/\rho, \\ f_\infty(v) &= \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(v - \mu)^T \Sigma^{-1} (v - \mu)\right), \\ \bar{\phi}_\infty(v) &= \|v - \mu\|^2. \end{aligned}$$

Proof. The assertion of the theorem is straightforward to verify by substituting $h_\infty(v)$, $f_\infty(v)$ and $\bar{\phi}_\infty(v)$ into the system (4.24)-(4.26). \square

We note that the stationary solution $f_\infty(v)$ is a Maxwellian distribution function with arbitrary μ bulk velocity. The vector μ may be chosen as the initial velocity population mean, i.e., $\mu = \int_{\mathbb{R}^{2n}} v f_0(x, v) dx dv$. The Maxwellian distribution function is a desirable steady-state solution of flocking models (see e.g. (1.5) in [50]).

REMARK 4.3. We note that the following weights satisfy the integrability condition: (i) the CS weights $a(\|x\|) = \frac{1}{(1+\|x\|^2)^\beta}$ in (4.4) for $\beta > 1$, and (ii) the Gaussian weights $a(\|x\|) = \exp(-\alpha\|x\|^2)$ for $\alpha > 0$.

4.6. ϵ -Nash Equilibrium Property

We shall assume that the MF system (4.17)-(4.19) has a unique solution

$$(h(\cdot, \cdot), f(\cdot, \cdot), \bar{\phi}(\cdot, \cdot)).$$

In a finite N population system we assume that the i^{th} agent applies the continuum (i.e., infinite population) based MF control input:

$$u_i^o(t) := u^o(z, t)|_{z=z_i(t)} = -\frac{1}{2}R^{-1}G^T D_z h(z, t)|_{z=z_i(t)}, \quad t \geq 0, \quad (4.27)$$

where $h(z, t)$ is the solution of the MF-HJB equation (4.17). Hence, the closed-loop MF dynamics of the i^{th} agent in the finite N population system is

$$dz_i^o(t) = (Fz_i^o(t) - \frac{1}{2}GR^{-1}G^T D_z h(z, t)|_{z=z_i(t)})dt + Ddw_i(t), \quad (4.28)$$

or

$$\begin{cases} dx_i^o(t) = v_i^o(t)dt, \\ dv_i^o(t) = -\frac{1}{2}R^{-1}\nabla_v h(v, t)|_{v=v_i^o(t)}dt + Cdw_i(t), \end{cases}$$

with initial conditions $x_i^o(0) = x_i(0)$ and $v_i^o(0) = v_i(0)$.

We recall the following definition from Chapter 2 (see Definition 2.2).

DEFINITION 4.1. [79] *Given $\epsilon > 0$, the set of controls $\{u_i^o \in \mathcal{U}_i : 1 \leq i \leq N\}$ for N agents generates an ϵ -Nash equilibrium with respect to the costs $\{J_i^N : 1 \leq i \leq N\}$, if $J_i^N(u_i^o, u_{-i}^o) - \epsilon \leq \inf_{u_i \in \mathcal{U}_i} J_i^N(u_i, u_{-i}^o) \leq J_i^N(u_i^o, u_{-i}^o)$, for any $1 \leq i \leq N$. \square*

Let

$$f_N(x, v, t) := \frac{1}{N} \sum_{i=1}^N \delta(x - x_i^o(t)) \delta(v - v_i^o(t)),$$

be the empirical density function associated with N agents with MF control inputs.

We denote

$$\begin{aligned} (\epsilon_N(x_i))^2 := E \int_0^\infty e^{-\rho t} & \left(\frac{\int_{\mathbb{R}^{2n}} a(\|x_i - x\|) v f_N(x, v, t) dx dv}{\int_{\mathbb{R}^{2n}} a(\|x_i - x\|) f_N(x, v, t) dx dv} \right. \\ & \left. - \frac{\int_{\mathbb{R}^{2n}} a(\|x_i - x\|) v f(x, v, t) dx dv}{\int_{\mathbb{R}^{2n}} a(\|x_i - x\|) f(x, v, t) dx dv} \right) dt, \end{aligned}$$

where x_i is the solution of the dynamics (4.2) under the admissible control $u_i \in \mathcal{U}_i$ such that $J_i^N(u_i, u_{-i}^o) \leq J_i^N(u_i^o, u_{-i}^o)$. This restriction causes no loss of generality in the Theorem below since, other wise the control u_i , will generate a cost higher than $J_i^N(u_i^o, u_{-i}^o)$. Under Assumption **(A4.1)** one can show that $\lim_{N \rightarrow \infty} \epsilon_N(x_i) = 0$ (i.e., a subsequence of f_N converges weakly to f) using the Prohorov's theorem [19]. This yields the following result:

THEOREM 4.3. *Assume **(A4.1)** holds. Then the set of MF control laws for the finite population system $\{u_i^o \in \mathcal{U}_i : 1 \leq i \leq N\}$ given in (4.27) generates an ϵ_N -Nash equilibrium such that*

$$J_i^N(u_i^o, u_{-i}^o) - \epsilon_N \leq \inf_{u_i \in \mathcal{U}_i} J_i^N(u_i, u_{-i}^o) \leq J_i^N(u_i^o, u_{-i}^o), \quad 1 \leq i \leq N,$$

where $\lim_{N \rightarrow \infty} \epsilon_N = 0$. □

4.7. Chapter Summary

This work is concerned with the synthesis of a controlled flocking model via Mean Field (MF) control theory. In this problem formulation the state of each agent consists of both its position and its controlled velocity such that: (i) all agents have similar stochastic dynamics, and (ii) each agent seeks to minimize by continuous state feedback its individual discounted cost functions involving a nonlinear (relative distance based) weighted mean of the velocity states of all other agents. The cost functions are based on the normalized Cucker-Smale (CS) flocking algorithm in its original uncontrolled formulation. For this dynamic game problem, the MF system of equations which consists of coupled deterministic HJB and FPK equations is derived approximating the stochastic system of agents as the population size goes to infinity.

Subject to the existence of a unique solution to the MF system of equations: (i) the stationary solution of the MF system of equations is a Maxwellian distribution function, (ii) the set of MF control laws for the system possesses an ϵ_N -Nash equilibrium property where ϵ_N goes to zero as the population size N approaches infinity.

This model may be regarded as a controlled game theoretic formulation of a flocking behaviour in which each agent, instead of responding to an ad-hoc algorithm, obtains its control law from a game theoretic Nash equilibrium.

A topic of future research is the stability analysis of the MF system near the Maxwellian distribution function based on the approach of Chapter 3 (see [132]).

CHAPTER 5

Mean Field LQG Control in Leader-Follower Stochastic Multi-Agent Systems: Likelihood Ratio Based Adaptation

This chapter studies large population leader-follower stochastic multi-agent systems where the agents have linear stochastic dynamics and are coupled via their quadratic cost functions. The cost of each leader is based on a trade-off between moving toward a certain reference trajectory which is unknown to the followers and staying near their own centroid. On the other hand, followers react by tracking a convex combination of their own centroid and the centroid of the leaders. We approach this large population dynamic game problem by mean field linear-quadratic-Gaussian (MF LQG) stochastic control theory. In this model, followers are adaptive in the sense that they use a likelihood ratio estimator (on a sample population of the leaders' trajectories) to identify the member of a given finite class of models which is generating the reference trajectory of the leaders. Under appropriate conditions, it is shown that the true reference trajectory model is identified by each follower in finite time with probability one as the leaders' population goes to infinity. Furthermore, we show that the resulting sets of mean field (MF) control laws for both leaders and adaptive followers

possess an almost sure ϵ_N -Nash equilibrium property for a system with population N where ϵ_N goes to zero as N goes to infinity. Numerical experiments are presented illustrating the results.

5.1. Introduction

Decision making and collective behaviour often involve some form of leader-follower behaviour. This behaviour is observed in humans [54] and many other species in nature [43, 122], and is studied in a variety of disciplines such as game theory [161], distributed networks [173], crowd flow dynamics [12] and biology [43], among others. Such behaviour in nature is often attributed to the fact that there exist some individuals in the group which have more information than others, for instance the location of resources or migratory routes [43].

In this chapter we study large population Leader-Follower (L-F) stochastic multi-agent systems where the agents have linear stochastic dynamics and are coupled via their quadratic cost functions. The cost of each leader is based on a trade-off between moving toward a certain reference trajectory which is unknown to the followers and staying near their own centroid. On the other hand, followers react by tracking a convex combination of their own centroid and the centroid of the leaders. Here, as in most practical leader-follower modelling of multi-agent systems, the leaders ignore the followers, but the followers' behaviours are influenced by the leaders. The model in this chapter is a generalization of that in [106] to the case of collective dynamics which include leaders, followers and an unknown (to the followers) reference trajectory for the leaders.

We approach the large population L-F model by mean field linear-quadratic-Gaussian (MF LQG) stochastic control theory. In this framework the computation of the followers' control laws requires knowledge of the complete reference trajectory of the leaders which is in general not known to the followers, hence a likelihood ratio based adaptation scheme is proposed. The main contributions of this chapter are as follows: (i) A likelihood ratio based adaptation algorithm (on a sample population of

the leaders' trajectories) is employed by the adaptive followers to identify the member of a given finite class of models which is generating the reference trajectory of the leaders. Under appropriate conditions, it is shown that the true reference trajectory model is identified in finite time with probability one by each follower as the leaders' population goes to infinity. (ii) A demonstration that the use of the resulting mean field (MF) control laws yields a set of leaders and adaptive followers' control laws possessing an almost sure (*a.s.*) ϵ_N -Nash equilibrium property, where ϵ_N goes to zero as the population size N goes to infinity.

The implementation of the overall MF control laws for the leaders and followers has the following form: (i) Each leader enacts an MF control law which consists of the feedback of its own local stochastic state and the precomputed leaders' deterministic mass effect. (ii) Each follower enacts an adaptive MF control law which consists of the feedback of its own local stochastic state and the estimation based mass effects of the leaders and followers.

In [140] we first developed a non-adaptive but general model with weighted couplings in the leaders and followers' cost functions (which depended on the locality parameters of the agents). [140] also presents the main adaptation result of the uniform cost coupling model in the case that the followers "only" track the centroid of the leaders. Subsequently, in [139] the optimality property of the (tracking like) adaptive followers' MF control laws is studied. In this chapter we present a complete analysis of a more general (and realistic) scenario where the followers are tracking a convex combination of their own centroid and the centroid of the leaders [137]. Hence, we have an ϵ -Nash equilibrium property for the adaptive followers' MF control laws.

In the standard consensus literature the agents have little a priori information but communicate over possibly time varying graphs, then under connectivity assumptions (e.g., the union of the interaction graphs for the system is connected frequently enough as the system evolves) the agents reach consensus (see [154], among many other papers). By contrast, the leader-follower agents in this chapter possess a priori data on the overall system; the leaders observe no one and the followers have limited

observations on the leaders and a priori data on the possible trajectory scenarios of the leaders. This permits the computation by each follower agent of different tracking scenarios amongst which it chooses one at each instant, depending upon the observations received. Therefore, the agents in the model considered in this chapter do not require communication except the limited observations on the leaders by the followers.

It is to be noted that the formulation in this chapter does not include any collision avoidance or formation control between the agents beyond the optimal tracking property since the states do not necessarily correspond to positions in space. We further note that if the states are given a spatial interpretation the inherently stochastic volatility of the dynamics prevents any state from converging onto another state.

The organization of this chapter is as follows. Section 5.2 is dedicated to the problem formulation, terminology and some applications of the model in multi-vehicle coordination control and economics (finance). The MF LQG systems of the L-F problem are derived and analyzed in Section 5.3. In Section 5.4 we present the estimation procedure for the followers. The stability analysis of the MF control laws and the adaptive MF algorithm for the followers are presented in Section 5.5. The optimality properties of the MF control laws for both the leaders and the adaptive followers are given in Section 5.6, and Section 5.7 presents sample numerical simulations of the model. Concluding remarks are stated in Section 5.8.

5.2. Problem Formulation, Terminology and some Applications

The following notation will be used in this chapter. We use the integer valued subscript as the label for a certain agent of the population and superscripts L and F for a leader and follower agent, respectively. In addition, an overline denotes the expected value of a random variable, i.e., $\bar{z}(t) := Ez(t)$. $\|\cdot\|$ denotes the 2-norm of vectors and $\|\cdot\|_\infty$ denotes the infinity or sup norm. $\|z\|_Q := (z^T Q z)^{1/2}$ for any appropriate dimension vector z and matrix $Q \geq 0$. A^T denotes the transpose of a

vector or matrix A and $\text{tr}(A)$ denotes the trace of a square matrix A . Let C_n be the family of all n -dimensional continuous functions on \mathbb{R}_+ and $C_n^b := \{f \in C_n : \|f\|_\infty := \sup_{t \geq 0} \|f(t)\| < \infty\}$. Note that C_n^b is a Banach space under the norm $\|\cdot\|_\infty$.

Let L denote the countably infinite set of leaders,

$$L^{N_L} := \{L_1, \dots, L_{N_L}\} \subset L,$$

of cardinality N_L , and similarly for the set of countably infinite followers F and the subset $F^{N_F} := \{F_1, \dots, F_{N_F}\} \subset F$ of cardinality N_F . We assume that $L \cap F = \emptyset$.

5.2.1. Leaders Stochastic LQG Dynamic Game Model. The dynamics for the N_L leaders are given by

$$dz_i^L = (A_i z_i^L + B_i u_i^L) dt + C_i dw_i^L, \quad t \geq 0, \quad 1 \leq i \leq N_L \quad (5.1)$$

where $z_i^L \in \mathbb{R}^n$ is the state, $u_i^L \in \mathbb{R}^m$ is the control input, and $\{w_i^L : 1 \leq i \leq N_L\}$ denotes a set of independent (i.e., independent) p -dimensional standard Wiener processes. The matrices A_i , B_i and C_i have compatible dimensions.

Let $\theta_i := [A_i, B_i, C_i]$ be defined as the dynamical parameter associated with leader i , $1 \leq i \leq N_L$, where we assume that θ_i , for all $1 \leq i \leq N_L$, are in the compact set Θ_L . The initial states $\{z_i^L(0) : 1 \leq i \leq N_L\}$ are assumed to be independent and also independent of $\{w_i^L : 1 \leq i \leq N_L\}$. In addition we assume that $\sup_{1 \leq i \leq N_L} E \|z_i^L(0)\|^2 < \infty$.

The admissible control set for leader i , $1 \leq i \leq N_L$, is given by

$$\mathcal{U}_i^L := \left\{ u_i^L : u_i^L(t) \text{ is adapted to sigma-field } \sigma(z_j^L(s), s \leq t, 1 \leq j \leq N_L), \right. \\ \left. \|z_i^L(T)\| = o(\sqrt{T}), \int_0^T \|z_i^L(t)\|^2 dt = O(T) \text{ a.s.} \right\}.$$

Intuitively, the admissible controls are the controls that do not use any information about future increments of the driving noise processes, which is essentially a causality requirement in stochastic systems.

The cost function of the leaders is based on a trade-off between moving towards a common reference trajectory and keeping cohesion of the flock of leaders by also tracking their centroid. We let

$$\phi^L(z^{L,N_L})(\cdot) := \lambda h(\cdot) + (1 - \lambda)z^{L,N_L}(\cdot) \quad (5.2)$$

where λ is a scalar in $(0, 1)$, $h \in C_n^b$ is a reference trajectory known to all the leaders, and $z^{L,N_L}(\cdot) := 1/N_L \sum_{i=1}^{N_L} z_i^L(\cdot)$ is the centroid of the leaders. The objective of each individual leader i ($1 \leq i \leq N_L$) is to minimize its Long Time Average (LTA) (i.e., ergodic) cost function given by

$$J_i^{L,N_L}(u_i^L; u_{-i}^L) := \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(\|z_i^L - \phi^L(z^{L,N_L})\|_Q^2 + \|u_i^L\|_R^2 \right) dt \quad (5.3)$$

where the matrices Q and R are symmetric positive semi-definite and symmetric positive definite, respectively, with compatible dimensions, and

$$u_{-i}^L := (u_1^L, \dots, u_{i-1}^L, u_{i+1}^L, \dots, u_{N_L}^L).$$

To indicate the dependence of J_i^L on $u_i^L(\cdot)$, $u_{-i}^L(\cdot)$ and the leaders' population size N_L , we write it as $J_i^{L,N_L}(u_i^L; u_{-i}^L)$. Note that the leaders' mean field cost coupling (5.2) is the same as mean field couplings in the basic models considered in [79], but with time-varying offset term $h(\cdot)$. If $\lambda = 1$ then the leaders become independent such that each leader is interested in optimally tracking $h(\cdot)$.

5.2.2. Followers Stochastic LQG Dynamic Game Model. Similarly, the dynamics for the N_F followers are given by

$$dz_i^F = (A_i z_i^F + B_i u_i^F) dt + C_i dw_i^F, \quad t \geq 0, \quad 1 \leq i \leq N_F \quad (5.4)$$

where $z_i^F \in \mathbb{R}^n$ is the state, $u_i^F \in \mathbb{R}^m$ is the control input, and $\{w_i^F : 1 \leq i \leq N_F\}$ denotes a set of independent p -dimensional standard Wiener processes independent of both $\{w_i^L : 1 \leq i \leq N_L\}$ and $\{z_i^L(0) : 1 \leq i \leq N_L\}$. The matrices A_i , B_i and C_i have compatible dimensions.

Let $\theta_i := [A_i, B_i, C_i]$ be defined as the dynamical parameter associated with follower i , $1 \leq i \leq N_F$, where we assume that θ_i , $1 \leq i \leq N_F$, are in the compact set Θ_F . The initial states $\{z_i^F(0) : 1 \leq i \leq N_F\}$ are assumed to be independent and also independent of the independent $\{w_i^F : 1 \leq i \leq N_F\}$, $\{w_i^L : 1 \leq i \leq N_L\}$, and $\{z_i^L(0) : 1 \leq i \leq N_L\}$. In addition we assume that $\sup_{1 \leq i \leq N_F} E \|z_i^F(0)\|^2 < \infty$.

The admissible control set for follower i , $1 \leq i \leq N_F$, is given by

$$\mathcal{U}_i^F := \left\{ u_i^F : u_i^F(t) \text{ is adapted to } \sigma(z_j^F(s), z_k^L(s), s \leq t, 1 \leq j \leq N_F, 1 \leq k \leq N_L), \right. \\ \left. \|z_i^F(T)\| = o(\sqrt{T}), \int_0^T \|z_i^F(t)\|^2 dt = O(T) \text{ a.s.} \right\}.$$

The followers react by tracking a convex combination of their own centroid and the centroid of the leaders. We let

$$\phi^F(z^{L, N_L}, z^{F, N_F})(\cdot) := \eta z^{L, N_L}(\cdot) + (1 - \eta) z^{F, N_F}(\cdot) \quad (5.5)$$

where η is a scalar in $(0, 1)$, $z^{F, N_F}(\cdot) := 1/N_F \sum_{i=1}^{N_F} z_i^F(\cdot)$ is the centroid of the followers, and z^{L, N_L} is the centroid of the leaders defined in (5.2). The LTA cost function for an individual follower i ($1 \leq i \leq N_F$) is given by

$$J_i^{F, N}(u_i^F; u_{-i}^F, u^L) := \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(\|z_i^F - \phi^F(z^{L, N_L}, z^{F, N_F})\|_Q^2 + \|u_i^F\|_R^2 \right) dt \quad (5.6)$$

where $u_{-i}^F := (u_1^F, \dots, u_{i-1}^F, u_{i+1}^F, \dots, u_{N_F}^F)$, $u^L := (u_1^L, \dots, u_{N_L}^L)$, $N = N_L + N_F$ is the population size of the system, and Q and R are defined in (5.3). To indicate the dependence of J_i^F on $u_i^F(\cdot)$, $u_{-i}^F(\cdot)$, $u^L(\cdot)$ and the population size of the system N , we write it as $J_i^{F, N}(u_i^F; u_{-i}^F, u^L)$.

We note that in this model: (i) the leaders are coupled to each other through their cost functions and respond to each other and their reference trajectory, and (ii) the followers attempt to track the convex combination of both their own centroid and the centroid of the leaders. These are captured in the two types of MF equation systems in Section 5.3.

5.2.3. Followers Observation Processes. We assume that all adaptive followers observe a random fraction of the leaders' trajectories with some added noise through their individual observation processes. More precisely, we assume each adaptive follower i , $1 \leq i \leq N_F < \infty$, observes a non-empty random subset $O_i \subset L$ of size $|O_i| := M_i \leq N_L < \infty$ of the leaders' trajectories through the process $y_i^{M_i}(\cdot)$ which is described in terms of the stochastic differential equations

$$dy_i^{M_i} = \left(\frac{1}{M_i} \sum_{j \in O_i} z_j^L \right) dt + \frac{1}{M_i} \sum_{j=1}^{M_i} D_j^i dv_j^i, \quad t \geq 0 \quad (5.7)$$

where $y_i^{M_i} \in \mathbb{R}^n$ and $\{v_j^i : 1 \leq i \leq N_F, 1 \leq j \leq M\}$ is a set of independent standard Wiener processes independent of $\{w_i^L : 1 \leq i \leq N_L\}$, $\{w_i^F : 1 \leq i \leq N_F\}$, $\{z_i^L(0) : 1 \leq i \leq N_L\}$ and $\{z_i^F(0) : 1 \leq i \leq N_F\}$. The set of constant matrices $\{D_j^i : 1 \leq i \leq N_F, 1 \leq j \leq M\}$ has compatible dimensions.

The sets $O_i \subset L$, $1 \leq i \leq N_F < \infty$, of cardinality M_i are chosen a priori with $M_i = M(N_L) = \lfloor \sqrt{N_L} \rfloor < N_L < \infty$ (where $\lfloor r \rfloor$ denotes greatest integer less than or equal to r) for $1 \leq i \leq N_F < \infty$, where $N_F \rightarrow \infty$ and $N_L \rightarrow \infty$ as the number of followers and leaders respectively goes to infinity. The sets $O_i \subset L$, $1 \leq i < \infty$, are chosen independently, and independent of all initial states and Wiener processes by uniformly distributed selections on the set of the leaders, L . We underline that the magnitude $\lfloor \sqrt{N_L} \rfloor$ is chosen simply for definition of modelling and exposition; any other integer value function $M(N_L)$ satisfying $M(N_L) \rightarrow \infty$ and $M(N_L)/N_L \rightarrow 0$, as $N_L \rightarrow \infty$, may be used for the theory in this chapter.

We also assume that, as the prior information of the followers, the reference trajectory of the leaders $h(\cdot)$ is parameterized by $\delta \in \Delta$ where Δ is a finite set and $h_\delta \in C_n^b$ for any $\delta \in \Delta$. This is only for the followers, and $h(\cdot)$ is fully known by the leaders.

5.2.4. Applications. The leader-follower modelling of this chapter is motivated by many practical problems in which some agents in a group have more information than the others.

A typical application is in multi-vehicle coordination control or in vehicle platooning (see [151]) where the aim is that the states (e.g., the velocities) of all vehicles approach a reference signal (which could be an exogenous signal or which could evolve according to a dynamic model). In many realistic situations some vehicles (the leaders) have complete access to the reference trajectory, and the other vehicles (followers) do not have access to this trajectory and need to estimate it. We can formulate this multi-agent model as a leader-follower MF LQG problem considered in this chapter where the followers need to identify the member of a given finite class of models which is generating the reference trajectory.

Another application of the model is leader-follower dynamic version of Keynes' beauty contest games in economics (finance). Keynes proposed beauty contest games where a newspaper would print some photographs and people would vote for the prettiest faces. Everyone who picked the most popular face automatically entered a lottery to win a prize. Keynes remarked that the stock market is similar to beauty contest games where each investor would like to guess the other investors' guesses (see Example 1 in [2]). A similar approach to MF stochastic control is considered in [2] to study large population static aggregative games such as Keynes' beauty contest games. Now we formulate a leader-follower LQG dynamic version of Keynes' beauty contest games. Here we consider a large population of players divided into two groups: (i) the leaders as large well-informed players (e.g., institutional investors in the stock market), and (ii) the followers (e.g., retail investors in the stock market). The state of each player is its publicly announced prediction of the prettiest face where $z_i^L(\cdot)$ denotes the state of the i -th leader ($1 \leq i \leq N_L$) and $z_i^F(\cdot)$ denotes the state of the i -th follower ($1 \leq i \leq N_F$). The leaders and followers have linear stochastic dynamics given in (5.1) and (5.4) with different classes of parameters Θ_L and Θ_F . The average prediction of the leaders and followers are given by their centroids $z^{L,N_L}(\cdot) := (1/N_L) \sum_{i=1}^{N_L} z_i^L(\cdot)$ and $z^{F,N_F}(\cdot) := (1/N_F) \sum_{i=1}^{N_F} z_i^F(\cdot)$, respectively. Based on the quadratic payoff functions considered in [2], we formulate cost functions of the agents as follows. The leaders would like to minimize their cost functions (5.3) based on a

trade-off between making guesses close to the exogenous private informative signal h (which is unknown to the followers) and guessing close to their own average prediction $z^{L,N_L}(\cdot)$. On other hand, the followers would like to guess close to some convex combination of their own average prediction $z^{F,N_F}(\cdot)$ and the average prediction of the leaders $z^{L,N_L}(\cdot)$ (see (5.6)).

There are many other similar applications of the model considered in this chapter in flocking [63], formation control [109], dynamic industry models [174], and social opinion models with a very large number of leaders (e.g., important members of a party) and followers [53].

5.3. Mean Field LQG Stochastic Control Theory

5.3.1. Preliminary LQG Optimal Control of a Single Agent. In this section first we consider a single agent with linear stochastic dynamics

$$dz = (Az + Bu)dt + Cdw, \quad t \geq 0 \quad (5.8)$$

where $z \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the control input, w denotes a p -dimensional standard Wiener process, and $z(0)$ is given. The matrices A , B and C have compatible dimensions. The initial condition $z(0)$ is independent of the process w .

Denote the admissible control set (see [106])

$$\mathcal{U} := \left\{ u : u(\cdot) \text{ is adapted to } \sigma(z(0), w(s), s \leq t), \right. \\ \left. \|z(T)\| = o(\sqrt{T}), \int_0^T \|z(t)\|^2 dt = O(T) \text{ a.s.} \right\}.$$

For $u(\cdot) \in \mathcal{U}$, let the LTA cost function be given by

$$J(u(\cdot)) := \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(\|z - \phi\|_Q^2 + \|u\|_R^2 \right) dt \quad (5.9)$$

where $\phi(\cdot) \in C_n^b$ is a known function, Q and R are, respectively, symmetric positive semi-definite and symmetric positive definite matrices with compatible dimensions.

THEOREM 5.1. (*Special case of [106]*) For the LQG optimal control problem (5.8)-(5.9), assume (i) $[A, B]$ is stabilizable, (ii) $[A, Q^{1/2}]$ is detectable, and (iii) $\phi(\cdot) \in C_n^b$. Then we have:

- (a) The algebraic Riccati equation $\Pi A + A^T \Pi - \Pi B R^{-1} B^T \Pi + Q = 0$ has a unique positive semi-definite solution Π .
- (b) $\Gamma := A - B R^{-1} B^T \Pi$ is asymptotically stable.
- (c) The differential equation $ds/dt = -(A - B R^{-1} B^T \Pi)^T s + Q \phi$ has a unique solution in C_n^b :

$$s(t) = - \int_t^\infty e^{-\Gamma(t-\tau)} Q \phi(\tau) d\tau, \quad t \geq 0.$$

- (d) The optimal control law: $u^o(\cdot) := \arg \inf_{u(\cdot) \in \mathcal{U}} J(u(\cdot)) = -R^{-1} B^T (\Pi z(\cdot) + s(\cdot))$.

Proof. See Theorem 3.1 in [106]. □

5.3.2. The MF LQG Systems. Considering the L-F model (5.1)-(5.6) we need to solve a set of tracking optimal control problems where the tracking trajectories are $\phi^L(z^{L, N_L})(\cdot)$ and $\phi^F(z^{L, N_L}, z^{F, N_F})(\cdot)$ for leaders and followers, respectively. However, these tracking trajectories cannot be known a priori and so cannot be used for constructing the control laws. Therefore, we take an MF stochastic control approach to approximate the coupling trajectory terms by purely deterministic processes called mass behaviours of leaders and followers.

In the MF methodology, each agent assumes that, in the large population limit, its individual action has no impact on the mass behaviour. In turn, the impact of the mass on the agent is captured in the limit via a posited deterministic but unknown trajectory which through consistency requirements is then shown to satisfy a fixed point equation system. Hence, each individual agent's control law may be viewed as a result of that agent having solved an optimal tracking problem for which the resulting control law is a combination of a local state feedback and a pre-computable mass dependent open loop deterministic component.

For any leader or follower representative agent with dynamical parameter $\theta = [A, B, C]$ (where $\theta \in \Theta_L$ for leaders and $\theta \in \Theta_F$ for followers) let Π_θ be the solution to the algebraic Riccati equation

$$\Pi_\theta A + A^T \Pi_\theta - \Pi_\theta B R^{-1} B^T \Pi_\theta + Q = 0 \quad (5.10)$$

and let $\Gamma_\theta := A - B R^{-1} B^T \Pi_\theta$.

We have the following assumptions for the model under consideration:

(A5.1) For each $\theta = [A, B, C]$ from the compact sets Θ_L or Θ_F , we assume that the pair $[A, B]$ is stabilizable and the pair $[A, Q^{1/2}]$ is detectable. \square

REMARK 5.1. *Let Assumption (A5.1) hold. Then for any $\theta = [A, B, C] \in \Theta_L \cup \Theta_F$, the algebraic Riccati equation (5.10) has a unique positive semi-definite solution, Γ_θ is asymptotically stable and there exist positive γ, ρ such that $\|e^{\Gamma_\theta t}\| \leq \gamma e^{-\rho t}$ for all $t \geq 0$ [48].*

For the sequence of the leaders and followers' dynamical parameters, $\{\theta_i \in \Theta_L : 1 \leq i \leq N_L\}$ and $\{\theta_i \in \Theta_F : 1 \leq i \leq N_F\}$, respectively, define the empirical distributions

$$F_N^L(\theta) := \frac{1}{N} \sum_{i=1}^N 1_{\{\theta_i \in \Theta_L: \theta_i \leq \theta\}}, \quad F_N^F(\theta) := \frac{1}{N} \sum_{i=1}^N 1_{\{\theta_i \in \Theta_F: \theta_i \leq \theta\}}$$

where $\theta_i \leq \theta$ means the component-wise inequality for the two vectors θ_i and θ , and $1_{\{\theta_i \in \Theta_L: \theta_i \leq \theta\}} = 1$ if $\Theta_L \ni \theta_i \leq \theta$ holds, and $1_{\{\theta_i \in \Theta_L: \theta_i \leq \theta\}} = 0$ otherwise. Similarly, $1_{\{\theta_i \in \Theta_F: \theta_i \leq \theta\}} = 1$ if $\Theta_F \ni \theta_i \leq \theta$ holds, and $1_{\{\theta_i \in \Theta_F: \theta_i \leq \theta\}} = 0$ otherwise.

(A5.2) There exist two probability distributions $F^L(\cdot)$ and $F^F(\cdot)$ such that F_N^L converges to F^L and F_N^F converges to F^F weakly, i.e., for any bounded and continuous function $\phi(\theta)$,

$$\begin{aligned} \lim_{N \rightarrow \infty} \int \phi(\theta) dF_N^L(\theta) &= \int \phi(\theta) dF^L(\theta), \\ \lim_{N \rightarrow \infty} \int \phi(\theta) dF_N^F(\theta) &= \int \phi(\theta) dF^F(\theta) \end{aligned}$$

where dF denotes the measure induced by the distribution functions F . \square

REMARK 5.2. *It is important to note that if the sequences $\{\theta_i \in \Theta_L : 1 \leq i \leq N_L\}$ and $\{\theta_i \in \Theta_F : 1 \leq i \leq N_F\}$ are generated by independent randomized observations on the distributions F^L and F^F , respectively, then **(A5.2)** holds with probability one by the Strong Law of Large Numbers or the Glivenko-Cantelli theorem [42].*

By the probability distributions $F^L(\cdot)$ and $F^F(\cdot)$ we define

$$\psi^{L,\infty}(\cdot) := \int_{\Theta_L} \bar{z}_\theta^L(\cdot) dF^L(\theta), \quad \psi^{F,\infty}(\cdot) := \int_{\Theta_F} \bar{z}_\theta^F(\cdot) dF^F(\theta)$$

as the centroid of the leaders and followers, respectively, in the infinite population limit. The functions $\psi^{L,\infty}(\cdot)$ and $\psi^{F,\infty}(\cdot)$ are intended to respectively approximate $z^{L,N_L}(\cdot)$ and $z^{F,N_F}(\cdot)$, defined in (5.2) and (5.5), in the infinite population limit.

Denote Γ_{θ_i} by Γ_i . Applying the MF LQG approach [79, 106] to the leaders' dynamic game model (5.1)-(5.3), we obtain the leaders' MF system of equations in the infinite leaders' population limit

$$\frac{ds_i^L}{dt} = -(A_i - B_i R^{-1} B_i^T \Pi_{\theta_i})^T s_i^L + Q \phi^{L,\infty} \quad (5.11)$$

$$\frac{d\bar{z}_i^L}{dt} = (A_i - B_i R^{-1} B_i^T \Pi_{\theta_i}) \bar{z}_i^L - B_i R^{-1} B_i^T s_i^L \quad (5.12)$$

$$\psi^{L,\infty}(\cdot) = \int_{\Theta_L} \bar{z}_{\theta_i}^L(\cdot) dF^L(\theta_i) \quad (5.13)$$

$$\phi^{L,\infty}(\cdot) = \lambda h(\cdot) + (1 - \lambda) \psi^{L,\infty}(\cdot) \quad (5.14)$$

where $\bar{z}_i^L(\cdot) := \bar{z}_{\theta_i}^L(\cdot)$, $s_i^L(\cdot) := s_{\theta_i}^L(\cdot)$, $s_i^L(0) = -\int_0^\infty e^{\Gamma_i \tau} Q \phi^{L,\infty}(\tau) d\tau$ and $\bar{z}_i^L(0)$ is given for $\theta_i \in \Theta_L$.

The equation system (5.11)-(5.14) is similar to the MF LQG system in [106] but with a time-varying reference trajectory $h(\cdot)$, and is constructed such that each leader carries out optimal tracking of the leaders' mass behavior $\phi^{L,\infty}(\cdot)$. More precisely, (i) (5.11) is the mass offset tracking equation, (ii) (5.12) is obtained by taking expectations of the closed-loop dynamics of a leader with dynamical parameter $\theta_i \in \Theta_L$ while the best response is $u_i^{L,\infty}(\cdot) := -R^{-1} B_i^T (\Pi_{\theta_i} z_i^L(\cdot) + s_i^L(\cdot))$, (iii) $\psi^{L,\infty}(\cdot)$ in (5.13) is

the centroid of the leaders in the infinite population limit, and (iv) $\phi^{L,\infty}(\cdot)$ in (5.14) is approximating $\phi^L(z^{L,N_L})(\cdot)$, defined in (5.2), in the infinite population limit.

Next we obtain the followers' MF system of equations associated with the followers' dynamic game model (5.4)-(5.6) in the infinite population limit

$$\frac{ds_i^F}{dt} = -(A_i - B_i R^{-1} B_i^T \Pi_{\theta_i})^T s_i^F + Q \phi^{F,\infty} \quad (5.15)$$

$$\frac{d\bar{z}_i^F}{dt} = (A_i - B_i R^{-1} B_i^T \Pi_{\theta_i}) \bar{z}_i^F - B_i R^{-1} B_i^T s_i^F \quad (5.16)$$

$$\psi^{L,\infty}(\cdot) = \int_{\Theta_L} \bar{z}_{\theta_i}^L(\cdot) dF^L(\theta_i) \quad (5.17)$$

$$\psi^{F,\infty}(\cdot) = \int_{\Theta_F} \bar{z}_{\theta_i}^F(\cdot) dF^F(\theta_i) \quad (5.18)$$

$$\phi^{F,\infty}(\cdot) = \eta \psi^{L,\infty}(\cdot) + (1 - \eta) \psi^{F,\infty}(\cdot) \quad (5.19)$$

where $\bar{z}_i^F(\cdot) := \bar{z}_{\theta_i}^F(\cdot)$, $s_i^F(\cdot) := s_{\theta_i}^F(\cdot)$, $s_i^F(0) = -\int_0^\infty e^{\Gamma_i \tau} Q \phi^{F,\infty}(\tau) d\tau$ and $\bar{z}_i^F(0)$ is given for $\theta_i \in \Theta_F$.

The equation system (5.15)-(5.19) is constructed such that each follower carries out optimal tracking of the followers' mass behavior $\phi^{F,\infty}(\cdot)$. More precisely, (i) (5.15) is the mass offset tracking equation, (ii) (5.16) is obtained by taking expectations of the closed-loop dynamics of a follower with dynamical parameter θ_i while the best response is $u_i^{F,\infty}(\cdot) := -R^{-1} B_i^T (\Pi_{\theta_i} \bar{z}_i^F(\cdot) + s_i^F(\cdot))$, (iii) $\psi^{L,\infty}(\cdot)$ in (5.17) is the centroid of the leaders, in the infinite population limit, computed by the followers from the leaders' MF equation system (5.11)-(5.14), (iv) $\psi^{F,\infty}(\cdot)$ in (5.18) is the centroid of the followers in the infinite population limit, and (v) $\phi^{F,\infty}(\cdot)$ in (5.19) is approximating $\phi^F(z^{L,N_L}, z^{F,N_F})(\cdot)$, defined in (5.5), in the infinite population limit.

It is important to note that the followers' MF system of equations is coupled to the leaders' MF system of equations due to appearance of leaders' centroid $\psi^{L,\infty}(\cdot)$ in the mass behavior of the followers $\phi^{F,\infty}(\cdot)$. We make the following assumption:

(A5.3) We assume that for all possible reference trajectories $h_\delta(\cdot)$ parameterized by $\delta \in \Delta$:

- (a) The leaders' MF system of equations (5.11)-(5.14) has a unique solution such that $\psi^{L,\infty}(\cdot) \in C_n^b$.
- (b) The followers' MF system of equations (5.15)-(5.19) (implicitly depending upon the solution of (5.11)-(5.14)) has a unique solution such that $\psi^{F,\infty}(\cdot) \in C_n^b$.

□

Generally, it seems difficult to verify this assumption. However, in the next subsection we provide sufficient conditions for the existence and uniqueness of solutions to these MF equation systems by using a contractive mapping argument (see [79, 106]).

5.3.3. Analysis of MF Systems for Leaders and Followers. First, we consider the leaders' MF system of equations (5.11)-(5.14). For given $h(\cdot)$, $\psi^{L,\infty}(\cdot)$ and hence $\phi^{L,\infty}(\cdot)$, the unique solution of the tracking offset equation (5.11) for a "generic" leader agent with dynamical parameter $\theta_i = [A_i, B_i, C_i] \in \Theta_L$ is

$$\begin{aligned} s_i^L(t) &= - \int_t^\infty e^{-\Gamma_i^T(t-\tau)} Q \phi^{L,\infty}(\tau) d\tau \\ &= - \int_t^\infty e^{-\Gamma_i^T(t-\tau)} Q (\lambda h(\tau) + (1-\lambda)\psi^{L,\infty}(\tau)) d\tau, \quad t \geq 0. \end{aligned} \quad (5.20)$$

Next, by solving (5.12) we have

$$\bar{z}_i^L(t) = e^{\Gamma_i t} \bar{z}_i^L(0) - \int_0^t e^{\Gamma_i(t-s)} B_i R^{-1} B_i^T s_i^L(s) ds, \quad t \geq 0 \quad (5.21)$$

which by substituting $s_i^L(\cdot)$ from (5.20) we get

$$\begin{aligned} \bar{z}_i^L(t) &= e^{\Gamma_i t} \bar{z}_i^L(0) + \int_0^t e^{\Gamma_i(t-s)} B_i R^{-1} B_i^T \\ &\quad \times \left(\int_s^\infty e^{-\Gamma_i^T(s-\tau)} Q (\lambda h(\tau) + (1-\lambda)\psi^{L,\infty}(\tau)) d\tau \right) ds \quad (5.22) \\ &:= (\Upsilon_i^L(\psi^{L,\infty}, h))(t), \quad t \geq 0 \end{aligned}$$

where Υ_i^L is an operator acting on bounded continuous functions. This and (5.13) result in

$$\begin{aligned}\psi^{L,\infty}(t) &\equiv \int_{\Theta_L} \bar{z}_i^L(t) dF^L(\theta_i) = \int_{\Theta_L} \Upsilon_i^L(\psi^{L,\infty}, h)(t) dF^L(\theta_i) \\ &=: \Upsilon^L(\psi^{L,\infty}, h)(t), \quad t \geq 0\end{aligned}$$

where we note $\psi^{L,\infty}(\cdot)$ is independent of $\theta_i \in \Theta_L$. Now for a given $h(\cdot)$ if the equation

$$\psi^{L,\infty} = \Upsilon^L(\psi^{L,\infty}, h) \quad (5.23)$$

has a unique solution $\psi^{L,\infty}(\cdot)$, then it can be used in (5.20) and (5.22) to compute the unique solution of the leaders' MF system of equations. In the following theorem we employ a contractive mapping argument to provide sufficient conditions under which equation (5.23) has a unique solution. We omit the proofs of the two following theorems which closely resemble that of Theorem 3.2 in [106].

THEOREM 5.2. *Assume (A5.1) and (A5.2) hold. For a given $h(\cdot) \in C_n^b$,*

(a) Υ^L defined in (5.23) is an operator from C_n^b to C_n^b .

(b) If

$$(1 - \lambda) \|R^{-1}\| \|Q\| \int_{\Theta_L} \|B_{\theta_i}\|^2 \left(\int_0^\infty \|e^{\Gamma_{\theta_i} t}\|^2 dt \right) dF^L(\theta_i) < 1 \quad (5.24)$$

then (5.23) has the unique solution $\psi^{L,\infty} \in C_n^b$, and so the leaders' MF system of equations, (5.11)-(5.14), has a unique solution which for a continuum of agents consists of the θ_i parameterized quadruple $(s_i^L(\cdot), \bar{z}_i^L(\cdot), \psi^{L,\infty}(\cdot), \phi^{L,\infty}(\cdot))$, $\theta_i \in \Theta_L$.

□

Second, we consider the followers' MF system of equations, (5.15)-(5.19). For given $\psi^{L,\infty}(\cdot)$, $\psi^{F,\infty}(\cdot)$ and hence $\phi^{F,\infty}(\cdot)$, we define the operator Υ_i^F for a "generic"

follower agent with dynamical parameter $\theta_i = [A_i, B_i, C_i] \in \Theta_F$ as

$$\begin{aligned}\bar{z}_{\theta_i}^F(t) &= e^{\Gamma_i t} \bar{z}_i^F(0) + \int_0^t e^{\Gamma_i(t-s)} B_i R^{-1} B_i^T \\ &\quad \times \left(\int_s^\infty e^{-\Gamma_i^T(s-\tau)} Q (\eta \psi^{L,\infty}(\tau) + (1-\eta) \psi^{F,\infty}(\tau)) d\tau \right) ds \\ &=: \Upsilon_i^F(\psi^{F,\infty}, \psi^{L,\infty})(t), \quad t \geq 0\end{aligned}$$

by using the solutions of (5.15) and (5.16) (similar to (5.22) for the i -th leader). This and (5.18) result in

$$\begin{aligned}\psi^{F,\infty}(t) &\equiv \int_{\Theta_F} \bar{z}_i^F(t) dF^F(\theta_i) = \int_{\Theta_F} \Upsilon_i^F(\psi^{F,\infty}, \psi^{L,\infty})(t) dF^F(\theta_i) \\ &=: \Upsilon^F(\psi^{F,\infty}, \psi^{L,\infty})(t), \quad t \geq 0\end{aligned}$$

where we note $\psi^{F,\infty}(\cdot)$ is independent of $\theta_i \in \Theta_F$. Now for a given $\psi^{L,\infty}(\cdot)$ (the solution of equation (5.23)) if the equation

$$\psi^{F,\infty} = \Upsilon^F(\psi^{F,\infty}, \psi^{L,\infty}) \quad (5.25)$$

has a unique solution $\psi^{F,\infty}(\cdot)$, then it can be used in (5.15) and (5.16) to compute the unique solution of the followers' MF system of equations. In the following theorem we employ the contractive mapping argument to provide sufficient conditions under which equation (5.25) has a unique solution.

THEOREM 5.3. *Assume (A5.1) and (A5.2) hold. For a given $\psi^{L,\infty}(\cdot) \in C_n^b$,*

(a) Υ^F defined in (5.25) is an operator from C_n^b to C_n^b .

(b) If

$$(1-\eta) \|R^{-1}\| \|Q\| \int_{\Theta_F} \|B_{\theta_i}\|^2 \left(\int_0^\infty \|e^{\Gamma_{\theta_i} t}\|^2 dt \right) dF^F(\theta_i) < 1 \quad (5.26)$$

then (5.25) has the unique solution $\psi^{F,\infty} \in C_n^b$, and so the followers' MF system of equations, (5.15)-(5.19), has a unique solution which for a continuum of agents consists of the θ_i parameterized quadruple $(s_i^F(\cdot), \bar{z}_i^F(\cdot), \psi^{F,\infty}(\cdot), \phi^{F,\infty}(\cdot))$, $\theta_i \in \Theta_F$.

□

Intuitively, (5.24) and (5.26) mean that λ and η should be reasonably close to 1, which means that the leaders should give sufficient attention to the reference trajectory, and on the other hand the followers should give sufficient attention to the group of leaders, so that at the end a desirable mean field behaviour (fixed point) can set in.

REMARK 5.3. *It is important to note that the conditions (5.24) and (5.26) do not depend on the reference trajectory of the leaders $h(\cdot)$.*

5.4. Estimation Procedure for the Adaptive Followers

The computation of the followers' MF system of equations (and hence followers' control laws) requires knowledge of the complete reference trajectory of the leaders $h(\cdot)$ which is in general not known to the followers. In this section we construct an adaptation procedure for a generic follower using a likelihood ratio based estimator (on a sample population of the leaders' trajectories) to identify the member generating the reference trajectory $h(\cdot)$ from a given finite set of possible parameters each identifying a single trajectory. Hence, followers are adaptive in the sense that they use an estimator to identify $h(\cdot)$. As stated earlier, we assume that the reference trajectory $h(\cdot)$ is parameterized with δ from a finite set Δ such that $h_\delta \in C_n^b$ for every $\delta \in \Delta$.

Likelihood ratio based estimation is a well-known method for generating estimates of an unknown stochastic model parameter (see e.g. [33, 62]). In this chapter we use the general result on the convergence of likelihood ratio estimators for stochastic processes parameterized by a finite set of alternative values which was established in [32] (see [33]).

5.4.1. The Likelihood Function. For a generic adaptive follower, we define the *likelihood function* (see [51, 52]) on (a subset of) the leaders' trajectories by ($0 < t < \infty$)

$$L_t^M(\delta) := \exp \left(\int_0^t (z_{\delta,s}^{L,M})^T dy_s^M - \frac{1}{2} \int_0^t \|z_{\delta,s}^{L,M}\|^2 ds \right) \quad (5.27)$$

where $z_{\delta,t}^{L,M} := 1/M \sum_{i=1}^M z_{i,\delta}^L(t)$ is the centroid of the leaders' states when the defining parameter of $h(\cdot)$ is assumed to be $\delta \in \Delta$, and $y^M(\cdot)$ is the observation process of the generic follower (which observes a non-empty random subset $O \subset L$ of cardinality $M \equiv M(N_L) = \lfloor \sqrt{N_L} \rfloor$ of the leaders' trajectories) as defined in (5.7).

Without loss of generality assume $\delta_1 \in \Delta$ is the true parameter of the reference trajectory $h(\cdot)$ in the rest of the chapter, that is to say, the parameter of the leaders generating the data. Therefore, the observation process of the generic follower is of the form

$$dy^M = \left(\frac{1}{M} \sum_{i \in O} z_{i,\delta_1}^L \right) dt + \frac{1}{M} \sum_{i=1}^M D_i dv_i, \quad t \geq 0 \quad (5.28)$$

where $\{v_i : 1 \leq i \leq N_L\}$ is a set of independent standard Wiener processes.

We define the *asymptotic (in population) likelihood function* of the generic adaptive follower to be the deterministic function ($0 < t < \infty$)

$$L_t^\infty(\delta) := \exp \left(\int_0^t (\psi_{\delta,s}^{L,\infty})^T \psi_{\delta_1,s}^{L,\infty} ds - \frac{1}{2} \int_0^t \|\psi_{\delta,s}^{L,\infty}\|^2 ds \right) \quad (5.29)$$

where $\psi_{\delta,t}^{L,\infty} := \psi_\delta^{L,\infty}(t)$ such that $\psi_\delta^{L,\infty}(\cdot)$ is the deterministic infinite population leaders' centroid computed from the leaders' MF system of equations (5.11)-(5.14) when the defining parameter of $h(\cdot)$ is assumed to be $\delta \in \Delta$.

Since (i) the processes $z_\delta^{L,M}(\cdot)$, $\delta \in \Delta$, in (5.27) are not computable or observable for the adaptive followers, and (ii) the true infinite population centroid of the leaders $\psi_{\delta_1}^{L,\infty}(\cdot)$ in (5.29) is not known to the followers, we introduce the following *hybrid likelihood function* for a generic adaptive follower with observation process $y^M(\cdot)$ ($0 < t < \infty$)

$$H_t^M(\delta) := \exp \left(\int_0^t (\psi_{\delta,s}^{L,\infty})^T dy_s^M - \frac{1}{2} \int_0^t \|\psi_{\delta,s}^{L,\infty}\|^2 ds \right). \quad (5.30)$$

It is important to note that the hybrid likelihood function (5.30) is computable for adaptive followers and includes: (i) the infinite population centroid of the leaders $\psi_\delta^{L,\infty}(\cdot)$ which can be computed by each adaptive follower from the leaders' MF system of equations (5.11)-(5.14) when the defining parameter of the reference trajectory $h(\cdot)$

is assumed to be $\delta \in \Delta$, and (ii) the observation process $y^M(\cdot)$ which is given in (5.28) for a generic adaptive follower.

PROPOSITION 5.1. *Assume (A5.1)-(A5.3-(a)) hold. Then*

(a) *For each $\delta \in \Delta$ and $t > 0$*

$$\lim_{M \rightarrow \infty} |L_t^M(\delta) - L_t^\infty(\delta)| = 0 \quad a.s.$$

i.e., for each $\delta \in \Delta$, $t > 0$ and $\epsilon > 0$ there exists a random $M_{\delta, \epsilon, t}$, $0 < M_{\delta, \epsilon, t}(\omega) < \infty$, such that $|L_t^m(\delta) - L_t^\infty(\delta)| < \epsilon$ a.s. for all $m > M_{\delta, \epsilon, t}$.

(b) *For each $\delta \in \Delta$ and $t > 0$,*

$$\lim_{M \rightarrow \infty} |H_t^M(\delta) - L_t^\infty(\delta)| = 0 \quad (a.s.).$$

Proof. See the Appendix. □

5.4.2. The Likelihood Ratio. The likelihood ratio test provides the means for comparing the likelihood of the observations under one hypothesis about the unknown parameters of the model against the likelihood of the observations under alternative hypotheses.

At each instant $0 < t < \infty$ the set of the *likelihood ratios* is

$$LR^M(t) := \left\{ \frac{L_t^M(\delta_i)}{L_t^M(\delta_j)} : \delta_i, \delta_j \in \Delta, \delta_i \neq \delta_j \right\}$$

in which each ratio $L_t^M(\delta_i)/L_t^M(\delta_j)$ depends explicitly upon the hypotheses δ_i and δ_j .

It is important to note that for any fixed δ_i in Δ the process

$$\left(\frac{L_t^M(\delta_i)}{L_t^M(\delta_1)} \right)_{t \geq 0}$$

is a positive martingale (with respect to the filtration $\{\mathcal{F}_t^{z^L, y^M}\}_{t \geq 0}$ where $\mathcal{F}_t^{z^L, y^M}$ is defined as the σ -field $\sigma(z_i^L(\tau), y^M(\tau) : 1 \leq i \leq M, \tau < t)$ [32, 51]. Therefore, for any fixed $\delta_i \in \Delta$, the L^1 -bounded martingale $(L_t^M(\delta_i)/L_t^M(\delta_1))_{t \geq 0}$ a.s. converges to a limiting random variable by the Martingale Convergence Theorem [42]. Hence, for

any fixed $\delta_i, \delta_j \in \Delta$, $\delta_i \neq \delta_j$, the process $(L_t^M(\delta_i)/L_t^M(\delta_j))_{t \geq 0}$ converges *a.s.* to a positive limiting random variable.

5.4.3. Main Estimation Theorem. We define the *Likelihood Ratio Estimator* (LRE) for a generic adaptive follower i , $1 \leq i \leq N_F$, with observation size m as

$$\hat{\delta}_i^m(t) := \left\{ \delta \in \Delta : \frac{H_{t_k}^m(\delta)}{H_{t_k}^m(\delta')} \geq 1, \forall \delta' \in \Delta, \delta' \neq \delta \right\} \quad (5.31)$$

where $t \in [t_k, t_k + \tau_i)$, τ_i is a pre-specific positive number and t_0, t_1, \dots is an infinite switching time sequence such that $t_{k+1} - t_k = \tau_i$, $k \geq 0$. $\hat{\delta}_i^m(\cdot)$ is a parameter selector between finite alternatives which is made at each instant.

At each time t if $\hat{\delta}_i^m(t)$ has more than one member, a tie-breaking rule measurable with respect to the σ -field $\sigma(y_i^m(\tau) : \tau < t)$ is employed.

We now enunciate the following verifiable identifiability condition.

(A5.4) (*Identifiability Condition*) there exist a deterministic real number $\alpha > 0$, and deterministic time T_α , $0 < T_\alpha < \infty$, such that

$$\int_0^t \|\psi_{\delta_i, s}^{L, \infty} - \psi_{\delta_j, s}^{L, \infty}\|^2 ds > \alpha, \quad \forall \delta_i, \delta_j \in \Delta, \delta_i \neq \delta_j, t > T_\alpha$$

where for any fixed $\delta \in \Delta$, $\psi_\delta^{L, \infty}(\cdot)$ is the deterministic infinite population leaders' centroid computed from (5.11)-(5.14), when the defining parameter of $h(\cdot)$ is assumed to be δ . □

REMARK 5.4. *This identifiability assumption implies that the corresponding centroid of the leaders for any two distinct parameters of the set Δ (which characterizes the reference trajectory $h(\cdot)$) is distinguishable, after some finite deterministic time.*

LEMMA 5.1. *Assume (A5.1)-(A5.3-(a)) and (A5.4) hold. Then there exist a deterministic ϵ , $0 < \epsilon < 1$, and a deterministic time T_ϵ , $0 < T_\epsilon < \infty$, such that*

$$\frac{L_t^\infty(\delta)}{L_t^\infty(\delta_1)} < 1 - \epsilon, \quad \forall \delta \in \Delta, \delta \neq \delta_1, t > T_\epsilon.$$

Proof. For any fixed $\delta \in \Delta$, $\delta \neq \delta_1$, and t , $0 < t < \infty$, by (5.29) we have

$$\begin{aligned} \frac{L_t^\infty(\delta)}{L_t^\infty(\delta_1)} &= \exp \left[\int_0^t \left(\psi_{\delta,s}^{L,\infty} - \psi_{\delta_1,s}^{L,\infty} \right)^T \psi_{\delta_1,s}^{L,\infty} ds \right. \\ &\quad \left. - \frac{1}{2} \int_0^t \left(\|\psi_{\delta,s}^{L,\infty}\|^2 - \|\psi_{\delta_1,s}^{L,\infty}\|^2 \right) ds \right] = \exp \left[-\frac{1}{2} \left(\int_0^t \|\psi_{\delta,s}^{L,\infty} - \psi_{\delta_1,s}^{L,\infty}\|^2 ds \right) \right]. \end{aligned}$$

But, by the Identifiability Condition **(A5.4)**, there exist a deterministic real number $\alpha > 0$, and deterministic time T_α , $0 < T_\alpha < \infty$, such that for all $t > T_\alpha$

$$\begin{aligned} &\exp \left[-\frac{1}{2} \left(\int_0^t \|\psi_{\delta,s}^{L,\infty} - \psi_{\delta_1,s}^{L,\infty}\|^2 ds \right) \right] \\ &\quad < \exp\left(-\frac{1}{2}\alpha\right), \quad \delta \in \Delta, \delta \neq \delta_1 \end{aligned}$$

and setting $1 - \epsilon = \exp(-\frac{1}{2}\alpha)$ yields the result with $0 < \epsilon < 1$. \square

LEMMA 5.2. *Assume **(A5.1)**-**(A5.3-(a))** and **(A5.4)** hold. Then there exist a deterministic η , $0 < \eta < 1$, a deterministic time T_η , $0 < T_\eta < \infty$, and, with probability one, a random M_η , $0 < M_\eta(\omega) < \infty$, such that for all $t > T_\eta$ and $m > M_\eta(\omega)$*

$$\frac{H_t^m(\delta)}{H_t^m(\delta_1)} < 1 - \eta, \quad \delta \in \Delta, \delta \neq \delta_1 \quad (a.s.).$$

Proof. By Lemma 5.1 there exist a deterministic ϵ , $0 < \epsilon < 1$, and a deterministic time T_ϵ , $0 < T_\epsilon < \infty$, such that for any $\delta \in \Delta$, $\delta \neq \delta_1$, and $t > T_\epsilon$

$$\frac{L_t^\infty(\delta)}{L_t^\infty(\delta_1)} < 1 - \epsilon. \quad (5.32)$$

Now choose $\eta := \eta_\epsilon$ such that $0 < \eta < \epsilon < 1$. Then at each instant t , $t > T_\eta := T_{\eta_\epsilon}$, by Proposition 5.1 with $\zeta := \frac{(\epsilon-\eta)L_t^\infty(\delta_1)}{2-\eta} > 0$, there exists, with probability one, $M_\eta(\omega) := M_{\zeta, T_\epsilon(\omega)}(\omega)$, $0 < M_\eta(\omega) < \infty$, where

$$M_\eta(\omega) = \max \left\{ M_{\delta_1, \eta}(\omega), M_{\delta_2, \eta}(\omega), \dots, M_{\delta_{|\Delta|}, \eta}(\omega) \right\}$$

such that for all $m > M_\eta$ and $\delta \in \Delta$

$$|H_t^m(\delta) - L_t^\infty(\delta)| < \zeta \quad (a.s.). \quad (5.33)$$

Hence, by (5.32) and (5.33), for all $t > T_\eta$ and $m > M_\eta(\omega)$ we have

$$\frac{H_t^m(\delta)}{H_t^m(\delta_1)} < \frac{L_t^\infty(\delta) + \zeta}{L_t^\infty(\delta_1) - \zeta} = \frac{\frac{L_t^\infty(\delta)}{L_t^\infty(\delta_1)} + \frac{\epsilon - \eta}{2 - \eta}}{1 - \frac{\epsilon - \eta}{2 - \eta}} < \frac{1 - \epsilon + \frac{\epsilon - \eta}{2 - \eta}}{1 - \frac{\epsilon - \eta}{2 - \eta}} = 1 - \eta, \quad \delta \in \Delta, \quad \delta \neq \delta_1 \quad (a.s.).$$

□

THEOREM 5.4. *Assume (A5.1)-(A5.3-(a)) and (A5.4) hold. Then for each generic adaptive follower i , $1 \leq i \leq N_F$, there exist a deterministic T_i , $0 < T_i < \infty$, and, with probability one, a random M_i , $0 < T_i$, $M_i(\omega) < \infty$, such that $\hat{\delta}_i^m(t) = \delta_1$ for all $t > T_i$ and $m > M_i(\omega)$.*

Proof. By Lemma 5.2 for any adaptive follower i , $1 \leq i \leq N_F$, there exist a deterministic η_i , $0 < \eta_i < 1$, a deterministic time $T_i := T_{\eta_i}$, $0 < T_i < \infty$, and a random $M_i := M_{\eta_i}$, $0 < M_i(\omega) < \infty$, such that for all $t > T_i$, $m > M_i(\omega)$ and $\delta \in \Delta$ where $\delta \neq \delta_1$

$$\frac{H_t^m(\delta_1)}{H_t^m(\delta)} = \left(\frac{H_t^m(\delta)}{H_t^m(\delta_1)} \right)^{-1} \geq (1 - \eta_i)^{-1} > 1 \quad a.s.$$

which implies that $\hat{\delta}_i^m(t) = \delta_1$ for all $t > T_i$ and $m > M_i(\omega)$, based on the definition of the LRE in (5.31). □

5.5. The Stability Analysis of the Mean Field Control Laws

DEFINITION 5.1. *In a large but finite population, the decentralized MF control laws for the leaders and followers, respectively, are as follows.*

- *Leaders' MF Control Laws: The control strategy of each generic leader i , $1 \leq i \leq N_L$, with dynamical parameter $\theta_i = [A_i, B_i, C_i] \in \Theta_L$ is defined as*

$$u_i^{L,\infty}(\cdot) := -R^{-1}B_i^T(\Pi_{\theta_i}z_i^L(\cdot) + s_i^L(\cdot)) \quad (5.34)$$

where $s_i^L(\cdot)$ is the solution of equation (5.11) presented in (5.20).

- *Followers' MF Control Laws: The control strategy of each generic adaptive follower i , $1 \leq i \leq N_F$, with dynamical parameter $\theta_i = [A_i, B_i, C_i] \in \Theta_F$,*

and observation size m is defined as

$$\hat{u}_i^{F,\infty}(\cdot) := -R^{-1}B_i^T(\Pi_{\theta_i}z_i^F(\cdot) + s_{\delta_i^m}^F(\cdot)) \quad (5.35)$$

where $s_{\delta_i^m}^F(\cdot)$ is the solution of equation (5.15), when the defining parameter of the reference trajectory $h(\cdot)$ is assumed to be $\hat{\delta}_i^m(\cdot)$ defined in (5.31). \square

In the construction of individual strategies, (i) each leader needs to know λ , $h(\cdot)$, the leader's population initial mean and the distribution of the leaders' dynamical parameters ($F^L(\cdot)$), where the term know denotes that the control law of the agent in question may be an explicit function of the indicated information; and (ii) each follower needs to know λ , η , the population initial means and the distributions of both leaders and followers' dynamical parameters ($F^L(\cdot)$ and $F^F(\cdot)$), but each adaptive follower does not know the reference trajectory of the leaders $h(\cdot)$ and estimates it by likelihood estimation from a finite set of predefined signals based upon its observation process. Note that it is not required for any leader or follower agent to know specific information (such as the dynamical parameter) of any other particular agent.

5.5.1. The Follower's Adaptive Mean Field Algorithm. The algorithm has the following two phases:

- (i) *Estimation Phase:* By observing a sample population of the leaders each follower computes the set of likelihood ratios (based on the hybrid likelihood functions defined in (5.30)) at each instant for alternative values of its hypothesis parameter $\delta \in \Delta$. Each follower also computes control laws by using the parameters in the finite set Δ . Therefore, each follower has a set of control strategies with respect to alternative defining parameters of the reference trajectory $h(\cdot)$, and at any instant uses the maximum likelihood ratio estimate (MLRE) without a guarantee that the MLRE has taken the true parameter value.
- (ii) *Lock-on Phase:* As the observation size of each adaptive follower i , $1 \leq i \leq N_F$, goes to infinity, its estimate converges to the true parameter of the

unknown reference trajectory at a deterministic time T_i (Theorem 5.4). In this phase, the control law of each adaptive follower i , $1 \leq i \leq N_F$, will necessarily be computed with the true parameter of the reference trajectory for all time $t > T_i$ and any sufficiently large random observation size.

Needless to say, an adaptive follower cannot deduce at which population size the lock-on phase has commenced since this occurs at some random size $M(\omega)$.

5.5.2. Stability Analysis. For the i -th leader, $1 \leq i \leq N_L$, with dynamical parameter $\theta_i = [A_i, B_i, C_i] \in \Theta_L$, denote $z_i^{L,\infty}(\cdot)$ as the closed-loop solution of its dynamics (5.1) while using the MF control law $u_i^{L,\infty}(\cdot)$ as defined in (5.34). In an analogous way let $\hat{z}_i^{F,\infty}(\cdot)$ be the closed-loop solution of the i -th adaptive follower's dynamics, $1 \leq i \leq N_F$, (5.4) where its dynamical parameter is $\theta_i = [A_i, B_i, C_i] \in \Theta_F$ and its adaptive control law is the MF control law $\hat{u}_i^{F,\infty}(\cdot)$ defined in (5.35).

THEOREM 5.5. (*Stability of the MF control laws in the sense of time average*)

(a) Assume **(A5.1)**-**(A5.3-(a))** hold. Then (a.s.)

$$\sup_{N_L \geq 1} \max_{1 \leq i \leq N_L} \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(\|z_i^{L,\infty}(s)\|^2 + \|u_i^{L,\infty}(s)\|^2 \right) ds < \infty. \quad (5.36)$$

(b) Assume **(A5.1)**-**(A5.4)** hold. Then (a.s.)

$$\sup_{N_F \geq 1} \max_{1 \leq i \leq N_F} \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(\|\hat{z}_i^{F,\infty}(s)\|^2 + \|\hat{u}_i^{F,\infty}(s)\|^2 \right) ds < \infty. \quad (5.37)$$

Proof. See the Appendix. □

It is important to note that these stability results depend on the boundedness of the infinite population centroids $\psi^{L,\infty}(\cdot)$ and $\psi^{F,\infty}(\cdot)$ computed from the leaders and followers' MF system of equations, respectively.

5.6. ϵ -Nash Equilibrium Property of the Mean Field Based Control Laws

In a system of population size N , let the admissible control set of agent i , $1 \leq i \leq N$, be $\mathcal{U}_i := \{u_i(\cdot) : u_i(t) \text{ is adapted to } \sigma(z_j(s), s \leq t, 1 \leq j \leq N), \|z_i(T)\| =$

$o(\sqrt{T})$, $\int_0^T \|z_i(t)\|^2 dt = O(T)$ a.s.}. Note that the strategies in \mathcal{U}_i may use full state information (i.e. $u_i(t)$ is a function of both time t and the system state at this time, $(z_1(t), \dots, z_N(t))$). Denote $u_{-i} = (u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_N)$. To indicate the dependence of the i -th agent's cost function J_i on u_i , u_{-i} and population size N , we write it as $J_i^N(u_i, u_{-i})$.

DEFINITION 5.2. *Given $\epsilon > 0$, a set of controls $u_k^o \in \mathcal{U}_k$, $1 \leq k \leq N$, for N agents generates an a.s. ϵ -Nash equilibrium with respect to the costs J_k , $1 \leq k \leq N$, if for any i , $1 \leq i \leq N$,*

$$J_i^N(u_i^o, u_{-i}^o) - \epsilon \leq \inf_{u_i \in \mathcal{U}_i} J_i^N(u_i, u_{-i}^o) \leq J_i^N(u_i^o, u_{-i}^o) \quad (a.s.). \quad \square$$

Let $z_i^{L,\infty}$ be the closed-loop solution of the i -th leader's dynamics (5.1) with the MF control input defined in (5.34), and $\psi^{L,\infty}(\cdot)$ be the infinite population centroid of the leaders (5.13), we denote

$$(\epsilon_{N_L})^2 := \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left\| \psi^{L,\infty}(t) - \frac{1}{N_L} \sum_{i=1}^{N_L} z_i^{L,\infty}(t) \right\|^2 dt. \quad (5.38)$$

The proof of the following theorem is similar to the one of Theorem 6.1 in [106].

THEOREM 5.6. *Assume (A5.1)-(A5.3-(a)) hold. Then the set of the leaders' MF control laws is an a.s. $O(\epsilon_{N_L})$ -Nash equilibrium, i.e., for any i , $1 \leq i \leq N_L$, we have*

$$J_i^{L,N_L}(u_i^{L,\infty}, u_{-i}^{L,\infty}) - O(\epsilon_{N_L}) \leq \inf_{u_i \in \mathcal{U}_i^L} J_i^{L,N_L}(u_i, u_{-i}^{L,\infty}) \leq J_i^{L,N_L}(u_i^{L,\infty}, u_{-i}^{L,\infty}) \quad a.s.$$

where $\lim_{N_L \rightarrow \infty} \epsilon_{N_L} = 0$ a.s., $(\epsilon_{N_L} \approx O(1/\sqrt{N_L}))$. □

For any adaptive follower i , $1 \leq i \leq N_F$, let $u_{i,\delta_1}^{F,\infty}(\cdot)$ be the MF control law when the defining parameter of the reference trajectory is assumed to be the true parameter $\delta_1 \in \Delta$ and let $\hat{z}_i^{F,\infty}(\cdot)$ and $z_{i,\delta_1}^{F,\infty}(\cdot)$ be the closed-loop solutions of dynamics (5.4) with control input $\hat{u}_i^{F,\infty}(\cdot)$ and $u_{i,\delta_1}^{F,\infty}(\cdot)$, respectively (the explicit form of $\hat{z}_i^{F,\infty}(\cdot)$ is

presented in (5.77)). Let $u^{L,\infty} := (u_1^{L,\infty}, \dots, u_{N_L}^{L,\infty})$, $\hat{u}^{F,\infty} := (\hat{u}_1^{F,\infty}, \dots, \hat{u}_{N_F}^{F,\infty})$ and $u_{\delta_1}^{F,\infty} := (u_{1,\delta_1}^{F,\infty}, \dots, u_{N_F,\delta_1}^{F,\infty})$.

LEMMA 5.3. *Assume (A5.1)-(A5.4) hold. Then there exists a random M , $0 < M(\omega) < \infty$, such that for $N_L \geq M(\omega)$ and observation sizes $m_i \geq M(\omega)$, $1 \leq i \leq N_F$, we have*

$$J_i^{F,N}(\hat{u}_i^{F,\infty}; \hat{u}_{-i}^{F,\infty}, u^{L,\infty}) \leq J_i^{F,N}(u_{i,\delta_1}^{F,\infty}; u_{-i,\delta_1}^{F,\infty}, u^{L,\infty}) \quad a.s.$$

for $1 \leq i \leq N_F$.

Proof. By Theorem 5.4 for each adaptive follower i , $1 \leq i \leq N_F$, there exist a deterministic time T_i and a random M_i , $0 < T_i$, $M_i(\omega) < \infty$, such that $\hat{\delta}_i^m(t) = \delta_1$ for all $t > T_i$ and $m > M_i(\omega)$. Let $T_F := \max\{T_1, \dots, T_{N_F}\} < \infty$ and $M := \max\{M_1, \dots, M_{N_F}\} < \infty$, then for $N_L \geq M(\omega)$ and observation sizes $m_i \geq M(\omega)$, $1 \leq i \leq N_F$, we have

$$\begin{aligned} & J_i^{F,N}(\hat{u}_i^{F,\infty}; \hat{u}_{-i}^{F,\infty}, u^{L,\infty}) \\ & \equiv \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(\left\| \hat{z}_i^{F,\infty} - \left(\frac{\eta}{N_L} \sum_{j=1}^{N_L} z_j^{L,\infty} + \frac{(1-\eta)}{N_F} \sum_{j=1}^{N_F} \hat{z}_j^{F,\infty} \right) \right\|_Q^2 + \|\hat{u}_i^{F,\infty}\|_R^2 \right) dt \\ & \leq \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^{T_F} \left(\left\| \hat{z}_i^{F,\infty} - \left(\frac{\eta}{N_L} \sum_{j=1}^{N_L} z_j^{L,\infty} + \frac{(1-\eta)}{N_F} \sum_{j=1}^{N_F} \hat{z}_j^{F,\infty} \right) \right\|_Q^2 + \|\hat{u}_i^{F,\infty}\|_R^2 \right) dt \\ & \quad + \limsup_{T \rightarrow \infty} \frac{1}{T} \int_{T_F}^T \left(\left\| z_{i,\delta_1}^{F,\infty} - \left(\frac{\eta}{N_L} \sum_{j=1}^{N_L} z_j^{L,\infty} + \frac{(1-\eta)}{N_F} \sum_{j=1}^{N_F} z_{j,\delta_1}^{F,\infty} \right) \right\|_Q^2 + \|u_{i,\delta_1}^{F,\infty}\|_R^2 \right) dt \\ & := I_1 + I_2. \end{aligned} \tag{5.39}$$

But, by the long time average stability of the MF control laws (5.37) there exists, with probability one, a real number k , $0 < k < \infty$, and independent of T such that

$$|I_1| \leq \limsup_{T \rightarrow \infty} \frac{kT_F}{T} = 0 \quad (a.s.). \tag{5.40}$$

Next consider $T' := T - T_F$, then

$$\begin{aligned}
 & J_i^{F,N}(\hat{u}_i^{F,\infty}; \hat{u}_{-i}^{F,\infty}, u^{L,\infty}) \\
 & \leq \limsup_{T \rightarrow \infty} \frac{1}{T} \int_{T_F}^T \left(\left\| z_{i,\delta_1}^{F,\infty} - \left(\frac{\eta}{N_L} \sum_{j=1}^{N_L} z_j^{L,\infty} + \frac{(1-\eta)}{N_F} \sum_{j=1}^{N_F} z_{j,\delta_1}^{F,\infty} \right) \right\|_Q^2 + \|u_{i,\delta_1}^{F,\infty}\|_R^2 \right) dt \\
 & = \limsup_{T' \rightarrow \infty} \frac{1}{T'} \int_0^{T'} \left(\left\| z_{i,\delta_1}^{F,\infty} - \left(\frac{\eta}{N_L} \sum_{j=1}^{N_L} z_j^{L,\infty} + \frac{(1-\eta)}{N_F} \sum_{j=1}^{N_F} z_{j,\delta_1}^{F,\infty} \right) \right\|_Q^2 + \|u_{i,\delta_1}^{F,\infty}\|_R^2 \right) dt \\
 & \equiv J_i^{F,N}(u_{i,\delta_1}^{F,\infty}; u_{-i,\delta_1}^{F,\infty}, u^{L,\infty}).
 \end{aligned}$$

by (5.39) and (5.40), which gives the result. \square

Denote

$$(\epsilon_{N_F})^2 := \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left\| \psi_{\delta_1}^{F,\infty} - \frac{1}{N_F} \sum_{i=1}^{N_F} z_{i,\delta_1}^{F,\infty} \right\|^2 dt \quad (5.41)$$

where $z_{i,\delta_1}^{F,\infty}$ is the closed-loop solution of the i -th follower's dynamics (5.4) with the MF control law (5.35), and $\psi_{\delta_1}^{F,\infty}(\cdot)$ is the infinite population centroid of the followers (5.18) when the defining parameter of the leaders reference trajectory $h(\cdot)$ is assumed to be the true one, $\delta_1 \in \Delta$.

LEMMA 5.4. *Assume (A5.1)-(A5.4) hold. Then we have $\lim_{N_F \rightarrow \infty} \epsilon_{N_F} = 0$ (a.s.).*

Proof. We have

$$\begin{aligned}
 (\epsilon_{N_F})^2 & \leq 2 \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left\| \frac{1}{N_F} \sum_{i=1}^{N_F} \bar{z}_{i,\delta_1}^{F,\infty} - \frac{1}{N_F} \sum_{i=1}^{N_F} z_{i,\delta_1}^{F,\infty} \right\|^2 dt \\
 & \quad + 2 \sup_{t \geq 0} \left\| \psi_{\delta_1}^{F,\infty}(t) - \frac{1}{N_F} \sum_{i=1}^{N_F} \bar{z}_{i,\delta_1}^{F,\infty}(t) \right\|^2.
 \end{aligned} \quad (5.42)$$

But, by Lemma 5.3 in [106] we have (a.s.)

$$\lim_{N_F \rightarrow \infty} \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left\| \frac{1}{N_F} \sum_{i=1}^{N_F} \left(\bar{z}_{i,\delta_1}^{F,\infty} - z_{i,\delta_1}^{F,\infty} \right) \right\|^2 dt = 0. \quad (5.43)$$

In addition,

$$\psi_{\delta_1}^{F,\infty}(\cdot) - \frac{1}{N_F} \sum_{i=1}^{N_F} \bar{z}_{i,\delta_1}^{F,\infty}(\cdot) = \int_{\Theta_F} \bar{z}_{\theta,\delta_1}^{F,\infty}(\cdot) dF^F(\theta) - \int_{\Theta_F} \bar{z}_{\theta,\delta_1}^{F,\infty}(\cdot) dF_{N_F}^F(\theta).$$

So by the weak convergence of F_N^F to F^F (Assumption **(A5.2)**) we have *a.s.* dF^F (see (33) in [106])

$$\lim_{N_F \rightarrow \infty} \sup_{t \geq 0} \left\| \psi_{\delta_1}^{F,\infty}(t) - \frac{1}{N_F} \sum_{i=1}^{N_F} \bar{z}_{i,\delta_1}^{F,\infty}(t) \right\| = 0. \quad (5.44)$$

Hence, by (5.42)-(5.44) we get $\lim_{N_F \rightarrow \infty} \epsilon_{N_F} = 0$ (*a.s.*). \square

THEOREM 5.7. *Assume **(A5.1)**-**(A5.4)** hold. Then there exists a random M , $0 < M(\omega) < \infty$, such that for $N_L \geq M(\omega)$ and observation sizes $m_i \geq M(\omega)$, $1 \leq i \leq N_F$, the set of the followers' MF control laws is an *a.s.* ϵ_N -Nash equilibrium, *i.e.*, for any i , $1 \leq i \leq N_F$, we have (*a.s.*)*

$$\begin{aligned} J_i^{F,N}(\hat{u}_i^{F,\infty}; \hat{u}_{-i}^{F,\infty}, u^{L,\infty}) - \epsilon_N &\leq \inf_{u_i \in \mathcal{U}_i^F} J_i^{F,N}(u_i; \hat{u}_{-i}^{F,\infty}, u^{L,\infty}) \\ &\leq J_i^{F,N}(\hat{u}_i^{F,\infty}; \hat{u}_{-i}^{F,\infty}, u^{L,\infty}) \end{aligned} \quad (5.45)$$

where $\epsilon_N = O(\epsilon_{N_L} + \epsilon_{N_F})$ *a.s.*, ($\epsilon_N \approx O(1/\sqrt{N})$).

Proof. The second inequality in (5.45) is trivial. Here, we shall prove the first inequality.

For any adaptive follower i , $1 \leq i \leq N$, by Lemma 5.3 there exists a random M , $0 < M(\omega) < \infty$, such that for $N_L \geq M(\omega)$ and observation sizes $m_i \geq M(\omega)$ we have

$$J_i^{F,N}(\hat{u}_i^{F,\infty}; \hat{u}_{-i}^{F,\infty}, u^{L,\infty}) \leq J_i^{F,N}(u_{i,\delta_1}^{F,\infty}; u_{-i,\delta_1}^{F,\infty}, u^{L,\infty}) \quad \textit{a.s.} \quad (5.46)$$

But,

$$\begin{aligned}
 & J_i^{F,N}(u_{i,\delta_1}^{F,\infty}; u_{-i,\delta_1}^{F,\infty}, u^{L,\infty}) \\
 & \equiv \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(\left\| z_{i,\delta_1}^{F,\infty} - \left(\frac{\eta}{N_L} \sum_{j=1}^{N_L} z_j^{L,\infty} + \frac{(1-\eta)}{N_F} \sum_{j=1}^{N_F} z_{j,\delta_1}^{F,\infty} \right) \right\|_Q^2 + \|u_{i,\delta_1}^{F,\infty}\|_R^2 \right) dt \\
 & \leq \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(\left\| z_{i,\delta_1}^{F,\infty} - \left(\eta \psi^{L,\infty} + (1-\eta) \psi_{\delta_1}^{L,\infty} \right) \right\|_Q^2 + \|u_{i,\delta_1}^{F,\infty}\|_R^2 \right) dt + I_N \quad (5.47)
 \end{aligned}$$

where I_N denotes the rest of the terms.

We now show that $I_N = O(\epsilon_{N_L} + \epsilon_{N_F})$.

We have

$$\begin{aligned}
 I_N & := \limsup_{T \rightarrow \infty} \frac{\eta^2}{T} \int_0^T \left(\left\| \psi^{L,\infty} - \frac{1}{N_L} \sum_{i=1}^{N_L} z_i^{L,\infty} \right\|_Q^2 \right) dt \\
 & + \limsup_{T \rightarrow \infty} \frac{(1-\eta)^2}{T} \int_0^T \left(\left\| \psi_{\delta_1}^{F,\infty} - \frac{1}{N_F} \sum_{i=1}^{N_F} z_{i,\delta_1}^{F,\infty} \right\|_Q^2 \right) dt \\
 & + 2\eta(1-\eta) \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(\psi^{L,\infty} - \frac{1}{N_L} \sum_{i=1}^{N_L} z_i^{L,\infty} \right)^T Q \left(\psi_{\delta_1}^{F,\infty} - \frac{1}{N_F} \sum_{i=1}^{N_F} z_{i,\delta_1}^{F,\infty} \right) dt \\
 & + 2\eta \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(z_{i,\delta_1}^{F,\infty} - \left(\eta \psi^{L,\infty} + (1-\eta) \psi_{\delta_1}^{L,\infty} \right) \right)^T Q \left(\psi^{L,\infty} - \frac{1}{N_L} \sum_{i=1}^{N_L} z_i^{L,\infty} \right) dt \\
 & + 2(1-\eta) \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(z_{i,\delta_1}^{F,\infty} - \left(\eta \psi^{L,\infty} + (1-\eta) \psi_{\delta_1}^{L,\infty} \right) \right)^T \\
 & \quad \times Q \left(\psi_{\delta_1}^{F,\infty} - \frac{1}{N_F} \sum_{i=1}^{N_F} z_{i,\delta_1}^{F,\infty} \right) dt \\
 & := I_{N_L}^1 + I_{N_F}^2 + I_{N_L, N_F}^3 + I_{N_L}^4 + I_{N_F}^5. \quad (5.48)
 \end{aligned}$$

But,

$$I_{N_L}^1 \leq \eta^2 \|Q\| \epsilon_{N_L}^2, \quad I_{N_F}^2 \leq (1-\eta)^2 \|Q\| \epsilon_{N_F}^2 \quad (5.49)$$

where ϵ_{N_L} and ϵ_{N_F} are defined in (5.38) and (5.41), respectively. By the Cauchy-Schwarz inequality we have

$$I_{N_L, N_F}^3 \leq 2\eta(1 - \eta)\|Q\|\epsilon_{N_L}\epsilon_{N_F}. \quad (5.50)$$

By the stability property of the MF control laws (5.36) there exists a real number k , $0 < k < \infty$, independent of both N_L and N_F such that (*a.s.*)

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \|z_{i, \delta_1}^{F, \infty} - (\eta\psi^{L, \infty} + (1 - \eta)\psi_{\delta_1}^{L, \infty})\|^2 dt \leq k.$$

This and the Cauchy-Schwarz inequality result in (*a.s.*)

$$I_{N_L}^4 \leq 2\eta\|Q\|\sqrt{k}\epsilon_{N_L}, \quad I_{N_F}^5 \leq 2(1 - \eta)\|Q\|\sqrt{k}\epsilon_{N_F}. \quad (5.51)$$

Hence, (5.49)-(5.51) imply that $I_N = I_{N_L}^1 + I_{N_F}^2 + I_{N_L, N_F}^3 + I_{N_L}^4 + I_{N_F}^5 = O(\epsilon_{N_L} + \epsilon_{N_F})$ (*a.s.*).

But, by the construction of the MF system of equations for the followers (5.15)-(5.19), $u_{i, \delta_1}^{F, \infty}(\cdot)$ is the optimal tracking control with respect to $u^{L, \infty}(\cdot)$ and $u_{-i, \delta_1}^{F, \infty}(\cdot)$ which collectively generate $\psi^{L, \infty}(\cdot)$ and $\psi^{F, \infty}(\cdot)$. Therefore,

$$\begin{aligned} & \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left(\|z_{i, \delta_1}^{F, \infty} - (\eta\psi^{L, \infty} + (1 - \eta)\psi_{\delta_1}^{L, \infty})\|_Q^2 + \|u_{i, \delta_1}^{F, \infty}\|_R^2 \right) dt \\ & \equiv J_i^{F, N}(u_{i, \delta_1}^{F, \infty}; \psi^{F, \infty}, \psi^{L, \infty}) \\ & \leq \inf_{u_i \in \mathcal{U}_i^F} J_i^{F, N}(u_i; \hat{u}_{-i}^{F, \infty}, u^{L, \infty}) + O(\epsilon_{N_L} + \epsilon_{N_F}) \quad a.s. \end{aligned} \quad (5.52)$$

where the last inequality follows from a similar argument as in (5.47)-(5.51). Hence, (5.46) and (5.52) imply the first inequality in (5.45). \square

5.7. Numerical Example

Consider a system of 50 agents with 20 leaders and 30 followers. For simplicity and clarity of the simulation, we assume that one follower, called the *adaptive follower*, needs to estimate the true reference trajectory but this reference trajectory is

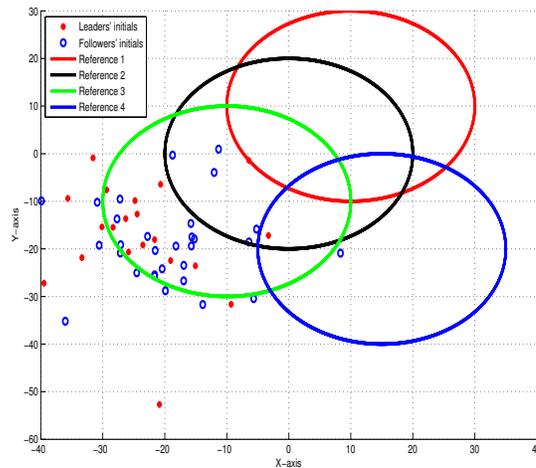


FIGURE 5.1. Four possible reference trajectories of the leaders (reference one is the true reference trajectory) corresponding to parameters $\delta_1, \delta_2, \delta_3, \delta_4$ in the set Δ , respectively, and the initial states of leaders and followers.

fully known to all other followers and all leaders. It is important to note that the followers are estimating the true reference trajectory independently. If all followers are adaptive the overall computational load increases linearly with the population size of the followers which is a manageable complexity. For the adaptive follower, the possible set of reference trajectories have the general form $[a_1 + b_1 \cos(wt) \ a_2 + b_2 \sin(wt)]$, $t \in [0, 10)$, where $\delta = (a_1, b_1, a_2, b_2, w) \in \Delta$. We assume this set Δ has four parameters including the true parameter of the reference trajectory which is reference one in Fig. 5.1. All the four possible reference trajectories generated by each parameter of the set Δ and the initial states of all the leaders and followers are shown in Fig. 5.1.

The dynamics of leaders and followers are given in (5.1) and (5.4), respectively, where the leaders and followers' MF control laws are given in (5.34) and (5.35). In this simulation we have the following parameters: (i) In (5.1) and (5.4) matrices A_i of the leaders and followers are chosen randomly from a normal probability distribution around matrix $((0.2, -0.3)^T, (-0.4, 0.2)^T)$ with identity covariance, while matrices B_i of both leaders and followers are identity matrices, and the noise intensity matrices of both leaders and followers is $C = 4I$; (ii) In (5.3) and (5.6) let $Q = I$ and $R = 0.01I$;

(iii) In (5.2) and (5.5) let $\lambda = 0.7$, $\eta = 0.6$ and the reference trajectory $h(\cdot)$ be reference one in Fig. 5.1; (iv) In (5.28) let the observation size of the adaptive follower be 15, and $D = 5I$; (v) In (5.31) let $t_0 = 0$, step-size $\tau = 1$ and the adaptive follower observes a non-empty subset of the leaders' trajectories of size 4.

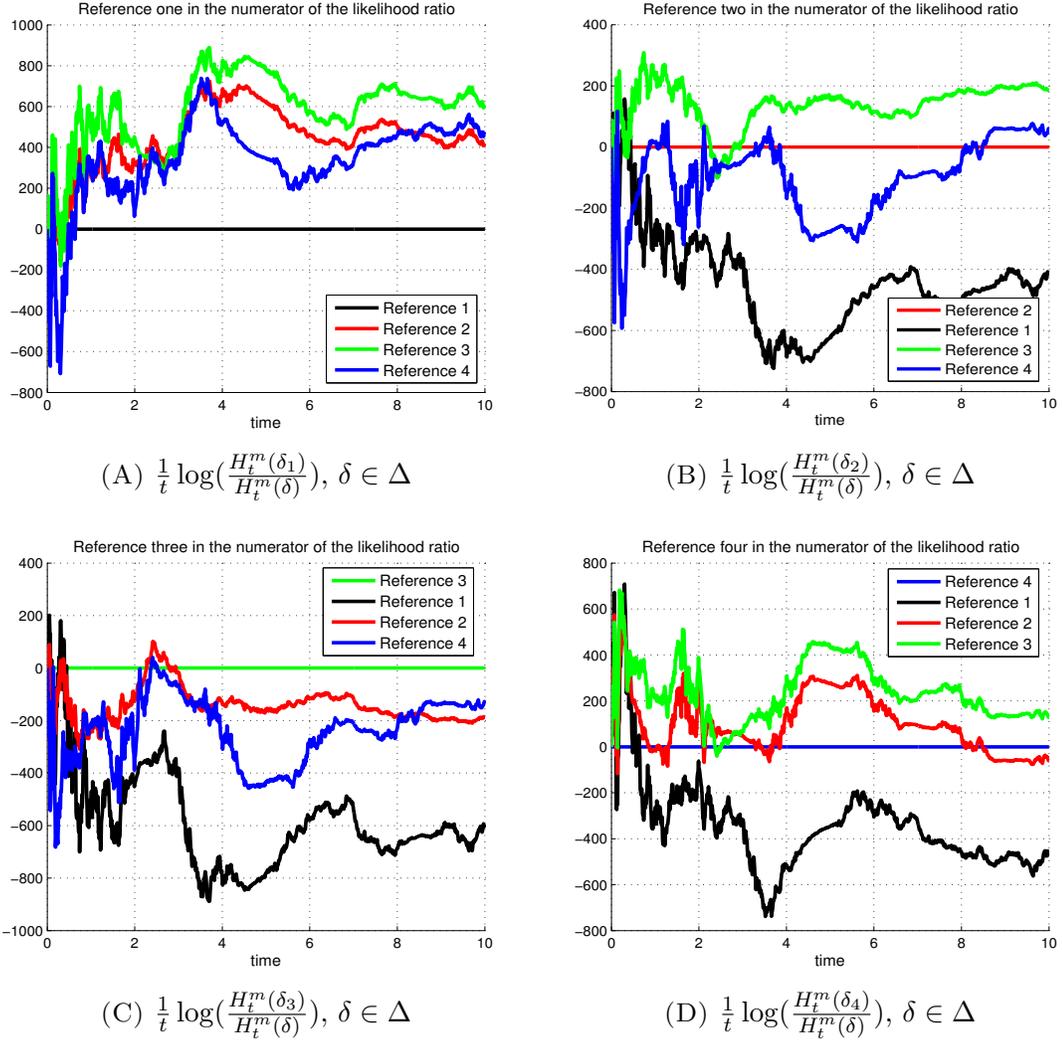


FIGURE 5.2. $(1/t) \log$ of likelihood ratios $((1/t) \log(H_t^m(\delta_i)/H_t^m(\delta_j)), \delta_i, \delta_j \in \Delta)$ such that in each figure there is a fixed parameter of set Δ in the numerator of the ratio and the parameter in the denominator changes in the set Δ .

It can be shown that this system satisfies the identifiability condition (A5.4). Figs. 5.2A, 5.2B, 5.2C and 5.2D correspond to $(1/t) \log(H_t^m(\delta_i)/H_t^m(\delta_j))$ for $\delta_i, \delta_j \in \Delta$

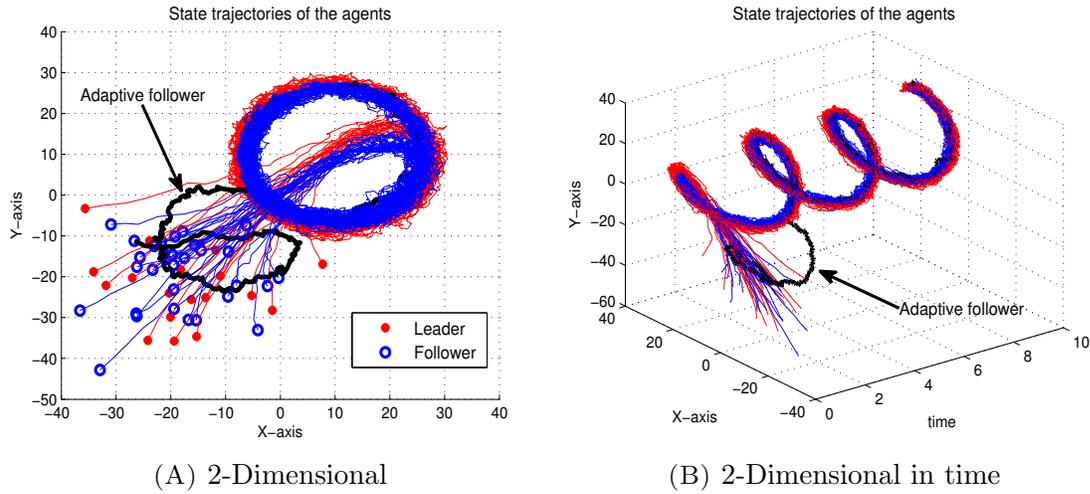


FIGURE 5.3. State trajectories of leaders (red), non-adaptive followers (blue), and the adaptive follower (black).

Δ such that, for instance, in Fig. 5.2A the parameter $\delta_1 \in \Delta$ corresponding to reference trajectory one is in the numerator and the plots in this figure display $(1/t) \log(H_t^m(\delta_1)/H_t^m(\delta_i))$ for $\delta_i \in \Delta$. Based on the LRE defined in (5.31), Figs. 5.2A, 5.2B, 5.2C and 5.2D show that for the adaptive follower $\hat{\delta}^m(t) = \delta_1$ for all $t \geq 1$ since $H_t^m(\delta_1)/H_t^m(\delta) > 1$ (i.e. $\log(H_t^m(\delta_1)/H_t^m(\delta)) > 0$) for all $\delta \in \Delta$, $\delta \neq \delta_1$.

In Figs. 5.3A, 5.3B the state trajectories of leaders, followers and the adaptive follower are shown. Based on $\hat{\delta}^m(\cdot)$ in (5.31) the adaptive follower initially considers the wrong reference trajectory signals until $t = 1$, reference 3 (generated by δ_3) from time zero to one, and then locks on the true reference trajectory (reference one generated by δ_1) as shown in Figs. 5.3A, 5.3B.

5.8. Chapter Summary

In this chapter we have developed a linear-quadratic-Gaussian (LQG) dynamic game based model of collective dynamics which include leaders, followers and an unknown (to the followers) reference trajectory for the leaders. The mean field LQG (MF LQG) equations characterizing the Nash equilibrium for infinite population systems were derived, and under appropriate conditions, they have a unique solution

leading to decentralized control laws. Furthermore, for large but finite population systems, such controls were shown to correspond to so-called ϵ -Nash equilibria.

The computation of the followers' control laws requires knowledge of the complete reference trajectory which is in general not known to the followers but is estimated by a likelihood ratio based adaptation scheme based on noisy observations taken by the followers on a random sample of leaders. Under appropriate identifiability conditions, it is established that this identification scheme is able to select the exact reference trajectory model within a finite class of candidates in a finite deterministic time almost surely as the number of samples goes to infinity. As a result, the estimation based adaptive mean field (MF) control laws of the followers together with the MF control laws of the leaders give rise to a dynamic stochastic Nash equilibrium for the overall leader-follower system.

The Leader-Follower (LF) model of this chapter is extended to the nonlinear Cucker-Smale (C-S) type cost coupling functions in the non-adaptive case [135]. In this model, the agents have similar dynamics and are coupled via their nonlinear individual cost functions which are based on the uncontrolled C-S flocking algorithm. The cost of each leader is based on a trade-off between moving its velocity toward a certain reference velocity and a C-S type weighted average of all the leaders' velocities. Followers react by tracking a C-S type weighted average of the velocities of all agents (leaders and followers).

For this controlled flocking dynamic game problem, similar to the analysis in Chapter 4 (see [?, 132]), we derive two sets of coupled deterministic equations approximating the stochastic model in the large population limit. These sets of equations consist of coupled Hamilton-Jacobi-Bellmann (HJB) and Fokker-Planck-Kolmogorov (FPK) equations in the control optimized form, and an infinite population cost coupling function.

Subject to the existence of unique solutions to these systems of equations we show that: (i) the set of MF control laws for the leaders possesses an ϵ_N -Nash equilibrium property with respect to all other leaders, (ii) the set of MF control laws for the

followers is almost surely ϵ_N -optimal with respect to all the other agents, and (iii) $\epsilon_N \rightarrow 0$ as the system's population size, N , goes to infinity. Furthermore, the MF system for the leaders and followers with linear coupling cost functions is analyzed similar to the analysis in Chapter 3.

5.9. Appendix

Proof of Proposition 5.1: Here, we shall prove part (a). Part (b) follows directly from part (a).

The terms dy^M and $z_\delta^{L,M}$:

The observation process (5.28) of a generic adaptive follower with the observation subset $O := \{L_1, \dots, L_M\} \subset L$ of cardinality M is given by

$$\begin{aligned} dy^M &= \left(\frac{1}{M} \sum_{i=1}^M z_{i,\delta_1}^L \right) dt + \frac{1}{M} \sum_{i=1}^M D_i dv_i \\ &\equiv \left(\frac{1}{M} \sum_{i=1}^M \bar{z}_{i,\delta_1}^L + \frac{1}{M} \sum_{i=1}^M \tilde{z}_{i,\delta_1}^L \right) dt + \frac{1}{M} \sum_{i=1}^M D_i dv_i \end{aligned} \quad (5.53)$$

where $\tilde{z}_{i,\delta_1}^L(\cdot) := z_{i,\delta_1}^L(\cdot) - \bar{z}_{i,\delta_1}^L(\cdot)$ in which $z_{i,\delta_1}^L(\cdot)$ and $\bar{z}_{i,\delta_1}^L(\cdot)$ are respectively the state trajectory and its expected value of the L_i -th leader, $1 \leq i \leq M$, where the defining parameter of $h(\cdot)$ is assumed to be $\delta_1 \in \Delta$. For each parameter $\delta \in \Delta$, the closed-loop solution of the L_i -th leader's dynamics, $1 \leq i \leq M$, in (5.1) with dynamical parameter $\theta_i = [A_i, B_i, C_i]$ is

$$z_{i,\delta}^L(t) = e^{\Gamma_i t} z_i^L(0) - \int_0^t e^{\Gamma_i(t-\tau)} B_i R^{-1} B_i^T s_{i,\delta}^L(\tau) d\tau + \int_0^t e^{\Gamma_i(t-\tau)} C_i dw_i^L$$

where $s_{i,\delta}^L(\cdot)$ is the solution of the leaders' offset tracking equation (5.11) given in (5.20) in which the defining parameter of $h(\cdot)$ is assumed to be $\delta \in \Delta$. Furthermore, the expected value of the corresponding closed-loop solution of the L_i -th leader, $1 \leq i \leq M$, as given in (5.21) is

$$\bar{z}_{i,\delta}^L(t) = e^{\Gamma_i t} \bar{z}^L(0) - \int_0^t e^{\Gamma_i(t-\tau)} B_i R^{-1} B_i^T s_{i,\delta}^L(\tau) d\tau.$$

Therefore, we have $\tilde{z}_{i,\delta_1}^L = z_{i,\delta_1}^L - \bar{z}_{i,\delta_1}^L = e^{\Gamma_i t}(z_i^L(0) - \bar{z}^L(0)) + \int_0^t e^{\Gamma_i(t-\tau)} C_i dw_i^L$. Hence, for any $\delta \in \Delta$ we may write

$$\begin{aligned} z_\delta^{L,M}(t) &:= \frac{1}{M} \sum_{i=1}^M z_{i,\delta}^L(t) = \psi_\delta^{L,\infty}(t) + (\bar{z}_\delta^{L,M}(t) - \psi_\delta^{L,\infty}(t)) \\ &\quad + \frac{1}{M} \sum_{i=1}^M e^{\Gamma_i t} (z_i^L(0) - \bar{z}^L(0)) + \frac{1}{M} \sum_{i=1}^M \int_0^t e^{\Gamma_i(t-\tau)} C_i dw_i^L \end{aligned} \quad (5.54)$$

where $\psi_\delta^{L,\infty}(\cdot)$ is the deterministic infinite population leaders' centroid computed from the leaders' MF system of equations, (5.11)-(5.14), when the defining parameter of $h(\cdot)$ is assumed to be $\delta \in \Delta$. In a similar way, we may write (5.53) as

$$\begin{aligned} dy^M &= \psi_{\delta_1}^{L,\infty} dt + (\bar{z}_{\delta_1}^{L,M} - \psi_{\delta_1}^{L,\infty}) dt + \left(\frac{1}{M} \sum_{i=1}^M e^{\Gamma_i t} (z_i^L(0) - \bar{z}^L(0)) \right. \\ &\quad \left. + \frac{1}{M} \sum_{i=1}^M \int_0^t e^{\Gamma_i(t-\tau)} C_i dw_i^L \right) dt + \frac{1}{M} \sum_{i=1}^M D_i dv_i \end{aligned} \quad (5.55)$$

where $\bar{z}_{\delta_1}^{L,M}(\cdot) := 1/M \sum_{i=1}^M \bar{z}_{i,\delta_1}^L(\cdot)$.

The term $\int_0^t (z_{\delta,s}^{L,M})^T dy_s^M$:

By (5.54) and (5.55) we have

$$\begin{aligned} \int_0^t (z_{\delta,s}^{L,M})^T dy_s^M &\equiv \int_0^t \left(\psi_{\delta,s}^{L,\infty} + (\bar{z}_{\delta,s}^{L,M} - \psi_{\delta,s}^{L,\infty}) \right. \\ &\quad \left. + \frac{1}{M} \sum_{i=1}^M e^{\Gamma_i s} (z_i^L(0) - \bar{z}^L(0)) + \frac{1}{M} \sum_{i=1}^M \int_0^s e^{\Gamma_i(s-\tau)} C_i dw_i^L \right)^T \\ &\quad \times \left(\psi_{\delta_1,s}^{L,\infty} ds + (\bar{z}_{\delta_1,s}^{L,M} - \psi_{\delta_1,s}^{L,\infty}) ds + \left(\frac{1}{M} \sum_{i=1}^M e^{\Gamma_i s} (z_i^L(0) - \bar{z}^L(0)) \right. \right. \\ &\quad \left. \left. + \frac{1}{M} \sum_{i=1}^M \int_0^s e^{\Gamma_i(s-\tau)} C_i dw_i^L \right) ds + \frac{1}{M} \sum_{i=1}^M D_i dv_i \right) \end{aligned} \quad (5.56)$$

where $z_{\delta,t}^{L,M} := z_\delta^{L,M}(t)$, $\bar{z}_{\delta,t}^{L,M} := \bar{z}_\delta^{L,M}(t)$, $y_t^M := y^M(t)$ and $\psi_{\delta,t}^{L,\infty} := \psi_\delta^{L,\infty}(t)$.

In the rest of the proof take fixed $\delta \in \Delta$ and $0 < T < \infty$. Let

$$\begin{aligned} \epsilon_{M,\delta} &:= \sup_{t \geq 0} \|\bar{z}_\delta^{L,M}(t) - \psi_\delta^{L,\infty}(t)\| \\ &= \sup_{t \geq 0} \left\| \int_{\Theta_L} \bar{z}_{\delta,\theta}^L(t) dF_M^L(\theta) - \int_{\Theta_L} \bar{z}_{\delta,\theta}^L(t) dF^L(\theta) \right\| \end{aligned}$$

where $F_M^L(x)$ is the empirical distribution associated with the leaders. By **(A5.2)** and similar to (33) in [106] we have $\lim_{M \rightarrow \infty} \epsilon_{M,\delta} = 0$ (*a.s.* dF^L). Therefore,

$$\int_0^T \|\bar{z}_{\delta,s}^{L,M} - \psi_{\delta,s}^{L,\infty}\|^2 ds \leq T \epsilon_{M,\delta}^2 \rightarrow 0 \quad \text{a.s. } dF^L \quad (5.57)$$

as M goes to infinity. Hence, (5.57) and the Cauchy-Schwarz inequality imply that

$$\begin{aligned} &\lim_{M \rightarrow \infty} \left| \int_0^T \left(\bar{z}_{\delta,s}^{L,M} - \psi_{\delta,s}^{L,\infty} \right)^T \left(\bar{z}_{\delta_1,s}^{L,M} - \psi_{\delta_1,s}^{L,\infty} \right) ds \right| \\ &\leq \lim_{M \rightarrow \infty} \left(\int_0^T \|\bar{z}_{\delta,s}^{L,M} - \psi_{\delta,s}^{L,\infty}\|^2 ds \right)^{1/2} \left(\int_0^T \|\bar{z}_{\delta_1,s}^{L,M} - \psi_{\delta_1,s}^{L,\infty}\|^2 ds \right)^{1/2} \\ &= \lim_{M \rightarrow \infty} T \epsilon_{M,\delta} \epsilon_{M,\delta_1} = 0 \quad (\text{a.s.}). \end{aligned} \quad (5.58)$$

Since for any $\delta \in \Delta$, $\psi_\delta^{L,\infty}(\cdot) \in C_n^b$, there exists a real number k , $0 < k < \infty$ independent of M such that

$$\int_0^T \|\psi_{\delta,s}^{L,\infty}\|^2 \leq kT. \quad (5.59)$$

By (5.57), (5.59) and the Cauchy-Schwarz inequality we have (*a.s.*)

$$\lim_{M \rightarrow \infty} \left| \int_0^T \left(\psi_{\delta,s}^{L,\infty} \right)^T \left(\bar{z}_{\delta_1,s}^{L,M} - \psi_{\delta_1,s}^{L,\infty} \right) ds \right| \leq \lim_{M \rightarrow \infty} T \sqrt{k} \epsilon_{M,\delta_1} = 0 \quad (5.60)$$

and

$$\lim_{M \rightarrow \infty} \left| \int_0^T \left(\bar{z}_{\delta,s}^{L,M} - \psi_{\delta,s}^{L,\infty} \right)^T \psi_{\delta_1,s}^{L,\infty} ds \right| \leq \lim_{M \rightarrow \infty} T \sqrt{k} \epsilon_{M,\delta} = 0 \quad (5.61)$$

where k is a fixed real number independent of M given in (5.59).

Analysis of disturbance terms via SLLN:

By the Strong Law of Large Numbers (SLLN) [42] we have

$$\begin{aligned} \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{i=1}^M \int_0^T e^{\Gamma_i(T-\tau)} C_i dw_i^L &= 0 \quad a.s. \\ \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{i=1}^M \int_0^T D_i dv_i &= 0 \quad a.s. \\ \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{i=1}^M e^{\Gamma_i T} (z_i^L(0) - \bar{z}^L(0)) &= 0 \quad a.s. \end{aligned} \quad (5.62)$$

and

$$\lim_{M \rightarrow \infty} \int_0^T (\psi_{\delta,s}^{L,\infty})^T \left(\frac{1}{M} \sum_{i=1}^M D_i dv_i \right) = 0 \quad a.s. \quad (5.63)$$

since for any $\delta \in \Delta$, $\psi_{\delta}^{L,\infty}(\cdot) \in C_n^b$. Similarly, we have (*a.s.*)

$$\lim_{M \rightarrow \infty} \int_0^T (\bar{z}_{\delta,s}^{L,M} - \psi_{\delta,s}^{L,\infty})^T \left(\frac{1}{M} \sum_{i=1}^M D_i dv_i \right) = 0 \quad (5.64)$$

$$\lim_{M \rightarrow \infty} \int_0^T \left(\frac{1}{M} \sum_{i=1}^M e^{\Gamma_i s} (z_i^L(0) - \bar{z}^L(0)) \right)^T \left(\frac{1}{M} \sum_{i=1}^M D_i dv_i \right) = 0 \quad (5.65)$$

$$\lim_{M \rightarrow \infty} \int_0^T \left(\frac{1}{M} \sum_{i=1}^M \int_0^s e^{\Gamma_i(s-\tau)} C_i dw_i^L \right)^T \left(\frac{1}{M} \sum_{i=1}^M D_i dv_i \right) = 0. \quad (5.66)$$

Analysis of influence of initial conditions:

Since for $t \geq 0$

$$\left\| \frac{1}{M} \sum_{i=1}^M e^{\Gamma_i t} (z_i^L(0) - \bar{z}^L(0)) \right\| \leq \gamma e^{-\rho t} \sup_{\theta_i \in \Theta_L} \|z_i^L(0) - \bar{z}^L(0)\| < \infty, \quad (5.67)$$

with the right hand side independent of M , and where positive γ and ρ are given in Remark 5.1, Lebesgue dominated convergence theorem [42] implies that

$$\begin{aligned} &\lim_{M \rightarrow \infty} \left(\int_0^T \left\| \frac{1}{M} \sum_{i=1}^M e^{\Gamma_i s} (z_i^L(0) - \bar{z}^L(0)) \right\|^2 ds \right) \\ &= \int_0^T \lim_{M \rightarrow \infty} \left\| \frac{1}{M} \sum_{i=1}^M e^{\Gamma_i s} (z_i^L(0) - \bar{z}^L(0)) \right\|^2 ds = 0 \quad a.s. \end{aligned} \quad (5.68)$$

where the last equality is obtained by (5.62). Again, by the Cauchy-Schwarz inequality we have (*a.s.*)

$$\lim_{M \rightarrow \infty} \int_0^T \left(\frac{1}{M} \sum_{i=1}^M e^{\Gamma_i s} (z_i^L(0) - \bar{z}^L(0)) \right)^T \psi_{\delta, s}^{L, \infty} ds = 0 \quad (5.69)$$

by (5.59) and (5.68); and (*a.s.*)

$$\lim_{M \rightarrow \infty} \int_0^T \left(\frac{1}{M} \sum_{i=1}^M e^{\Gamma_i s} (z_i^L(0) - \bar{z}^L(0)) \right)^T \left(\bar{z}_{\delta, s}^{L, M} - \psi_{\delta, s}^{L, \infty} \right) ds = 0 \quad (5.70)$$

by (5.57) and (5.68). By the same argument as in proving Theorem 5.1 in [106] (see (31) in [106]) we get

$$\lim_{M \rightarrow \infty} \int_0^T \left\| \frac{1}{M} \sum_{i=1}^M \int_0^s e^{\Gamma_i (s-\tau)} C_i dw_i^L \right\|^2 ds = 0 \quad (a.s.). \quad (5.71)$$

Again by the Cauchy-Schwarz inequality we have (*a.s.*)

$$\lim_{M \rightarrow \infty} \int_0^T \left(\psi_{\delta, s}^{L, \infty} \right)^T \left(\frac{1}{M} \sum_{i=1}^M \int_0^s e^{\Gamma_i (s-\tau)} C_i dw_i^L \right) ds = 0 \quad (5.72)$$

by (5.59) and (5.71),

$$\lim_{M \rightarrow \infty} \int_0^T \left(\bar{z}_{\delta, s}^{L, M} - \psi_{\delta, s}^{L, \infty} \right)^T \left(\frac{1}{M} \sum_{i=1}^M \int_0^s e^{\Gamma_i (s-\tau)} C_i dw_i^L \right) ds = 0 \quad (5.73)$$

by (5.57) and (5.71), and

$$\begin{aligned} & \lim_{M \rightarrow \infty} \int_0^T \left(\frac{1}{M} \sum_{i=1}^M e^{\Gamma_i s} (z_i^L(0) - \bar{z}^L(0)) \right)^T \\ & \quad \times \left(\frac{1}{M} \sum_{i=1}^M \int_0^s e^{\Gamma_i (s-\tau)} C_i dw_i^L \right) ds = 0 \end{aligned} \quad (5.74)$$

by (5.68) and (5.71).

Conclusion of the asymptotic analysis:

By (5.58), (5.60)-(5.61), (5.63)-(5.66), (5.68)-(5.70) and (5.71)-(5.74) we obtain (a.s.)

$$\lim_{M \rightarrow \infty} \exp \left(\int_0^T (z_{\delta,s}^{L,M})^T dy_s^M \right) = \exp \left(\int_0^T (\psi_{\delta,s}^{L,\infty})^T \psi_{\delta_1,s}^{L,\infty} ds \right) \quad (5.75)$$

from (5.56) for any fixed $\delta \in \Delta$ and T , $0 < T < \infty$, and in an analogous way one can show that (a.s.)

$$\lim_{M \rightarrow \infty} \exp \left(\int_0^T \|z_{\delta,s}^{L,M}\|^2 ds \right) = \exp \left(\int_0^T \|\psi_{\delta,s}^{L,\infty}\|^2 ds \right). \quad (5.76)$$

Finally, (5.75) and (5.76) imply that for any fixed $\delta \in \Delta$, and t , $0 < t < \infty$, we have (a.s.)

$$\begin{aligned} \lim_{M \rightarrow \infty} L_t^M(\delta) &\equiv \lim_{M \rightarrow \infty} \exp \left(\int_0^t (z_{\delta,s}^{L,M})^T dy_s^M - \frac{1}{2} \int_0^t \|z_{\delta,s}^{L,M}\|^2 ds \right) \\ &= L_t^\infty(\delta) \equiv \exp \left(\int_0^t (\psi_{\delta,s}^{L,\infty})^T \psi_{\delta_1,s}^{L,\infty} ds - \frac{1}{2} \int_0^t \|\psi_{\delta,s}^{L,\infty}\|^2 ds \right). \end{aligned}$$

□

Proof of Theorem 5.5: Part (a) is a special case of Theorem 4.1 in [106]. Here, we broadly follow the same approach to prove Part (b). For an adaptive follower i , $1 \leq i \leq N_F$, with dynamical parameter $\theta_i = [A_i, B_i, C_i] \in \Theta_L$ by application of the adaptive MF control law (5.35) we have the closed-loop solution

$$\begin{aligned} \hat{z}_i^{F,\infty}(t) &= e^{\Gamma_i t} z_i^F(0) - \int_0^t e^{\Gamma_i(t-\tau)} B_i R^{-1} B_i^T \hat{s}_i^{F,\infty}(\tau) d\tau \\ &\quad + \int_0^t e^{\Gamma_i(t-\tau)} C_i dw_i^F(\tau), \quad t \geq 0 \end{aligned} \quad (5.77)$$

where $\hat{s}_i^{F,\infty}(\cdot)$ is the solution of the tracking offset equation (5.15) given by

$$\hat{s}_i^{F,\infty}(t) = - \int_t^\infty e^{-\Gamma_i^T(t-\tau)} Q \left(\eta \psi_{\delta}^{L,\infty}(\tau) + (1-\eta) \psi_{\delta}^{F,\infty}(\tau) \right) d\tau. \quad (5.78)$$

Denote $k_L := \sup_{\delta \in \Delta} \|\psi_{\delta}^{L,\infty}\|_\infty$, $k_F := \sup_{\delta \in \Delta} \|\psi_{\delta}^{F,\infty}\|_\infty$, and $k' = \max(k_L, k_F)$ then by Assumption (A5.3) we have $k' < \infty$. Subsequently, from (5.78), we get $\|\hat{s}_i^{F,\infty}\|_\infty \leq$

$\gamma\|Q\|k'/\rho =: k_s < \infty$, where positive γ and ρ are given in Remark 5.1, and hence,

$$\left\| \int_0^t e^{\Gamma_i(t-\tau)} B_i R^{-1} B_i^T \hat{s}_i^{F,\infty}(\tau) d\tau \right\| \leq \gamma k_s \|R^{-1}\| M_B^2 / \rho =: k_1 < \infty,$$

where $M_B := \sup_{\theta_i \in \Theta_F} \|B_i\| < \infty$ since Θ_F is compact. Therefore, we have (a.s.)

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left\| \int_0^t e^{\Gamma_i(t-\tau)} B_i R^{-1} B_i^T \hat{s}_i^{F,\infty}(\tau) d\tau \right\|^2 dt \leq k_1 \quad (5.79)$$

and

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left\| \int_0^t e^{\Gamma_i(t-\tau)} B_i R^{-1} B_i^T \hat{s}_i^{F,\infty}(\tau) d\tau \right\|^2 dt \leq k_1^2. \quad (5.80)$$

Since Γ_i is asymptotically stable (Remark 5.1) we have

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \|e^{\Gamma_i t} z_i^F(0)\|^2 dt = \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \|e^{\Gamma_i t} z_i^F(0)\|^2 dt = 0 \quad (a.s.). \quad (5.81)$$

By the same argument as in proving Theorem 4.1 in [106] we get (a.s.)

$$\begin{aligned} & \limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \left\| \int_0^t e^{\Gamma_i(t-\tau)} C_i dw_i^F \right\|^2 dt \\ &= \int_0^\infty \text{tr}(e^{\Gamma_i t} C_i C_i^T e^{\Gamma_i^T t}) dt \leq \frac{\gamma^2}{2\rho} \sup_{\theta_i \in \Theta_F} \|C_i\|^2 =: k_2. \end{aligned} \quad (5.82)$$

Thus, it follows from (5.79)-(5.82) that (a.s.)

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \|\hat{z}_i^{F,\infty}(t)\|^2 dt \leq k_1^2 + k_2 =: k < \infty. \quad (5.83)$$

Since $\hat{u}_i^{F,\infty}(\cdot) = -R^{-1} B_i^T \left(\Pi_i \hat{z}_i^{F,\infty}(\cdot) + \hat{s}_i^{F,\infty}(\cdot) \right)$ we have

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \int_0^T \|\hat{u}_i^{F,\infty}(t)\|^2 dt \leq \|R^{-1}\|^2 M_B^2 (M_P^2 k + k_s^2 + 2M_P k k_s) < \infty \quad a.s. \quad (5.84)$$

where $M_P := \sup_{\theta_i \in \Theta_F} \|\Pi_i\| < \infty$ since Θ_F is compact. Since k , k_s , M_B and M_P are independent of i and N_F , by (5.83) and (5.84) we obtain (5.37). \square

CHAPTER 6

Mean Field Game Theory for Nonlinear Stochastic Dynamical Systems with Major and Minor Agents

This chapter studies a large population dynamic game involving nonlinear stochastic dynamical systems with agents of the following mixed types: (i) a major agent, and (ii) a population of N minor agents where N is very large. The major and minor (MM) agents are coupled via both: (i) their individual nonlinear stochastic dynamics, and (ii) their individual finite time horizon nonlinear cost functions. This problem is approached by the so-called ϵ -Nash Mean Field Game (ϵ -NMFG) theory.

A distinct feature of the mixed agent MFG problem is that even asymptotically (as the population size N approaches infinity) the noise process of the major agent causes random fluctuation of the mean field behaviour of the minor agents. To deal with this, the overall asymptotic ($N \rightarrow \infty$) mean field game problem is decomposed into: (i) two non-standard stochastic optimal control problems with random coefficient processes which yield forward adapted stochastic best response control processes determined from the solution of (backward in time) stochastic Hamilton-Jacobi-Bellman (SHJB) equations, and (ii) two stochastic (coefficient) McKean-Vlasov (SMV) equations which characterize the state of the major agent and the measure determining the mean field behaviour of the minor agents. (i) and (ii) are coupled in the

following way: the forward adapted stochastic best response control processes in (i) involve the state of the major agent and the distribution measure corresponding to the mean field behaviour of the minor agents in (ii) where these in turn depend upon the best response control processes themselves. By introducing density functions corresponding to the state distribution measures of the agents the SMV equations may be expressed in the form of stochastic (coefficient) Fokker-Planck-Kolmogorov (SFPK) equations.

Existence and uniqueness of the solution to the Stochastic Mean Field (SMF) system (SHJB and SMV equations) is established by a fixed point argument in the Wasserstein space of random probability measures. In the case that minor agents are coupled to the major agent only through their cost functions, the ϵ_N -Nash equilibrium property of the SMF best response control possess is shown for a finite N population system where $\epsilon_N = O(1/\sqrt{N})$.

As a particular but important case, the results of Nguyen and Huang [124] for MM-SMF linear-quadratic-Gaussian (LQG) systems with homogeneous population are retrieved, and, in addition, the results of this chapter are illustrated with a major and minor agent version of a game model of the synchronization of coupled nonlinear oscillators.

6.1. Introduction

Recently, Huang [75] introduced a large population LQG dynamic game model with mean field couplings which involves not only a large number of multi-class minor agents but also a major agent with a significant influence on minor agents (see [70, 71, 123] for static cooperative games of agents with different influences or so-called “mixed agents”). Since all minor agents respond to the same major agent, the mean field behaviour of minor agents in each class is directly impacted by the major agent and hence is a random process [75]. This is in contrast to the situation in the standard MFG models with only minor agents. A state-space augmentation approach for the approximation of the mean field behaviour of the minor agents is taken in order

to Markovianize the problem and hence to obtain ϵ -NMFG equilibrium strategies [75]. An extension of the model in [75] to the systems of agents with Markov jump parameters in their dynamics and random parameters in their cost functions is studied in [171] in a discrete-time setting. See also [91] for the extension of the model in [75] to the case of systems with egoistic and altruistic agents.

The model of [75] with finite classes of minor agents is extended in [124] to the case of minor agents parameterized by an infinite set of dynamical parameters where the state augmentation trick cannot be applied to obtain a finite dimensional Markov model. Due to the LQ structure of the problem an appropriate representation for the mean field behaviour of the minor agents as a random process is assumed which depends linearly on the random initial state and Brownian motion of the major agent. Appropriate approximation of the model by LQG control problems with random parameters in the dynamics and costs yields non-Markovian forward adapted ϵ -NMFG strategies resulting from backward stochastic differential equations (BSDEs) obtained by a stochastic maximum principle [124].

In this chapter we extend the LQG model for major and minor (MM) agents [75] to the case of a nonlinear stochastic dynamic games formulation of controlled McKean-Vlasov (MV) type [85]. Specifically, we consider a large population dynamic game involving nonlinear stochastic dynamical systems with agents of the following mixed types: (i) a major agent, and (ii) a population of N minor agents where N is very large. The MM agents are coupled via both: (i) their individual nonlinear stochastic dynamics, and (ii) their individual finite time horizon nonlinear cost functions. Applications of the major and minor formulation may be found in charging control of plug-in electric vehicles [117, 178], social opinion models [53] with a finite number of leaders, and power markets involving large consumers and large utilities together with many domestic consumers represented by smart meter agents and possibly large numbers of renewable energy based generators [93].

A distinctive feature of the mixed agent MFG problem is that even asymptotically (as the population size N approaches infinity) the noise process of the major agent

causes random fluctuation of the mean field behaviour of the minor agents [75, 124]. To deal with this, the overall asymptotic ($N \rightarrow \infty$) mean field game problem is decomposed into: (i) two non-standard Stochastic Optimal Control Problems (SOCPs) with random coefficient processes which yield forward adapted stochastic best response control processes determined from the solution of (backward in time) stochastic Hamilton-Jacobi-Bellman (SHJB) equations, and (ii) two stochastic (coefficient) McKean-Vlasov (SMV) equations which characterize the state of the major agent and the measure determining the mean field behaviour of the minor agents. (i) and (ii) are coupled in the following way: the forward adapted stochastic best response control processes in (i) involve the state of the major agent and the distribution measure corresponding to the mean field behaviour of the minor agents in (ii) where these in turn depend upon the best response control processes themselves.

Existence and uniqueness of the solution to the Stochastic Mean Field (SMF) system (SHJB and SMV equations) is established by a fixed point argument in the Wasserstein space of random probability measures. In the case that minor agents are coupled to the major agent only through their cost functions, the ϵ_N -Nash equilibrium property of the SMF best response control possess is shown for a finite N population system where $\epsilon_N = O(1/\sqrt{N})$. As a particular but important case, the results of Nguyen and Huang [124] for MM-SMF LQG systems with homogeneous population are retrieved. In addition, the results of this chapter are illustrated with a major and minor agent version of a game model of the synchronization of coupled nonlinear oscillators [177].

It is to be emphasized that the non-standard nature of the SOCPs in (i), which consists of the coupling through the SMV equations in (ii), arises from a distinct feature of the problem formulation. The source of this non-standard nature is the game structure whereby the minor agents are (through the Principle of Optimality) optimizing with respect to the future stochastic evolution of the major agent's state which is partly a result of that agent's future best response control actions. Not only this feature vanishes in the non-game theoretic setting of one controller with

one cost function with respect to the trajectories of all the system components (in the game situation called agents), but also in the infinite population limit of the standard ϵ -NMFG models with no major agents. This is true for completely and partially observed SOCPs. The nonstandard feature of the SOCPs here give rise to the analysis of systems with (non necessarily Markovian) stochastic parameters. Here, as in [124, 179], the theory of BSDEs (see in particular [20, 143, 145, 146]) is used in the resulting stochastic dynamic game theory. More specifically, we utilize techniques from [145] which applies the Principle of Optimality to a stochastic nonlinear control problem with random coefficients; this leads to a formulation of a SHJB equation by use of (i) a semi-martingale representation for the corresponding stochastic value function, and (ii) the Itô-Kunita formula. An application of Peng results to portfolio-consumption optimization under habit formation in complete markets is studied in [55].

The organization of the chapter is as follows. Section 6.2 is dedicated to the problem formulation. A major-minor (MM) agents mean field (MF) convergence theorem is presented in Section 6.3. A preliminary nonlinear Stochastic Optimal Control Problem (SOCP) with random parameters is studied in Section 6.4. The stochastic mean field (SMF) system of the MM agents is given in Section 6.5, and the existence and uniqueness of its solution is established in Section 6.6. Section 6.7 presents two applications of the MM MFG theory in the MM LQG formulation of Nguyen and Huang [124], and major and minor agent version of the synchronization of coupled nonlinear oscillators game model. The ϵ -Nash equilibrium property of the resulting SMF control laws is studied in Section 6.8. Finally, Section 6.9 concludes the chapter.

6.1.1. Notation and Terminology. The following notation will be used throughout the chapter. Let \mathbb{R}^n denote the n -dimensional real Euclidean space with the standard Euclidean norm $|\cdot|$ and the standard Euclidean inner product $\langle \cdot, \cdot \rangle$. The transpose of a vector (or matrix) x is denoted by x^T . $\text{tr}(A)$ denotes the trace of a square matrix A . Let $\mathbb{R}^{n \times m}$ be the Hilbert space consisting of all $(n \times m)$ -matrices

with the inner product $\langle A, B \rangle := \text{tr}(AB^T)$ and the norm $|A| := \langle A, A \rangle^{1/2}$. The set of non-negative real numbers is denoted by \mathbb{R}_+ . $T \in [0, \infty)$ is reserved to denote the terminal time. The integer N is reserved to designate the population size of the minor agents. The superscript N for a process (such as state, control or cost function) is used to indicate the dependence on the population size N . We use the subscript 0 for the major agent \mathcal{A}_0 and an integer valued subscript for an individual minor agent $\{\mathcal{A}_i : 1 \leq i \leq N\}$. At time $t \geq 0$, (i) the states of agents \mathcal{A}_0 and \mathcal{A}_i are respectively denoted by $z_0^N(t)$ and $z_i^N(t)$, $1 \leq i \leq N$, and (ii) for the system configuration of minor agents $(z_1^N(t), \dots, z_N^N(t))$ the empirical distribution δ_t^N is defined as the normalized sum of Dirac's masses, i.e., $\delta_t^N := (1/N) \sum_{i=1}^N \delta_{z_i^N(t)}$ where $\delta_{(\cdot)}$ is the Dirac measure. $C(S)$ is the set of continuous functions and $C^k(S)$ the set of k -times continuously differentiable functions on S . The symbol ∂_t denotes the partial derivative with respect to variables t . We denote D_x and D_{xx}^2 as the gradient and Hessian operators with respect to the variable x . These are respectively denoted by ∂_x and ∂_{xx}^2 when applied to a function defined on a one-dimensional domain. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$ be a complete filtered probability space. E denotes the expectation. The conditional expectation with respect to the σ -field \mathcal{V} is denoted by $E_{\mathcal{V}}$. We note that we may not display the dependence of random variables or stochastic processes on the sample point $\omega \in \Omega$. For an Euclidean space H we denote by $L_{\mathcal{G}}^2([0, T]; H)$ the space of all $\{\mathcal{G}_t\}_{t \geq 0}$ -adapted H -valued processes $f(t, \omega)$ such that $E \int_0^T |f(t, \omega)|^2 dt < \infty$.

6.2. Problem Formulation

We consider a dynamic game involving: (i) a major agent \mathcal{A}_0 , and (ii) a population of N minor agents $\{\mathcal{A}_i : 1 \leq i \leq N\}$ where N is very large. We assume homogenous minor agents although the modelling may be generalized to the case of multi-class heterogeneous minor agents [75, 85] (see [127]).

The dynamics of the agents are given by the following controlled Itô stochastic differential equations (SDEs) on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$:

$$\begin{aligned} dz_0^N(t) &= \frac{1}{N} \sum_{j=1}^N f_0[t, z_0^N(t), u_0^N(t), z_j^N(t)] dt \\ &+ \frac{1}{N} \sum_{j=1}^N \sigma_0[t, z_0^N(t), z_j^N(t)] dw_0(t), \quad z_0^N(0) = z_0(0), \quad 0 \leq t \leq T, \end{aligned} \quad (6.1)$$

$$\begin{aligned} dz_i^N(t) &= \frac{1}{N} \sum_{j=1}^N f[t, z_i^N(t), u_i^N(t), z_0^N(t), z_j^N(t)] dt \\ &+ \frac{1}{N} \sum_{j=1}^N \sigma[t, z_i^N(t), z_0^N(t), z_j^N(t)] dw_i(t), \quad z_i^N(0) = z_i(0), \quad 1 \leq i \leq N, \end{aligned} \quad (6.2)$$

with terminal time $T \in (0, \infty)$ where (i) $z_0^N : [0, T] \rightarrow \mathbb{R}^n$ is the state of the major agent \mathcal{A}_0 and $z_i^N : [0, T] \rightarrow \mathbb{R}^n$ is the state of the minor agent \mathcal{A}_i ; (ii) $u_0^N : [0, T] \rightarrow U_0$ and $u_i^N : [0, T] \rightarrow U$ are respectively the control inputs of \mathcal{A}_0 and \mathcal{A}_i ; (iii) $f_0 : [0, T] \times \mathbb{R}^n \times U_0 \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, $\sigma_0 : [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$, $f : [0, T] \times \mathbb{R}^n \times U \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $\sigma : [0, T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$; (iv) the set of initial states is given by $\{z_j^N(0) = z_j(0) : 0 \leq j \leq N\}$, and (v) the sequence $\{(w_j(t))_{t \geq 0} : 0 \leq j \leq N\}$ denotes $N + 1$ mutually independent standard Brownian motions in \mathbb{R}^m . We denote the filtration \mathcal{F}_t as the σ -field generated by the initial states and the Brownian motions up to time t , i.e., $\mathcal{F}_t := \sigma\{z_j(0), w_j(s) : 0 \leq j \leq N, 0 \leq s \leq t\}$. We also set $\mathcal{F}_t^{w_0} = \sigma\{z_0(0), w_0(s) : 0 \leq s \leq t\}$. These filtrations are augmented by all the P -null sets in \mathcal{F} .

For $0 \leq j \leq N$, $u_{-j}^N := \{u_0^N, \dots, u_{j-1}^N, u_{j+1}^N, \dots, u_N^N\}$. The objective of each agent is to minimize its finite time horizon nonlinear cost function given by

$$J_0^N(u_0^N; u_{-0}^N) := E \int_0^T \left((1/N) \sum_{j=1}^N L_0[t, z_0^N(t), u_0^N(t), z_j^N(t)] \right) dt, \quad (6.3)$$

$$J_i^N(u_i^N; u_{-i}^N) := E \int_0^T \left((1/N) \sum_{j=1}^N L[t, z_i^N(t), u_i^N(t), z_0^N(t), z_j^N(t)] \right) dt, \quad (6.4)$$

for $1 \leq i \leq N$, where $L_0 : [0, T] \times \mathbb{R}^n \times U_0 \times \mathbb{R}^n \rightarrow \mathbb{R}_+$ and $L(z_i, u_i, z_0, x) : [0, T] \times \mathbb{R}^n \times U \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+$ are the nonlinear cost-coupling functions of the major and minor agents. For $0 \leq j \leq N$, we indicate the dependence of J_j on u_j^N , u_{-j}^N and the population size N by $J_j^N(u_j^N; u_{-j}^N)$.

REMARK 6.1. *Under suitable conditions, the results of this chapter may be adapted to deal with cost-couplings of the form:*

$$L_0[t, z_0^N(t), u_0^N(t), z_j^N(t), \frac{1}{N} \sum_{j=1}^N z_j^N(t)], \quad L[t, z_i^N(t), u_i^N(t), z_0^N(t), z_j^N(t), \frac{1}{N} \sum_{j=1}^N z_j^N(t)],$$

in (6.3)-(6.4) (see Section 6.7.1).

We note that in the modelling (6.1)-(6.4) the major agent \mathcal{A}_0 has a significant influence on minor agents while each minor agent has an asymptotically negligible impact on other agents in a large N population system. The major and minor (MM) agents are coupled via both: (i) their individual nonlinear stochastic dynamics (6.1)-(6.2), and (ii) their individual finite time horizon nonlinear cost functions (6.3)-(6.4).

We note that the coupling terms may be written as functionals of the empirical distribution $\delta_{(\cdot)}^N$ by the formula $\int_{\mathbb{R}^n} \phi(x) \delta_t^N(dx) = (1/N) \sum_{i=1}^N \phi(x_i(t))$ for a bounded continuous function ϕ in \mathbb{R}^n .

6.2.1. Assumptions. Let the empirical distribution of N minor agents' initial states be defined by $F_N(x) = (1/N) \sum_{i=1}^N 1_{\{Ez_i(0) < x\}}$, where $1_{\{Ez_i(0) < x\}} = 1$ if $Ez_i(0) < x$, and $1_{\{Ez_i(0) < x\}} = 0$ otherwise. We enunciate the following assumptions:

(A6.1) The initial states $\{z_j(0) : 0 \leq j \leq N\}$ are \mathcal{F}_0 -adapted random variables mutually independent and independent of all Brownian motions $\{(w_j(t))_{t \geq 0} : 0 \leq j \leq N\}$, and there exists a constant k independent of N such that $\sup_{0 \leq j \leq N} E|z_j(0)|^2 \leq k < \infty$.

(A6.2) $\{F_N : N \geq 1\}$ converges to a probability distribution F weakly, i.e., for any bounded and continuous function ϕ on \mathbb{R}^n we have $\lim_{N \rightarrow \infty} \int_{\mathbb{R}^n} \phi(x) dF_N(x) = \int_{\mathbb{R}^n} \phi(x) dF(x)$.

(A6.3) U_0 and U are compact metric spaces.

(A6.4) The functions $f_0[t, x, u, y]$, $\sigma_0[t, x, y]$, $f[t, x, u, y, z]$ and $\sigma[t, x, y, z]$ are continuous and bounded with respect to all their parameters, and Lipschitz continuous in (x, y, z) . In addition, their first order derivatives (w.r.t. x) are all uniformly continuous and bounded with respect to all their parameters, and Lipschitz continuous in (y, z) .

(A6.5) $f_0[t, x, u, y]$ and $f[t, x, u, y, z]$ are Lipschitz continuous in u .

(A6.6) $L_0[t, x, u, y]$ and $L[t, x, u, y, z]$ are continuous and bounded with respect to all their parameters, and Lipschitz continuous in (x, y, z) . In addition, their first order derivatives (w.r.t. x) are all uniformly continuous and bounded with respect to all their parameters, and Lipschitz continuous in (y, z) .

(A6.7) (*Non-degeneracy Assumption*) There exists a positive constant α such that

$$\sigma_0[t, x, y]\sigma_0^T[t, x, y] \geq \alpha I, \quad \sigma[t, x, y, z]\sigma^T(t, x, y, z) \geq \alpha I, \quad \forall(t, x, y, z).$$

We note that if we relax the non-degeneracy assumption (A6.7) then a notion of “viscosity” like solutions seems necessary [180].

6.3. Major and Minor Agents Mean Field Convergence Theorem

We take a probabilistic approach to show a “decoupling effect” result such that a generic minor agent’s statistical properties can effectively approximate the distribution produced by all minor agents as the number of minor agents N goes to infinity (this is motivated by the analysis in Section I.1 of [162] and in Section 8.1 of [85]). Therefore, each minor agent’s state will be an independent copy of a non-linear McKean-Vlasov (MV) process as N approaches infinity.

Let $\varphi_0(\omega, t, x) : \Omega \times [0, T] \times \mathbb{R} \rightarrow U_0$ and $\varphi(\omega, t, x) : \Omega \times [0, T] \times \mathbb{R} \rightarrow U$ be two arbitrary $\mathcal{F}_t^{w_0}$ -measurable stochastic processes for which we introduce the following assumption:

(H6.4) $\varphi_0(\omega, t, x)$ and $\varphi(\omega, t, x)$ are Lipschitz continuous in x , and $\varphi_0(\omega, t, 0) \in L^2_{\mathcal{F}_t^{w_0}}([0, T]; U_0)$ and $\varphi(\omega, t, 0) \in L^2_{\mathcal{F}_t^{w_0}}([0, T]; U)$.

We assume that $\varphi_0(t, x) := \varphi_0(\omega, t, x)$ and $\varphi(t, x) := \varphi(\omega, t, x)$ are respectively used by the major and minor agents as their control laws in (6.1) and (6.2) (i.e., $u_0 = \varphi_0$ and $u_i = \varphi$ for $1 \leq i \leq N$). Then we have the following closed-loop equations with random coefficients:

$$\begin{aligned} d\hat{z}_0^N(t) &= \frac{1}{N} \sum_{j=1}^N f_0[t, \hat{z}_0^N(t), \varphi_0(t, \hat{z}_0^N(t)), \hat{z}_j^N(t)] dt \\ &\quad + \frac{1}{N} \sum_{j=1}^N \sigma_0[t, \hat{z}_0^N(t), \hat{z}_j^N(t)] dw_0(t), \quad \hat{z}_0^N(0) = z_0(0), \quad 0 \leq t \leq T, \\ d\hat{z}_i^N(t) &= \frac{1}{N} \sum_{j=1}^N f[t, \hat{z}_i^N(t), \varphi(t, \hat{z}_i^N(t)), \hat{z}_0^N(t), \hat{z}_j^N(t)] dt \\ &\quad + \frac{1}{N} \sum_{j=1}^N \sigma[t, \hat{z}_i^N(t), \hat{z}_0^N(t), \hat{z}_j^N(t)] dw_i(t), \quad \hat{z}_i^N(0) = z_i(0), \quad 1 \leq i \leq N. \end{aligned}$$

Under **(A6.4)**-**(A6.5)** and **(H6.4)** there exists a unique solution $(z_0^N(\cdot), \dots, z_N^N(\cdot))$ to the above system (see Theorem 6.16, Chapter 1 of [180], page 49).

We now introduce the McKean-Vlasov (MV) SDE system

$$\begin{aligned} d\bar{z}_0(t) &= f_0[t, \bar{z}_0(t), \varphi_0(t, \bar{z}_0(t)), \mu_t] dt + \sigma_0[t, \bar{z}_0(t), \mu_t] dw_0(t), \quad 0 \leq t \leq T, \\ d\bar{z}(t) &= f[t, \bar{z}(t), \varphi(t, \bar{z}(t)), \mu_t^0, \mu_t] dt + \sigma[t, \bar{z}(t), \mu_t^0, \mu_t] dw(t), \end{aligned}$$

with initial condition $(\bar{z}_0(0), \bar{z}(0))$, where for an arbitrary function $g \in C(\mathbb{R}^s)$ for appropriate s , and probability distributions μ_t and μ_t^0 in \mathbb{R}^n , we set

$$g[t, z, \mu_t] = \int_{\mathbb{R}^n} g[t, z, x] \mu_t(dx), \quad g[t, z, \mu_t^0, \mu_t] = \int_{\mathbb{R}^n \times \mathbb{R}^n} g[t, z, x, y] \mu_t^0(dx) \mu_t(dy),$$

when the indicated integrals converge. In the above MV system $(\bar{z}_0(\cdot), \bar{z}(\cdot), \mu_{(\cdot)}^0, \mu_{(\cdot)})$ is a “consistent solution” if $(\bar{z}_0(\cdot), \bar{z}(\cdot))$ is a solution to the above SDE system, μ_t , $0 \leq t \leq T$, is the conditional law of $\bar{z}(t)$ given $\mathcal{F}_t^{w_0}$ (i.e., $\mu_t := \mathcal{L}(\bar{z}(t) | \mathcal{F}_t^{w_0})$), and μ_t^0 , $0 \leq t \leq T$, is the unit mass measure concentrated at $\bar{z}_0(t)$ (i.e., $\mu_t^0 = \delta_{\bar{z}_0(t)}$).

Under **(A6.4)**-**(A6.5)** and **(H6.4)** it can be shown by a fixed point argument that there exists a unique solution $(\bar{z}_0(\cdot), \bar{z}(\cdot), \mu_{(\cdot)}^0, \mu_{(\cdot)})$ to the above system (see Theorem 1.1 in [162] or Theorem 6.8 below).

We also introduce the equations

$$\begin{aligned} d\bar{z}_0(t) &= f_0[t, \bar{z}_0(t), \varphi_0(t, \bar{z}_0(t)), \mu_t]dt + \sigma_0[t, \bar{z}_0(t), \mu_t]dw_0(t), \quad 0 \leq t \leq T, \\ d\bar{z}_i(t) &= f[t, \bar{z}_i(t), \varphi(t, \bar{z}_i(t)), \mu_t^0, \mu_t]dt + \sigma[t, \bar{z}_i(t), \mu_t^0, \mu_t]dw_i(t), \quad 1 \leq i \leq N, \end{aligned}$$

with initial conditions $\bar{z}_j(0) = z_j(0)$, $0 \leq j \leq N$, which can be viewed as N independent samples of the MV SDE system above. We develop a decoupling result below such that each \hat{z}_i^N , $1 \leq i \leq N$, has the natural limit \bar{z}_i in the infinite population limit (see Theorem 12 in [85]).

THEOREM 6.1. *Assume **(A6.1)**, **(A6.3)**-**(A6.5)** and **(H6.4)** hold. Then we have*

$$\sup_{0 \leq j \leq N} \sup_{0 \leq t \leq T} E|\hat{z}_j^N(t) - \bar{z}_j(t)| = O(1/\sqrt{N}), \quad (6.5)$$

where the right hand side may depend upon the terminal time T .

Proof: We will show

$$\sup_{0 \leq j \leq N} \sup_{0 \leq t \leq T} E|\hat{z}_j^N(t) - \bar{z}_j(t)|^2 = O(1/N),$$

which implies the result of the theorem by the Cauchy-Schwarz inequality. First by the inequality $(x + y)^2 \leq 2x^2 + 2y^2$, we have

$$\begin{aligned} &E|\hat{z}_0^N(t) - \bar{z}_0(t)|^2 \\ &\leq 2E \left| \int_0^t \left(\frac{1}{N} \sum_{j=1}^N f_0[s, \hat{z}_0^N, \varphi_0(s, \hat{z}_0^N), \hat{z}_j^N] - f_0[s, \bar{z}_0, \varphi_0(s, \bar{z}_0), \mu_s] \right) ds \right|^2 \\ &\quad + 2E \left| \int_0^t \left(\frac{1}{N} \sum_{j=1}^N \sigma_0[s, \hat{z}_0^N, \hat{z}_j^N] - \sigma_0[s, \bar{z}_0, \mu_s] \right) dw_0(s) \right|^2. \end{aligned}$$

By the Cauchy-Schwarz inequality and the properties of Itô integrals we then obtain

$$\begin{aligned}
 & E|\hat{z}_0^N(t) - \bar{z}_0(t)|^2 \\
 & \leq 2tE\left(\int_0^t \left|\frac{1}{N} \sum_{j=1}^N f_0[s, \hat{z}_0^N, \varphi_0(s, \hat{z}_0^N), \hat{z}_j^N] - f_0[s, \bar{z}_0, \varphi_0(s, \bar{z}_0), \mu_s]\right|^2 ds\right) \\
 & \quad + 2E\left(\int_0^t \left|\frac{1}{N} \sum_{j=1}^N \sigma_0[s, \hat{z}_0^N, \hat{z}_j^N] - \sigma_0[s, \bar{z}_0, \mu_s]\right|^2 ds\right). \tag{6.6}
 \end{aligned}$$

Clearly,

$$\begin{aligned}
 & \frac{1}{N} \sum_{j=1}^N f_0[s, \hat{z}_0^N, \varphi_0(s, \hat{z}_0^N), \hat{z}_j^N] - f_0[s, \bar{z}_0, \varphi_0(s, \bar{z}_0), \mu_s] \\
 & = \left(\frac{1}{N} \sum_{j=1}^N f_0[s, \hat{z}_0^N, \varphi_0(s, \hat{z}_0^N), \hat{z}_j^N] - \frac{1}{N} \sum_{j=1}^N f_0[s, \bar{z}_0, \varphi_0(s, \bar{z}_0), \hat{z}_j^N]\right) \\
 & \quad + \left(\frac{1}{N} \sum_{j=1}^N f_0[s, \bar{z}_0, \varphi_0(s, \bar{z}_0), \hat{z}_j^N] - \frac{1}{N} \sum_{j=1}^N f_0[s, \bar{z}_0, \varphi_0(s, \bar{z}_0), \bar{z}_j]\right) \\
 & \quad + \left(\frac{1}{N} \sum_{j=1}^N f_0[s, \bar{z}_0, \varphi_0(s, \bar{z}_0), \bar{z}_j] - f_0[s, \bar{z}_0, \varphi_0(s, \bar{z}_0), \mu_s]\right), \tag{6.7}
 \end{aligned}$$

and

$$\begin{aligned}
 & \frac{1}{N} \sum_{j=1}^N \sigma_0[s, \hat{z}_0^N, \hat{z}_j^N] - \sigma_0[s, \bar{z}_0, \mu_s] = \left(\frac{1}{N} \sum_{j=1}^N \sigma_0[s, \hat{z}_0^N, \hat{z}_j^N] - \frac{1}{N} \sum_{j=1}^N \sigma_0[s, \bar{z}_0, \hat{z}_j^N]\right) \\
 & \quad + \left(\frac{1}{N} \sum_{j=1}^N \sigma_0[s, \bar{z}_0, \hat{z}_j^N] - \frac{1}{N} \sum_{j=1}^N \sigma_0[s, \bar{z}_0, \bar{z}_j]\right) + \left(\frac{1}{N} \sum_{j=1}^N \sigma_0[s, \bar{z}_0, \bar{z}_j] - \sigma_0[s, \bar{z}_0, \mu_s]\right).
 \end{aligned}$$

Applying the inequality $(x+y+z)^2 \leq 3(x^2+y^2+z^2)$, and the Lipschitz continuity conditions of f_0 and φ_0 to (6.7) we obtain

$$\begin{aligned} & E \left(\int_0^t \left| \frac{1}{N} \sum_{j=1}^N f_0[s, \hat{z}_0^N, \varphi_0(s, \hat{z}_0^N), \hat{z}_j^N] - f_0[s, \bar{z}_0, \varphi_0(s, \bar{z}_0), \mu_s] \right|^2 ds \right. \\ & \leq 3C \int_0^t E |\hat{z}_0^N(s) - \bar{z}_0(s)|^2 ds + 3C \int_0^t E \left| \frac{1}{N} \sum_{j=1}^N \hat{z}_j^N(s) - \bar{z}_j(s) \right|^2 ds \\ & \left. + 3C \int_0^t E \left| \frac{1}{N} \sum_{j=1}^N f_0[s, \bar{z}_0, \varphi_0(s, \bar{z}_0), \bar{z}_j] - f_0[s, \bar{z}_0, \varphi_0(s, \bar{z}_0), \mu_s] \right|^2 ds, \quad (6.8) \right. \end{aligned}$$

where $C > 0$ is a constant independent of N . Due to the centring of $g_s[s, \bar{z}_0, x] := f_0[s, \bar{z}_0, \varphi_0(s, \bar{z}_0), x] - f_0[s, \bar{z}_0, \varphi_0(s, \bar{z}_0), \mu_s]$ with respect to x and the independence of \bar{z}_j and $\bar{z}_{j'}$ when $j \neq j'$, there are no cross terms in the expansion of the last term in (6.8), i.e., $E(g_s[s, \bar{z}_0, \bar{z}_j]g_s[s, \bar{z}_0, \bar{z}_{j'}]) = EE_{\mathcal{F}_t^{w_0}}(g_s[s, \bar{z}_0, \bar{z}_j]g_s[s, \bar{z}_0, \bar{z}_{j'}]) = 0$ for $j \neq j'$ (see [162], Page 175). This property together with (6.8), the boundedness of f_0 and the inequality $(\sum_{i=1}^N x_i)^2 \leq N \sum_{i=1}^N x_i^2$ yields

$$\begin{aligned} & E \left(\int_0^t \left| \frac{1}{N} \sum_{j=1}^N f_0[s, \hat{z}_0^N, \varphi_0(s, \hat{z}_0^N), \hat{z}_j^N] - f_0[s, \bar{z}_0, \varphi_0(s, \bar{z}_0), \mu_s] \right|^2 ds \right. \\ & \leq 3C \int_0^t E |\hat{z}_0^N(s) - \bar{z}_0(s)|^2 ds + \frac{3C}{N} \int_0^t \sum_{j=1}^N E |\hat{z}_j^N(s) - \bar{z}_j(s)|^2 ds + \frac{k_1(t)}{N}, \quad (6.9) \end{aligned}$$

where $k_1(t) \geq 0$ is an increasing function independent of N . Similarly, for the second term on the right hand side of (6.6) we have

$$\begin{aligned} & E \left(\int_0^t \left| \frac{1}{N} \sum_{j=1}^N \sigma_0[s, \hat{z}_0^N, \hat{z}_j^N] - \sigma_0[s, \bar{z}_0, \mu_s] \right|^2 ds \right) \\ & \leq 3C \int_0^t E |\hat{z}_0^N(s) - \bar{z}_0(s)|^2 ds + \frac{3C}{N} \int_0^t \sum_{j=1}^N E |\hat{z}_j^N(s) - \bar{z}_j(s)|^2 ds + \frac{k_1(t)}{N}. \quad (6.10) \end{aligned}$$

The inequalities (6.6), (6.9) and (6.10) imply that

$$\begin{aligned} \sup_{0 \leq t \leq T} E|\hat{z}_0^N(t) - \bar{z}_0(t)|^2 &\leq 6C(T+1) \int_0^T E|\hat{z}_0^N(s) - \bar{z}_0(s)|^2 ds \\ &+ \frac{6C(T+1)}{N} \int_0^T \sum_{j=1}^N E|\hat{z}_j^N(s) - \bar{z}_j(s)|^2 ds + \frac{2(T+1)k_1(T)}{N}. \end{aligned} \quad (6.11)$$

Second, by taking a similar approach for the i^{th} minor agent ($1 \leq i \leq N$) we get

$$\begin{aligned} \sup_{0 \leq t \leq T} E|\hat{z}_i^N(t) - \bar{z}_i(t)|^2 &\leq 8C(T+1) \int_0^T E|\hat{z}_i^N(s) - \bar{z}_i(s)|^2 ds + \frac{k(T)}{N} \\ &+ 8C(T+1) \left(\int_0^T E|\hat{z}_0^N(s) - \bar{z}_0(s)|^2 ds + \frac{1}{N} \int_0^T \sum_{j=1}^N E|\hat{z}_j^N(s) - \bar{z}_j(s)|^2 ds \right), \end{aligned} \quad (6.12)$$

where $k(T) > 0$ is independent of N .

The inequalities (6.11) and (6.12) yield

$$\begin{aligned} g^N(T) &:= \sup_{0 \leq t \leq T} E|\hat{z}_0^N(t) - \bar{z}_0(t)|^2 + \frac{1}{N} \sum_{j=1}^N \sup_{0 \leq t \leq T} E|\hat{z}_j^N(t) - \bar{z}_j(t)|^2 \\ &\leq 22C(T+1) \int_0^T \left(E|\hat{z}_0^N(s) - \bar{z}_0(s)|^2 + \frac{1}{N} \sum_{j=1}^N E|\hat{z}_j^N(s) - \bar{z}_j(s)|^2 \right) ds \\ &+ \frac{k_0(T) + k(T)}{N} \leq 22C(T+1) \int_0^T g(s) ds + \frac{k_0(T) + k(T)}{N}. \end{aligned} \quad (6.13)$$

It follows from Gronwall's Lemma that

$$g^N(T) \leq \frac{k_0(T) + k(T)}{N} \left(\exp(22C(T+1)T) \right) = O(1/N), \quad (6.14)$$

where the right hand side may only depend upon the terminal time T . This yields

$$\sup_{0 \leq t \leq T} E|\hat{z}_0^N(t) - \bar{z}_0(t)|^2 = O(1/N).$$

The inequalities (6.12) and (6.14) combined with Gronwall's Lemma imply that

$$\sup_{1 \leq i \leq N} \sup_{0 \leq t \leq T} E|\hat{z}_i^N(t) - \bar{z}_i(t)|^2 = O(1/N).$$

This completes the proof. \square

6.4. A Preliminary Nonlinear Stochastic Optimal Control Problem

Let $(W(t))_{t \geq 0}$ and $(B(t))_{t \geq 0}$ be mutually independent standard Brownian motions in \mathbb{R}^m , with $\mathcal{F}_t^{W,B} := \sigma\{W(s), B(s) : s \leq t\}$ and $\mathcal{F}_t^W := \sigma\{W(s) : s \leq t\}$ where both are augmented by all the P -null sets in \mathcal{F} .

We now consider the following “single agent” nonlinear stochastic optimal control problem (SOCP) on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, P)$:

$$dz(t, \omega) = f[t, \omega, z, u]dt + \sigma[t, \omega, z]dW(t) + \varsigma[t, \omega, z]dB(t), \quad 0 \leq t \leq T, \quad (6.15)$$

$$\inf_{u \in \mathcal{U}} J(u) := \inf_{u \in \mathcal{U}} E \left[\int_0^T L[t, \omega, z(t), u(t)]dt \right], \quad (6.16)$$

where the coefficients f, σ, ς and L are random depending on $\omega \in \Omega$ explicitly. In (6.15)-(6.16): (i) $z : [0, T] \times \Omega \rightarrow \mathbb{R}^n$ is the state of the agent with $\mathcal{F}_0^{W,B}$ -adapted random initial state $z(0)$ such that $E|z(0)|^2 < \infty$; (ii) $u : [0, T] \times \Omega \rightarrow U$ is the control input where U is a compact metric space; (iii) the functions $f : [0, T] \times \Omega \times \mathbb{R}^n \times U \rightarrow \mathbb{R}^n$, $\sigma, \varsigma : [0, T] \times \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$ are \mathcal{F}_t^W -adapted stochastic processes; (iv) the admissible control set \mathcal{U} is taken as $\mathcal{U} := \{u(\cdot) \in U : u(t) \text{ is adapted to } \sigma\text{-field } \mathcal{F}_t^{W,B} \text{ and } E \int_0^T |u(t)|^2 dt < \infty\}$. We introduce the following assumptions (see [145]).

(H6.1) $f[t, x, u]$ and $L[t, x, u]$ are a.s. continuous in (x, u) for each t , a.s. continuous in t for each (x, u) , $f[t, 0, 0] \in L^2_{\mathcal{F}_t}([0, T]; \mathbb{R}^n)$ and $L[t, 0, 0] \in L^2_{\mathcal{F}_t}([0, T]; \mathbb{R}_+)$. In addition, they and all their first derivatives (w.r.t. x) are a.s. continuous and bounded.

(H6.2) $\sigma[t, x]$ and $\varsigma[t, x]$ are a.s. continuous in x for each t , a.s. continuous in t for each x and $\sigma[t, 0], \varsigma[t, 0] \in L^2_{\mathcal{F}_t}([0, T]; \mathbb{R}^{n \times m})$. In addition, they and all their first derivatives (w.r.t. x) are a.s. continuous and bounded.

(H6.3) (*Non-degeneracy Assumption*) There exist non-negative constants α_1 and α_2 such that

$$\sigma[t, \omega, x]\sigma^T[t, \omega, x] \geq \alpha_1 I, \quad \varsigma[t, \omega, x]\varsigma^T(t, \omega, x) \geq \alpha_2 I, \quad \text{a.s.}, \quad \forall(t, \omega, x),$$

where α_1 or α_2 (but not both) can be zero.

The value function for the SOCP (6.15)-(6.16) is defined by (see [145])

$$\phi(t, x(t)) = \inf_{u \in \mathcal{U}} E_{\mathcal{F}_t^W} \int_t^T L[s, \omega, z(s), u(s)] ds, \quad (6.17)$$

where $x(t)$ is the initial condition for the process $x(s) := z(s)$, $t \leq s$. We note that $\phi(t, x(t))$ is an \mathcal{F}_t^W -adapted process which is sample path continuous a.s. under the assumptions (H6.1)-(H6.2). We assume that there exists an optimal control law $u^\circ \in \mathcal{U}$ such that

$$\phi(t, x(t)) = E_{\mathcal{F}_t^W} \int_t^T L[s, \omega, x(s), u^\circ(s, \omega, x(s))] ds,$$

where $x(\cdot)$ is the closed-loop solution when the control law u° is applied. By the Principle of Optimality, it can be shown that the process

$$\zeta(t) := \phi(t, x(t)) + \int_0^t L[s, \omega, x(s), u^\circ(s, x(s))] ds, \quad (6.18)$$

is an $\{\mathcal{F}_t^W\}_{0 \leq t \leq T}$ -martingale (see [21]). Next, by the martingale representation theorem (see Theorem 5.7, Chapter 1, [180]) there exists an \mathcal{F}_t^W -adapted process $\psi(\cdot, x(\cdot))$ such that

$$\zeta(t) = \phi(0, x(0)) + \int_0^t \psi^T(s, x(s)) dW(s), \quad t \in [0, T]. \quad (6.19)$$

From (6.18)-(6.19) and the fact that $\phi(T, x(T)) = 0$, it follows that

$$\zeta(T) = \int_0^T L[s, \omega, x(s), u^\circ(s, x(s))] ds = \phi(0, x(0)) + \int_0^T \psi^T(s, x(s)) dW(s),$$

which gives

$$\phi(0, x(0)) = \int_0^T L[s, \omega, x(s), u^\circ(s, x(s))] ds - \int_0^T \psi^T(s, x(s)) dW(s). \quad (6.20)$$

Hence, combining (6.18)-(6.20) yields

$$\begin{aligned}\phi(t, x(t)) &= \int_t^T L[s, \omega, x(s), u^o(s, x(s))]ds - \int_t^T \psi^T(s, x(s))dW(s) \\ &=: \int_t^T \Gamma(s, x(s))ds - \int_t^T \psi^T(s, x(s))dW(s), \quad t \in [0, T],\end{aligned}\tag{6.21}$$

where $\phi(s, x(s))$, $\Gamma(s, x(s))$ and $\psi(s, x(s))$ are \mathcal{F}_s^W -adapted stochastic processes (see the assumed “semi-martingale representation” form (3.5) in [145]).

We recall an extended version of the Itô-Kunita formula [99] for the composition of stochastic processes (see Theorem 2.3 in [145]).

THEOREM 6.2. *Let $\phi(t, x)$ be a stochastic process a.s. continuous in (t, x) such that (i) for each t , $\phi(t, \cdot)$ is a $C^2(\mathbb{R}^n)$ map a.s., (ii) for each x , $\phi(\cdot, x)$ is a continuous semi-martingale represented by*

$$d\phi(t, x) = -\Gamma(t, x)dt + \sum_{k=1}^m \psi_k(t, x)dW_k(t), \quad (t, x) \in [0, T] \times \mathbb{R}^n,$$

where $\Gamma(t, x)$ and $\psi_k(t, x)$, $1 \leq k \leq m$, are \mathcal{F}_t^W -adapted stochastic processes which are continuous in (t, x) a.s., such that for each t , $\Gamma(t, \cdot)$ is a $C^1(\mathbb{R}^n)$ map a.s., and $\psi_k(t, \cdot)$, $1 \leq k \leq m$, are $C^2(\mathbb{R}^n)$ maps (a.s.).

Let $x(\cdot) = (x^1(\cdot), \dots, x^n(\cdot))$ be a continuous semi-martingale of the form

$$dx^i(t) = f_i(t)dt + \sum_{k=1}^m \sigma_{ik}(t)dW_k(t) + \sum_{k=1}^m \varsigma_{ik}(t)dB_k(t), \quad 1 \leq i \leq n,$$

where f_i , $\sigma_i = (\sigma_{i1}, \dots, \sigma_{im})$ and $\varsigma_i = (\varsigma_{i1}, \dots, \varsigma_{im})$, $1 \leq i \leq n$, are \mathcal{F}_t^W -adapted stochastic processes such that (i) f_i is an integrable process a.s., and (ii) σ_i and ς_i are square integrable processes (a.s.).

Then the composition map $\phi(\cdot, x(\cdot))$ is also a continuous semi-martingale which has the form

$$\begin{aligned}
 d\phi(t, x(t)) &= -\Gamma(t, x(t))dt + \sum_{k=1}^m \psi_k(t, x(t))dW_k(t) + \sum_{i=1}^n \partial_{x_i} \phi(t, x(t)) f_i(t)dt \\
 &+ \sum_{i=1}^n \sum_{k=1}^m \partial_{x_i} \phi(t, x(t)) \sigma_{ik}(t) dW_k(t) + \sum_{i=1}^n \sum_{k=1}^m \partial_{x_i} \phi(t, x(t)) \varsigma_{ik}(t) dB_k(t) \\
 &+ \sum_{i=1}^n \sum_{k=1}^m \partial_{x_i} \psi_k(t, x(t)) \sigma_{ik}(t) dt + \frac{1}{2} \sum_{i,j=1}^n \sum_{k=1}^m \partial_{x_i x_j}^2 \phi(t, x(t)) \sigma_{ik}(t) \sigma_{jk}(t) dt \\
 &+ \frac{1}{2} \sum_{i,j=1}^n \sum_{k=1}^m \partial_{x_i x_j}^2 \phi(t, x(t)) \varsigma_{ik}(t) \varsigma_{jk}(t) dt. \tag{6.22}
 \end{aligned}$$

□

Using the extended Itô-Kunita formula (6.22) and the Principle of Optimality, Peng [145] showed that since $\phi(t, x)$ can be expressed in the semi-martingale form (6.21), and if $\phi(t, x)$, $\psi(t, x)$, $D_x \phi(t, x)$, $D_{xx}^2 \phi(t, x)$ and $D_x \psi(t, x)$ are a.s. continuous in (t, x) , then the pair $(\phi(s, x), \psi(s, x))$ satisfies the following backward in time stochastic Hamilton-Jacobi-Bellman (SHJB) equation:

$$\begin{aligned}
 -d\phi(t, \omega, x) &= \left[H[t, \omega, x, D_x \phi(t, \omega, x)] + \langle \sigma[t, \omega, x], D_x \psi(t, \omega, x) \rangle \right. \\
 &\left. + \frac{1}{2} \text{tr}[a[t, \omega, x] D_{xx}^2 \phi(t, \omega, x)] \right] dt - \psi^T(t, \omega, x) dW(t, \omega), \quad \phi(T, x) = 0, \tag{6.23}
 \end{aligned}$$

where $(t, x) \in [0, T] \times \mathbb{R}^n$, $a[t, \omega, x] := \sigma[t, \omega, x] \sigma^T[t, \omega, x] + \varsigma[t, \omega, x] \varsigma^T[t, \omega, x]$, and the stochastic Hamiltonian $H : [0, T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ is given by

$$H[t, \omega, x, p] := \inf_{u \in \mathcal{U}} \{ \langle f[t, \omega, x, u], p \rangle + L[t, \omega, x, u] \}.$$

We note that the appearance of the term $\langle \sigma[t, \omega, x], D_x \psi(t, \omega, x) \rangle$ in equation (6.23) corresponds to the Brownian motion $W(\cdot)$ in the extended Itô-Kunita formula (6.22) for the composition of \mathcal{F}_t^W -adapted stochastic processes $\phi(t, \omega, x)$ and $z(t, \omega)$ given in (6.21) and (6.15), respectively.

The solution to the backward in time SHJB equation (6.23) is a unique forward in time \mathcal{F}_t^W -adapted pair $(\phi, \psi)(t, x) \equiv (\phi(t, \omega, x), \psi(t, \omega, x))$ (see [145, 180]). We omit the proof of the following theorem which closely resembles that of Theorem 4.1 in [145].

THEOREM 6.3. *Assume (H6.1)-(H6.3) hold. Then the SHJB equation (6.23) has a unique solution $(\phi(t, x), \psi(t, x))$ in $(L_{\mathcal{F}_t}^2([0, T]; \mathbb{R}), L_{\mathcal{F}_t}^2([0, T]; \mathbb{R}^m))$. \square*

The forward in time \mathcal{F}_t^W -adapted optimal control process of the SOCP (6.15)-(6.16) is given by (see [145])

$$\begin{aligned} u^o(t, \omega, x) &:= \arg \inf_{u \in U} H^u[t, \omega, x, D_x \phi(t, \omega, x), u] \\ &= \arg \inf_{u \in U} \{ \langle f[t, \omega, x, u], D_x \phi(t, \omega, x) \rangle + L[t, \omega, x, u] \}. \end{aligned} \quad (6.24)$$

By a verification theorem approach, Peng [145] showed that if the unique solution $(\phi, \psi)(t, x)$ of the SHJB equation (6.23) satisfies:

- (i) for each t , $(\phi, \psi)(t, \cdot)$ is a $C^2(\mathbb{R}^n)$ map from \mathbb{R}^n into $\mathbb{R} \times \mathbb{R}^m$,
- (ii) for each x , $(\phi, \psi)(t, x)$ and $(D_x \phi, D_{xx}^2 \phi, D_x \psi)(t, x)$ are continuous F_t^W -adapted stochastic processes, then $\phi(x, t)$ coincides with the value function (6.17) of the SOCP (6.15)-(6.16).

6.5. The Major and Minor Agents' Stochastic Mean Field System

In the formulation (6.1)-(6.4) all minor agents are reacting to the same major agent and hence the major agent has non-negligible influence on the mean field behaviour of the minor agents. In other words, the noise process of the major agent w_0 causes random fluctuation of the mean-field behaviour of the minor agents and makes it stochastic (see the discussion in Section 2 of [75] for the major and minor (MM) linear-quadratic-Gaussian (LQG) model).

In this section, we first construct two auxiliary stochastic optimal control problems (SOCP) with random coefficients for the major and a generic minor agent in

Sections 6.5.1 and 6.5.2, respectively. Then, we present the stochastic mean field system for the major and minor agents game formulation (6.1)-(6.4) via the mean field game consistency condition in Section 6.5.3.

6.5.1. Stochastic Optimal Control Problem of the Major Agent. By the decoupling result in Theorem 6.1 which indicates that a single minor agent's statistical properties can effectively approximate the empirical distribution produced by all minor agents, we may approximate the empirical distribution of minor agents $\delta_{(\cdot)}^N$ with a stochastic probability measure $\mu_{(\cdot)}$ which depends on the noise process of the major agent w_0 .

In this section, let $\mu_t(\omega)$, $0 \leq t \leq T$, be an exogenous nominal minor agent stochastic measure process such that $\mu_0(dx) := dF(x)$ where F is defined in (A6.2). Note that in Section 6.5.3 $\mu_t(\omega)$ will be characterized via the mean field game consistency condition as the random measure of minor agents' mean field behaviour.

We define the following SOCP (6.15)-(6.16) with $\mathcal{F}_t^{w_0}$ -adapted random coefficients from the major agent's model (6.1) and (6.3) in the infinite population limit:

$$dz_0(t) = f_0[t, z_0(t), u_0(t), \mu_t(\omega)]dt + \sigma_0[t, z_0(t), \mu_t(\omega)]dw_0(t, \omega), \quad z_0(0), \quad (6.25)$$

$$\inf_{u_0 \in \mathcal{U}_0} J_0(u_0) := \inf_{u_0 \in \mathcal{U}_0} E \left[\int_0^T L_0[t, z_0(t), u_0(t), \mu_t(\omega)]dt \right], \quad (6.26)$$

where we explicitly indicate the dependence of the random measure $\mu_{(\cdot)}$ on the sample point $\omega \in \Omega$.

Step I (*Major Agent's Stochastic Hamilton-Jacobi-Bellman (SHJB) Equation*):

The value function of the major agent's SOCP (6.25)-(6.26) is defined by

$$\phi_0(t, x(t)) = \inf_{u_0 \in \mathcal{U}_0} E_{\mathcal{F}_t^{w_0}} \int_t^T L_0[s, z_0(s), u_0(s), \mu_s(\omega)]ds, \quad (6.27)$$

where $x(t)$ is the initial condition for the process $x(s) := z_0(s)$, $t \leq s$ (see (6.17)). As in Section 6.4, $\phi_0(t, x(t))$ has the form (see (6.21))

$$\phi_0(t, x(t)) = \int_t^T \Gamma_0(s, x(s))ds - \int_t^T \psi_0^T(s, x(s))dw_0(s), \quad t \in [0, T],$$

where $\phi_0(s, x(s))$, $\Gamma_0(s, x(s))$ and $\psi_0(s, x(s))$ are $\mathcal{F}_s^{w_0}$ -adapted stochastic processes. If $\phi_0(t, x)$, $\psi_0(t, x)$, $D_x\phi_0(t, x)$, $D_{xx}^2\phi_0(t, x)$ and $D_x\psi_0(t, x)$ are a.s. continuous in (t, x) , then (see [145]) the pair $(\phi_0(s, x), \psi_0(s, x))$ satisfies the following stochastic Hamilton-Jacobi-Bellman (SHJB) equation:

$$\begin{aligned} -d\phi_0(t, \omega, x) &= \left[H_0[t, \omega, x, D_x\phi_0(t, \omega, x)] + \langle \sigma_0[t, x, \mu_t(\omega)], D_x\psi_0(t, \omega, x) \rangle \right. \\ &\quad \left. + \frac{1}{2} \text{tr}(a_0[t, \omega, x] D_{xx}^2\phi_0(t, \omega, x)) \right] dt - \psi_0^T(t, \omega, x) dw_0(t, \omega), \quad \phi_0(T, x) = 0, \end{aligned} \quad (6.28)$$

where $(t, x) \in [0, T] \times \mathbb{R}^n$, $a_0[t, \omega, x] := \sigma_0[t, x, \mu_t(\omega)]\sigma_0^T[t, x, \mu_t(\omega)]$, and the stochastic Hamiltonian $H_0 : [0, T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ is given by

$$H_0[t, \omega, x, p] := \inf_{u \in \mathcal{U}_0} \left\{ \langle f_0[t, x, u, \mu_t(\omega)], p \rangle + L_0[t, x, u, \mu_t(\omega)] \right\}.$$

The solution to the backward in time SHJB equation (6.28) is a forward in time $\mathcal{F}_t^{w_0}$ -adapted pair $(\phi_0(t, x), \psi_0(t, x)) \equiv (\phi_0(t, \omega, x), \psi_0(t, \omega, x))$ (see [145]).

We note that the appearance of the term $\langle \sigma_0[t, x, \mu_t(\omega)], D_x\psi_0(t, \omega, x) \rangle$ in equation (6.28) corresponds to the major agent's Brownian motion $w_0(\cdot)$ in the extended Itô-Kunita formula (6.22) for the composition of $\mathcal{F}_t^{w_0}$ -adapted processes $\phi_0(t, \omega, x)$ and $z_0(t, \omega)$ in (6.25).

The best response control process of the major agent's SOCP (6.25)-(6.26) is given by

$$\begin{aligned} u_0^o(t, \omega, x) &\equiv u_0^o(t, x | \{\mu_s(\omega)\}_{0 \leq s \leq T}) := \arg \inf_{u_0 \in U_0} H_0^{u_0}[t, \omega, x, u_0, D_x\phi_0(t, \omega, x)] \\ &\equiv \arg \inf_{u_0 \in U_0} \left\{ \langle f_0[t, x, u_0, \mu_t(\omega)], D_x\phi_0(t, \omega, x) \rangle + L_0[t, x, u_0, \mu_t(\omega)] \right\}, \end{aligned} \quad (6.29)$$

where the infimum exists a.s. here and in all analogous infimizations in the chapter due to the continuity of all functions appearing in $H_0^{u_0}$ and the compactness of U_0 . It should be noted that the stochastic best response (SBR) control u_0^o is a forward in time $\mathcal{F}_t^{w_0}$ -adapted process which depends on the Brownian motion w_0 via the stochastic measure $\mu_t(\omega)$, $0 \leq t \leq T$. The notation in (6.29) indicates that u_0^o at time t depends upon the stochastic measure $\mu_s(\omega)$ on the whole interval $0 \leq s \leq T$.

Step II (*Major Agent's Stochastic Coefficient McKean-Vlasov (SMV) and Stochastic Fokker-Planck-Kolmogorov (SFPK) Equations*): By substituting the best response control process u_0^o (6.29) into the major agent's dynamics (6.25) we get the following stochastic McKean-Vlasov (SMV) dynamics with random coefficients:

$$dz_0^o(t, \omega) = f_0[t, z_0^o, u_0^o(t, \omega, z_0^o), \mu_t(\omega)]dt + \sigma_0[t, z_0^o, \mu_t(\omega)]dw_0(t, \omega), \quad (6.30)$$

with $z_0^o(0) = z_0(0)$, where f_0 and σ_0 are random processes via the stochastic measure μ and u_0^o . The random measure of the major agent $\mu_t^0(\omega)$, $0 \leq t \leq T$, is denoted as the unit mass measure concentrated at $z_0^o(t, \omega)$ (i.e., $\mu_t^0(\omega) = \delta_{z_0^o(t, \omega)}$).

An equivalent method to characterize the SMV of the major agent is to express (6.30) in the form of stochastic Fokker-Planck-Kolmogorov (SFPK) equation with random coefficients:

$$\begin{aligned} dp_s^0(t, \omega, x) = & \left(- \langle D_x, f_0[t, x, u_0^o(t, \omega, x), \mu_t(\omega)]p_s^0(t, \omega, x) \rangle \right. \\ & + \frac{1}{2} \text{tr} \langle D_{xx}^2, a_0[t, \omega, x]p_s^0(t, \omega, x) \rangle \Big) dt \\ & - \langle D_x, \sigma_0[t, x, \mu_t(\omega)]p_s^0(t, \omega, x)dw_0(t, \omega) \rangle, \quad p_s^0(s, \omega, x) = \delta_{z_0^o(s, \omega)}(dx), \end{aligned} \quad (6.31)$$

where $0 \leq s \leq t \leq T$. $p_s^0(t, \omega, x)$ is the conditional probability density of $z_0^o(t, \omega)$ given $\mathcal{F}_t^{w_0}$ and has the initial condition $p_s^0(s, \omega, x) = \delta_{z_0^o(s, \omega)}(dx)$.

6.5.2. Stochastic Optimal Control Problem of the Generic Minor Agent.

As in Section 6.5.1 let μ_t , $0 \leq t \leq T$, be the exogenous nominal minor agent stochastic measure process approximating the empirical distribution produced by all minor agents in the infinite population limit such that $\mu_0(dx) = dF(x)$ where F is defined in (A6.2). We let $\mu_t^0(\omega)$, $0 \leq t \leq T$, be the unit mass measure concentrated at the major agent's state $z_0^o(t, \omega)$ obtained from the major agent's SMV equation (6.30).

We define the following SOCP (6.15)-(6.16) with $\mathcal{F}_t^{w_0}$ -adapted random coefficients from the i^{th} generic minor agent's model (6.2), (6.4) in the infinite population limit:

$$dz_i(t) = f[t, z_i(t), u_i(t), \mu_t^0(\omega), \mu_t(\omega)]dt + \sigma[t, z_i(t), \mu_t^0(\omega), \mu_t(\omega)]dw_i(t, \omega'), \quad (6.32)$$

$$\inf_{u_i \in \mathcal{U}} J_i(u_i) := \inf_{u_i \in \mathcal{U}} E \left[\int_0^T L[t, z_i(t), u_i(t), \mu_t^0(\omega), \mu_t(\omega)]dt \right], \quad z_i(0), \quad (6.33)$$

where we explicitly indicate the dependence of the random measures $\mu_{(\cdot)}^0$ and $\mu_{(\cdot)}$ on the sample point $\omega \in \Omega$.

Step I (*Generic Minor Agent's Stochastic Hamilton-Jacobi-Bellman (SHJB) Equation*):

The value function of the generic minor agent's SOCP (6.32)-(6.33) is defined by

$$\phi_i(t, x(t)) = \inf_{u_i \in \mathcal{U}_0} E_{\mathcal{F}_t^{w_0}} \int_t^T L[s, z_i(s), u_i(s), \mu_s^0(\omega), \mu_s(\omega)]ds, \quad (6.34)$$

where $x(t)$ is the initial condition for the process $x(s) := z_i(s)$, $t \leq s$ (see (6.17)). As in Section 6.4, $\phi_i(t, x(t))$ has the form (see (6.21))

$$\phi_i(t, x(t)) = \int_t^T \Gamma_i(s, x(s))ds - \int_t^T \psi_i^T(s, x(s))dw_0(s), \quad t \in [0, T],$$

where $\phi_i(s, x(s))$, $\Gamma_i(s, x(s))$ and $\psi_i(s, x(s))$ are $\mathcal{F}_s^{w_0}$ -adapted stochastic processes. If $\phi_i(t, x)$, $\psi_i(t, x)$, $D_x \phi_i(t, x)$ and $D_{xx}^2 \phi_i(t, x)$ are a.s. continuous in (x, t) , then the pair $(\phi_i(s, x), \psi_i(s, x))$ satisfies the following backward in time stochastic Hamilton-Jacobi-Bellman (SHJB) equation (see (6.23)):

$$\begin{aligned} -d\phi_i(t, \omega, x) &= \left[H[t, \omega, x, D_x \phi_i(t, \omega, x)] + \frac{1}{2} \text{tr}(a[t, \omega, x] D_{xx}^2 \phi_i(t, \omega, x)) \right] dt \\ &\quad - \psi_i^T(t, \omega, x) dw_0(t, \omega), \quad \phi_i(T, x) = 0, \end{aligned} \quad (6.35)$$

where $(t, x) \in [0, T] \times \mathbb{R}^n$, $a[t, \omega, x] := \sigma[t, x, \mu_t^0(\omega), \mu_t(\omega)]\sigma^T[t, x, \mu_t^0(\omega), \mu_t(\omega)]$, and the stochastic Hamiltonian $H : [0, T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ is given by

$$H[t, \omega, x, p] := \inf_{u \in \mathcal{U}} \left\{ \langle f[t, x, u, \mu_t^0(\omega), \mu_t(\omega)], p \rangle + L[t, x, u, \mu_t^0(\omega), \mu_t(\omega)] \right\}.$$

The solution to the backward in time SHJB equation (6.35) is a forward in time $\mathcal{F}_t^{w_0}$ -adapted pair $(\phi_i(t, x), \psi_i(t, x)) \equiv (\phi_i(t, \omega, x), \psi_i(t, \omega, x))$ (see [145]). We note that since the coefficients of the SOCP (6.32)-(6.33) are $\mathcal{F}_t^{w_0}$ -adapted random processes we have the major agent's Brownian motion w_0 in (6.35) which allows us to seek for a forward in time adapted solution to the backward in time SHJB equation (6.35).

It is important to note that in (6.35) unlike the major agent's SHJB equation (6.28) we do not have the term $\langle \sigma[t, x, \mu_t^0(\omega), \mu_t(\omega)] D_x \psi_i(t, \omega, x) \rangle$ since the coefficients in the minor agent's model (6.32)-(6.33) are $\mathcal{F}_t^{w_0}$ -adapted random processes depending upon the major agent's Brownian motion (w_0) which is independent of the minor agent's Brownian motion (w_i) (see the extended Itô-Kunita formula (6.22)).

As in Section 6.5.1, the best response control process of the minor agent's SOCP (6.32)-(6.33) is

$$\begin{aligned} u_i^o(t, \omega, x) &\equiv u_i^o(t, x | \{\mu_s^0(\omega), \mu_s(\omega)\}_{0 \leq s \leq T}) := \arg \inf_{u \in U} H^u[t, \omega, x, u, D_x \phi_i(t, \omega, x)] \\ &\equiv \arg \inf_{u \in U} \{ \langle f[t, x, u, \mu_t^0(\omega), \mu_t(\omega)], D_x \phi_i(t, \omega, x) \rangle + L[t, x, u, \mu_t^0(\omega), \mu_t(\omega)] \}, \end{aligned} \quad (6.36)$$

where the infimum exists a.s. here and in all analogous infimizations in the chapter due to the continuity of all functions appearing in H^u and the compactness of U . It should be noted that the stochastic best response (SBR) control of the generic minor agent u_i^o is a forward in time $\mathcal{F}_t^{w_0}$ -adapted random process which depends on the Brownian motion w_0 via the stochastic measures $\mu_t^0(\omega)$ and $\mu_t(\omega)$, $0 \leq t \leq T$. The notation in (6.36) indicates that u_i^o at time t depends upon the stochastic measures $\mu_s^0(\omega)$ and $\mu_s(\omega)$ on the whole interval $0 \leq s \leq T$.

Step II (*Minor Agent's Stochastic Coefficient McKean-Vlasov (SMV) and Stochastic Fokker-Planck-Kolmogorov (SFPK) Equations*): By substituting the best response control process u_i^o (6.36) into the minor agent's dynamics (6.32) we get the following stochastic McKean-Vlasov (SMV) dynamics with random coefficients:

$$\begin{aligned} dz_i^o(t, \omega, \omega') &= f[t, z_i^o, u_i^o(t, \omega, z_i), \mu_t^0(\omega), \mu_t(\omega)] dt \\ &\quad + \sigma[t, z_i^o, \mu_t^0(\omega), \mu_t(\omega)] dw_i(t, \omega'), \quad z_i^o(0) = z_i(0), \end{aligned} \quad (6.37)$$

where f and σ are random processes via the stochastic measures μ^0 and μ , and the best response control process u_i^o which all depend on the Brownian motion of the major agent (w_0).

Based on the decoupling effect (see Section 6.3 or the theory of propagation of chaos in [162]) the generic agent's statistical properties can effectively approximate the empirical distribution produced by all minor agents in a large population system. Hence, we obtain a new stochastic measure $\hat{\mu}_t(\omega)$ for the mean field behaviour of minor agents as the conditional law of the generic minor agent's process $z_i^o(t, \omega)$ given $\mathcal{F}_t^{w_0}$. We characterize $\hat{\mu}_t(\omega)$, $0 \leq t \leq T$, by $P(z_i^o(t, \omega) \leq \alpha | \mathcal{F}_t^{w_0}) = \int_{-\infty}^{\alpha} \hat{\mu}(t, \omega, dx)$ a.s. for all $\alpha \in \mathbb{R}^n$ and $0 \leq t \leq T$, with $\hat{\mu}_0(dx) = \mu_0(dx) = dF(x)$ where F is defined in (A6.2).

An equivalent method to characterize the SMV of the generic minor agent is to express (6.37) in the form of stochastic Fokker-Planck-Kolmogorov (SFPK) equation with random coefficients:

$$d\hat{p}(t, \omega, x) = \left(- \langle D_x, f[t, x, u_i^o(t, \omega, x), \mu_t^0(\omega), \mu_t(\omega)] \hat{p}(t, \omega, x) \rangle + \frac{1}{2} \text{tr} \langle D_{xx}^2, a[t, \omega, x] \hat{p}(t, \omega, x) \rangle \right) dt, \quad \hat{p}(0, x) = p_0(x), \quad (6.38)$$

in $[0, T] \times \mathbb{R}^n$ where $p(t, \omega, x)$ is the conditional probability density of $z_i^o(t, \omega)$ given $\mathcal{F}_t^{w_0}$. By the decoupling effect (see Section 6.3) it is possible to characterize the mean field behaviour of minor agents in terms of generic agent's density function $\hat{p}(t, \omega, x)$. The reason that the generic minor agent's FPK equation (6.38) does not include the Itô integral term with respect to w_i is due to the fact that $p(t, \omega, x)$ is the conditional probability density given $\mathcal{F}_t^{w_0}$, and the independence of the Brownian motions w_0 and w_i , $1 \leq i \leq N$.

The density function $\hat{p}(t, \omega, x)$ generates the random measure of the minor agent's mean field behaviour $\hat{\mu}_t(\omega)$ such that $\hat{\mu}(t, \omega, dx) = \hat{p}(t, \omega, x) dx$ (a.s.), $0 \leq t \leq T$.

We note that the major agent's SOCP (6.25)-(6.26) and minor agent's SOCP (6.32)-(6.33) may be written with respect to the random density $p(t, \omega, x)$ of the stochastic measure $\mu(t, \omega, dx)$ by $\mu(t, \omega, dx) = p(t, \omega, x) dx$ (a.s.), $0 \leq t \leq T$.

6.5.3. The Mean Field Game Consistency Condition. Based on the mean field game (MFG) or Nash certainty equivalence (NCE) consistency (see [85] and [103]), we close the “measure and control” mapping loop by setting $\hat{\mu}_t(\omega) = \mu_t(\omega)$ a.s., $0 \leq t \leq T$, or $\hat{p}(t, \omega, x) = p(t, \omega, x)$ a.s. for $(t, x) \in [0, T] \times \mathbb{R}^n$. The MFG consistency is demonstrated in: (i) the major agent’s stochastic mean field (SMF) system

$$\begin{aligned}
 \text{[MF-SHJB]} \quad & -d\phi_0(t, \omega, x) = \left[H_0[t, \omega, x, D_x\phi_0(t, \omega, x)] \right. \\
 & \left. + \langle \sigma_0[t, x, \mu_t(\omega)], D_x\psi_0(t, \omega, x) \rangle + \frac{1}{2} \text{tr}(a_0[t, \omega, x] D_{xx}^2\phi_0(t, \omega, x)) \right] dt \\
 & - \psi_0^T(t, \omega, x) dw_0(t, \omega), \quad \phi_0(T, x) = 0,
 \end{aligned} \tag{6.39}$$

$$\begin{aligned}
 \text{[MF-SBR]} \quad & u_0^o(t, \omega, x) \equiv u_0^o(t, x | \{\mu_s(\omega)\}_{0 \leq s \leq T}) \\
 & := \arg \inf_{u_0 \in U_0} \left\{ \langle f_0[t, x, u_0, \mu_t(\omega)], D_x\phi_0(t, \omega, x) \rangle + L_0[t, x, u_0, \mu_t(\omega)] \right\},
 \end{aligned} \tag{6.40}$$

$$\begin{aligned}
 \text{[MF-SMV]} \quad & dz_0^o(t, \omega) = f_0[t, z_0^o, u_0^o(t, \omega, z_0^o), \mu_t(\omega)] dt \\
 & + \sigma_0[t, z_0^o, \mu_t(\omega)] dw_0(t, \omega), \quad z_0^o(0) = z_0(0),
 \end{aligned} \tag{6.41}$$

together with (ii) the minor agents’ SMF system

$$\begin{aligned}
 \text{[MF-SHJB]} \quad & -d\phi(t, \omega, x) = \left[H[t, \omega, x, D_x\phi(t, \omega, x)] \right. \\
 & \left. + \frac{1}{2} \text{tr}(a[t, \omega, x] D_{xx}^2\phi(t, \omega, x)) \right] dt - \psi^T(t, \omega, x) dw_0(t, \omega), \quad \phi(T, x) = 0,
 \end{aligned} \tag{6.42}$$

$$\begin{aligned}
 \text{[MF-SBR]} \quad & u^o(t, \omega, x) \equiv u^o(t, x | \{\mu_s^0(\omega), \mu_s(\omega)\}_{0 \leq s \leq T}) \\
 & \equiv \arg \inf_{u \in U} \left\{ \langle f[t, x, u, \mu_t^0(\omega), \mu_t(\omega)], D_x\phi(t, \omega, x) \rangle + L[t, x, u, \mu_t^0(\omega), \mu_t(\omega)] \right\},
 \end{aligned} \tag{6.43}$$

$$\begin{aligned}
 \text{[MF-SMV]} \quad & dz^o(t, \omega, \omega') = f[t, z^o, u^o(t, \omega, z^o), \mu_t^0(\omega), \mu_t(\omega)] dt \\
 & + \sigma[t, z^o, \mu_t^0(\omega), \mu_t(\omega)] dw(t, \omega'),
 \end{aligned} \tag{6.44}$$

where $(t, x) \in [0, T] \times \mathbb{R}^n$, and $z^o(0)$ has the measure $\mu_0(dx) = dF(x)$ where F is defined in (A6.2). We note that in the minor agents’ SMF system (6.42)-(6.44) we dropped index i from the generic minor agent’s equations (6.32)-(6.37). The MM

SMF system is given by the MM agents' coupled SMF systems (6.39)-(6.41) and (6.42)-(6.44).

The solution of the MM SMF system consists of 8-tuple $\mathcal{F}_t^{w_0}$ -adapted random processes

$$(\phi_0(t, \omega, x), \psi_0(t, \omega, x), u_0^o(t, \omega, x), z_0^o(t, \omega), \phi(t, \omega, x), \psi(t, \omega, x), u^o(t, \omega, x), z^o(t, \omega)),$$

where the random measure $\mu_t^0(\omega)$ is the unit mass measure concentrated at the major agent's state $z_0^o(t, \omega)$, i.e., $\mu_t^0(\omega) = \delta_{z_0^o(t, \omega)}$, and $z^o(t, \omega)$ generates the conditional random law $\mu_t(\omega)$, i.e., $P(z^o(t, \omega) \leq \alpha | \mathcal{F}_t^{w_0}) = \int_{-\infty}^{\alpha} \mu_t(\omega, dx)$ for all $\alpha \in \mathbb{R}^n$ and $0 \leq t \leq T$. Note that the major and minor (MM) agents' SMF systems (6.39)-(6.41) and (6.42)-(6.44) are coupled together through the stochastic measures $\mu_t^0(\omega)$ and $\mu_t(\omega)$.

We note that the solution to the MM SMF system is a “stochastic mean field” in contrast to the deterministic mean field of the standard mean field game problems in [78, 85, 101–103]. If the noise process of the major agent vanishes then the MM SMF system reduces to a deterministic MF system (see (6)-(9) in [78]).

6.6. Existence and Uniqueness of Solutions to the MM SMF System

In this section we establish existence and uniqueness for the solution of the joint major and minor (MM) agents' SMF system (6.39)-(6.41) and (6.42)-(6.44). The analysis is based on providing sufficient conditions for a map that goes from the random measure of minor agents $\mu_{(\cdot)}(\omega)$ back to itself, through the equations (6.39)-(6.41) and (6.42)-(6.44), to be a contraction operator on the space of random probability measures (see the diagram below).

$$\begin{array}{ccccc} \mu_{(\cdot)}(\omega) & \xrightarrow{(6.39)} & (\phi_0(\cdot, \omega, x), \psi_0(\cdot, \omega, x)) & \xrightarrow{(6.40)} & u_0^o(\cdot, \omega, x) \\ \uparrow(6.44) & & & & \downarrow(6.41) \\ u^o(\cdot, \omega, x) & \xleftarrow{(6.43)} & (\phi(\cdot, \omega, x), \psi(\cdot, \omega, x)) & \xleftarrow{(6.42)} & \mu_{(\cdot)}^0(\omega) \end{array}$$

In this section we first introduce some preliminary material about the Wasserstein space of probability measures. Second, we analyze the SHJB and SMV equations of the major agent and minor agents in Sections 6.6.1 and 6.6.2, respectively. Third, the analysis of the joint major and minor agents' SMF system is carried out in Section 6.6.3 which consists of two parts: (i) a sensitivity analysis of the SHJB equations for obtaining the feedback regularity conditions (Section 6.6.3.1), and (ii) the main theorem which provides sufficient conditions for a contraction operator map that goes from the random measure of minor agents $\mu_{(\cdot)}(\omega)$ back to itself (Section 6.6.3.2).

On the Banach space $C([0, T]; \mathbb{R}^n)$ we define the metric $\rho_T(x, y) = \sup_{0 \leq t \leq T} |x(t) - y(t)|^2 \wedge 1$, where \wedge denotes minimum. It can be shown that $C_\rho := (C([0, T]; \mathbb{R}^n), \rho_T)$ forms a separable complete metric space (i.e., a Polish space). Let $\mathcal{M}(C_\rho)$ be the space of all Borel probability measures μ on $C([0, T]; \mathbb{R}^n)$ such that $\int |x|^2 d\mu(x) < \infty$. We also denote $\mathcal{M}(C_\rho \times C_\rho)$ as the space of probability measures on the product space $C([0, T]; \mathbb{R}^n) \times C([0, T]; \mathbb{R}^n)$. As in [85] the process x is defined to be a generic random process with the sample space $C([0, T]; \mathbb{R}^n)$, i.e., $x(t, \omega) = \omega(t)$ for $\omega \in C([0, T]; \mathbb{R}^n)$.

Based on the metric ρ_T , we introduce the Wasserstein metric on $\mathcal{M}(C_\rho)$:

$$D_T^\rho(\mu, \nu) = \inf_{\gamma \in \Pi(\mu, \nu)} \left[\int_{C_\rho \times C_\rho} \rho_T(x(\omega_1), x(\omega_2)) d\gamma(\omega_1, \omega_2) \right]^{1/2},$$

where $\Pi(\mu, \nu) \subset \mathcal{M}(C_\rho \times C_\rho)$ is the set of Borel probability measures γ such that $\gamma(A \times C([0, T]; \mathbb{R}^n)) = \mu(A)$ and $\gamma(C([0, T]; \mathbb{R}^n) \times A) = \nu(A)$ for any Borel set $A \in C([0, T]; \mathbb{R}^n)$. The metric space $\mathcal{M}_\rho := (\mathcal{M}(C_\rho), D_T^\rho)$ is a Polish space since $C_\rho \equiv (C([0, T]; \mathbb{R}^n), \rho_T)$ is a Polish space.

We also introduce the class \mathcal{M}_ρ^β of stochastic measures in the space \mathcal{M}_ρ with a.s. Hölder continuity of exponent β , $0 < \beta < 1$ (see Definition 3 in [85] for the non-stochastic case).

DEFINITION 6.1. *A stochastic probability measure $\mu_t(\omega)$, $0 \leq t \leq T$, in the space \mathcal{M}_ρ is in \mathcal{M}_ρ^β if μ is a.s. uniformly Hölder continuous with exponent $0 < \beta < 1$, i.e., there exists $\beta \in (0, 1)$ and constant c such that for any bounded and Lipschitz*

continuous function ϕ on \mathbb{R}^n ,

$$\left| \int_{\mathbb{R}^n} \phi(x) \mu_t(\omega, dx) - \int_{\mathbb{R}^n} \phi(x) \mu_s(\omega, dx) \right| \leq c(\omega) |t - s|^\beta, \quad a.s.,$$

for all $0 \leq s < t \leq T$, where c may depend upon the Lipschitz constant of ϕ and the sample point $\omega \in \Omega$.

As in [85], we may take μ_t , $0 \leq t \leq T$, to be a Dirac measure at any constant $x \in \mathbb{R}^n$ to show that the set \mathcal{M}_ρ^β is nonempty. We introduce the following assumption.

(A6.8) For any $p \in \mathbb{R}^n$ and $\mu, \mu^0 \in \mathcal{M}_\rho^\beta$, the sets

$$S_0(t, \omega, x, p) := \arg \inf_{u_0 \in U_0} H_0^{u_0}[t, \omega, x, u_0, p],$$

$$S(t, \omega, x, p) := \arg \inf_{u \in U} H^u[t, \omega, x, u, p],$$

where $H_0^{u_0}$ and H^u are respectively defined in (6.29) and (6.36), are singletons and the resulting u and u_0 as functions of $[t, \omega, x, p]$ are a.s. continuous in t , Lipschitz continuous in (x, p) , uniformly with respect to t and $\mu, \mu^0 \in \mathcal{M}_\rho^\beta$. In addition, $u_0[t, \omega, 0, 0]$ and $u[t, \omega, 0, 0]$ are in the space $L^2_{\mathcal{F}_t}([0, T]; \mathbb{R}^n)$.

The first part of (A6.8) may be satisfied under suitable convexity conditions with respect to u_0 and u (see [85]).

6.6.1. Analysis of the Major Agent's SMF System. Let $\mu_t(\omega)$, $0 \leq t \leq T$, be a fixed stochastic measure in the set \mathcal{M}_ρ^β with $0 < \beta < 1$ such that $\mu_0(dx) := dF(x)$ where F is defined in (A6.2). Then, the functionals of $\mu_{(\cdot)}(\omega)$ in (6.25)-(6.26) become random functions which we write as

$$\begin{aligned} f_0^*[t, \omega, z_0, u_0] &:= f_0[t, z_0, u_0, \mu_t(\omega)], & \sigma_0^*[t, \omega, z_0] &:= \sigma_0[t, z_0, \mu_t(\omega)], \\ L_0^*[t, \omega, z_0, u_0] &:= L_0[t, z_0, u_0, \mu_t(\omega)]. \end{aligned} \tag{6.45}$$

We have the following result which broadly follows Proposition 4 in [85].

PROPOSITION 6.1. Assume **(A6.3)** holds for U_0 . Let $\mu_t(\omega)$, $0 \leq t \leq T$, be a fixed stochastic measure in the set \mathcal{M}_ρ^β with $0 < \beta < 1$. For f_0^* , σ_0^* and L_0^* defined in (6.45) it is the case that:

(i) Under **(A6.4)** for f_0 and σ_0 , the functions $f_0^*[t, \omega, z_0, u_0]$ and $\sigma_0^*[t, \omega, z_0]$ and their first order derivatives (w.r.t z_0) are a.s. continuous and bounded on $[0, T] \times \mathbb{R}^n \times U_0$ and $[0, T] \times \mathbb{R}^n$. $f_0^*[t, \omega, z_0, u_0]$ and $\sigma_0^*[t, \omega, z_0]$ are a.s. Lipschitz continuous in z_0 . In addition, $f_0^*[t, \omega, 0, 0]$ is in the space $L_{\mathcal{F}_t}^2([0, T]; \mathbb{R}^n)$ and $\sigma_0^*[t, \omega, 0]$ is in the space $L_{\mathcal{F}_t}^2([0, T]; \mathbb{R}^{n \times m})$.

(ii) Under **(A6.5)** for f_0 , the function $f_0^*[t, \omega, z_0, u_0]$ is a.s. Lipschitz continuous in $u_0 \in U_0$, i.e., there exist a constant $c > 0$ such that

$$\sup_{t \in [0, T], z_0 \in \mathbb{R}^n} |f_0^*[t, \omega, z_0, u_0] - f_0^*[t, \omega, z_0, u'_0]| \leq c(\omega)|u_0 - u'_0|, \quad (a.s.).$$

(iii) Under **(A6.6)** for L_0 , the function $L_0^*[t, \omega, z_0, u_0]$ and its first order derivative (w.r.t z_0) is a.s. continuous and bounded on $[0, T] \times \mathbb{R}^n \times U_0$. $L_0^*[t, \omega, z_0, u_0]$ is a.s. Lipschitz continuous in z_0 . In addition, $L_0^*[t, \omega, 0, 0]$ is in the space $L_{\mathcal{F}_t}^2([0, T]; \mathbb{R}_+)$.

(iv) Under **(A6.8)** for $H_0^{u_0}$, the set of minimizers

$$\arg \inf_{u_0 \in U_0} \{ \langle f_0^*[t, \omega, z_0, u_0], p \rangle + L_0^*[t, \omega, z_0, u_0] \},$$

is a singleton for any $p \in \mathbb{R}^n$, and the resulting u_0 as a function of $[t, \omega, z_0, p]$ is a.s. continuous in t , a.s. Lipschitz continuous in (z_0, p) , uniformly with respect to t . In addition, $u_0[t, \omega, 0, 0]$ is in the space $L_{\mathcal{F}_t}^2([0, T]; \mathbb{R}^n)$.

Proof: (i) We only show the results for f_0^* , the analysis for σ_0^* is similar. For $\omega \in \Omega$, we take (t, z, u) and (s, z', u') both from $[0, T] \times \mathbb{R}^n \times U_0$. We have

$$\begin{aligned} & |f_0^*[t, \omega, z, u] - f_0^*[s, \omega, z', u']| \equiv |f_0[t, z, u, \mu_t(\omega)] - f_0[s, z', u', \mu_s(\omega)]| \\ & \leq |f_0[t, z, u, \mu_t(\omega)] - f_0[s, z', u', \mu_t(\omega)]| + |f_0[s, z', u', \mu_t(\omega)] - f_0[s, z', u', \mu_s(\omega)]| \\ & \leq |f_0[t, z, u, \mu_t(\omega)] - f_0[s, z, u, \mu_t(\omega)]| + |f_0[s, z, u, \mu_t(\omega)] - f_0[s, z', u', \mu_t(\omega)]| \\ & \quad + |f_0[s, z', u', \mu_t(\omega)] - f_0[s, z', u', \mu_s(\omega)]|. \end{aligned}$$

By **(A6.4)**, $f_0[t, \omega, z, u]$ is continuous with respect to (t, z, u) and therefore

$$|f_0[t, z, u, \mu_t(\omega)] - f_0[s, z, u, \mu_t(\omega)]| + |f_0[s, z, u, \mu_t(\omega)] - f_0[s, z', u', \mu_t(\omega)]| \rightarrow 0,$$

as $|t - s| + |z - z'| + |u - u'| \rightarrow 0$. Since $\mu_{(\cdot)}(\omega)$ is in the set \mathcal{M}_ρ^β , $0 < \beta < 1$, and by **(A6.4)** there exists a constant $k > 0$ independent of (s, z, u) such that

$$|f_0[s, z, u, y] - f_0[s, z, u, y']| \leq k|y - y'|,$$

we get $|f_0[s, z', u', \mu_t(\omega)] - f_0[s, z', u', \mu_s(\omega)]| \rightarrow 0$ as $|t - s| \rightarrow 0$. This concludes the a.s. continuity of $f_0^*[t, \omega, z_0, u_0]$ on $[0, T] \times \mathbb{R}^n \times U_0$.

Using the Leibniz rule we have

$$D_{z_0} f_0^*[t, \omega, z_0, u_0] = \int D_{z_0} f_0[t, z_0, u_0, x] \mu_t(\omega)(dx), \quad a.s.,$$

where the partial derivative exists due to the boundedness of the first order derivative (w.r.t z_0) of f_0 by **(A6.4)**. The a.s. continuity of $D_{z_0} f_0^*$ on $[0, T] \times \mathbb{R}^n \times U_0$ may be proved by a similar argument above for f_0^* . Other results of the Proposition follow directly from **(A6.4)**.

(ii) This is a direct result of **(A6.5)**.

(iii) The proofs are similar to the proofs for f_0^* in part (i).

(iv) This is a direct result of **(A6.8)** for S_0 using the measure $\mu_{(\cdot)}(\omega) \in \mathcal{M}_\rho^\beta$, $0 < \beta < 1$. □

Employing the results of Section 6.4, we analyze the SHJB equation (6.39) where the probability measure $\mu_{(\cdot)}(\omega)$ is in the set \mathcal{M}_ρ^β , $0 < \beta < 1$.

THEOREM 6.4. *Assume (A6.3)-(A6.7) for U_0 , f_0 , σ_0 and L_0 hold, and the probability measure $\mu_{(\cdot)}(\omega)$ is in the set \mathcal{M}_ρ^β , $0 < \beta < 1$. Then the SHJB equation for the major agent (6.39) has a unique solution $(\phi_0(t, x), \psi_0(t, x))$ in the set $(L^2_{\mathcal{F}_t}([0, T]; \mathbb{R}), L^2_{\mathcal{F}_t}([0, T]; \mathbb{R}^m))$.*

Proof: Proposition 6.1 indicates that the SOCP of the major agent (6.25)-(6.26) satisfies the Assumptions (H6.1)-(H6.3) of Section 6.4 with $\varsigma[t, x] = 0$. The result follows directly from Theorem 6.3. \square

Let $\mu_{(\cdot)}(\omega) \in \mathcal{M}_\rho^\beta$, $0 < \beta < 1$, be given. We assume that the unique solution $(\phi_0, \psi_0)(t, x)$ to the SHJB equation (6.39) satisfies the regularity properties: (i) for each t , $(\phi_0, \psi_0)(t, x)$ is a $C^2(\mathbb{R}^n)$ map from \mathbb{R}^n into $\mathbb{R} \times \mathbb{R}^m$, (ii) for each x , (ϕ_0, ψ_0) and $(D_x \phi_0, D_{xx}^2 \phi_0, D_x \psi_0)$ are continuous F_t^W -adapted stochastic processes. Then, $\phi_0(x, t)$ coincides with the value function (6.27) [145], and under (A6.8) for $H_0^{u_0}$ we get the best response control process (6.29):

$$u_0^o(t, \omega, x) \equiv u_0^o(t, x | \{\mu_s(\omega)\}_{0 \leq s \leq T}) := \arg \inf_{u_0 \in U_0} H_0^{u_0}[t, \omega, x, u_0, D_x \phi_0(t, \omega, x)], \quad (6.46)$$

where $(t, x) \in [0, T] \times \mathbb{R}^n$.

We introduce the following assumption (see (H6) in [85]).

(A6.9) For any $\mu_{(\cdot)}(\omega) \in \mathcal{M}_\rho^\beta$, $0 < \beta < 1$, the best response control $u_0^o(t, \omega, x)$ is a.s. continuous in (t, x) and a.s. Lipschitz continuous in x .

We denote $C_{\text{Lip}(x)}([0, T] \times \Omega \times \mathbb{R}^n; H)$ be the class of a.s. continuous functions from $[0, T] \times \Omega \times \mathbb{R}^n$ to H , which are a.s. Lipschitz continuous in x [85]. We introduce the following well-defined map:

$$\begin{aligned} \Upsilon_0^{\text{SHJB}} : \mathcal{M}_\rho^\beta &\longrightarrow C_{\text{Lip}(x)}([0, T] \times \Omega \times \mathbb{R}^n; U_0), & 0 < \beta < 1, \\ \Upsilon_0^{\text{SHJB}}(\mu_{(\cdot)}(\omega)) &= u_0^o(t, \omega, x) \equiv u_0^o(t, x | \{\mu_s(\omega)\}_{0 \leq s \leq T}). \end{aligned} \quad (6.47)$$

We now analyze the major agent's SMV equation (6.41) with $\mu_{(\cdot)}(\omega) \in \mathcal{M}_\rho^\beta$ where $0 < \beta < 1$, and $u_0^o(t, \omega, x) \in C_{\text{Lip}(x)}([0, T] \times \Omega \times \mathbb{R}^n; U_0)$ be given in (6.46).

THEOREM 6.5. *Assume **(A6.3)**-**(A6.7)** for U_0 , f_0 and σ_0 , and **(A6.9)** hold. Let $\mu_{(\cdot)}(\omega) \in \mathcal{M}_\rho^\beta$ where $0 < \beta < 1$, and $u_0^o(t, \omega, x)$ be given in (6.46). Then, there exists a unique solution z_0^o on $[0, T] \times \Omega$ to the major agent's SMV equation (6.41).*

Proof: Proposition 6.1 indicates that the major agent's SMV equation (6.41) satisfies the Assumption **(RC)** in [180], page 49. The result follows directly from Theorem 6.16, Chapter 1 of [180], page 49. \square

THEOREM 6.6. *Assume **(A6.3)**-**(A6.7)** for U_0 , f_0 and σ_0 , and **(A6.9)** hold. Let $\mu_{(\cdot)}(\omega) \in \mathcal{M}_\rho^\beta$ where $0 < \beta < 1$, and $u_0^o(t, \omega, x)$ be given in (6.46). Then, the probability measure $\mu_{(\cdot)}^0(\omega)$ obtained from the major agent's SMV equation (6.41) is in the class \mathcal{M}_ρ^γ where $0 < \gamma < 1/2$.*

Proof: We take $0 \leq s < t \leq T$. Since $\mu_t^0(\omega) = \delta_{z_0^o(t, \omega)}$, for any bounded and Lipschitz continuous function ϕ on \mathbb{R}^n with a Lipschitz constant $K > 0$, we have

$$\begin{aligned} E \left| \int_{\mathbb{R}^n} \phi(x) \mu_t^0(\omega, dx) - \int_{\mathbb{R}^n} \phi(x) \mu_s^0(\omega, dx) \right| &= E \left| \phi(z_0^o(t, \omega)) - \phi(z_0^o(s, \omega)) \right| \\ &\leq K E \left| z_0^o(t, \omega) - z_0^o(s, \omega) \right|. \end{aligned}$$

On the other hand, Theorem 6.5 indicates that there exists a unique solution to the SMV equation (6.41) such that

$$z_0^o(t, \omega) - z_0^o(s, \omega) = \int_s^t f_0[\tau, z_0^o, u_0^o, \mu_\tau(\omega)] d\tau + \int_s^t \sigma_0[\tau, z_0^o, \mu_\tau(\omega)] dw_0(\tau).$$

Boundedness of f_0 and σ_0 (see **(A6.4)**), the Cauchy-Schwarz inequality and the property of Itô integral yield

$$E \left| z_0^o(t, \omega) - z_0^o(s, \omega) \right|^2 \leq 2C_1^2 |t - s|^2 + 2C_2^2 |t - s|,$$

where C_1 and C_2 are upper bounds for f_0 and σ_0 , respectively. Hence,

$$\begin{aligned} E \left| \int_{\mathbb{R}^n} \phi(x) \mu_t^0(\omega, dx) - \int_{\mathbb{R}^n} \phi(x) \mu_s^0(\omega, dx) \right| &\leq \sqrt{2}K(C_1|t-s| + C_2|t-s|^{1/2}) \\ &\leq \sqrt{2}K(C_1\sqrt{T} + C_2)|t-s|^{1/2}. \end{aligned}$$

By Kolmogorov's Theorem (Theorem 18.19, Page 266, [98]), for each $0 < \gamma < 1/2$, $T > 0$, and almost every $\omega \in \Omega$, there exists a constant $c(\omega, \gamma, K, T)$ such that

$$\left| \int_{\mathbb{R}^n} \phi(x) \mu_t^0(\omega, dx) - \int_{\mathbb{R}^n} \phi(x) \mu_s^0(\omega, dx) \right| \leq c(\omega, \gamma, K, T)|t-s|^\gamma,$$

for all $0 \leq s < t \leq T$. Hence, $\mu_{(\cdot)}^0(\omega)$ is in the class \mathcal{M}_ρ^γ where $0 < \gamma < 1/2$. \square

By Theorems 6.5 and 6.6 we may now introduce the following well-defined map:

$$\begin{aligned} \Upsilon_0^{\text{SMV}} : M_\rho^\beta \times C_{Lip(x)}([0, T] \times \Omega \times \mathbb{R}^n; U_0) &\longrightarrow M_\rho^\gamma, \quad 0 < \beta < 1, \quad 0 < \gamma < 1/2, \\ \Upsilon_0^{\text{SMV}}(\mu_{(\cdot)}(\omega), u_0^0(t, \omega, x)) &= \mu_{(\cdot)}^0(\omega). \end{aligned} \quad (6.48)$$

6.6.2. Analysis of the Minor Agents' SMF System. Let $\mu_{(\cdot)}(\omega) \in \mathcal{M}_\rho^\beta$, $0 < \beta < 1$, be the fixed stochastic measure assumed in Section 6.6.1. In this section we assume that $\mu_{(\cdot)}^0(\omega) \in \mathcal{M}_\rho^\gamma$, $0 < \gamma < 1/2$, is the obtained stochastic measure of the major agent from the composite map:

$$\begin{aligned} \Upsilon_0 : M_\rho^\beta &\longrightarrow M_\rho^\gamma, \quad 0 < \beta < 1, \quad 0 < \gamma < 1/2, \\ \Upsilon_0(\mu_{(\cdot)}(\omega)) &:= \Upsilon_0^{\text{SMV}}\left(\mu_{(\cdot)}(\omega), \Upsilon_0^{\text{SHJB}}(\mu_{(\cdot)}(\omega))\right) = \mu_{(\cdot)}^0(\omega), \end{aligned} \quad (6.49)$$

where Υ_0^{SHJB} and Υ_0^{SMV} are given in (6.47) and (6.48), respectively.

We may write the functionals of $\mu_{(\cdot)}^0(\omega)$ and $\mu_{(\cdot)}(\omega)$ in (6.32)-(6.33) as random functions:

$$\begin{aligned} f^*[t, \omega, z_i, u_i] &:= f[t, z_i, u_i, \mu_t^0(\omega), \mu_t(\omega)], \quad \sigma^*[t, \omega, z_i] := \sigma[t, z_i, \mu_t^0(\omega), \mu_t(\omega)], \\ L^*[t, \omega, z_i, u_i] &:= L[t, z_i, u_i, \mu_t^0(\omega), \mu_t(\omega)]. \end{aligned} \quad (6.50)$$

We have the following proposition where its proof closely resembles that of Proposition 6.1 (see Proposition 4 in [85]).

PROPOSITION 6.2. *Assume (A6.3) holds for U . Let $\mu_t(\omega)$, $0 \leq t \leq T$, be a fixed stochastic measure in the set \mathcal{M}_ρ^β with $0 < \beta < 1$, and $\mu_{(\cdot)}^0(\omega) = \Upsilon_0(\mu_{(\cdot)}(\omega)) \in \mathcal{M}_\rho^\gamma$, $0 < \gamma < 1/2$, be the obtained probability measure of the major agent in Section 6.6.1. For f^* , σ^* and L^* defined in (6.50) we have:*

- (i) *Under (A6.4) for f and σ , the functions $f^*[t, \omega, z_i, u_i]$ and $\sigma^*[t, \omega, z_i]$ and their first order derivatives (w.r.t z_i) are a.s. continuous and bounded on $[0, T] \times \mathbb{R}^n \times U$ and $[0, T] \times \mathbb{R}^n$. $f^*[t, \omega, z_i, u_i]$ and $\sigma^*[t, \omega, z_i]$ are a.s. Lipschitz continuous in z_i . In addition, $f^*[t, \omega, 0, 0]$ is in the space $L_{\mathcal{F}_t}^2([0, T]; \mathbb{R}^n)$ and $\sigma^*[t, \omega, 0]$ is in the space $L_{\mathcal{F}_t}^2([0, T]; \mathbb{R}^{n \times m})$.*
- (ii) *Under (A6.5) for f , the function $f^*[t, \omega, z_i, u_i]$ is a.s. Lipschitz continuous in $u_i \in U$, i.e., there exist a constant $c > 0$ such that*

$$\sup_{t \in [0, T], z_i \in \mathbb{R}^n} |f^*[t, \omega, z_i, u_i] - f^*[t, \omega, z_i, u'_i]| \leq c(\omega) |u_i - u'_i|, \quad (a.s.).$$

- (iii) *Under (A6.6) for L , the function $L^*[t, \omega, z_i, u_i]$ and its first order derivative (w.r.t z_i) is a.s. continuous and bounded on $[0, T] \times \mathbb{R}^n \times U$. It is a.s. Lipschitz continuous in z_i . In addition, $L^*[t, \omega, 0, 0] \in L_{\mathcal{F}_t}^2([0, T]; \mathbb{R}_+)$.*
- (iv) *Under (A6.8) for H^u , the set of minimizers*

$$\arg \inf_{u_i \in U} \{ \langle f^*[t, \omega, z_i, u_i], p \rangle + L^*[t, \omega, z_i, u_i] \},$$

is a singleton for any $p \in \mathbb{R}^n$, and the resulting u_i as a function of $[t, \omega, z_i, p]$ is a.s. continuous in t , a.s. Lipschitz continuous in (z_i, p) , uniformly with respect to t . In addition, $u_i[t, \omega, 0, 0]$ is in the space $L_{\mathcal{F}_t}^2([0, T]; \mathbb{R}^n)$.

□

Following arguments exactly parallel to those used in Section 6.6.1, we analyze the SHJB equation (6.42) where the probability measures $\mu_{(\cdot)}(\omega) \in \mathcal{M}_\rho^\beta$, $0 < \beta < 1$ and $\mu_{(\cdot)}^0(\omega) \in \mathcal{M}_\rho^\gamma$, $0 < \gamma < 1/2$.

THEOREM 6.7. *Assume (A6.3)-(A6.7) for U , f , σ and L hold, and $\mu_{(\cdot)}(\omega) \in \mathcal{M}_\rho^\beta$, $0 < \beta < 1$ and $\mu_{(\cdot)}^0(\omega)$ is in the set \mathcal{M}_ρ^γ , $0 < \gamma < 1/2$. Then the SHJB*

equation for the generic minor agent (6.35) has a unique solution $(\phi_i(t, x), \psi_i(t, x))$ in $(L^2_{\mathcal{F}_t}([0, T]; \mathbb{R}), L^2_{\mathcal{F}_t}([0, T]; \mathbb{R}^m))$.

Proof: Proposition 6.2 indicates that the SOCP of the generic minor agent (6.32)-(6.33) satisfies the Assumptions **(H6.1)**-(**H6.3**) of Section 6.4 with $\sigma[t, x] = 0$. The result follows directly from Theorem 6.3. \square

For the probability measure $\mu_{(\cdot)}(\omega) \in \mathcal{M}_\rho^\beta$, $0 < \beta < 1$, and $\mu_{(\cdot)}^0(\omega) \in \mathcal{M}_\rho^\gamma$, $0 < \gamma < 1/2$, we assume that the unique solution $(\phi_i, \psi_i)(t, x)$ to the SHJB equation (6.35) satisfies the regularity properties: (i) for each t , $(\phi_i, \psi_i)(t, x)$ is a $C^2(\mathbb{R}^n)$ map from \mathbb{R}^n into $\mathbb{R} \times \mathbb{R}^m$, (ii) for each x , (ϕ_i, ψ_i) and $(D_x \phi_i, D_{xx}^2 \phi_i, D_x \psi_i)$ are continuous F_t^W -adapted stochastic processes. Then, $\phi_i(x, t)$ coincides with the value function (6.34) [145], and under **(A6.8)** for H^u we get the best response control process (6.36):

$$\begin{aligned} u_i^o(t, \omega, x) &\equiv u_i^o(t, x | \{\mu_s^0(\omega), \mu_s(\omega)\}_{0 \leq s \leq T}) \\ &:= \arg \inf_{u_i \in U} H^u[t, \omega, x, u_i, D_x \phi_i(t, \omega, x)], \end{aligned} \quad (6.51)$$

where $(t, x) \in [0, T] \times \mathbb{R}^n$.

We introduce the following assumption (see **(A6.9)** or (H6) in [85]).

(A6.10) For any $\mu_{(\cdot)}(\omega) \in \mathcal{M}_\rho^\beta$, $0 < \beta < 1$, and $\mu_{(\cdot)}^0(\omega) \in \mathcal{M}_\rho^\gamma$, $0 < \gamma < 1/2$, the best response control process $u_i^o(t, \omega, x)$ is a.s. continuous in (t, x) and a.s. Lipschitz continuous in x .

We introduce the following well-defined map for the generic minor agent i :

$$\begin{aligned} \Upsilon_i^{\text{SHJB}} : M_\rho^\beta \times M_\rho^\gamma &\longrightarrow C_{\text{Lip}(x)}([0, T] \times \Omega \times \mathbb{R}^n; U), \quad 0 < \beta < 1, \quad 0 < \gamma < 1/2, \\ \Upsilon_i^{\text{SHJB}}(\mu_{(\cdot)}(\omega), \mu_{(\cdot)}^0(\omega)) &= u_i^o(t, \omega, x) \equiv u_i^o(t, x | \{\mu_s^0(\omega), \mu_s(\omega)\}_{0 \leq s \leq T}). \end{aligned} \quad (6.52)$$

For given probability measure $\mu_{(\cdot)}^0(\omega) \in \mathcal{M}_\rho^\gamma$, $0 < \gamma < 1/2$, we analyze the generic minor agent's SMV equation (6.37):

$$\begin{aligned} dz_i^o(t, \omega, \omega') &= f[t, z_i^o, u_i^o(t, \omega, z_i^o), \mu_t^0(\omega), \mu_t(\omega)]dt \\ &\quad + \sigma[t, z_i^o, \mu_t^0(\omega), \mu_t(\omega)]dw_i(t, \omega'), \quad z_i^o(0) = z_i(0), \end{aligned} \quad (6.53)$$

where $u_i^o(t, \omega, x) \in C_{\text{Lip}(x)}([0, T] \times \Omega \times \mathbb{R}^n; U)$ is given in (6.51). We call the pair $(z_i^o(\cdot, \omega, \omega'), \mu_{(\cdot)}(\omega))$ a consistent solution of the generic minor agent's SMV equation (6.53) if $(z_i^o(\cdot, \omega, \omega'), \mu_{(\cdot)}(\omega))$ solves (6.53) and $\mu_{(\cdot)}(\omega)$ be the the law of the process $z_i^o(\cdot, \omega, \omega')$, i.e., $\mu_{(\cdot)} = \mathcal{L}(z_i^o(\cdot, \omega, \omega'))$. We define Λ as the map which associates to $\mu_{(\cdot)}(\omega) \in \mathcal{M}_\rho^\beta$, $0 < \beta < 1/2$, the law of the process $z_i^o(\cdot, \omega, \omega')$ in (6.53):

$$\begin{aligned} z_i^o(t, \omega, \omega') &= z_i^o(0) + \int_0^t \left(\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f[s, z_i^o, u_i^o, y, z] d\mu_s^0(\omega)(y) d\mu_s(\omega)(z) \right) ds \\ &\quad + \int_0^t \left(\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \sigma[s, z_i^o, y, z] d\mu_s^0(\omega)(y) d\mu_s(\omega)(z) \right) dw_i(s, \omega'), \end{aligned} \quad (6.54)$$

where we observe that the law Λ depends on the sample point $\omega \in \Omega$.

We now show that there exists a unique $\mu_{(\cdot)}(\omega) \in \mathcal{M}_\rho^\beta$, $0 < \beta < 1$, such that $\mu(\omega) = \Lambda(\mu(\omega))$. The proof of the following theorem is based upon a fixed point argument with random parameters (see Theorem 6 in [85] and Theorem 1.1 in [162]).

THEOREM 6.8. *Assume (A6.3)-(A6.7) for U , f and σ , and (A6.10) hold. Let $\mu_{(\cdot)}^0(\omega)$ be in the set \mathcal{M}_ρ^γ where $0 < \gamma < 1/2$, and $u_i^o(t, \omega, x)$ be given in (6.51). Then, there exists a unique consistent solution pair $(z_i^o(\cdot, \omega, \omega'), \mu_{(\cdot)}(\omega))$ to the generic minor agent's SMV equation (6.53) where $\mu_{(\cdot)}(\omega) = \mathcal{L}(z_i^o(\cdot, \omega, \omega'))$.*

Proof: Let $\omega \in \Omega$ be fixed. For given probability measure $\mu_{(\cdot)}(\omega) \in \mathcal{M}_\rho^\beta$, $0 < \beta < 1$, one can show that the law of the process $z_i^o(\cdot, \omega, \omega')$ given in (6.54), $\Lambda(z_i^o(\cdot, \omega, \omega'))$, belongs to \mathcal{M}_ρ^β , $0 < \beta < 1$ (see Theorem 6.9).

We take $\mu_{(\cdot)}(\omega)$, $\nu_{(\cdot)}(\omega) \in \mathcal{M}_\rho^\beta$, $0 < \beta < 1$. Let $z_i^o(\cdot, \omega, \omega')$ be defined by (6.54), and similarly $x_i^o(\cdot, \omega, \omega')$ be defined by (6.54) after replacing $\mu_{(\cdot)}(\omega)$ by $\nu_{(\cdot)}(\omega)$. We

have

$$\begin{aligned}
 & E_{\mathcal{F}_t^{w_0}} \sup_{0 \leq s \leq t} |z_i^o(s, \omega) - x_i^o(s, \omega)|^2 \\
 & \leq 2t \int_0^t \left| \int_{\mathbb{R}^n \times \mathbb{R}^n} f[s, z_i^o, u_i^o, y, z] d\mu_s^0(\omega)(y) d\mu_s(\omega)(z) \right. \\
 & \quad \left. - \int_{\mathbb{R}^n \times \mathbb{R}^n} f[s, x_i^o, u_i^o, y, z] d\mu_s^0(\omega)(y) d\nu_s(\omega)(z) \right|^2 ds \\
 & \quad + 2 \int_0^t \left| \int_{\mathbb{R}^n \times \mathbb{R}^n} \sigma[s, z_i^o, y, z] d\mu_s^0(\omega)(y) d\mu_s(\omega)(z) \right. \\
 & \quad \left. - \int_{\mathbb{R}^n \times \mathbb{R}^n} \sigma[s, x_i^o, y, z] d\mu_s^0(\omega)(y) d\nu_s(\omega)(z) \right|^2 ds. \quad (6.55)
 \end{aligned}$$

But,

$$\begin{aligned}
 & \left| \int f[s, z_i^o, u_i^o, y, z] d\mu_s^0(\omega)(y) d\mu_s(\omega)(z) - \int f[s, x_i^o, u_i^o, y, z] d\mu_s^0(\omega)(y) d\nu_s(\omega)(z) \right|^2 \\
 & \leq 2C \left(|z_i^o(s) - x_i^o(s)|^2 + \int_{C_\rho \times C_\rho} |z_s(\omega_1) - z_s(\omega_2)|^2 d\gamma(\omega_1, \omega_2) \right),
 \end{aligned}$$

where C is obtained from the boundedness and Lipschitz continuity of both f and u^o , and $\gamma \in \mathcal{M}(C_\rho \times C_\rho)$ is any coupling of μ and ν where $\gamma(A \times C([0, T]; \mathbb{R}^n)) = \mu(A)$ and $\gamma(C([0, T]; \mathbb{R}^n) \times A) = \nu(A)$ for any Borel set $A \in C([0, T]; \mathbb{R}^n)$. Taking the infimum over all such γ couplings and then using the definition of metrics $\rho_{(\cdot)}$ and $D_{(\cdot)}^\rho$ yields

$$\begin{aligned}
 & \left| \int f[s, z_i^o, u_i^o, y, z] d\mu_s^0(\omega)(y) d\mu_s(\omega)(z) - \int f[s, x_i^o, u_i^o, y, z] d\mu_s^0(\omega)(y) d\nu_s(\omega)(z) \right|^2 \\
 & \leq 2C \left(\rho_s(z_i^o(s), x_i^o(s)) + (D_s^\rho(\mu, \nu))^2 \right). \quad (6.56)
 \end{aligned}$$

Similarly we have

$$\begin{aligned}
 & \left| \int \sigma[s, z_i^o, y, z] d\mu_s^0(\omega)(y) d\mu_s(\omega)(z) - \int \sigma[s, x_i^o, y, z] d\mu_s^0(\omega)(y) d\nu_s(\omega)(z) \right|^2 \\
 & \leq 2C_1 \left(\rho_s(z_i^o(s), x_i^o(s)) + (D_s^\rho(\mu, \nu))^2 \right), \quad (6.57)
 \end{aligned}$$

where C_1 is obtained from the boundedness and Lipschitz continuity of both σ .

It follows from (6.55)-(6.57) that

$$\begin{aligned} E_{\mathcal{F}_t^{w_0}} \rho_t(z_i^o(\omega), x_i^o(\omega)) &\equiv E_{\mathcal{F}_t^{w_0}} \sup_{0 \leq s \leq t} |z_i^o(s, \omega) - x_i^o(s, \omega)|^2 \wedge 1 \\ &\leq 2(Ct + C_1) \int_0^t \left(\rho_s(z_i^o(\omega), x_i^o(\omega)) + (D_s^\rho(\mu(\omega), \nu(\omega)))^2 \right) ds, \end{aligned} \quad (6.58)$$

which by Gronwall's lemma yields

$$E_{\mathcal{F}_t^{w_0}} \rho_t(z_i^o(\omega), x_i^o(\omega)) \leq 2(CT + C_1) \exp(2(CT + C_1)) \int_0^t \left(D_s^\rho(\mu(\omega), \nu(\omega)) \right)^2 ds.$$

This together with the definition of the Wasserstein metric $D_{(\cdot)}^\rho$ leads to the contraction inequality:

$$\left(D_t^\rho(\mu(\omega), \nu(\omega)) \right)^2 \leq 2(CT + C_1) \exp(2(CT + C_1)) \int_0^t \left(D_s^\rho(\mu(\omega), \nu(\omega)) \right)^2 ds.$$

By following a similar argument as in [162] (Theorem 1.1), one can show that $\{\Lambda^k(\mu(\omega)) : k \geq 1\}$ forms a Cauchy sequence a.s. in the complete metric space \mathcal{M}_ρ^β , $0 < \beta < 1$, and converges a.s. to a unique (a.s.) fixed point of Λ . \square

THEOREM 6.9. *Assume (A6.3)-(A6.7) for U , f and σ , and (A6.10) hold. Let $\mu_{(\cdot)}^0(\omega)$ be in the set \mathcal{M}_ρ^γ where $0 < \gamma < 1/2$. For given $u_i^o(t, \omega, x)$ in (6.51), let $(z_i^o(\cdot, \omega, \omega'), \mu_{(\cdot)}(\omega))$ be the consistent solution pair of the SMV equation (6.53). Then, the probability measure $\mu_{(\cdot)}(\omega)$ is in the class \mathcal{M}_ρ^β where $0 < \beta < 1$.*

Proof: We take $0 \leq s < t \leq T$. For any bounded and Lipschitz continuous function ϕ on \mathbb{R}^n with a Lipschitz constant $K > 0$, we have

$$\begin{aligned} E \left| \int_{\mathbb{R}^n} \phi(x) \mu_t(\omega, dx) - \int_{\mathbb{R}^n} \phi(x) \mu_s(\omega, dx) \right| &= E \left| E_\omega(\phi(z_i^o(t, \omega, \omega')) - \phi(z_i^o(s, \omega, \omega'))) \right| \\ &\leq K E \left| E_\omega(z_i^o(t, \omega, \omega') - z_i^o(s, \omega, \omega')) \right|. \end{aligned}$$

On the other hand, Theorem 6.8 indicates that there exists a unique solution to the SMV equation (6.53) such that

$$E_\omega(z_i^o(t, \omega, \omega') - z_i^o(s, \omega, \omega')) = \int_s^t f[\tau, z_i^o, u_0^i, \mu_\tau^0(\omega), \mu_\tau(\omega)] d\tau,$$

where we note that $E_\omega \int_0^t \sigma[\tau, z_i^o, \mu_\tau^0(\omega), \mu_\tau(\omega)] dw_i(\tau, \omega') = 0$ for $0 \leq t \leq T$. Boundedness of f (see **(A6.4)**) yields

$$E|E_\omega(z_i^o(t, \omega, \omega') - z_i^o(s, \omega, \omega'))| \leq C_1|t - s|,$$

where C_1 is the upper bound for f .

By Kolmogorov's Theorem (Theorem 18.19, Page 266, [98], Page 266), for each $0 < \gamma < 1$, $T > 0$, and almost every $\omega \in \Omega$, there exists a constant $c(\omega, \gamma, K, T)$ such that

$$\left| \int_{\mathbb{R}^n} \phi(x) \mu_t(\omega, dx) - \int_{\mathbb{R}^n} \phi(x) \mu_s(\omega, dx) \right| \leq c(\omega, \gamma, K, T) |t - s|^\gamma,$$

for all $0 \leq s < t \leq T$. Hence, $\mu_{(\cdot)}(\omega)$ is in the class \mathcal{M}_ρ^β where $0 < \beta < 1$. \square

By Theorems 6.8 and 6.9 we may now introduce the following well-defined map:

$$\begin{aligned} \Upsilon_i^{\text{SMV}} : M_\rho^\gamma \times C_{\text{Lip}(x)}([0, T] \times \Omega \times \mathbb{R}^n; U_0) &\longrightarrow M_\rho^\beta, \quad 0 < \beta < 1, \quad 0 < \gamma < 1/2, \\ \Upsilon_i^{\text{SMV}}(\mu_{(\cdot)}^0(\omega), u_i^o(t, \omega, x)) &= \mu_{(\cdot)}(\omega). \end{aligned} \quad (6.59)$$

6.6.3. Analysis of the Joint Major and Minor Agents' SMF System.

Based on the analysis of Sections 6.6.1 and 6.6.2 we obtain the following well-defined map:

$$\begin{aligned} \Upsilon : M_\rho^\beta &\longrightarrow M_\rho^\beta, \quad 0 < \beta < 1, \\ \Upsilon(\mu_{(\cdot)}(\omega)) &= \Upsilon_i^{\text{SMV}}\left(\Upsilon_0(\mu_{(\cdot)}(\omega)), \Upsilon_i^{\text{SHJB}}(\mu_{(\cdot)}(\omega)), \Upsilon_0(\mu_{(\cdot)}(\omega))\right), \end{aligned} \quad (6.60)$$

which is the composition of the maps Υ_0 , Υ_i^{SHJB} and Υ_i^{SMV} introduced in (6.49), (6.52) and (6.59), respectively. Subsequently, the problem of existence and uniqueness of solution to the MM SMV system (6.39)-(6.41) and (6.42)-(6.44) is translated into a fixed point problem with random parameters for the map Υ on the Polish space \mathcal{M}_ρ^β , $0 < \beta < 1$.

We introduce the following assumption without which one needs to work with the "expectation" of the Wasserstein metric $D_{(\cdot)}^o$ of stochastic measure.

(A6.11) We assume that the diffusion coefficient of the major agent σ_0 in (6.1) does not depend on its own state z_0^N and the states of the minor agents z_i^N , $1 \leq i \leq N$.

LEMMA 6.1. (i) Assume **(A6.3)**-**(A6.7)** for U_0 , f_0 and σ_0 , and **(A6.11)** hold. Let $\mu_{(\cdot)}(\omega)$ be in the set \mathcal{M}_ρ^β where $0 < \beta < 1$. Then, for given $u_0, u'_0 \in C_{Lip(x)}([0, T] \times \Omega \times \mathbb{R}^n; U_0)$ there exists a constant c_0 such that

$$\left(D_T^\rho(\mu^0(\omega), \nu^0(\omega)) \right)^2 \leq c_0 \sup_{(t,x) \in [0,T] \times \mathbb{R}^n} |u_0(t, \omega, x) - u'_0(t, \omega, x)|^2, \quad a.s., \quad (6.61)$$

where $\mu^0(\omega), \nu^0(\omega) \in \mathcal{M}_\rho^\gamma$, $0 < \gamma < 1/2$, are induced by the map Υ_0^{SMV} in (6.48) using the two control processes u_0 and u'_0 , respectively.

(ii) Assume **(A6.3)**-**(A6.7)** for U_0 , f_0 and σ_0 , and **(A6.11)** hold. Let u_0^o be in the space $C_{Lip(x)}([0, T] \times \Omega \times \mathbb{R}^n; U_0)$. Then, for given $\mu(\omega), \nu(\omega) \in \mathcal{M}_\rho^\beta$, $0 < \beta < 1$, there exists a constant c_1 such that

$$\left(D_T^\rho(\mu^0(\omega), \nu^0(\omega)) \right)^2 \leq c_1 \left(D_T^\rho(\mu(\omega), \nu(\omega)) \right)^2, \quad a.s., \quad (6.62)$$

where $\mu^0(\omega), \nu^0(\omega) \in \mathcal{M}_\rho^\gamma$, $0 < \gamma < 1/2$, are induced by the map Υ_0^{SMV} in (6.48) using the stochastic measures $\mu(\omega)$ and $\nu(\omega)$, respectively.

(iii) Assume **(A6.3)**-**(A6.7)** for U , f and σ hold. Let $\mu_{(\cdot)}^0(\omega)$ be in the set \mathcal{M}_ρ^γ where $0 < \gamma < 1/2$. Then, for given $u, u' \in C_{Lip(x)}([0, T] \times \Omega \times \mathbb{R}^n; U)$ there exists a constant c_2 such that

$$\left(D_T^\rho(\mu(\omega), \nu(\omega)) \right)^2 \leq c_2 \sup_{(t,x) \in [0,T] \times \mathbb{R}^n} |u(t, \omega, x) - u'(t, \omega, x)|^2, \quad a.s., \quad (6.63)$$

where $\mu(\omega), \nu(\omega) \in \mathcal{M}_\rho^\beta$, $0 < \beta < 1$, are induced by the map Υ_i^{SMV} in (6.59) using the two control processes u and u' , respectively.

(iv) Assume **(A6.3)**-**(A6.7)** for U , f and σ hold. Let u_i^o be in the space $C_{Lip(x)}([0, T] \times \Omega \times \mathbb{R}^n; U)$. Then, for given $\mu^0(\omega), \nu^0(\omega) \in \mathcal{M}_\rho^\gamma$, $0 < \gamma < 1/2$, there exists a constant c_3 such that

$$\left(D_T^\rho(\mu(\omega), \nu(\omega)) \right)^2 \leq c_3 \left(D_T^\rho(\mu^0(\omega), \nu^0(\omega)) \right)^2, \quad a.s., \quad (6.64)$$

where $\mu(\omega), \nu(\omega) \in \mathcal{M}_\rho^\beta$, $0 < \beta < 1$, are induced by the map Υ_i^{SMV} in (6.59) using the stochastic measures $\mu^0(\omega)$ and $\nu^0(\omega)$, respectively.

Proof: (i) (6.41) gives

$$\begin{aligned} z_0(s, \omega) &= z_0(0) + \int_0^s \left(\int_{\mathbb{R}^n} f_0[\tau, z_0, u_0, y] d\mu_\tau(\omega)(y) \right) d\tau + \int_0^s \sigma_0[\tau] dw_0(\tau, \omega), \\ z'_0(s, \omega) &= z_0(0) + \int_0^s \left(\int_{\mathbb{R}^n} f_0[\tau, z'_0, u'_0, y] d\mu_\tau(\omega)(y) \right) d\tau + \int_0^s \sigma_0[\tau] dw_0(\tau, \omega), \end{aligned}$$

corresponding to the control processes u_0 and u'_0 in $C_{\text{Lip}(x)}([0, T] \times \Omega \times \mathbb{R}^n; U_0)$. By the Lipschitz continuity of f_0 (see **(A6.4)** and **(A6.5)**) there are positive constants C_0 and C_1 such that

$$\begin{aligned} |z_0(s, \omega) - z'_0(s, \omega)|^2 &\leq 2C_0 s \int_0^s |z_0(\tau, \omega) - z'_0(\tau, \omega)|^2 d\tau \\ &\quad + 2C_1 s^2 \sup_{(t,x) \in [0,T] \times \mathbb{R}^n} |u_0(t, \omega, x) - u'_0(t, \omega, x)|^2. \end{aligned}$$

The Gronwall's lemma yields

$$\rho_t(z_0(\omega), z'_0(\omega)) \leq 2C_1 t^2 \exp(2C_0 t) \sup_{t,x} |u_0(t, \omega, x) - u'_0(t, \omega, x)|^2.$$

This together with the fact that $\mu_t^0(\omega) = \delta_{z_0(t, \omega)}$ and $\nu_t^0(\omega) = \delta_{z'_0(t, \omega)}$, and the definition of the Wasserstein metric $D_{(\cdot)}^\rho$ leads to (6.61) where $c_0 := 2C_1 T^2 \exp(2C_0 T)$.

(ii) We have

$$\begin{aligned} z_0(s, \omega) &= z_0(0) + \int_0^s \left(\int_{\mathbb{R}^n} f_0[\tau, z_0, u_0^o, y] d\mu_\tau(\omega)(y) \right) d\tau + \int_0^s \sigma_0[\tau] dw_0(\tau, \omega), \\ z'_0(s, \omega) &= z_0(0) + \int_0^s \left(\int_{\mathbb{R}^n} f_0[\tau, z'_0, u_0^o, y] d\nu_\tau(\omega)(y) \right) d\tau + \int_0^s \sigma_0[\tau] dw_0(\tau, \omega), \end{aligned}$$

corresponding to the stochastic measures $\mu(\omega), \nu(\omega) \in \mathcal{M}_\rho^\beta$, $0 < \beta < 1$. By the Lipschitz continuity of f_0 (see **(A6.4)** and **(A6.5)**) and u_0^o there are positive constants

C_0 and C_1 such that

$$\begin{aligned} |z_0(s, \omega) - z'_0(s, \omega)|^2 &\leq 2C_0s \int_0^s |z_0(\tau, \omega) - z'_0(\tau, \omega)|^2 d\tau \\ &\quad + 2C_1s^2 \left(D_T^\rho(\mu(\omega), \nu(\omega)) \right)^2. \end{aligned}$$

The Gronwall's lemma yields

$$\rho_t(z_0(\omega), z'_0(\omega)) \leq 2C_1t^2 \exp(2C_0t) \left(D_T^\rho(\mu(\omega), \nu(\omega)) \right)^2.$$

This together with the fact that $\mu_t^0(\omega) = \delta_{z_0(t, \omega)}$ and $\nu_t^0(\omega) = \delta_{z'_0(t, \omega)}$, and the definition of the Wasserstein metric $D_{(\cdot)}^\rho$ leads to (6.62) where $c_1 := 2C_1T^2 \exp(2C_0T)$.

(iii) (6.44) gives

$$\begin{aligned} z_i(s, \omega, \omega') &= z_i(0) + \int_0^t \left(\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f[s, z_i, u, y, z] d\mu_s^0(\omega)(y) d\mu_s(\omega)(z) \right) ds \\ &\quad + \int_0^t \left(\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \sigma[s, z_i, y, z] d\mu_s^0(\omega)(y) d\mu_s(\omega)(z) \right) dw_i(s, \omega'), \\ z'_i(s, \omega, \omega') &= z_i(0) + \int_0^t \left(\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f[s, z'_i, u', y, z] d\mu_s^0(\omega)(y) d\nu_s(\omega)(z) \right) ds \\ &\quad + \int_0^t \left(\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \sigma[s, z'_i, y, z] d\mu_s^0(\omega)(y) d\nu_s(\omega)(z) \right) dw_i(s, \omega'), \end{aligned}$$

corresponding to the control processes u and u' in $C_{\text{Lip}(x)}([0, T] \times \Omega \times \mathbb{R}^n; U)$. By the Lipschitz continuity of f and σ (see **(A6.4)** and **(A6.5)**) there are positive constants C_0, C_1 and C_2 such that

$$\begin{aligned} E_\omega |z_i(s, \omega, \omega') - z'_i(s, \omega, \omega')|^2 &\leq 2(3C_0s + 2C_1)E_\omega \int_0^s |z_0(\tau, \omega) - z'_0(\tau, \omega)|^2 d\tau \\ &\quad + 2(3C_0s + 2C_1)E_\omega \int_0^s \left(D_\tau^\rho(\mu(\omega), \nu(\omega)) \right)^2 d\tau \\ &\quad + 6C_2s^2 \sup_{t, x} E_\omega |u(t, \omega, x) - u'(t, \omega, x)|^2. \end{aligned}$$

The Gronwall's lemma yields

$$\begin{aligned} \rho_t(z_i(s, \omega), z'_i(s, \omega)) &\leq 2(3C_0t + 2C_1) \exp(2(3C_0t + 2C_1)) \int_0^t \left(D_\tau^\rho(\mu(\omega), \nu(\omega)) \right)^2 d\tau \\ &\quad + 6C_2t^2 \exp(2(3C_0t + 2C_1)) \sup_{t,x} |u(t, \omega, x) - u'(t, \omega, x)|^2. \end{aligned}$$

This together with the definition of the Wasserstein metric $D_{(\cdot)}^\rho$ leads to

$$\begin{aligned} \left(D_T^\rho(\mu(\omega), \nu(\omega)) \right)^2 &\leq K(T) \int_0^T \left(D_\tau^\rho(\mu(\omega), \nu(\omega)) \right)^2 d\tau \\ &\quad + K'(T) \sup_{t,x} |u(t, \omega, x) - u'(t, \omega, x)|^2, \end{aligned}$$

where $K(T) := 2(3C_0T + 2C_1) \exp(2(3C_0T + 2C_1))$ and $K'(T) := 6C_2T^2 \exp(2(3C_0T + 2C_1))$. Applying the Gronwall's lemma gives (6.63) with $c_2 := K'(T) \exp(K(T))$.

(iv) The proof of this part closely resembles that of Part (iii). \square

6.6.3.1. *The Sensitivity Analysis of the SHJB Equations.* In this section we study the sensitivity of the major and minor agents' SHJB equations (6.39) and (6.42) to the stochastic measures $\mu_{(\cdot)}(\omega)$ and $\mu_{(\cdot)}^0(\omega)$ in order to show the feedback regularity conditions. The analysis of this section is based on the framework of Section 6 of [97].

First we consider a family of stochastic optimal control problems (SOCP) (6.15)-(6.16) parameterized by $\alpha \in \mathbb{R}$. In this α -parameterized formulation called (SOCP) $_\alpha$: (i) the dynamics of the states $z^\alpha(t, \omega)$, denoted by (6.15) $_\alpha$, are of the form (6.15) with $f[t, \omega, z, u]$, $\sigma[t, \omega, z]$ and $\varsigma[t, \omega, z]$ replaced by $f^\alpha[t, \omega, z^\alpha, u^\alpha]$, $\sigma^\alpha[t, \omega, z^\alpha]$ and $\varsigma^\alpha[t, \omega, z^\alpha]$, respectively, and (ii) the cost functions $J^\alpha(u^\alpha)$, denoted by (6.16) $_\alpha$, are of the form (6.16) with $L[t, \omega, z, u]$ replaced by $L^\alpha[t, \omega, z^\alpha, u^\alpha]$.

The value functions $\phi^\alpha(\cdot, x(\cdot))$ correspond to the (SOCP) $_\alpha$ are defined similar to (6.17) with $L[t, \omega, z, u]$ replaced by $L^\alpha[t, \omega, z^\alpha, u^\alpha]$. Based on [145] we shall restrict to the case where $\phi^\alpha(\cdot, x(\cdot))$ are semi-martingales of the form (6.21) with $\Gamma(\cdot, x(\cdot))$ and $\psi(\cdot, x(\cdot))$ are replaced by $\Gamma^\alpha(\cdot, x(\cdot))$ and $\psi^\alpha(\cdot, x(\cdot))$, respectively.

If the α -parameterized family of processes $\phi^\alpha(t, x)$, $\Gamma^\alpha(t, x)$ and $\psi^\alpha(t, x)$ are a.s. continuous in (x, t) and are smooth enough with respect to x , then by using the

analysis in [145] one can show that the pairs $(\phi^\alpha(s, x), \psi^\alpha(s, x))$ satisfy the following backward in time α -parameterized stochastic Hamilton-Jacobi-Bellman (SHJB) $_\alpha$ equations:

$$\begin{aligned} -d\phi^\alpha(t, \omega, x) &= \left[H^\alpha[t, \omega, x, D_x\phi^\alpha(t, \omega, x)] + \langle \sigma^\alpha[t, \omega, x], D_x\psi^\alpha(t, \omega, x) \rangle \right. \\ &\left. + \frac{1}{2}\text{tr}(a^\alpha[t, \omega, x]D_{xx}^2\phi^\alpha(t, \omega, x)) \right] dt - (\psi^\alpha)^T(t, \omega, x)dW(t, \omega), \quad \phi^\alpha(T, x) = 0, \end{aligned} \quad (6.65)$$

where $a^\alpha[t, \omega, x] := \sigma^\alpha[t, \omega, x](\sigma^\alpha[t, \omega, x])^T + \varsigma^\alpha[t, \omega, x](\varsigma^\alpha[t, \omega, x])^T$, and the stochastic Hamiltonians $H^\alpha : [0, T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ are given by

$$H^\alpha[t, \omega, x, p] := \inf_{u^\alpha \in \mathcal{U}} \left\{ \langle f^\alpha[t, \omega, x, u], p \rangle + L^\alpha[t, \omega, x, u] \right\}.$$

Suppose the assumptions **(H6.1)**-**(H6.3)** hold for $(f^\alpha, L^\alpha, \sigma^\alpha, \varsigma^\alpha)$. Then the (SHJB) $_\alpha$ equations (6.65) have unique solutions (see Theorem 6.3 or Theorem 4.1 in [145]):

$$(\phi^\alpha(t, x), \psi^\alpha(t, x)) \in (L_{\mathcal{F}_t}^2([0, T]; \mathbb{R}), L_{\mathcal{F}_t}^2([0, T]; \mathbb{R}^m)), \quad \forall \alpha \in \mathbb{R}.$$

The forward in time \mathcal{F}_t^W -adapted optimal control processes of the (SOCP) $_\alpha$ (6.15) $_\alpha$ - (6.16) $_\alpha$ are given by (see [145])

$$\begin{aligned} u^{\alpha,o}(t, \omega, x) &:= \arg \inf_{u^\alpha \in U} H^{\alpha,u}[t, \omega, x, D_x\phi^\alpha(t, \omega, x), u^\alpha] \\ &= \arg \inf_{u^\alpha \in U} \left\{ \langle f^\alpha[t, \omega, x, u^\alpha], D_x\phi^\alpha(t, \omega, x) \rangle + L^\alpha[t, \omega, x, u^\alpha] \right\}. \end{aligned} \quad (6.66)$$

We set

$$\begin{aligned} g^\alpha[t, \omega, x, \phi^\alpha(t, \omega, x), \psi^\alpha(t, \omega, x)] &:= H^\alpha[t, \omega, x, D_x\phi^\alpha(t, \omega, x)] \\ &\quad + \langle \sigma^\alpha[t, \omega, x], D_x\psi^\alpha(t, \omega, x) \rangle, \\ A^\alpha(t, \omega, x)(\cdot) &:= \frac{1}{2}\text{tr}(a^\alpha[t, \omega, x]D_{xx}^2(\cdot)), \end{aligned}$$

where A^α in $[0, T] \times \Omega \times \mathbb{R}^n$ is an operator on $C^2(\mathbb{R}^n)$. We may now rewrite the backward in time α -parameterized (SHJB) $_\alpha$ equations (6.65) as

$$\begin{aligned} d\phi^\alpha(t, \omega, x) + A^\alpha(t, \omega, x)(\phi^\alpha(t, \omega, x))dt \\ = -g^\alpha[t, \omega, x, \phi^\alpha(t, \omega, x), \psi^\alpha(t, \omega, x)]dt + (\psi^\alpha)^T(t, \omega, x)dW(t, \omega), \end{aligned} \quad (6.67)$$

with $\phi^\alpha(T, x) = 0$.

At this point we introduce the mild form of (6.67) because this form is more suitable for the sensitivity analysis of this section. We note that it is sufficient to consider the mild solution in the analysis of existence and uniqueness of solutions to the SMF system.

If the pair $(\phi^\alpha(t, x), \psi^\alpha(t, x))$ is a smooth solution to (6.67) that satisfies the following mild form by a Duhamel Principle [97]:

$$\begin{aligned} \phi^\alpha(t, \omega, x) = \int_t^T \exp\left(\int_t^s A^\alpha(\tau, \omega, x)d\tau\right) \left(g^\alpha[s, \omega, x, \phi^\alpha(s, \omega, x), \psi^\alpha(s, \omega, x)]\right) ds \\ - \int_t^T \exp\left(\int_t^s A^\alpha(\tau, \omega, x)d\tau\right) \left((\psi^\alpha)^T(s, \omega, x)\right) dW(s, \omega). \end{aligned} \quad (6.68)$$

We define the operators:

$$\begin{aligned} \Phi^\alpha(t, s, \omega, x)(\cdot) &= \exp\left(\int_t^s A^\alpha(\tau, \omega, x)(\cdot)d\tau\right) \equiv \exp\left(\int_t^s \frac{1}{2}\text{tr}(a^\alpha[\tau, \omega, x]D_{xx}^2(\cdot))d\tau\right), \\ \Psi^\alpha(t, s, \omega, x)(\cdot) &= \int_t^s \partial_\alpha A^\alpha(\tau, \omega, x)(\cdot)d\tau \equiv \int_t^s \frac{1}{2}\text{tr}(\partial_\alpha a^\alpha[\tau, \omega, x]D_{xx}^2(\cdot))d\tau, \end{aligned}$$

in $[0, T] \times \Omega \times \mathbb{R}^n$ which are maps on $C^\infty(\mathbb{R}^n)$ and $C^2(\mathbb{R}^n)$, respectively.

Differentiating (6.68) with respect to α gives

$$\begin{aligned}
\partial_\alpha \phi^\alpha(t, \omega, x) &= \int_t^T (\Phi^\alpha(t, s, \omega, x)) (\Psi^\alpha(t, s, \omega, x)) \\
&\quad \left(g^\alpha[s, \omega, x, \phi^\alpha(s, \omega, x), \psi^\alpha(s, \omega, x)] \right) ds \\
&+ \int_t^T (\Phi^\alpha(t, s, \omega, x)) \left(\partial_\alpha g^\alpha[s, \omega, x, \phi^\alpha(s, \omega, x), \psi^\alpha(s, \omega, x)] \right) ds \\
&- \int_t^T (\Phi^\alpha(t, s, \omega, x)) (\Psi^\alpha(t, s, \omega, x)) \left((\psi^\alpha)^T(s, \omega, x) \right) dW(s, \omega) \\
&- \int_t^T (\Phi^\alpha(t, s, \omega, x)) \left((\partial_\alpha \psi^\alpha)^T(s, \omega, x) \right) dW(s, \omega), \tag{6.69}
\end{aligned}$$

where

$$\begin{aligned}
\partial_\alpha g^\alpha[t, \omega, x, \phi^\alpha(t, \omega, x), \psi^\alpha(t, \omega, x)] &\equiv \partial_\alpha H^\alpha[t, \omega, x, D_x \phi^\alpha(t, \omega, x)] \\
&+ \partial_p H^\alpha[t, \omega, x, D_x \phi^\alpha(t, \omega, x)] D_x (\partial_\alpha \phi^\alpha(t, \omega, x)) \\
&+ \langle \partial_\alpha \sigma^\alpha[t, \omega, x], D_x \psi^\alpha(t, \omega, x) \rangle + \langle \sigma^\alpha[t, \omega, x], D_x (\partial_\alpha \psi^\alpha(t, \omega, x)) \rangle.
\end{aligned}$$

We may rewrite (6.69) as

$$\begin{aligned}
\partial_\alpha \phi^\alpha(t, \omega, x) &= \int_t^T (\Phi^\alpha(t, s, \omega, x)) A_1^\alpha(s, \omega, x) (\partial_\alpha \phi^\alpha(t, \omega, x)) ds \\
&+ \int_t^T (\Phi^\alpha(t, s, \omega, x)) \left(h_1^\alpha[t, s, \omega, x, \partial_\alpha \psi^\alpha] \right) ds \\
&- \int_t^T (\Phi^\alpha(t, s, \omega, x)) \left((\partial_\alpha \psi^\alpha)^T(s, \omega, x) \right) dW(s, \omega), \\
&- \int_t^T (\Phi^\alpha(t, s, \omega, x)) \left(h_2^\alpha[t, s, \omega, x] \right) dW(s, \omega), \tag{6.70}
\end{aligned}$$

where

$$\begin{aligned}
 A_1^\alpha(s, \omega, x)(\cdot) &:= \partial_p H^\alpha[s, \omega, x, D_x \phi^\alpha(s, \omega, x)] D_x(\cdot), \\
 h_1^\alpha[t, s, \omega, x, \partial_\alpha \psi^\alpha] &:= (\Psi^\alpha(t, s, \omega, x)) \left(g^\alpha[s, \omega, x, \phi^\alpha(s, \omega, x), \psi^\alpha(s, \omega, x)] \right) \\
 &\quad + \partial_\alpha H^\alpha[s, \omega, x, D_x \phi^\alpha(s, \omega, x)] + \langle \partial_\alpha \sigma^\alpha[s, \omega, x], D_x \psi^\alpha(s, \omega, x) \rangle \\
 &\quad + \langle \sigma^\alpha[s, \omega, x], D_x(\partial_\alpha \psi^\alpha) \rangle, \\
 h_2^\alpha[t, s, \omega, x] &:= (\Psi^\alpha(t, s, \omega, x)) \left((\psi^\alpha)^T(s, \omega, x) \right).
 \end{aligned}$$

We introduce the following assumption:

(H6.5) $\partial_\alpha f^\alpha[t, x, u]$, $\partial_\alpha L^\alpha[t, x, u]$, $\partial_\alpha \sigma^\alpha[t, x]$ and $\partial_\alpha \varsigma^\alpha[t, x]$ exist and are $C^\infty(\mathbb{R}^n)$.

Assume **(H6.1)**-**(H6.3)** hold where $(f, L, \sigma, \varsigma)$ are replaced by $(\partial_\alpha f^\alpha, \partial_\alpha L^\alpha, \partial_\alpha \sigma^\alpha, \partial_\alpha \varsigma^\alpha)$, and all the boundedness assumptions are uniform.

PROPOSITION 6.3. *Assume **(H11)**-**(H6.3)** hold for $(f^\alpha, L^\alpha, \sigma^\alpha, \varsigma^\alpha)$. Let the pair $(\phi^\alpha(t, x), \psi^\alpha(t, x))$ be the unique solution to (6.65) which are $C^\infty(\mathbb{R}^n)$ and a.s. uniformly bounded. In addition, we assume **(H6.5)** holds. Then, the equation (6.69) has a unique solution*

$$(\partial_\alpha \phi(t, x), \partial_\alpha \psi(t, x)) \in (L^2_{\mathcal{F}_t}([0, T]; \mathbb{R}), L^2_{\mathcal{F}_t}([0, T]; \mathbb{R}^m))$$

such that $\sup_{0 \leq t \leq T} |D_x \partial_\alpha \phi(t, \cdot)| < \infty$ (a.s.).

Proof. The proof of existence and uniqueness of solution to (6.70) follows from Theorem 4.1 in [73] (see the proof of Theorem 4.1 in [145], see also [72, 113, 115] or Chapter 5 of [114]). By taking the conditional expectation $E_{\mathcal{F}_t^{w_0}}$ of the square of both sides of (6.70) and the boundedness assumptions in the theorem, one can show $\sup_{0 \leq t \leq T} |\partial_\alpha \phi(t, \cdot)| < \infty$ (a.s.) (see the proof of Theorem 2.1 in [145]). Using this in equation (6.70) implies the boundedness of $D_x \partial_\alpha \phi(t, \cdot)$. \square

We define the Gâteaux derivative of the function $F(t, x, \mu)$ with respect to the measure $\mu(y)$ as [97]

$$\partial_{\mu(y)} F(t, x, \mu) = \lim_{\epsilon \rightarrow 0} \frac{F(t, x, \mu + \epsilon \delta(y)) - F(t, x, \mu)}{\epsilon},$$

where δ is the Dirac delta function. We introduce the following assumptions:

(A6.12) (i) In (6.25)-(6.26) the Gâteaux derivative of f_0 , σ_0 and L_0 with respect to μ exist, are $C^\infty(\mathbb{R}^n)$ and a.s. uniformly bounded. (ii) In (6.32)-(6.33) the partial derivatives of f , σ and L with respect to μ^0 and μ exist, are $C^\infty(\mathbb{R}^n)$ and a.s. uniformly bounded.

The proof of the following lemma is based on the sensitivity analysis of the SHJB equations (6.39) and (6.42) to the stochastic measures $\mu_{(\cdot)}(\omega)$ and $\mu_{(\cdot)}^0(\omega)$.

LEMMA 6.2. (i) Assume **(A6.3)**-**(A6.7)** for U_0 , f_0 , σ_0 , L_0 , and **(A6.12)**-(i) hold. Let $(\phi_0(t, x), \psi_0(t, x))$ be the unique solution pair to (6.39) which is $C^\infty(\mathbb{R}^n)$ and is a.s. uniformly bounded. In addition, we assume **(A6.8)** holds for S_0 and the resulting u_0 is also a.s. Lipschitz continuous in μ . Then, for $\mu_{(\cdot)}(\omega)$ and $\nu_{(\cdot)}(\omega)$ in the set \mathcal{M}_ρ^β , $0 < \beta < 1$, there exists a constant c_4 such that

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^n} |u_0(t, \omega, x) - u'_0(t, \omega, x)|^2 \leq c_4 \left(D_T^\rho(\mu(\omega), \nu(\omega)) \right)^2, \quad a.s., \quad (6.71)$$

where $u_0, u'_0 \in C_{Lip(x)}([0, T] \times \Omega \times \mathbb{R}^n; U_0)$ are induced by the map Υ_0^{SHJB} in (6.47) using two stochastic measures $\mu_{(\cdot)}(\omega)$ and $\nu_{(\cdot)}(\omega)$, respectively.

(ii) Assume **(A6.3)**-**(A6.7)** for U , f , σ , L , and **(A6.12)**-(ii) hold. Let $(\phi(t, x), \psi(t, x))$ be the unique solution pair to (6.42) which is $C^\infty(\mathbb{R}^n)$ and is a.s. uniformly bounded. In addition, we assume **(A6.8)** holds for S and the resulting u is also a.s. Lipschitz continuous in μ . Then, for $\mu_{(\cdot)}^0(\omega) \in \mathcal{M}_\rho^\gamma$, $0 < \gamma < 1/2$, and $\mu_{(\cdot)}(\omega)$ and $\nu_{(\cdot)}(\omega)$ in the set \mathcal{M}_ρ^β , $0 < \beta < 1$, there exists a constant c_5 such that

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^n} |u(t, \omega, x) - u'(t, \omega, x)|^2 \leq c_5 \left(D_T^\rho(\mu(\omega), \nu(\omega)) \right)^2, \quad a.s., \quad (6.72)$$

where $u, u' \in C_{Lip(x)}([0, T] \times \Omega \times \mathbb{R}^n; U)$ are induced by the map Υ_i^{SHJB} in (6.52) using two stochastic measures $\mu_{(\cdot)}(\omega)$ and $\nu_{(\cdot)}(\omega)$, respectively.

(iii) Assume (A6.3)-(A6.7) for U, f, σ, L , and (A6.12)-(ii) hold. Let $(\phi(t, x), \psi(t, x))$ be the unique solution pair to (6.42) which is $C^\infty(\mathbb{R}^n)$ and is a.s. uniformly bounded. In addition, we assume (A6.8) holds for S and the resulting u is also a.s. Lipschitz continuous in μ^0 . Then, for $\mu_{(\cdot)}(\omega) \in \mathcal{M}_\rho^\beta$, $0 < \beta < 1$, and $\mu_{(\cdot)}^0(\omega)$ and $\nu_{(\cdot)}^0(\omega)$ in the set \mathcal{M}_ρ^γ , $0 < \gamma < 1/2$, there exists a constant c_6 such that

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^n} |u(t, \omega, x) - u'(t, \omega, x)|^2 \leq c_6 \left(D_T^\rho(\mu^0(\omega), \nu^0(\omega)) \right)^2, \quad \text{a.s.}, \quad (6.73)$$

where $u, u' \in C_{Lip(x)}([0, T] \times \Omega \times \mathbb{R}^n; U)$ are induced by the map Υ_i^{SHJB} in (6.52) using the two stochastic measures $\mu_{(\cdot)}^0(\omega)$ and $\nu_{(\cdot)}^0(\omega)$, respectively.

Proof: (i) Assumption (A6.8) for S_0 together with the fact that the resulting u_0 in (A6.8) is also a.s. Lipschitz continuous in μ yields

$$\begin{aligned} |u_0(t, \omega, x) - u'_0(t, \omega, x)| &\leq k_1 D_t^\rho(\mu(\omega), \nu(\omega)) \\ &\quad + k_2 |D_x \phi_0^\mu(t, \omega, x) - D_x \phi_0^\nu(t, \omega, x)|, \end{aligned} \quad (6.74)$$

with positive constants k_1, k_2 , where we indicate the dependence of ϕ_0 on measures μ and ν by ϕ_0^μ and ϕ_0^ν , respectively.

We consider the Gâteaux derivative of ϕ_0 with respect to the measure μ . The assumptions of the theorem imply that the conditions for Proposition 6.3 hold. Therefore, Proposition 6.3 yields that the Gâteaux derivative of $D_x \phi_0$ with respect to measure μ is a.s. uniformly bounded. This together with the mean value theorem yields

$$|D_x \phi_0^\mu(t, \omega, x) - D_x \phi_0^\nu(t, \omega, x)| \leq k_3 D_t^\rho(\mu(\omega), \nu(\omega)), \quad (6.75)$$

with positive constant k_3 . (6.74) and (6.75) give

$$|u_0(t, \omega, x) - u'_0(t, \omega, x)| \leq k D_t^\rho(\mu(\omega), \nu(\omega)),$$

with $k := k_1 + k_2 k_3$, which yields the result. \square

REMARK 6.2. *In the standard mean field game model of [85] a similar condition to (6.71)-(6.73) is taken as an assumption (see the feedback regularity condition (37) in [85]). Following the argument in Section 7.1 of [85], one can show that the inequalities (6.71)-(6.73) hold in the linear-quadratic-Gaussian (LQG) model with Lipschitz continuous nonlinear couplings.*

6.6.3.2. *Main Result.* We recall the map Υ given in (6.60) which is the composition of the maps Υ_0^{SHJB} , Υ_0^{SMV} , Υ_i^{SHJB} and Υ_i^{SMV} introduced in (6.47), (6.48), (6.52), and (6.59), respectively (see the diagram below).

$$\begin{array}{ccc} \mu_{(\cdot)}(\omega) & \xrightarrow{\Upsilon_0^{\text{SHJB}}} & u_0^o(\cdot, \omega, x) \\ \uparrow \Upsilon_i^{\text{SMV}} & & \downarrow \Upsilon_0^{\text{SMV}} \\ u^o(\cdot, \omega, x) & \xleftarrow{\Upsilon_i^{\text{SHJB}}} & \mu_{(\cdot)}^0(\omega) \end{array}$$

THEOREM 6.10. *Let the assumptions of both Lemma 6.1 and Lemma 6.2 hold. If the constants $\{c_i : 0 \leq i \leq 6\}$ for (6.61)-(6.64) and (6.71)-(6.73) satisfy the gain condition*

$$\max \{c_2c_5, c_2c_6c_0, c_2c_6c_1, c_3c_1, c_3c_0c_4\} < 1,$$

then there exists a unique solution for the map Υ , and hence a unique solution to the MM SMF system (6.39)-(6.41) and (6.42)-(6.44).

Proof: The result follows from the Banach fixed point theorem for the map Υ given in (6.60) on the Polish space \mathcal{M}_ρ^β , $0 < \beta < 1$. We note that the gain condition ensures that Υ is a contraction. \square

6.7. Applications

6.7.1. **The MM SMF LQG System.** We consider the Major and Minor (MM) Linear-Quadratic-Gaussian (LQG) dynamic game problem of [75]. In this case

all functions in (6.1)-(6.4) are given by (see Remark 6.1)

$$f_0[t, z_0^N(t), u_0^N(t), z_j^N(t)] = A_0 z_0^N(t) + B_0 u_0^N(t) + F_0 z_j^N(t),$$

$$f[t, z_i^N(t), u_i^N(t), z_0^N(t), z_j^N(t)] = A z_i^N(t) + B u_i^N(t) + F z_j^N(t) + G z_0^N(t),$$

$$\sigma_0[t, z_0^N(t), z_j^N(t)] = S_0, \quad \sigma[t, z_i^N(t), z_0^N(t), z_j^N(t)] = S,$$

$$\begin{aligned} L_0[t, z_0^N(t), u_0^N(t), z_j^N(t)] &= \left[z_0^N(t) - \left(H_0 \left(\frac{1}{N} \sum_{j=1}^N z_j^N(t) \right) + \eta_0 \right) \right]^T Q_0 \\ &\quad \times \left[z_0^N(t) - \left(H_0 \left(\frac{1}{N} \sum_{j=1}^N z_j^N(t) \right) + \eta_0 \right) \right]^T + (u_0^N(t))^T R_0 u_0^N(t), \end{aligned}$$

$$\begin{aligned} L[t, z_i^N(t), u_i^N(t), z_0^N(t), z_j^N(t)] &= \left[z_i^N(t) - \left(H z_0^N(t) + \hat{H} \left(\frac{1}{N} \sum_{j=1}^N z_j^N(t) \right) + \eta \right) \right]^T Q \\ &\quad \times \left[z_i^N(t) - \left(H z_0^N(t) + \hat{H} \left(\frac{1}{N} \sum_{j=1}^N z_j^N(t) \right) + \eta \right) \right]^T + (u_i^N(t))^T R u_i^N(t), \end{aligned}$$

with the deterministic constant matrices: (i) $A_0, F_0, A, F, G, H_0, H$ and \hat{H} in $\mathbb{R}^{n \times n}$, (ii) B_0 and B in $\mathbb{R}^{n \times k}$, (iii) S_0 and S in $\mathbb{R}^{n \times m}$, (iv) the symmetric nonnegative definite matrices Q_0 and Q in $\mathbb{R}^{n \times n}$, (v) the symmetric positive definite matrices R_0 and R in $\mathbb{R}^{k \times k}$, and the deterministic constant vectors η and η_0 are in \mathbb{R}^n .

In this formulation the major agent's SMF system (6.39)-(6.41) is of the form

$$\begin{aligned} -d\phi_0(t, \omega, x) &= \left[\left\langle A_0 x - \frac{1}{4} B_0 R_0^{-1} B_0^T D_x \phi_0(t, \omega, x) + F_0 z^o(t, \omega), D_x \phi_0(t, \omega, x) \right\rangle \right. \\ &\quad \left. + \left\langle x - (H_0 z^o(t, \omega) + \eta_0), Q_0 (x - (H_0 z^o(t, \omega) + \eta_0)) \right\rangle \right. \\ &\quad \left. + \left\langle S_0, D_x \psi_0(t, \omega, x) \right\rangle + \frac{1}{2} \text{tr}((S_0^T S_0) D_{xx}^2 \phi_0(t, \omega, x)) \right] dt \\ &\quad - \psi_0^T(t, \omega, x) d\omega_0(t, \omega), \quad \phi_0(T, x) = 0, \end{aligned} \quad (6.76)$$

$$u_0^o(t, \omega, x) = -\frac{1}{2} R_0^{-1} B_0^T D_x \phi_0(t, \omega, x), \quad (6.77)$$

$$\begin{aligned} dz_0^o(t, \omega) &= \left[A_0 z_0^o(t, \omega) + B_0 u_0^o(t, \omega, z_0^o) + F_0 z^o(t, \omega) \right] dt \\ &\quad + S_0 d\omega_0(t, \omega), \quad z_0^o(0) = z_0(0), \end{aligned} \quad (6.78)$$

and the minor agents' SMF system (6.42)-(6.44) is given by

$$\begin{aligned}
-d\phi(t, \omega, x) = & \left[\left\langle Ax - \frac{1}{4}BR^{-1}B^T D_x\phi(t, \omega, x) + Fx + Gz_0^o(t, \omega), D_x\phi(t, \omega, x) \right\rangle \right. \\
& + \left\langle x - (Hz_0^o(t, \omega) + \hat{H}x + \eta), Q(x - (Hz_0^o(t, \omega) + \hat{H}x + \eta)) \right\rangle \\
& \left. + \frac{1}{2}\text{tr}((S^T S)D_{xx}^2\phi(t, \omega, x)) \right] dt - \psi^T(t, \omega, x)dw(t, \omega), \quad \phi_0(T, x) = 0, \quad (6.79)
\end{aligned}$$

$$u^o(t, \omega, x) = -\frac{1}{2}R^{-1}B^T D_x\phi(t, \omega, x), \quad (6.80)$$

$$\begin{aligned}
dz^o(t, \omega) = & \left[Az^o(t, \omega) + Bu^o(t, \omega, z^o) + F_0z^o(t, \omega) + Gz_0^o(t, \omega) \right] dt \\
& + Sdw(t, \omega), \quad z_0^o(0) = z_0(0). \quad (6.81)
\end{aligned}$$

Let $\Pi_0(\cdot) \geq 0$ be the unique solution of the deterministic Riccati equation

$$\partial_t \Pi_0(t) + \Pi_0(t)A_0 + A_0^T \Pi_0(t) - \Pi_0(t)B_0R_0^{-1}B_0^T \Pi_0(t) + Q_0 = 0, \quad \Pi_0(T) = 0.$$

We denote $\mathbb{A}_0(\cdot) = A_0 - B_0R_0^{-1}B_0^T \Pi_0(\cdot)$. It can be verified that the pair $(\phi_0, \psi_0)(t, \omega, x)$ in (6.39) is given by

$$\begin{aligned}
\phi_0(t, \omega, x) &= x^T \Pi_0(t)x + 2x^T s_0(t, \omega) + g_0(t, \omega), \\
\psi_0^T(t, \omega, x) &= 2x^T q_0(t, \omega) + h_0(t, \omega),
\end{aligned}$$

where $(s_0, q_0)(t, \omega)$ and $(g_0, h_0)(t, \omega)$ are unique solutions of the following Backward Stochastic Differential Equations (BSDEs):

$$\begin{aligned}
-ds_0(t, \omega) &= \left[\mathbb{A}_0^T(t)s_0(t, \omega) + (\Pi_0(t)F_0 - Q_0H_0)z^o(t, \omega) - Q_0\eta_0 \right] dt \\
& - q_0(t, \omega)dw_0(t, \omega), \quad s_0(T) = 0, \\
-dg_0(t, \omega) &= \left[-s_0^T(t, \omega)B_0R_0^{-1}B_0^T s_0(t, \omega) + 2F_0z^o(t, \omega) + 2\text{tr}(S_0^T q_0(t, \omega)) \right. \\
& \left. + (H_0z^o(t, \omega) + \eta_0)^T Q_0(H_0z^o(t, \omega) + \eta_0) + \text{tr}(S_0^T S_0 \Pi_0(t)) \right] dt \\
& - h_0(t, \omega)dw_0(t, \omega), \quad g_0(T) = 0.
\end{aligned}$$

We may now express the major agent's SMF-LQG system (6.76)-(6.78) in the following form:

$$\begin{aligned}
 -ds_0(t, \omega) &= \left[\mathbb{A}_0^T(t)s_0(t, \omega) + (\Pi_0(t)F_0 - Q_0H_0)z^o(t, \omega) - Q_0\eta_0 \right] dt \\
 &\quad - q_0(t, \omega)dw_0(t, \omega), \quad s_0(T) = 0, \\
 u_0^o(t, \omega) &= -R_0^{-1}B_0^T(\Pi_0(t)z_0^o(t, \omega) + s_0(t, \omega)), \\
 dz_0^o(t, \omega) &= \left[\mathbb{A}_0(t)z_0^o(t, \omega) - B_0R_0^{-1}B_0^T\Pi_0(t)s_0(t, \omega) + F_0z^o(t, \omega) \right] dt \\
 &\quad + S_0dw_0(t, \omega), \quad z_0^o(0) = z_0(0),
 \end{aligned}$$

where $z^o(t, \omega)$ is the mean field behaviour of the minor agents (see the minor agents' SMF-LQG system below).

In a similar way, let $\Pi(\cdot) \geq 0$ be the unique solution of the deterministic Riccati equation

$$\partial_t \Pi(t) + \Pi(t)A + A^T \Pi(t) - \Pi(t)BR^{-1}B^T \Pi(t) + Q = 0, \quad \Pi(T) = 0.$$

We denote $\mathbb{A}(\cdot) = A - BR^{-1}B^T \Pi(\cdot)$. It can be verified that the pair $(\phi, \psi)(t, \omega, x)$ in (6.42) is given by

$$\begin{aligned}
 \phi(t, \omega, x) &= x^T \Pi(t)x + 2x^T s(t, \omega) + g(t, \omega), \\
 \psi^T(t, \omega, x) &= 2x^T q(t, \omega) + h(t, \omega),
 \end{aligned}$$

where $(s, q)(t, \omega)$ and $(g, h)(t, \omega)$ are unique solutions of the following BSDEs:

$$\begin{aligned}
 -ds(t, \omega) &= \left[\mathbb{A}^T(t)s(t, \omega) + (\Pi(t)F - Q\hat{H})z^o(t, \omega) + (\Pi(t)G - QH)z_0^o(t, \omega) \right. \\
 &\quad \left. - Q\eta \right] dt - q(t, \omega)dw_0(t, \omega), \quad s(T) = 0, \\
 -dg(t, \omega) &= \left[-s^T(t, \omega)BR^{-1}B^T s(t, \omega) + 2Fz^o(t, \omega) + 2Gz_0^o(t, \omega) \right. \\
 &\quad \left. + (\hat{H}z^0(t, \omega) + Hz_0^o(t, \omega) + \eta)^T Q_0(\hat{H}z^0(t, \omega) + Hz_0^o(t, \omega) + \eta) \right. \\
 &\quad \left. + \text{tr}(S^T S \Pi(t)) \right] dt - h(t, \omega)dw_0(t, \omega), \quad g(T) = 0.
 \end{aligned}$$

We may now express the minor agents' SMF-LQG system (6.79)-(6.81) in the following form:

$$\begin{aligned}
-ds(t, \omega) &= \left[\mathbb{A}^T(t)s(t, \omega) + (\Pi(t)F - Q\hat{H})z^o(t, \omega) + (\Pi(t)G - QH)z_0^o(t, \omega) \right. \\
&\quad \left. - Q\eta \right] dt - q(t, \omega)dw_0(t, \omega), & s(T) &= 0, \\
u^o(t, \omega) &= -R^{-1}B^T(\Pi(t)z^o(t, \omega) + s(t, \omega)), \\
dz^o(t, \omega) &= \left[(\mathbb{A}(t) + F)z^o(t, \omega) - BR^{-1}B^T\Pi(t)s(t, \omega) + Gz_0^o(t, \omega) \right] dt \\
&\quad + Sdw(t, \omega), & z^o(0) &= z(0).
\end{aligned}$$

So we retrieve the SMF system for the MM LQG dynamic games model of [124] for minor agents with uniform parameters (see equations (2.10) and (2.22) in [124], see also [75]). The reader is referred to [124] for an explicit representation of a solution to the SMF-LQG system under some appropriate conditions.

6.7.2. Synchronization of Coupled Nonlinear Oscillators Game.

In this section we present a major and minor version of the synchronization of coupled nonlinear oscillators game model [177]. Consider a population of $N + 1$ oscillators with dynamics

$$d\theta_j^N(t) = u_j^N(t)dt + \sigma dw_j(t) \pmod{2\pi} \quad 0 \leq j \leq N, \quad t \geq 0,$$

where $\theta_j(t) \in [0, 2\pi]$ is the phase of the j^{th} oscillator at time t , $u_j(\cdot)$ is the control input, σ is a non-negative scalar, and $\{w_j : 0 \leq j \leq N\}$ denotes a sequence of independent standard scalar Wiener processes. It is assumed that the initial states $\{\theta_j(0)\}$ are chosen independently on $[0, 2\pi]$. The objective of the j^{th} oscillator is to

minimize its own cost function

$$\begin{aligned}
 J_0^N(u_0^N, u_{-0}^N) &:= E \int_0^T \left(\frac{1}{N} \sum_{k=1}^N \sin^2 [\theta_0^N(t) - \theta_k^N(t)] + r(u_0^N(t))^2 \right) dt, \\
 J_i^N(u_i^N, u_{-i}^N) &:= E \int_0^T \left(\frac{1}{N} \sum_{k=1}^N \sin^2 [\theta_i^N(t) - (\lambda\theta_0^N(t) + (1-\lambda)\theta_k^N(t))] \right. \\
 &\quad \left. + r(u_i^N(t))^2 \right) dt, \quad 1 \leq i \leq N,
 \end{aligned}$$

where r is a positive scalar and $\lambda \in (0, 1)$.

Similar arguments in previous section yield the following major agent's SMF system (6.31) and (6.39)-(6.40):

$$\begin{aligned}
 -d\phi_0(t, \omega, x) &= \left[-\frac{1}{4r} (\partial_x \phi_0(t, \omega, x))^2 + m_0(t, \omega, x) + \sigma \partial_x \psi_0(t, \omega, x) \right. \\
 &\quad \left. + \frac{\sigma^2}{2} \partial_{xx}^2 \phi_0(t, \omega, x) \right] dt - \psi_0(t, \omega, x) dw_0(t, \omega), \quad \phi_0(T, x) = 0, \\
 u_0^o(t, \omega, x) &= -\frac{1}{2r} \partial_x \phi_0(t, \omega, x), \\
 dp_s^o(t, \omega, x) &= \left[\frac{1}{2r} \partial_x \left((\partial_x \phi_0(t, \omega, x)) p_s^o(t, \omega, x) \right) + \frac{\sigma^2}{2} \partial_{xx}^2 p_s^o(t, \omega, x) \right] dt \\
 &\quad - \sigma \partial_x p_s^o(t, \omega, x) dw_0(t, \omega), \quad p_s^o(s, x) = \delta_{\theta_0^o(s)}(dx), \\
 m_0(t, \omega, x) &= \int_0^{2\pi} \sin^2(x - \theta) p(t, \omega, \theta) d\theta,
 \end{aligned}$$

where $m_0(t, \omega, x)$ is called the infinite population cost-coupling of the major agent, and $\theta_0^o(\cdot)$ is the solution of the closed-loop equation

$$d\theta_0^o(t) = u_0^o(t, \theta_0^o(t)) dt + \sigma dw_0(t) \pmod{2\pi} \quad t \geq 0.$$

In a similar way, the minor agents' SMF system (6.38) and (6.42)-(6.43) is given by

$$\begin{aligned}
-d\phi(t, \omega, x) &= \left[-\frac{1}{4r} (\partial_x \phi(t, \omega, x))^2 + m(t, \omega, x) + \frac{\sigma^2}{2} \partial_{xx}^2 \phi(t, \omega, x) \right] dt \\
&\quad - \psi(t, \omega, x) dw(t, \omega), \quad \phi(T, x) = 0, \\
u^o(t, \omega, x) &= -\frac{1}{2r} \partial_x \phi(t, \omega, x), \\
dp(t, \omega, x) &= \left[\frac{1}{2r} \partial_x \left((\partial_x \phi(t, \omega, x)) p(t, \omega, x) \right) + \frac{\sigma^2}{2} \partial_{xx}^2 p(t, \omega, x) \right] dt, \quad p(0, x) \\
m(t, \omega, x) &= \int_0^{2\pi} \int_0^{2\pi} \sin^2(x - (\lambda\theta_0 + (1-\lambda)\theta)) p_0^0(t, \omega, \theta_0) p(t, \omega, \theta) d\theta_0 d\theta,
\end{aligned}$$

where $m(t, \omega, x)$ is called the infinite population cost-coupling of the major agent. The reader is referred to the deterministic mean field system (14a)-(14c) in [177] for the synchronization of coupled nonlinear oscillators game model with only minor agents.

6.8. ϵ -Nash Equilibrium Property of the SMF Control Laws

We let

$$(\phi_0(t, \omega, x), \psi_0(t, \omega, x), u_0^o(t, \omega, x), z_0^o(t, \omega), \phi(t, \omega, x), \psi(t, \omega, x), u^o(t, \omega, x), z^o(t, \omega)),$$

be the unique solution of the MM SMF system (6.39)-(6.41) and (6.42)-(6.44) such that the best response SMF control processes $u_0^o(t, \omega, x)$ and $u^o(t, \omega, x)$ are a.s. continuous in (t, x) and a.s. Lipschitz continuous in x .

We now apply the SMF control processes $u_0^o(t, \omega, x)$ and $u^o(t, \omega, x)$ into a finite $N + 1$ major and minor population (6.1)-(6.2). This yields the following closed loop

individual dynamics:

$$\begin{aligned} dz_0^{o,N}(t) &= \frac{1}{N} \sum_{j=1}^N f_0[t, z_0^{o,N}(t), u_0^o(t, z_0^{o,N}(t)), z_j^{o,N}(t)] dt \\ &+ \frac{1}{N} \sum_{j=1}^N \sigma_0[t, z_0^{o,N}(t), z_j^{o,N}(t)] dw_0(t), \quad z_0^{o,N}(0) = z_0(0), \quad 0 \leq t \leq T, \end{aligned} \quad (6.82)$$

$$\begin{aligned} dz_i^{o,N}(t) &= \frac{1}{N} \sum_{j=1}^N f[t, z_i^{o,N}(t), u^o(t, z_i^{o,N}(t)), z_0^{o,N}(t), z_j^{o,N}(t)] dt \\ &+ \frac{1}{N} \sum_{j=1}^N \sigma[t, z_i^{o,N}(t), z_0^{o,N}(t), z_j^{o,N}(t)] dw_i(t), \quad z_i^{o,N}(0) = z_i(0), \quad 1 \leq i \leq N, \end{aligned} \quad (6.83)$$

We set the admissible control set of agent \mathcal{A}_j , $0 \leq j \leq N$, as

$\mathcal{U}_j = \left\{ u_j(\cdot, \omega) := u_j(\cdot, \omega, z_0(\cdot, \omega), \dots, z_N(\cdot, \omega)) \in C_{\text{Lip}(z_0, \dots, z_N)} : u_j(t, \omega) \text{ is a}$

$\mathcal{F}_t^{w_0}$ -measurable process adapted to sigma-field $\sigma\{z_i(\tau, \omega) : 0 \leq i \leq N, 0 \leq \tau \leq t\}$

such that $E \int_0^T |u_j(t, \omega)|^2 dt < \infty \}$.

We note that \mathcal{U}_j , $0 \leq j \leq N$, are the full information admissible control which are not restricted to be decentralized.

DEFINITION 6.2. *Given $\epsilon > 0$, the admissible control laws (u_0^o, \dots, u_N^o) for $N + 1$ agents generates an ϵ -Nash equilibrium with respect to the costs J_j^N , $0 \leq j \leq N$, if $J_j^N(u_j^o; u_{-j}^o) - \epsilon \leq \inf_{u_j \in \mathcal{U}_j} J_j^N(u_j; u_{-j}^o) \leq J_j^N(u_j^o; u_{-j}^o)$, for any $0 \leq j \leq N$. \square*

We now show that the SMF control processes for a finite $N + 1$ major and minor population system (6.82)-(6.83) is an ϵ -Nash equilibrium with respect to the cost functions (6.3)-(6.4) in the case that minor agents are coupled to the major agent only through their cost functions (see the MM LQG model in [124]).

(A6.13) Assume the functions f and σ in (6.2) (and hence in (6.83)) do not contain the state of major agent z_0^N .

Note that in the case of assumption **(A6.13)** the major agent \mathcal{A}_0 has a significant influence on the minor agents through their cost functions (6.4). We note that an

analysis based on the anticipative variational calculations used in the MM LQG case [125] is required for establishing the ϵ -Nash equilibrium property of the SMF in the general case.

THEOREM 6.11. *Assume (A6.1)-(A6.6) and (A6.13) hold, and there exists a unique solution to the MM SMF system (6.39)-(6.41) and (6.42)-(6.44) such that the SMF best response control processes $u_0^o(t, \omega, x)$ and $u^o(t, \omega, x)$ are a.s. continuous in (t, x) and a.s. Lipschitz continuous in x . Then $(u_0^o, u_1^o, \dots, u_N^o)$ where $u_i^o \equiv u^o$, $1 \leq i \leq N$, generates an $O(\epsilon_N + 1/\sqrt{N})$ -Nash equilibrium with respect to the cost functions (6.3)-(6.4) such that $\lim_{N \rightarrow \infty} \epsilon_N = 0$.*

Proof. Under (A6.13) we have the the following closed loop individual dynamics under the SMF best response control processes:

$$\begin{aligned} dz_0^{o,N}(t) &= \frac{1}{N} \sum_{j=1}^N f_0[t, z_0^{o,N}(t), u_0^o(t, z_0^{o,N}(t)), z_j^{o,N}(t)] dt \\ &\quad + \frac{1}{N} \sum_{j=1}^N \sigma_0[t, z_0^{o,N}(t), z_j^{o,N}(t)] dw_0(t), \quad z_0^{o,N}(0) = z_0(0), \quad 0 \leq t \leq T, \\ dz_i^{o,N}(t) &= \frac{1}{N} \sum_{j=1}^N f[t, z_i^{o,N}(t), u^o(t, z_i^{o,N}(t)), z_j^{o,N}(t)] dt \\ &\quad + \frac{1}{N} \sum_{j=1}^N \sigma[t, z_i^{o,N}(t), z_j^{o,N}(t)] dw_i(t), \quad z_i^{o,N}(0) = z_i(0), \quad 1 \leq i \leq N. \end{aligned}$$

We also introduce the associated McKean-Vlasov (MV) SDE system

$$\begin{aligned} dz_0^o(t) &= f_0[t, z_0^o(t), u_0^o(t, z_0^o), \mu_t] dt + \sigma_0[t, z_0^o(t), \mu_t] dw_0(t), \\ dz_i^o(t) &= f[t, z_i^o(t), u^o(t, z_i^o), \mu_t] dt + \sigma[t, z_i^o, \mu_t] dw_i(t), \end{aligned} \tag{6.84}$$

with the initial condition $z_j^o(0) = z_j(0)$, $0 \leq j \leq N$. In the above MV equation μ_t , $0 \leq t \leq T$, is the conditional law of $z_i^o(t)$, $1 \leq i \leq N$, given $\mathcal{F}_t^{w_0}$ (i.e., $\mu_t :=$

$\mathcal{L}(z_i^o(t)|\mathcal{F}_t^{w_0})$, $1 \leq i \leq N$). Theorem 6.1 implies that

$$\sup_{0 \leq j \leq N} \sup_{0 \leq t \leq T} E|z_j^{o,N}(t) - z_j^o(t)| = O(1/\sqrt{N}), \quad (6.85)$$

where the right hand side may depend upon the terminal time T .

Let $z(0) = \int_{\mathbb{R}^n} x dF(x)$ be the mean value of the minor agents' initial states (see **(A6.2)**). We denote

$$(\epsilon_N)^2 = \left| \int_{\mathbb{R}^N} x^T x dF_N(x) - 2z^T(0) \int_{\mathbb{R}^N} x dF_N(x) + z^T(0)z(0) \right|.$$

It is evident from **(A6.2)** that $\lim_{N \rightarrow \infty} \epsilon_N = 0$. To prove the ϵ -Nash equilibrium property we consider two cases as follows.

Case I (strategy change for the major agent \mathcal{A}_0): While the minor agents are using the SMF best response control law $u^0(t, \omega, x)$, a strategy change from $u_0^0(t, \omega, x)$ to the $\mathcal{F}_t^{w_0}$ -adapted process $u_0(t, \omega, x, z_{-0}^{o,N}(t, \omega)) \in \mathcal{U}_0$ for the major agent yields

$$\begin{aligned} dz_0^N(t) &= \frac{1}{N} \sum_{j=1}^N f_0[t, z_0^N(t), u_0(t, z_0^N(t), z_{-0}^{o,N}(t)), z_j^{o,N}(t)] dt \\ &+ \frac{1}{N} \sum_{j=1}^N \sigma_0[t, z_0^N(t), z_j^{o,N}(t)] dw_0(t), \quad z_0^N(0) = z_0(0), \quad 0 \leq t \leq T, \end{aligned}$$

where $z_{-0}^{o,N} \equiv (z_1^{o,N}, \dots, z_N^{o,N})$. Since minor agents are coupled to the major agent only through their cost functions (see **(A6.13)**) the strategy change of the major agent does not affect the the minor agents' states $z_i^{o,N}$ and z_i^o , $1 \leq i \leq N$, above.

Let $\hat{z}_0^N(\cdot)$ be the solution of the SDE:

$$\begin{aligned} d\hat{z}_0^N(t) &= \frac{1}{N} \sum_{j=1}^N f_0[t, \hat{z}_0^N(t), u_0(t, \hat{z}_0^N(t), z_{-0}^o(t)), z_j^o(t)] dt \\ &+ \frac{1}{N} \sum_{j=1}^N \sigma_0[t, \hat{z}_0^N(t), z_j^o(t)] dw_0(t), \quad \hat{z}_0^N(0) = z_0(0), \quad 0 \leq t \leq T, \end{aligned}$$

where $z_{-0}^o \equiv (z_1^o, \dots, z_N^o)$ is given by the MV SDE system above. Theorem 6.1 and the Gronwall's lemma imply that

$$\sup_{0 \leq t \leq T} E|z_0^N(t) - \hat{z}_0^N(t)| = O(1/\sqrt{N}). \quad (6.86)$$

We also introduce the SDE

$$d\hat{z}_0(t) = f_0[t, \hat{z}_0(t), u_0(t, \hat{z}_0(t), z_{-0}^o(t)), \mu_t]dt + \sigma_0[t, \hat{z}_0(t), \mu_t]dw_0(t),$$

with initial condition $\hat{z}_0(0) = z_0(0)$, where $\mu_{(\cdot)}$ is the minor agents' measure given by the MV SDE system above. Again, by Theorem 6.1 and the Gronwall's lemma It can be shown that

$$\sup_{0 \leq t \leq T} E|\hat{z}_0^N(t) - \hat{z}_0(t)| = O(1/\sqrt{N}). \quad (6.87)$$

(A6.3), (A6.6), (6.85)-(6.87) and Theorem 6.1 yield

$$\begin{aligned} J_0^N(u_0; u_{-0}^o) &\equiv E \int_0^T \left((1/N) \sum_{j=1}^N L_0[t, z_0^N(t), u_0(t, z_0^N, z_{-0}^{o,N}), z_j^{o,N}(t)] \right) dt \\ &\stackrel{(6.85)}{\geq} E \int_0^T \left((1/N) \sum_{j=1}^N L_0[t, z_0^N(t), u_0(t, z_0^N, z_{-0}^o), z_j^o(t)] \right) dt - O(\epsilon_N + 1/\sqrt{N}) \\ &\stackrel{(6.86)}{\geq} E \int_0^T \left((1/N) \sum_{j=1}^N L_0[t, \hat{z}_0^N(t), u_0(t, \hat{z}_0^N, z_{-0}^o), z_j^o(t)] \right) dt - O(\epsilon_N + 1/\sqrt{N}) \\ &\stackrel{(6.87)}{\geq} E \int_0^T \left((1/N) \sum_{j=1}^N L_0[t, \hat{z}_0(t), u_0(t, \hat{z}_0, z_{-0}^o), z_j^o(t)] \right) dt - O(\epsilon_N + 1/\sqrt{N}) \\ &\stackrel{(6.5)}{\geq} E \int_0^T L_0[t, \hat{z}_0(t), u_0(t, \hat{z}_0, z_{-0}^o), \mu_t] dt - O(\epsilon_N + 1/\sqrt{N}), \end{aligned} \quad (6.88)$$

where the appearance of the ϵ_N term in the first inequality of (6.88) is due to the fact that here the sequence of minor agents' initials $\{z_j^o(0) : 1 \leq j \leq N\}$ in the SMV system (6.84) is generated by independent randomized observations on the distribution F given in (A6.2).

Furthermore, by the construction of the major agent's SMF system (6.39)-(6.41) (see the major agent's SOCP (6.25)-(6.26)) we have

$$E \int_0^T L_0[t, \hat{z}_0(t), u_0(t, \hat{z}_0, z_{-0}^o), \mu_t] dt \geq E \int_0^T L_0[t, z_0^o(t), u_0^o(t, z_0^o), \mu_t] dt. \quad (6.89)$$

But, Theorem 6.1 and (6.85) imply

$$\begin{aligned} & E \int_0^T L_0[t, z_0^o(t), u_0^o(t, z_0^o), \mu_t] dt \\ & \stackrel{(6.5)}{\geq} E \int_0^T \left((1/N) \sum_{j=1}^N L_0[t, z_0^o(t), u_0(t, z_0^o), z_j^o(t)] \right) dt - O(\epsilon_N + 1/\sqrt{N}) \\ & \stackrel{(6.85)}{\geq} E \int_0^T \left((1/N) \sum_{j=1}^N L_0[t, z_0^{o,N}(t), u_0(t, z_0^{o,N}), z_j^{o,N}(t)] \right) dt - O(\epsilon_N + 1/\sqrt{N}) \\ & \equiv J_0^N(u_0^o; u_{-0}^o) - O(\epsilon_N + 1/\sqrt{N}). \end{aligned} \quad (6.90)$$

It follows from (6.88)-(6.90) that $J_0^N(u_0^o; u_{-0}^o) - O(\epsilon_N + 1/\sqrt{N}) \leq \inf_{u_0 \in \mathcal{U}_0} J_0^N(u_0; u_{-0}^o)$.

Case II (strategy change for the minor agents): Without loss of generality, we assume that the first minor agent changes its MF best response control strategy $u^o(t, \omega, x)$ to $u_1(t, \omega, x, z_{-1}(t, \omega)) \in \mathcal{U}_1$. This leads to

$$\begin{aligned} dz_0^N(t) &= \frac{1}{N} \sum_{j=1}^N f_0[t, z_0^N, u_0^o(t, z_0^N), z_j^N] dt + \frac{1}{N} \sum_{j=1}^N \sigma_0[t, z_0^N, z_j^N] dw_0(t), \\ dz_1^N(t) &= \frac{1}{N} \sum_{j=1}^N f[t, z_1^N, u_1(t, z_1^N, z_{-1}^N), z_j^N] dt + \frac{1}{N} \sum_{j=1}^N \sigma[t, z_1^N, z_j^N] dw_1(t), \\ dz_2^N(t) &= \frac{1}{N} \sum_{j=1}^N f[t, z_2^N, u^o(t, z_2^N), z_j^N] dt + \frac{1}{N} \sum_{j=1}^N \sigma[t, z_2^N, z_j^N] dw_2(t), \\ &\vdots \\ dz_N^N(t) &= \frac{1}{N} \sum_{j=1}^N f[t, z_N^N, u^o(t, z_N^N), z_j^N] dt + \frac{1}{N} \sum_{j=1}^N \sigma[t, z_N^N, z_j^N] dw_N(t). \end{aligned}$$

By the same argument as in proving Theorem 6.1, it can be shown that

$$\begin{aligned} \sup_{j=0,2,\dots,N} \sup_{0 \leq t \leq T} E|z_j^{o,N}(t) - z_j^N(t)| &= O(1/\sqrt{N}), \\ \sup_{j=0,2,\dots,N} \sup_{0 \leq t \leq T} E|z_j^o(t) - z_j^N(t)| &= O(1/\sqrt{N}). \end{aligned}$$

Let $\hat{z}_1^N(\cdot)$ be the solution of the SDE:

$$\begin{aligned} d\hat{z}_1^N(t) &= \frac{1}{N} \sum_{j=1}^N f[t, \hat{z}_1^N(t), u_1(t, \hat{z}_1^N(t), z_{-1}^o(t)), z_j^o(t)] dt \\ &+ \frac{1}{N} \sum_{j=1}^N \sigma[t, \hat{z}_1^N(t), z_j^o(t)] dw_1(t), \quad \hat{z}_1^N(0) = z_1(0), \quad 0 \leq t \leq T, \end{aligned}$$

where $z_{-1}^o \equiv (z_1^o, \dots, z_N^o)$ is given by the MV SDE system above. Theorem 6.1 and the Gronwall's lemma implies that

$$\sup_{0 \leq t \leq T} E|z_1^N(t) - \hat{z}_1^N(t)| = O(1/\sqrt{N}). \quad (6.91)$$

We also introduce the SDE

$$d\hat{z}_1(t) = f[t, \hat{z}_1(t), u_1(t, \hat{z}_1(t), z_{-1}^o(t)), \mu_t] dt + \sigma[t, \hat{z}_1(t), \mu_t] dw_1(t),$$

with initial condition $\hat{z}_1(0) = z_1(0)$, where $\mu_{(\cdot)}$ is the minor agents' measure given by the MV SDE system above. Again, Theorem 6.1 and the Gronwall's lemma yield

$$\sup_{0 \leq t \leq T} E|\hat{z}_1^N(t) - \hat{z}_1(t)| = O(1/\sqrt{N}). \quad (6.92)$$

Using (6.85) and (6.91)-(6.92), and by the same argument as in (6.88)-(6.90) one can show that $J_1^N(u_1^o; u_{-1}^o) - O(\epsilon_N + 1/\sqrt{N}) \leq \inf_{u \in \mathcal{U}_1} J_1^N(u_1; u_{-1}^o)$. \square

6.9. Chapter Summary

This chapter studies a stochastic mean field (SMF) system for a class of dynamic games involving nonlinear stochastic dynamical systems with major and minor (MM) agents. The SMF system consists of coupled (i) backward in time stochastic Hamilton-Jacobi-Bellman (SHJB) equations, and (ii) forward in time stochastic McKean-Vlasov (SMV) or stochastic Fokker-Planck-Kolmogorov (SFPK) equations. Existence and uniqueness of the solution to the MM SMF system is established by a fixed point argument in the Wasserstein space of random probability measures. In the case that minor agents are coupled to the major agent only through their cost functions, the ϵ_N -Nash equilibrium property of the SMF best response control possess is shown for a finite N population system where $\epsilon_N = O(1/\sqrt{N})$. As a particular but important case, the results of Nguyen and Huang [124] for MM-SMF linear-quadratic-Gaussian (LQG) systems with homogeneous population are retrieved, and, in addition, the results of this chapter are illustrated with a major and minor agent version of a game model of the synchronization of coupled nonlinear oscillators.

CHAPTER 7

Conclusion and Future Research

This thesis studied Mean Field Game (MFG) theory with applications to consensus, flocking, leader-follower and major-minor agent systems. The MFG methodology addresses a class of dynamic games with a large number of minor agents in which each agent interacts with the average or so-called mean field effect of other agents via couplings in their individual dynamics and cost functions. A minor agent is an agent which, asymptotically as the population size goes to infinity, has a negligible influence on the overall system while the overall populations effect on it is significant.

The thesis is presented in three main parts. The first part is concerned with applications of the MFG methodology to large population consensus and flocking behaviour. The second part is focused on the extension of the mean field linear-quadratic-Gaussian (MF LQG) framework so as to model the collective system dynamics which include large population of leaders and followers, and an unknown (to the followers) reference trajectory for the leaders. The final part investigates dynamic games with nonlinear stochastic dynamical systems of controlled McKean-Vlasov type involving agents of the following mixed types: (i) a major agent, and (ii) a large population of minor agents, where the major agent has a significant influence on minor agents while each minor agent has a negligible impact on other agents.

We now conclude by outlining some possible future research directions.

Applications of the Major and Minor MFG Theory in Finance. One of the fascinating features of backward stochastic differential equations (BSDEs) is their applications in finance with deep interpretations (see [56]). Therefore, it is of significant interest to apply the Major and Minor MFG theory to so-called “option pricing problems” with dynamic games formulations (see Section 6 in [180]).

A Major and Minor MF Nonlinear Markov Systems Theory. In Chapter 6 the MF LQG model for major and minor agents is extended to nonlinear stochastic dynamical systems of controlled McKean-Vlasov type. A more general case of major and minor MFG theory for nonlinear Markov systems (in the sense of [97]) surely merits study.

A MFG Theory for Partially Observed Stochastic Systems. In all the MFG models so far it is assumed that the controller of each individual agent is able to completely observe its own state. However, in many situations, the state of each agent can only partially be observed to itself via other variables plus some noise (see e.g., [18, 180]). Therefore, a MFG theory for partially observed stochastic systems based on Duncan-Mortensen-Zakai filter [22] is an open and challenging area (see [77] for the LQG case).

A Hybrid MFG Theory. The hybrid systems (see e.g., [23, 159]) are systems with both continuous and discrete states and corresponding continuous and discrete dynamics. These systems arise in a wide range of areas including communication networks, traffic control and queueing models whose fluid model limits involve differential equations of differential inclusion type. Since such systems may often have large populations of agents or messages it is of interest to generalize the MFG idea to systems of agents with hybrid differential inclusion dynamics.

REFERENCES

- [1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover: New York, NY, 1964.
- [2] D. Acemoglu and M. K. Jensen. Robust comparative statics in large static games. In *Proc. 49th IEEE CDC*, pages 3133–3139, Atlanta, Dec., 2010.
- [3] Y. Achdou, F. Camilli, and I. Capuzzo-Dolcetta. Mean field games: Numerical methods for the planning problem. *SIAM J. Control Optim.*, 50(1):77–109, 2012.
- [4] Y. Achdou and I. Capuzzo-Dolcetta. Mean field games: Numerical methods. *SIAM J. Numer. Anal.*, 48(3):1136–1162, 2010.
- [5] S. Adlakha, R. Johari, G. Weintraub, and A. Goldsmith. Oblivious equilibrium for large-scale stochastic games with unbounded costs. In *Proc. 48th IEEE CDC*, pages 5531–5538, Shanghai, China, Dec. 2008.
- [6] S. Adlakha, R. Johari, and G. Y. Weintraub. Equilibria of dynamic games with many players: existence, approximation, and market structure. Working paper, Stanford University, Stanford, CA, Dec. 2011.
- [7] T. Alpcan, T. Basar, R. Srikant, and E. Altman. CDMA uplink power control as a noncooperative game. *Wireless Networks*, 8(6):659–670, 2002.
- [8] E. Altman and Z. Altman. S-modular games and power control in wireless networks. *IEEE Trans. Autom. Control*, 48(5):839–842, May 2003.

REFERENCES

- [9] E. Altman, T. Boulogne, R. El-Azouzi, T. Jiménez, and L. Wynter. A survey on networking games in telecommunications. *Computers & Operations Research*, 33(2):286 – 311, 2006.
- [10] D. Andersson and B. Djehiche. A maximum principle for sdes of mean-field type. *Appl. Math. Optim.*, 63(3):341–356, 2011.
- [11] A. Arapostathis, V. S. Borkar, and M. K. Ghosh. *Ergodic control of diffusion processes*, volume 143. Cambridge Univ. Press: Cambridge, UK, 2011.
- [12] F. Aubé and R. Shield. Modeling the effect of leadership on crowd flow dynamics. *Cellular Automata*, 3305:601–611, 2004.
- [13] R. J. Aumann. Markets with a continuum of traders. *Econometrica*, 32:39–50, 1964.
- [14] M. Bardi. Explicit solutions of some linear-quadratic mean field games. *Networks and Heterogeneous Media*, 7(2):243–261, 2012.
- [15] T. Başar and G. J. Olsder. *Dynamic noncooperative game theory*. SIAM: Philadelphia, PA, 1999.
- [16] C. T. Bauch and D. J. D. Earn. Vaccination and the theory of games. *Proc. Natl. Acad. Sci. USA*, 101(36):13391–13394, 2004.
- [17] D. Bauso, L. Giarré, and R. Pesenti. Non-linear protocols for optimal distributed consensus in networks of dynamic agents. *Systems Control Lett.*, 55(11):918–928, 2006.
- [18] A. Bensoussan. *Stochastic Control of Partially Observable Systems*. Cambridge Univ. Press: Cambridge, UK, 1992.
- [19] P. Billingsley. Convergence of probability measures. *John Wiley and Sons: New York, NY*, 1999.
- [20] J. M. Bismut. Linear quadratic optimal stochastic control with random coefficients. *SIAM J. Control Optim.*, 14(3):419–444, 1976.

- [21] R. Boel and P. Varaiya. Optimal control of jump processes. *SIAM J. Control Optim.*, 15(1):92–119, 1977.
- [22] V. S. Borkar. Ergodic control of diffusion processes. pages 1299–1310, Madrid, Spain, 2006.
- [23] M. S. Branicky, V. S. Borkar, and S. K. Mitter. A unified framework for hybrid control: Model and optimal control theory. *IEEE Trans. Autom. Control*, 43(1):31–45, Jan. 1998.
- [24] R. Breban, R. Vardavas, and S. Blower. Mean-field analysis of an inductive reasoning game: Application to influenza vaccination. *Phys. Rev. E*, 76(3):031127, 2007.
- [25] C. M. Breder. Equations descriptive of fish schools and other animal aggregations. *Ecology*, 35(3):361–370, 1954.
- [26] W. A. Brock and S. N. Durlauf. Discrete choice with social interactions. *Rev. Economic studies*, 68(2):235–260, 2001.
- [27] W. A. Brock and A. Haurie. On existence of overtaking optimal trajectories over an infinite time horizon. *Math. Oper. Res.*, 1(4):337–346, 1976.
- [28] R. Buckdahn, P. Cardaliaguet, and M. Quincampoix. Some recent aspects of differential game theory. *Dynamic Games and Applications*, 1(1):74–114, 2011.
- [29] R. Buckdahn, B. Djehiche, J. Li, and S. Peng. Mean-field backward stochastic differential equations: A limit approach. *The Annals of Prob.*, 37(4):1524–1565, 2009.
- [30] R. Buckdahn, J. Li, and S. Peng. Mean-field backward stochastic differential equations and related partial differential equations. *Stoch. Proc. & their App.*, 119(10):3133–3154, 2009.

REFERENCES

- [31] P. E. Caines. Bode lecture: mean field stochastic control. in *48th IEEE CDC*, Shanghai, China, Dec. 2009 (<http://www.ieeecss-011.org/lectures/2009/mean-field-stochastic-control>).
- [32] P. E. Caines. A note on the consistency of maximum likelihood estimates for finite families of stochastic processes. *The Annals of Stat.*, 3(2):539–546, 1975.
- [33] P. E. Caines. *Linear Stochastic Systems*. John Wiley and Sons: New York, NY, 1987.
- [34] Y. Cao and W. Ren. Optimal linear-consensus algorithms: an LQR perspective. *IEEE Trans. Syst., Man, Cybern., Part B: Cybern.*, 40(3):819–830, 2010.
- [35] Y. U. Cao, A. S. Fukunaga, and A. Kahng. Cooperative mobile robotics: Antecedents and directions. *Autonomous Robots*, 4(1):7–27, 1997.
- [36] P. Cardaliaguet. *Notes on mean field games*. Université Paris-Dauphine, Cermade, France, Jan. 2012.
- [37] P. Cardaliaguet, J.-M. Lasry, P.-L. Lions, and A. Porretta. Long time average of mean field games. *Networks and Heterogeneous Media*, 7(2):279–301, Jun. 2012.
- [38] G. Carmona. Nash equilibria of games with a continuum of players. Technical report, EconWPA, 2004.
- [39] J. A. Carrillo, M. Fornasier, J. Rosado, and G. Toscani. Asymptotic flocking dynamics for the kinetic Cucker–Smale model. *SIAM J. Math. Anal.*, 42(1):218–236, 2010.
- [40] J. A. Carrillo, M. Fornasier, G. Toscani, and F. Vecil. Particle, kinetic, and hydrodynamic models of swarming. *Math. Model. Collec. Beh. in Socio-Econ. and Life Sci.*, G. Naldi, L. Pareschi and G. Toscani Eds., pp. 297–336, Springer: Berlin, 2010.

- [41] C. Y. Chong and S. P. Kumar. Sensor networks: Evolution, opportunities, and challenges. *Proc. IEEE*, 91(8):1247–1256, 2003.
- [42] K. L. Chung. *A Course in Probability Theory*. Harcourt, Brace and World Inc: San Diego, CA, 1968.
- [43] I. D. Couzin, J. Krause, N. R. Franks, and S. A. Levin. Effective leadership and decision-making in animal groups on the move. *Nature*, 433(7025):513–516, Feb. 2005.
- [44] F. Cucker and J. Dong. Avoiding collisions in flocks. *IEEE Trans. Autom. Control*, 55(5):1238–1243, May 2010.
- [45] F. Cucker and E. Mordecki. Flocking in noisy environments. *J. Math. & App.*, 89(3):278–296, 2008.
- [46] F. Cucker and S. Smale. Emergent behavior in flocks. *IEEE Trans. on Autom. Control*, 52(5):852–862, May 2007.
- [47] M. H. DeGroot. Reaching a consensus. *J. of the American Stat. Asso.*, 69(345):118–121, 1974.
- [48] D. Delchamps. Analytic feedback control and the algebraic Riccati equation. *IEEE Trans. Autom. Control*, 29(11):1031–1033, Nov. 1984.
- [49] C. Dogbé. Modeling crowd dynamics by the mean-field limit approach. *Math. & Comput. Model.*, 52(9-10):1506–1520, 2010.
- [50] R. Duan, M. Fornasier, and G. Toscani. A kinetic flocking model with diffusion. *Communications in Math. Physics*, 300(1):95–145, 2010.
- [51] T. Duncan. Evaluation of likelihood functions. *Inf. and Cont.*, 13(1):62–74, Jul. 1968.
- [52] T. Duncan. Likelihood functions for stochastic signals in white noise. *Inf. and Cont.*, 16(4):303–310, 1970.

REFERENCES

- [53] B. Düring, P. Markowich, J. F. Pietschmann, and M. T. Wolfram. Boltzmann and Fokker–Planck equations modelling opinion formation in the presence of strong leaders. *in Proc. Royal Soc. A*, 465(2112):3687–3707, 2009.
- [54] J. R. G. Dyer, A. Johansson, D. Helbing, I. D. Couzin, and J. Krause. Leadership, consensus decision making and collective behaviour in humans. *Phil. Tran. Royal Soc. B: Bio. Sci.*, 364(1518):781–789, 2009.
- [55] N. Egglezos. *Aspects of Utility Maximization With Habit Formation: Dynamic Programming And Stochastic PDE’s*. PhD thesis, Columbia University, 2007.
- [56] N. El Karoui, S. Peng, and M. C. Quenez. Backward stochastic differential equations in finance. *Mathematical finance*, 7(1):1–71, 1997.
- [57] G. M. Erickson. Differential game models of advertising competition. *European J. Oper. Res.*, 83(3):431–438, 1995.
- [58] J. A. Fax and R. M. Murray. Information flow and cooperative control of vehicle formations. *IEEE Trans. Autom. Control*, 49(9):1465–1476, Sep. 2004.
- [59] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*. Springer-Verlag: New York, NY, 2006.
- [60] F. Garin and L. Schenato. A survey on distributed estimation and control applications using linear consensus algorithms. *Netw. Cont. Sys.*, A. Bemporad, M. Heemels, and M. Johansson Eds., pp. 75-107, Springer: Berlin, 2011.
- [61] D. A. Gomes, J. Mohr, and R. R. Souza. Discrete time, finite state space mean field games. *J. Math. Pures & App.*, 93(3):308–328, 2010.
- [62] G. C. Goodwin and K. S. Sin. *Adaptive Filtering Prediction and Control*. Prentice-Hall: Englewood Cliffs, NJ, 1984.
- [63] D. Gu and Z. Wang. Leader-follower flocking: Algorithms and experiments. *IEEE Trans. Control Syst. Technol.*, 17(5):1211–1219, 2009.

- [64] O. Guéant. A reference case for mean field games models. *J. Math. Pures & App.*, 92(3):276–294, 2009.
- [65] O. Guéant, J.-M. Lasry, and P.-L. Lions. Mean field games and oil production. *Preprint*, 2010.
- [66] O. Guéant, J.-M. Lasry, and P.-L. Lions. Mean field games and applications. *Paris-Princeton Lectures on Mathematical Finance*, pages 205–266, 2011.
- [67] S.-Y. Ha, K. Lee, and D. Levy. Emergence of time-asymptotic flocking in a stochastic Cucker-Smale system. *Communications in Math. Physics*, 7(2):453–469, 2009.
- [68] S.-Y. Ha and J. G. Liu. A simple proof of the Cucker-Smale flocking dynamics and mean-field limit. *Communications in Mathematical Sciences*, 7(2), 2009.
- [69] S.-Y. Ha and E. Tadmor. From particle to kinetic and hydrodynamic descriptions of flocking. *Kinetic and Related Models*, 1(3):415–435, 2008.
- [70] O. Haimanko. Nonsymmetric values of nonatomic and mixed games. *Math. Oper. Res.*, 25:591–605, 2000.
- [71] S. Hart. Values of mixed games. *Int. J. Game Theory*, 2(1):69–85, 1973.
- [72] Y. Hu, J. Ma, and J. Yong. On semi-linear degenerate backward stochastic partial differential equations. *Probability Theory and Related Fields*, 123(3):381–411, 2002.
- [73] Y. Hu and S. Peng. Adapted solution of a backward semilinear stochastic evolution equation. *Stochastic Analysis and Applications*, 9(4):445–459, 1991.
- [74] M. Huang. *Stochastic control for distributed systems with applications to wireless communications*. PhD thesis, McGill University, Montreal, Canada, June 2003.
- [75] M. Huang. Large-population LQG games involving a major player: the Nash certainty equivalence principle. *SIAM J. Control Optim.*, 48(5):3318–3353, 2010.

REFERENCES

- [76] M. Huang, P. E. Caines, and R. P. Malhamé. Individual and mass behaviour in large population stochastic wireless power control problems: centralized and Nash equilibrium solutions. In *Proc. 42nd IEEE CDC*, pages 98–103, Maui, HI, Dec. 2003.
- [77] M. Huang, P. E. Caines, and R. P. Malhamé. Distributed multi-agent decision-making with partial observations: Asymptotic nash equilibria. In *Proc. the 17th Internat. Symposium on Math. Theory on Networks and Systems (MTNS'06), Kyoto, Japan*, pages 2725–2730, Kyoto, Japan, Jul. 2006.
- [78] M. Huang, P. E. Caines, and R. P. Malhamé. An invariance principle in large population stochastic dynamic games. *J. Sys. Sci. & Complex.*, 20(2):162–172, 2007.
- [79] M. Huang, P. E. Caines, and R. P. Malhamé. Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ϵ -Nash equilibria. *IEEE Trans. Autom. Control*, 52(9):1560–1571, Sep. 2007.
- [80] M. Huang, P. E. Caines, and R. P. Malhamé. The NCE (mean field) principle with locality dependent cost interactions. *IEEE Trans. Autom. Control*, 55(12):2799–2805, Dec. 2010.
- [81] M. Huang, P. E. Caines, and R. P. Malhamé. Social dynamics in mean field LQG control: Egoistic and altruistic agents. In *Proc. 49th IEEE CDC*, pages 3140–3145, Atlanta, GA, Dec. 2010.
- [82] M. Huang, P. E. Caines, and R. P. Malhamé. Social optima in mean field LQG control: Centralized and decentralized strategies. *IEEE Trans. Autom. Control*, 57(7):1736 – 1751, Jul. 2012.
- [83] M. Huang, P. E. Caines, and R. P. Malhamé. A locality generalization of the NCE (Mean Field) principle: Agent specific cost interactions. In *Proc. 47th IEEE CDC*, pages 5539–5544, Cancun, Mexico, Dec. 2008.

- [84] M. Huang, S. Dey, G. N. Nair, and J. H. Manton. Stochastic consensus over noisy networks with Markovian and arbitrary switches. *Automatica*, 46(10):1571–1583, 2010.
- [85] M. Huang, R. P. Malhamé, and P. E. Caines. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Inf. and Sys.*, 6(3):221 – 252, 2006.
- [86] M. Huang and J. H. Manton. Coordination and consensus of networked agents with noisy measurements: Stochastic algorithms and asymptotic behavior. *SIAM J. Control Optim.*, 48(1):134–161, 2009.
- [87] M. Huang and J. H. Manton. Stochastic consensus seeking with noisy and directed inter-agent communication: Fixed and randomly varying topologies. *IEEE Trans. Autom. Control*, 55(1):235–241, Jan. 2010.
- [88] A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Trans. Autom. Control*, 48(6):988–1001, Jun. 2003.
- [89] M. A. Khan and Y. Sun. Non-cooperative games with many players. *Handbook of Game Theory with Economic Appl.*, 3:1761–1808, 2002.
- [90] A. C. Kizilkale and P. E. Caines. Stochastic adaptive Nash certainty equivalence control: Population parameter distribution estimation. In *Proc. 49th IEEE CDC*, pages 6169–6176, Atlanta, GA, 2010.
- [91] A. C. Kizilkale and P. E. Caines. Mean field (NCE) stochastic control: Populations of major and egoist-altruist agents. In *Proc. 50th IEEE CDC-ECC*, pages 5547–5552, Orlando, FL, 2011.
- [92] A. C. Kizilkale and P. E. Caines. Mean field stochastic adaptive control. *IEEE Trans. Autom. Control*, 58(4), Apr. 2012.

REFERENCES

- [93] A. C. Kizilkale, S. Mannor, and P. E. Caines. Large scale real-time bidding in the smart grid: A mean field framework. To be presented at the *51st IEEE CDC*, Maui, HI, Dec. 2012.
- [94] J. K. Kok, C. J. Warmer, and I. G. Kamphuis. Powermatcher: Multiagent control in the electricity infrastructure. In *Proc. of the 4th Conference on AAMAS*, pages 75–82, Utrecht, Netherlands, Jul. 2005. ACM.
- [95] V. N. Kolokoltsov. *Nonlinear Markov Processes and Kinetic Equations*. Cambridge Univ. Press: Cambridge, UK, 2010.
- [96] V. N. Kolokoltsov. Nonlinear markov games on a finite state space (mean-field and binary interactions). *International J. Statistics and Probability*, 1(1):77–91, 2012.
- [97] V. N. Kolokoltsov, J. Li, and W. Yang. Mean field games and nonlinear markov processes. <http://arxiv.org/abs/1112.3744>, 2011.
- [98] L. B. Koralov and Y. G. Sinai. *Theory of probability and random processes*. Springer-Verlag: Heidelberg, Germany, 2007.
- [99] H. Kunita. Some extensions of Ito’s formula. *Séminaire de Probabilités XV 1979/80*, pages 118–141, 1981.
- [100] V. Lambson. Self-enforcing collusion in large dynamic markets. *J. Econ. Theory*, 34(2):282–291, Dec. 1984.
- [101] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. I - Le cas stationnaire. *C. Rendus Math.*, 343(9):619 – 625, 2006.
- [102] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. II Horizon fini et contrôle optimal. *C. Rendus Math.*, 343(10):679 – 684, 2006.
- [103] J.-M. Lasry and P.-L. Lions. Mean field games. *Jap. J. Math.*, 2(1):229–260, 2007.

- [104] A. Lazarte. Large population stochastic dynamic games with nonuniform large population stochastic dynamic games with nonuniform agents and localized coupling. Master's thesis, Dept. Elect. Comp. Eng., McGill Univ., Montreal, QC, Canada, 2009.
- [105] S. A. Levin. Complex adaptive systems: Exploring the known, the unknown and the unknowable. *Bulletin-American Mathematical Society*, 40(1):3–20, 2003.
- [106] T. Li and J.-F. Zhang. Asymptotically optimal decentralized control for large population stochastic multiagent systems. *IEEE Trans. Autom. Control*, 53(7):1643–1660, Aug. 2008.
- [107] T. Li and J.-F. Zhang. Consensus conditions of multi-agent systems with time-varying topologies and stochastic communication noises. *IEEE Trans. Autom. Control*, 55(9):2043–2057, Sep. 2010.
- [108] Z. Lin, M. Broucke, and B. Francis. Local control strategies for groups of mobile autonomous agents. *IEEE Trans. Autom. Control*, 49(4):622–629, Apr. 2004.
- [109] H. T. Liu, J. Shan, and D. Sun. Adaptive synchronization control of multiple spacecraft formation flying. *J. Dyn. Sys., Meas., and Cont.*, 129:337, 2007.
- [110] Y. Liu and K. M. Passino. Stable social foraging swarms in a noisy environment. *IEEE Trans. Autom. Control*, 49(1):30–44, Jan. 2004.
- [111] J. Lorenz. Continuous opinion dynamics under bounded confidence: A survey. *Int. J. Modern Phy. C*, 18(12):1819–1838, 2007.
- [112] T. Lux and M. Marchesi. Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*, 397(6719):498–500, 1999.
- [113] J. Ma and J. Yong. Adapted solution of a degenerate backward SPDE with applications. *Stoch. Proc. & their App.*, 70(1):59–84, 1997.

REFERENCES

- [114] J. Ma and J. Yong. *Forward-backward stochastic differential equations and their applications*. Number 1702. Springer-Verlag: Berlin, Germany, 1999.
- [115] J. Ma and J. Yong. On linear, degenerate backward stochastic partial differential equations. *Probability Theory and Related Fields*, 113(2):135–170, 1999.
- [116] Z. Ma, D. Callaway, and I. Hiskens. Decentralized charging control for large populations of plug-in electric vehicles: Application of the Nash certainty equivalence principle. In *Proc. IEEE CCA*, pages 191–195, Yokohama, Japan, Sep. 2010.
- [117] Z. Ma, D. Callaway, and I. Hiskens. Optimal charging control for plug-in electric vehicles. *Control and Optimization Methods for Electric Smart Grids*, 3(3):259–273, 2012.
- [118] Z. Ma, D. S. Callaway, and I. A. Hiskens. Decentralized charging control of large populations of plug-in electric vehicles. *IEEE Trans. Control Syst. Technol.*, (99):1–12, 2011.
- [119] C. D. Meyer. *Matrix analysis and applied linear algebra*. SIAM: Philadelphia, PA, 2000.
- [120] L. Moreau. Stability of multiagent systems with time-dependent communication links. *IEEE Trans. Autom. Control*, 50(2):169–182, Feb. 2005.
- [121] S. Motsch and E. Tadmor. A new model for self-organized dynamics and its flocking behavior. *J Stat. Phys.*, 144(5):923–947, 02 2011.
- [122] M. Nagy, Z. Ákos, D. Biro, and T. Vicsek. Hierarchical group dynamics in pigeon flocks. *Nature*, 464(7290):890–893, 2010.
- [123] A. Neyman. Values of games with infinitely many players. *Handbook of Game Theory with Economic App.*, 3:2121–2167, 2002.

- [124] S. L. Nguyen and M. Huang. Linear-quadratic-Gaussian mixed games with continuum-parametrized minor players. *SIAM J. Control Optim.*, 50(5):2907–2937, 2012.
- [125] S. L. Nguyen and M. Huang. Mean field LQG games with mass behavior responsive to a major player. To be presented at the *51st IEEE CDC*, Maui, HI, Dec. 2012.
- [126] M. Nourian and P. E. Caines. ϵ -Nash mean field game theory for nonlinear stochastic games with major and minor agents. Under revision in *SIAM J. Control Optim.*, submitted: Aug. 2012 (available in ArXive: <http://arxiv.org/abs/1209.5684>).
- [127] M. Nourian and P. E. Caines. ϵ -Nash mean field theory for nonlinear stochastic dynamic games with major-minor agents. In the *20th MTNS*, Melbourne, Australia, July 2012.
- [128] M. Nourian and P. E. Caines. ϵ -Nash mean field theory for nonlinear stochastic games with mixed agents. To be presented at the *51st IEEE CDC*, Maui, HI, Dec. 2012.
- [129] M. Nourian, P. E. Caines, and R. P. Malhamé. A mean field game synthesis of initial mean consensus problems: A continuum approach for non-gaussian behaviour. Under revision in *IEEE Trans. Autom. Cont.* (submitted: Nov. 2012, revised: Sep. 2012).
- [130] M. Nourian, P. E. Caines, and R. P. Malhamé. Synthesis of Cucker-Smale type flocking via mean field stochastic control theory: Nash equilibria. In *Proc. 48th Allerton*, pages 814–819, Monticello, IL, Sep. 2010.
- [131] M. Nourian, P. E. Caines, and R. P. Malhamé. An evolution mean field equation system for initial mean consensus behaviour: A stability analysis. In *Proc. 50th IEEE CDC-ECC*, pages 5029–5034, Orlando, FL, Dec. 2011.

REFERENCES

- [132] M. Nourian, P. E. Caines, and R. P. Malhamé. Mean field analysis of controlled Cucker-Smale type flocking: Linear analysis and perturbation equations. In *Proc. 18th IFAC WC*, pages 4471–4476, Milan, Italy, Aug. 2011.
- [133] M. Nourian, P. E. Caines, and R. P. Malhamé. A solution to the initial mean consensus problem via a continuum based mean field control approach. In *Proc. 50th IEEE CDC-ECC*, pages 5708–5713, Orlando, FL, Dec. 2011.
- [134] M. Nourian, P. E. Caines, R. P. Malhamé, and M. Huang. Derivation of consensus algorithm dynamics from mean-field stochastic control NCE equations. In *1st IFAC NecSys Workshop*, pages 13–18, Venice, Italy, Sep. 2009.
- [135] M. Nourian, P. E. Caines, R. P. Malhamé, and M. Huang. Leader-follower Cucker-Smale type flocking synthesized via mean field stochastic control theory. In J. Angeles et. al., editor, *Brain, Body and Machine*, chapter 9, pages 283–298. Springer, Berlin, Germany, 2010.
- [136] M. Nourian, P. E. Caines, R. P. Malhamé, and M. Huang. A solution to the consensus problem via stochastic mean field control. In *2nd IFAC NecSys Workshop*, pages 323–328, Annecy, France, Sep. 2010.
- [137] M. Nourian, P. E. Caines, R. P. Malhamé, and M. Huang. Mean field LQG control in leader-follower stochastic multi-agent systems: Likelihood ratio based adaptation. *IEEE Trans. Autom. Control*, 57(11), Nov. 2012.
- [138] M. Nourian, P. E. Caines, R. P. Malhamé, and M. Huang. Nash, social and centralized solutions to consensus problems via mean field control theory. *IEEE Trans. Autom. Control*, 58(3), Mar. 2013.
- [139] M. Nourian, R. P. Malhamé, M. Huang, and P. E. Caines. Optimality of adaptation based mean field control laws in leader-follower stochastic collective dynamics. In *Proc. 49th IEEE CDC*, pages 2270–2275, Atlanta, GA, Dec. 2010.

- [140] M. Nourian, R. P. Malhamé, M. Huang, and P. E. Caines. Mean field (NCE) formulation of estimation based leader-follower collective dynamics. *Int. J. Robotics & Aut.*, 26(1):120–129, 2011.
- [141] R. Olfati-Saber, J. A. Fax, and R. M. Murray. Consensus and cooperation in networked multi-agent systems. *Proc. IEEE*, 95(1):215–233, 2007.
- [142] R. Olfati-Saber and R. M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Trans. Autom. Control*, 49(9):1520–1533, Sep. 2004.
- [143] E. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equation. *Systems Control Lett.*, 14(1):55–61, 1990.
- [144] J. Park, H. J. Kim, and S.-Y. Ha. Cucker-Smale flocking with inter-particle bonding forces. *IEEE Trans. Autom. Control*, 55(11):2617–2623, nov. 2010.
- [145] S. Peng. Stochastic Hamilton–Jacobi–Bellman equations. *SIAM J. Control Optim.*, 30(2):284–304, 1992.
- [146] S. Peng. Backward stochastic differential equations and applications to optimal control. *Appl. Math. Optim.*, 27(2):125–144, 1993.
- [147] L. Perea, G. Gómez, and P. Elosegui. Extension of the Cucker-Smale control law to space flight formations. *J. Guid., Control, and Dyn.*, 32(2):527–537, 2009.
- [148] M. Porfiri and D. J. Stilwell. Consensus seeking over random weighted directed graphs. *IEEE Trans. Autom. Control*, 52(9):1767–1773, Sep. 2007.
- [149] Z. Qu. *Cooperative Control of Dynamical Systems: Applications to Autonomous Vehicles*. Springer-Verlag: London, UK, 2009.
- [150] T. C. Reluga and S. Viscido. Simulated evolution of selfish herd behavior. *J. Theoretical Biology*, 234(2):213–225, 2005.
- [151] W. Ren. Multi-vehicle consensus with a time-varying reference state. *Systems Control Lett.*, 56(7-8):474–483, 2007.

REFERENCES

- [152] W. Ren and R. W. Beard. Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE Trans. Autom. Control*, 50(5):655–661, May 2005.
- [153] W. Ren and R. W. Beard. *Distributed Consensus in Multi-Vehicle Cooperative Control: Theory and Applications*. Springer-Verlag: London, UK, 2007.
- [154] W. Ren, R. W. Beard, and E. M. Atkins. A survey of consensus problems in multi-agent coordination. In *Proc. ACC*, pages 1859–1864, Portland, OR, Jun. 2005.
- [155] C. W. Reynolds. Flocks, herds, and schools: A distributed behavioral model. *Computer Graphics*, 21(4):25–34, 1987.
- [156] H. Schaub, S. R. Vadali, J. L. Junkins, and K. T. Alfriend. Spacecraft formation flying control using mean orbital spacecraft formation flying control using mean orbital elements. *J. Astronaut. Sci.*, 40(1):69–87, 2000.
- [157] L. Schenato and S. Zampieri. On rendezvous control with randomly switching communication graphs. *Networks and Heterogeneous Media*, 2(4):627–646, 2007.
- [158] E. Semsar and K. Khorasani. Optimal control and game theoretic approaches to cooperative control of a team of multi-vehicle unmanned systems. In *IEEE Int. Conf. in Net., Sens. & Cont.*, pages 628–633, London, UK, Jun. 2007.
- [159] M. S. Shaikh and P. E. Caines. On the hybrid optimal control problem: theory and algorithms. *IEEE Trans. Autom. Control*, 52(9):1587–1603, Sep. 2007.
- [160] J. Shen. Cucker-Smale flocking under hierarchical leadership. *SIAM J. Appl. Math.*, 68(8):694–719, 2007.
- [161] M. Simaan and J. B. Cruz. On the Stackelberg strategy in nonzero-sum games. *J. Opt. Theory & App.*, 11(5):533–555, 1973.
- [162] A. S. Sznitman. Topics in propagation of chaos. *Ecole d’Eté de Probabilités de Saint-Flour XIXó1989*, pages 165–251, 1991.

- [163] H. G. Tanner, A. Jadbabaie, and G. J. Pappas. Flocking in fixed and switching networks. *IEEE Trans. Autom. Control*, 52(5):863–868, May 2007.
- [164] H. Tembine, S. Lasaulce, and M. Jungers. Joint power control-allocation for green cognitive wireless networks using mean field theory. In *Proc. 5th IEEE Intl. Conf. CROWNCOM*, pages 1–5, Fornebu, Norway, Jun. 2010.
- [165] H. Tembine, J. Y. Le Boudec, R. El-Azouzi, and E. Altman. Mean field asymptotics of markov decision evolutionary games and teams. In *Int. Conference on Game Theory for Networks (GameNets' 09)*, pages 140–150, Istanbul, Turkey, May 2009.
- [166] H. Tembine, Q. Zhu, and T. Basar. Risk-sensitive mean-field stochastic differential games. In *Proc. 18th IFAC World Congress*, number 3222-3227, Milano, Italy, Aug. 2011.
- [167] C. M. Topaz and A. L. Bertozzi. Swarming patterns in a two-dimensional kinematic model for biological groups. *SIAM J. Appl. Math.*, 65(1):152–174, 2004.
- [168] J. Tsitsiklis, D. Bertsekas, and M. Athans. Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE Trans. Autom. Control*, 31(9):803–812, Sep. 1986.
- [169] T. Vicsek, A. Czirók, E. B. Jacob, I. Cohen, and O. Shochet. Novel type of phase transition in a system of self-driven particles. *Phys. Rev. Lett.*, 75(6):1226–1229, 1995.
- [170] J. Von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton Univ. Press: Princeton, NJ, 3th edition, 1953.
- [171] B. C. Wang and J.-F. Zhang. Distributed control of multi-agent systems with random parameters and a major agent. *Automatica*, 48(9):2093–2106, Sep. 2012.

REFERENCES

- [172] B. C. Wang and J.-F. Zhang. Mean field games for large-population multi-agent systems with Markov jump parameters. *SIAM J. Control Optim.*, 50(4):2308–2334, 2012.
- [173] W. Wang and J. E. Slotine. A theoretical study of different leader roles in networks. *IEEE Trans. Autom. Control*, 51(7):1156–1161, Jul. 2006.
- [174] G. Y. Weintraub, C. L. Benkard, and B. Van Roy. Oblivious equilibrium: A mean field approximation for large-scale dynamic games. In *Advances in Neural Inf. Proc. Sys.*, MIT Press, 2006.
- [175] G. Y. Weintraub, C. L. Benkard, and B. Van Roy. Markov perfect industry dynamics with many firms. *Econometrica*, 76(6):1375–1411, 2008.
- [176] T. Yang, P. G. Mehta, and S. P. Meyn. Feedback particle filter with mean-field coupling. In *Proc. 50th IEEE CDC-ECC*, pages 7909 –7916, Orlando, FL, Dec. 2011.
- [177] H. Yin, P. G. Mehta, S. P. Meyn, and U. V. Shanbhag. Synchronization of coupled oscillators is a game. *IEEE Trans. Autom. Control*, 57(4):920–935, Apr. 2012.
- [178] X. Yin, Z. Ma, and L. Dong. Decentralized loading coordinations for large-population plug-in electric vehicles and a few controllable bulk loads. In *Proc. 50th IEEE CDC*, pages 3092–3097, Orlando, FL, Dec. 2011.
- [179] J. Yong. A leader-follower stochastic linear quadratic differential game. *SIAM J. Control Optim.*, 41(4):1015–1041, 2002.
- [180] J. Yong and X. Y. Zhou. *Stochastic controls: Hamiltonian systems and HJB equations*. Springer-Verlag: New York, NY, 1999.

Document Log:

Manuscript Version 1 — 14 September 2012
Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$ - $\mathcal{L}\mathcal{A}\mathcal{T}\mathcal{E}\mathcal{X}$ — 14 September 2012

MOJTABA NOURIAN

CENTER FOR INTELLIGENT MACHINES, MCGILL UNIVERSITY, 3480 UNIVERSITY STREET,
MONTREAL (QUEBEC) H3A 2A7, CANADA

E-mail address: mnourian@cim.mcgill.ca

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$ - $\mathcal{L}\mathcal{A}\mathcal{T}\mathcal{E}\mathcal{X}$