FINITE STATE MEAN FIELD GAMES

Peiqi Wang

A DISSERTATION
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Operations Research and Financial Engineering
Adviser: Professor René Carmona

April 2019
Abstract

Mean field game is a powerful framework for studying the strategic interactions within a large population of rational agents. Although existing research has predominantly relied on diffusion models to depict agents’ states, numerous applications, such as epidemic control and botnet defense, can best be modeled by systems of particles in discrete state space. This thesis tackles finite state mean field games. In the first part of the thesis, we develop a probabilistic approach for finite state mean field games. Based on the weak formulation of optimal control, the approach accommodates the interactions through the players’ strategies and flexible information structures.

The second part of the thesis is devoted to finite state mean field games involving a player possessing dominating influence. Two different mechanisms are explored. We first study a form of Stackelberg games, in which the dominating player, referred to as principal, moves first and chooses its strategy which impacts the dynamics and objective functions of every remaining player, referred to as agent. Having observed the principal’s strategy, the agents reach a Nash equilibrium. We seek optimal strategies of the principal, whose objective function depends on the statistical distribution of the agents’ states in equilibrium. Using the weak formulation of finite state mean field games developed previously in the thesis, we transform the principal’s optimization problem into a McKean-Vlasov control problem, and provide a semi-explicit solution under the assumptions of linear transition rate, quadratic cost and risk-neutral utility.

In the second model, we assume that all players move simultaneously and we study Nash equilibria formed jointly by major and minor players. We introduce finite player games and derive mean field game formulation in the limit of infinitely many minor players. In this limit, we characterize the best responses of major and minor players via viscosity solutions of HJB equations, and we prove existence of Nash equilibria under reasonable assumptions. We also derive approximate Nash equilibria for the finite player game from the solution of the mean field game.
Acknowledgements

I am profoundly indebted to my thesis advisor René Carmona. I am deeply grateful that René believed in my potential and accepted me as his student. During the three years I worked with him, I have come to know René as a brilliant scholar with deep and broad knowledge, acute sense of intuition and infinite enthusiasm for inquisition and creation. René taught me a great deal of mathematics, probability and statistics, but the most invaluable qualities René has instilled upon me are perhaps the ability to identify a meaningful problem and the tenacity of cracking the problem even after a series of frustrations and failures. His passion, optimism and diligence towards research are extremely contagious, which was a great source of inspiration during the course of my thesis research, and will continue to inspire me to be a better researcher.

I would like to express my sincere gratitude to a group of great professors and researchers both inside and outside Princeton for the advice and support I received from them: Patrick Cheridito, Miklos Racz, Beatrice Acciaio, Julio Daniel Backhoff Veraguas, Gaoyue Guo, Ronnie Sircar, Mykhaylo Shkolnikov and Ludovic Tangpi. I would like to thank my thesis committee, Miklos, Mykhaylo and Ludovic for providing invaluable advices and comments regarding the material and the presentation of this thesis. It was a great pleasure working with Patrick during my first and second years at Princeton. He helped me transition from an undergraduate learner to an independent researcher and instructor, and I am grateful for his valuable time offered to me, for the inspiring discussion on research and for the friendly encouragement and support. My heartfelt appreciation goes to Beatrice, Julio and Gaoyue for their enthusiastic discussions with me and sharing enlightening and novel ideas regarding the theory of optimal transport. I would also like to thank them for their warm hospitality during my visit to LSE and Oxford.

I will miss the time I spent with my officemates and fellow PhD students, Mark Cerenza, Xinyuan Fang, Christy Vaughn Graves, Juan Sagredo, Zongjun Tan and
Xiuneng Zhu. I shared the same office with Juan and Xinyuan for two years and the
fact that we all work on seemingly unrelated topics did not prevent us from learning
from each other. As more senior members of graduate students in the department,
they were so generous in sharing their perspective and experience. Christy, Mark,
Xiuneng and Zongjun, on the other hand, are my academic sisters and brothers,
having the same passion for mathematics, probability and game theory as I do. I
thank them for the stimulating discussion, camaraderie and caring they provided.

I wish to thank the amazing group of staffs in the department, Michael Bino,
Connie Brown, Lauren Coleman, Kimberly Lupinacci, Melissa Thomas, Tara Zigler,
who worked so hard to take good care of me and provided generous help for research
and teaching duties. My special thanks go to Tabitha Mischler for her warm and
smiling friendship.

I would not have completed the thesis if it were not for the unconditional love
and support from my wonderful parents. They always believe in me, want the best
of me and stand ready to advice and comfort in time of need. So, thank you, Mom
and Dad, for being the most loving and supportive family one could every wish for.
To my parents, Hailan Liu and Shunxing Wang
# Contents

Abstract ................................................................. iii
Acknowledgements ...................................................... iv

1 Introduction .............................................................. 1
  1.1 Past Work on Mean Field Games ................................. 2
  1.2 Mean Field Games in Finite State Space ........................ 7
  1.3 Organization of the Thesis ....................................... 12
  1.4 Notations .......................................................... 14

2 Preliminaries on Finite State Mean Field Games ................. 15
  2.1 Controlled Continuous-Time Markov Chain ..................... 15
    2.1.1 Definition via Infinitesimal Generator .................. 18
    2.1.2 A Semimartingale Representation ....................... 21
  2.2 Optimal Control of Continuous-Time Markov Chains .......... 27
    2.2.1 Analytical Approach via the Dynamic Programming Principle 27
    2.2.2 Probabilistic Approach by BSDE .......................... 30
  2.3 Analytical Approach to Finite State Mean Field Games ....... 36
    2.3.1 Model Setup ............................................... 36
    2.3.2 Existence and Uniqueness of Nash Equilibria ............. 39
    2.3.3 Game With Finitely Many Players ........................ 44
    2.3.4 Approximate Nash Equilibria .............................. 49
3 Probabilistic Analysis of Finite State Mean Field Games

3.1 Weak Formulation of Finite State Mean Field Games .............. 58
  3.1.1 Mean Field Games with Interaction through States and Controls 58
  3.1.2 Representative Player’s Optimization Problem ............... 62
3.2 Existence of Nash Equilibria ......................................... 66
  3.2.1 Topology for the Mean Field ..................................... 67
  3.2.2 Mapping for the Fixed Point Argument .......................... 76
  3.2.3 Proving the Existence of Nash Equilibria .................... 79
3.3 Uniqueness of the Nash Equilibrium ................................. 89
3.4 Approximate Nash Equilibria for Games with Finite Many Players 92
3.5 Appendix: BSDE Driven by Multiple Independent Continuous-Time
  Markov Chains .......................................................... 100

4 Finite State Principal Agent Problem ................................. 108

4.1 Model Setup .............................................................. 112
  4.1.1 Agent’s State Process .............................................. 112
  4.1.2 Principal’s Contract and Agent’s Objective Function ........ 114
  4.1.3 Principal’s Objective Function .................................. 114
  4.1.4 Agents’ Mean Field Nash Equilibria ............................ 115
  4.1.5 Principal’s Optimal Contracting Problem ...................... 115
4.2 A McKean-Vlasov Control Problem ................................. 116
  4.2.1 Representation of Nash Equilibria Based on BSDEs .......... 116
  4.2.2 Principal’s Optimal Contracting Problem ...................... 120
4.3 Solving the Linear-Quadratic Model ................................. 121
  4.3.1 Model Setup ........................................................ 122
  4.3.2 Reduction to Optimal Control on Flows of Probability Measures 123
5 Finite State Mean Field Game with Major and Minor Players

5.1 Game Model with Finitely Many Players
5.2 Mean Field Game Formulation
5.3 Convergence of Large Finite Player Games
  5.3.1 Major Player’s Problem with Finitely Many Minor Players
  5.3.2 Deviating Minor Player’s Problem
5.4 Optimization Problem for Individual Players
5.5 Existence of Nash Equilibria
5.6 The Master Equation for Nash Equilibria
5.7 Propagation of Chaos and Approximate Nash Equilibria
  5.7.1 Back to the (N + 1)-Player Game
  5.7.2 Propagation of Chaos
  5.7.3 Approximate Nash Equilibria
5.8 Appendix
  5.8.1 Proof of Lemma 5.4.3
  5.8.2 Proof of Theorem 5.4.4
  5.8.3 Proof of Lemma 5.5.8
  5.8.4 Proof of Propositions 5.7.3 & 5.7.4

6 Conclusion

Bibliography
Chapter 1

Introduction

The effort to depict and analyze a large system of particles dates back at least to the era of ancient Greek philosophy, when Leucippus and Democritus, among other theorists, invented the notion of atom. They posited that the origin of everything is rooted in the interaction of indivisible bodies, as these atoms with simple intrinsic properties like size and shape collide with one another and interlock in an infinite void. Accordingly, properties of macroscopic bodies perceivable by man is the aggregation of these same atomic interactions. Over the course of many centuries that follows, philosophers and physicists would see these infinitesimal particles with their own eyes (invention of microscope), uncover the fundamental rules of interactions between two individual particles (classical and quantum mechanics) and gain empirical insights into the macroscopic effect of countless interacting particles (thermodynamics and universal gravity). However, how to reconcile the microscopic behavior of an individual particle with the macroscopic phenomenon of a large ensemble remained obscure.

It was not until the advent of statistical mechanics and mean field theory that the gap could finally be bridged. The groundbreaking idea is to consider the limit scenario of infinitely many identical particles and study the statistical distribution of their
states. The aggregation of the effects exerted on any given particle by all the others can be approximated by a single averaged effect as a function of the distribution. In turn, the evolution of the distribution over time can be deduced from the dynamics governing an individual particle’s state. The potential of this innovative paradigm quickly emerged beyond the realm of physics, as people soon realized the possibility to deduce the macroscopic behavior of any large system of particles, provided that the microscopic behaviors of individual particle is known. Indeed, by the term particle, we might refer to substantial particles such as atoms or molecules, but we might also refer to certain particles of more abstract forms: a bird in a migrating flock, a trader in the stock market, a user of a social network or even a neuron in the human brain. However, unlike substantial particles among which the interactions obey fundamental principles of classical or quantum mechanics, there is no single canonical method to depict the dynamics for these abstract particles.

Fortunately, game theory in modern mathematics and economics provides the missing piece of the puzzle by endowing these imaginary particles with rationality. Each particle is assumed to optimize an objective, which depends on its own action and state as well as the actions and states of all the other particles. Consequently, the dynamics of the particles are endogenously determined by each particle’s strategic choice on the one hand and a certain type of equilibrium formed among the particles on the other hand. With the rules of mechanics in physics replaced by the rules of equilibrium in game theory, classical statistical mechanics readily evolved into the theory of Mean Field Games as we know today.

1.1 Past Work on Mean Field Games

The theory of Mean Field Games was initiated by Lasry and Lions [2007], in an attempt to analyze the seemingly intractable behavior of a large population of rational
agents with mutual impact in the context of stochastic differential games. At about the same time, the same notion of mean field game was independently introduced by Huang et al. [2006] in the community of electrical engineering under the name of Nash Certainty Equivalence. Since then, the research on mean field game have attracted interests from multiple disciplines and rapidly advanced on both theoretical and applicational fronts.

Tremendous efforts have been made on building the theoretical foundation of mean field games. Among them, a voluminous body of the literature is concerned with the fundamental question of existence and uniqueness of the Nash equilibrium. Named after the Nobel laureate John Forbes Nash Jr. who invented it, Nash equilibrium is perhaps one of the most fundamental notion of equilibrium in non-cooperative games. It describes a situation where no player can benefit by deviating from its current strategy while the others keep theirs unchanged. In the trailblazing work Nash [1950], the authors showed that any game with a finite number of players in which each player can choose from a finite set of pure strategies has at least one mixed-strategy Nash equilibrium.

Starting from the pioneering work of Lasry and Lions [2007] and Huang et al. [2006], these questions are first investigated via analytical methods, where the Nash equilibrium is characterized by a system of partial differential equations (PDEs) composed of a Hamilton-Jacobi-Bellman (HJB) equation coupled with a Fokker-Planck equation for the dynamics of the mean field. Weak solutions to this system of PDEs were studied in Lasry and Lions [2006b] with finite-horizon setup and in Lasry and Lions [2006a] with stationary setup, whereas the ergodic setup was considered in Cardaliaguet et al. [2012] as the limit of long-time average of mean field game. Later Gomes et al. [2015], Gomes et al. [2016] and Gomes and Pimentel [2015] investigated the existence and regularity of classical solutions under a variety of assumptions on the nonlinearity. Inspired by Lions [2007], Porretta [2014] introduced the optimal
planning problem by modifying the boundary condition of the mean field game PDEs and Achdou et al. [2012] studied the related numerical problem through the penalization method. An explicitly solvable instance, namely the linear-quadratic mean field game, was examined in Bensoussan et al. [2016b].

The probabilistic framework of mean field games was initially developed in Carmona and Delarue [2013], where the authors used the Pontryagin maximum principle to treat the player’s optimization problem. The Nash equilibrium can then be depicted by a system of McKean-Vlasov forward-backward stochastic differential equations (FBSDEs), for which topological arguments on spaces of measures were used to show the existence of solutions. A more general form of McKean-Vlasov FBSDE was closely studied in Carmona et al. [2015], as it can also be applied to the optimal control problem of McKean-Vlasov type. The connection between mean field games and McKean-Vlasov control problems was further explored in Carmona et al. [2013a]. A different probabilistic approach based on the weak formulation of optimal control was proposed in Carmona and Lacker [2015] and was probed in a more general setup in Lacker [2015].

Recently there has been renewed interest in the master equation of the mean field game. First introduced in Lions [2007] as a PDE involving the space of probability measures, the master equation emerges as the limit of the coupled system of HJB equations which describes the Nash equilibrium of a symmetric $N$-player stochastic differential game. In Cardaliaguet et al. [2015], the authors revealed the connection between the master equation and the mean field game PDEs, and the existence of a classical solution to the master equation was established. Relying on the classical solution of the master equation, the authors also showed the convergence of the Nash equilibrium of the $N$-player game to the Nash equilibrium of the mean field game as $N$ tends to infinity. The well-posedness of the master equation was also investigated in Chassagneux et al. [2014]. From the probabilistic prospective, Carmona and Delarue
[2014] showed that the master equation is the decoupling field of the McKean-Vlasov FBSDE which describes the Nash equilibrium of the mean field game. This interpretation was also exploited in Chassagneux et al. [2017] to develop a numerical scheme for the master equation. In Bensoussan et al. [2015], the master equation was interpreted as the value function of McKean-Vlasov control problems, and recently Pham and Wei [2017] studied the underlying dynamic programming principle which leads to the master equation.

Meanwhile, inspired by practical applications, contributions have also been made on a plethora of meaningful extensions. Carmona et al. [2016] studied the existence and uniqueness of Nash equilibrium in the presence of common noise by introducing the notion of weak Nash equilibrium of mean field games. Under the assumptions of linear dynamics, quadratic running cost and convex terminal cost, Ahuja [2016] showed a unique Nash equilibrium exists in the strong sense. An explicit solution to a mean field game with common noise was obtained in Carmona et al. [2013b]. The effect of a dominant player’s presence in the mean field game was first studied in Huang [2010] in the case of linear quadratic model via the PDE approach. Later, Carmona and Zhu [2016] performed probabilistic analysis on mean field games with major and minor players. The authors highlighted the fact that even in the limit of infinite number of minor players, the major player’s influence does not average out and behaves as a common noise. They further illustrated this by establishing the result of conditional propagation of chaos for the minor players. In Carmona and Wang [2017], a different scheme of fixed point based on strategies was proposed to solve mean field games with major and minor players. They applied the scheme to the linear quadratic model to derive semi-explicit solutions for both open-loop and Markov controls and illustrated the conditional propagation of chaos through a flocking model. Meanwhile, inspired by applications in bank run models, Carmona et al. [2017] examined mean field interaction in the context of an optimal stopping game. The authors showed
equilibria exist for pure stopping strategies under the assumption of complementarity on cost functionals, while equilibria exist for randomized stopping time under a more general setting. A different mean field game of stopping was formulated in Nutz [2016] where the author derived the explicit solution for equilibrium strategy.

While mean field games have been mostly formulated in continuous state spaces, some works considered the cases involving state processes in discrete space. Gomes et al. [2013] applied the analytical framework to study finite-state mean field games and established existence and uniqueness of the Nash equilibrium. Doncel et al. [2016] looked into a similar model with a finite set of actions. Later Cecchin and Fischer [2017] gave a probabilistic formulation based on stochastic differential equations (SDEs) driven by Poisson random measure and investigated existence of Nash equilibria with relaxed open-loop and feedback controls. Most recently, the master equation for finite state mean field game was investigated in Bayraktar and Cohen [2017], where the authors showed the regularity of the solution and established convergence of the value function of the $N$-player game towards the solution of the master equation.

All of the theoretical efforts mentioned above have been greatly inspired and invigorated by the ever-expanding applications of the mean field game theory in biology, engineering, economics and finance, among other fields. Early illustrations of mean field Nash equilibrium include the famous pedagogical models ‘Towels on Beach’ and ‘When do the meeting start?‘ in Lions [2007]. In the engineering community, mean field game proved to be a versatile framework in modeling power grid management Mériaux et al. [2013], Kizilkale and Malhamé [2013], Ma et al. [2012], cellular networks Aziz and Caines [2017], crowd congestion control Lachapelle and Wolfram [2011] and botnet defense Kolokoltsov and Bensoussan [2016]. Among the many applications of mean field game in economic theory, Lasry et al. [2008] studied the growth theory, Chan and Sircar [2015] examined price formation by Bertrand and Cournot equilibria,
Shen and Turinici [2012] models the trading volume of a given asset by traders’ heterogeneous beliefs and costly estimations, and Carmona et al. [2013b] looked into the systemic risk of banking system. Achdou et al. [2014] provided a number of examples of PDEs that naturally arise in macroeconomics theory and Lachapelle et al. [2010] tackled the numerical methods of mean field equilibria in economics. Meanwhile in mathematical biology, Nourian et al. [2011] used mean field game theory in providing an alternative approach to Cucker and Smale Model of flocking. Recently Carmona and Graves [2018] explained the mechanism of jet lag by a mean field game model on synchronization of circadian rhythm among neurons.

Finally, let us mention that interested readers can refer to Lions [2007] and Carmona and Delarue [2017] for an exhaustive presentation of the mean field game theory from the analytical perspective and probabilistic perspective respectively.

1.2 Mean Field Games in Finite State Space

Although existing applications of mean field games have been predominantly focused on diffusion-based model, there is a plethora of real life examples of interacting particle systems in which the states of the particles belong to a discrete space. Epidemic models are perhaps the most representative examples of this kind. From the modeling perspective, the mechanism of an epidemic is extremely generic and versatile to fit in a variety of scenarios such as cyber security in Kolokoltsov and Bensoussan [2016] and opinion formation in Stella et al. [2013]. In its most simplest form, each particle is either infected or not infected, and the dynamics of states are modeled as a Markov chain, of which the transition rate is assumed to be a function of the fraction of infected population. In a more realistic version of the model, a graph can be used to depict the location of each particle. Accordingly, the transition rate is assumed
to include the effect of only the neighboring particles, which allows to study how
infection spreads in the dimension of space.

Since we have just introduced the idea of a particle system on a graph, we would
like to mention the model of congestion studied in Guéant [2015]. The author con-
sidered a continuum of identical players distributed on a directed graph consisting
of $N$ nodes. Naturally, the state of each player at any given moment is the node
where the player is located. Each player’s movement on the graph is modeled as a
continuous-time Markov chain. Each player can strategically choose the transition
rate from one node to another wherever the passage is allowed according to the edge
of the graph, and it pays a cost to make such choice. Each player then tries to max-
imizes its earning which consisted of an instantaneous and a terminal reward derived
from the location of the player. In order to model the effect of the congestion, the
reward is assumed to be a decreasing function of the fraction of players located at a
given node. Under some structural assumptions on the reward functions, the author
shows a unique equilibrium of spatial distribution exists for this congestion model.

Another interesting example we would like to curate here is the model of informa-
tion percolation proposed in Duffie et al. [2009]. Let us imagine a large population of
agents who wish to uncover some information of common concern. Each agent is en-
dowed with a signal regarding the information, which is characterized by its precision
$N_t$, assumed to be an integer. For a certain period of time, agents randomly meet in
pairs to share their signal. When two agents meet, the precisions of both agents’ sig-
nals after the meeting becomes the sum of the precisions before the meeting. During
the process, each agent chooses a search effort and faces a tradeoff between reaching
the maximal precision at the end of the game and reducing the cost incurred by its
effort. The key part of the model consists of the matching mechanism of the agents.
The authors assumed that an agent exerting a search effort $\alpha_1$ has an intensity of
$\alpha_1\alpha_2q(\alpha_2)$ of being matched to some agent from the set of agents currently using
effort level \( \alpha_2 \) at time \( t \), where \( q(\alpha_2) \) is the proportion of the agents using effort \( \alpha_2 \). This can be interpreted as the complementarity of search and matching technology, which means that the more effort that an agent makes to be found, the more effective are the efforts of his counterpart to find him.

Although the model was not explicitly formulated as a mean field game with discrete state space in Duffie et al. [2009], we believe that it can be cast in such a framework. Indeed, we may consider the precision of the signal \( N_t \) as the agent’s state, which belongs to the discrete space of positive integers. When we assume that the agent’s search effort is a function of time and the current precision of its signal \( \alpha(t, n) \), it can be shown that \( N_t \) evolves as a continuous-time Markov chain. Indeed, according to the search and matching mechanism described above, an agent with current precision \( n \) has an intensity \( \alpha(t, n)\hat{\alpha}(t, m)\hat{\mu}(t, m) \) of being matched to another agent with current precision \( m \), given that \( \hat{\mu}(t, m) \) is the current fraction of agents with precision \( m \), and \( \hat{\alpha} \) is the search effort adopted by all the other agents. This means that the transition rate from state \( m \) to state \( m + n \) equals \( \alpha(t, n)\hat{\alpha}(t, m)\hat{\mu}(t, m) \). It is worth noticing that the agent feels the presence of the other agents through the effort-weighted measure \( \hat{\mu} \hat{\alpha}(n) := \hat{\alpha}(t, m)\hat{\mu}(t, m) \), instead of the distribution of the states \( \hat{\mu} \) itself. Fortunately, this is the only singularity of the model compared to other conventional mean field game models. Once the effort-weighted measure \( \hat{\mu} \hat{\alpha} \) is regarded as the mean field and fixed, the optimization problem of the agent is straightforward to solve. The Nash equilibrium can be obtained by matching \( \hat{\mu} \hat{\alpha} \) with the effort-weighted measure resulting from the optimal search effort.

Inspired by the myriad of existing and potential applications, this thesis explores several theoretical aspects of mean field games in finite state space. In the first part of the thesis, we propose a probabilistic framework to analyze finite state mean field game. Compared with the existing analytical approach in Gomes et al. [2013], this new approach allows us not only to incorporate the interaction through agents’
actions into the mean field game, but also to obtain equilibria among a broader class of strategies. In the second part of the thesis, we study the mean field game in which a player with dominating influence on the rest of the players is present. As the terminology we shall use repeatedly in the thesis, we call this dominating player the *major player*, while we refer to all the remaining players as the *minor players*. Our quest is motivated by a list of fundamental questions prevailing in the theory of mean field game:

1. **Formulation of games with finitely many players.** What are the different types of equilibria for a game involving a major player? What are the information structure and admissible strategies for each player? Accordingly what is the proper way to model the dynamics of the states and the interactions among the players?

2. **Passage to the mean field limit.** When the number of minor players goes to infinity, does the empirical distribution of the minor players’ states converge and in what sense does it converge? What about the convergence of the state processes of the players? How does the dominating player influence the dynamics of the entire system in the limit? Can we formulate a mean field game based on the limit?

3. **Equilibrium.** Does the equilibrium exist and is it unique? Does the finite nature of the state space facilitate the numerical computation of the equilibria? What is the master equation characterizing the equilibrium?

4. **Approximation.** How well does the equilibrium of the mean field limit approximate the equilibrium of the game with finitely many players?

When we investigate these problems in the context of the dynamic games in finite state space, we quickly realize that the technical tools developed for stochastic differential mean field games cannot be applied here. Let us mention some of the concerns specific to the games in finite state space. The first subtlety lies in the choice of proper information structure for the players. Recall that in the case of stochastic differential games, we might want to consider open-loop strategies (each player can
observe the past history of noise processes of all the other players), Markov strategies (each player can observe the current states of all the other players), or distributed strategies (each player can only observe its own noise process or state). Despite different information structures to be considered, the choice does not interfere with how we describe the state processes of the agents. Indeed, the admissible control corresponding to each of these information structures can be plugged in the controlled SDE and produce a well-defined state process. Unfortunately, it is not the case for the games in finite state space. As we shall see in Chapter 2, the most tractable way to describe the dynamics in a finite state space is to rely on continuous-time Markov chains and specify the transition rates between the states. This means that much of the time Markovian strategy will be the only viable choice granting a well-defined state process. Moreover, if we choose the Markovian information structure, we still need to decide what state variables should be made available to the players, which boils down to identifying the so-called sufficient statistics of the states of every players in the game.

The second difficulty lies in how we can model the independence of the state processes of different players. The assumption of independence regarding the noise processes of different players is essential in any mean field game model, since it guarantees the the convergence of the empirical distribution as the number of players tends to infinity. In diffusion-based mean field games, this requirement can be intuitively fulfilled by assigning independent Wiener processes to the dynamics of the players. However in the case of continuous-time Markov chains, it is more delicate to translate independence into the transition rates between the states.

Due to the delicacies mentioned above, it is not immediately clear what the correct formulation of mean field game with major and minor player should be. Therefore we begin by formulating the game with finite number of players, and then study how the states of each player evolve in the limit of infinity many players in the game.
More specifically, we need to show the limiting state processes for both the major and minor player are well-defined. Notice that this is very different from the more commonly used top-down procedure to study mean field games, in which we first formulate the mean field game based on the intuitions of the propagation of chaos and then reverse-engineer the formulation of the game with finite many players.

1.3 Organization of the Thesis

We start with some preliminary results on finite state mean field games in Chapter 2. Since the optimization problem is the centerpiece in analyzing the strategic behaviors in any type of game, in Section 2.2 we review the optimal control of continuous-time Markov chains based on the dynamic programming principal, and then introduce a new probabilistic approach based on backward stochastic differential equation (BSDE). When implemented in finite state mean field games, the first approach naturally leads to the analytical formulation developed in Gomes et al. [2013]. In Section 2.3, we will give a detailed review of some important results of this approach, including existence and uniqueness of the Nash equilibrium for the mean field game and the convergence of the $N$-player game to the mean field game. In particular, the way how we formulate the $N$-player game will be generalized to the game with major and minor players in Chapter 5.

In Chapter 3, we present the weak formulation of finite state mean field games. The probabilistic approach to the optimal control problem introduced in Chapter 2 serves as the backbone of this formulation, and its flexibility allows us to incorporate the interaction through the strategies of players into the game, which is otherwise intractable using the analytical approach. We give the definition of Nash equilibria in the weak formulation in Section 3.1 and study the existence and uniqueness of the equilibrium in Section 3.2 and Section 3.3. Then in Section 3.4, we formulate
the finite-player counterpart of the game. We show that when being assigned the distributed strategies derived from the Nash equilibrium of the mean field game, the players reach an approximative Nash equilibrium in the game with $N$ players.

In the next two chapters, we study the scenario where a strong influential player is present in the mean field game. We envisage two types of equilibria. In Chapter 4, we study a type of Stackelberg equilibrium formulated as the so-called principal agent problem. In term of the terminology, in order to be consistent with the existing literatures dealing with similar models, we refer to the player with dominating influence as the *principal*, while we call the remaining players as the *agent*. After introducing the model in Section 4.1, we show in Section 4.2 that the principal’s optimal contracting problem is equivalent to an optimal control problem of McKean-Vlasov SDE. This results relies on the weak formulation of finite state mean field games developed in Chapter 3. In Section 4.3, we treat the special case of the linear-quadratic model where we further simplify the optimal contracting problem into the deterministic control on the distribution of agents’ states. We illustrate the linear-quadratic model with an example of epidemic containment in Section 4.4.

In Chapter 5, we study an entirely different game model with major and minor players. Compared to the principal agent model studied in Chapter 4, we endow the major player with not only an objective function, but also a state process and we aim to study the Nash equilibrium formed by the major player and the minor players. We start by formulating the finite player game in Section 5.1, in which we pay special attention to the information structure of each player and the dynamics of the major and minor player’s state process. Then in Section 5.3 we rigorously derive the limit of the state processes of both major and minor players. Combined with a fixed-point scheme for mean field games with major and minor player based on the strategies of the players which we describe in Section 5.2, we formulate the mean field game. In Section 5.4, we deal with the optimization problem of the major and minor player and
derive their optimal response, respectively. Section 5.5 is devoted to the existence of Nash equilibria and we derive the master equation in Section 5.6. In Section 5.7 we investigate how the solution of the mean field game can provide an approximative Nash equilibrium for the finite player game.

We conclude the thesis in Chapter 6 by discussing a few unsolved problems emerging during our investigation of finite state mean field games, which, we hope, will be helpful in pointing out new directions for future research on this subject.

The thesis is built on material from four papers by the author and his thesis advisor, some of which are under review or revision at the time this thesis is written. Chapter 2 and Chapter 3 interpolate material from Carmona and Wang [2018b]. Meanwhile, Chapter 4 is based on Carmona and Wang [2018a]. Finally, Chapter 5 is based on Carmona and Wang [2016], and uses ideas from Carmona and Wang [2017].

1.4 Notations

We introduce some notations that will be used throughout the rest of the thesis. For a real matrix \( M \), we denote its transpose by \( M^\ast \), and we denote by \( M^+ \) the Moore-Penrose pseudo inverse of \( M \). The multiplication operator is denoted by \( \cdot \), and is always understood as matrix multiplications if the operands are vectors or matrices of proper dimensions. For a random variable \( \gamma \) in a probability space \((\Omega, \mathcal{F}, \mathbb{P})\), we denote by \( \mathbb{P}_{\gamma} := \mathbb{P} \circ \gamma^{-1} \) the push-forward measure of \( \mathbb{P} \) by \( \gamma \). We denote by \( \mathbb{I}(\cdot) \) the indicator function. In some of the proofs, we use \( C \) to denote a positive constant that might change its value from line to line.

Finally, considering that the players or agents in the game models we are about to present are sometimes abstract objects, and in order to avoid the confusion of the genders when referring to the players or agents, we decide to make the players genderless and use the pronouns it/its throughout this thesis.
Chapter 2

Preliminaries on Finite State Mean Field Games

In this chapter, we start with some basic facts about continuous-time Markov chain and present two different approaches to the related stochastic control problem: one based on dynamic programming principal and the other based on a probabilistic representation by backward stochastic differential equation (BSDE). The analytical approach naturally leads to the existing framework of finite state mean field game proposed by Gomes et al. [2013], for which we shall give a detailed account at the end of this chapter. The probabilistic approach, on the other hand, will provide us with the necessary tools to develop the weak formulation of the finite state mean field game in Chapter 3.

2.1 Controlled Continuous-Time Markov Chain

In preparation for finite state mean field games, in this section, we introduce the controlled continuous-time Markov chain. This is a stochastic process taking values in a discrete space, of which the dynamics, controlled by external actions, can be
characterized in different fashions depending on the admissible set of controls, as well as the techniques used to tackle the optimal control problem.

Throughout the thesis, we denote by $E$ the state space of the process and we assume that $E$ is finite. The elements in $E$ are denoted by $1, 2, \ldots, m$, however we sometimes identify these states with the $m$ basis vectors $e_1, e_2, \ldots, e_m$ of $\mathbb{R}^m$. We fix a finite time horizon $T$. Let us first recall the definition of continuous-time Markov chains.

**Definition 2.1.1.** A stochastic process $(X_t)_{t \in [0,T]}$ in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called a continuous-time Markov chain, if for any set of time $0 \leq t_0 < t_1 < \cdots < t_n \leq T$ and any set of states $i_0, i_1, \ldots, i_n \in E$, we have:

$$\mathbb{P}[X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}, \ldots, X_{t_0} = i_0] = \mathbb{P}[X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}].$$

For any continuous-time Markov chain $X$, we can associate it with a probability transition matrix $P(s, t)$. For $s, t \in [0, T]$ we define the component of $P(s, t)$ on $i$-th row and $j$-th column by $P_{ij}(s, t) := \mathbb{P}[X_t = j | X_s = i]$. Combined with the Markov property described in Definition 2.1.1, the probability transition matrix allows us to compute the marginal law of the stochastic process $X$. However, a better way to describe the law of a continuous-time Markov chain is to look at the derivative of the probability transition matrix, which is called the transition rate matrix. For $t \in [0, T]$, the element of the transition rate matrix $Q(t)$ is defined as:

$$Q_{ii}(t) := \lim_{h \to 0} \frac{1}{h} (\mathbb{P}[X_{t+h} = i | X_t = i] - 1) = \lim_{h \to 0} \frac{1}{h} (P_{ii}(t, t + h) - 1),$$

$$Q_{ij}(t) := \lim_{h \to 0} \frac{1}{h} \mathbb{P}[X_{t+h} = j | X_t = i] = \lim_{h \to 0} \frac{1}{h} P_{ij}(t, t + h).$$
It can be shown that the transition rate matrix $Q(t)$ of any continuous-time Markov chain satisfies several properties. We summarize these properties in the following definition, and we call any matrix satisfying these properties the Q-matrix.

**Definition 2.1.2.** We say a real-valued matrix $Q = [Q_{ij}]_{i,j \in E}$ is a Q-matrix if $Q_{ij} \geq 0$ for $i \neq j$ and

$$\sum_{j \neq i, j \in E} Q_{ij} = -Q_{ii}, \quad \text{for all } i \in E.$$  

Conversely, given a measurable mapping $[0, T] \ni t \rightarrow Q(t)$ such that $Q(t)$ is a Q-matrix for all $t \in [0, T]$, we can construct the corresponding probability transition matrix $P(s, t)$ satisfying (2.1) and (2.2). Moreover, we can construct a probability space and a stochastic process $X$ such that $X$ is a continuous-time Markov chain with $P(s, t)$ as its probability transition matrix.

**Theorem 2.1.3.** Let $\mu$ be a probability measure on $E$ and $[0, T] \ni t \rightarrow Q(t)$ be a measurable mapping such that $Q(t)$ is a Q-matrix for all $t \in [0, T]$ and the components $Q_{ij}(t)$ are bounded uniformly in $i, j \in E$ and $t \in [0, T]$. Let $p^0$ be a probability distribution on the state space $E$. Then there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a continuous-time Markov chain $X$ such that the distribution of $X_0$ is $p^0$ and we have for all $t \in [0, T]$:

$$Q_{ii}(t) = \lim_{h \to 0} \frac{1}{h} (\mathbb{P}[X_{t+h} = i | X_t = i] - 1),$$

$$Q_{ij}(t) = \lim_{h \to 0} \frac{1}{h} \mathbb{P}[X_{t+h} = j | X_t = i].$$

**Proof.** From Section 8.9.2 in Iosifescu [2014], we see that there exists a mapping $[0, T]^2 \ni (s, t) \rightarrow P(s, t)$ solving the forward and backward Kolmogorov equations:

$$P(s, t) = I_m + \int_s^t Q(u) \cdot P(u, t) du,$$

$$P(s, t) = I_m + \int_s^t P(s, u) \cdot Q(u) du,$$
where \( I_m \) is the identity matrix of size \( m \). In addition, for all \((s, t) \in [0, T]^2\), all the components of \( P(s, t) \) are positive and \( \sum_{j \in E} P_{ij}(s, t) = 1 \) for all \( i \in E \).

Now for each \( k \in \mathbb{N} \), and finite sequence of distinct times \( t_1, \ldots, t_k \in [0, T] \), we define \( \nu_{t_1, \ldots, t_k} \) the probability measure on \( E^k \) as follows:

\[
\nu_{t_1, \ldots, t_k}(F_1 \times \cdots \times F_k) := \sum_{i_0 \in E} \sum_{i_1 \in F_{\sigma(1)}} \cdots \sum_{i_k \in F_{\sigma(k)}} p_{i_0} P_{i_0 i_1}(0, t_{\sigma(1)}) P_{i_1 i_2}(t_{\sigma(1)}, t_{\sigma(2)}) \cdots P_{i_{k-1} i_k}(t_{\sigma(k-1)}, t_{\sigma(k)}).
\] (2.3)

Here \( \sigma \) is the permutation which puts \( t_1, \ldots, t_k \) in increasing order, i.e. \( t_{\sigma(1)} < \cdots < t_{\sigma(k)} \). It is plain to check that the family of measures \( \nu_{t_1, \ldots, t_k} \) is consistent. We can therefore apply Kolmogorov extension theorem, which allow us to construct a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) and a stochastic process \( X \) such that \( \nu_{t_1, \ldots, t_k}(F_1 \times \cdots \times F_k) = \mathbb{P}[X_{t_1} \in F_1, \ldots, X_{t_k} \in F_k] \) for all \( k \in \mathbb{N}, t_1, \ldots, t_k \in [0, T] \) and \( F_1, \ldots, F_k \subset E \).

Now from the definition of the measure \( \nu_{t_1, \ldots, t_k} \), we easily check that \( X \) is a continuous-time Markov chain with probability transition matrix \( P(\cdot, \cdot) \) (see Definition 2.1.1), and the distribution of \( X_0 \) is \( p^0 \). Finally, from the Kolmogorov equation we verify that \( Q \) is the transition rate function of \( X \).

From Theorem 2.1.3, we see that in order to describe the continuous-time Markov chain, we only need to specify its transition rate function and its initial distribution. This provides a convenient way to model controlled continuous-time Markov chain, since we only need to specify how the transition rate function can be controlled.

### 2.1.1 Definition via Infinitesimal Generator

Recall the case of controlled diffusion process, in which the dynamics of the process are completely determined by a drift function \( b : (t, x, \alpha) \to b(t, x, \alpha) \), a volatility function \( \sigma : (t, x, \alpha) \to \sigma(t, x, \alpha) \), and finally, a control process \( t \to \alpha_t \) which satisfies a certain information structure. This is a pathwise characterization of the dynamics,
as we consider the strong solution to the following stochastic differential equation (SDE):

\[
dY_t = b(t, Y_t, \alpha_t)dt + \sigma(t, Y_t, \alpha_t)dW_t.
\]

However, another interesting perspective is to regard \( b \) and \( \sigma \) as the local mean and standard deviation of the increment of the process, during an infinitesimally small interval of time. This amounts to defining the infinitesimal generator \( \mathcal{G} \) of a Markov process which characterizes how the probability distribution of the process - starting from a certain point in space-time - will evolve onward. Due to the Markovian nature of the process, we need to restrict the admissible control to be Markovian control, which is of the form \( \alpha_t = \phi(t, Y_t) \). Formally, we define \( \mathcal{G} \) as an operator on the space of functions:

\[
\mathcal{G}F(t, y) := \lim_{h \to 0} \frac{1}{h} \mathbb{E}[F(t+h, Y_{t+h}) - F(t, Y_t)|Y_t = y].
\]

Under suitable condition of regularity on \( F \) and by Ito’s formula we obtain:

\[
\mathcal{G}F(t, y) = \partial_t F(t, y) + b(t, y, \phi(t, y))\partial_y F(t, y) + \frac{1}{2} \sigma^2(t, y, \phi(t, y))\partial^2_{yy} F(t, y).
\]

When it comes to the controlled continuous-time Markov chain, it is still possible to define the dynamics of the process in a pathwise fashion by using the jump process. However, it is more intuitive to directly specify the infinitesimal generator of the process, which grants an interpretation in terms of the jump rate and the probability of transitions between the states. Let \( A \) be a compact subset of \( \mathbb{R}^l \) in which we pick the control. We introduce a function \( q \):

\[
[0, T] \times E^2 \times A \to \mathbb{R}
\]

\[
(t, i, j, \alpha) \to q(t, i, j, \alpha),
\]
and for later use, we denote by \( Q(t, \alpha) \) the matrix \([q(t, i, j, \alpha)]_{1 \leq i, j \leq m}\). Throughout this chapter, we make the following assumptions on \( q \):

**Assumption 2.1.4.** For all \((t, \alpha) \in [0, T] \times A\), the matrix \( Q(t, \alpha) \) is a Q-matrix.

**Assumption 2.1.5.** There exists \( C_1, C_2 > 0 \) such that for all \((t, i, j, \alpha) \in [0, T] \times E^2 \times A\) such that \( i \neq j \), we have \( 0 < C_1 < q(t, i, j, \alpha) < C_2 \).

We consider a Markovian control \( \alpha \) of the form \( \alpha_t = \phi(t, X_t) \), where \( \phi : [0, T] \times E \to A \) is a measurable mapping. \( \phi \) is called the feedback function associated with the Markovian control \( \alpha \). As we have seen in Theorem 2.1.3, we can define the dynamics of the process \( X \) by specifying the transition probability in an infinitesimal interval of time \( h \), or the transition rate function:

\[
\begin{align*}
\mathbb{P}[X_{t+h} = j | X_t = i, \{X_u, 0 \leq u < t\}] &= q(t, i, j, \phi(t, i)) h + o(h), \quad \text{if } j \neq i \quad (2.4) \\
\mathbb{P}[X_{t+h} = i | X_t = i, \{X_u, 0 \leq u < t\}] &= 1 - \sum_{j \neq i} q(t, i, j, \phi(t, i)) h + o(h) \quad (2.5)
\end{align*}
\]

This is equivalent to defining the following infinitesimal generator \( \mathcal{G}_\phi \), acting on the space of mappings \( F : [0, T] \times E \to \mathbb{R} \) which are \( C^1 \) in \( t \):

\[
\mathcal{G}_\phi F(t, i) := \partial_t F(t, i) + \sum_{j \neq i} (F(t, j) - F(t, i)) q(t, i, j, \phi(t, i)). \quad (2.6)
\]

A more intuitive way to interpret the dynamics of \( X \) described by \( \mathcal{G}_\phi \) is to consider the hazard rate of the jumps. More specifically, conditioned on the past path of \( X \) and given that \( X_t = i \), the event that \( X \) jumps from \( i \) to \( j \) has a hazard rate \( q(t, i, j, \alpha(t, i)) \). If we assume that \( q \) and \( \phi \) does not depend on \( t \), i.e. the dynamics are time-homogeneous, then conditioned on the past path of \( X \) and given that \( X_t = i \), the time until the next jump follows an exponential distribution of parameter \( \sum_{j \neq i} q(i, j, \alpha(i)) \) and the probability of jumping from \( i \) to \( j \) is given by \([q(i, j, \alpha(i))/\sum_{j \neq i} q(t, i, j, \alpha(i))].\)
2.1.2 A Semimartingale Representation

We now present an alternative formalism of finite state continuous-time Markov chain based on semimartingales, which was first introduced in Elliott et al. [1995] and later developed in Cohen and Elliott [2008] and Cohen and Elliott [2010]. We shall see that by an argument of change of measure, such a representation can accommodate not only the controlled continuous-time Markov chains, but also a larger class of controlled point processes.

Consider $X = (X_t)_{t \in [0,T]}$ a continuous-time Markov chain with $m$ states. We identify these states with the unit vector $e_i$ in $\mathbb{R}^m$ and therefore we denote $E := \{e_1, \ldots, e_m\}$. We assume that the sample path $t \to X_t$ is càdlàg, i.e., right continuous and admits left limit. In addition, we assume that $X$ is continuous at $T$.

We now construct a canonical probability space for $X$. Let $\Omega$ be the space of càdlàg functions from $[0,T]$ to $E$ which are continuous at $T$ and let $X$ be the canonical process. We denote by $\mathcal{F}_t := \sigma(\{X_s, s \leq t\})$ the natural filtration generated by $X$ and $\bar{\mathcal{F}} := \mathcal{F}_T$. Now let us fix a probability measure $p^\circ$ on $E$. On $(\Omega, (\mathcal{F}_t)_{t \in [0,T]}, \bar{\mathcal{F}})$ we consider the probability measure $\mathbb{P}$ under which $X$ is a continuous-time Markov chain of which the law of $X_0$ is $p^\circ$ and the transition rate between any two different states is $1$. This means that for $i, j \in \{1, \ldots, m\}$, $i \neq j$ and $\Delta t > 0$, we have $\mathbb{P}[X_{t+\Delta t} = e_j | \mathcal{F}_t] = \mathbb{P}[X_{t+\Delta t} = e_j | X_t] \text{ and } \mathbb{P}[X_{t+\Delta t} = e_j | X_t = e_i] = \Delta t + O(\Delta t)$. By Appendix B in Elliott et al. [1995], the process $X$ has the following representation:

$$X_t = X_0 + \int_{(0,t]} Q^0 \cdot X_{t-} dt + M_t.$$  (2.7)

Here and in the rest of this chapter, $Q^0$ is the $m$ by $m$ square matrix with diagonal elements all equal to $-(m-1)$ and off-diagonal elements all equal to $1$. $M$ is a $\mathbb{R}^m$-valued $\mathbb{P}$-martingale. The multiplication $\cdot$ is understood as matrix multiplication.
Indeed, $Q^0$ is the transition rate matrix of $X$ under the probability measure $\mathbb{P}$. We shall refer to the probability measure $\mathbb{P}$ as the reference measure of the sample space.

**Remark 2.1.6.** The representation originally proposed in Elliott et al. [1995] is:

$$X_t = X_0 + \int_{(0,t]} Q^0 \cdot X_t \, dt + M_t.$$  

However, since we know that almost surely $X_t$ is only discontinuous on a countable set, we can replace $X_t$ by $X_{t-}$ in the integral. The reason for this slight change of representation is to make the integrand a predictable process, which will be suitable for the change of measure argument in the following.

Let us recall the definitions of predictable quadratic (co)variation $\langle \cdot, \cdot \rangle$ and quadratic (co)variation $[\cdot, \cdot]$. See Chapter II.6 in Protter [2005].

**Definition 2.1.7.** Let $(\Omega, \bar{\mathcal{F}}, \mathbb{P})$ be a probability space supporting a filtration $\mathbb{F} := (\mathcal{F}_t)_{t \geq 0}$ and a square integrable $\mathbb{F}$-martingale $X$. Then by Doob-Meyer decomposition theorem, there exists a unique increasing predictable process $(A_t)_{t \geq 0}$ such that $A_0 = 0$ and the process $t \to (X_t^2 - A_t)$ is a martingale. The process $(A_t)_{t \geq 0}$ is called the predictable quadratic variation of $X$, which we denote by $\langle X, X \rangle_t := A_t$. In addition, if $X, Y$ are two square integrable $\mathbb{F}$-martingales, the predictable quadratic covaration of $X$ and $Y$, denoted by $(\langle X, Y \rangle_t)_{t \geq 0}$ is defined by:

$$\langle X, Y \rangle_t := \frac{1}{4}(\langle X + Y, X + Y \rangle_t - \langle X - Y, X - Y \rangle_t) \quad (2.8)$$

**Definition 2.1.8.** Let $X, Y$ be semimartingales. The quadratic variation of $X$, denoted by $([X, X]_t)_{t \geq 0}$, is defined by:

$$[X, X]_t := X_t^2 - 2 \int_0^t X_s \, dX_s \quad (2.9)$$
The quadratic covariation of $X$ and $Y$, denoted by $([X, Y]_t)_{t \geq 0}$, is defined by:

$$[X, Y]_t := \frac{1}{4}([X + Y, X + Y]_t - [X - Y, X - Y]_t)$$  \hfill (2.10)

The predictable quadratic variation of the martingale $M$, which appears in the canonical decomposition of the continuous-time Markov chain in equation (2.7), is given in the following lemma. The proof can be found in Cohen and Elliott [2008]:

**Lemma 2.1.9.** The predictable quadratic variation of the martingale $M$ under $\mathbb{P}$ is:

$$\langle M, M \rangle_t = \int_0^t \psi_s ds,$$  \hfill (2.11)

where the process $(\psi_t)_{t \in [0, T]}$ is given by:

$$\psi_t := \text{diag}(Q^0 \cdot X_t) - Q^0 \cdot \text{diag}(X_t) - \text{diag}(X_t) \cdot Q^0.$$  \hfill (2.12)

Let us define the matrix $\psi^i := \text{diag}(Q^0 e_i) - Q^0 \text{diag}(e_i) - \text{diag}(e_i) Q^0$. Clearly we have $\psi_t = \sum_{i=1}^m 1(X_t = e_i) \psi^i$. Since each $\psi^i$ is a semi-definite positive matrix, so is $\psi_t$.

We define the corresponding (stochastic) seminorm $\| \cdot \|_{X_t}$ on $\mathbb{R}^m$ by:

$$\|Z\|_{X_t}^2 := Z^* \cdot \psi_t \cdot Z.$$  \hfill (2.13)

The seminorm $\| \cdot \|_{X_t}$ can be rewritten in a more explicit way. For $i \in \{1, \ldots, m\}$, let us define the seminorm $\| \cdot \|_{e_i}$ on $\mathbb{R}^m$ by:

$$\|Z\|_{e_i}^2 := Z^* \cdot \psi^i \cdot Z = \sum_{j \neq i} |Z_j - Z_i|^2.$$  \hfill (2.14)

Then it is easy to verify that $\|Z\|_{X_t} = \sum_{i=1}^m 1(X_t = i) \|Z\|_{e_i}$. 

23
For later use, we also denote by $\psi^+_t$ the Moore-Penrose generalized inverse of the matrix $\psi_t$. Since $\psi_t$ is symmetric, we have $(\psi^+_t)^* = \psi^+_t$. On the other hand, it is straightforward to verify that for all $t \leq T$ and $\omega \in \Omega$, the image of the matrix $\psi_t$ (i.e. the linear space spanned by the columns of $\psi_t$) is the space $\{q \in \mathbb{R}^m, \sum q_i = 0\}$. Therefore for all $q \in \mathbb{R}^m$ with $\sum q_i = 0$, we have $\psi_t \cdot \psi^+_t \cdot q = q$. This holds in particular for any row vector from any $Q$-matrix, or any vectors of the form $e_j - e_i$.

Let us also recall the definition of the Doléans-Dade exponential.

**Definition 2.1.10.** For a semimartingale $L$ with $L_0 = 0$, the Doléans-Dade exponential, denoted as $\mathcal{E}(L)$, is the (unique) semimartingale $Z$ such that for all $t \geq 0$, we have $Z_t = 1 + \int_0^t Z_s - dX_s$.

To define the controlled probability measure for $X$, we shall use the following version of Girsanov Theorem regarding the change of probability measure. See Theorem III.41 in Protter [2005] or Lemma 4.3 in Sokol and Hansen [2015].

**Theorem 2.1.11.** Let $T > 0$ and $L$ be a martingale defined on $[0, T]$ with $\Delta L \geq -1$. Assume that $\mathcal{E}(L)$, which is the Doléans-Dade exponential of $L$, is a uniformly integrable martingale and let $Q$ be the probability measure having Radon-Nikodym derivative $\mathcal{E}(L)_T$ with respect to $\mathbb{P}$. If $[\mathcal{M}, L]$ is integrable under $\mathbb{P}$, then $\mathcal{M} - \langle \mathcal{M}, L \rangle$ is a martingale under $Q$, where the predictable quadratic covariation process $\langle \mathcal{M}, L \rangle$ is computed under the measure $\mathbb{P}$.

We now describe how the control process $\alpha$ affects the dynamics of $X$ through the probability measure of its sample path. Let $q : [0, T] \times E^2 \times A \rightarrow \mathbb{R}$ be a transition rate function that satisfies Assumption 2.1.4 and Assumption 2.1.5 and denote by $Q(t, \alpha)$ the matrix $q(t, i, j, \alpha)$. Compared to the previous characterization of controlled continuous-time Markov chains based on the infinitesimal generator where we only allow Markovian controls, here we allow a larger class of admissible controls. Let us
define the set $A$ to be the collection of $\mathcal{F}$-predictable processes $(\alpha_t)_{0 \leq t \leq T}$ such that $\alpha_t \in A$ for $0 \leq t \leq T$. For $\alpha \in A$, we define the scalar martingale $L^{(\alpha)}$ under $\mathbb{P}$:

$$L_t^{(\alpha)} := \int_0^t X_{s^-}^* \cdot (Q(s, \alpha_s) - Q^0) \cdot \psi_s^+ \cdot d\mathcal{M}_s. \quad (2.15)$$

Clearly, the jumps of the martingale $L^{(\alpha)}$ are given by:

$$\Delta L_t^{(\alpha)} = X_{t^-}^* \cdot (Q(t, \alpha_t) - Q^0) \cdot \psi_t^+ \cdot \Delta X_t. \quad (2.16)$$

It can be easily verified that $\psi_t^+ \cdot (e_j - X_{t-}) = \frac{m-1}{m} e_j - \sum_{i \neq j} \frac{1}{m} e_i$ when $X_{t-} = e_i \neq e_j$.

Therefore when $X_{t-} = e_i \neq e_j = X_t$, we have:

$$\Delta L_t^{(\alpha)} = X_{t^-}^* \cdot (Q(t, \alpha_t) - Q^0) \cdot \psi_t^+ \cdot (e_j - X_{t-}) = e_t^* \cdot (Q(t, \alpha_t) - Q^0) \cdot \left( \frac{m-1}{m} e_j - \sum_{k \neq j} \frac{1}{m} e_k \right)$$

$$= \frac{m-1}{m} (q(t, i, j, \alpha_t) - q_t^{0,j}) - \frac{1}{m} \sum_{k \neq j} (q(t, i, k, \alpha_t) - q_t^{0,k})$$

$$= q(t, i, j, \alpha_t) - q_t^{0,j} = q(t, i, j, \alpha_t) - 1,$$

where the last equality is due to the fact that $\sum_k (q(t, i, k, \alpha_t) - q_t^{0,k}) = 0$. Therefore we have $\Delta L_t^{(\alpha)} \geq -1$. According to Theorem III.45 in Protter [2005] and the remark that follows, in order to show that $\mathcal{E}(L^{(\alpha)})$ is uniformly integrable, it suffices to show $\mathbb{E}[\exp(\langle L^{(\alpha)}, L^{(\alpha)} \rangle_T)] < \infty$. This is straightforward since we have:

$$\langle L^{(\alpha)}, L^{(\alpha)} \rangle_T = \int_0^T X^*_{s^-} \cdot (Q(s, \alpha_s) - Q^0) \cdot \psi^+_s \cdot d\langle \mathcal{M}, \mathcal{M} \rangle_s \cdot (X^*_{s^-} \cdot (Q(s, \alpha_s) - Q^0) \cdot \psi^+_s)^* ds$$

$$= \int_0^T X^*_{s^-} \cdot (Q(s, \alpha_s) - Q^0) \cdot \psi^+_s \cdot (Q^*(s, \alpha_s) - Q^0) \cdot X_{s^-} ds,$$

and the integrand is bounded by some constant according to Assumption 2.1.5. Now using the definition of $L^{(\alpha)}$ in equation (2.15), we obtain:

$$\langle \mathcal{M}, L^{(\alpha)} \rangle_t = \int_0^t d\langle \mathcal{M}, \mathcal{M} \rangle_s \cdot (\psi^+_s)^* \cdot (Q^*(s, \alpha_s) - Q^0) \cdot X_{s^-} = \int_0^t \psi_s \cdot \psi^+_s \cdot (Q^*(s, \alpha_s) - Q^0) \cdot X_{s^-} ds$$
\[
= \int_0^t (Q^*(s, \alpha_s) - Q^0) \cdot X_s \, ds.
\]

In the last equality, we use the fact that \((Q^*(s, \alpha_s) - Q^0) \cdot X_s\) is the difference between two row vectors coming from \(Q\)-matrices, and therefore is invariant by \(\psi_s \cdot \psi_s^+\). Let us define the measure \(Q^{(\alpha)}\) by:

\[
\frac{dQ^{(\alpha)}}{d\mathbb{P}} := \mathcal{E}(L^{(\alpha)})_T. \tag{2.17}
\]

By Theorem 2.1.11, we know that the process \(\mathcal{M}^{(\alpha)}\), defined as:

\[
\mathcal{M}^{(\alpha)}_t := \mathcal{M}_t - \int_0^t (Q^*(s, \alpha_s) - Q^0) \cdot X_s \, ds \tag{2.18}
\]

is a \(Q^{(\alpha)}\)-martingale. Therefore the canonical decomposition of \(X\) can be rewritten as:

\[
X_t = X_0 + \int_{[0,t]} Q^*(s, \alpha_s) \cdot X_s \, ds + \mathcal{M}^{(\alpha)}_t. \tag{2.19}
\]

This means that under the measure \(Q^{(\alpha)}\), the stochastic intensity rate of \(X\) is given by \(Q(t, \alpha_t)\). In addition, since \(Q^{(\alpha)}\) and \(\mathbb{P}\) coincide on \(\mathcal{F}_0\), the law of \(X_0\) under \(Q^{(\alpha)}\) is \(\mathbb{P}^0\). In particular, when \(\alpha\) is a Markovian control, i.e. of the form \(\alpha_t = \phi(t, X_{t-})\) for some measurable feedback function \(\phi\), \(X\) is a continuous-time Markov chain with intensity rate \(q(t, i, j, \phi(t, i))\) and initial distribution \(\mathbb{P}^0\) under the measure \(Q^{(\alpha)}\).

**Remark 2.1.12.** In classical literatures of optimal control, admissible controls are often classified into the categories of open-loop controls and closed-loop controls. Open-loop controls are refers to controls adapted to the underlying filtration, which is often generated by the noise process. Closed-loop controls, on the other hand, are controls that are adapted to the filtration generated by the history of the state process. In our setup, however, we see that the underlying filtration is indeed the one generated by the past path of state process and therefore this difference vanishes.
2.2 Optimal Control of Continuous-Time Markov Chains

2.2.1 Analytical Approach via the Dynamic Programming Principle

We first recall the analytical approach to optimal control of continuous-time Markov chain. Let \( A \) be a compact subset of \( \mathbb{R}^l \). Let \( f : [0, T] \times E \times A \to \mathbb{R} \) and \( g : E \to \mathbb{R} \) be respectively the running cost function and the terminal cost function. In the rest of Section 2.2, we let Assumption 2.1.4 and Assumption 2.1.5 hold. In addition, we make the following assumption on the regularity of the cost functions and the transition rate:

**Assumption 2.2.1.** For all \( i, j \in E \), the mappings \((t, \alpha) \to q(t, i, j, \alpha)\) and \((t, \alpha) \to f(t, i, \alpha)\) are continuous.

Let \( A^0 \) be the collection of Markovian controls, i.e.

\[
A^0 = \{ t \to \phi(t, X_t) | \phi : [0, T] \times E \to A \text{ is measurable.} \}. \tag{2.20}
\]

We use the notation \( \alpha \leftrightarrow \phi \) to denote an element in \( A^0 \), emphasizing the feedback function \( \phi \). When we pick a control \( \alpha \leftrightarrow \phi \in A^0 \), the total expected cost is given by:

\[
J(\alpha, t, i) := \mathbb{E} \left[ \int_t^T f(s, X_s, \phi(s, X_s))ds + g(X_T)|X_t = i \right], \tag{2.21}
\]

where the dynamics of \( X \) are determined by the infinitesimal generator \( \mathcal{G}_\phi \):

\[
\mathcal{G}_\phi F(t, i) := \partial_t F(t, i) + \sum_{j \neq i} (F(t, j) - F(t, i))q(t, i, j, \phi(t, i)). \tag{2.22}
\]
We consider the optimization problem:

\[ V(t, i) := \inf_{\alpha \leftrightarrow \phi \in A^0} J(\alpha, t, i). \]  

(2.23)

For \( i = 1, \ldots, M \), we define the reduced Hamiltonian \( H_i \) by:

\[ (0, T] \times \mathbb{R}^m \times A \rightarrow \mathbb{R} \]
\[ (t, z, \alpha) \rightarrow H_i(t, z, \alpha) := f(t, i, \alpha) + \sum_{j \neq i} (z_j - z_i) q(t, i, j, \alpha). \]  

(2.24)

Let us denote the infimum of the Hamiltonian by \( \hat{H}_i(t, z) := \inf_{\alpha \in A} H_i(t, z, \alpha) \). We have the following result on the minimizer of the Hamiltonian.

**Lemma 2.2.2.** For all \( i \in E \), there exists a measurable mapping \( \hat{a}_i : [0, T] \times \mathbb{R}^m \rightarrow A \) such that \( H_i(t, z, \hat{a}_i(t, z)) = \inf_{\alpha \in A} H_i(t, z, \alpha) \) for all \( (i, t, z) \in E \times [0, T] \times A \).

**Proof.** This is a direct consequence of the measurable selection theorem (see Theorem 18.19 in Aliprantis and Border [2006]). \( \square \)

Recall the definition of the seminorm \( \| \cdot \|_{e_i} \) defined in (2.14). By the regularity of \( q \) and \( f \), we can show the following lemma:

**Lemma 2.2.3.** For all \( i \in E \), the mapping \((t, z) \rightarrow \hat{H}_i(t, z)\) is continuous. In addition, there exists a constant \( C > 0 \) such that for all \( t \in [0, T] \) and \( z, z' \in \mathbb{R}^m \), we have:

\[ |\hat{H}_i(t, z) - \hat{H}_i(t, z')| \leq C \| z - z' \|_{e_i}. \]  

(2.25)

**Proof.** By Berge’s maximum theorem, the continuity of \( H_i \) and the compactness of \( A \) implies the continuity of \( \hat{H}_i \). Now let \( z, z' \in \mathbb{R}^m \). Since the transition rate function \( q \) is bounded uniformly in \( \alpha \), we have:

\[ \inf_{\alpha \in A} H_i(t, z, \alpha) - H_i(t, z', \alpha) \leq H_i(t, z, \alpha) - H_i(t, z', \alpha) \]
\[
\sum_{j \neq i} [(z_j - z_j') - (z_i - z_i')] g(t, i, j, \alpha) \leq C \|z - z'\|_{e_i}.
\]

Since the above is true for all \(\alpha\), taking supremum of the left-hand side, we obtain
\[
\hat{H}_i(t, z) - \hat{H}_i(t, z') \leq C \|z - z'\|_{e_i}.
\]
Exchanging the roles of \(z\) and \(z'\), we obtain the desired inequality. \(\Box\)

The Markovian setting allows us to apply the Dynamic Programming Principle to the optimal control problem (2.23), which leads to the following Hamilton-Jacobi equation:

\[
\begin{align*}
\partial_t v(t, i) &= -\hat{H}_i(t, v(t, \cdot)), \quad i \in E \\
v(T, i) &= g(i),
\end{align*}
\] (2.26)

where the unknown is a mapping \(v : [0, T] \times E \to \mathbb{R}\). Since \(E\) is finite, \(v(t, \cdot)\) can be identified with an element in \(\mathbb{R}^m\), and the above equation is nothing but a coupled system of \(m\) ordinary differential equations. Using the Lipschitz regularity of \(\hat{H}_i\) we just proved, the existence and uniqueness of a local solution can be obtained by applying the Cauchy-Lipschitz theorem. In order to extend the existence of the solution to the entire interval \([0, T]\), it suffices to find a uniform bound for the solution. This can be done by linking the solution to the total expected cost associated with the Markovian control with feedback function \(\hat{\phi}(t, i) := \hat{a}_i(t, v(t, \cdot))\), where \(\hat{a}_i\) is the measurable minimizer of the reduced Hamiltonian \(H_i\) as mentioned in Lemma 2.2.2. It is plain to check that the feedback function \(\hat{\phi}\) is measurable, therefore it is an admissible Markov control. Since the control belongs to a compact set and the running cost is continuous on the control, the total expected cost can be bounded by a constant. This shows existence of the solution on the entire interval \([0, T]\) for any \(T > 0\). The uniqueness of the solution on \([0, T]\) follows immediately from the uniqueness of the local solution.

Once we show the HJB equation (2.26) admits a unique solution \(v\), by applying the Dynkin formula, it is then straightforward to verify that \(v\) is the value function.
and $\hat{\phi}(t,i) := \hat{a}_i(t,v(t,\cdot))$ is an optimal control. This leads to the following result of the verification argument.

**Theorem 2.2.4.** Equation (2.26) admits a unique solution, i.e. a mapping $v : [0,T] \times E \to \mathbb{R}$ that is $C^1$ in $t$ and satisfies (2.26), and the solution $v(t,i)$ equals the value function of the optimal control problem $V(t,i)$ defined in 2.23. In addition, given any measurable minimizer $\hat{a}_i$ of the reduced Hamiltonian $H_i$ as defined in Lemma 2.2.2, the Markovian control with the feedback function $\hat{\phi}$ defined by $\hat{\phi}(t,i) := \hat{a}_i(t,v(t,\cdot))$ is optimal.

### 2.2.2 Probabilistic Approach by BSDE

We now introduce the weak formulation of stochastic optimal control of continuous-time Markov chains which is based on the semimartingale representation. Our approach is based on the theory of BSDE driven by Markov chain, developed in Cohen and Elliott [2008] and Cohen and Elliott [2010]. For the sake of completeness, we collect a few useful results in the appendix of this chapter.

We reuse the notations of transition rate function and cost functions as in Section 2.1.2 and Section 2.2.1. Let $A$ be a compact subset of $\mathbb{R}^m$. Let $f : [0,T] \times E \times A \to \mathbb{R}$ and $g : E \to \mathbb{R}$ be, respectively, the running cost function and the terminal cost function. As is mentioned in Section 2.1.2, we shall consider all the closed-loop controls as the admissible controls in the weak formulation. In other words, the set of admissible controls $A$ is the set of all $\mathbb{F}$-predictable processes taking values in $A$. When the control is $\alpha \in A$, the total expected cost is defined as:

$$J(\alpha) := \mathbb{E}^{Q^{(\alpha)}} \left[ \int_0^T f(t,X_t,\alpha_t) dt + g(X_T) \right]$$ (2.27)
where $Q^{(\alpha)}$ is the probability measure defined in (2.17). We aim to solve the following optimization problem:

$$V := \inf_{\alpha \in A} J(\alpha).$$  \hspace{1cm} (2.28)

We now show that this optimization problem can be characterized by a BSDE driven by the continuous-time Markov chain $X$. To begin with, we define the Hamiltonian $H : [0,T] \times E \times \mathbb{R}^m \times A \to \mathbb{R}$ by

$$H(t,x,z,\alpha) := f(t,x,\alpha) + x^* \cdot (Q(t,\alpha) - Q^0) \cdot z$$  \hspace{1cm} (2.29)

Recall the definition of reduced Hamiltonian $H_i$ in (2.24) and the fact that $x$ takes values in the set $\{e_1, \ldots, e_m\}$. It is straightforward to verify that $H(t,x,z,\alpha) = \sum_{i \in E} 1(x = e_i)[H_i(t,z,\alpha) - \sum_{j \neq i}(z_j - z_i)]$.

In the following, we let Assumption 2.1.4, 2.1.5 and 2.2.1 hold. We have shown in Lemma 2.2.2 that there exists a measurable mapping $\hat{a}_i$ such that $H_i(t,z,\hat{a}_i(t,z)) = \inf_{\alpha \in A} H_i(t,z,\alpha)$. We then define the mapping $\hat{H}$ and $\hat{a}$ by:

$$\hat{H}(t,x,z) := \sum_{i = 1}^m 1(x = e_i)[\hat{H}_i(t,z) - \sum_{j \neq i}(z_j - z_i)],$$  \hspace{1cm} (2.30)

$$\hat{a}(t,x,z) := \sum_{i = 1}^m 1(x = e_i)\hat{a}_i(t,z).$$  \hspace{1cm} (2.31)

It is clear that for all $t \in [0,T]$, $x \in E$ and $z \in \mathbb{R}^m$, $\hat{a}(t,x,z)$ is a minimizer of the mapping $\alpha \to H(t,x,z,\alpha)$, and the minimum is given by $\hat{H}(t,x,z)$. Combining Lemma 2.2.3 and the definition of the seminorm $\| \cdot \|_{X_{t-}}$ (see equation (2.14)), we obtain:

**Lemma 2.2.5.** There exists a constant $C > 0$ such that for all $(\omega,t,\alpha) \in \Omega \times ]0,T] \times A$, and $z,z' \in \mathbb{R}^m$, we have:

$$|H(t,X_{t-},z,\alpha) - H(t,X_{t-},z',\alpha)| \leq C \|z - z'\|_{X_{t-}},$$  \hspace{1cm} (2.32)
|\hat{H}(t, X_{t-}, z) - \hat{H}(t, X_{t-}, z')| \leq C\|z - z'\|_{X_{t-}}. \tag{2.33}

Let us fix an admissible control \(\alpha \in \mathbb{A}\) and consider the following BSDE:

\[
Y_t = g(X_T, p_T) + \int_t^T H(s, X_{s-}, Z_s, \alpha_s)ds - \int_t^T Z_s^* \cdot d\mathcal{M}_s. \tag{2.34}
\]

**Lemma 2.2.6.** Given \(\alpha \in \mathbb{A}\), BSDE (2.34) admits a unique solution \((Y, Z)\) and we have \(J(\alpha) = \mathbb{E}^P[Y_0]\).

**Proof.** By Lemma 2.2.5, the driver function \(H\) of the BSDE (2.34) is Lipschitz in \(z\) with regard to the stochastic semi norm \(\|\cdot\|_{X_{t-}}\), therefore by Lemma 2.4.1, it admits a unique solution \((Y, Z)\). We have:

\[
Y_0 = g(X_T, p_T) + \int_0^T H(t, X_{t-}, Z_t, \alpha_t)dt - \int_t^T Z_s^* \cdot d\mathcal{M}_s
\]

\[
= g(X_T, p_T) + \int_0^T f(t, X_{t-}, \alpha_t)dt - \int_t^T Z_s^* \cdot (d\mathcal{M}_t - (Q^*(t, \alpha_t) - Q^0) \cdot X_{t-}dt)
\]

\[
= g(X_T, p_T) + \int_0^T f(t, X_{t-}, \alpha_t)dt - \int_0^T Z_t^* \cdot d\mathcal{M}_t^{(\alpha)}
\]

Since \(\mathcal{M}^{(\alpha)}\) is a martingale under the measure \(Q^{(\alpha)}\), we take the expectation under \(Q^{(\alpha)}\) and obtain that \(\mathbb{E}^{Q^{(\alpha)}}[Y_0] = J(\alpha)\). Since \(Y_0\) is \(\mathcal{F}_0\)-measurable, and \(P\) coincides with \(Q^{(\alpha)}\) on \(\mathcal{F}_0\), we have \(\mathbb{E}^P[Y_0] = J(\alpha)\). \(\square\)

Now let us consider the following BSDE:

\[
Y_t = g(X_T) + \int_t^T \tilde{H}(s, X_{s-}, Z_s)ds - \int_t^T Z_s^* \cdot d\mathcal{M}_s. \tag{2.35}
\]

We show that the above BSDE characterizes the optimality of the control problem (2.28):

**Proposition 2.2.7.** BSDE (2.35) admits a unique solution \((Y, Z)\). In addition, the value function of the optimal control problem (2.28) is given by \(V = \mathbb{E}[Y_0]\) and the

\[
\]

32
control $\hat{\alpha}$ defined by:

$$\hat{\alpha}_t = \hat{a}(t, X_{t-}, Z_t)$$

(2.36)

is an optimal control.

Proof. The existence and uniqueness of solution to (2.35) is easily verified by using the Lipschitz property of $\hat{H}$ stated in Lemma 2.2.5. Let $(Y, Z)$ be the solution and define the process $\hat{\alpha}$ by $\hat{\alpha}_t := \hat{a}(t, X_{t-}, Z_t)$. We first show that $\hat{\alpha}$ is an admissible control. Recall the definition of $\hat{a}$ in (2.31). We have:

$$\hat{a}(t, X_{t-}, Z_t) = \sum_{i=1}^{m} \mathbb{1}(X_{t-} = e_i)\hat{a}_i(t, Z_t) = X_{t-}^* \cdot \left(\sum_{i=1}^{m} \hat{a}_i(t, Z_t)e_i\right)$$

Since $\hat{a}_i$ is measurable for each $i \in \{1, \ldots, m\}$, we see that $\hat{a}$ is a measurable mapping from $[0, T] \times \mathbb{R}^m \times \mathbb{R}^m$ to $A$. Since both the processes $t \to X_{t-}$ and $Z$ are predictable, we conclude that $\hat{\alpha}$ is a predictable process and therefore an admissible control.

Now let us fix an arbitrary control $\alpha \in A$. Let us denote by $(Y^{(\alpha)}, Z^{(\alpha)})$ the solution to BSDE (2.34) corresponding to the control $\alpha$. Moreover, $(Y, Z)$ is also the solution to BSDE (2.34) corresponding to the control $\hat{\alpha}$. We denote $\Delta Y := Y^{(\alpha)} - Y$ and $\Delta Z := Z^{(\alpha)} - Z$. Computing the difference of the two BSDEs, we notice that $\Delta Y$ and $\Delta Z$ are the solutions to the following BSDE:

$$\Delta Y_t = \int_t^T [H(s, X_{s-}, Z_s^{(\alpha)}, \alpha_s) - H(s, X_{s-}, Z_s, \hat{\alpha}_s)]ds - \int_t^T \Delta Z_s^* \cdot dM_s.$$ 

We can further decompose the driver of the above BSDE as:

$$H(s, X_{s-}, Z_s^{(\alpha)}, \alpha_s) - H(s, X_{s-}, Z_s, \hat{\alpha}_s) = H(s, X_{s-}, Z_s^{(\alpha)}, \alpha_s) - H(s, X_{s-}, Z_s, \alpha_s) + H(s, X_{s-}, Z_s, \alpha_s) - H(s, X_{s-}, Z_s, \hat{\alpha}_s)$$

$$= [H(s, X_{s-}, Z_s, \alpha_s) - H(s, X_{s-}, Z_s, \hat{\alpha}_s)] + X_{s-}^* \cdot (Q(s, \alpha_s) - Q^0) \cdot \Delta Z.$$
Define the processes 
\[ \psi_t := H(t, X_{t-}, Z_t, \alpha_t) - H(t, X_{t-}, Z_t, \hat{\alpha}_t) \]
and
\[ \gamma_t := (Q^*(t, \alpha_t) - Q^0) \cdot X_{t-} \]. Therefore \((\Delta Y, \Delta Z)\) is the solution to the linear BSDE of the form \((2.69)\) with \(\psi\) and \(\gamma\) defined previously and \(\beta = 0\). Clearly \(\psi\) and \(\gamma\) are both predictable. Since \(\hat{\alpha}_t\) minimizes the Hamiltonian, \(\psi\) is nonnegative. The boundedness of \(\gamma\) follows from the boundedness of the transition rate function \(q\), which is a consequence of the boundedness of \(A\) and the Lipschitz property of \(q\) as stated in Assumption 2.2.1. It remains to check that \(1 + \gamma_t^* \cdot \psi_t^+ \cdot (e_j - X_{t-}) \geq 0\).

Indeed, when \(X_{t-} = e_j\), it is trivial that the above inequality holds. We now assume that \(X_{t-} = e_i \neq e_j\). We have
\[ \psi_t^+ \cdot (e_j - X_{t-}) = \frac{m-1}{m} e_j - \sum_{i \neq j} \frac{1}{m} e_i. \]

Therefore when \(X_{t-} = e_i \neq e_j\), we have:
\[ \gamma_t^* \cdot \psi_t^+ \cdot (e_j - X_{t-}) = X_{t-}^* \cdot (Q(t, \alpha_t) - Q^0) \cdot \psi_t^+ \cdot (e_j - X_{t-}) = e_i^* \cdot (Q(t, \alpha_t) - Q^0) \cdot \left( \frac{m-1}{m} e_j - \sum_{k \neq j} \frac{1}{m} e_k \right) \]
\[ = \frac{m-1}{m} (q(t, i, j, \alpha_t) - q^0_{i,j}) - \frac{1}{m} \sum_{k \neq j} (q(t, i, k, \alpha_t) - q^0_{i,k}) = q(t, i, j, \alpha_t) - q^0_{i,j}, \]

where the last equality is due to the fact that \(\sum_{k} (q(t, i, k, \alpha_t) - q^0_{i,k}) = 0\). Therefore we have:
\[ 1 + \gamma_t^* \cdot \psi_t^+ \cdot (e_j - X_{t-}) = 1 + q(t, i, j, \alpha_t) - q^0_{i,j} = q(t, i, j, \alpha_t) \geq 0. \]

By Lemma 2.4.3, we conclude that \(\Delta Y\) is nonnegative and in particular \(\mathbb{E}[Y_0^{(\alpha)}] \geq \mathbb{E}[Y_0]\). Since \(\alpha\) is an arbitrary admissible control, in light of Lemma 2.2.6, this means that \(\mathbb{E}[Y_0] \leq \inf_{\alpha \in \mathcal{A}} J(\alpha) = V\). Finally, we notice that \(\mathbb{E}[Y_0]\) is the total expected cost when the control is \(\hat{\alpha}\). We conclude that \(\hat{\alpha}\) is an optimal control and \(\mathbb{E}[Y_0] = V\). $\square$

**Remark 2.2.8.** In both the analytical and probabilistic approach we presented above, we chose to work under weak assumptions of regularity on the transition rate function and the cost function (see Assumption 2.1.4, 2.1.5 and 2.2.1). As a result, we have to
rely on an argument of measurable selection, in order to obtain the regularity for the minimizer of the Hamiltonian. Indeed, we can construct an optimal control from any minimizer of the Hamiltonian. However, if we assume in addition that the minimizer of the Hamiltonian is unique, then the optimal control is unique.

**Proposition 2.2.9.** Assume in addition that for all \((i, t, z) \in E \times [0, T] \times \mathbb{R}^m\), the reduced Hamiltonian \(A \ni \alpha \rightarrow H_i(t, z, \alpha)\) admits a unique minimizer \(\hat{\alpha}_i(t,z)\). Then for any optimal control \(\alpha' \in \mathbb{A}\), we have \(\alpha'_i = \hat{\alpha}_i := \hat{\alpha}(t, X_{t-}, Z_t), dt \otimes d\mathbb{P}\text{-a.e.},\) where \((Y, Z)\) is the solution to BSDE (2.35) and \(\hat{\alpha}\) is defined in (2.31).

**Proof.** We consider the solution \((Y', Z')\) to the following BSDE:

\[
Y'_t = \int_t^T H(s, X_{s-}, Z_s', \alpha'_s)ds - \int_t^T (Z'_s)^* \cdot dM_s. \tag{2.37}
\]

Since \(\alpha'\) is optimal, we have \(\mathbb{E}^\mathbb{P}[Y'_0] = J(\alpha', p, \nu) = V(p, \nu) = \mathbb{E}^\mathbb{P}[Y_0].\) Now taking the difference of the BSDE (2.35) and (2.37), we obtain:

\[
Y_0 - Y'_0 = \int_0^T \left[H(t, X_{t-}, Z_t, \hat{\alpha}_t) - H(t, X_{t-}, Z'_t, \alpha'_t)\right] dt - \int_0^T (Z_t - Z'_t)^* \cdot dM_t
\]

\[
= \int_0^T \left[X^*_t \cdot (Q(t, \hat{\alpha}_t) - Q^0) \cdot Z_t - X^*_t \cdot (Q(t, \alpha'_t) - Q^0) \cdot Z'_t\right] dt
\]

\[
+ \int_0^T \left[f(t, X_{t-}, \hat{\alpha}_t) - f(t, X_{t-}, \alpha'_t)\right] dt - \int_0^T (Z_t - Z'_t)^* \cdot dM_t
\]

\[
= \int_0^T \left[f(t, X_{t-}, \hat{\alpha}_t) - f(t, X_{t-}, \alpha'_t) + X^*_t \cdot (Q(t, \hat{\alpha}_t) - Q(t, \alpha'_t)) \cdot Z_t\right] dt
\]

\[
- \int_0^T (Z_t - Z'_t)^* \cdot [dM_t - (Q^* (t, \alpha'_t) - Q^0) \cdot X_{t-} dt]
\]

\[
= \int_0^T \left[H(t, X_{t-}, Z_t, \hat{\alpha}_t) - H(t, X_{t-}, Z_t, \alpha'_t)\right] dt - \int_0^T (Z_t - Z'_t)^* \cdot dM_t^{(\alpha')}. \tag{2.36}
\]

Taking the expectation under the measure \(\mathbb{Q}^{(\alpha')}\) and using the fact that \(\mathbb{Q}^{(\alpha')}\) coincides with \(\mathbb{P}\) in \(\mathcal{F}_0\), we have:

\[
0 = \mathbb{E}^\mathbb{P}[Y_0 - Y'_0] = \mathbb{E}^\mathbb{Q}^{(\alpha')}[Y_0 - Y'_0] = \mathbb{E}^\mathbb{Q}^{(\alpha')}\left[\int_0^T \left[H(t, X_{t-}, Z_t, \hat{\alpha}_t) - H(t, X_{t-}, Z_t, \alpha'_t)\right] dt\right] \leq 0,
\]

35
where the last inequality is due to the fact that $\hat{\alpha}_t$ minimizes the Hamiltonian. In fact, we have $\hat{\alpha}_t = \alpha'_t, \ dt \otimes dQ^{(\alpha')} - a.e.$ If we assume otherwise, the last inequality would be strict, since the minimizer of the Hamiltonian is unique. Since $\mathbb{P}$ is equivalent to $Q^{(\alpha)}$, we have $\hat{\alpha}_t = \alpha'_t, \ dt \otimes d\mathbb{P} - a.e.$

2.3 Analytical Approach to Finite State Mean Field Games

In this section, we shall revisit the analytical formulation of finite state mean field games. We shall see that Nash equilibria of finite state mean field games can be characterized by a system of two strongly coupled ordinary differential equations. The first equation - the Hamilton-Jacobi equation - translates a representative player’s optimal response facing the other players in the game, while the second equation - the Kolmogorov equation - depicts the evolution of the distribution of players’ states. We will provide a version of Lasry-Lions monotonicity condition for finite state mean field games, which guarantees the uniqueness of the Nash equilibrium. We shall also formulate the $N$-player game and explained how the solution of the mean field game consists of an approximate Nash equilibrium for the $N$-player game. For the results and their proofs in this section, we will loosely follow the presentation in Gomes et al. [2013]. We also refer the readers to Chapter 7, Volume I of Carmona and Delarue [2017] for a detailed account of analytical formulation to finite state mean field games and numerical examples.

2.3.1 Model Setup

We consider a continuous-time game of duration $T$ with a large population of players, whose states can be in one of $m$ possible states. As mentioned previously, the dynamics according to which each player switches their states are determined by a
transition rate matrix. To simplify the model, we assume that each player is allowed
to control directly the transition rate. Using the notations of Section 2.2.1, we let $A$
be a compact subset of $(\mathbb{R}^+)^m$. For a control $\alpha = (\alpha_1, \ldots, \alpha_m) \in A$, we define the
transition rate $q$:

\[
q(t, i, j, \alpha) := \alpha_j, \quad \text{for } j \neq i,
\]

\[
q(t, i, i, \alpha) := -\sum_{j \neq i} \alpha_j. \tag{2.38}
\]

We assume that each player can only observe its own state and chooses a Marko-
vian control $\alpha \leftrightarrow \phi \in \mathcal{A}^0$ of the form $\alpha_t = \phi(t, X_t)$, where $\phi$ is the feedback function.
Each player then incurs a running cost and a terminal cost which depend on its state,
its control, and the statistical distribution of all the players’ states. Since there are
only $m$ possible states, the distribution of all the players’ states at any given time $t$
can be described by an element in the $m$-dimensional simplex, which we denote by $\mathcal{S}$:

\[
\mathcal{S} := \{ p \in \mathbb{R}^m, \sum p_i = 1, p_i \geq 0 \}. \tag{2.39}
\]

We consider the mean field limit where the number of players tends to infinity
and each player has the same cost functions and uses the Markov control with the
same feedback function $\phi$. Under such a limit, a version of law of large number
applies and the empirical distribution of all the players’ states converges to the law
of any single player’s state, which can be represented by a deterministic mapping
$\mu : [0, T] \ni t \rightarrow \mu(t) \in \mathcal{S}$. Assuming that all the players use the control $\alpha \leftrightarrow \phi \in \mathcal{A}^0$,
the flow of distribution $\mu$ evolves according to the Kolmogorov equation:

\[
\frac{d\mu_i(t)}{dt} = \sum_{j \in E} \mu_j(t) \cdot q(t, j, i, \phi(t, j)), \quad \mu(0) = p^0. \tag{2.40}
\]

Here $p^0$ is the distribution of players’ states at $t = 0$. Now we single out one player
which we call the representative player. When it picks a control $\alpha \leftrightarrow \phi \in \mathcal{A}^0$, and
the distribution of all the players’ states at time $t$ is given by $\mu(t)$ for $t \leq T$, the
representative player incurs a total expected cost which it tries to minimize:

$$J(\alpha, \mu) := \mathbb{E} \left[ \int_0^T f(t, X_t, \alpha_t, \mu(t)) + g(X_T, \mu(T)) \right].$$  \hspace{1cm} (2.41)

Our goal is to identify Nash equilibria of the mean field game. In such an equi-
librium which is characterized by a control $\hat{\alpha} \in \mathbb{A}^0$ and a flow of distribution $\hat{\mu}$ of
players’ states, the representative player cannot do better by deviating from the con-
trol $\hat{\alpha}$, provided that all the other players maintain their controls $\hat{\alpha}$. Notice that in
the limit of infinitely many players, the effect on the distribution of players’ states
resulting from the unilateral deviation of the representative player is indeed negli-
gible. Therefore, when we derive the optimal response of the representative player, we
can treat the distribution of players’ states as fixed. We give the formal definition of
Nash equilibrium below:

**Definition 2.3.1.** Given $p^0 \in S$. We say a tuple $(\hat{\alpha} \leftrightarrow \hat{\phi}, \hat{\mu})$ is a Nash equilibrium
of the finite state mean field game with initial distribution $p^0$ if we have:

$$\hat{\alpha} = \arg \inf_{\alpha \leftrightarrow \phi \in \mathbb{A}^0} \mathbb{E} \left[ \int_0^T f(t, X_t, \alpha_t, \hat{\mu}(t)) + g(X_T, \hat{\mu}(T)) \right],$$  \hspace{1cm} (2.42)

$$\frac{d\hat{\mu}_i(t)}{dt} = \sum_{j \in E} \hat{\mu}_j(t) \cdot q(t, j, i, \hat{\phi}(t, j)), \quad \hat{\mu}(0) = p^0.$$  \hspace{1cm} (2.43)

Quite naturally, the search of Nash equilibrium can be carried out in these two
steps:

**Step 1** (Representative player’s optimization problem)

Given a flow of measures $\mu$, solve the optimization problem:

$$\inf_{\alpha \leftrightarrow \phi \in \mathbb{A}^0} \mathbb{E} \left[ \int_0^T f(t, X_t, \alpha_t, \mu(t)) + g(X_T, \mu(T)) \right],$$  \hspace{1cm} (2.44)
where \(X_t\) has the transition rate \(q(t, i, j, \phi(t, i))\). We denote by \(\alpha^{(\mu)} \leftrightarrow \phi^{(\mu)}\) the optimal control of the representative player facing the mean field \(\mu\).

**Step 2** (Consistency of the flow of measures)
Find the measure flow \(\mu\), such that for all \(t \in [0, T]\), the distribution of \(X_t\), which is the state of the player using the optimal control \(\alpha^{(\mu)}\), coincides with \(\mu(t)\). This can be translated into the Kolmogorov equation:

\[
\frac{d\mu_i(t)}{dt} = \sum_{j \in E} \mu_j(t) \cdot \phi_i^{(\mu)}(t, j), \quad \mu(0) = \mathbf{p}^\circ.
\]  

(2.45)

where \(\phi_i^{(\mu)}(t, j)\) is the \(i\)-th component of \(\phi^{(\mu)}(t, j)\).

### 2.3.2 Existence and Uniqueness of Nash Equilibria

We now implement the above scheme to identify the Nash equilibrium. We make the following assumption on the cost functions \(f\) and \(g\).

**Assumption 2.3.2.** (i) \(f\) is Lipschitz in \(\mu\), uniformly in \(t, \alpha\) and \(i\).

(ii) For \(i \in E\), \(\alpha \to f(t, i, \alpha, \mu)\) does not depend on the \(i\)-th component of \(\alpha\) and is differentiable with respect to \(\alpha\). In addition, the gradient of \(f\) with regard to \(\alpha\) is Lipschitz in \(\mu\), uniformly in \(t, i\) and \(\alpha\).

(iii) \(f\) is strongly convex in \(\alpha\), in the sense that there exists a constant \(\gamma > 0\) such that for all \((t, i, \mu) \in [0, T] \times E \times \mathcal{S}\) and all \(\alpha, \alpha' \in A\) we have:

\[
f(t, i, \alpha', \mu) - f(t, i, \alpha, \mu) - \Delta_\alpha f(t, i, \alpha, \mu) \cdot (\alpha' - \alpha) \geq \gamma \|\alpha - \alpha'\|^2.
\]  

(2.46)

(iv) \(g\) is Lipschitz in \(\mu\), uniformly in \(i\).

**Representative player’s optimization problem**

We now fix a flow of measures \(\mu : [0, T] \rightarrow \mathcal{S}\) and solve the optimization problem (2.44). Using the notation introduced in Section 2.2.1, the reduced Hamiltonian \(H_i\)
becomes:

\[ H_i(t, z, \alpha, \mu) := f(t, i, \alpha, \mu) + \sum_{j \neq i} (z_j - z_i)\alpha_j. \]

Since \( A \) is convex and compact and \( f \) is uniformly convex in \( \alpha \), we can prove that the mapping \([\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_m] \to H_i(t, z, \alpha, \mu)\) admits a unique minimizer \( \hat{a}(t, i, z, \mu) \), of which the components are denoted by \( \hat{a}_j(t, i, z, \mu) \) for \( j \neq i \). We denote the minimum of the Hamiltonian by \( \hat{H}_i(t, z, \mu) \). In addition, we define \( \hat{a}_i(t, i, z, \mu) := -\sum_{j \neq i} \hat{a}_j(t, i, z, \mu) \). Using Assumption 2.3.2, in particular the uniform convexity of \( f \) on \( \alpha \) and the Lipschitz property of \( \Delta_\alpha f \), we can show the following result:

**Lemma 2.3.3.** \( \hat{a}_j(t, i, z, \mu) \) is Lipschitz in \( z \) and \( \mu \), uniformly in \( t, i, j \). \( \hat{H}_i \) is locally Lipschitz in \( z \) and \( \mu \).

From Theorem 2.2.4, we see that the optimal control of the representative player is characterized by the following Hamilton-Jacobi equation:

\[
\partial_t v(t, i) = -\hat{H}_i(t, v(t, \cdot), \mu(t)), \quad i \in E, \quad (2.47)
\]

\[ v(T, i) = g(i, \mu(T)). \]

More specifically, the ODE (2.47) admits a unique solution \( v^{(\mu)} \) and the optimal control is given by \( \varphi^{(\mu)}_j(t, i) = \hat{a}_j(t, i, v^{(\mu)}(t, \cdot), \mu(t)) \).

**Consistency of the flow of measures**

The second step of the search for Nash equilibrium can be formulated as a fixed-point problem. Let us consider the following Kolmogorov equation:

\[
\frac{dp_i(t)}{dt} = \sum_{j \in E} p_j(t) \cdot \hat{a}_i(t, j, v^{(\mu)}(t, \cdot), p(t)), \quad i \in \{1, \ldots, m\}, \quad (2.48)
\]

with the initial condition \( p(0) = p^0 \). By the Lipschitz property and boundedness of the mapping \( \hat{a} \), the ODE (2.48) admits a unique solution which we denote by \( p^{(\mu)} \).

We then consider the mapping \( \Phi : \mu \to p^{(\mu)} \). It is clear that in order to show the
existence of a Nash equilibrium, we need to show that the mapping $\Phi$ admits a fixed point.

Let us denote $C := C([0, T], S)$ the space of continuous mapping from $[0, T]$ to the simplex $S$ endowed with the uniform norm $\| \cdot \|_\infty$. This will be the space where the flow of measures lives and clearly it is a Banach space. Indeed, the Lipschitz property and boundedness of the mapping $\hat{a}$ implies that any solution $p^{(\mu)}$ of (2.48) is Lipschitz in $t$ with Lipschitz constant bounded uniformly in $\mu$. Therefore by the Arzela-Ascoli theorem, we can construct a convex and compact subset of $C$ which is stable by the mapping $\Phi$. In addition, the Lipschitz continuity of the mappings $\hat{a}_i$ and $\hat{H}_i$ leads to the continuity of the mapping $\Phi$ with regard to the norm $\| \cdot \|_\infty$. Finally, we apply Schauder’s fixed point theorem to the mapping $C$. To sum it up, we have just proved:

**Theorem 2.3.4.** Under Assumption 2.3.2, the following system of forward-backward ODEs admits a solution:

\[
\begin{align*}
\partial_t v(t, i) &= \hat{H}_i(t, v(t, \cdot), p(t)), \quad v(T, i) = g(i, p(T)), \quad i \in \{1, \ldots, m\} \\
\frac{dp_i(t)}{dt} &= \sum_{j=1}^{m} p_j(t) \cdot \hat{a}_i(t, j, v(t, \cdot), p(t)), \quad p(0) = p^\circ, \quad i \in \{1, \ldots, m\}. 
\end{align*}
\]

The finite state mean field game admits a Nash equilibrium in the sense of Definition 2.3.1, and at the equilibrium, the distribution of the states of the players is given by $(p(t))_{t \in [0, T]}$ and the strategies of the players are given by the feedback function $\phi_j(t, i) := \hat{a}_i(t, j, v(t, \cdot), p(t))$.

We now investigate the uniqueness of the Nash equilibrium. In reminiscence of the monotonicity condition proposed in Lasry and Lions [2006a], Lasry and Lions [2006b] and Lasry and Lions [2007], we work under the following assumption:

**Assumption 2.3.5.** (i) The cost functional $f$ takes the following form:

\[
f(t, X, \alpha, p, \nu) = f_0(t, X, p) + f_1(t, X, \alpha).
\]
(ii) For all $t \in [0, T]$, $i \in \{1, \ldots, m\}$ and $p, p' \in S$, we have:

$$\sum_{i=1}^{m} (g(e_i, p) - g(e_i, p'))(p_i - p'_i) \geq 0,$$

(2.52)

$$\sum_{i=1}^{m} (f_0(t, e_i, p) - f_0(t, e_i, p'))(p_i - p'_i) \geq 0,$$

(2.53)

and the inequality is strict if $p \neq p'$.

**Theorem 2.3.6.** Under Assumption 2.3.5, the system of forward-backward ODEs (2.49)-(2.50) has at most one solution.

**Proof.** Let $(p^{(1)}, v^{(1)})$ and $(p^{(2)}, v^{(2)})$ be two solutions to the system (2.49)-(2.50). We denote by $\alpha_j^{(1)}(t, i) := \hat{a}_j(t, v^{(1)}(t, \cdot), p^{(1)}(t))$ and $\alpha_j^{(2)}(t, i) := \hat{a}_j(t, v^{(2)}(t, \cdot), p^{(2)}(t))$ the optimal control corresponding to the two solutions respectively. Observe that:

$$\frac{d}{dt} \left[ \sum_{i=1}^{m} (p_i^{(1)}(t) - p_i^{(2)}(t))(v_i^{(1)}(t) - v_i^{(2)}(t)) \right]$$

$$= \sum_{i=1}^{m} (p_i^{(1)}(t) - p_i^{(2)}(t))(-\tilde{H}_i(t, v_i^{(1)}(t, \cdot), p_i^{(1)}(t)) + \tilde{H}_i(t, v_i^{(2)}(t, \cdot), p_i^{(2)}(t)))$$

$$+ \sum_{i=1}^{m} \sum_{j=1}^{m} \left( p_j^{(1)}(t)\alpha_j^{(1)}(t, j) - p_j^{(2)}(t)\alpha_j^{(2)}(t, j) \right) (v_i^{(1)}(t) - v_i^{(2)}(t))$$

$$= \sum_{i=1}^{m} (p_i^{(1)}(t) - p_i^{(2)}(t)) \left[ f(t, i, \alpha^{(2)}(t, i), p_i^{(2)}(t)) + \sum_{j=1}^{m} v_j^{(2)}(t, j)\alpha_j^{(2)}(t, i) \right]$$

$$- \sum_{i=1}^{m} (p_i^{(1)}(t) - p_i^{(2)}(t)) \left[ f(t, i, \alpha^{(1)}(t, i), p_i^{(1)}(t)) + \sum_{j=1}^{m} v_j^{(1)}(t, j)\alpha_j^{(1)}(t, i) \right]$$

$$+ \sum_{i=1}^{m} \sum_{j=1}^{m} \left( p_j^{(1)}(t)\alpha_j^{(1)}(t, j) - p_j^{(2)}(t)\alpha_j^{(2)}(t, j) \right) (v_i^{(1)}(t) - v_i^{(2)}(t)).$$

Interchanging the index $i$ and $j$ in the summation and rearranging the terms, we obtain:

$$\frac{d}{dt} \left[ \sum_{i=1}^{m} (p_i^{(1)}(t) - p_i^{(2)}(t))(v_i^{(1)}(t) - v_i^{(2)}(t)) \right]$$
we minimize the hamiltonian
\[ H \]
Interchanging the indices, we also have:
\[ \text{Injecting these two inequalities into the computation of the derivative, we have:} \]
\[
\begin{align*}
= \sum_{i=1}^{m} p_{i}^{(2)}(t) \left[ f(t, i, \alpha^{(1)}(t, i), p^{(1)}(t)) - f(t, i, \alpha^{(2)}(t, i), p^{(2)}(t)) + \sum_{j=1}^{m} (\alpha_{j}^{(1)}(t, i) - \alpha_{j}^{(2)}(t, i))v^{(1)}(t, j) \right] \\
- \sum_{i=1}^{m} p_{i}^{(1)}(t) \left[ f(t, i, \alpha^{(1)}(t, i), p^{(1)}(t)) - f(t, i, \alpha^{(2)}(t, i), p^{(2)}(t)) + \sum_{j=1}^{m} (\alpha_{j}^{(1)}(t, i) - \alpha_{j}^{(2)}(t, i))v^{(2)}(t, j) \right] \\
= \sum_{i=1}^{m} p_{i}^{(2)}(t) \left[ H_{i}(t, v^{(1)}(t, \cdot), \alpha^{(1)}(t, i), p^{(1)}(t)) - H_{i}(t, v^{(1)}(t, \cdot), \alpha^{(2)}(t, i), p^{(2)}(t)) \right] \\
- \sum_{i=1}^{m} p_{i}^{(1)}(t) \left[ H_{i}(t, v^{(2)}(t, \cdot), \alpha^{(1)}(t, i), p^{(1)}(t)) - H_{i}(t, v^{(2)}(t, \cdot), \alpha^{(2)}(t, i), p^{(2)}(t)) \right].
\end{align*}
\]

We have assumed that \( f \) is of the form \( f(t, X, \alpha, p) = f_{0}(t, X, p) + f_{1}(t, X, \alpha) \), when we minimize the hamiltonian \( H_{i}(t, z, \alpha, p) \) over \( \alpha \), the minimizer does not depend on \( p \). Therefore, since \( \alpha^{(1)}(t, i) \) minimizes \( \alpha \to H_{i}(t, v^{(1)}(t, \cdot), \alpha, p^{(1)}(t)) \), it also minimizes \( \alpha \to H_{i}(t, v^{(1)}(t, \cdot), \alpha, p^{(2)}(t)) \), which implies that
\[
H_{i}(t, v^{(1)}(t, \cdot), \alpha^{(2)}(t, i), p^{(2)}(t)) \geq H_{i}(t, v^{(1)}(t, \cdot), \alpha^{(1)}(t, i), p^{(2)}(t)).
\]

Interchanging the indices, we also have:
\[
H_{i}(t, v^{(2)}(t, \cdot), \alpha^{(1)}(t, i), p^{(1)}(t)) \geq H_{i}(t, v^{(2)}(t, \cdot), \alpha^{(2)}(t, i), p^{(1)}(t)).
\]

Injecting these two inequalities into the computation of the derivative, we have:
\[
\begin{align*}
\frac{d}{dt} \left[ \sum_{i=1}^{m} (p_{i}^{(1)}(t) - p_{i}^{(2)}(t))(v^{(1)}(t, i) - v^{(2)}(t, i)) \right] \\
\leq \sum_{i=1}^{m} p_{i}^{(2)}(t) \left[ H_{i}(t, v^{(1)}(t, \cdot), \alpha^{(1)}(t, i), p^{(1)}(t)) - H_{i}(t, v^{(1)}(t, \cdot), \alpha^{(1)}(t, i), p^{(2)}(t)) \right] \\
- \sum_{i=1}^{m} p_{i}^{(1)}(t) \left[ H_{i}(t, v^{(2)}(t, \cdot), \alpha^{(2)}(t, i), p^{(1)}(t)) - H_{i}(t, v^{(2)}(t, \cdot), \alpha^{(2)}(t, i), p^{(2)}(t)) \right] \\
= - \sum_{i=1}^{m} (p_{i}^{(1)}(t) - p_{i}^{(2)}(t))(f_{0}(t, e_{i}, p^{(1)}(t)) - f_{0}(t, e_{i}, p^{(2)}(t)).
\end{align*}
\]
Using the initial and terminal conditions $p^{(1)}(0) - p^{(2)}(0) = 0$, $v^{(1)}(T, i) = g(e_i, p^{(1)}(T))$ and $v^{(2)}(T, i) = g(e_i, p^{(2)}(T))$, we integrate the inequality and obtain:

\[
0 \geq \int_0^T \sum_{i=1}^m (p_i^{(2)}(t) - p_i^{(1)}(t))(f_0(t, e_i, p^{(2)}(t)) - f_0(t, e_i, p^{(1)}(t)))dt \\
+ \sum_{i=1}^m (p_i^{(2)}(T) - p_i^{(1)}(T))(g(e_i, p^{(2)}(T)) - g(e_i, p^{(1)}(T))).
\]

Now we assume by contradiction that there exists $j \in E$ and $t \in [0, T]$ such that $p_j^{(1)}(t) \neq p_j^{(2)}(t)$. Then by the continuity of $p^{(1)}$ and $p^{(2)}$, there exists $\epsilon > 0$ such that $p_j^{(1)}(s) \neq p_j^{(2)}(s)$ for $s \in [t - \epsilon, t + \epsilon] \cap [0, T]$. By the monotonicity condition stated in Assumption 2.3.5, we deduce that the integrand in the above inequality is positive. In addition, we have $(p_i^{(2)}(s) - p_i^{(1)}(s))(f_0(s, e_i, p^{(2)}(s)) - f_0(s, e_i, p^{(1)}(s))) > 0$ for $s \in [t - \epsilon, t + \epsilon] \cap [0, T]$. This implies that the integral on the right hand side is strictly positive, which leads to a contradiction. It follows that $p^{(1)} = p^{(2)}$ and subsequently $v^{(1)} = v^{(2)}$.

**2.3.3 Game With Finitely Many Players**

We now present the model for the game with $N$ players in the finite state space $E$. Let us denote by $X_{t}^{n,N}$ the state of player $n$ at time $t$. We assume that each player $n$ can observe its own state $X_{t}^{n,N}$ as well as the empirical distribution of the states of all players in the game $\mu_{t}^{N}$. Clearly, $\mu_{t}^{N}$ is given by the following vector:

\[
\mu_{t}^{N} := \frac{1}{N} \left[ \sum_{n=1}^N 1(X_{t}^{n,N} = 1), \sum_{n=1}^N 1(X_{t}^{n,N} = 2), \ldots, \sum_{n=1}^N 1(X_{t}^{n,N} = m) \right],
\]  
(2.54)

which belongs to $S^N$, a finite subset of the $m$-dimensional simplex $S$:

\[
S^N := \left\{ \left[ \frac{n_1}{N}, \ldots, \frac{n_m}{N} \right] \middle| n_i \in \mathbb{N}, \sum_{i=1}^m n_i = N \right\}.
\]  
(2.55)
As in the mean field game, we assume that agents can directly control the transition rates of their states. Accordingly, the control space is $A = (\mathbb{R}^+)^m$ and the admissible strategy of player $n$ at time $t$ is a function of $t$, $X_{t}^{n,N}$ and $\mu_{t}^{N}$ taking values in $\mathbb{R}^m$. We define the set of admissible strategies for player $n$ as:

$$A_{n,N} := \{t \rightarrow \phi^n(t, X_{t}^{n,N}, \mu_{t}^{N})|\phi^n : [0,T] \times E \times S \rightarrow A\}. \quad (2.56)$$

Here $\phi^n$ is the feedback function of player $n$ and we denote by $\phi^n_j(t, i, p)$ the $j$-th component of $\phi^n(t, i, p)$. We now describe the joint dynamics for the states of all the players in the game. We assume that the process $t \rightarrow (X_1^t, \ldots, X_N^t)$ is a continuous-time Markov chain and conditioned on the current states of the players in the game, the changes of state are independent for different players. Accordingly we define the transition rate to be:

$$P[X_{t+\Delta t}^1 = j^1, \ldots, X_{t+\Delta t}^N = j^N | X_t^1 = i^1, \ldots, X_t^N = i^N] := \prod_{n=1}^N 1(i_n \neq j_n)[\phi^n_j(t, i_n, \mu_{t}^{N})\Delta t + o(\Delta t)]$$

$$\times 1(i_n = j_n)[1 - \sum_{j \neq i_n} \phi^n_j(t, i_n, \mu_{t}^{N})\Delta t + o(\Delta t)]. \quad (2.57)$$

where we recall the definition of the empirical measure $\mu_{t}^{N}$ in (2.54). Notice that by expanding the above product and retaining only the term of the same order as $\Delta t$, we obtain the transition rate matrix for the the continuous-time Markov chain $(X^1, \ldots, X^N)$, which is a sparse matrix of dimension $m^N$. Moreover, due to the independence assumption, at most one player changes its state at any given time. We then define the total expected cost of player $n$:

$$\mathbb{E} \left[ \int_0^T f(t, X_t^{n,N}, \phi^n(t, X_t^{n,N}, \mu_{t}^{N}), \mu_{t}^{N}) + g(X_T^{n,N}, \mu_{T}^{N}) \right]. \quad (2.58)$$

Since our objective is to explore the connection with the mean field game, we shall consider only symmetric Nash equilibria, in which the players use Markovian
control with the same feedback function $\phi$. Without loss of generality, we single out player $n = 1$ as the deviating player who uses a different strategy $\bar{\phi}$ and we assume that all the other players use the strategy $\phi$. We now show that $(X_{t}^{1,N}, \mu_{t}^{N})$ is a continuous-time Markov chain. Let us denote $(\mathcal{F}_{t}^{N})_{t \in [0,T]}$ the natural filtration of $(X_{1,N}, \ldots, X_{N,N})$. For any $j \in E$, $\pi \in \mathcal{S}^{N}$ and $t \leq T$ and $h > 0$ such that $t + h \leq T$, by the Markovian property, we have:

$$\mathbb{P}[X_{t+h}^{1,N} = j, \mu_{t+h}^{N} = \pi | \mathcal{F}_{t}^{N}] = \mathbb{P}[X_{t+\Delta t}^{1,N} = j, \mu_{t+\Delta t}^{N} = \pi | X_{t}^{1,N}, \ldots, X_{t}^{N,N}].$$

Now assuming that at time $t$, we have $X_{t}^{n,N} = i^{n}$ for $n = 1, \ldots, N$ and we denote $p \in \mathcal{S}^{N}$ the empirical measure corresponds to the states $(i^{1}, \ldots, i^{N})$. Clearly we have:

$$\mathbb{P}[X_{t+h}^{1,N} = j, \mu_{t+h}^{N} = \pi | \cap_{n=1}^{N} \{X_{t}^{n,N} = i^{n}\}] = \mathbb{P}[X_{t+h}^{1,N} = j, \mu_{t+h}^{N} = \pi | \cap_{n=2}^{N} \{X_{t}^{n,N} = i^{n}\} \cap \{X_{t}^{1,N} = i^{1}\} \cap \{\mu_{t}^{N} = p\}].$$

Since all the players except player 1 use the same strategy, the above probability will remain the same if we permute the states among players 2 to $N$, provided that the empirical distribution of all players remains to be $p$. In fact, all these permutations constitute the event $\{X_{t}^{1} = i^{1}\} \cap \{\mu_{t}^{N} = p\}$. Therefore we have:

$$\mathbb{P}[X_{t+h}^{1,N} = j, \mu_{t+h}^{N} = \pi | \cap_{n=1}^{N} \{X_{t}^{n,N} = i^{n}\}] = \mathbb{P}[X_{t+\Delta t}^{1,N} = j, \mu_{t+\Delta t}^{N} = \pi | X_{t}^{1,N} = i^{1}, \mu_{t}^{N} = p],$$

and subsequently:

$$\mathbb{P}[X_{t+h}^{1,N} = j, \mu_{t+h}^{N} = \pi | \mathcal{F}_{t}^{N}] = \mathbb{P}[X_{t+\Delta t}^{1,N} = j, \mu_{t+\Delta t}^{N} = \pi | X_{t}^{1,N}, \mu_{t}^{N}].$$

Now that we have showed the Markovian property for the process $(X_{t}^{1,N}, \mu_{t}^{N})$, we determine its dynamics by computing the infinitesimal generator. To this end, let us fix a function $F : [0,T] \times E \times \mathcal{S}^{N} \rightarrow \mathbb{R}$. Let $i \in E$ and $p \in \mathcal{S}^{N}$. Denote $n_{j} := N \cdot p_{j} - 1(i = j)$ the number of players in state $j$ for $i \in E$, without counting
player 1. Using the same argument of symmetry as above, we have:

\[
\mathbb{E}[F(t + h, X_{t+h}^{1,N}, \mu_{t+h}^N) | X_t^1 = i, \mu_t^N = p]
\]

\[
= \mathbb{E}[F(t + h, X_{t+h}^{1,N}, \mu_{t+h}^N) | X_t^1 = i, X_t^{2:(n_1+1),N} = 1, \ldots, X_t^{(n_1+\cdots+n_{m-1}+2):(n_1+\cdots+n_m+1),N} = m].
\]

Using the transition rate defined in equation (2.57), we obtain the infinitesimal generator \( G_{\phi,\phi} \) evaluated based on the following limit:

\[
G_{\phi,\phi} F(t, i, p) = \lim_{h \to 0} \frac{1}{h} \left[ \mathbb{E}[F(t + h, X_{t+h}^{1,N}, \mu_{t+h}^N) | X_t^{1,N} = i, \mu_t^N = p] - F(t, i, p) \right]
\]

\[
= \frac{d}{dt} F(t, i, p) + \sum_{j,j \neq i} (F(t, j, p + \frac{1}{N} (e_j - e_i))) - F(t, i, p)) \phi_j(t, i, p)
\]

\[
+ \sum_{(j,k), j \neq k} (F(t, i, p + \frac{1}{N} (e_j - e_k)) - F(t, i, p)) (N \cdot p_k - 1 (k = i)) \phi_j(t, k, p).
\]

Here the \( e_i \)'s are the basis vector of \( \mathbb{R}^m \). On the right hand side, the first summation corresponds to the change of states of player 1, while the second summation corresponds to the change of states of the remaining players. Having specified the dynamics of the deviating player and the empirical distribution of players’ states, we define the total expected cost of the deviating players:

\[
J_{\phi,\phi}^N(t, i, p) := \mathbb{E} \left[ \int_t^T f(s, X_s^{1,N}, \phi(s, X_s^{1,N}, \mu_s^N), \mu_s^N) + g(X_T^{1,N}, \mu_T^N) \right| X_t^{1,N} = i, \mu_t^N = p \]
\]

(2.59)

where \( f \) and \( g \) are respectively the running cost and terminal cost. To sum it up, given that all the other players use the Markov strategy \( \phi \), the deviating player solves the following control problem:

\[
V_{\phi}^N(t, i, p) := \inf_{\tilde{\alpha} \to \phi \in \mathcal{A}_{1,N}} J_{\phi,\phi}^N(t, i, p).
\]

(2.60)

Notice that the Markov process \( (X^{1,N}, \mu^N) \) takes values in a finite set \( \{1, \ldots, m\} \times S^N \) and therefore it is a continuous-time finite state Markov chain. The analytical approach to optimal control problem presented in Section 2.2.1 applies. Using the
infinitesimal generator of the process \((X^1, \mu^N)\) we just derived, it is straightforward to write down the HJB equation, which is a system of coupled ODEs:

\[
\frac{d}{dt} V^N(t, i, p) = -\inf_{\alpha \in A} \left\{ \sum_{j,j \neq i} \left[ V^N(t, j, p + \frac{1}{N}(e_j - e_i)) - V^N(t, i, p) \right] \alpha_j + f(t, i, \alpha, p) \right\} \\
- \sum_{(j,k), j \neq k} \left[ V^N(t, i, p + \frac{1}{N}(e_j - e_k)) - V^N(t, i, p) \right] (N \cdot p_k - 1(k = i)) \hat{\phi}_j(t, k, p),
\]

(2.61)

\[
V^N_T(t, i, p) = g(i, p).
\]

(2.62)

To identify the symmetric Nash equilibria, we search for a feedback function \(\hat{\phi}\) such that when all the remaining players use Markovian controls given by feedback function \(\hat{\phi}\), the optimal strategy of the deviating player is also \(\hat{\phi}\). Therefore if the feedback function \(\hat{\phi}\) leads to a symmetric Nash equilibrium, we must have:

\[
\hat{\phi}_j(t, i, p) = \hat{a}_j(t, i, V^N(t, \cdot, p + \frac{1}{N}(e_i - e_i)), p),
\]

(2.63)

where by abuse of notation, we denote by \(V^N(t, \cdot, p + \frac{1}{N}(e_i - e_i))\) the \(m\)-dimensional vector of which the \(j\)-th component is \(V^N(t, j, p + \frac{1}{N}(e_j - e_i))\). This leads to the following system of non-linear coupled ODEs:

\[
\frac{d}{dt} v^N(t, i, p) = -\sum_{j,j \neq i} [v^N(t, j, p + \frac{1}{N}(e_j - e_i)) - v^N(t, i, p)] \hat{\phi}_j(t, i, p) - f(t, i, \alpha, p)
\]

(2.64)

\[
- \sum_{(j,k), j \neq k} [v^N(t, i, p + \frac{1}{N}(e_j - e_k)) - v^N(t, i, p)] (N \cdot p_k - 1(k = i)) \hat{\phi}_j(t, k, p),
\]

\[
v^N_T(t, i, p) = g(i, p),
\]

(2.65)

\[
\hat{\phi}_j(t, i, p) = \hat{a}_j(t, i, v^N(t, \cdot, p + \frac{1}{N}(e_i - e_i)), p).
\]

(2.66)

To show the well-posedness of the system (2.64)-(2.66), we verify that the right hand side of equation (2.64) is Lipschitz in \(v^N\), and any local solution of the system (2.64)-
(2.66) is bounded by a uniform constant. Finally, we obtain the existence of the Nash equilibrium by a verification argument.

**Theorem 2.3.7.** There exists a symmetric Nash equilibrium for the \(N\)-player game. Let \(v^N\) be the unique solution to the system (2.64)-(2.66). Then for each \(n \in \{1, \ldots, N\}\), the total expected cost of the player \(n\) in equilibrium is given by \(v^N(0, i, p^o)\) if the player is in state \(i\) and the empirical distribution of players’s states is \(p^o\) at time 0, i.e. \(X_0^{n,N} = i\) and \(\frac{1}{N} \sum_{n=1}^{N} \mathbbm{1}(X_0^{n,N} = j) = p^o_j\). Moreover, the optimal strategy of the players in equilibrium is given by the feedback function:

\[
\hat{\phi}_j(t, i, p) = \hat{a}_j(t, i, v^N(t, \cdot, p + \frac{1}{N} (e_i - e_i)), p).
\]

**Remark 2.3.8.** In full analogy with \(N\)-player stochastic differential games studied in Chassagneux et al. [2014], when we let \(N\) tend to infinity in the system of ODEs (2.64)-(2.66), formally we should obtain the master equation of the finite state mean field game. The solution to the master equation gives the expected total cost of the Nash equilibrium at any given time, state and the distribution of the populations, and provides a way to compute the Nash equilibrium without searching for a fixed point. In the case of finite state mean field game, the master equation is a first-order partial differential equation, thus is much easier to handle numerically than the master equation for diffusion-based mean field game, where the space of distribution is of infinite dimension. When studying finite state mean field games with major and minor players in Chapter 5, we shall derive the master equation using the same technique.

### 2.3.4 Approximate Nash Equilibria

To conclude our presentation on the analytical approach to finite state mean field games, we state without proof the result on the convergence of the Nash equilibrium
of the $N$-player game to the Nash equilibrium of the mean field game. We refer the interested readers to Section 4 in Gomes et al. [2013] for detailed proofs.

Let us fix $p^o \in S$. At time $t = 0$, we assign states to the $N$ players randomly according to the distribution $p^o$ and these random assignments are independent. We denote $\hat{\mu}_0^N$ the resulting empirical distribution at time 0. These $N$ players then enter the game described in Section 2.3.3. Clearly, a symmetric Nash equilibrium is formed when each agent adopts the Markovian strategy specified in Theorem 2.3.7. Using the same argument of symmetry when we show the Markov property of the deviating player’s state and the empirical distribution, we can show that the flow of the empirical distribution $(\hat{\mu}_t^N)_{t \in [0,T]}$ in a symmetric Nash equilibrium is a continuous-time Markov chain. Now let us denote by $(v, p)$ the solution to the system (2.49) - (2.50) which characterizes the Nash equilibria of the mean field game. We also recall that $v^N$ is the solution to the system (2.64) - (2.66). We state the following result of convergence:

**Theorem 2.3.9.** There exists two positive constants $T^*$ and $C$, independent of $N$, such that if $T < T^*$ and $TC < 1$, then for all $t \in [0,T]$ we have:

$$
\mathbb{E} \left[ \|\hat{\mu}_t^N - p(t)\|_2^2 \right] + \mathbb{E} \left[ \|v^N(t, \cdot, \hat{\mu}_t^N) - v(t, \cdot)\|_2^2 \right] \leq \frac{C}{(1 - TC)N},
$$

(2.67)

where $\| \cdot \|_2$ is the Euclidean norm on $\mathbb{R}^m$.

**Remark 2.3.10.** Theorem 2.3.9 only provides the result of approximate Nash equilibria for games with duration smaller than $T^*$. A closer look at the proof of Theorem 2.3.9 suggests that the upper bound of the duration $T^*$ is a decreasing function of the Lipschitz constant of the terminal cost function $g$ (with regard to the argument of the measure $p$). It is not clear whether this upper bound is optimal, or whether the result of approximate Nash equilibria hold for any duration of the game.
2.4 Appendix: BSDE Driven by Single Continuous-Time Markov Chain

We recall several useful results on BSDEs driven by continuous-time Markov chain, which were developed in Cohen and Elliott [2008] and Cohen and Elliott [2010]. Recall that $M$ is the $\mathbb{P}$-martingale in the canonical decomposition of the Markov chain $X$ in (2.7). We consider the following BSDE with unknown $(Y,Z)$, where $Y$ is an adapted and càdlàg process in $\mathbb{R}$, and $Z$ is an adapted and left continuous process in $\mathbb{R}^m$:

$$Y_t = \xi + \int_t^T F(w,s,Y_s,Z_s)ds - \int_t^T Z_s^*dM_s. \quad (2.68)$$

Here $\xi$ is a $\mathcal{F}_T$-measurable $\mathbb{P}$-square integrable random variable and $F$ is the driver function such that the process $t \to F(w,t,y,z)$ is predictable for all $y,z$.

Recall the stochastic seminorm $\| \cdot \|_{X_{t-}}$ defined in (2.13). We have the following result on existence and uniqueness of the solution to the BSDE. See Theorem 1.1 in Cohen and Elliott [2010].

**Lemma 2.4.1.** Assume that there exists $C > 0$ such that $d\mathbb{P} \otimes dt$-a.e., for all $y_1, y_2 \in \mathbb{R}$ and $z_1, z_2 \in \mathbb{R}^m$ we have:

$$|F(w,t,y_1,z_1) - F(w,t,y_2,z_2)| \leq C(|y_1 - y_2| + \|z_1 - z_2\|_{X_{t-}}).$$

Then the BSDE (2.68) admits a solution $(Y,Z)$ satisfying

$$\mathbb{E}\left[ \int_0^T |Y_t|^2 dt \right] < +\infty, \quad \mathbb{E}\left[ \int_0^T \|Z_t\|_{X_{t-}}^2 dt \right] < +\infty.$$

In addition, the solution is unique in the sense that if $(Y^{(1)},Z^{(1)})$ and $(Y^{(2)},Z^{(2)})$ are two solutions, then $Y^{(1)}$ and $Y^{(2)}$ are indistinguishable and we have $\mathbb{E}[\int_0^T |Z_t^{(1)} - Z_t^{(2)}|^2_{X_{t-}} dt] = 0$. 

51
We also have the following stability property of the solution, which can be proved by mimicking the argument used in the proof of Theorem 2.1 in Hu and Peng [1997].

**Lemma 2.4.2.** For \( n \geq 0 \), let \((Y^{(n)}, Z^{(n)})\) be the solution to the BSDE with driver \( F^{(n)} \) and terminal condition \( \xi^{(n)} \). Assume that for all \( n \), \( F^{(n)} \) satisfies the Lipschitz condition in Lemma 2.4.1 and the following conditions hold:

(i) \( \lim_{n \to +\infty} \mathbb{E}[|\xi^{(n)} - \xi^{(0)}|^2] = 0 \).

(ii) For all \( t \leq T \), we have:

\[
\lim_{n \to +\infty} \mathbb{E}\left[ \left( \int_t^T (F^{(n)}(w, s, Y_s^{(0)}, Z_s^{(0)}) - F^{(0)}(w, s, Y_s^{(0)}, Z_s^{(0)}))ds \right)^2 \right] = 0.
\]

(iii) There exists \( C > 0 \) such that for all \( t \leq T \) and \( n \geq 0 \), we have:

\[
\mathbb{E}\left[ \left( \int_t^T (F^{(n)}(w, s, Y_s^{(0)}, Z_s^{(0)}) - F^{(0)}(w, s, Y_s^{(0)}, Z_s^{(0)}))ds \right)^2 \right] \leq C
\]

Then we have:

\[
\lim_{n \to +\infty} \mathbb{E}\left[ \int_t^T \|Z_s^{(n)} - Z_s^{(0)}\|_{\mathcal{X}_s}^2 ds \right] + \mathbb{E}[\|Y_t^{(n)} - Y_t^{(0)}\|^2] = 0.
\]

Finally, we state a result of comparison for linear BSDEs. See Theorem 3.16 in Cohen and Elliott [2010].

**Lemma 2.4.3.** Let \( \gamma \) be a bounded predictable process in \( \mathbb{R}^m \) and \( \beta \) be a bounded predictable process in \( \mathbb{R} \). Let \( \phi \) be a non-negative predictable process in \( \mathbb{R} \) such that \( \mathbb{E}[\int_0^T \|\phi_t\|^2 dt] < +\infty \) and \( \xi \) a non-negative square-integrable \( \mathcal{F}_T \) measurable random variable in \( \mathbb{R} \). Let \((Y, Z)\) be the solution of the linear BSDE:

\[
Y_t = \xi + \int_t^T (\phi_u + \beta_u Y_u + \gamma_u^* \cdot Z_u)du - \int_t^T Z_u^* \cdot d\mathcal{M}_u.
\]
Assume that for all $t \in (0, T]$ and $j$ such that $e_j^* \cdot Q^0 \cdot X_{t-} > 0$, we have:

$$1 + \gamma_t^* \cdot \psi_t^+ \cdot (e_j - X_{t-}) \geq 0,$$  

(2.70)

where $\psi_t^+$ is the Moore-Penrose inverse of the matrix $\psi_t$ defined in (2.12). Then $Y$ is nonnegative.
Chapter 3

Probabilistic Analysis of Finite State Mean Field Games

In this chapter, we construct a probabilistic framework for continuous-time mean field games in finite state space. We draw our inspiration from the weak formulation of stochastic differential mean field games developed in Carmona and Lacker [2015]. Evoking the notion of weak solutions of stochastic differential equations in which the probability measure and the underlying Wiener process both become part of the solution, the key idea is to view the optimization problem of a deviating player as controlling the probability law of the state process, as opposed to controlling the state process in a pathwise fashion. This perspective seems to be tailor-made to continuous-time Markov chains. Indeed, as we have seen in Chapter 2, the most convenient way to describe the dynamics of a continuous-time Markov chain is to specify the transition rates between the states, which is essentially to specify the probability distribution of the time until the next change of state, and the probability of each new possible state after the change.

Similar to the approach taken in Carmona and Lacker [2015], our starting point is the semimartingale representation of controlled continuous-time Markov chain in-
roduced in Section 2.1.2 and the weak formulation of optimal control presented in Section 2.2.2 of Chapter 2. As a reminder, we constructed a probability space of sample paths under which the canonical process $X$ has the representation $X_t = X_0 + \int_{[0,t]} Q^0 \cdot X_s - ds + M_t$ where $Q^0$ is the transition rate matrix of a continuous-time Markov chain with constant transition rate 1. By applying Girsanov theorem, we showed how to construct an equivalent probability measure under which the process $X$ has the desired controlled transition rates. We then identify the expected total cost under the controlled probability measure with the solution to a BSDE based on Markov chain, and the optimality of the control problem can be obtained by the comparison principle for linear BSDEs.

As we have already seen in Section 2.3 of Chapter 2, the paradigm of mean field games can be summarized as a two-step procedure. In the first step, we fix the mean field, i.e. the distribution of all players’ states, and try to derive the optimal strategy of a representative player in the game. Fixing the mean field makes perfect sense in the limit of infinitely many players, because the unilateral deviation of any single player has negligible impact on the mean field. In the second step, we solve the fixed point problem of Nash equilibrium by matching the mean field generated by the optimal strategy to the one we fix in the first step. Different from the analytical approach presented in Section 2.3, in the probabilistic approach we are about to present in this chapter, we achieve the first step through the weak formulation of optimal control. We then recast the search for the Nash equilibrium into a fixed point problem on the space of probability measures of sample of players’ state processes.

The advantage of the weak formulation over the traditional analytical approach are three-fold. First of all, the flexibility of the probabilistic approach allows us to consider mean field games where the interactions are not only realized through the states, but also through the strategies of the players. Mean field games with interactions through
the strategies of the players are known to be notoriously intractable via PDE/ODE methods due to the difficulty in deriving the equation obeyed by the flow of probability measure of the optimal control. Under the probabilistic framework however, the mean field of state and the mean field of control can be treated in similar manners, although the treatment of the mean field of control is more involved in terms of topological arguments. Secondly, using the weak formulation we are able to show Nash equilibria exist among all closed-loop strategies, including the strategies depending on the past history of player’s states, whereas the analytical approach presented in Chapter 2 can only accommodate Markovian controls.

Lastly, the weak formulation we develop for finite state mean field games will play an indispensable role when we tackle the finite state mean field principal agent problems in Chapter 4. This is a special model of Stackelberg game in which the principal fixes a contract first and a large population of competitive agents form a Nash equilibrium according to the contract proposed by the principal. Mathematically speaking, the contract is a control chosen by the principal which enters into each agent’s dynamics and cost functions. One meaningful direction in probing mean field principal agent problems is to understand how the principal can choose the optimal contract in order to minimize its own cost function which depends on the distribution of the agents’ states. To the best of our knowledge, this type of problems was first investigated in Elie et al. [2016] where agent’s dynamics are given by a controlled diffusion process. The main idea is to formulate the optimal contract problem as a McKean-Vlasov optimal control problem, in which the state process to be controlled is the McKean-Vlasov BSDE characterizing the Nash equilibrium in the weak formulation of the mean field game. We shall see in Chapter 4 that with the help of the weak formulation developed in this chapter, mean field principal agent problems in finite state space can be tackled in an analogous way. Moreover, thanks to the finite nature of the state space, the optimal contract can be computed with straightforward
numerical schemes, which could lead to potential applications in epidemic control and cyber security.

We would also like to mention some literature related to the weak probabilistic approach we present in this chapter. In Cecchin and Fischer [2017] the authors proposed a probabilistic framework for finite state mean field games where the state dynamics of each player are represented by stochastic differential equations driven by Poisson random measures. By using Ky Fan’s fixed point theorem, the authors obtained existence and uniqueness of Nash equilibrium in relaxed open-loop controls as well as relaxed feedback controls. Then under a stronger assumption that guarantees uniqueness of optimal non-relaxed feedback control, the authors deduced existence of Nash equilibria in non-relaxed feedback form. In Doncel et al. [2016], continuous-time mean field games with finite state space and finite action space were studied. The authors proved existence of Nash equilibria among relaxed feedback controls. In Benazzoli et al. [2017] the authors investigated mean field games where each player’s state follows a jump-diffusion process and the player controls the sizes of the jumps. Their approach is based on weak formulation of stochastic controls and martingale problems. Existence of Nash equilibria among relaxed controls and Markov controls was established.

The rest of this chapter is organized as follows. In Section 3.1, we introduce the weak formulation of finite state mean field games, which is based on the semimartingale representation of continuous-time Markov chains and an argument of change of measure. We state the assumptions of the model and give a precise definition of Nash equilibria in the weak formulation. We also analyze the representative player’s optimal control problem when facing a fixed mean field of state and control, as a straightforward application of weak formulation of optimal control introduced in Section 2.2 of Chapter 2. Section 3.2 and Section 3.3 are devoted to the existence and
the uniqueness of Nash equilibria. Finally in Section 3.4, we formulate the game with finitely many players and show the Nash equilibrium of the mean field game is an approximative Nash equilibrium of the game with finite number of players.

3.1 Weak Formulation of Finite State Mean Field Games

3.1.1 Mean Field Games with Interaction through States and Controls

We consider a large population of players whose states belong to a finite state space $E$ constituted of $m$ elements. These states are identified with the $m$ basis vectors of $\mathbb{R}^m$, which are denoted by $e_1, \ldots, e_m$. We denote by $\mathcal{S}$ the $m$-dimensional simplex $\mathcal{S} := \{p \in \mathbb{R}^m, \sum p_i = 1, p_i \geq 0\}$ equipped with the euclidean norm $\| \cdot \|$, which we identify with the space of probability distribution on $E$.

Let $\Omega$ be the space of càdlàg functions from $[0, T]$ to $E$ that are continuous on $T$, and $X$ be the canonical process. Later we will use $X$ to represent the state process of a representative player. We denote by $\mathcal{F}_t := \sigma(\{X_s, s \leq t\})$ the natural filtration generated by $X$ and $\bar{\mathcal{F}} := \mathcal{F}_T$. Now let us fix $p^0 \in \mathcal{S}$ a probability measure on $E$. On $(\Omega, (\mathcal{F}_t)_{t \in [0, T]}, \bar{\mathcal{F}})$, we consider the probability measure $\mathbb{P}$ under which $X$ is the continuous-time Markov chain such that the law of $X_0$ is $p^0$ and the transition rate between any two different states is 1. We recall that under $\mathbb{P}$ the process $X$ has the following representation:

$$X_t = X_0 + \int_{[0,t]} Q^0 \cdot X_s - ds + \mathcal{M}_t, \quad (3.1)$$

where $\mathcal{M}$ is a $\mathbb{P}$-martingale and $Q^0$ is a $m$ by $m$ matrix with diagonal elements all equal to $-(m - 1)$ and off-diagonal elements all equal to 1.
Let $A$ be a compact subset of $\mathbb{R}^l$ in which the players can choose their controls. Denote by $\mathcal{P}(A)$ the space of probability measures on $A$. We introduce a function $q$:

$$[0, T] \times \{1, \ldots, m\}^2 \times A \times \mathcal{S} \times \mathcal{P}(A) \to q(t, i, j, \alpha, p, \nu),$$

(3.2)

and we denote by $Q(t, \alpha, p, \nu)$ the matrix $[q(t, i, j, \alpha, p, \nu)]_{1 \leq i, j \leq m}$. Throughout this chapter, we make the following assumption on $q$:

**Assumption 3.1.1.** For all $(t, \alpha, p, \nu) \in [0, T] \times A \times \mathcal{S} \times \mathcal{P}(A)$, the matrix $Q(t, \alpha, p, \nu)$ is a Q-matrix. (See Definition 2.1.2)

We now describe how the player’s control and the distribution of all the players’ states affect the probability law of $X$ in our model. Let us define the player’s strategy set $A$ to be the collection of $\mathcal{F}$-predictable processes $(\alpha_t)_{t \in [0, T]}$ such that $\alpha_t \in A$ for $0 < t \leq T$. Given $(p_t)_{t \in [0, T]}$, a flow of measures on $\{1, \ldots, m\}$, and $(\nu_t)_{t \in [0, T]}$, a flow of measures on $A$, we define the scalar martingale $L^{(\alpha, p, \nu)}$ under $\mathbb{P}$ by:

$$L^{(\alpha, p, \nu)}_t := \int_0^t X^*_s \cdot (Q(s, \alpha_s, p_s, \nu_s) - Q^0) \cdot \psi^+_s \cdot d\mathcal{M}_s,$$

(3.3)

where we recall that $\psi_t$ is the density of the predictable quadratic variation of the martingale $\mathcal{M}$ under $\mathbb{P}$ as defined in equation (2.12), and $\psi^+_t$ is the Moore-Penrose inverse of $\psi_t$. We then define the measure $Q^{(\alpha, p, \nu)}$ through the Doléans-Dade exponential of $L^{(\alpha, p, \nu)}$:

$$\frac{dQ^{(\alpha, p, \nu)}}{d\mathbb{P}} := \mathcal{E}(L^{(\alpha, p, \nu)})_T.$$

(3.4)

Following the same arguments developed in Section 2.1.2, Chapter 2, we know that the process $\mathcal{M}^{(\alpha, p, \nu)}$, defined as:

$$\mathcal{M}^{(\alpha, p, \nu)}_t := \mathcal{M}_t - \int_0^t (Q^*(s, \alpha_s, p_s, \nu_s) - Q^0) \cdot X_{s-} ds,$$

(3.5)
is a $Q^{(\alpha,p,\nu)}$-martingale. Therefore the canonical decomposition of $X$ can be rewritten as:

$$X_t = X_0 + \int_{[0,t]} Q^*(s, \alpha_s, p_s, \nu_s) \cdot X_s \, ds + M^{(\alpha,p,\nu)}_t.$$ (3.6)

This means that under the measure $Q^{(\alpha,p,\nu)}$, the stochastic intensity rate of $X$ is given by $Q(t, \alpha_t, p_t, \nu_t)$. In particular, when $\alpha$ is a Markovian control, i.e. of the form $\alpha_t = \phi(t, X_t)$ for some measurable feedback function $\phi$, $X$ is a continuous-time Markov chain with intensity rate $q(t, i, j, \phi(t, i), p_t, \nu_t)$ under the measure $Q^{(\alpha,p,\nu)}$. Moreover, we have $Q^{(\alpha,p,\nu)} = \mathbb{P}$ on $\mathcal{F}_0$, and therefore the distribution of $X_0$ under $Q^{(\alpha,p,\nu)}$ is $p^0$.

Let $f : [0,T] \times E \times A \times S \times \mathcal{P}(A) \to \mathbb{R}$ and $g : E \times S \to \mathbb{R}$ be respectively the running cost function and the terminal cost function. When the player picks a control $\alpha \in \mathbb{A}$ and the mean field is $(p, \nu)$, its cost is:

$$J(\alpha, p, \nu) := \mathbb{E}^{Q^{(\alpha,p,\nu)}} \left[ \int_0^T f(t, X_t, \alpha_t, p_t, \nu_t) \, dt + g(X_T, p_T) \right]$$ (3.7)

Each player aims at minimizing $J(\alpha, p, \nu)$, that is, it solves the optimization problem:

$$V(p, \nu) := \inf_{\alpha \in \mathbb{A}} \mathbb{E}^{Q^{(\alpha,p,\nu)}} \left[ \int_0^T f(t, X_t, \alpha_t, p_t, \nu_t) \, dt + g(X_T, p_T) \right]$$ (3.8)

In the classical setup of mean field games, a single player’s strategy $\alpha$ does not alter the mean field $(p, \nu)$. Therefore when each player solves its own optimization problem, it considers $(p, \nu)$ as given. A Nash equilibrium is then reached when the law of $X_t$ under the player’s controlled probability measure, along with the law of its control at time $t$ under the same probability measure, coincides with $(p_t, \nu_t)$ for all $t \leq T$. Accordingly, we have the following definition of a Nash equilibrium for the weak formulation of finite state mean field games.
**Definition 3.1.2.** Let $p^* : [0, T] \to \mathcal{S}$, $\nu^* : [0, T] \to \mathcal{P}(A)$ be two measurable functions and $\alpha^* \in \mathbb{A}$. We say that the tuple $(\alpha^*, p^*, \nu^*)$ is a Nash equilibrium for the weak formulation of finite state mean field game if:

(i) $\alpha^*$ minimizes the cost when the mean field is given by $(p^*, \nu^*)$:

$$
\alpha^* \in \arg \inf_{\alpha \in \mathbb{A}} \mathbb{E}^{Q(\alpha, p^*, \nu^*)} \left[ \int_0^T f(t, X_t, \alpha_t, p^*_t, \nu^*_t) dt + g(X_T, p^*_T) \right].
$$

(ii) $(\alpha^*, p^*, \nu^*)$ satisfies the consistency conditions:

\begin{align}
 p^*_t &= Q(\alpha^*, p^*, \nu^*)[X_t = e_i], \forall t \leq T, \quad (3.10) \\
 \nu^*_t &= Q(\alpha^*, p^*, \nu^*)#\alpha^*_t, \forall t \leq T. \quad (3.11)
\end{align}

Throughout the rest of this chapter, we make the following assumptions on the transition rate $q$ and the cost functionals $f$ and $g$.

**Assumption 3.1.3.** There exists $C_1, C_2 > 0$ such that for all $t, i, j, \alpha, p, \nu \in [0, T] \times \mathbb{E}^2 \times \mathbb{A} \times \mathcal{S} \times \mathcal{P}(A)$ with $i \neq j$, we have $0 < C_1 < q(t, i, j, \alpha, p, \nu) < C_2$.

**Assumption 3.1.4.** There exists a constant $C > 0$ such that for all $(t, i, j) \in [0, T] \times \mathbb{E}^2$, $\alpha, \alpha' \in \mathbb{A}$, $p, p' \in \mathcal{S}$ and $\nu, \nu' \in \mathcal{P}(A)$, we have:

\begin{align}
|q(t, i, j, \alpha, p, \nu) - q(t, i, j, \alpha', p', \nu')| &\leq C(\|\alpha - \alpha'\| + \|p - p'\| + \mathcal{W}_1(\nu, \nu')), \quad (3.12) \\
|f(t, e_i, \alpha, p, \nu) - f(t, e_i, \alpha', p', \nu')| &\leq C(\|\alpha - \alpha'\| + \|p - p'\| + \mathcal{W}_1(\nu, \nu')), \quad (3.13) \\
|g(e_i, p) - g(e_i, p')| &\leq C\|p - p'\|. \quad (3.14)
\end{align}

**Remark 3.1.5.** Assumption 3.1.3 is analogous to the non-degeneracy condition in the diffusion-based mean field game model. It guarantees that the probability measure $Q(\alpha, p, \nu)$ is equivalent to the reference measure $\mathbb{P}$.

**Remark 3.1.6.** In certain application of continuous-time Markov chain models, it is possible that the jumps from some states to others are forbidden. That is we consider
a transition rate function $q$ such that $q(t, i, j, \alpha, p, \nu) \equiv 0$ for some tuples $(i, j)$. This is indeed the case in the botnet defense model proposed by Kolokoltsov and Bensoussan [2016] as well as in an extended version of model including an attacker studied in Carmona and Wang [2016]. If this were the case, we need to define a different reference probability $P$: we set the transition rate to 1 for the jumps, except those jumps that are forbidden, for which we set the transition rate to 0. Fortunately, this is the only modification we need to make in order to accommodate for this kind of special case. The arguments presented in the following can be trivially extended to be compatible with this modified reference probability.

### 3.1.2 Representative Player’s Optimization Problem

We now apply the weak formulation introduced in Chapter 2 to tackle a representative player’s optimization problem. As before, we define the Hamiltonian $H : [0, T] \times E \times \mathbb{R}^m \times A \times S \times \mathcal{P}(A) \to \mathbb{R}$ by:

$$H(t, x, z, \alpha, p, \nu) := f(t, x, \alpha, p, \nu) + x^* \cdot (Q(t, \alpha, p, \nu) - Q^0) \cdot z,$$  \hspace{1cm} (3.15)

as well as the reduced Hamiltonian:

$$H_i(t, z, \alpha, p, \nu) := H(t, e_i, z, \alpha, p, \nu).$$  \hspace{1cm} (3.16)

We make the following assumption on the minimizer of the reduced Hamiltonian.

**Assumption 3.1.7.** (i) For any $t \in [0, T]$, $i \in \{1, \ldots, m\}$, $z \in \mathbb{R}^m$, $p \in S$ and $\nu \in \mathcal{P}(A)$, the mapping $\alpha \to H_i(t, z, \alpha, p, \nu)$ admits a unique minimizer which does not depend on the mean field of control $\nu$. We denote the minimizer by $\hat{a}_i(t, z, p)$.
(ii) \( \hat{a}_i \) is measurable in \([0, T] \times \mathbb{R}^m \times \mathcal{S} \). There exist constants \( C_1 > 0 \) and \( C_2 \geq 0 \) such that for all \( i \in \{1, \ldots, m\} \), \( z, z' \in \mathbb{R}^m \), \( p, p' \in \mathcal{S} \):

\[
\| \hat{a}_i(t, z, p) - \hat{a}_i(t, z', p') \| \leq C_1 \| z - z' \| e_i + (C_1 + C_2 \| z \| e_i) \| p - p' \|. \tag{3.17}
\]

We denote by \( \hat{H}_i \) the minimum of the reduced Hamiltonian: \( \hat{H}_i(t, Z, p, \nu) := \hat{H}_i(t, Z, \hat{\alpha}(t, Z, p), p, \nu) \). Now we define the mapping \( \hat{H} \) and \( \hat{a} \) by:

\[
\hat{H}(t, X, Z, p, \nu) := \sum_{i=1}^{m} 1(X = e_i) \cdot \hat{H}_i(t, Z, p, \nu), \tag{3.18}
\]

\[
\hat{a}(t, X, Z, p) := \sum_{i=1}^{m} 1(X = e_i) \cdot \hat{a}_i(t, Z, p). \tag{3.19}
\]

**Remark 3.1.8.** For the sake of convenience and generality, we choose to make the assumption directly on the uniqueness and the regularity of the minimizer of the Hamiltonian. Alternatively, we could also impose structural constraints on the transition rate function \( q \) and require strong convexity for the running cost function \( f \). For example, we can make the following assumption:

**Assumption 3.1.9.** (i) \( A \) is a convex and compact subset of \( \mathbb{R}^l \)

(ii) The transition rate function \( q \) takes the form \( q(t, i, j, \alpha, p, \nu) = q_0(t, i, j, p, \nu) + q_1(t, i, j, p) \cdot \alpha \), where \( q_0 : [0, T] \times E^2 \times \mathcal{S} \times \mathcal{P}(A) \to \mathbb{R} \) and \( q_1 : [0, T] \times E^2 \times \mathcal{S} \to \mathbb{R}^l \) are two continuous mapping.

(iii) The running cost function \( f \) is of the form \( f(t, X, \alpha, p, \nu) = f_0(t, X, \alpha, p) + f_1(t, X, p, \nu) \), where for each \( i \in \{1, \ldots, m\} \), the mapping \( f_0(\cdot, e_i, \cdot, \cdot) \) (resp. \( f_1(\cdot, e_i, \cdot, \cdot) \)) is continuous in \([0, T] \times A \times \mathcal{P} \) (resp. \([0, T] \times \mathcal{S} \times \mathcal{P}(A) \)).

(iv) For all \((t, i, p) \in [0, T] \times E \times \mathcal{S} \), the mapping \( \alpha \to f_0(t, e_i, \alpha, p) \) is once continuous differentiable. Moreover there exists a constant \( C > 0 \) such that:

\[
\| \nabla_{\alpha} f_0(t, e_i, \alpha, p) - \nabla_{\alpha} f_0(t, e_i, \alpha, p') \| \leq C \| p - p' \|. \tag{3.20}
\]
(v) \( f_0 \) is \( \gamma \)-strongly convex in \( \alpha \), i.e., for all \((t, i, p) \in [0, T] \times E \times S\) and \( \alpha, \alpha' \in A \), we have:

\[
f_0(t, e_i, \alpha, p) - f_0(t, e_i, \alpha', p) - (\alpha - \alpha') \cdot \nabla_\alpha f_0(t, e_i, \alpha, p) \geq \gamma \| \alpha' - \alpha \|^2
\]  

(3.21)

Under Assumption 3.1.7, it is clear from the definition of \( \hat{a} \) in equation (3.19) that \( \hat{a}(t, x, z, p) \) is the unique minimizer of the mapping \( \alpha \rightarrow H(t, x, z, \alpha, p, \nu) \), and the minimum is given by \( \hat{H}(t, x, z, p, \nu) \). In addition, from Assumption 3.1.3 and Assumption 3.1.4, we can show the following result on the Lipschitz property of \( \hat{H} \):

**Lemma 3.1.10.** There exists a constant \( C > 0 \) such that for all \((\omega, t) \in \Omega \times (0, T] \), \( p, p' \in S \), \( \nu, \nu' \in \mathcal{P}(A) \) and \( z, z' \in \mathbb{R}^m \), we have:

\[
|\hat{H}(t, X_{t-}, z, p, \nu) - \hat{H}(t, X_{t-}, z', p', \nu')| \leq C\|z - z'\|_{X_{t-}} + C(1 + \|z\|_{X_{t-}})(\|p - p'\| + W_1(\nu, \nu')),
\]  

(3.22)

\[
|\hat{a}(t, X_{t-}, z, p) - \hat{a}(t, X_{t-}, z', p')| \leq C\|z - z'\|_{X_{t-}} + C(1 + \|z\|_{X_{t-}})\|p - p'\|.
\]  

(3.23)

**Proof.** Inequality (3.23) is an easy consequence of Assumption 3.1.7 and the definition of the stochastic seminorm \( \| \cdot \|_{X_{t-}} \). We now deal with the regularity of \( \hat{H} \). By Berge's maximum theorem, the continuity of \( H_i \) and the compactness of \( A \) imply the continuity of \( \hat{H}_i \). Let \( z, z' \in \mathbb{R}^m \), \( p, p' \in S \) and \( \nu, \nu' \in \mathcal{P}(A) \). For any \( \alpha \in A \), we have:

\[
\hat{H}_i(t, z, p, \nu) - H_i(t, z', \alpha, p', \nu') 
\leq H_i(t, z, \alpha, p, \nu) - H_i(t, z', \alpha, p', \nu') 
\]

\[
= f(t, e_i, \alpha, p, \nu) - f(t, e_i, \alpha, p', \nu') + \sum_{j \neq i} [(z_j - z_i) - (z'_j - z'_i)]q(t, i, j, \alpha, p', \nu') 
\]

\[
+ \sum_{j \neq i} (z_j - z_i)[q_0(t, i, j, p, \nu) - q_0(t, i, j, p', \nu')] + (z_j - z_i)[q_1(t, i, j, p) - q_1(t, i, j, p')],
\]

\[
\leq C\|z - z'\|_{e_i} + C(1 + \|z\|_{e_i})(\|p - p'\| + W_1(\nu, \nu')).
\]  

64
where we have used the Lipschitz property of \( f \) and \( q \), and the boundedness of \( A \) and \( q \). Since the above is true for all \( \alpha \in A \), taking supremum of the left-hand side, we obtain:

\[
    \hat{H}_i(t, z, p, \nu) - \hat{H}_i(t, z', p', \nu') \leq C\|z - z'\|_{\epsilon_i} + C(1 + \|z\|_{\epsilon_i})(\|p - p'\| + \mathcal{W}_1(\nu, \nu')).
\]

Exchanging the roles of \( z \) and \( z' \), we obtain:

\[
    |\hat{H}_i(t, z, p, \nu) - \hat{H}_i(t, z', p', \nu')| \leq C\|z - z'\|_{\epsilon_i} + C(1 + \|z\|_{\epsilon_i})(\|p - p'\| + \mathcal{W}_1(\nu, \nu')),
\]

and (3.22) follows immediately from the definition of the seminorm \( \|\cdot\|_{X_t} \).

Let us fix measurable functions \( p : [0, T] \to S \) and \( \nu : [0, T] \to \mathcal{P}(A) \). We consider the following BSDE:

\[
    Y_t = g(X_T, p_T) + \int_t^T \hat{H}(s, X_s, Z_s, p_s, \nu_s)ds - \int_t^T Z_s^* \cdot dM_s.
\]

Piggybacking on the argument developed in Section 2.2.2 of Chapter 2, in particular Proposition 2.2.7 and 2.2.9, we show that the above BSDE characterizes the optimality of the control problem (3.8), and the optimal control is unique.

**Proposition 3.1.11.** For any measurable function \( p \) from \([0, T]\) to \( S \) and any measurable function \( \nu \) from \([0, T]\) to \( \mathcal{P}(A) \), the BSDE (3.24) admits a unique solution \((Y, Z)\). The value function of the optimal control problem (3.8) is given by \( V(p, \nu) = \mathbb{E}^\mathbb{P}[Y_0] \) and the control \( \hat{\alpha}^{(p, \nu)} \) defined by:

\[
    \hat{\alpha}_t^{(p, \nu)} = \hat{a}(t, X_t, Z_t, p_t).
\]

is optimal. In addition, if \( \alpha' \in \mathbb{A} \) is an optimal control, we have \( \alpha'_t = \hat{\alpha}_t^{(p, \nu)}, \, dt \otimes d\mathbb{P} - \text{a.e.} \)
3.2 Existence of Nash Equilibria

**Theorem 3.2.1.** Under Assumption 3.1.3, 3.1.4 and 3.1.7, there exists a Nash equilibrium \((\alpha^*, p^*, \nu^*)\) for the weak formulation of the finite-state mean field game in the sense of Definition 3.1.2.

The rest of this section is devoted to the proof of Theorem 3.2.1. As in the case of diffusion-based mean field games, we shall rely on an argument of fixed point to show the existence of Nash equilibria. We start from a measurable function \(p : [0, T] \to S\) and a measurable function \(\nu : [0, T] \to \mathcal{P}(A)\) where we recall that \(S\) is the \(m\)-dimensional simplex isomorphic to the space of probability measures on \(E\), while \(\mathcal{P}(A)\) is the space of probability measures on \(A\). We then solve the BSDE (3.24) which allows us to obtain the solution \((Y^{(p, \nu)}, Z^{(p, \nu)})\) as well as the optimal control \(\hat{\alpha}^{(p, \nu)}\) given by (3.25). Finally, we compute the probability measure \(\hat{Q}^{(p, \nu)} := Q^{(\hat{\alpha}^{(p, \nu)}, p, \nu)}\) as defined in (3.4) and consider the push-forward measures of \(\hat{Q}^{(p, \nu)}\) by \(X_t\) as well as \(\hat{\alpha}^{(p, \nu)}_t\). Clearly, a Nash equilibrium exists if the mapping \((p, \nu) \to (\hat{Q}^{(p, \nu)}_{\#X_t}, \hat{Q}^{(p, \nu)}_{\#\hat{\alpha}^{(p, \nu)}_t})\) admits a fixed-point.

In practice however, the implementation of the fixed-point argument mentioned above is prone to several difficulties. The foremost challenge lies in the fact that the space of the mean fields of states and controls, or more specifically the space of measurable mappings from \([0, T]\) to \(S\) or to \(\mathcal{P}(A)\), where \(p\) or \(\nu\) belong respectively, does not have a convenient topological description to grant the compactness. This makes it difficult to apply Schauder’s theorem or similar versions of fixed point theorems.

Due to the reasons mentioned above, we shall resort to different descriptions of the mean fields for the states and the controls. For the mean field of the states, since we have assumed from the very beginning that \(X\) is a càdlàg process, we will directly deal with the probability measures on \(D\), which is the space of all càdlàg functions from \([0, T]\) to \(E = \{e_1, \ldots, e_m\}\) endowed with the Skorokhod topology. The
space of probability measures on $D$ and its topological properties have been studied thoroughly (see Jacod and Shiryaev [1987] for a detailed account). In particular a simple criterion for compactness is available.

On the other hand, the construction of the space for mean field on control is more involved. Here, we adopt the same technique as used in Carmona and Lacker [2015] based on the stable topology. Indeed, a measurable function from $[0, T]$ to $\mathcal{P}(A)$ can be viewed as a random variable defined on the space $([0, T], \mathcal{B}([0, T]), \mathcal{L})$ taking values in $\mathcal{P}(A)$, where $\mathcal{B}([0, T])$ is the collection of Borel subsets of $[0, T]$, $\mathcal{L}$ is the uniform distribution on $[0, T]$ and $\mathcal{P}(A)$ is endowed with the Wasserstein-1 distance. To obtain compactness, the idea is to use randomization. We consider the space of probability measures on $[0, T] \times \mathcal{P}(A)$, denoted by $\mathcal{P}([0, T] \times \mathcal{P}(A))$. Then for each measurable function $\nu$ from $[0, T]$ to $\mathcal{P}(A)$, we consider the measure $\eta$ on $[0, T] \times \mathcal{P}(A)$ given by $\eta(dt, dm) := \mathcal{L}(dt) \times \delta_{\nu_t}(dm)$ where $\delta$ is the Dirac measure. We may endow the space $\mathcal{P}([0, T] \times \mathcal{P}(A))$ with the so-called stable topology proposed by Jacod and Méméin [1981], for which convenient results on compactness are readily available.

In the following, we detail the steps that lead to the existence of Nash equilibria. We start by specifying the topology in the space of mean fields of the states as well as the controls. We then properly define the mapping compatible with the definition of the Nash equilibrium. We show the continuity and construct a compact subset that is stable by the mapping. Once these ingredients are in place, we apply Schauder’s fixed point theorem to conclude.

### 3.2.1 Topology for the Mean Field

We start by defining the topological space for the mean field of the states. To this end, we first endow the space of states $E := \{e_1, \ldots, e_m\}$ with the discrete metric $d_E(x, y) := 1(x \neq y)$. It is well known that $(E, d_E)$ is a Polish space. Now let’s define
the Skorokhod space $D$:

$$D := \{ x : [0, T] \to E, x \text{ is càdlàg and left continuous on } T \}. \quad (3.26)$$

We endow $D$ with the $J_1$ metric defined as follows:

$$d_D(x, y) := \inf_{\lambda \in \Lambda} \max \left\{ \sup_{t \in [0, T]} |\lambda(t) - t|, \sup_{t \in [0, T]} |y(\lambda(t)) - x(t)| \right\}, \quad (3.27)$$

where $\Lambda$ is the set of all strictly increasing, continuous bijections from $[0, T]$ to itself.

It can be proved that $d_D$ is a metric on $D$ and the metric space $(D, d_D)$ is a Polish space. Let us denote by $\mathcal{P}$ the collection of probability measures on $(D, d_D)$ endowed with the weak topology. Recall that the reference measure $\mathbb{P}$ defined in the very beginning is an element of $\mathcal{P}$. Consider $\mathcal{P}_0$ a subset of $\mathcal{P}$ defined as:

$$\mathcal{P}_0 := \{ Q : \frac{dQ}{d\mathbb{P}} = L, \text{ with } \mathbb{E}^\mathbb{P}[L^2] \leq C_0 \}, \quad (3.28)$$

where $C_0$ is a constant which we will choose later (see the proof of Proposition 3.2.10).

We have the following result:

**Proposition 3.2.2.** $\mathcal{P}_0$ is convex and relatively compact in $\mathcal{P}$.

**Proof.** The convexity of $\mathcal{P}_0$ is trivial. Let us show that $\mathcal{P}_0$ is relatively compact. We proceed in three steps.

**Step 1** For $K \in \mathbb{N}$ and $\delta > 0$, we define $D_{\delta,K}$ the collection of paths in $D$ which meets the following criteria: (a) the path has no more than $K$ discontinuities, (b) the first jump time if there is any happens on or after $\delta$, (c) the last jump happens on or before $T - \delta$, and (d) the time between jumps are greater or equal than $\delta$. We now show that $D_{\delta,K}$ is compact in $D$. Since $D$ is a Polish space it is enough to show the sequential compactness. Let us fix a sequence $x_n$ in $D_{\delta,K}$. For each $x_n$, we use the following notation: $k_n$ is the number of its jumps, $\delta \leq t_1^n < t_2^n < \cdots < t_{k_n}^n \leq T - \delta$ are the moments of its jumps. $\Delta t_1^n := t_1^n$ and $\Delta t_i^n := t_i^n - t_{i-1}^n$ for $i = 2, \ldots, k_n$ are the
time elapsed between consecutive jumps and \( x_n^0, x_n^1, \ldots, t_n^{k_n} \) are the value taken by \( x_n \) in each interval defined by the jumps. Then we can represent \( x_n \) using the vector \( y_n \) of dimension \( 2(K + 1) \):

\[
y_n = [k_n, \Delta t_n^1, \Delta t_n^2, \ldots, \Delta t_n^{k_n}, 0, \ldots, 0, x_n^0, x_n^1, \ldots, x_n^{k_n}, 0, \ldots, 0].
\]

In the above representation, the first coordinate of \( y_n \) is the number of jumps. Coordinate 2 to \( K + 1 \) are the time elapsed between jumps defined above, and if there are fewer than \( K \) jumps, we complete the vector by 0. Coordinates \( K + 2 \) to \( 2(K + 1) \) are the values taken by the path \( x \) and completed with 0. Clearly there is a bijection from \( x_n \) to \( y_n \) by this representation. By the definition of the set \( D_{\delta,K} \), we have \( \Delta t_n^i \in [\delta, T] \) for \( i \leq k_n \) and \( \sum_{i=1}^{k_n} \Delta t_n^i \leq T - \delta \), whereas the rest of the coordinates of \( y_n \) belong to a finite set. This implies that \( y_n \) lives in a compact set and therefore we can extract a converging subsequence which we still denote by \( y_n \). Again, since \( k_n \) and the last \( K + 1 \) components can only take finitely many values by their definition, there exists \( N_0 \) such that for \( n \geq N_0 \), we have \( k_n = k \) and \( x_n^i = x^i \) for all \( i \leq k \). In addition we have \( \Delta t_n^i \) converges to \( \Delta t^i \) for all \( i \leq k \), where \( \Delta t^i \geq \delta \) for all \( i \leq k \) and \( \sum_{i=1}^{k} \Delta t^i \leq T - \epsilon \). We consider the path represented by the vector \( y \):

\[
y = [k, \Delta t^1, \Delta t^2, \ldots, \Delta t^k, 0, \ldots, 0, x^0, x^1, \ldots, x^k, 0, \ldots, 0]
\]

Clearly \( x \) belongs to the set \( D_{\delta,K} \) and it is straightforward to verify that \( x_n \) converges to \( x \) in the \( J_1 \) metric, where \( x_n \) is the path represented by the vector \( y_n \). This implies that \( D_{\delta,K} \) is compact.

**Step 2** Now we show that for any \( \epsilon > 0 \), there exists \( \delta > 0 \) and \( K \in \mathbb{N} \) such that \( \mathbb{P}(D_{\delta,K}) \geq 1 - \epsilon \). Recall that \( \mathbb{P} \) is the reference measure and under \( \mathbb{P} \) the canonical process \( X \) is a continuous-time Markov chain with transition rate matrix \( Q^0 \). Therefore the time of the first jump, as well as the time between consecutive jumps thereafter, which we denote by \( \Delta t_1, \Delta_2, \ldots \) are i.i.d. exponential random
variables of parameter \((m - 1)\) under the measure \(\mathbb{P}\). We have:

\[
\mathbb{P}(D_{\delta,K}) = \mathbb{P}[\Delta t_1 > T] + \sum_{k=1}^{K} \mathbb{P} \left[ \{\Delta t_1 \geq \delta\} \cap \cdots \cap \{\Delta t_k \geq \delta\} \cap \left\{ \sum_{i=1}^{k+1} \Delta t_i > T \right\} \cap \left\{ \sum_{i=1}^{k} \Delta t_i \leq T - \delta \right\} \right].
\]

For each \(k = 1, \ldots, K\), we have:

\[
\mathbb{P} \left[ \{\Delta t_1 \geq \delta\} \cap \cdots \cap \{\Delta t_k \geq \delta\} \cap \left\{ \sum_{i=1}^{k+1} \Delta t_i > T \right\} \cap \left\{ \sum_{i=1}^{k} \Delta t_i \leq T - \delta \right\} \right] \geq \mathbb{P} \left[ \{\Delta t_1 \geq \delta\} \cap \cdots \cap \{\Delta t_k \geq \delta\} \right] + \mathbb{P} \left[ \left\{ \sum_{i=1}^{k+1} \Delta t_i > T \right\} \cap \left\{ \sum_{i=1}^{k} \Delta t_i \leq T - \delta \right\} \right] - 1
\]

\[
= \left( \mathbb{P}[\Delta t_1 \geq \delta] \right)^k + \mathbb{P} \left[ \left\{ \sum_{i=1}^{k+1} \Delta t_i > T \right\} \cap \left\{ \sum_{i=1}^{k} \Delta t_i \leq T - \delta \right\} \right] - 1
\]

\[
= (\exp(-k(m - 1)\delta) - 1) + \exp(-(m - 1)T) \frac{(m - 1)^k(T - \delta)^k}{k!}.
\]

It follows that:

\[
\mathbb{P}(D_{\delta,K}) \geq \sum_{k=1}^{K} (\exp(-k(m - 1)\delta) - 1) - 1 + \exp(-(m - 1)T) \sum_{k=0}^{K} \frac{(m - 1)^k(T - \delta)^k}{k!}
\]

\[
\geq \sum_{k=1}^{K} (\exp(-k(m - 1)\delta) - 1) + \exp(-(m - 1)T) \sum_{k=0}^{K} \frac{(m - 1)^kT^k}{k!} - (1 - \exp(-(m - 1)\delta)).
\]

We can first pick \(K\) large enough such that \((\exp(-(m - 1)T) \sum_{k=0}^{K} \frac{(m - 1)^kT^k}{k!})\) is greater than \(1 - \epsilon/2\) and then pick \(\delta\) small enough to make the rest of the terms greater than \(-\epsilon/2\), which eventually makes \(\mathbb{P}(D_{\delta,K})\) greater than \(1 - \epsilon\).

**Step 3** Finally, we show that \(\mathcal{P}_0\) is tight. For any \(\epsilon > 0\), by Step 2 we can pick \(\delta > 0\) and \(K \in \mathbb{N}\) such that \(\mathbb{P}(D \setminus D_{\delta,K}) \leq (\epsilon/C_0)^2\). For all \(Q \in \mathcal{P}_0\), we have \(dQ/d\mathbb{P} = L\) and \((\mathbb{E}[L^2])^{1/2} \leq C_0\) and by Cauchy-Schwartz inequality we obtain:

\[
Q(D \setminus D_{\delta,K}) = \mathbb{E}[L \cdot 1(x \in D \setminus D_{\delta,K})] \leq (\mathbb{E}[L^2])^{1/2} \mathbb{P}(D \setminus D_{\delta,K})^{1/2} \leq \epsilon.
\]
This implies the tightness of $\mathcal{P}_0$. Finally by Prokhorov’s Theorem we conclude that $\mathcal{P}_0$ is relatively compact. \hfill \Box

We now need to link the convergence of measures on path space to the convergence in $\mathcal{S}$, i.e. measures on state space. We define the function $\pi$:

$$
\pi : [0, T] \times \mathcal{P}(\mathbb{E}) \to \mathcal{S} \\
(t, \mu) \to [\mu_{\#X_t}(\{e_1\}), \mu_{\#X_t}(\{e_2\}), \ldots, \mu_{\#X_t}(\{e_m\})].
$$

We have the following result:

**Lemma 3.2.3.** Let $\mu^n$ converge to $\mu$ in $\mathcal{P}$. Then there exists a subset $\mathcal{D}(\mu)$ of $[0, T]$ at most countable such that for all $t \not\in \mathcal{D}(\mu)$:

$$
\pi(t, \mu^n) \to \pi(t, \mu), \quad n \to +\infty. \quad (3.29)
$$

**Proof.** Denote the set $\mathcal{D}(\mu) := \{0 \leq t \leq T, \mu(X_t - X_{t-} \neq 0) > 0\}$. By Lemma 3.12 in Jacod and Shiryaev [1987], the set $\mathcal{D}(\mu)$ is at most countable. In addition, we have $T \not\in \mathcal{D}(\mu)$ since all the paths in $D$ are left-continuous on $T$. In light of Proposition 3.14 in Jacod and Shiryaev [1987], we have $\mu^n_{\#X_t}$ converges to $\mu_{\#X_t}$ weakly for all $t \not\in \mathcal{D}(\mu)$. To conclude, we use the fact that $\mu^n_{\#X_t}$ for all $t \in [0, T]$ and $n$ are counting measures on the discrete set $E$. \hfill \Box

We now turn to the mean field of control. Let $(\mathcal{P}(A), \mathcal{W}_1)$ be the space of probability measures on the compact set $A \subset \mathbb{R}^l$ endowed with the weak topology and metrized by the Wasserstein-1 distance. $(\mathcal{P}(A), \mathcal{W}_1)$ is therefore a Polish space. Since $A$ is compact, it is easy to show that $\mathcal{P}(A)$ is tight and therefore by Prokhorov’s theorem $(\mathcal{P}(A), \mathcal{W}_1)$ is compact. We endow $\mathcal{P}(A)$ with the Borel $\sigma-$algebra denoted by $\mathcal{B}(\mathcal{P}(A))$. For the set $[0, T]$, we endow it with the Borel $\sigma-$algebra $\mathcal{B}([0, T])$ and the
uniform distribution on $[0,T]$. Finally, we construct the product space $[0,T] \times \mathcal{P}(A)$ endowed with the $\sigma$-algebra $\mathcal{B}([0,T]) \otimes \mathcal{B}(\mathcal{P}(A))$. The space of probability measures on $[0,T] \times \mathcal{P}(A)$ can be viewed as a randomized version of the space of mean field of control. We introduce the stable topology on this space:

**Definition 3.2.4.** Let us denote by $\mathcal{R}$ the space of probability measures on $([0,T] \times \mathcal{P}(A), \mathcal{B}([0,T]) \otimes \mathcal{B}(\mathcal{P}(A)))$. We call the stable topology of $\mathcal{R}$ the coarsest topology such that the mapping $\eta \rightarrow \int g(t,m)\eta(dt,dm)$ is continuous for all bounded and measurable mappings $g$ defined on $[0,T] \times \mathcal{P}(A)$ such that $m \rightarrow g(t,m)$ is continuous for all $t \in [0,T]$.

We collect a few useful results on the space $\mathcal{R}$ endowed with the stable topology.

**Proposition 3.2.5.** The topological space $\mathcal{R}$ is metrizable and Polish.

*Proof.* Notice that both $[0,T]$ and $\mathcal{P}(A)$ are Polish with the topology we endow with them. This implies that the $\sigma$-algebra $\mathcal{B}([0,T]) \otimes \mathcal{B}(\mathcal{P}(A)))$ is separable. It follows from Proposition 2.10 in Jacod and Mémin [1981] that $\mathcal{R}$ is metrizable.

We now show that $\mathcal{R}$ is compact. Notice that for an element $\eta$ in $\mathcal{R}$, its first marginal is a probability measure on $[0,T]$ and its second marginal is a probability measure on $\mathcal{P}(A)$. It is trivial to see that both the spaces of probability measures on $[0,T]$ and on $\mathcal{P}(A)$ are tight and therefore relatively compact by Prokhorov’s theorem. We then apply Theorem 2.8 in Jacod and Mémin [1981] and obtain the compactness of $\mathcal{R}$.

Having showed that $\mathcal{R}$ is compact and metrizable, we see that $\mathcal{R}$ is separable. Compactness also leads to completeness. Therefore $\mathcal{R}$ is Polish space. Finally, we notice that $\mathcal{R}$ is also sequential compact since $\mathcal{R}$ is metrizable.

The following result provides a more convenient way to characterize the convergence in the stable topology.
Lemma 3.2.6. Denote by $\mathcal{H}$ the collection of mappings $f$ of the form $f(t, \nu) = 1_B(t) \cdot g(\nu)$ where $B$ is a Borel subset of $[0, T]$ and $g : \mathcal{P}(A) \to \mathbb{R}$ is a bounded Lipschitz function (with regard to the Wasserstein-1 distance on $\mathcal{P}(A)$). Then the stable topology introduced in Definition 3.2.4 is the coarsest topology which makes the mapping $\eta \to \int_{[0,T] \times \mathcal{P}(A)} f(t, \nu) \eta(dt, d\nu)$ continuous for all $f \in \mathcal{H}$.

Proof. Let $\mathcal{H}_0$ be the collection of mappings $f$ of the form $f(t, \nu) = 1_B(t) \cdot g(\nu)$ where $B$ is a Borel subset of $[0, T]$ and $g : \mathcal{P}(A) \to \mathbb{R}$ is a bounded and uniformly continuous function. Then clearly we have $\mathcal{H} \subset \mathcal{H}_0$. By Proposition 2.4 in Jacod and Mémin [1981], the stable topology is the coarsest topology under which the mapping $\eta \to \int_{[0,T] \times \mathcal{P}(A)} f(t, \nu) \eta(dt, d\nu)$ is continuous for all $f \in \mathcal{H}_0$. Therefore, we only need to show that if $\eta^n$ is a sequence of elements in $\mathcal{R}$ such that $\int f(t, \nu) \eta^n(dt, d\nu) \to \int f(t, \nu) \eta^0(dt, d\nu)$ for all $f \in \mathcal{H}$, then we have $\int f(t, \nu) \eta^n(dt, d\nu) \to \int f(t, \nu) \eta^0(dt, d\nu)$ for all $f \in \mathcal{H}_0$ as well.

Now let us fix $f \in \mathcal{H}_0$ with $f(t, \nu) = 1_B(t) \cdot g(\nu)$. Note that $\mathcal{P}(A)$ is a compact metric space and $g$ is a bounded, uniformly continuous and real-valued function. A famous result from Georganopoulos [1967] (see also Miculescu [2000]) shows that $g$ can be approximated uniformly by bounded Lipschitz continuous function. That is, for all $\epsilon > 0$, we can find $g_\epsilon \in \mathcal{H}$ such that $\sup_{\nu \in \mathcal{P}(A)} |g_\epsilon(\nu) - g(\nu)| \leq \epsilon/3$. By our assumption we have $\int 1_B(t) g_\epsilon(\nu) \eta^n(dt, d\nu) \to \int 1_B(t) g_\epsilon(\nu) \eta^0(dt, d\nu)$. Therefore there exists $N_0$ such that $| \int 1_B(t) g_\epsilon(\nu) \eta^n(dt, d\nu) - \int 1_B(t) g_\epsilon(\nu) \eta^0(dt, d\nu) | \leq \epsilon/3$ for all $n \geq N_0$. Combining these facts we have, for $n \geq N_0$:

$$\left| \int 1_B(t) g(\nu) \eta^n(dt, d\nu) - \int 1_B(t) g(\nu) \eta^0(dt, d\nu) \right| \leq \left| \int 1_B(t) g_\epsilon(\nu) \eta^n(dt, d\nu) - \int 1_B(t) g_\epsilon(\nu) \eta^0(dt, d\nu) \right| + \int 1_B(t) |g_\epsilon(\nu) - g(\nu)| \eta^n(dt, d\nu)$$

$$+ \int 1_B(t) |g_\epsilon(\nu) - g(\nu)| \eta^0(dt, d\nu) \leq \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon,$$
which shows that \( \int f(t, \nu) \eta^n(dt, d\nu) \rightarrow \int f(t, \nu) \eta^0(dt, d\nu). \)

Now consider the following subset of \( \mathcal{R} \):

\[
\mathcal{R}_0 := \{ \eta \in \mathcal{R}, \text{the marginal distribution of } \eta \text{ on } [0, T] \text{ is } \mathcal{L} \}.
\] (3.30)

where \( \mathcal{L} \) is the uniformly distributed measure on \([0, T]\). We have the following result:

**Lemma 3.2.7.** \( \mathcal{R}_0 \) is a convex and compact subset of \( \mathcal{R} \).

**Proof.** We apply Theorem 2.8 in Jacod and Mémin [1981]. In particular, we verify without difficulty that \( \{ \eta^{[0,T]}, \eta \in \mathcal{R}_0 \} = \{ \mathcal{L} \} \) is compact and \( \{ \eta^{\mathcal{P}(A)}, \eta \in \mathcal{R}_0 \} \) is a subset of \( \mathcal{P}(\mathcal{P}(A)) \), which is relatively compact as well.

For any \( \eta \in \mathcal{R}_0 \), since the distribution on its first marginal is \( \mathcal{L} \), by the disintegration theorem, we can write \( \eta(dt, dm) = \mathcal{L}(dt) \times \eta_t(dm) \) where the mapping \( [0, T] \ni t \rightarrow \eta_t(\cdot) \in \mathcal{P}(\mathcal{P}(A)) \) is a measurable mapping and the decomposition is unique up to almost everywhere equality. On the other hand, for any measurable function \( \nu : [0, T] \rightarrow \mathcal{P}(A) \), we may construct an element \( \Psi(\nu) \) in \( \mathcal{R}_0 \) by:

\[
\Psi(\nu)(dt, dm) := \mathcal{L}(dt) \times \delta_{\nu_t}(dm).
\] (3.31)

Since we have changed the way we represent the mean field of control, we need to modify accordingly the definition of transition rate matrix as well as the cost functionals in order to make them compatible with the randomization procedure. For any function \( F : \mathcal{P}(A) \rightarrow \mathbb{R} \) possibly containing other arguments, we denote \( \mathcal{E} : \mathcal{P}(\mathcal{P}(A)) \rightarrow \mathbb{R} \) by \( \mathcal{E}(m) := \int_{\nu \in \mathcal{P}(A)} F(\nu)m(d\nu) \), which we call the randomized version of \( F \). Obviously we have \( \mathcal{E}(\delta\nu) = F(\nu) \). In this way, without any ambiguity, we define \( \underline{q} \) the randomized version of the rate function \( q \), as well as its matrix representation \( \underline{Q} \). We also define \( \underline{f} \) the randomized version of cost functional \( f \).
Since the terminal cost \( g \) does not depend on the mean field of control, we do not need to consider its randomized version.

Recall from Assumption 3.1.7 that the minimizer \( \hat{a}_i \) of the reduced Hamiltonian is only a function of \( t, z \) and \( p \). Consequently, for \( \hat{H}, \hat{H}_i, \bar{H} \) and \( \bar{H}_i \), which are respectively the randomized version of respectively \( H, H_i, \bar{H} \) and \( \bar{H}_i \), we still have:

\[
\hat{H}(t, x, z, p, m) = \inf_{\alpha \in A} H(t, x, z, \alpha, p, m),
\]

\[
\hat{a}(t, x, z, p) = \arg\inf_{\alpha \in A} H(t, x, z, \alpha, p, m).
\]

In addition, we have the following result on the Lipschitz property of \( \hat{H} \) and \( \hat{a} \):

**Lemma 3.2.8.** There exists a constant \( C > 0 \) such that for all \((\omega, t) \in \Omega \times (0, T], p, p' \in S, \alpha, \alpha' \in A, Z, Z' \in \mathbb{R}^m \) and \( m, m' \in \mathcal{P}(\mathcal{P}(A)) \), we have:

\[
|q(t, i, j, \alpha, p, m) - q(t, i, j, \alpha', p', m')| \leq C(\|\alpha - \alpha'\| + \|p - p'\| + \bar{W}_1(m, m')),
\]

\[
|\hat{H}(t, X_{t \rightarrow}, z, p, m) - \hat{H}(t, X_{t \rightarrow}, z', p', m')| \leq C\|z - z'\|_{X_{t \rightarrow}} + C(1 + \|z\|_{X_{t \rightarrow}})(\|p - p'\| + \bar{W}_1(m, m')).
\]

**Proof.** For the Lipschitz property of \( \hat{H} \), we have:

\[
|\hat{H}(t, X_{t \rightarrow}, z, p, m) - \hat{H}(t, X_{t \rightarrow}, z', p', m')| \leq \left| \int_{\nu \in \mathcal{P}(A)} \hat{H}(t, X_{t \rightarrow}, z, p, \nu) (m(d\nu) - m'(d\nu)) \right| + C \|z - z'\|_{X_{t \rightarrow}} + C(1 + \|z\|_{X_{t \rightarrow}})(\|p - p'\| + \bar{W}_1(m, m')).
\]
where the last inequality is due to Lemma 3.1.10. Since the space $\mathcal{P}(A)$ is compact and the mapping $\nu \rightarrow \hat{H}(t, X_{t-}, z, p, \nu)$ is Lipschitz, with Lipschitz constant $C(1 + \|z\|_{X_{t-}})$, the Kantorovich-Rubinstein duality theorem implies that:

$$
\left| \int_{\nu \in \mathcal{P}(A)} \hat{H}(t, X_{t-}, z, p, \nu)(m(d\nu) - m'(d\nu)) \right| \leq C(1 + \|z\|_{X_{t-}})\tilde{W}_1(m, m'),
$$

where $\tilde{W}_1(m, m')$ is the Wasserstein-1 distance on the space of probability measure on $\mathcal{P}(A)$ defined as followed:

$$
\tilde{W}_1(m, m') := \inf_{\pi \in \mathcal{P}(\mathcal{P}(A) \times \mathcal{P}(A))} \int_{\mathcal{P}(A) \times \mathcal{P}(A)} W_1(\nu, \nu') \pi(d\nu, d\nu'). \tag{3.34}
$$

Combined with the estimation above, we obtain the desired inequality for $\hat{H}$. The Lipschitz property for $q$ can be proved in the same way.

**3.2.2 Mapping for the Fixed Point Argument**

We now define the mapping of which the fixed points characterize Nash equilibria of the mean field game in the weak formulation. For any $(\mu, \eta) \in \mathcal{P} \times \mathcal{R}_0$, where $\eta$ has the disintegration $\eta(dt, dm) = L(dt) \times \eta_t(dm)$, we consider the solution $(Y^{(\mu, \eta)}, Z^{(\mu, \eta)})$ to the BSDE:

$$
Y_t = g(X_T, p_T) + \int_t^T \hat{H}(s, X_{s-}, Z_s, \pi(s, \mu), \eta_s) ds - \int_t^T Z_s^* \cdot dM_s. \tag{3.35}
$$

Denote by $\hat{\alpha}^{(\mu, \eta)}$ the predictable process $t \rightarrow \hat{\alpha}(t, X_{t-}, Z_t^{(\mu, \eta)}, \pi(t, \mu))$, which is the optimal control of the player faced with the mean field $(\mu, \eta) \in \mathcal{P} \times \mathcal{R}_0$. Next, we consider the scalar martingale $L^{(\mu, \eta)}$ defined by:

$$
L_t^{(\mu, \eta)} := \int_0^t X_{s-}^* \cdot (Q(s, \hat{\alpha}^{(\mu, \eta)}_s, \pi(s, \mu), \eta_s) - Q^0) \cdot \psi^+_s \cdot dM_s. \tag{3.36}
$$
Define the probability measure \( \mathbb{P}^{(\mu, \eta)} \) by:

\[
\frac{d\mathbb{P}^{(\mu, \eta)}}{d\mathbb{P}} := \mathcal{E}(L^{(\mu, \eta)})_T,
\]

where \( \mathcal{E}(L^{(\mu, \eta)}) \) is the Doléans-Dade exponential of the martingale \( L^{(\mu, \eta)} \). Finally we define the mappings \( \Phi^\mu, \Phi^\eta \) and \( \Phi \) respectively by:

\[
\Phi^\mu : \mathcal{P} \times \mathcal{R}_0 \rightarrow \mathcal{P}
\]

\[
(\mu, \eta) \rightarrow \mathbb{P}^{(\mu, \eta)},
\]

\[
\Phi^\eta : \mathcal{P} \times \mathcal{R}_0 \rightarrow \mathcal{R}_0
\]

\[
(\mu, \eta) \rightarrow \mathcal{L}(dt) \times \delta_{\mathbb{P}^{(\mu, \eta)}}(d\nu),
\]

\[
\Phi : \mathcal{P} \times \mathcal{R}_0 \rightarrow \mathcal{P} \times \mathcal{R}_0
\]

\[
(\mu, \eta) \rightarrow (\Phi^\mu(\mu, \eta), \Phi^\eta(\mu, \eta)).
\]

**Remark 3.2.9.** Before delving into the properties of \( \Phi \), we need to show that the mapping \( \Phi \) is well-defined. More specifically, we need to show that given \((\mu, \eta) \in \mathcal{P} \times \mathcal{R}_0\), the output of the mapping \((\mathbb{P}^{(\mu, \eta)}\) and \(\mathcal{L}(dt) \times \delta_{\mathbb{P}^{(\mu, \eta)}}(d\nu)\)) does not depend on which solution to the BSDE (3.35) we use to construct successively \( \hat{\alpha}^{(\mu, \eta)} \), \( L^{(\mu, \eta)} \) and \( \mathcal{E}(L^{(\mu, \eta)}) \). To this end, let us consider \((Y, Z)\) and \((Y', Z')\) two solutions to BSDE (3.35), \( \hat{\alpha} \) and \( \hat{\alpha}' \) the corresponding optimal controls, \( L \) and \( L' \) the corresponding martingales defined in (3.36), and \( Q \) and \( Q' \) the resulting probability measures defined in (3.37). By uniqueness of the solution to (3.35), we have \( \mathbb{E} \left[ \int_0^T \|Z'_t - Z_t\|^2_{X_t} dt \right] = 0 \). Using the Lipschitz continuity of \( \hat{\alpha} \) and \( q \), it is straightforward to show \( \mathbb{E} \left[ \int_0^T \|\hat{\alpha}'_t - \alpha_t\|^2 dt \right] = 0 \) and eventually \( Q = Q' \).

**Proposition 3.2.10.** Denote \( \mathcal{P}_0 \) the closure of the set \( \mathcal{P}_0 \) defined in (3.28). Then the set \( \mathcal{P}_0 \times \mathcal{R}_0 \) is stable for the mapping \( \Phi \).

**Proof.** It suffices to show that for all \((\mu, \eta) \in \mathcal{P} \times \mathcal{R}_0\), we have \( \Phi^\mu(\mu, \eta) \in \mathcal{P}_0 \). By the definition of \( \mathcal{P}_0 \) in (3.28), this boils down to showing that there exists a constant
\( C_0 > 0 \) such that for all \((\mu, \eta)\), we have:

\[
\mathbb{E}^P[(\mathcal{E}(L^{(\mu, \eta)})_T)^2] \leq C_0.
\]

Let us denote \( W_t := \mathcal{E}(L^{(\mu, \eta)})_t \). By Ito’s lemma we have:

\[
d(W^2_t) = 2W_t\,dW_t + d[W, W]_t,
\]

since \( dL^{(\mu, \eta)}_t = X^*_t \cdot (Q(t, \hat{\alpha}^{(\mu, \eta)}_t, \pi(t, \mu), \eta_t) - Q^0) \cdot \psi_t \cdot d\mathcal{M}_t \) and \( dW_t = W_t\,dL^{(\mu, \eta)}_t \).

Denoting \( I_t := \psi_t \cdot (Q^*(t, \hat{\alpha}^{(\mu, \eta)}_t, \pi(t, \mu), \eta_t) - Q^0) \cdot X^*_t \), we have:

\[
d(W^2_t) = 2W^2_t\,dL^{(\mu, \eta)}_t + W^2_t\,I_t \cdot d\mathcal{M}_t \cdot I_t.
\]

We know that the optional quadratic variation of \( \mathcal{M} \) can be decomposed as:

\[
[\mathcal{M}, \mathcal{M}]_t = G_t + \langle \mathcal{M}, \mathcal{M} \rangle_t = G_t + \int_0^t \psi_s ds,
\]

where \( G \) is a martingale. Therefore we have:

\[
d(W^2_t) = 2W^2_t\,dL^{(\mu, \eta)}_t + W^2_t\,I_t^* \cdot I_t + W^2_t\,I_t^* \cdot \psi_t \cdot I_t dt.
\]

Let \( T_n \) be a sequence of stopping times converging to +\( \infty \) which localize both the local martingales \( \int_0^T W^2_s \,dL^2_s^{(\mu, \eta)} \) and \( \int_0^T W^2_s \,I^*_s \cdot dG_s \cdot I_s \). Then integrating the above SDE between 0 and \( T \wedge T_n \) and taking the expectation under \( \mathbb{P} \) we obtain:

\[
\mathbb{E}^P[W^2_{T \wedge T_n}] = 1 + \mathbb{E}^P \left[ \int_0^{T \wedge T_n} W^2_t \,I^*_t \cdot \psi_t \cdot I_t dt \right] = 1 + \mathbb{E}^P \left[ \int_0^{T \wedge T_n} W^2_t \,I^*_t \cdot \psi_t \cdot I_t dt \right] \\
\leq 1 + \int_0^T \mathbb{E}^P \left[ W^2_{T \wedge T_n} I^*_T \cdot \psi_{T \wedge T_n} \cdot I_{T \wedge T_n} \right] \,dt \leq 1 + C_0 \int_0^T \mathbb{E}^P[W^2_{T \wedge T_n}].
\]

Here we have used Tonelli’s theorem as well as the fact that \( I^*_s \cdot \psi_s \cdot I_s \) is bounded by a constant \( C_0 \) independent of \( \mu, \eta \) and \( n \), which is a consequence of the boundedness of the transition rate function \( q \). Now applying Gronwall’s lemma we obtain
$\mathbb{E}^P[W_{T \wedge T_n}^2] \leq C_0$, where the constant $C_0$ does not depend on $n$, $\mu$ or $\eta$. Notice that $W_{T \wedge T_n}^2$ converges to $W_1^2$ almost surely. We then apply Fatou’s lemma and obtain $\mathbb{E}^P[W_1^2] \leq C_0$.

### 3.2.3 Proving the Existence of Nash Equilibria

The last missing piece in applying Schauder’s fixed point theorem is to show the continuity of the mapping $\Phi$ in $\mathcal{P} \times \mathcal{R}_0$, which is the product topology space of $\mathcal{P}$ and $\mathcal{R}_0$. To this end, we show the continuity of the mapping $\Phi^\mu$ and $\Phi^\eta$, respectively.

Notice that both $\mathcal{P}$ and $\mathcal{R}_0$ are metrizable, so we only need to show the sequential continuity.

Let us fix a sequence $(\mu^{(n)}, \eta^{(n)})_{n \geq 1}$ converging to $(\mu^{(0)}, \eta^{(0)})$ in $\mathcal{P} \times \mathcal{R}_0$, with the decomposition $\eta^{(n)}(dt, d\nu) = \mathcal{L}(dt) \times \eta^{(n)}_t(d\nu)$. To simplify the notation we denote $Y^{\mu^{(n)}, \eta^{(n)}}(t) = \mathcal{L}(dt) \times \eta^{(n)}_t([0, T])$. We also denote by $E^{(n)}$ the expectation evaluated under the measure $Q^{(n)}$ and $p_i^{(n)} = \pi(t, \mu^{(n)})$, whereas $E$ still denotes the expectation under the reference measure $P$.

Let us start with the proof of continuity of $\Phi^\mu$, or equivalently the convergence of $Q^{(n)}$ to $Q^{(0)}$. We shall divide the proof into several intermediary results.

**Lemma 3.2.11.** *Without loss of generality, we may assume that there exists a constant $C$ such that the mapping $\|Z_{t_i}^{(n)}\|_{X_{t_i}} \leq C$ for all $(\omega, t) \in \Omega \times [0, T]$.*

**Proof.** We consider the following ODE with unknown $V_i = [V_1(t), \ldots, V_m(t)] \in \mathbb{R}^m$:

\[
\begin{aligned}
0 &= \frac{dV_i(t)}{dt} + \dot{H}_i(t, V(t), p_i^{(0)}, \eta_i^{(0)}) + \sum_{j \neq i} [V_j(t) - V_i(t)], \\
V_i(T) &= g(e_i, p_T^{(0)}), \quad i = 1, \ldots, m.
\end{aligned}
\]  

(3.41)

Denote $\zeta : [0, T] \times \mathbb{R}^m \ni (t, v) \mapsto [\zeta_1(t, v), \ldots, \zeta_m(t, v)] \in \mathbb{R}^m$ where $\zeta_i(t, v) := \dot{H}_i(t, v, p_i^{(0)}, \eta_i^{(0)}) + \sum_{j \neq i} [v_j - v_i]$. By Lemma 3.2.8, we see that $t \rightarrow \zeta(t, v)$ is measurable.
for all $v \in \mathbb{R}^m$ and $v \to \zeta(t,v)$ is Lipschitz in $v$ uniformly in $t$. By Theorem 1 and Theorem 2 in Filippov [2013], the ODE (3.41) admits a unique solution on the interval $[0, T]$, which is absolutely continuous. Now we define $Y_t = \sum_{i=1}^m 1(X_t = e_i)V_i(t)$ and $Z_t = V_t$. By continuity of $V$, we have $\Delta Y_t := Y_t - Y_{t-} = V^*_t \cdot (X_t - X_{t-}) = Z^*_t \cdot dX_t$.

Applying Ito’s formula to $Y$, we obtain:

$$Y_t = Y_T - \int_t^T \sum_{i=1}^m 1(X_t = e_i)\frac{dV_i}{dt}(s)ds - \sum_{t<s\leq T} \Delta Y_s$$

$$= g(X_T, p_T^{(0)}) + \int_t^T \sum_{i=1}^m 1(X_s = e_i)\tilde{H}_i(s, V(s), p_s^{(0)}, \eta_s^{(0)})$$

$$+ \int_t^T \sum_{i=1}^m 1(X_s = e_i)\sum_{j \neq i}[V_j(s) - V_i(s)] - \int_t^T Z^*_s \cdot dX_s$$

$$= g(X_T, p_T^{(0)}) + \int_t^T \tilde{H}_i(s, X_s, Z_s, p_s^{(0)}, \eta_s^{(0)})ds - \int_t^T Z^*_s \cdot dM_s,$$

where in the last equality we used the fact that $dX_s = Q^0_{X_s-}ds + dM_s$ and $V_t = Z_t$.

Therefore $(Y, Z)$ and $(Y^{(0)}, Z^{(0)})$ solve the same BSDE. As we have discussed in Remark 3.2.9, we may assume that $Z^{(0)} = Z$. Therefore $Z^{(0)}_t = V(t)$. It follows from the continuity of $t \to V(t)$ that $t \to \|Z^{(0)}_t\|_{X_{t-}}$ is bounded for all $\omega \in \Omega$. 

Now we show $Z^{(n)}$ converges to $Z^{(0)}$.

**Proposition 3.2.12.** We have:

$$\lim_{n \to +\infty} \mathbb{E}\left[\int_0^T \|Z^{(n)}_t - Z^{(0)}_t\|_{X_{t-}}^2 dt\right] = 0. \quad (3.42)$$

**Proof.** By Lemma 2.4.2, we need to check

$$I_n(t) := \mathbb{E}\left[\left(\int_t^T \tilde{H}(s, X_{s-}, Z^{(0)}_s, p_s^{(n)}, \eta_s^{(n)} - \tilde{H}(s, X_{s-}, Z^{(0)}_s, p_s^{(0)}, \eta_s^{(0)})ds\right)^2\right]$$

converges to 0 for all $t \leq T$ and $I_n(t)$ is bounded by $C$ uniformly in $t$ and $n$. We also need to check $J_n := \mathbb{E}[|g(X_T, p_T^{(n)}) - g(X_T, p_T^{(0)})|^2]$ converges to 0. By the Lipschitz...
continuity of the cost functional $g$ and Lemma 3.2.3, we have:

$$J_n \leq C\|p_T^{(n)} - p_T^{(0)}\|^2 = C\|\pi(\mu^{(n)}, T) - \pi(\mu^{(0)}, T)\|^2 \to 0, n \to +\infty.$$  

To check the uniform boundedness for $I_n(t)$, we recall from Lemma 3.1.10 that

$$|\tilde{H}(t, X_{t-}, Z_{t}^{(0)}, p_{t}^{(n)}, \eta_{t}^{n}) - \tilde{H}(t, X_{t-}, Z_{t}^{(0)}, p_{t}^{(0)}, \eta_{t}^{(0)})|$$

$$\leq C(1 + \|Z_{t}^{(0)}\|_{X_{t-}})(\|p_{t}^{(n)} - p_{t}^{(0)}\| + \tilde{W}_1(\eta_{t}^{(n)}, \eta_{t}^{(0)})),$$

where $\tilde{W}_1$ is the Wasserstein distance on the space $\mathcal{P}(\mathcal{P}(A))$. Clearly $\|p_{t}^{(n)} - p_{t}^{(0)}\|$ can be bounded by a constant since $p_{t}^{(n)}$ is in the simplex $\mathcal{S}$. On the other hand, we have:

$$\tilde{W}_1(\eta_{t}^{(n)}, \eta_{t}^{(0)}) \leq \int_{(\nu_1, \nu_2) \in \mathcal{P}(A)^2} \tilde{W}_1(\nu_1, \nu_2)\eta_{t}^{(n)}(d\nu_1)\eta_{t}^{(0)}(d\nu_2).$$

Since $A$ is compact, $\tilde{W}_1(\nu_1, \nu_2)$ for $(\nu_1, \nu_2) \in \mathcal{P}(A)^2$ is bounded, which implies that $\tilde{W}_1(\eta_{t}^{(n)}, \eta_{t}^{(0)})$ is also bounded by a constant uniformly in $n$ and $t$. This implies:

$$I_n(t) \leq C \mathbb{E} \left[ \int_t^T (1 + \|Z_{s}^{(0)}\|_{X_{s-}})ds \right] \leq C \left( \mathbb{E} \left[ \int_0^T \|Z_{s}^{(0)}\|_{X_{s-}}^2 ds \right] \right)^{1/2} < +\infty,$$

which means that $I_n(t)$ is uniformly bounded in $n$ and $t$. To show that $I_n(t)$ converges to 0, we write:

$$I_n(t) \leq 2\mathbb{E} \left[ \left( \int_t^T (\tilde{H}(s, X_{s-}, Z_{s}^{(0)}, p_{s}^{(n)}, \eta_{s}^{(n)}) - \tilde{H}(s, X_{s-}, Z_{s}^{(0)}, p_{s}^{(0)}, \eta_{s}^{(0)}))ds \right)^2 \right]$$

$$+ 2\mathbb{E} \left[ \left( \int_t^T (\tilde{H}(s, X_{s-}, Z_{s}^{(0)}, p_{s}^{(0)}, \eta_{s}^{(n)}) - \tilde{H}(s, X_{s-}, Z_{s}^{(0)}, p_{s}^{(0)}, \eta_{s}^{(0)}))ds \right)^2 \right]$$

$$\leq 2C \mathbb{E} \left[ \int_t^T (1 + \|Z_{s}^{(0)}\|_{X_{s-}})^2\|p_{s}^{(n)} - p_{s}^{(0)}\|^2 ds \right]$$

$$+ 2\mathbb{E} \left[ \left( \int_t^T (\tilde{H}(s, X_{s-}, Z_{s}^{(0)}, p_{s}^{(0)}, \eta_{s}^{(n)}) - \tilde{H}(s, X_{s-}, Z_{s}^{(0)}, p_{s}^{(0)}, \eta_{s}^{(0)}))ds \right)^2 \right]$$.
By Lemma 3.29, we have $(1 + \|Z_s^{(0)}\|_{X_s})^2\|p_s^{(n)} - p_s^{(0)}\|^2 \to 0$, $ds \times \mathbb{P}$-almost surely.

On the other hand, we have:

$$(1 + \|Z_s^{(0)}\|_{X_s})^2\|p_s^{(n)} - p_s^{(0)}\|^2 \leq C(1 + \|Z_s^{(0)}\|_{X_s})^2,$$

where the right hand side is $ds \times \mathbb{P}$-integrable. Therefore by dominated convergence theorem, we obtain:

$$\lim_{n \to +\infty} E\left[\int_t^T (1 + \|Z_s^{(0)}\|_{X_s})^2\|p_s^{(n)} - p_s^{(0)}\|^2 ds\right] = 0.$$

It remains to show that

$$K_n := E\left[\left(\int_t^T (\hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(n)}) - \hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(0)}))ds\right)^2\right].$$

converges to 0. For fixed $w \in \Omega$ and $s \leq T$, we have:

$$\int_t^T (\hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(n)}) - \hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(0)}))ds$$

$$= \int_t^T \int_{\nu \in \mathcal{P}(A)} \hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \nu)(\eta_s^{(n)} - \eta_s^{(0)})(d\nu)ds$$

$$= \int_{[0,T] \times \mathcal{P}(A)} \kappa(s, \nu)\eta^{(n)}(ds, d\nu) - \int_{[0,T] \times \mathcal{P}(A)} \kappa(s, \nu)\eta^{(0)}(ds, d\nu),$$

where we have defined the mapping $\kappa(s, \nu) := 1(t \leq s \leq T)H(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \nu)$. Clearly $\kappa$ is continuous in $\nu$ for all $s$. On the other hand, by inequality (3.22) in Lemma 3.1.10, for all $t \leq s \leq T$ and $\nu \in \mathcal{P}(A)$ we have:

$$|H(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \nu)|$$

$$\leq |H(s, X_{s-}, 0, 0, 0)| + C\|Z_s^{(0)}\|_{X_{s-}} + C(1 + \|Z_s^{(0)}\|_{X_{s-}})(\|p_s^{(0)}\| + \mathcal{W}_1(\nu, 0)).$$

Therefore by Lemma 3.2.11 and the boundedness of $\mathcal{P}(A)$, we conclude that the mapping $(s, \nu) \to \kappa(s, \nu)$ is bounded. By the definition of stable topology, the fact
that \( \eta^{(n)} \to \eta^{(0)} \) implies:

\[
\lim_{n \to +\infty} \int_t^T |\hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(n)}) - \hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(0)})| ds = 0
\]

for all \( \omega \in \Omega \). In addition, we have:

\[
\left( \int_t^T (\hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(n)}) - \hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(0)})) ds \right)^2 \\
\leq (T - t) \int_t^T |\hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(n)}) - \hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(0)})|^2 ds \\
\leq C \int_t^T (1 + \|Z_s^{(0)}\| X_{s-})^2 (\hat{W}_1(\eta_s^{(n)}, \eta_s^{(0)}))^2 ds \leq C \int_t^T (1 + \|Z_s^{(0)}\| X_{s-})^2 ds,
\]

and \( \int_t^T (1 + \|Z_s^{(0)}\| X_{s-})^2 ds \) is integrable. Applying once again the dominated convergence theorem, we conclude that \( K_n \) converges to 0. This completes the proof. \( \square \)

We will also need a result on a more convenient representation of the Doléans-Dade exponential of \( L^{(n)} \).

**Lemma 3.2.13.** Denote by \( W^{(n)} \) the Doléans-Dade exponential of \( L^{(n)} \). Then the Itô differential of \( \log(W^{(n)}) \) satisfies:

\[
d[\log(W_t^{(n)})] = X_t^* \cdot (Q(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)}) - Q^0 + U(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)}) \cdot Q^0) \cdot X_t \cdot dt \\
+ X_t^* \cdot U(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)}) \cdot dM_t,
\]

where \( U(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)}) \) is the matrix with \( \log(q(t, i, j, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)})) \) as off-diagonal elements and zeros on the diagonal.

**Proof.** Since \( W^{(n)} \) is the Doléans-Dade exponential of \( L^{(n)} \), \( W^{(n)} \) satisfies the SDE

\[
dW_t^{(n)} = W_t^{(n)} dL_t^{(n)}.
\]

Applying Ito’s formula and noticing that the continuous martingale part of \( L^n \) is zero, we have:

\[
d\log(W_t^{(n)}) = dL_t^{(n)} - \Delta L_t^{(n)} + \log(1 + \Delta L_t^{(n)}).
\]

83
Then using $dL_t^{(n)} = X_t^* \cdot (Q(t, a_t^{(n)}, p_t^{(n)}, \eta_t^{(n)}) - Q^0) \cdot \psi_t^+ \cdot dM_t$ and noticing that the jumps of $L^n$ are driven by the jumps of $M$, and hence $X$, we obtain:

$$
d\log(W_t^{(n)}) = -X_t^* \cdot (Q(t, \dot{a}_t^{(n)}, \dot{p}_t^{(n)}, \dot{\eta}_t^{(n)}) - Q^0) \cdot \psi_t^+ \cdot Q^0 \cdot X_t \, dt + \log(1 + L_t^{(n)})
$$

$$
= X_t^* \cdot (Q(t, \dot{a}_t^{(n)}, \dot{p}_t^{(n)}, \dot{\eta}_t^{(n)}) - Q^0) \cdot X_t \, dt + \log(1 + L_t^{(n)}),
$$

where we have used the fact that for all $q$-matrices $A$, we have $X_t^* \cdot A \cdot \psi_t^+ \cdot Q^0 \cdot X_t = -X_t^* \cdot A \cdot X_t$. Piggybacking on the derivation following equation (2.16) in Section 2.1.2, for $X_t = e_i$ and $X_t = e_j$ we have:

$$
\log(1 + L_t^{(n)}) = \log(q(t, i, j, \dot{a}_t^{(n)}, \dot{p}_t^{(n)}, \dot{\eta}_t^{(n)})).
$$

Using matrix notation and recalling the definition of $\bar{U}$ in the statement of Lemma 3.2.13, we may write:

$$
\log(1 + L_t^{(n)}) = X_t^* \cdot \bar{U}(t, \dot{a}_t^{(n)}, \dot{p}_t^{(n)}, \dot{\eta}_t^{(n)}) \cdot \Delta X_t.
$$

Using again the equality $\Delta X_t = dX_t = Q^0 \cdot X_t \, dt + dM_t$, we arrive at the desired representation of the differential of $\log(W_t)$.

We now show that the mapping $\Phi^\mu$ is sequentially continuous.

**Proposition 3.2.14.** $Q^{(n)}$ converges to $Q^{(0)}$ in $\mathcal{P}$.

**Proof.** For two probability measures $Q$ and $Q'$ in $\mathcal{P}$, we define the total variation distance $d_{TV}$ between $Q$ and $Q'$:

$$
d_{TV}(Q, Q') := \sup\{|Q(A) - Q'(A)|, A \in \mathcal{B}(D)\},
$$

(3.43)

where $\mathcal{B}(D)$ is the $\sigma$-algebra generated by the Borel sets of $D$. It is well-known that convergence of total variation distance implies the weak convergence of the measure
and hence the convergence in the topology space $\mathcal{P}$. Therefore our aim is to show that $d_{TV}(Q^{(n)}, Q^{(0)}) \to 0$ as $n \to +\infty$.

By Pinsker’s inequality, we have:

$$d_{TV}^2(Q^{(0)}, Q^{(n)}) \leq \frac{1}{2} \mathbb{E}^{(0)} \left[ \log \left( \frac{dQ^{(0)}}{dQ^{(n)}} \right) \right].$$

Since $\frac{dQ^n}{d\mathbb{P}} = \mathcal{E}(L^n)_T$, we have:

$$d_{TV}^2(Q^{(0)}, Q^{(n)}) \leq \mathbb{E}^{(0)}[\log(\mathcal{E}(L^{(0)}))_T] - \log(\mathcal{E}(L^{(n)}))_T].$$

Using Lemma 3.2.13, we have:

$$\mathbb{E}^{(0)}[\log(\mathcal{E}(L^{(0)}))_T] - \log(\mathcal{E}(L^{(n)}))_T] = \mathbb{E}^{(0)} \left[ \int_0^T X_t^* \cdot (Q(t, \hat{\alpha}^{(0)}_t, \hat{\nu}^{(0)}_t) - Q(t, \hat{\alpha}^{(n)}_t, \hat{\nu}^{(n)}_t)) \cdot dM_t \right].$$

By Assumption 3.1.3, the process $t \to \int_0^t X^*_s - (O(s, \hat{\alpha}^{(0)}_s, \hat{\nu}^{(0)}_s) - O(s, \hat{\alpha}^{(n)}_s, \hat{\nu}^{(n)}_s))$. $dM_s$ is a martingale, and therefore has zero expectation. We now deal with the convergence of the term $\mathbb{E}^{(0)}[\int_0^T X_t^* \cdot (Q(t, \hat{\alpha}^{(0)}_t, \hat{\nu}^{(0)}_t) - Q(t, \hat{\alpha}^{(n)}_t, \hat{\nu}^{(n)}_t)) \cdot X_t \cdot dM_t]$.

while the term $\mathbb{E}^{(0)}[\int_0^T X_t^* \cdot (U(t, \hat{\alpha}^{(0)}_t, \hat{\nu}^{(0)}_t) - U(t, \hat{\alpha}^{(n)}_t, \hat{\nu}^{(n)}_t)) \cdot Q^0 \cdot X_t \cdot dM_t]$ can be dealt with in the exact the same way. Using the Lipschitz property in Lemma 3.2.8, we obtain:

$$\mathbb{E}^{(0)} \left[ \int_0^T X_t^* \cdot (Q(t, \hat{\alpha}^{(0)}_t, \hat{\nu}^{(0)}_t) - Q(t, \hat{\alpha}^{(n)}_t, \hat{\nu}^{(n)}_t)) \cdot X_t \cdot dM_t \right] \leq \mathbb{E}^{(0)} \left[ \int_0^T X_t^* \cdot (Q(t, \hat{\alpha}^{(0)}_t, \hat{\nu}^{(0)}_t) - Q(t, \hat{\alpha}^{(n)}_t, \hat{\nu}^{(n)}_t)) \cdot X_t \cdot dM_t \right]$$

$$+ \mathbb{E}^{(0)} \left[ \int_0^T X_t^* \cdot (Q(t, \hat{\alpha}^{(0)}_t, \hat{\nu}^{(0)}_t) - Q(t, \hat{\alpha}^{(n)}_t, \hat{\nu}^{(n)}_t)) \cdot X_t \cdot dM_t \right]$$

85
This converges to 0, since $\eta$ is bounded according to Assumption 3.1.3. Then by dominated convergence theorem. Finally from Lemma 3.2.11 that $E\left[\left(\int_0^T C(t, \alpha_t^{(n)}, \eta_t^{(n)}) \cdot X_t \, dt\right)^2\right] \leq C\int_0^T E\left[\left(\int_0^T C(t, \alpha_t^{(n)}, \eta_t^{(n)}) \cdot X_t \, dt\right)^2\right] \leq C\int_0^T \left(\int_0^T C(t, \alpha_t^{(n)}, \eta_t^{(n)}) \cdot X_t \, dt\right)^2 \leq C\int_0^T \left(\int_0^T C(t, \alpha_t^{(n)}, \eta_t^{(n)}) \cdot X_t \, dt\right)^2 dt.

This converges to 0 by Proposition 3.2.12. For the second expectation, we notice from Lemma 3.2.11 that $\|Z_t^{(0)}\|_{X_{t-}}$ is bounded by a constant for all $(\omega, t) \in \Omega \times [0, T]$. Therefore we have:

$$E\left[\int_0^T C(1 + \|Z_t^{(0)}\|_{X_{t-}}) \|p_t^{(n)} - p_t^{(0)}\| \, dt\right] \leq C\int_0^T \|p_t^{(n)} - p_t^{(0)}\| \, dt,$$

where the right-hand side converges to 0 by dominated convergence theorem. Finally for the third expectation, we rewrite the integrand as:

$$\int_0^T X_{t-} \cdot (Q(t, \alpha_t^{(n)}, p_t^{(0)}, \eta_t^{(n)}) - Q_t^{(0)}(\cdot, \cdot)) \cdot X_t \, dt$$

This converges to 0, since $\eta^{(n)}$ converges to $\eta^{(0)}$ in stable topology and the mapping $\nu \rightarrow Q(t, \hat{\alpha}_t^{(n)}, p_t^{(0)}, \nu) \cdot X_{t-}(\eta^{(n)}(dt, d\nu) - \eta^{(0)}(dt, d\nu))$ is continuous for all $t$. Notice also that the integrand is bounded by a constant, since $q$ is bounded according to Assumption 3.1.3. Then by dominated convergence theorem. Therefore we have:

$$\int_0^T X_{t-} \cdot (Q(t, \alpha_t^{(n)}, p_t^{(0)}, \eta_t^{(n)}) - Q_t^{(0)}(\cdot, \cdot)) \cdot X_t \, dt$$

This converges to 0, since $\eta^{(n)}$ converges to $\eta^{(0)}$ in stable topology and the mapping $\nu \rightarrow Q(t, \hat{\alpha}_t^{(n)}, p_t^{(0)}, \nu) \cdot X_{t-}(\eta^{(n)}(dt, d\nu) - \eta^{(0)}(dt, d\nu))$ is continuous for all $t$. Notice also that the integrand is bounded by a constant, since $q$ is bounded according to Assumption 3.1.3. Then by dominated convergence theorem.
converges theorem the third expectation converges to 0 as well. This completes the proof.

To show the continuity of $\Phi^n$, we need the following lemma:

**Lemma 3.2.15.** Let $(\nu^{(n)})_{n \in \mathbb{N}}$ be a sequence of measurable functions from $[0, T]$ to $\mathcal{P}(A)$ such that $\int_0^T \mathcal{W}_1(\nu_t^{(n)}, \nu_t^{(0)}) \to 0$. Then $\mathcal{L}(dt) \times \delta_{\nu_t^{(n)}}(d\nu)$ converges to $\mathcal{L}(dt) \times \delta_{\nu_t^{(0)}}(d\nu)$ in $\mathcal{R}_0$ in the sense of the stable topology.

**Proof.** Denote $\lambda^{(n)}(dt, d\nu) := \mathcal{L}(dt) \times \delta_{\nu_t^{(n)}}(d\nu)$ for $n \geq 0$. Let $f : [0, T] \times \mathcal{P}(A) \to \mathbb{R}$ be a mapping of the form $f(t, \nu) = 1(t \in B) \cdot g(\nu)$ where $B$ is measurable subset of $[0, T]$ and $g$ a bounded Lipschitz function on $\mathcal{P}(A)$. We then have:

$$
\left| \int_{[0,T] \times \mathcal{P}(A)} f(t, \nu) \lambda^{(n)}(dt, d\nu) - \int_{[0,T] \times \mathcal{P}(A)} f(t, \nu) \lambda^{(0)}(dt, d\nu) \right| \\
\leq \frac{1}{T} \int_{t \in B} |g(\nu_t^{(n)}) - g(\nu_t^{(0)})| dt \leq C \int_0^T \mathcal{W}_1(\nu_t^{(n)}, \nu_t^{(0)}) dt.
$$

By Lemma 3.2.6, we conclude that $\lambda^{(n)}$ converges to $\lambda^{(0)}$ in the stable topology. \[\]

**Proposition 3.2.16.** $\mathcal{L}(\cdot) \times \delta_{\#\hat{\alpha}_t^{(n)}}(\cdot)$ converges to $\mathcal{L}(\cdot) \times \delta_{\#\hat{\alpha}_t^{(0)}}(\cdot)$ in $\mathcal{R}_0$ in the sense of the stable topology.

**Proof.** By Lemma 3.2.15, we only need to show $\int_0^T \mathcal{W}_1(\frac{Q_n^{(n)}}{\#\hat{\alpha}_t^{(n)}}, \frac{Q_0^{(0)}}{\#\hat{\alpha}_t^{(0)}}) dt$ converges to 0. Notice that:

$$
\int_0^T \mathcal{W}_1(\frac{Q_n^{(n)}}{\#\hat{\alpha}_t^{(n)}}, \frac{Q_0^{(0)}}{\#\hat{\alpha}_t^{(0)}}) dt \leq \int_0^T \mathcal{W}_1(\frac{Q_n^{(n)}}{\#\hat{\alpha}_t^{(n)}}, \frac{Q_0^{(0)}}{\#\hat{\alpha}_t^{(0)}}) dt + \int_0^T \mathcal{W}_1(\frac{Q_0^{(0)}}{\#\hat{\alpha}_t^{(0)}}, \frac{Q_0^{(0)}}{\#\hat{\alpha}_t^{(0)}}) dt.
$$

By the definition of total variation distance (see equation 3.43), we clearly have:

$$
d_{TV}(\frac{Q_n^{(n)}}{\#\hat{\alpha}_t^{(n)}}, \frac{Q_0^{(0)}}{\#\hat{\alpha}_t^{(0)}}) \leq d_{TV}(Q_n^{(n)}, Q_0^{(0)}),
$$

which converges to 0 according to the proof of Proposition 3.2.14. By Theorem 6.16 in Villani [2008], since $A$ is bounded and $\frac{Q_n^{(n)}}{\#\hat{\alpha}_t^{(n)}} \in \mathcal{P}(A)$, there exists a constant $C$
such that:
\[ \mathcal{W}_1(\mathbb{Q}^{(n)}_{\#\hat{\alpha}_t^{(n)}}, \mathbb{Q}^{(0)}_{\#\hat{\alpha}_t^{(0)}}) \leq C \cdot d_{TV}(\mathbb{Q}^{(n)}_{\#\hat{\alpha}_t^{(n)}}, \mathbb{Q}^{(0)}_{\#\hat{\alpha}_t^{(0)}}). \]

This shows that \( \mathcal{W}_1(\mathbb{Q}^{(n)}_{\#\hat{\alpha}_t^{(n)}}, \mathbb{Q}^{(0)}_{\#\hat{\alpha}_t^{(0)}}) \) converges to 0. In addition, it is also bounded since \( A \) is bounded. The dominated convergence theorem then implies that:
\[
\lim_{n \to +\infty} \int_0^T \mathcal{W}_1(\mathbb{Q}^{(n)}_{\#\hat{\alpha}_t^{(n)}}, \mathbb{Q}^{(0)}_{\#\hat{\alpha}_t^{(0)}}) dt = 0.
\]

Now for the other term, we have:
\[
\int_0^T \mathcal{W}_1(\mathbb{Q}^{(0)}_{\#\hat{\alpha}_t^{(0)}}, \mathbb{Q}^{(0)}_{\#\hat{\alpha}_t^{(0)}}) dt \leq \int_0^T \mathbb{E}^0[||\dot{\alpha}_t^{(n)} - \dot{\alpha}_t^{(0)}||] dt
\]
\[
= \mathbb{E}^0 \left[ \int_0^T ||\dot{\alpha}_t^{(n)} - \dot{\alpha}_t^{(0)}|| dt \right] \leq \left( \mathbb{E}^0 \left[ \left( \int_0^T ||\dot{\alpha}_t^{(n)} - \dot{\alpha}_t^{(0)}||^2 dt \right)^{1/2} \right] \right)^{1/2}.
\]

Using the Lipschitz property of \( \hat{a} \) and Proposition 3.2.12, we verify easily that
\[
\mathbb{E}^0 \left[ \int_0^T ||\dot{\alpha}_t^{(n)} - \dot{\alpha}_t^{(0)}||^2 dt \right] \text{ converges to } 0.
\]

We are now ready to show the existence of Nash equilibria.

Proof. (of Theorem 3.2.1) Consider the product space \( \Gamma := \mathcal{P} \times \mathcal{R} \) endowed with the product topology of the weak topology on \( \mathcal{P} \) and the stable topology on \( \mathcal{R} \). By Proposition 3.2.5, \( \Gamma \) is a metrizable Polish space. By Proposition 3.2.2 and Lemma 3.2.7, \( \Gamma_0 := \bar{\mathcal{P}}_0 \times \mathcal{R}_0 \) is a compact convex subset of \( \Gamma \). By Proposition 3.2.10, \( \Gamma_0 \) is stable for the mapping \( \Phi \) defined in (3.40). In addition, we see from Proposition 3.2.14 and Proposition 3.2.16 that \( \Phi \) is continuous on \( \Gamma_0 \). Therefore, applying Schauder’s fixed point theorem, we conclude that \( \Phi \) admits a fixed point \((\mu^*, \eta^*) \in \bar{\mathcal{P}}_0 \times \mathcal{R}_0 \).

Now let us define \( p_t^* := \pi(t, \mu^*) \in \mathcal{S} \) and \( \alpha_t^* := \hat{a}(t, X_t, Z_t^*, \pi(t, \mu^*)) \) where \((Y^*, Z^*) \) is the solution to the BSDE (3.35) with \( \mu = \mu^* \) and \( \eta = \eta^* \). We then define \( \mathbb{P}^* := \mathbb{P}^{\mu^*, \eta^*} \) and \( \nu_t^* := \mathbb{P}^*_{\#\alpha_t^*} \). Since \((\mu^*, \eta^*) \) is the fixed point of the mapping \( \Phi \), we have \( \eta_t^* = \delta_{\nu_t^*} \) and \( p^* = \mathbb{P}^* \). It follows that \( p_t^* = \pi(t, \mathbb{P}^*) = [\mathbb{P}^*(X_t = e_i)]_{1 \leq i \leq m-1} \).

By Proposition 3.1.11, we see that \( \alpha^* \) is the solution to the optimal control problem
3.8 when the mean field of state is \( p^* \) and the mean field of control is \( \nu^* \). This implies that \((\alpha^*, p^*, \nu^*)\) is a Nash equilibrium.

\[ \square \]

### 3.3 Uniqueness of the Nash Equilibrium

**Assumption 3.3.1.** The transition rate function \( q \) does not depend on the mean field of state \( p \) and the mean field of control \( \nu \). The cost functional \( f \) is of the form:

\[
f(t, X, \alpha, p, \nu) = f_0(t, X, p) + f_1(t, X, \alpha) + f_2(t, p, \nu).
\]

(3.44)

For all \( t \in [0, T] \), \( i \in \{1, \ldots, m\} \), \( z \in \mathbb{R}^m \), \( p \in \mathcal{S} \) and \( \nu \in \mathcal{P}(A) \), the mapping \( \alpha \to H_i(t, z, \alpha, p, \nu) \) admits a unique minimizer denoted by \( \hat{a}_i(t, z) \). In addition, there exists a constant \( C > 0 \) such that for all \( i \in \{1, \ldots, m\} \), \( z, z' \in \mathbb{R}^m \):

\[
\|\hat{a}_i(t, z) - \hat{a}_i(t, z')\| \leq C\|z - z\|_{e_i}.
\]

(3.45)

Here is a version of the Lasry-Lions monotonicity condition:

**Assumption 3.3.2.** For all \( p, p' \in \mathcal{S} \) and \( t \in [0, T] \), we have:

\[
\sum_{i=1}^{m} (g(e_i, p) - g(e_i, p'))(p_i - p'_i) \geq 0,
\]

(3.46)

\[
\sum_{i=1}^{m} (f_0(t, e_i, p) - f_0(t, e_i, p'))(p_i - p'_i) \geq 0.
\]

(3.47)

**Theorem 3.3.3.** Under Assumption 3.1.3, 3.1.4, 3.3.1 and 3.3.2, there exists at most one Nash equilibrium for the weak formulation of the finite-state mean field game.

**Proof.** Let \((\alpha^{(1)}, p^{(1)}, \nu^{(1)})\) and \((\alpha^{(2)}, p^{(2)}, \nu^{(2)})\) be two Nash equilibria of the mean field game. For \( i = 1, 2 \), we denote by \((Y^{(i)}, Z^{(i)})\) the solution to the BSDE (3.24) with \( p = p^{(i)}, \nu = \nu^{(i)} \), which is written as:

\[
Y^{(i)}_0 = g(X_T, p^{(i)}_T) + \int_0^T \dot{H}(t, X_{t-}, Z^{(i)}_t, p^{(i)}_t, \nu^{(i)}_t) dt - \int_0^T (Z^{(i)}_t)^* \cdot dM_t.
\]

89
Let us denote $\hat{\alpha}_t^{(i)} := \hat{a}(t, Z_t^{(i)})$ the optimal control, $Q^{(i)} := Q(\hat{\alpha}_t^{(i)}, p_t^{(i)}, \nu_t^{(i)})$ the controlled probability measure defined in (3.4), under which $M^{(i)} := M(\hat{\alpha}_t^{(i)}, p_t^{(i)}, \nu_t^{(i)})$ is a martingale. In addition, we use the abbreviation $f_t^{(i)} := f(t, X_{t-}, \hat{\alpha}_t^{(i)}, p_t^{(i)}, \nu_t^{(i)})$, $g^{(i)} = g(X_T, p_T^{(i)})$ and $Q_t^{(i)} := Q(t, \hat{\alpha}_t^{(i)})$. Taking the difference of the BSDEs we obtain:

$$Y_0^{(1)} - Y_0^{(2)} = \int_0^T (f_1^{(1)} - f_1^{(2)} + X_{t-}^* \cdot (Q_t^{(1)} - Q_t^{(2)}) \cdot Z_t^{(1)} - X_{t-}^* \cdot (Q_t^{(2)} - Q_t^{(1)}) \cdot Z_t^{(2)}) \, dt$$

$$= g^{(1)} - g^{(2)} + \int_0^T (f_1^{(1)} - f_1^{(2)} + X_{t-}^* \cdot (Q_t^{(1)} - Q_t^{(2)}) \cdot Z_t^{(1)}) \, dt + \int_0^T (Z_t^{(1)} - Z_t^{(2)})^* \, dM_t$$

$$= g^{(1)} - g^{(2)} + \int_0^T (f_1^{(1)} - f_1^{(2)} + X_{t-}^* \cdot (Q_t^{(1)} - Q_t^{(2)}) \cdot Z_t^{(1)}) \, dt + \int_0^T (Z_t^{(1)} - Z_t^{(2)})^* \, dM_t^{(1)}.$$

We then take the expectation under $Q^{(1)}$ and $Q^{(2)}$ respectively. Notice that $\mathbb{E}^P[Y_0^{(1)} - Y_0^{(2)}] = \mathbb{E}^{Q^{(1)}}[Y_0^{(1)} - Y_0^{(2)}] = \mathbb{E}^{Q^{(2)}}[Y_0^{(1)} - Y_0^{(2)}]$, we obtain the following equality:

$$\mathbb{E}^{Q^{(1)}} \left[ g^{(1)} - g^{(2)} + \int_0^T (f_1^{(1)} - f_1^{(2)} + X_{t-}^* \cdot (Q_t^{(1)} - Q_t^{(2)}) \cdot Z_t^{(1)}) \, dt \right]$$

$$= \mathbb{E}^{Q^{(2)}} \left[ g^{(1)} - g^{(2)} + \int_0^T (f_1^{(1)} - f_1^{(2)} + X_{t-}^* \cdot (Q_t^{(1)} - Q_t^{(2)}) \cdot Z_t^{(1)}) \, dt \right].$$

Now we take a closer look at the term $f_1^{(1)} - f_1^{(2)} + X_{t-}^* \cdot (Q_t^{(1)} - Q_t^{(2)}) \cdot Z_t^{(2)}$. We have:

$$f_1^{(1)} + X_{t-}^* \cdot Q_t^{(1)} \cdot Z_t^{(2)} = f(t, X_{t-}, \hat{\alpha}_t^{(1)}, p_t^{(1)}, \nu_t^{(1)}) + X_{t-}^* \cdot Q(t, \hat{\alpha}_t^{(1)}) \cdot Z_t^{(2)}$$

$$= H(t, X_{t-}, \hat{\alpha}_t^{(1)}, Z_t^{(2)}, p_t^{(1)}, \nu_t^{(1)})$$

$$\geq H(t, X_{t-}, \hat{\alpha}_t^{(2)}, Z_t^{(2)}, p_t^{(1)}, \nu_t^{(1)})$$

$$= H(t, X_{t-}, \hat{\alpha}_t^{(2)}, Z_t^{(2)}, p_t^{(1)}, \nu_t^{(2)}) + (f_1(t, X_{t-}, p_t^{(1)}) - f_1(t, X_{t-}, p_t^{(2)}))$$

$$+ (f_2(t, p_t^{(1)}, \nu_t^{(1)}) - f_2(t, p_t^{(2)}, \nu_t^{(2)})).$$
Here we use the inequality:

\[ H(t, X_{t-}, \hat{\alpha}_t^{(1)}, Z_t^{(2)}, p_t^{(1)}, \nu_t^{(1)}) \geq H(t, X_{t-}, \hat{\alpha}_t^{(2)}, Z_t^{(2)}, p_t^{(1)}, \nu_t^{(1)}) , \]

which is due to the fact that $\hat{\alpha}_t^{(2)}$ minimizes the Hamiltonian $\alpha \to H(t, X_{t-}, \alpha, Z_t^{(2)}, p_t^{(1)}, \nu_t^{(1)})$ and Assumption 3.3.1 that the minimizer does not depend on the mean field terms.

It follows that:

\[ f_t^{(1)} - f_t^{(2)} + X_{t-}^*(Q_t^{(1)} - Q_t^{(2)})Z_t^{(2)} \geq (f_1(t, X_{t-}, p_t^{(1)}) - f_1(t, X_{t-}, p_t^{(2)})) + (f_2(t, p_t^{(1)}, \nu_t^{(1)}) - f_2(t, p_t^{(2)}, \nu_t^{(2)})). \]

We can interchange the indices and obtain:

\[ f_t^{(2)} - f_t^{(1)} + X_{t-}^*(Q_t^{(2)} - Q_t^{(1)})Z_t^{(1)} \geq (f_1(t, X_{t-}, p_t^{(1)}) - f_1(t, X_{t-}, p_t^{(2)})) + (f_2(t, p_t^{(2)}, \nu_t^{(2)}) - f_2(t, p_t^{(1)}, \nu_t^{(1)})). \]

Injecting these inequalities into equation (3.48) we have:

\[
0 = \mathbb{E}^{Q^{(1)}} \left[ g^{(1)} - g^{(2)} + \int_0^T (f_t^{(1)} - f_t^{(2)} + X_{t-}^* \cdot (Q_t^{(1)} - Q_t^{(2)}) \cdot Z_t^{(2)}) dt \right] \\
- \mathbb{E}^{Q^{(2)}} \left[ g^{(1)} - g^{(2)} + \int_0^T (f_t^{(1)} - f_t^{(2)} + X_{t-}^* \cdot (Q_t^{(1)} - Q_t^{(2)}) \cdot Z_t^{(1)}) dt \right] \\
\geq \mathbb{E}^{Q^{(1)}} \left[ g^{(1)} - g^{(2)} + \int_0^T (f_1(t, X_{t-}, p_t^{(1)}) - f_1(t, X_{t-}, p_t^{(2)}) + f_2(t, p_t^{(1)}, \nu_t^{(1)}) - f_2(t, p_t^{(2)}, \nu_t^{(2)})) dt \right] \\
- \mathbb{E}^{Q^{(2)}} \left[ g^{(1)} - g^{(2)} + \int_0^T (f_1(t, X_{t-}, p_t^{(1)}) - f_1(t, X_{t-}, p_t^{(2)}) + f_2(t, p_t^{(2)}, \nu_t^{(2)}) - f_2(t, p_t^{(1)}, \nu_t^{(1)})) dt \right] \\
= \mathbb{E}^{Q^{(1)}} \left[ g^{(1)} - g^{(2)} + \int_0^T (f_1(t, X_{t-}, p_t^{(1)}) - f_1(t, X_{t-}, p_t^{(2)})) dt \right] \\
- \mathbb{E}^{Q^{(2)}} \left[ g^{(1)} - g^{(2)} + \int_0^T (f_1(t, X_{t-}, p_t^{(1)}) - f_1(t, X_{t-}, p_t^{(2)})) dt \right],
\]

where the last equality is due to the fact that $[f_2(t, p_t^{(2)}, \nu_t^{(2)}) - f_2(t, p_t^{(1)}, \nu_t^{(1)})]$ is deterministic.

From Proposition 3.1.11, since $\alpha^{(i)}$ is the optimal control with regard to the mean field $p^{(i)}$ and $\nu^{(i)}$, we have $\alpha_t^{(i)} = \hat{\alpha}_t^{(i)}$, $dt \otimes d\mathbb{P}$-a.e. This implies that $Q^{(i)}[X_{t-} = e_k] = Q^{(\alpha^{(i)}, p^{(i)}, \nu^{(i)})}[X_{t-} = e_k]$ for all $i = 1, 2$ and $k = 1, \ldots, m$. Since $(\alpha^{(i)}, p^{(i)}, \nu^{(i)})$
is a Nash equilibrium, we have $Q^{(i)}[X_{t-} = e_k] = [p_t^{(i)}]_k$. Therefore we obtain $Q^{(i)}[X_{t-} = e_k] = [p_t^{(i)}]_k$ for all $i = 1, 2$ and $k = 1, \ldots, m$. Now using Assumption 3.3.2, we have:

$$0 \geq \sum_{i=1}^{m} (g(e_i, p_T^{(1)}) - g(e_i, p_T^{(2)}))(|p_T^{(1)}|_i - |p_T^{(2)}|_i)$$

$$+ \int_0^T \sum_{i=1}^{m} (f_0(t, e_i, p_t^{(1)}) - f_0(t, e_i, p_t^{(2)}))(|p_t^{(1)}|_i - |p_t^{(2)}|_i) dt \geq 0. \tag{3.49}$$

Assume that there exists a measurable subset $N$ of $[0, T] \times \Omega$ with strictly positive $dt \otimes dQ^{(1)}$ measure, such that $\hat{\alpha}_t^{(1)} \neq \hat{\alpha}_t^{(2)}$ on $N$. By Assumption 3.3.1, the mapping $\alpha \to H(t, X_{t-}, \alpha, Z_t^{(2)}, p_t^{(1)}, \nu_t^{(1)})$ admits a unique minimizer and therefore for all $(t, w) \in N$, we have:

$$H(t, X_{t-}, \hat{\alpha}_t^{(1)}, Z_t^{(2)}, p_t^{(1)}, \nu_t^{(1)}) > H(t, X_{t-}, \hat{\alpha}_t^{(2)}, Z_t^{(2)}, p_t^{(1)}, \nu_t^{(1)})$$

Piggybacking on the argument laid out above, we see that the first inequality is strict in (3.49) which leads to a contradiction. Therefore we have $\hat{\alpha}_t^{(1)} = \hat{\alpha}_t^{(2)}$, $dt \otimes dQ^{(1)}$-a.e., and $dt \otimes dP$-a.e., since $P$ is equivalent to $Q^{(1)}$. It follows that $\alpha_t^{(1)} = \alpha_t^{(2)}$, $dt \otimes dP$-a.e. Finally, using the same type of argument as in the proof of Proposition 3.2.14, we obtain $Q^{(1)} = Q^{(2)}$ which finally leads to $(p^{(1)}, \mu^{(1)}) = (p^{(2)}, \mu^{(2)})$. This completes the proof of the uniqueness.

\[\square\]

### 3.4 Approximate Nash Equilibria for Games with Finite Many Players

In this section we show that the solution of the mean field game is an approximate Nash equilibrium for the game of finite number of players. We first set the stage for the weak formulation of the game with $N$ players in finite state. Recall that $\Omega$ is the space of càdlàg mappings from $[0, T]$ to $E = \{e_1, \ldots, e_m\}$ which are continuous on $T$, $t \to X_t$ is the canonical process and $\mathcal{F}$ is the natural filtration generated by
Let us fix $p^0 \in \mathcal{S}$ a distribution on state space $E$. Let $\mathbb{P}$ be the probability on $(\Omega, \mathcal{F}_T)$ under which $X$ is a continuous-time Markov chain with transition rate matrix $Q^0$ and initial distribution $p^0$. Let $\Omega^N$ be the product space of $N$ copies of $\Omega$, and $\mathbb{P}^N$ be the product probability measure of $N$ identical copies of $\mathbb{P}$. For $n = 1, \ldots, N$, define the process $X^n_t(w) := w^n_t$ of which the natural filtration is denoted by $(\mathcal{F}^n_t)_{t \in [0,T]}$. We also denote by $(\mathcal{F}^N_t)_{t \in [0,T]}$ the natural filtration generated by the process $(X^1, X^2, \ldots, X^N)$. Denote $\mathcal{M}^n_t := X^n_t - X^0_t - \int_0^t Q^0 \cdot X^n_s \, ds$. It is clear that under $\mathbb{P}^N$, $X^1, \ldots, X^N$ are $N$ independent continuous-time Markov chains with initial distribution $p^0$ and $Q^0$ as the transition rate matrix, and $\mathcal{M}^1, \ldots, \mathcal{M}^N$ are independent $\mathcal{F}^N$-martingales.

Throughout this section, we adopt the following assumption:

**Assumption 3.4.1.** The transition rate function $q$ does not depend on either the mean field of state or the mean field of control.

We assume that each player can observe the entire past history of every player’s state. We denote by $\mathbb{A}^N$ the collection of $\mathbb{P}^N$-predictable processes taking values in $A$. Each player $n$ chooses a strategy $\alpha^n \in \mathbb{A}^N$. We consider the martingale $L^{(\alpha^1, \ldots, \alpha^N)}$ by:

$$L^{(\alpha^1, \ldots, \alpha^N)}_t := \int_0^t \sum_{n=1}^N (X^n_s)^* \cdot (Q(s, \alpha^n_s) - Q^0) \cdot (\psi^n_s)^+ \, d\mathcal{M}^n_s.$$  \hfill (3.50)

We define the probability measure $\mathbb{Q}^{(\alpha^1, \ldots, \alpha^N)}$ by:

$$\frac{d\mathbb{Q}^{(\alpha^1, \ldots, \alpha^N)}}{d\mathbb{P}^N} = \mathcal{E}^{(\alpha^1, \ldots, \alpha^N)},$$  \hfill (3.51)

where we denote by $\mathcal{E}^{(\alpha^1, \ldots, \alpha^N)}$ the Doléans-Dade exponential of $L^{(\alpha^1, \ldots, \alpha^N)}$. Finally we introduce the empirical distribution of the states:

$$p^n_t := \frac{1}{N} \left[ \sum_{n=1}^N \mathbb{1}(X^n_t = e_1), \sum_{n=1}^N \mathbb{1}(X^n_t = e_2), \ldots, \sum_{n=1}^N \mathbb{1}(X^n_t = e_m) \right] \in \mathcal{S}, \hfill (3.52)$$

93
as well as the empirical distribution of the controls:

$$
\nu(\alpha^1_t, \ldots, \alpha^N_t) := \frac{1}{N} \sum_{n=1}^{N} \delta_{\alpha^i_t} \in \mathcal{P}(A), \quad (3.53)
$$

where $\delta_a(\cdot)$ is the Dirac measure on $a$. The total expected cost of player $n$ in the game with $N$ players, denoted by $J^{n,N}(\alpha^1, \ldots, \alpha^N)$, is defined as:

$$
J^{n,N}(\alpha^1, \ldots, \alpha^N) := \mathbb{E}_{Q^1 \cdots Q^N} \left[ \int_0^T f(t, X^n_t, \alpha^n_t, p^n_T, \nu(\alpha^1_t, \ldots, \alpha^N_t))dt + g(X^n_T, p^n_T) \right]. \quad (3.54)
$$

Now let us consider an equilibrium of the mean field game $(\alpha^*, p^*, \nu^*)$ as in Definition 3.1.2. Recall that $\alpha^*$ is a predictable process with respect to the natural filtration generated by the canonical process $X$. For each $n = 1, \ldots, N$, we may define the control $\hat{\alpha}^n$ of player $n$ by:

$$
\hat{\alpha}^n(w^1, \ldots, w^N) := \alpha^*(w^n). \quad (3.55)
$$

Clearly, $\hat{\alpha}^n$ is $\mathcal{F}_{n,N}$-predictable. In other words, it only depends on the observation of player $n$’s own path. Accordingly, the strategy profile $\hat{\alpha}^N := (\hat{\alpha}^1, \ldots, \hat{\alpha}^N)$ is a distributed strategy profile, which means that every player’s strategy is only based on the observation of its own path.

In the following, we will show that $\hat{\alpha}$ is an approximate Nash equilibrium. To this end, we first give a result on the propagation of chaos, which compares players $n$’s total expected cost in the mean field game with its total expected cost in the finite player game. To simplify the notation, we use the abbreviation $(\beta, \hat{\alpha}^{-n,N}) := (\hat{\alpha}^1, \ldots, \hat{\alpha}^{n-1}, \beta, \hat{\alpha}^{n+1}, \ldots, \hat{\alpha}^N)$, $\hat{Q}^N := Q^{\hat{\alpha}^N}, \hat{\mathbb{E}}^N := \mathbb{E}^{\hat{Q}^N}$ and $\hat{\mathcal{E}}^N := \mathcal{E}^{\hat{\alpha}^N}$. We start from the following lemmas:
Lemma 3.4.2. There exists a sequence \((\delta_N)_{N \geq 0}\) such that \(\delta_N \to 0\), \(N \to +\infty\) and for all \(N \geq 1\), \(n \leq N\) and \(t \leq T\) we have:

\[
\max \{ \hat{\mathbb{E}}^N [W_t^2(\nu(\beta_t, \hat{\alpha}_t^{-1,N}), \nu_t^*)], \ \hat{\mathbb{E}}^N [\|p_t^N - p_t^*\|^2] \} \leq \delta_N. \tag{3.56}
\]

Proof. By the definition of \(\hat{Q}^N = \hat{Q}^{\hat{\alpha}N}\) as well as the fact that \((\alpha^*, p^*, \nu^*)\) is an equilibrium of the mean field game, we deduce that under the measure \(\hat{Q}^N\), the states of players \(X^1_t, \ldots, X^N_t\) are independent and have the same distribution characterized by \(p_t^*\) and the controls of players \(\alpha_1^t, \ldots, \alpha_i^N\) are independent and have the same distribution \(\nu_t^*\). Therefore, for \(i \in \{1, \ldots, m\}\), we have:

\[
\hat{\mathbb{E}}^N \left[ \left( \frac{1}{N} \sum_{n=1}^{N} 1(X^n_t = e_i) - \hat{Q}^N[X^1_t = e_i] \right)^2 \right] = \frac{1}{N} (\hat{Q}^N[X^1_t = e_i] - (\hat{Q}^N[X^1_t = e_i])^2) \leq \frac{1}{4N},
\]

which leads to:

\[
\hat{\mathbb{E}}^N [\|p_t^N - p_t^*\|^2] = \sum_{i=1}^{M} \hat{\mathbb{E}}^N \left[ \left( \frac{1}{N} \sum_{n=1}^{N} 1(X^n_t = e_i) - \hat{Q}^N[X^1_t = e_i] \right)^2 \right] \leq \frac{M}{4N}.
\]

On the other hand, \(\nu(\beta_t, \hat{\alpha}_t^{-1,N})\) and \(\nu_t^*\) are in \(\mathcal{P}(A)\) with \(A\) being a compact subset of \(\mathbb{R}^d\). We have:

\[
\hat{\mathbb{E}}^N [W_t^2(\nu(\beta_t, \hat{\alpha}_t^{-1,N}), \nu_t^*)] \\
\leq C \hat{\mathbb{E}}^N [W_1(\nu(\beta_t, \hat{\alpha}_t^{-1,N}), \nu_t^*)] \leq C \hat{\mathbb{E}}^N [W_1(\nu(\beta_t, \hat{\alpha}_t^{-1,N}), \nu(\hat{\alpha}^N)) + W_1(\nu(\hat{\alpha}^N), \nu_t^*)] \\
\leq C(\hat{\mathbb{E}}^N [\frac{1}{N} \|\beta_t - \hat{\alpha}_t^{1,N}\|] + \hat{\mathbb{E}}^N [W_1(\nu(\hat{\alpha}^N), \nu_t^*)]) \leq C(\frac{1}{N} + \hat{\mathbb{E}}^N [W_1(\nu(\hat{\alpha}^N), \nu_t^*)]),
\]

where \(C\) is a constant only depending on \(\sup_{a \in A} \|a\|\) which changes its value from line to line. Now applying Theorem 1 in Fournier and Guillin [2015], we have:

\[
\hat{\mathbb{E}}^N [W_1(\nu(\hat{\alpha}^N), \nu_t^*)] \leq \sup_{a \in A} \|a\| \cdot [1(d \leq 2)(N^{-1/2}\log(1+N)+N^{-2/3})+1(d > 2)(N^{-1/d}+N^{-1/2})].
\]

Combining with the estimates previously shown, we obtain the desired result. \(\square\)
Lemma 3.4.3. There exists a constant $C$ which only depends on the bound of the transition rate $q$, such that for all $N > 0$ and $\beta \in \mathcal{A}^N$ we have:

$$
\hat{E}^N \left[ \left( \frac{\mathcal{E}^{(\beta, \beta^{-1}, N)}_t}{\hat{E}^N_t} \right)^2 \right] \leq C.
$$

(3.57)

Proof. Let us denote $W_t := \mathcal{E}^{(\beta, \beta^{-1}, N)}_t / \hat{E}^N_t$. By Ito’s formula we have:

$$
dW_t = d \left( \frac{\mathcal{E}^{(\beta, \beta^{-1}, N)}_t}{\hat{E}^N_t} \right) = \frac{d\mathcal{E}^{(\beta, \beta^{-1}, N)}_t}{\hat{E}^N_t} - \frac{\Delta \mathcal{E}^{(\beta, \beta^{-1}, N)}_t}{\hat{E}^N_t} - \frac{\hat{E}^{(\beta, \beta^{-1}, N)}_t}{(\hat{E}^N_t)^2} \Delta W_t + \Delta W_t
$$

$$
= W_{t-} \left( \frac{d\mathcal{E}^{(\beta, \beta^{-1}, N)}_t}{\mathcal{E}^{(\beta, \beta^{-1}, N)}_{t-}} - \frac{\hat{E}^{(\beta, \beta^{-1}, N)}_t}{\mathcal{E}^{(\beta, \beta^{-1}, N)}_{t-}} \right) + \Delta W_t
$$

Recall that:

$$
\frac{d\hat{E}^N}{\hat{E}^N} = \sum_{n=1}^N (X^n_{t-})^* \cdot (Q(t, \hat{\alpha}^n_t) - Q^0) \cdot (\psi^n_t)^+ \cdot d\mathcal{M}^n_t,
$$

$$
\frac{d\mathcal{E}^{(\beta, \beta^{-1}, N)}_t}{\mathcal{E}^{(\beta, \beta^{-1}, N)}_{t-}} = (X^1_{t-})^* \cdot (Q(t, \beta_t) - Q^0) \cdot (\psi^1_t)^+ \cdot d\mathcal{M}^1_t + \sum_{n=2}^N (X^n_{t-})^* \cdot (Q(t, \hat{\alpha}^n_t) - Q^0) \cdot (\psi^n_t)^+ \cdot d\mathcal{M}^n_t,
$$

and $d\mathcal{M}^n_t = \Delta \mathcal{M}^n_t - Q^0 \cdot X^n_{t-} dt$. Noticing that for $n \neq 1$, the jumps of $\mathcal{M}^n_t$ do not result in the jumps of $W_t$, we obtain:

$$
\Delta W_t = \Delta \left( \frac{\mathcal{E}^{(\beta, \beta^{-1}, N)}_t}{\mathcal{E}^{(\beta, \beta^{-1}, N)}_{t-}} \right) = \frac{\mathcal{E}^{(\beta, \beta^{-1}, N)}_{t-}}{\mathcal{E}^{(\beta, \beta^{-1}, N)}_{t-}} \cdot \frac{1 + (X^1_{t-})^* \cdot (Q(t, \beta_t) - Q^0) \cdot (\psi^1_t)^+ \cdot \Delta \mathcal{M}^1_t}{1 + (X^1_{t-})^* \cdot (Q(t, \hat{\alpha}_t^1) - Q^0) \cdot (\psi^1_t)^+ \cdot \Delta \mathcal{M}^1_t} - 1
$$

$$
= W_{t-} \frac{(X^1_{t-})^* \cdot (Q(t, \beta_t) - Q(t, \hat{\alpha}_t^1)) \cdot (\psi^1_t)^+ \cdot \Delta \mathcal{M}^1_t}{1 + (X^1_{t-})^* \cdot (Q(t, \hat{\alpha}_t^1) - Q^0) \cdot (\psi^1_t)^+ \cdot \Delta \mathcal{M}^1_t}.
$$

Piggybacking on the computation in equation (2.16), we see that when $X^1_{t-} = e_i \neq e_j = X_t$, we have $\Delta \mathcal{M}^1_t = \Delta X^1_t = e_j - e_i$ and:

$$
\Delta W_t = W_{t-} \frac{q(t, i, j, \beta_t) - q(t, i, j, \hat{\alpha}_t^1)}{q(t, i, j, \hat{\alpha}_t^1)}.
$$

Let us define $\Xi^\beta_t$ to be an $m$ by $m$ matrix where the diagonal elements are 0 and the element on the $i$-th row and the $j$-th column is $\frac{q(t, i, j, \beta_t) - q(t, i, j, \hat{\alpha}_t^1)}{q(t, i, j, \hat{\alpha}_t^1)}$. Then it is clear that
\[ \Delta W_t = e_i^* \cdot \Xi_t^\beta \cdot (e_j - e_i). \] It follows that:

\[ \Delta W_t = W_{t-} \cdot (X_{t-}^1)^* \cdot \Xi_t^\beta \cdot \Delta M_t^1. \]

Injecting the above equation into the Itô decomposition of \( W_t \), we obtain:

\[
dW_t = W_{t-}[(Q(t, \hat{\alpha}_t^1) - Q(t, \beta_t)) \cdot (\psi_t^1)^+ \cdot Q^0 \cdot X_{t-}^1.dt + (X_{t-}^1)^* \cdot \Xi_t^\beta \cdot \Delta M_t^1]
\]

In the second equality, we use the fact that under the measure \( \hat{Q}^N \), the state process \( X_t^1 \) has the canonical decomposition \( dX_t^1 = d\hat{M}_t^1 + Q^*(t, \hat{\alpha}_t^1) \cdot X_t^1 dt \) where \( \hat{M}_t^1 \) is a \( \hat{Q}^N \)-martingale. We also use the equality \( \Delta M_t^1 = \Delta X_t^1 = dX_t^1 \). In addition, by replacing \( X_t^1 \) with \( e_i \) for \( i = 1, \ldots, M \), it is plain to check the following equality:

\[
(Q(t, \hat{\alpha}_t^1) - Q(t, \beta_t)) \cdot (\psi_t^1)^+ \cdot Q^0 \cdot X_{t-}^1 dt + (X_{t-}^1)^* \cdot \Xi_t^\beta \cdot Q^*(t, \hat{\alpha}_t^1) \cdot X_{t-}^1 \]

This leads to the following representation of \( W_t \):

\[
dW_t = W_{t-} \cdot (X_{t-}^1)^* \cdot \Xi_t^\beta \cdot d\hat{M}_t^1,
\]

which is a local martingale under the measure \( \hat{Q}^N \). At this stage, the rest of the proof is exactly the same as the proof of Proposition 3.2.10. In particular, we make use of Assumption 3.1.3, that is the transition rate \( q \) being bounded uniformly with regard to the controls.

We are now ready to show the propagation of chaos.

**Proposition 3.4.4.** There exists a sequence \( (\epsilon_N)_{N \geq 0} \) such that \( \epsilon_N \to 0, N \to +\infty \) and for all \( N \geq 0, n \leq N \) and \( \beta \in A^N \):

\[
\left| J^{n,N}_{\beta, \hat{\alpha}^{n,N}}(\hat{\beta}, \hat{\alpha}^{n,N}) - \mathbb{E}^{Q(\beta, \hat{\alpha}^{n,N})} \left[ \int_0^T f(t, X_t^n, \hat{\beta}_t, p_t^n, \nu_t^n)dt + g(X_T^n, p_T^n) \right] \right| \leq \epsilon_N. \quad (3.58)
\]
Proof. Due to symmetry, we only need to show the claim for \( n = 1 \). Let \( N > 0 \) and \( \beta \in \mathcal{A}^N \). Using successively Cauchy-Schwartz inequality, Assumption 3.1.4, Lemma 3.4.2 and Lemma 3.4.3, we have:

\[
\left| J_{n,N}^{n}(\beta, \hat{\alpha}^{-1,n,N}) - \mathbb{E}_{Q}^{(\beta, \hat{\alpha}^{-1,n,N})} \left[ \int_{0}^{T} f(t, X_{1}^{1}, \beta, p_{t}^{*}, \nu_{t}^{*}) dt + g(X_{1}^{1}, p_{T}^{*}) \right] \right| \\
\leq \mathbb{E}_{Q}^{(\beta, \hat{\alpha}^{-1,n,N})} \left[ \int_{0}^{T} |f(t, X_{1}^{1}, \beta, p_{t}^{*}, \nu_{t}^{*}) - f(t, X_{1}^{1}, \beta, p_{t}^{N}, \nu(\beta, \hat{\alpha}^{-1,n,N}))| dt + |g(X_{1}^{1}, p_{T}^{*}) - g(X_{1}^{1}, p_{T}^{N})| \right] \\
= \hat{E}^{N} \left[ \frac{C_{T}^{(\beta, \hat{\alpha}^{-1,n,N})}}{\hat{E}_{T}^{N}} \int_{0}^{T} |f(t, X_{1}^{1}, \beta, p_{t}^{*}, \nu_{t}^{*}) - f(t, X_{1}^{1}, \beta, p_{t}^{N}, \nu(\beta, \hat{\alpha}^{-1,n,N}))| dt + |g(X_{1}^{1}, p_{T}^{*}) - g(X_{1}^{1}, p_{T}^{N})| \right] \\
\leq \hat{E}^{N} \left[ \left( \frac{C_{T}^{(\beta, \hat{\alpha}^{-1,n,N})}}{\hat{E}_{T}^{N}} \right)^{1/2} \left( \int_{0}^{T} |f(t, X_{1}^{1}, \beta, p_{t}^{*}, \nu_{t}^{*}) - f(t, X_{1}^{1}, \beta, p_{t}^{N}, \nu(\beta, \hat{\alpha}^{-1,n,N}))| dt + |g(X_{1}^{1}, p_{T}^{*}) - g(X_{1}^{1}, p_{T}^{N})| \right)^{2} \right]^{1/2} \\
\leq C \hat{E}^{N} \left[ \left( \frac{C_{T}^{(\beta, \hat{\alpha}^{-1,n,N})}}{\hat{E}_{T}^{N}} \right)^{1/2} \left( \int_{0}^{T} (\hat{E}^{N}[|p_{t}^{N} - p_{t}^{*}|^{2}] + \hat{E}^{N}[W_{\mathcal{A}}^{2}(\nu(\beta, \hat{\alpha}^{-1,n,N}), \nu_{t}^{*})]) dt + \hat{E}^{N}[|p_{T}^{N} - p_{T}^{*}|^{2}] \right)^{1/2} \right] \\
\leq C \sqrt{\delta_{N}},
\]

where \( \delta_{N} \) is as appeared in Lemma 3.4.2, and \( C \) is a constant only depending on \( T \), the Lipschitz constant of \( f \) and \( g \) and the constant appearing in Lemma 3.4.3. this gives us the desired inequality.

As a direct consequence of the above result on the propagation of chaos, we show that the Nash equilibrium of the mean field game consists of an approximate Nash equilibrium for the game with finite many players.

**Theorem 3.4.5.** There exists a sequence \( \epsilon_{N} \) converging to 0 such that for all \( N > 0 \), \( \beta \in \mathcal{A}^{N} \) and \( n \leq N \), we have:

\[
J_{n,N}^{n}(\beta, \hat{\alpha}^{-n,N}) \leq J_{n,N}^{n}(\hat{\alpha}) + \epsilon_{N}.
\]

98
Proof. Recall that the strategy profile is $\hat{\alpha}^N = (\hat{\alpha}^1, \ldots, \hat{\alpha}^N)$ is defined as:

$$\hat{\alpha}^n(w^1, w^2, \ldots, w^N) := \alpha^*(w^n),$$

where $\alpha^*$ is the strategy of the mean field game equilibrium, together with $p^*$ as the mean field of states and $\nu^*$ as the mean field of control. For a strategy profile $(\alpha^1, \ldots, \alpha^N)$ we denote $K^{n,N}(\alpha^1, \ldots, \alpha^N)$ by:

$$K^{n,N}(\alpha^1, \ldots, \alpha^N) := E^Q(\alpha^1, \ldots, \alpha^N) \left[ \int_0^T f(t, X^1_t, \alpha^1_t, p^*_t, \nu^*_t) dt + g(X^1_T, p^*_T) \right].$$

Now taking $n = 1$, we observe that $K^{1,N}(\hat{\alpha}^N) = E^P_N[Y(\hat{\alpha}^N)]$, where $Y(\hat{\alpha}^N)$ is the solution (at time $t = 0$) of the following BSDE:

$$Y_t = g(X^1_T, p^*_T) + \int_t^T H(s, X^1_s, Z^1_s, \hat{\alpha}^1_s, p^*_s, \nu^*_s) ds - \int_t^T (Z^1_s)^* \cdot dM^1_s. \quad (3.59)$$

By the optimality of the equilibrium, we know that for all $t \in [0, T]$, $\hat{\alpha}^1_t$ minimizes the mapping $\alpha \to H(t, X^1_t, Z^1_t, \alpha, p^*_t, \nu^*_t)$. Clearly, the solution of the above BSDE (3.59) is also the unique solution to the following BSDE:

$$Y_t = g(X^1_T, p^*_T) + \int_t^T H(s, X^1_s, Z^1_s, \hat{\alpha}^1_s, p^*_s, \nu^*_s) ds + \sum_{n=2}^N (X^n_t)^* \cdot (Q(s, \hat{\alpha}^n_s) - Q^0) \cdot Z^n_t ds - \int_t^T \sum_{n=1}^N \int_t^T (Z^n_s)^* \cdot dM^n_s,$$

(3.60)

with $Z^n_t = 0$ for $n = 2, \ldots, N$. Indeed, the existence and uniqueness of the BSDE (3.60) can be checked easily by applying Theorem 3.5.2. On the other hand, by following exactly the same argument as in the proof of Lemma 2.2.6, we can show that $K^{1,N}(\beta, \hat{\alpha}^{-1,N}) = E^P_N[Y(\beta^{\hat{\alpha}^{-1,N}})]$, where $Y(\beta^{\hat{\alpha}^{-1,N}})$ is the solution (at time $t = 0$)
of the following BSDE:

\[ Y_t = g(X^1_T, p^*_T) + \int_t^T (H(s, X^1_s, Z^1_s, \beta_s, p^*_s, \nu^*_s) + \sum_{n=2}^N (X^n_s)^* \cdot (Q(s, \alpha^n_s) - Q^0) \cdot Z^n_s \cdot dM^n_s \]

(3.61)

Notice that

\[ H(s, X^1_s, Z^1_s, \alpha, p^*_s, \nu^*_s) = f(s, X^1_s, \alpha, p^*_s, \nu^*_s) + (X^1_s)^* \cdot (Q(s, \alpha) - Q^0) \cdot Z^1_s, \]

and

\[ H(s, X^1_s, Z^1_s, \hat{\alpha}^1_s, p^*_s, \nu^*_s) \geq H(s, X^1_s, Z^1_s, \beta_s, p^*_s, \nu^*_s). \]

Applying the comparison principle as stated in Theorem 3.5.3 to the BSDEs (3.60) and (3.61), we conclude that

\[ K^{1,N}(\beta, \hat{\alpha}^{-1,N}) \leq K^{1,N}(\hat{\alpha}^N) \text{ for all } \beta \in \mathcal{A}^N. \]

Now thanks to symmetry, we have

\[ K^{n,N}(\beta, \hat{\alpha}^{-n,N}) \leq K^{1,N}(\hat{\alpha}^N) \text{ for all } \beta \in \mathcal{A}^N \text{ and } n = 1, \ldots, N. \]

The desired results immediately follows by applying Proposition 3.4.4.

\[ \square \]

### 3.5 Appendix: BSDE Driven by Multiple Independent Continuous-Time Markov Chains

Let us consider a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) supporting \(N\) independent continuous-time Markov chains \(X^1_t, \ldots, X^N_t\). For each \(n = 1, \ldots, N\), \(X^n\) has \(m_n\) states, which are represented by the basis vectors of the space \(\mathbb{R}^{m_n}\). We assume that under \(\mathbb{P}\), the transition rate matrix of \(X^n\) is \(Q^{0,n}\), which is an \(m_n \times m_n\) matrix where all the diagonal elements equal \(-(m_n - 1)\) and all the off-diagonal elements equal 1. We denote by \((\mathcal{F}_t)_{t \in [0,T]}\) the natural filtration generated by \((X^1_t, \ldots, X^N_t)\). It is clear that for each \(n\), we can decompose the Markov chain \(X^n\) as

\[ X^n_t = X^n_0 + \int_0^t Q^{0,n} \cdot X^n_s ds + dM^n_t, \]

where \(M^n\) is an \(\mathcal{F}\)-martingale. In addition, due to the independence of the Markov chains, for all \(n_1 \neq n_2\) and \(t \leq T\), \(\mathbb{P}\)-almost surely we have \(\Delta X^{n_1}_t = 0\) or \(\Delta X^{n_2}_t = 0\). In other words, any two Markov chains cannot jump simultaneously.

Let us consider the process \(X_t := X^1_t \otimes X^2_t \otimes \cdots \otimes X^N_t\) where \(\otimes\) stands for the Kronecker product. Indeed, \(X\) is a Markov chain encoding the joint states of the the
$N$ independent Markov chains, and $X$ only takes values among the unit vectors of the space $\mathbb{R}^{m_1 \times \cdots \times m_N}$. We have the following result on the decomposition of $X$.

**Lemma 3.5.1.** $X$ is a continuous-time Markov chain with transition rate matrix $Q^0$ given by:

$$Q^0 := \sum_{n=1}^{N} I_{m_1} \otimes \cdots \otimes I_{m_n} \otimes Q_{0,n} \otimes I_{m_{n+1}} \otimes \cdots \otimes I_{m_N}.$$  \hspace{1cm} (3.62)

In addition it has the canonical decomposition:

$$dX_t = Q^0 \cdot X_{t-} dt + d\mathcal{M}_t,$$ \hspace{1cm} (3.63)

where $\mathcal{M}$ is a $\mathcal{F}$-martingale and satisfies:

$$d\mathcal{M}_t = \sum_{n=1}^{N} (X_{t-}^1 \otimes \cdots \otimes X_{t-}^{n-1} \otimes I_{m_n} \otimes X_{t-}^{n+1} \otimes \cdots \otimes X_{t-}^N) \cdot d\mathcal{M}_t^n.$$ \hspace{1cm} (3.64)

**Proof.** We show the claim for $N = 2$. Applying Itô’s formula to $X^1_t \otimes X^2_t$ and noticing that $X^1_t$ and $X^2_t$ have no simultaneous jumps, we obtain:

$$d(X^1_t \otimes X^2_t) = dX^1_t \otimes X^2_{t-} + X^1_{t-} \otimes dX^2_t$$

$$= (Q^{0,1} \cdot X^1_{t-}) \otimes X^2_{t-} dt + d\mathcal{M}^1_t \otimes X^2_{t-} + X^1_{t-} \otimes (Q^{0,2} \cdot X^2_{t-}) dt + X^1_{t-} \otimes d\mathcal{M}^2_t.$$  

Using the properties of the Kronecker product, we have:

$$(Q^{0,1} \cdot X^1_{t-}) \otimes X^2_{t-} = (Q^{0,1} \cdot X^1_{t-}) \otimes (I_{m_2} \cdot X^2_{t-}) = (Q^{0,1} \otimes I_{m_2}) \cdot (X^1_{t-} \otimes X^2_{t-})$$

$$d\mathcal{M}^1_t \otimes X^2_{t-} = (I_{m_1} \cdot d\mathcal{M}^1_t) \otimes (X^2_{t-} \cdot 1)$$

$$= (I_{m_1} \otimes X^2_{t-}) \cdot (d\mathcal{M}^1_t \otimes 1) = (I_{m_1} \otimes X^2_{t-}) \cdot d\mathcal{M}^1_t$$

$$X^1_{t-} \otimes (Q^{0,2} \cdot X^2_{t-}) = (I_{m_1} \cdot X^1_{t-}) \otimes (Q^{0,2} \cdot X^2_{t-}) = (I_{m_1} \otimes Q^{0,2}) \cdot (X^1_{t-} \otimes X^2_{t-})$$

$$X^1_{t-} \otimes d\mathcal{M}^2_t = (X^1_{t-} \cdot 1) \otimes (I_{m_2} \cdot d\mathcal{M}^2_t)$$

$$= (X^1_{t-} \otimes I_{m_2}) \cdot (1 \otimes d\mathcal{M}^2_t) = (X^1_{t-} \otimes I_{m_2}) \cdot d\mathcal{M}^2_t.$$
Plugging the above equalities into the Itô decomposition yields the desired result for $N = 2$. The case $N > 2$ can be treated by applying a simple argument of induction, which we will not detail here.

As in the case of a single Markov chain, we define the stochastic matrix $\psi_t^n := \text{diag}(Q^0 \cdot X^n_t) - Q^0 \cdot \text{diag}(X^n_t) - \text{diag}(Q^0 \cdot X^n_t) \cdot Q^0$ for $n = 1, \ldots, N$ as well as $\psi_t := \text{diag}(Q^0 \cdot X_t) - Q^0 \cdot \text{diag}(X_t) - \text{diag}(Q^0 \cdot X_t) \cdot Q^0$. For $n = 1, \ldots, N$, we define the stochastic seminorm $\| \cdot \|_{X^n_t}$ by $\| Z \|_{X^n_t}^2 := Z^* \cdot \psi^n_t \cdot Z$ where $Z \in \mathbb{R}^{m \cdot 1 \times \cdots \times m \cdot N}$. For $n = 1, \ldots, N$, we define the stochastic seminorm $\| \cdot \|_{X^n_t}$ by $\| Z \|_{X^n_t}^2 := Z^* \cdot \psi_t \cdot Z$ where $Z \in \mathbb{R}^{m \cdot 1 \times \cdots \times m \cdot N}$.

Our objective is to show the existence and uniqueness of the following BSDE:

$$Y_t = \xi + \int_t^T F(w, s, Y_s, Z^1_s, \ldots, Z^n_s)ds - \sum_{n=1}^N \int_t^T (Z^n_s)^* \cdot dM^n_s. \quad (3.65)$$

Here $\xi$ is a $\mathcal{F}_T$-measurable $\mathbb{P}$-square integrable random variable and $F : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^{m \cdot 1} \times \cdots \times \mathbb{R}^{m \cdot N} \rightarrow \mathbb{R}$ is the driver function such that the process $t \rightarrow F(w, t, y, z_1, \ldots, z_N)$ is predictable for all $y, z_1, \ldots, z_N \in \mathbb{R} \times \mathbb{R}^{m \cdot 1 \times \cdots \times m \cdot N}$. The unknowns of the equation are a càdlàg process $Y$ taking values in $\mathbb{R}$ and predictable processes $Z^1, \ldots, Z^N$ taking values in $\mathbb{R}^{m \cdot 1}, \ldots, \mathbb{R}^{m \cdot N}$ respectively.

**Theorem 3.5.2.** Assume that there exists a constant $C > 0$ such that $dt \times \mathbb{P}$-a.s., we have:

$$|F(w, t, y, z_1, \ldots, z_N) - F(w, t, \tilde{y}, \tilde{z}_1, \ldots, \tilde{z}_N)| \leq C \left( |y - \tilde{y}| + \sum_{n=1}^N \| z_n - \tilde{z}_n \|_{X^n_t} \right) \quad (3.66)$$

Then the BSDE (3.65) admits a solution $(Y, Z^1, \ldots, Z^N)$ satisfying:

$$\mathbb{E} \left[ \int_0^T |Y_t|^2 dt \right] < +\infty, \quad \mathbb{E} \left[ \sum_{n=1}^N \int_0^T \| Z^n_t \|_{X^n_{t-}}^2 dt \right] < +\infty$$

Moreover, the solution is unique in the sense that if $(Y^{(1)}, Z^{(1)})$ and $(Y^{(2)}, Z^{(2)})$ are two solutions, then $Y^{(1)}$ and $Y^{(2)}$ are indistinguishable and we have $\mathbb{E} \left[ \int_0^T \| Z^{(1)}_t - Z^{(2)}_t \|_{X_{t-}}^2 dt \right] = 0$. 

102
Proof. For simplicity of the presentation, we give the proof for $N = 2$, which can be trivially generalized to any $N > 2$. Our first step is to show that the following equality holds for all $Z \in \mathbb{R}^{m_1 \times m_2}$:

$$
\|Z\|^2_{X_t^{-}} = \|(I_{m_1} \otimes (X_{t_-}^2)^*) \cdot Z\|^2_{X_t^{-}} + \|(X_{t_-}^1)^* \otimes I_{m_2}\) \cdot Z\|^2_{X_t^{-}} \quad (3.67)
$$

By the definition of the semi-norm $\| \cdot \|_{X_t^{-}}$, we have:

$$
\|(I_{m_1} \otimes (X_{t_-}^2)^*) \cdot Z\|^2_{X_t^{-}} = Z^* \cdot (I_{m_1} \otimes X_{t_-}^2) \cdot \psi_t^1 \cdot (I_{m_1} \otimes (X_{t_-}^2)^*) \cdot Z
$$

$$
= Z^* \cdot (I_{m_1} \otimes X_{t_-}^2) \cdot (\psi_t^1 \otimes 1) \cdot (I_{m_1} \otimes (X_{t_-}^2)^*) \cdot Z = Z^* \cdot (\psi_t^1 \otimes X_{t_-}^2) \cdot (I_{m_1} \otimes (X_{t_-}^2)^*) \cdot Z
$$

$$
= Z^* \cdot [\psi_t^1 \otimes (X_{t_-}^2 \cdot (X_{t_-}^2)^*)] \cdot Z = Z^* \cdot (\psi_t^1 \otimes \text{diag}(X_{t_-}^2)) \cdot Z.
$$

Similarly we have $\|(X_{t_-}^1)^* \otimes I_{m_2}\) \cdot Z\|^2_{X_t^{-}} = Z^* \cdot (\text{diag}(X_{t_-}^1) \otimes \psi_t^2) \cdot Z$. Now by the definition of $\psi_t$, we have:

$$
\psi_t = \text{diag}(Q^0 \cdot X_t) - Q^0 \cdot \text{diag}(X_t) - \text{diag}(X_t) \cdot Q^0
$$

$$
= \text{diag}((I_{m_1} \otimes Q^{0,2} + Q^{0,1} \otimes I_{m_2}) \cdot (X_{t_-}^1 \otimes X_{t_-}^2)) - (I_{m_1} \otimes Q^{0,2} + Q^{0,1} \otimes I_{m_2}) \cdot \text{diag}(X_{t_-}^1 \otimes X_{t_-}^2)
$$

$$
- \text{diag}(X_{t_-}^1 \otimes X_{t_-}^2) \cdot (I_{m_1} \otimes Q^{0,2} + Q^{0,1} \otimes I_{m_2})
$$

$$
= \text{diag}(X_{t_-}^1 \otimes (Q^{0,2} \cdot X_{t_-}^2)) + \text{diag}((Q^{0,1} \cdot X_{t_-}^1) \otimes X_{t_-}^2)
$$

$$
- \text{diag}(X_{t_-}^1 \otimes (Q^{0,2} \cdot \text{diag}(X_{t_-}^2))) - (Q^{0,1} \cdot \text{diag}(X_{t_-}^1)) \otimes \text{diag}(X_{t_-}^2)
$$

$$
- \text{diag}(X_{t_-}^1 \otimes (\text{diag}(X_{t_-}^2) \cdot Q^{0,2}) - (\text{diag}(X_{t_-}^1) \cdot Q^{0,1}) \otimes \text{diag}(X_{t_-}^2)
$$

$$
= \psi_t^1 \otimes \text{diag}(X_{t_-}^2) + \text{diag}(X_{t_-}^1) \otimes \psi_t^2,
$$

where we have used the fact that for any two vectors $X^1, X^2$ we have $\text{diag}(X^1 \otimes X^2) = \text{diag}(X^1) \otimes \text{diag}(X^2)$. This immediately leads to the equality $(3.67)$. Now we consider the BSDE driven by the continuous-time Markov chain $X$ with terminal condition $\xi$ and the driver function $F$ defined by:

$$
F(w, t, Y, Z) := F(w, t, Y, (I_{m_1} \otimes (X_{t_-}^2)^*) \cdot Z, ((X_{t_-}^1)^* \otimes I_{m_2}) \cdot Z).
$$
By equality (3.67) and the assumption on the regularity of $F$, we have:

$$|F(w,t,Y_1,Z_1) - F(w,t,Y_2,Z_2)|$$
$$\leq C(|Y_1 - Y_2| + \|(I_{m_1} \otimes (X_{t-}^2)^*) \cdot (Z_1 - Z_2)\|_{X_{t-}^1} + \|(X_{t-}^1)^* \otimes I_{m_2} \cdot (Z_1 - Z_2)\|_{X_{t-}^1})$$
$$\leq C(|Y_1 - Y_2| + \sqrt{2}\|Z_1 - Z_2\|_{X_{t-}}).$$

Applying Lemma 2.4.1 we obtain the existence of the solution to the BSDE:

$$Y_t = \xi + \int_t^T F(s,Y_s,Z_s)ds + \int_t^T Z^*_s \cdot dM_s$$

Now we set $Z^1_t := (I_{m_1} \otimes (X_{s-}^2)^*) \cdot Z_s$ and $Z^2_t := ((X_{s-}^1)^* \otimes I_{m_2}) \cdot Z_s$. From the definition of the driver $F$ and $M$ in equation (3.64), we see that:

$$Y_t = \xi + \int_t^T F(w,s,Y_s,(I_{m_1} \otimes (X_{s-}^2)^*) \cdot Z_s,((X_{s-}^1)^* \otimes I_{m_2}) \cdot Z_s)ds$$
$$+ \int_t^T Z^*_s \cdot [(I_{m_1} \otimes X_{s-}^2) \cdot dM^1_s + (X_{s-}^1 \otimes I_{m_2}) \cdot dM^2_s]$$
$$= \xi + \int_t^T F(w,s,Y_s,Z^1_s,Z^2_s)ds + \int_t^T (Z^1_s)^* \cdot dM^1_s + \int_t^T (Z^2_s)^* \cdot dM^2_s$$

This shows that $(Y, Z^1, Z^2)$ is a solution to BSDE (3.65). \qed

We also state a comparison principle for linear BSDEs driven by multiple independent Markov chains.

**Theorem 3.5.3.** For each $n \in \{1, \ldots, N\}$, let $\gamma^n$ be a bounded predictable process in $\mathbb{R}^{m_n}$ such that $\sum_{i=1}^{m_n}[\gamma^n_i] = 0$ for all $t \in [0,T]$, and $\beta$ a bounded predictable process in $\mathbb{R}$. Let $\phi$ be a non-negative predictable process in $\mathbb{R}$ such that $\mathbb{E}[\int_0^T \|\phi_t\|^2 dt] < +\infty$ and $\xi$ a non-negative square-integrable $\mathcal{F}_T$ measurable random variable in $\mathbb{R}$. Let $(Y, Z)$ be the solution of the linear BSDE:

$$Y_t = \xi + \int_t^T (\phi_s + \beta_s Y_s + \sum_{n=1}^N (\gamma^n_s)^* \cdot Z^n_s)ds - \sum_{n=1}^N \int_t^T (Z^n_s)^* \cdot dM^n_s.$$

(3.68)
Assume that for all \( n = 1, \ldots, N, \ t \in [0, T] \) and \( j \) such that \((e_1^n)^* \cdot Q^{0,n} \cdot X_{1}^{n} > 0\), we have \( 1 + (\gamma_1^n)^* \cdot (\psi_1^n)^+ \cdot (e_1^n - X_{1}^{n}) \geq 0 \) where \((\psi_1^n)^+\) is the Moore-Penrose inverse of the matrix \(\psi_1^n\). Then \(Y\) is nonnegative.

Proof. As before we treat the case for \( N = 2 \), for which the argument can be trivially generalized to any \( N > 2 \). Since \( \gamma \) and \( \beta \) are bounded processes and \( \sum_{i=1}^{m_n} [\gamma_i^n]_i = 0 \) for all \( t \leq T \) and \( n \leq N \), we easily verify that the Lipschitz condition (3.66) as stated in Theorem 3.5.2 is satisfied and therefore the BSDE (3.68) admits a unique solution.

Now consider the following BSDE driven by \( \mathcal{M} \):

\[
Y_t = \xi + \int_t^T (\phi_s + \beta_s Y_s + \gamma_s^* \cdot Z_s) ds - \sum_{n=1}^{2} \int_t^T Z_s^* \cdot d \mathcal{M}_s, \tag{3.69}
\]

where \( \gamma_t := (\gamma_1^1 \otimes X_{1}^{2} - \otimes \gamma_2^1) \). It is easy to verify the BSDE (3.69) admits a unique solution \((Y, Z)\) and following the same argument as in the proof of Theorem 3.5.2, we verify that \((Y_t, Z_1^1, Z_1^2) := (Y_t, (I_{m_1} \otimes (X_{2}^{2})^*) \cdot Z_s, ((X_{2}^{1})^* \otimes I_{m_2}) \cdot Z_s)\) solves the BSDE (3.68), which is also its unique solution. Therefore we only need to show that the solution \(Y\) to BSDE (3.68) is nonnegative. To this ends, we need to apply the comparison principal for the case of a single Markov chain, as is stated in Lemma 2.4.3. Note that \(X^1\) and \(X^2\) do not jump simultaneously and \(X_t = X_{1}^1 \otimes X_{2}^1\). For the jump of \(X\) resulting from the jump of \(X^1\), we need to show that for \( k = 1, \ldots, m_1 \):

\[
1 + \gamma_k^1 \cdot \psi_k^1 \cdot (e_k^1 \otimes X_{1}^{2} - X_{1}^{1} \otimes X_{2}^{2}) \geq 0. \tag{3.70}
\]

Let us assume that \(X_{1}^{1} = e_1^1, X_{1}^{2} = e_2^2\). If \( k = i \), the above equality is trivial. In the following, we consider the case \( k \neq i \). Then by the assumption of the theorem, we have:

\[
1 + (\gamma_k^1)^* \cdot (\psi_k^1)^+ \cdot (e_k^1 - e_i^1) \geq 0. \tag{3.71}
\]
It can be easily verified that:

\[
(diag(e_1^1) \otimes \psi_1^2 + \psi_1 \otimes diag(e_2^2)) \cdot \left[ (m_1 + m_2 - 2)e_k^1 \otimes e_j^2 - \sum_{k_0 \neq k} e_{k_0}^1 \otimes e_j^2 - \sum_{j_0 \neq j} e_i^1 \otimes e_{j_0}^2 \right] = e_k^1 \otimes e_j^2 - e_i^1 \otimes e_j^2,
\]

so that we have:

\[
\psi_t^+ \cdot (e_k^1 \otimes X_{t_-}^2 - X_{t_-}^1 \otimes X_{t_-}^2)
\]

\[
= \frac{1}{m_1 + m_2 - 1} \left[ (m_1 + m_2 - 2)e_k^1 \otimes e_j^2 - \sum_{k_0 \neq k} e_{k_0}^1 \otimes e_j^2 - \sum_{j_0 \neq j} e_i^1 \otimes e_{j_0}^2 \right].
\]

It follows that:

\[
\gamma_t^1 \cdot \psi_t^+ \cdot (e_k^1 \otimes X_{t_-}^2 - X_{t_-}^1 \otimes X_{t_-}^2)
\]

\[
= \frac{1}{m_1 + m_2 - 1} (\gamma_t^1 \otimes e_j^2 + e_i^1 \otimes \gamma_t^2)^* \cdot \left[ (m_1 + m_2 - 2)e_k^1 \otimes e_j^2 - \sum_{k_0 \neq k} e_{k_0}^1 \otimes e_j^2 - \sum_{j_0 \neq j} e_i^1 \otimes e_{j_0}^2 \right]
\]

\[
= \frac{1}{m_1 + m_2 - 1} \left[ (m_1 + m_2 - 2)(e_k^1)^* \cdot \gamma_t^1 - \sum_{k_0 \neq k} (e_{k_0}^1)^* \cdot \gamma_t^1 - (e_j^2)^* \cdot \gamma_t^2 - \sum_{j_0 \neq j} (e_{j_0}^2)^* \cdot \gamma_t^2 \right]
\]

\[
= \frac{1}{m_1 + m_2 - 1} \left[ (m_1 + m_2 - 1)(e_k^1)^* \cdot \gamma_t^1 - \sum_{k_0} (e_{k_0}^1)^* \cdot \gamma_t^1 - (e_j^2)^* \cdot \gamma_t^2 - \sum_{j_0} (e_{j_0}^2)^* \cdot \gamma_t^2 \right]
\]

\[
= (e_k^1)^* \cdot \gamma_t^1,
\]

where in the last equality we used the assumption that \( \sum_{i=1}^{m_n} [\gamma_t^n]_i \) = 0 for \( n = 1, 2 \).

Now noticing that \((e_k^1)^* \cdot \gamma_t^1 = (\gamma_t^1)^* \cdot (\psi_t^1)^+ \cdot (e_k^1 - e_1^1)\), we obtain:

\[
1 + \gamma_t^1 \cdot \psi_t^+ \cdot (e_k^1 \otimes X_{t_-}^2 - X_{t_-}^1 \otimes X_{t_-}^2) = 1 + (\gamma_t^1)^* \cdot (\psi_t^1)^+ \cdot (e_k^1 - e_1^1).
\]
Combining this with the inequality (3.71), we obtain the inequality (3.70). Proceeding in a similar way we can also show that for $k = 1, \ldots, m_2$:

$$1 + \gamma_i^* \cdot \psi_i^* \cdot (X_{t-}^1 \otimes e_k^2 - X_{t-}^1 \otimes X_{t-}^2) \geq 0.$$ 

Applying Lemma 2.4.3 to the BSDE (3.69), we obtain the desired result. \hfill \Box
Chapter 4

Finite State Principal Agent Problem

In many real life situations, not only are we interested in understanding whether and how a system of many interacting particles form equilibria, but we also attempt to control or manage such equilibria so that the macroscopic behavior of these particles reflects certain preferences. How should the government encourage citizens to vaccinate during the flu season to battle the flu outbreak? How should tax be collected to influence people’s decision on consumption, saving and investment? How can an employer compensate its employees in order to boost their productivity? Present in all these scenarios are two parties: the principal - who devices a contract according to which incentives are given to, or penalties are imposed on - the agents. Agents are assumed to be rational, in the sense that they optimally choose their action to maximize their utility function, which translates the tradeoff between the rewards/penalties they received and the effort they put in. The principal’s problem is therefore to conceive a contract that maximizes the principal’s own utility, which is often a function of the population’s states and the transactions with the agents. To make the model even more realistic, we can consider a situation where the principal
can only partially observe the agent’s action. We can also model the agent’s choice of accepting or declining the contract by introducing a reservation utility.

Historically, principal agent problems were first studied under the framework of contract theory in economics and most contributions dealt with models involving a principal and a single agent. In the seminal work Holmstrom and Milgrom [1991], the authors considered a discrete-time model in which the agent’s effort influences the drift of the state process. By assuming both the principal and the agent have a constant absolute risk aversion (CARA) utility function, the authors showed that the optimal contract is linear. This model was further extended in Schättler and Sung [1993], Sung [1995], Müller [1998], and Hellwig and Schmidt [2002]. The breakthrough in understanding the continuous-time principal agent problem can be attributed to Sannikov [2008] and Sannikov [2012], where the author exploited the dynamic nature of the agent’s value function stemming from the dynamic programming principal. This remarkable observation allows the principal’s optimal contracting problem to be formulated as a tractable optimal control problem. This approach was further investigated and generalized in Cvitanić et al. [2015] and Cvitanić et al. [2016] with the help of the second-order BSDE.

It is however a prevailing situation in real life applications that the principal is faced with multiple, if not a large population of agents. Needless to say, the interactions or competition among the agents are likely to play a key role in shaping their behaviors. This adds an extra layer of complexity since the principal not only needs to figure out the optimal response of the agents, but also needs to deduce the equilibrium resulting from the interaction among the agents. Equilibria formed by a large population of agents with complex dynamics used to be notoriously intractable, not to mention how to solve an optimization problem on the equilibria which are parameterized by principal’s contracts. To circumvent these difficulties, Elie et al. [2016] borrowed the idea from the theory of mean field games and considered the
limit scenario of infinitely many agents. Relying on the weak formulation of stochastic differential mean field games developed in Carmona and Lacker [2015], the authors obtained a succinct representation of the equilibrium strategies of individual agents as well as the evolution of the equilibrium distribution across the population. Based on this convenient description of the equilibria, the author managed to obtain a tractable formulation of the optimal contracting problem, which can be solved by the technique of McKean-Vlasov optimal control.

Harnessing the probabilistic approach of finite state mean field games developed in Chapter 3, we believe that the approach adopted in Elie et al. [2016] can be extended to principal agent problem in finite state space. In this chapter, we consider a game involving one principal and a population of infinitely many agents whose states belong to a finite space. Similar to the mean field game model discussed in Chapter 2 and Chapter 3, we use a continuous-time Markov chain to model the evolution of the agents’ states, and we assume that the transition rate between different states is a function of the agent’s control and the statistical distribution of all the agents’ states. At the beginning of the game, the principal chooses a continuous reward rate and a terminal reward to be paid to each agent, which is assumed to be a function of the past history of the agent’s state. Each agent then maximizes their objective function by choosing their optimal effort, accounting for the principal’s reward and the states of all the other agents. Finally, the principal records a utility which depends on the total reward offered to the agents as well as the states of the population of the agents. Assuming that a Nash equilibrium is formed among the population of the agents, we would like to investigate the principal’s optimal choice of rewards to achieve the maximal payoff.

Relying on the weak formulation developed in Chapter 3, the Nash equilibrium can be readily described by a McKean-Vlasov backward stochastic differential equation (BSDE). This BSDE is parameterized by the principal’s continuous payment stream,
as well as the principal’s terminal payment as the terminal condition. Using a similar technique in Cvitanić et al. [2016] and Elie et al. [2016], we show that controlling the terminal condition of the BSDE is indeed equivalent to controlling the initial condition and the martingale term of the corresponding SDE. This allows us to transform the rather intractable formulation of the principal’s optimal contracting problem into an optimal control problem of McKean-Vlasov SDE.

We then focus on a special case where the agents are risk-neutral in the terminal reward, their state transition rates are linear and their instantaneous cost is quadratic in the control. We show that the optimal contracting problem can be further simplified to a deterministic control problem on the flow of distribution of the agents’ states, and the agents use Markovian strategies when the principal announces the optimal contract. By applying the Pontryagin maximum principle, the optimal strategies of the agents and the resulting dynamics of the states under the optimal contract can be obtained by solving a system of forward-backward ODEs.

The rest of this chapter is organized as follow. We introduce the model in Section 4.1, where we will revisit some key elements of the weak formulation of finite state mean field games and formulate the principal’s optimal contracting problem. Our first main result is stated in Section 4.2, where we show the equivalence between the optimal contracting problem and a McKean-Vlasov control problem. In Section 4.3, we deal with the linear-quadratic model, where the search for the optimal contract can be reduced to a deterministic control problem on the flow of distribution of agents’ states. We illustrate the numerical aspect of this linear-quadratic model in Section 4.4 through an example of epidemic containment.
4.1 Model Setup

4.1.1 Agent’s State Process

Given a finite time horizon $T$, we assume that at any time $t \in [0, T]$ each agent finds itself in one of $m$ different states. We denote by $E := \{e_1, e_2, \ldots, e_m\}$ the space of these possible states, where the $e_i$’s are the basis vectors in $\mathbb{R}^m$. As previously, we denote by $S$ the $m$-dimensional simplex $S := \{p \in \mathbb{R}^m, \sum p_i = 1, p_i \geq 0\}$, which we identify with the space of probability measures on $E$.

Let $\Omega$ be the space of càdlàg mappings from $[0, T]$ to $E$. Let $X$ be the canonical process, where $X_t$ represents a representative agent’s state at time $t$. We denote by $\mathcal{F}_t := \sigma(\{X_s, s \leq t\})$ the natural filtration generated by $X$ and $\bar{\mathcal{F}} := \mathcal{F}_T$. Now let us fix an initial distribution for the agents’ states $p^o \in S$. On $(\Omega, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$ we consider the probability measure $\mathbb{P}$ under which the law of $X_0$ is $p^o$ and the canonical process $(X_t)_{t \in [0, T]}$ is a continuous-time Markov chain where the transition rate between any two different states is 1. Recall that the process $X$ has the following representation:

$$X_t = X_0 + \int_0^t Q^0 \cdot X_s - ds + M_t,$$  \hspace{1cm} (4.1)

where $\mathcal{M}$ is a $\mathbb{R}^m$-valued $\mathbb{P}$-martingale, and $Q^0$ is the square matrix with diagonal elements all equal to $-(m - 1)$ and off-diagonal elements all equal to 1.

Let $A$ be a compact subset of $\mathbb{R}^d$ in which the players can choose their controls. Denote by $A$ the collection of $\mathcal{F}$-predictable processes $(\alpha_t)_{t \in [0, T]}$ taking values in $A$. We introduce the transition rate function $q$:

$$[0, T] \times \{1, \ldots, m\}^2 \times A \times S \rightarrow \mathbb{R}$$

$$q(t, i, j, \alpha, p) \rightarrow q(t, i, j, \alpha, p),$$

and we denote by $Q(t, \alpha, p)$ the matrix $[q(t, i, j, \alpha, p)]_{1 \leq i, j \leq m}$. In the rest of the chapter, we make the following assumption on the transition rate:
**Assumption 4.1.1.** (i) For all \((t, \alpha, p) \in [0, T] \times A \times S\), the matrix \(Q(t, \alpha, p)\) is a Q-matrix.

(ii) There exists \(C_1, C_2 > 0\) such that for all \(t, i, j, \alpha, p, \nu \in [0, T] \times E^2 \times A \times S\) with \(i \neq j\), we have \(0 < C_1 < q(t, i, j, \alpha, p) < C_2\).

(iii) There exists a constant \(C > 0\) such that for all \((t, i, j) \in [0, T] \times E^2, \alpha, \alpha' \in A, p, p' \in S\), we have:

\[
|q(t, i, j, \alpha, p) - q(t, i, j, \alpha', p')| \leq C(\|\alpha - \alpha'\| + \|p - p'\|). \tag{4.2}
\]

Now we fix \((p(t))_{0 \leq t \leq T}\) a flow of distributions in \(E\) such that \(p(0) = p^°\), with \(p(t)\) representing the probability distribution of the agents’ states at time \(t\). Following the presentation in Section 2.1.2 of Chapter 2, we construct a probability measure \(Q^{(\alpha, p)}\) which is absolutely continuous with regard to \(\mathbb{P}\) such that the process \(X\) admits the following representation:

\[
X_t = X_0 + \int_0^t Q^*(s, \alpha_s, p(s)) \cdot X_{s-} ds + \mathcal{M}_t^{(\alpha, p)}, \tag{4.3}
\]

where \(\mathcal{M}_t^{(\alpha, p)}\) is a martingale under the measure \(Q^{(\alpha, p)}\). The Radon-Nikodym derivative of \(Q^{(\alpha, p)}\) with regard to \(\mathbb{P}\) is given by \(\frac{dQ^{(\alpha, p)}}{d\mathbb{P}} = \mathcal{E}(L^{(\alpha, p)})_T\), where \(\mathcal{E}(L^{(\alpha, p)})\) is the Doléans-Dade exponential (see Definition 2.1.10) of the martingale \(L^{(\alpha, p)}\) defined as follows:

\[
dL_t^{(\alpha, p)} = X_{t-}^* \cdot (Q(t, \alpha_t, p(t)) - Q^0) \cdot \psi_t^+ \cdot d\mathcal{M}_t, \quad L_0^{(\alpha, p)} = 0.
\]

The decomposition of \(X\) in equation (4.3) means that under the measure \(Q^{(\alpha, p)}\), the stochastic intensity rate of jumps for the process \(X\) is given by the matrix \(Q(t, \alpha_t, p(t))\). In addition, since \(Q^{(\alpha, p)}\) and the reference measure \(\mathbb{P}\) coincide on \(\mathcal{F}_0\), the distributions of \(X_0\) under \(Q^{(\alpha, p)}\) and under \(\mathbb{P}\) are both \(p^°\).
4.1.2 Principal’s Contract and Agent’s Objective Function

The principal fixes a contract consisting of a payment stream \((r_t)_{t \in [0,T]}\) which is a \(\mathcal{F}\)-predictable process and a terminal payment \(\xi\) which is a \(\mathcal{F}_T\)-measurable \(\mathbb{P}\)-square integrable random variable. Let \(c : [0,T] \times E \times A \times \mathcal{S} \rightarrow \mathbb{R}\) be the running cost function of the agent. Let \(u : \mathbb{R} \rightarrow \mathbb{R}\) (resp. \(U : \mathbb{R} \rightarrow \mathbb{R}\)) be the agent’s utility function with regard to continuous payments (resp. terminal payment). When an agent enters into the contract \((r,\xi)\) with the principal and the distribution of all the agents’ states at time \(t\) is \(p(t)\) for \(t \leq T\), the agent’s total expected cost, \(J^{r,\xi}(\alpha,p)\), is defined as:

\[
J^{r,\xi}(\alpha,p) := \mathbb{E}_{Q(\alpha,p)}\left[\int_0^T \left[c(t,X_t,\alpha_t,p(t)) - u(r_t)\right] dt - U(\xi)\right]. \tag{4.4}
\]

4.1.3 Principal’s Objective Function

The principal’s objective function depends on the distribution of the agents’ states and the payments it makes to the agents. Let \(c_0 : [0,T] \times \mathcal{S} \rightarrow \mathbb{R}\) be the running cost function and \(C_0 : \mathcal{S} \rightarrow \mathbb{R}\) the terminal cost function resulting from the distribution of the agents’ states. Given that all the agents choose \(\alpha\) as their controls and the distribution of the agents’ states is given by \((p(t))_{0 \leq t \leq T}\) and the principal offers a contract \((r,\xi)\), the principal records a total expected cost, \(J_0^{\alpha,p}(r,\xi)\), defined as:

\[
J_0^{\alpha,p}(r,\xi) := \mathbb{E}_{Q(\alpha,p)}\left[\int_0^T \left[c_0(t,p(t)) + r_t\right] dt + C_0(p(T)) + \xi\right]. \tag{4.5}
\]
4.1.4 Agents’ Mean Field Nash Equilibria

We assume that the population of infinitely many agents reach a symmetric Nash equilibrium for a given contract \((r, \xi)\) proposed by the principal. We recall the following definition of the Nash equilibrium introduced in Chapter 3 and adapted to the present situation. See Definition 3.1.2 in Section 3.1.1.

**Definition 4.1.2.** Let \((r, \xi)\) be a contract, \(\hat{p} : [0, T] \to S\) be a measurable mapping such that \(\hat{p}(0) = p^0\), and \(\hat{\alpha} \in \mathbb{A}\) be an admissible control for the agents. We say that the tuple \((\hat{\alpha}, \hat{p})\) is a Nash equilibrium of the agents for the contract \((r, \xi)\), denoted by \((\hat{\alpha}, \hat{p}) \in \mathcal{N}(r, \xi)\), if:

(i) \(\hat{\alpha}\) minimizes the cost when the agent is committed to the contract \((r, \xi)\) and the distribution of all the agents is given by the flow \(\hat{p}\):

\[
\hat{\alpha} = \arg \inf_{\alpha \in A} \mathbb{E}_{Q}^{(\alpha, \hat{p})} \left[ \int_{0}^{T} \left[ c(t, X_t, \alpha_t, \hat{p}(t)) - u(r_t) \right] dt - U(\xi) \right].
\]

(ii) \((\hat{\alpha}, \hat{p})\) satisfies the consistency conditions:

\[
\hat{p}(t) = \mathbb{E}_{Q}^{(\hat{\alpha}, \hat{p})}[X_t], \quad \forall t \in [0, T].
\]

Note that equation (4.7) is equivalent to \(\hat{p}_i(t) = Q^{(\hat{\alpha}, \hat{p})}[X_t = e_i], \forall t \in [0, T], i \in \{1, \ldots, m\}\).

4.1.5 Principal’s Optimal Contracting Problem

We now turn to the principal’s optimal choice of the contract, which consists of minimizing the objective function computed based on the Nash equilibria formed by the agents. Of course we need to address the existence of Nash equilibria. However, in the following formulation of the optimal contracting problem, we shall avoid the problem of existence by only considering the contracts \((r, \xi)\) that will result in at least one Nash equilibrium. We call such contracts *admissible contracts*, and we denote by
the collection of all admissible contracts. In addition, among all the possible Nash equilibria, we would like to disregard the equilibria in which the agent’s expected total cost is above a given threshold \( \kappa \). The motivation for imposing this additional constraint is to model the take-it-or-leave-it behavior of the agents in the contract theory: if the agent’s expected total cost exceeds a certain amount, it will turn down the contract. Summarizing the constraints mentioned above, we propose the following optimization problem for the principal:

\[
V(\kappa) := \inf_{(r,\xi) \in \mathcal{C}} \inf_{(\alpha,p) \in \mathcal{N}(r,\xi)} \mathbb{E}^{Q(\alpha,p)} \left[ \int_0^T \left[ c_0(t, p(t)) + r_t \right] dt + C_0(p(T)) + \xi \right],
\]

where we adopt the convention that the infimum over an empty set equals \(+\infty\).

### 4.2 A McKean-Vlasov Control Problem

The original formulation of the principal’s optimal contract is intuitive but far from tractable. In this section we shall provide an equivalent formulation of the principal’s optimal contracting problem which turns out to be an optimal control problem of McKean-Vlasov type. To this end, we rely on the probabilistic characterization of the agents’ Nash equilibria developed in Chapter 3.

#### 4.2.1 Representation of Nash Equilibria Based on BSDEs

We define \( H : [0, T] \times E \times \mathbb{R}^m \times A \times S \times \mathbb{R} \to \mathbb{R} \) the Hamiltonian function for the agent:

\[
H(t, x, z, \alpha, p, R) := c(t, x, \alpha, p) - R + x^* \cdot (Q(t, \alpha, p) - Q^0) \cdot z,
\]

as well as the reduced Hamiltonian \( H_i(t, z, \alpha, p, R) := H(t, e_\xi, z, \alpha, p, R) \).

We make the following assumption on the minimizer of the Hamiltonian:
**Assumption 4.2.1.** For all \( t \in [0,T] \), \( i \in \{1,\ldots,m\} \), \( z \in \mathbb{R}^m \) and \( p \in S \), the mapping \( \alpha \rightarrow H_i(t,z,\alpha,p,R) \) admits a unique minimizer. We denote the minimizer by \( \hat{a}_i(t,z,p) \). In addition, for all \( i \in \{1,\ldots,m\} \), \( \hat{a}_i \) is a measurable mapping on \([0,T] \times \mathbb{R}^m \times S \), and there exists a constant \( C > 0 \) such that for all \( i \in \{1,\ldots,m\} \), \( z,z' \in \mathbb{R}^m \) and \( p \in S \):

\[
\|\hat{a}_i(t,z,p) - \hat{a}_i(t,z',p)\| \leq C\|z - z'\|_{e_i}
\] (4.10)

The above assumption holds, for example, when the cost function \( c \) is strongly convex in \( \alpha \) and the transition rate function is linear in \( \alpha \) (see Remark 3.1.8 in Chapter 3).

We denote by \( \hat{H}_i \) the minimum of the reduced Hamiltonian:

\[
\hat{H}_i(t,z,p,R) := \hat{H}_i(t,z,\hat{a}_i(t,z,p),p,R).
\] (4.11)

Now we define the mapping \( \hat{H} \) and \( \hat{a} \) by:

\[
\hat{H}(t,x,z,p,R) := \sum_{i=1}^{m} \hat{H}_i(t,z,p,R)\mathbb{1}(x = e_i),
\] (4.12)

\[
\hat{a}(t,x,z,p) := \sum_{i=1}^{m} \hat{a}_i(t,z,p)\mathbb{1}(x = e_i).
\] (4.13)

Under Assumption 4.2.1, it is clear that \( \hat{a}(t,x,z,p) \) is the unique minimizer of the mapping \( \alpha \rightarrow H(t,x,z,\alpha,p,R) \), and the minimum is given by \( \hat{H}(t,x,z,p,R) \). In addition, from Assumption 4.1.1 and Assumption 4.2.1, we can show the following result on the Lipschitz property of \( \hat{H} \) and \( \hat{a} \).

**Lemma 4.2.2.** There exists a constant \( C > 0 \) such that for all \( (\omega,t) \in \Omega \times (0,T] \), \( R \in \mathbb{R} \), \( p \in S \) and \( z,z' \in \mathbb{R}^m \), we have:

\[
|\hat{H}(t,X_{t-},z,p,R) - \hat{H}(t,X_{t-},z',p,R)| \leq C\|z - z'\|_{X_{t-}}
\] (4.14)
\[
|\hat{a}(t, X_{t-}, z, p) - \hat{a}(t, X_{t-}, z', p)| \leq C\|z - z'\|_{X_t^-}.
\] 

(4.15)

Now we consider the following McKean-Vlasov BSDE:

\[
Y_t = -U(\xi) + \int_t^T \hat{H}(s, X_{s-}, Z_s, p(s), u(r_s))ds - \int_t^T Z_s^* \cdot dM_s, \quad (4.16)
\]

\[
\mathcal{E}_t = 1 + \int_0^t \mathcal{E}_{s-} X_{s-}^* \cdot (Q(s, \alpha_s, p(s)) - Q^0) \cdot \psi_s^+ \cdot dM_s, \quad (4.17)
\]

\[
\alpha_t = \hat{a}(t, X_{t-}, Z_t, p(t)), \quad (4.18)
\]

\[
p(t) = \mathbb{E}^Q[X_t], \quad \frac{dQ}{d\mathbb{P}} = \mathcal{E}_T. \quad (4.19)
\]

**Definition 4.2.3.** We say that a tuple \((Y, Z, \alpha, p, Q)\) is a solution to the McKean-Vlasov BSDE (4.16)-(4.19) if

- \(Y\) is càdlàg and adapted such that \(\mathbb{E}^\mathbb{P}[\int_0^T Y_t^2 dt] < +\infty\),
- \(Z\) is left-continuous and adapted such that \(\mathbb{E}^\mathbb{P}[\int_0^T \|Z_t\|^2_{X_t^-} dt] < +\infty\), \(\alpha \in \mathbb{A}\), \(p : [0, T] \to \mathcal{S}\) is a measurable mapping, \(Q\) is a probability measure on \(\Omega\) and equations (4.16)-(4.19) are satisfied \(\mathbb{P}\)-a.s. for all \(t \leq T\).

The following result links the solution of the McKean-Vlasov BSDE (4.16)-(4.19) to the Nash equilibrium of the agents.

**Theorem 4.2.4.** Let Assumption 4.1.1 and Assumption 4.2.1 hold. Let \((r, \xi)\) be a contract. Assuming that the BSDE (4.16)-(4.19) admits a solution \((Y, Z, \alpha, p, Q)\), then \((\alpha, p)\) is a Nash equilibrium. Conversely if \((\hat{\alpha}, \hat{p})\) is a Nash equilibrium, then the BSDE (4.16)-(4.19) admits a solution \((Y, Z, \alpha, p, Q)\) such that \(\alpha = \hat{\alpha}, d\mathbb{P} \otimes dt\)-a.e. and \(p(t) = \hat{p}(t), dt\)-a.e..

**Proof.** The first part of the theorem is proved in Chapter 3 (see the end of Section 3.2.3). We now show the second part of the claim. Let \((\hat{\alpha}, \hat{p})\) be a Nash equilibrium and set \(\hat{Q} := Q^{(\hat{\alpha}, \hat{p})}\). By item (ii) in Definition 4.1.2, we see that (4.17) and (4.19) are satisfied. Therefore it remains to show (4.16) and (4.18). Let us consider the BSDE:

\[
Y_t = -U(\xi) + \int_t^T H(s, X_{s-}, Z_s, \hat{\alpha}_s, \hat{p}(s), u(r_s))ds - \int_t^T Z_s^* \cdot dM_s, \quad (4.20)
\]
By the regularity of the cost function and transition rate, it can be shown that the above BSDE admits a unique solution which we denote by \((Y^0, Z^0)\). In addition, we have \(\mathbb{E}^\mathbb{P}[Y_0^0] = J^{r, \xi}(\hat{\alpha}, \hat{p})\). On the other hand, we consider the following BSDE:

\[
Y_t = -U(\xi) + \int_t^T \hat{H}(s, X_{s-}, Z_s, \hat{p}(s), u(r_s))ds - \int_t^T Z_s^* \cdot dM_s. \tag{4.21}
\]

From the regularity of \(\hat{H}\), we can show that BSDE (4.21) also admits a unique solution which we denote by \((Y^1, Z^1)\). In addition, the expectation of \(Y_0^1\) equals the value function of the agent’s optimization problem facing a population of agents with distribution \(\hat{p}\), i.e. \(\mathbb{E}^\mathbb{P}[Y_0^1] = \inf_{\alpha \in \mathbb{A}} J^{r, \xi}(\alpha, \hat{p})\). By the definition of the Nash equilibrium, we have:

\[
\hat{\alpha} \in \arg \inf_{\alpha \in \mathbb{A}} J^{r, \xi}(\alpha, \hat{p}),
\]

which implies that \(\mathbb{E}^\mathbb{P}[Y_0^1] = \mathbb{E}^\mathbb{P}[Y_0^0]\). Note that \(Y_0^1\) and \(Y_0^0\) are \(\mathcal{F}_0\)-measurable and \(\mathbb{P}\) coincides with \(\hat{Q}\) on \(\mathcal{F}_0\). Therefore we have \(\mathbb{E}^\hat{Q}[Y_0^1 - Y_0^0] = 0\). Now we set \(\hat{\alpha}_t' := \hat{\alpha}(t, X_{t-}, Z_t^1, \hat{p}(t))\) which minimizes the mapping \(\alpha \to H(s, X_{t-}, Z_t^1, \alpha, \hat{p}(t), u(r_t))\). Since \((Y^1, Z^1)\) solves the BSDE (4.21), we have:

\[
Y_t^1 = -U(\xi) + \int_t^T H(s, X_{s-}, Z_s^1, \hat{\alpha}_t', \hat{p}(s), u(r_s))ds - \int_t^T (Z_s^1)^* \cdot dM_s. \tag{4.22}
\]

Taking the difference of the BSDEs (4.22) and (4.20), we obtain:

\[
Y_0^0 - Y_0^1 = \int_0^T \left[ H(t, X_{t-}, Z_t^0, \hat{\alpha}_t, \hat{p}(t), u(r_t)) - H(t, X_{t-}, Z_t^1, \hat{\alpha}_t', \hat{p}(t), u(r_t)) \right] dt - \int_0^T (Z_t^0 - Z_t^1)^* \cdot dM_t
\]

\[
= \int_0^T \left[ X_{t-}^* \cdot (Q(t, \hat{\alpha}_t, \hat{p}(t)) - Q^0) \cdot Z_t^0 - X_{t-}^* \cdot (Q(t, \hat{\alpha}_t', \hat{p}(t)) - Q^0) \cdot Z_t^1 \right] dt
\]

\[
+ \int_0^T \left[ c(t, X_{t-}, \hat{\alpha}_t, \hat{p}(t)) - c(t, X_{t-}, \hat{\alpha}_t', \hat{p}(t)) \right] dt - \int_0^T (Z_t^0 - Z_t^1)^* \cdot dM_t
\]

\[
= \int_0^T \left[ c(t, X_{t-}, \hat{\alpha}_t, \hat{p}(t)) - c(t, X_{t-}, \hat{\alpha}_t', \hat{p}(t)) + X_{t-}^* \cdot (Q(t, \hat{\alpha}_t, \hat{p}(t)) - Q(t, \hat{\alpha}_t', \hat{p}(t))) \cdot Z_t^1 \right] dt
\]

\[- \int_0^T (Z_t^0 - Z_t^1)^* \cdot dM_t^{(\hat{\alpha}, \hat{p})}.
\]
Using the optimality of $\hat{\alpha}_t$ and taking the expectation under the measure $Q^{(\hat{\alpha}, \hat{p})}$, we obtain that:

$$0 = \mathbb{E}^{Q^{(\hat{\alpha}, \hat{p})}}[H(t, X_{t-}, Z^1_t, \hat{\alpha}_t, \hat{p}(t), u(r_t)) - H(t, X_{t-}, Z^1_t, \hat{\alpha}'_t, \hat{p}(t), u(r_t))] \geq 0.$$ 

Since the Hamiltonian admits a unique minimizer, we deduce from an argument of contradiction that $\hat{\alpha}'_t = \hat{\alpha}_t$, $dQ^{(\hat{\alpha}, \hat{p})} \otimes dt$-a.e., which implies that $\hat{\alpha}'_t = \hat{\alpha}_t$, $dP \otimes dt$-a.e..

Comparing BSDE (4.20) and BSDE (4.22) and using the uniqueness of the solution, we obtain $\mathbb{E}[\int_0^T \|Z^1_t - Z^0_t\|_X^2 dt] = 0$. Now by the regularity of $\hat{a}$ in Assumption 4.2.1, we have:

$$\mathbb{E}\left[\int_0^T \|\hat{\alpha}_t - \hat{a}(t, X_{t-}, Z^0_t, \hat{p}(t))\|^2 dt\right] = \mathbb{E}\left[\int_0^T \|\hat{\alpha}'_t - \hat{a}(t, X_{t-}, Z^0_t, \hat{p}(t))\|^2 dt\right]$$

$$= \mathbb{E}\left[\int_0^T \|\hat{a}(t, X_{t-}, Z^1_t, \hat{p}(t)) - \hat{a}(t, X_{t-}, Z^0_t, \hat{p}(t))\|^2 dt\right] \leq C\mathbb{E}[\int_0^T \|Z^1_t - Z^0_t\|^2 dt] = 0.$$

Therefore we have $\hat{\alpha}_t = \hat{a}(t, X_{t-}, Z^0_t, \hat{p}(t))$, $dP \otimes dt$-a.e., which immediately implies (4.16) and (4.18). This completes the proof.

4.2.2 Principal’s Optimal Contracting Problem

We now turn to the principal’s optimal choice of the contract. Recall that we have defined $C$ as the collection of all contracts that result in at least one Nash equilibrium. In addition, in order to model the fact that agents are not accepting contracts that will incur a cost above a certain threshold, we further restrict our optimization to the collection of Nash equilibria in which the agent’s expected total cost is below a given threshold $\kappa$. Therefore we propose to solve the following optimization problem for the principal:

$$V(\kappa) := \inf_{(r, \xi) \in C(\alpha, p) \in \mathcal{N}(r, \xi)} \inf_{J^{\alpha}(\alpha, p) \leq \kappa} \mathbb{E}^{Q^{(\alpha, p)}}\left[\int_0^T [c_0(t, p(t)) + r_t] dt + C_0(p(T)) + \xi\right], \quad (4.23)$$

where we adopt the convention that the infimum over an empty set equals $+\infty$. 120
Formulated in this way, the problem seems rather intractable. However, thanks to the probabilistic characterization of the agents’ equilibria stated in Theorem 4.2.4, it is possible to transform it into a McKean-Vlasov control problem. Let us denote by \( \mathcal{H}_X^2 \) the collection of \( \mathcal{F} \)-adapted and left-continuous processes \( Z \) such that \( Z_t \in \mathbb{R}^m \) for \( t \in [0, T] \) and \( \mathbb{E}[\int_0^T ||Z_t||_{X_t}dt] < +\infty \). We also denote by \( \mathcal{R} \) the collection of \( \mathbb{F} \)-predictable process taking values in \( \mathbb{R}^n \). Given \( Z \in \mathcal{H}_X^2 \), \( r \in \mathcal{R} \) and \( Y_0 \) a \( \mathcal{F}_0 \)-measurable random variable, we consider the following SDE of McKean-Vlasov type:

\[
Y_t = Y_0 - \int_0^t \hat{H}(s, X_{s-}, Z_s, p(s, u(r_s)))ds + \int_0^t Z_s^* \cdot dM_s, \tag{4.24}
\]

\[
\mathcal{E}_t = 1 + \int_0^t \mathcal{E}_{s-}X_s^* \cdot (Q(s, \alpha_s, p(s)) - Q^0) \cdot \psi_s^* \cdot dM_s, \tag{4.25}
\]

\[
\alpha_t = \hat{a}(t, X_{t-}, Z_t, p(t)), \tag{4.26}
\]

\[
p(t) = \mathbb{E}^Q[X_t], \quad \frac{dQ}{dP} = \mathcal{E}_T. \tag{4.27}
\]

These are exactly the same equations as in the McKean-Vlasov BSDE (4.16)-(4.19), except that we write the dynamic of \( Y \) in the forward manner. We denote the solution by \( (Y(Z, r, Y_0), \mathcal{E}(Z, r, Y_0), \alpha(Z, r, Y_0), p(Z, r, Y_0), \mathbb{P}(Z, r, Y_0)) \) and the expectation under \( \mathbb{P}(Z, r, Y_0) \) by \( \mathbb{E}^{(Z, r, Y_0)} \). We consider the following optimal control problem:

\[
\tilde{V}(\kappa) := \inf_{\mathbb{E}^Q[Y_0] \leq \kappa} \inf_{Z \in \mathcal{H}_X^2, r \in \mathcal{R}} \mathbb{E}^{(Z, r, Y_0)} \left[ \int_0^T [c_0(t, p(Z, r, Y_0)(t)) + r_t]dt + C_0(p(Z, r, Y_0)(T)) + U^{-1}(-Y_T^{(Z, r, Y_0)}) \right]. \tag{4.28}
\]

As a direct consequence of Theorem 4.2.4, we have the following result:

**Theorem 4.2.5.** Let Assumption 4.1.1 and Assumption 4.2.1 hold. Then \( \tilde{V}(\kappa) = V(\kappa) \).

### 4.3 Solving the Linear-Quadratic Model

In this section, we focus on a particular setup of the principal agent problem where the transition rate has a linear structure and the cost function is quadratic in the
control. When the agent is risk-neutral in terms of their utility function with regard to the terminal reward, we show that the principal’s optimal contracting problem can be further reduced to a deterministic control problem on the probability distribution of agents’ states.

4.3.1 Model Setup

We set the initial distribution of the agents to be \( p^0 \in S \). Each agent is allowed to pick a control \( \alpha \) which belongs to the set \( A = [\underline{\alpha}, \bar{\alpha}] \subset \mathbb{R} \). We assume that the transition rate is a linear function of the control and we define:

\[
q(t, i, j, \alpha, p) := \bar{q}_{i,j}(t, p) + \lambda_{i,j}(\alpha - \underline{\alpha}), \quad \text{for } i \neq j,
\]

\[
q(t, i, i, \alpha, p) := -\sum_{j \neq i} q(t, i, j, \alpha, p),
\]

where \( \lambda_{i,j} \in \mathbb{R}^+, \sum_{j \neq i} \lambda_{i,j} > 0 \) for all \( i \in \{1, \ldots, m\} \) and \( \bar{q}_{i,j} : [0, T] \times S \to \mathbb{R}^+ \) are continuous mappings. We define the agent’s cost function (not including the utility derived from the payment stream):

\[
c(t, e_i, \alpha, p) := c_1(t, e_i, p) + \frac{\gamma_i}{2} \alpha^2,
\]

with \( \gamma_i > 0 \) for all \( i \in \{1, \ldots, m\} \). Finally we define the agent’s utility function of terminal reward to be \( U(\xi) = \xi \) and the utility function of continuous reward \( u \) to be a concave and increasing function.

Under the setup outlined above, it is straightforward to compute the optimizer of the Hamiltonian, which takes the form:

\[
\hat{a}(e_i, Z) = b \left( -\frac{1}{\gamma_i} \sum_{j \neq i} \lambda_{i,j}(Z_j - Z_i) \right),
\]

for \( i \in \{1, \ldots, m\} \), where we have defined \( b(z) := \min\{\max\{z, \underline{\alpha}\}, \bar{\alpha}\} \).
4.3.2 Reduction to Optimal Control on Flows of Probability Measures

We now proceed to reduce the principal’s optimal contracting problem to a deterministic control problem on flows of probability measures corresponding to a continuous-time Markov chain. From the SDE (4.24) and the definition of \( \hat{H} \), we have:

\[
Y_T(Z,r,Y_0) = Y_0 - \int_0^T [c(t, X_t, \hat{a}(X_{t-}, Z_t), p^{(Z,r,Y_0)}(t)) - u(r_t)] dt + \int_0^T Z^*_t \cdot d\mathcal{M}^{(Z,r,Y_0)}_t,
\]

where \( \mathcal{M}^{(Z,r,Y_0)} \) is a martingale under the measure \( \mathbb{P}^{(Z,r,Y_0)} \). Now using \( U^{-1}(-Y_T^{(Z,r,Y_0)}) = -Y_T^{(Z,r,Y_0)} \) and injecting the SDE into the objective function of the principal defined in (4.23), we may rewrite the principal’s optimal contracting problem into:

\[
V(\kappa) = \inf_{\mathbb{E}[Y_0] \leq \kappa} \inf_{Z \in \mathcal{H}_2^X} \mathbb{E}^{(Z,r,Y_0)} \left[ \int_0^T \left[ c_0(t, p^{(Z,r,Y_0)}(t)) + c(t, X_t, \hat{a}(X_{t-}, Z_t), p^{(Z,r,Y_0)}(t)) + r_t - u(r_t) \right] dt + C_0(p^{(Z,r,Y_0)}(T)) - Y_0 \right]
\]

\[
= -\kappa + \inf_{Z \in \mathcal{H}_2^X} \mathbb{E}^{(Z,r,Y_0)} \left[ \int_0^T \left[ c_0(t, p^{(Z,r,Y_0)}(t)) + c(t, X_t, \hat{a}(X_{t-}, Z_t), p^{(Z,r,Y_0)}(t)) + r_t - u(r_t) \right] dt + C_0(p^{(Z,r,Y_0)}(T)) \right],
\]

where we have used the equality \( \mathbb{E}^{\mathbb{F}}[Y_0] = \mathbb{E}^{(Z,r,Y_0)}[Y_0] \). Notice that both the transition rate matrix \( Q \) and the optimizer \( \hat{a} \) do not depend on the reward \( r \) or the agent’s total expected cost \( Y_0 \), and thus we can drop the dependency of \( p^{(Z,r,Y_0)} \), \( \mathbb{P}^{(Z,r,Y_0)} \) and \( \alpha^{(Z,r,Y_0)} \) on \( r \) and \( Y_0 \). We can also isolate the optimal choice of \( r \) from the principal’s optimization problem. Indeed, let \( \hat{r} \) be the minimizer of the mapping \( r \rightarrow r - u(r) \). It is immediately clear that for the principal it is optimal to choose \( r_t = \hat{r} \) for all \( t \in [0, T] \). It follows that:

\[
V(\kappa) = \inf_{Z \in \mathcal{H}_2^X} \mathbb{E}^{(Z)} \left[ \int_0^T \left( c_0(t, p^{(Z)}(t)) + \sum_{i=1}^m 1(X_t = e_i)[c_1(t, e_i, p^{(Z)}(t)) + \frac{\gamma_i}{2} \hat{a}^2(e_i, Z_t)] \right) dt + C_0(p^{(Z)}(T)) \right]
\]

\[
- \kappa + T(\hat{r} - u(\hat{r})�).
\]

(4.33)
We now focus on:

\[ W := \inf_{Z \in \mathcal{H}_X^2} \mathbb{E}^{(Z)} \left[ \int_0^T \left( c_0(t, p^{(Z)}(t)) + \sum_{i=1}^m \mathbb{1} (X_t = e_i) [c_1(t, e_i, p^{(Z)}(t)) + \frac{\gamma_i}{2} \hat{a}^2(e_i, Z_t)] \right) dt + C_0(p^{(Z)}(T)) \right], \]

(4.34)

where \( p^{(Z)}(t) = \mathbb{P}^{(Z)}[X_t] \) and under \( \mathbb{P}^{(Z)} \), \( X \) has the following decomposition with \( \mathcal{M}^{(Z)} \) being a \( \mathbb{P}^{(Z)} \)-martingale:

\[ X_t = X_0 + \int_0^t Q^*(s, \hat{a}(X_{s-}, Z_s), p^{(Z)}(s)) \cdot X_s ds + \mathcal{M}_t^{(Z)}. \]

The key observation is that the control \( Z \) affects the value function only through the optimal control \( \alpha^{(Z)}_t := \hat{a}(t, X_{t-}, Z_t) \) and the mapping \( \mathcal{H}_X^2 \ni Z \mapsto \alpha^{(Z)} \in \mathbb{A} \) is surjective. For any \( Z \in \mathcal{H}_X^2 \), it is clear from the definition of \( \hat{a} \) that \( \alpha^{(Z)} \in \mathbb{A} \). On the other hand, given an arbitrary \( \alpha \in \mathbb{A} \), we set the \( j \)-th component of the process \( Z \) to be:

\[ Z^j_t := \sum_{i \neq j} \mathbb{1} (X_{t-} = i) \frac{\gamma_i \alpha_t}{(m-1) \sum_{k \neq i} \lambda_{i,k}}. \]

(4.35)

By the boundedness of \( \alpha \) we have \( Z \in \mathcal{H}_X^2 \) and it is easy to verify that \( \hat{a}(t, X_{t-}, Z_t) = \alpha_t \). Therefore we can transform the optimization problem \( W \) into \( W = \inf_{\alpha \in \mathbb{A}} I(\alpha) \), where we define:

\[ I(\alpha) := \mathbb{E}^{(\alpha)} \left[ \int_0^T \left( c_0(t, p^{(\alpha)}(t)) + \sum_{i=1}^m \mathbb{1} (X_t = e_i) [c_1(t, e_i, p^{(\alpha)}(t)) + \frac{\gamma_i}{2} \alpha_t^2] \right) dt + C_0(p^{(\alpha)}(T)) \right]. \]

(4.36)

Here with a mild abuse of notation, we denote by \( \mathbb{E}^{(\alpha)} \) the expectation under \( \mathbb{P}^{(\alpha)} \), and \( \mathbb{P}^{(\alpha)} \) is the probability measure defined by the following McKean-Vlasov SDE:

\[
\frac{d\mathbb{P}^{(\alpha)}}{d\mathbb{P}} = \mathcal{E}_T^{(\alpha)}, \quad p^{(\alpha)}(t) = \mathbb{E}^{(\alpha)}[X_t],
\]

\[
\mathcal{E}_t^{(\alpha)} = 1 + \int_0^t \mathcal{E}_{s-} X^*_s \cdot (Q(s, \alpha_s, p^{(\alpha)}(s)) - Q^0) \cdot \psi^+_s \cdot d\mathcal{M}_s.
\]

(4.37)
Recall that under \( \mathbb{P}^{(\alpha)} \), \( X \) has the decomposition:

\[
X_t = X_0 + \int_0^t Q^*(s, \alpha_s, p^{(\alpha)}(s)) \cdot X_{s-} ds + \mathcal{M}^{(\alpha)}_t,
\]

where \( \mathcal{M}^{(\alpha)} \) is a \( \mathbb{P}^{(\alpha)} \)-martingale. Taking expectation under \( \mathbb{P}^{(\alpha)} \), we have:

\[
p^{(\alpha)}(t) = p^{(\alpha)}(0) + \int_0^t \mathbb{E}^{(\alpha)}[Q^*(s, \alpha_s, p^{(\alpha)}(s)) \cdot X_{s-}] ds \\
= p^\circ + \int_0^t \sum_{j=1}^m p^{(\alpha)}_j(s) \mathbb{E}^{(\alpha)}[Q^*(s, \alpha_s, p^{(\alpha)}(s)) \cdot e_j | X_{s-} = e_j] ds.
\]

Expanding further the coordinates \( p^{(\alpha)}_j(t) \) of \( p^{(\alpha)}(t) \) and using the linear structure of \( Q \), we have:

\[
p^{(\alpha)}_i(t) = p^\circ_i + \int_0^t \sum_{j=1}^m p^{(\alpha)}_j(s) \mathbb{E}^{(\alpha)}[e_j \cdot Q(s, \alpha_s, p^{(\alpha)}(s)) \cdot e_i | X_{s-} = e_j] ds \\
= p^\circ_i + \int_0^t \sum_{j=1}^m p^{(\alpha)}_j(s) \mathbb{E}^{(\alpha)}[q(s, j, i, \alpha_s, p^{(\alpha)}(s)) | X_{s-} = e_j] ds \\
= p^\circ_i + \int_0^t \sum_{j \neq i} p^{(\alpha)}_j(s) \mathbb{E}^{(\alpha)}[q(s, j, i, \alpha_s, p^{(\alpha)}(s)) | X_{s-} = e_j] ds \\
- \int_0^t \sum_{j \neq i} p^{(\alpha)}_i(s) \mathbb{E}^{(\alpha)}[q(s, i, j, \alpha_s, p^{(\alpha)}(s)) | X_{s-} = e_i] ds \\
= p^\circ_i + \int_0^t \sum_{j \neq i} p^{(\alpha)}_j(s) \left[ \lambda_{i,j}(\mathbb{E}^{(\alpha)}[\alpha_s | X_{s-} = e_j] - \alpha) + \bar{q}_{i,j}(s, p^{(\alpha)}(s)) \right] ds \\
- \int_0^t \sum_{j \neq i} p^{(\alpha)}_i(s) \left[ \lambda_{i,j}(\mathbb{E}^{(\alpha)}[\alpha_s | X_{s-} = e_i] - \alpha) + \bar{q}_{i,j}(s, p^{(\alpha)}(s)) \right] ds.
\]

We see that the dynamic of \( p^{(\alpha)} \) is completely driven by the deterministic processes \( t \rightarrow \mathbb{E}^{(\alpha)}[\alpha_t | X_{t-} = e_i] \) for \( i \in E \). We shall denote \( \tilde{\alpha}_i^t := \mathbb{E}^{(\alpha)}[\alpha_t | X_{t-} = e_i] \) and \( \tilde{\alpha}_t := [\tilde{\alpha}_1^t, \ldots, \tilde{\alpha}_m^t] \) in the following. On the other hand, by Jensen’s inequality, for all \( \alpha \in \mathbb{A} \) we have:

\[
I(\alpha) = \int_0^T \left[ c_0(t, p^{(\alpha)}(t)) + \sum_{i=1}^m (c_1(t, e_i, p^{(\alpha)}(t)) + \frac{\gamma_i}{2} \mathbb{E}^{(\alpha)}[\alpha_t^2 | X_{t-} = e_i]) p^{(\alpha)}_i(t) \right] dt + C_0(p^{(\alpha)}(T)) \\
\geq \int_0^T \left[ c_0(t, p^{(\alpha)}(t)) + \sum_{i=1}^m (c_1(t, e_i, p^{(\alpha)}(t)) + \frac{\gamma_i}{2} \mathbb{E}^{(\alpha)}[\alpha_t | X_{t-} = e_i]^2 p^{(\alpha)}_i(t) \right] dt + C_0(p^{(\alpha)}(T))
\]

125
This leads to the deterministic control problem:

\[ \tilde{W} := \inf_{\tilde{\alpha} \in \tilde{\mathbb{A}}} \tilde{I}(\tilde{\alpha}) \]  

(4.38)

where we define \( \tilde{\mathbb{A}} \) as the collection of all measurable mappings from \([0, T]\) to \( A^m \), and \( \pi^{(\tilde{\alpha})} \) as the solution to the following system of coupled ODEs:

\[
\frac{d\pi_i^{(\tilde{\alpha})}(t)}{dt} = \sum_{j \neq i} \pi_j^{(\tilde{\alpha})}(t) \left[ \lambda_{j,i}(\tilde{\alpha}_i^j - \alpha) + \tilde{q}_{j,i}(t, \pi^{(\tilde{\alpha})}(t)) \right] \\
- \sum_{j \neq i} \pi_i^{(\tilde{\alpha})}(t) \left[ \lambda_{i,j}(\tilde{\alpha}_i^i - \alpha) + \tilde{q}_{i,j}(t, \pi^{(\tilde{\alpha})}(t)) \right], \quad i \in \{1, \ldots, m\}, \quad (4.39)
\]

\[ \pi^{(\tilde{\alpha})}(0) = p^0, \]

and finally the objective function \( \tilde{I} \) is given by:

\[ \tilde{I}(\tilde{\alpha}) := \int_0^T \left[ c_0(t, \pi^{(\tilde{\alpha})}(t)) + \sum_{i=1}^m c_1(t, e_i, \pi^{(\tilde{\alpha})}(t)) + \frac{\gamma_i}{2} (\tilde{\alpha}_i^i)^2 \pi_i^{(\tilde{\alpha})}(t) \right] dt + C_0(\pi^{(\tilde{\alpha})}(T)). \quad (4.40) \]

The following results show that \( W \) and \( \tilde{W} \) are two equivalent optimization problems.

**Proposition 4.3.1.** We have \( W = \tilde{W} \). If \( \tilde{\alpha} \) is a solution to the optimization problem \( \tilde{W} \), then the predictable process \( \alpha \) defined by \( \alpha_t := \sum_{i=1}^m \mathbb{1}(X_{t-} = e_i)\tilde{\alpha}_i^i \) is an optimal control of \( W \). Moreover, under the probability measure at optimum, the agent’s state evolves as a continuous-time Markov chain.

**Proof.** From the derivation above, we see that \( W \geq \tilde{W} \). We now show that \( \tilde{W} \geq W \).

Given any \( \tilde{\alpha} \in \tilde{\mathbb{A}} \), we set:

\[ \alpha_t := \sum_{i=1}^m \mathbb{1}(X_{t-} = e_i)\tilde{\alpha}_i^i. \quad (4.41) \]

Clearly we have \( \alpha \in \mathbb{A} \). By standard results on ordinary differential equation, equation (4.39) admits a unique solution \( (\pi^{(\tilde{\alpha})}(t))_{t \in [0, T]} \). Now let us consider the
probability measure ˜\(\mathbb{P}\) defined by:

\[
\frac{d\hat{\mathbb{P}}}{d\mathbb{P}} = \hat{\mathcal{E}}_T, \quad (4.42)
\]

\[
d\hat{\mathcal{E}}_t = \hat{\mathcal{E}}_{t-} \cdot (Q(t, \alpha_t, \pi^{(\hat{\alpha})}(t)) - Q^0) \cdot \psi^+_t \cdot d\mathcal{M}_t, \quad \hat{\mathcal{E}}_0 = 1.
\]

It is easy to see that under ˜\(\mathbb{P}\), \(X\) has the decomposition:

\[
X_t = X_0 + \int_0^t Q^*(s, \alpha_s, \pi^{(\hat{\alpha})}(s)) \cdot X_s \cdot ds + \hat{\mathcal{M}}_t = X_0 + \int_0^t \hat{Q}^*(s) \cdot X_s \cdot ds + \hat{\mathcal{M}}_t,
\]

where \(\hat{\mathcal{M}}_t\) is a martingale and \(\hat{Q}(t)\) is the q-matrix with components \(\hat{Q}_{ij}(t) = q(t, i, j, \hat{\alpha}_t, \pi^{(\hat{\alpha})}(t))\). This implies that \(X\) is a continuous-time Markov chain under ˜\(\mathbb{P}\) and it is then straightforward to write the Kolmogorov equation for the law of \(X\) under ˜\(\mathbb{P}\). Comparing this Kolmogorov equation with the ODE (4.39), we conclude by the uniqueness of the solution that ˜\(\mathbb{P}\)[\(X_t\)] = \(\pi^{(\hat{\alpha})}(t)\), where ˜\(\mathbb{E}\) stands for the expectation under ˜\(\mathbb{P}\). Now in light of equation (4.42), we conclude that ˜\(\mathbb{P}\) is the solution to the McKean-Vlasov SDE defined in (4.37) corresponding to the control \(\alpha\). It follows that ˜\(\mathbb{P}\) = \(\mathbb{P}^{(\alpha)}\) and \(\pi^{(\hat{\alpha})}(t) = p^{(\alpha)}(t)\). We now compute  \(I(\alpha)\):

\[
I(\alpha) = \mathbb{E}^{(\alpha)} \left[ \int_0^T \left[ c_0(t, p^{(\alpha)}(t)) + \sum_{i=1}^m 1(X_t = e_i) [c_1(t, e_i, p^{(\alpha)}(t)) + \frac{\gamma_i}{2} \alpha_t^2] \right] dt + C_0(p^{(\alpha)}(T)) \right]
\]

\[
= \int_0^T \left[ c_0(t, p^{(\alpha)}(t)) + \sum_{i=1}^m c_1(t, e_i, p^{(\alpha)}(t)) \mathbb{P}^{(\alpha)}[X_t = e_i] + \sum_{i=1}^m \frac{\gamma_i}{2} (\hat{\alpha}_t)^2 \mathbb{P}^{(\alpha)}[X_t = e_i] \right] dt
\]

\[
+ C_0(p^{(\alpha)}(T))
\]

\[
= \int_0^T \left[ c_0(t, \pi^{(\hat{\alpha})}(t)) + \sum_{i=1}^m c_1(t, e_i, \pi^{(\hat{\alpha})}(t)) \pi^{(\hat{\alpha})}_t(t) + \sum_{i=1}^m \frac{\gamma_i}{2} (\hat{\alpha}_t)^2 \pi^{(\hat{\alpha})}_t(t) \right] dt + C_0(\pi^{(\hat{\alpha})}(T))
\]

\[
= \hat{I}(\hat{\alpha}).
\]

From this, we deduce that \(\hat{I}(\hat{\alpha}) = I(\alpha) \geq W\) and thus \(\tilde{W} \geq W\). Therefore we have \(\tilde{W} = W\). Let \(\hat{\alpha} \in \tilde{\mathcal{A}}\) be the optimizer of \(\tilde{W}\) and set \(\alpha \in \mathcal{A}\) as in (4.41). Then from the computations above, we see that \(\tilde{W} = \hat{I}(\hat{\alpha}) = I(\alpha)\). This immediately implies \(I(\alpha) = W\) and \(\alpha\) is the optimal control of \(W\). \(\square\)
4.3.3 Construction of the Optimal Contract

We continue to investigate the deterministic control problem $\tilde{W}$ as defined in equations (4.38)-(4.40). Once we have identified the optimal strategy of the agents at the equilibrium from the control problem $\tilde{W}$, we can then provide a semi-explicit construction for the optimal contract.

**Lemma 4.3.2.** An optimal control exists for the deterministic control problem $\tilde{W}$.

*Proof.* It is straightforward to verify that: (1) The space of controls is convex and compact. (2) The right-hand side of the ODE (4.39) is $C^1$ and is linear in $\tilde{\alpha}$. (3) The running cost is $C^1$ and convex in $\alpha$ for all $(t, \pi) \in [0, T] \times S$ and the terminal cost is $C^1$. This allows us to apply Theorem I.11.1 in Fleming and Soner [2006] and obtain the existence of the optimal control. \hfill \qed

Having verified that $\tilde{W}$ admits an optimal solution, we now apply the necessary part of the Pontryagin maximum principle (see for example Theorem I.6.3 in Fleming and Soner [2006]), and derive a system of ODEs that characterizes the optimal control and the corresponding flow of probability measures. The Hamiltonian of the control problem $\tilde{W}$ is a mapping from $[0, T] \times S \times \mathbb{R}^m \times A^m$ to $\mathbb{R}$ defined by:

$$
\tilde{H}(t, \pi, y, \alpha) := \sum_{i=1}^{m} \sum_{j \neq i} y_i \left[ \pi_j \left( \lambda_{i,j} (\alpha_j - \bar{\alpha}) + \bar{q}_{i,j}(t, \pi) \right) - \pi_i \left( \lambda_{i,j}(\alpha_i - \bar{\alpha}) + \bar{q}_{i,j}(t, \pi) \right) \right] + c_0(t, \pi) + \sum_{i=1}^{m} \pi_i \left[ c_1(t, e_i, \pi) + \frac{\gamma_i}{2}(\alpha_i)^2 \right].
$$

It is straightforward to obtain:

$$
\partial_{\alpha_i} \tilde{H}(t, \pi, y, \alpha) = \pi_i \left( \sum_{k \neq i} (y_k - y_i) \lambda_{i,k} + \gamma_i \alpha_i \right),
$$

128
and the minimizer of $\alpha \to \tilde{H}(t, \pi, y, \alpha)$ is a vector $[\hat{a}_1(y), \ldots, \hat{a}_m(y)]^*$ where $\hat{a}_i(y)$ is given by:

$$\hat{a}_i(y) := b(-\sum_{k \neq i} \lambda_{i,k}(y_k - y_i)/\gamma_i),$$

and $b(z) := \min\{\max\{z, \alpha\}, \bar{\alpha}\}$. By the necessary condition of Pontryagin’s maximum principle, if $(\pi(t))_{t \in [0,T]}$ is the flow of measures associated with the optimal control, and $(y(t))_{t \in [0,T]}$ is the adjoint process associated with $\pi$, then $(\pi(t), y(t))_{t \in [0,T]}$ is the solution to the following system of forward-backward ODEs:

$$\frac{d\pi(t)}{dt} = \partial_y \tilde{H}(t, \pi(t), y(t), \hat{a}(y(t))), \quad \pi(0) = p^\circ, \quad (4.43)$$

$$\frac{dy(t)}{dt} = -\partial_\pi \tilde{H}(t, \pi(t), y(t), \hat{a}(y(t))), \quad y(T) = \nabla C_0(\pi(T)). \quad (4.44)$$

Summarizing the above discussion, as well as the arguments which allows us to reduce the optimal contracting problem to the deterministic control problem we just solve, we provide a semi-explicit construction of the optimal contract and the optimal strategy of the agents at the equilibrium.

**Theorem 4.3.3.** Let $(\hat{\pi}, \hat{y})$ be the solution to the system (4.43)-(4.44). We define the process $\hat{\alpha} \in \mathbb{A}$ and $\hat{Z} \in \mathcal{H}_X^2$ by:

$$\hat{\alpha}_t := \sum_{i=1}^m \mathbb{1}(X_{t-} = e_i)\hat{a}_i(\hat{y}(t)), \quad (4.45)$$

$$\hat{Z}_t := [\hat{Z}_1^t, \ldots, \hat{Z}_m^t]^*, \quad \hat{Z}_i^t := \sum_{j \neq i} \mathbb{1}(X_{t-} = e_j)\frac{\gamma_j\hat{a}_j(\hat{y}(t))}{(m-1)\sum_{k \neq j} \lambda_{j,k}}. \quad (4.46)$$

Let $Y_0 \in \mathbb{R}^m$ such that $(p^\circ)^* \cdot Y_0 \leq \kappa$ and $\hat{r}$ be the minimizer of the mapping $r \to r - u(r)$. We define the random variable $\hat{\xi}$ by the following Stieltjes integral:

$$\hat{\xi} := -X_0^* \cdot Y_0 + \int_0^T [c(t, X_{t-}, \hat{\alpha}_t, \hat{\pi}(t)) - \hat{r} + X_{t-}^* \cdot Q(t, \hat{\alpha}_t, \hat{\pi}(t)) \cdot \hat{Z}_t]dt - \int_0^T \hat{Z}_t^* \cdot dX_{t-}. \quad (4.47)$$
Then \((\hat{r}, \hat{\xi})\) is an optimal contract. Moreover, under the optimal contract, every agent adopts the Markovian strategy where they pick the control \(\hat{a}_i(t, \hat{y}(t))\) when in the state \(i\) at time \(t\), and the flow of distributions of agents’ states is given by \((\hat{\pi}(t))_{t \in [0,T]}\).

### 4.4 Application to Epidemic Containment

To illustrate the linear quadratic model studied above, we consider an example of epidemic containment. We imagine a disease control authority aims at containing the spread of a virus over a time period \(T\) within its jurisdiction, which consists of two cities \(A\) and \(B\). Each individual’s state is encoded by whether it is infected (denoted by \(I\)) or healthy (denoted by \(H\)), and by its location (denoted by \(A\) or \(B\)). Therefore the state space is \(E = \{AI, AH, BI, BH\}\), and we use \(\pi_{AI}, \pi_{AH}, \pi_{BI}, \pi_{BH}\) to denote the proportion of individuals in each of these 4 states. We assume that each individual’s state evolves as a continuous-time Markov chain, and our modeling of the transition rate accounts for the following set of mechanisms regarding the contraction of the virus and the migration of the individuals between the two cities:

1. Within each city, the rate of contracting the virus depends on the proportion of infected individuals in the city, and accordingly we assume that the transition rate from state \(AH\) to state \(AI\) is \(\theta_A^-(\frac{\pi_{AI}}{\pi_{AI} + \pi_{AH}})\) and the transition rate from state \(BH\) to state \(BI\) is \(\theta_B^-(\frac{\pi_{BI}}{\pi_{BI} + \pi_{BH}})\). Here \(\theta_A^-\) and \(\theta_B^-\) are two increasing, positive and differentiable mappings from \([0,1]\) to \(\mathbb{R}^+\), and are related to the quality of health care in city \(A\) and \(B\), respectively.

2. Likewise, the rate of recovery in each city is a function of the proportion of healthy individuals in the city. We thus assume that the transition rate from state \(AI\) to state \(AH\) is \(\theta_A^+\left(\frac{\pi_{AH}}{\pi_{AI} + \pi_{AH}}\right)\) and the transition rate from state \(BH\) to state \(BI\) is \(\theta_B^+\left(\frac{\pi_{BI}}{\pi_{BI} + \pi_{BH}}\right)\). Similarly, \(\theta_A^+\) and \(\theta_B^+\) are two increasing, positive and differentiable
mappings from $[0, 1]$ to $\mathbb{R}^+$, characterizing the quality of health care in city $A$ and $B$ respectively.

(3) Each individual can choose an effort $\alpha$ which will give it the chance to move to the other city. We assume that the efficacy of that effort depends on whether the individual is healthy or infected. Accordingly, we set $\nu_I\alpha$ as the transition rates between the states $AI$ and $BI$, and we set $\nu_H\alpha$ as the transition rates between the states $AH$ and $BH$.

(4) To model the inflow of infection, we assume that each individual’s status of infection does not change when it moves between the cities. This means that we set the transition rates between the state $AI$ and $BH$, and the transition rates between the state $AH$ and $BI$ to 0.

To summarize, we define the transition rate matrix $Q$ to be:

$$Q(t, \alpha, \pi) := \begin{bmatrix} \cdots & \theta_A^+\left(\frac{\pi_{AH}}{\pi_{AI} + \pi_{AH}}\right) & \nu_I\alpha & 0 \\ \theta_A\left(\frac{\pi_{AI}}{\pi_{AI} + \pi_{AH}}\right) & \cdots & 0 & \nu_H\alpha \\ \nu_I\alpha & 0 & \cdots & \theta_B^+\left(\frac{\pi_{BH}}{\pi_{BI} + \pi_{BH}}\right) \\ 0 & \nu_H\alpha & \theta_B\left(\frac{\pi_{BI}}{\pi_{BI} + \pi_{BH}}\right) & \cdots \end{bmatrix}. \quad (4.48)$$

Here the transition rate matrix is written with the order of states $AI, AH, BI, BH$. For simplicity of the notation, we also omit the diagonal elements. Indeed, since $Q$ is a transition rate matrix, each diagonal element equals the opposite of the sum of off-diagonal elements in the same row.

We resume the description of our model in terms of the cost functions for the individuals and the disease control authority. We assume that individuals living in the same city incur the same cost, which depends on the proportion of infected individuals in that city. On the other hand, the cost for exerting the effort to move depends on the status of infection. Using the notations of the cost function for the
linear quadratic model as in (4.31), we define:

\[
c_1(t, AI, \pi) = c_1(t, AH, \pi) := \phi_A \left( \frac{\pi_{AI}}{\pi_{AI} + \pi_{AH}} \right), \tag{4.49}
\]

\[
c_1(t, BI, \pi) = c_1(t, BH, \pi) := \phi_B \left( \frac{\pi_{BI}}{\pi_{BI} + \pi_{BH}} \right), \tag{4.50}
\]

\[
\gamma_{AI} = \gamma_{BI} := \gamma_I, \quad \gamma_{AH} = \gamma_{BH} := \gamma_H, \tag{4.51}
\]

where \( \phi_A \) and \( \phi_B \) are two increasing mappings on \( \mathbb{R} \). For the authority, we propose to use the following running cost and terminal cost:

\[
c_0(t, \pi) = \exp(\sigma_A \pi_{AI} + \sigma_B \pi_{BI}), \tag{4.52}
\]

\[
C_0(\pi) = \sigma_P \cdot (\pi_{AI} + \pi_{AH} - \pi_{A}^0)^2, \tag{4.53}
\]

where \( \pi_{A}^0 \) is the population of city \( A \) at time 0. Intuitively speaking, the above cost function translates the trade-off between the epidemic control and the population planning. On the one hand, the authority attempts to minimize the infection rate of both cities. On the other hand, as we shall see shortly in the numerical simulation, individuals tends to move away from the city with higher infection rate and poorer health care, which might result in the overpopulation of the other city. Therefore, the authority also wishes to maintain the population of both cities at a steady level. The coefficients \( \sigma_A, \sigma_B \) and \( \sigma_P \) reflect the importance the authority attributes to each of these objectives.

It can be easily verified that the setup outlined above satisfies the assumptions of the linear quadratic model studied in Section 4.3. Although the forward-backward system of ODEs (4.43) - (4.44) which characterizes the optimal contract can be readily derived, for the sake of completeness, we shall give the details of the equations to be solved in the appendix of this chapter.
For the numerical simulations, we consider a scenario in which city $A$ has a higher quality of health care than city $B$. Accordingly we set:

$$
\theta_A^+(q) := 0.4 \cdot q, \quad \theta_A^-(q) := 0.1 \cdot q, \quad \theta_B^+(q) := 0.2 \cdot q, \quad \theta_B^-(q) := 0.2 \cdot q. \quad (4.54)
$$

This means that it is easier to recover and harder to get infected in city $A$ than in city $B$. In addition, we assume that individuals suffer a higher cost associated with the epidemic in city $B$ than in city $A$, and we set the cost function of individuals in each city to be:

$$
\phi_A(q) := q, \quad \phi_B(q) := 2 \cdot q. \quad (4.55)
$$

We set the maximal possible effort of individuals to $\bar{\alpha} := 10$ and the coefficients for quadratic cost of efforts to $\gamma_I := 2.0$ and $\gamma_H = 0.5$. Finally, the parameters for the cost of the authority in equation (4.52) and (4.53) are set to:

$$
\sigma_A = \sigma_B := 1, \quad \sigma_P := 0. \quad (4.56)
$$

Notice that the authority has attributed the same importance to the infection rates of city $A$ and city $B$ while disregarding the problem of overpopulation.

In the following, we shall visualize the effect of the disease control authority’s intervention by comparing the equilibrium computed from the principal agent problem with the equilibrium from the mean field game of anarchy. By the term anarchy, we refer to the situation where the states of individuals are still governed by the same transition rate, but the individuals do not receive any rewards or penalties from the authority. More specifically, the total expected cost of each individual is given by:

$$
\mathbb{E}^{Q(\alpha, \pi)} \left[ \int_0^T c(t, X_t, \alpha_t, \pi(t)) dt \right],
$$

133
with the instantaneous cost $c$ is given by:

$$c(t, X, \alpha, \pi) := \mathbb{1}(X \in \{AI, AH\}) \cdot \phi_A \left( \frac{\pi_{AI}}{\pi_{AI} + \pi_{AH}} \right) + \mathbb{1}(X \in \{BI, BH\}) \cdot \phi_B \left( \frac{\pi_{BI}}{\pi_{BI} + \pi_{BH}} \right) + (\mathbb{1}(X \in \{AI, BI\})\gamma_I + \mathbb{1}(X \in \{AH, BH\})\gamma_H) \cdot \frac{\alpha^2}{2}.$$

Applying the analytical approach of finite-state mean field games which we have presented in Section 2.3 of Chapter 2, it is straightforward to derive the system of forward-backward ODEs characterizing the Nash equilibrium (see equations (2.49)-(2.50)). We display this system of ODEs in the appendix of this chapter (see Section 4.5.2).

The effect of the authority’s intervention is conspicuous in diminishing the infection rate among the entire population, as is shown in the upper panel of Figure 4.1. However, when we visualize the respective infection rate of each city in the lower panels of Figure 4.1, we do observe a surge of infection in city $A$ as a result of the authority’s intervention, although the epidemic eventually dies down. This is due to the inflow of infected individuals from city $B$ in the early stage of the epidemic. Indeed, since city $A$ provides better health care and maintains a lower rate of infection compared to city $B$, an infected individual has a better chance of recovery if it moves to city $A$. The cost of moving prevents individuals from seeking better health care in the scenario of anarchy, whereas individuals seem to receive subsidy to move when the authority tries to intervene. This outcome of the authority’s intervention is further corroborated when we visualize an individual’s optimal strategy in Figure 4.2. We observe that individuals have a greater propensity to move when the authority provides incentives.
Figure 4.1: Evolution of infection rate with and without authority's intervention.

Figure 4.2: Optimal effort of moving for an individual in different states.
Recall that in the above numerical computation, by setting $\sigma_P = 0$ in its terminal cost function (4.53), the authority does not attempt to maintain the balance of population between the two cities. We now investigate how the behavior of the individuals changes when the authority seeks to prevent the occurrence of overpopulation. To this end, we redo the computations with $\sigma_P = 1.5$ and all the other parameters unchanged.

![Graph](image)

**Figure 4.3:** Evolution of population in city A and the total infection rate. The solid curve corresponds to intervention without population planning ($\sigma_P = 0$), the dashed-dotted curve corresponds to intervention with population planning ($\sigma_P = 1.5$), and the dashed curve corresponds to the absence of intervention.

In Figure 4.3, we compare the evolution of the population in city A as well as the total infection rate with and without the population planning. When the authority does not try to control the flow of the population, the entire population ends up in city A. However, when a terminal cost related to population planning is introduced, we see a more balanced population distribution while the infection rate is still well managed. This can be explained by Figure 4.4, from which we have a more detailed perspective on the change of individual behavior when the authority implements the population planning. We see that healthy individuals in city A are now encouraged...
to move to city $B$, in order to compensate the exodus caused by the epidemic in city $B$. On the other hand, healthy individuals in city $B$ are now incentivized to stay in place.

Figure 4.4: Individual’s optimal effort of moving with and without population planning.

4.5 Appendix: System of ODEs for Epidemic Containment

4.5.1 Authority’s optimal planning

Using the transition rate (4.48) and the cost functions (4.49)-(4.53), the system of ODEs (4.43)-(4.44) becomes:

$$
\dot{y}_0(t) = - \frac{[y_0(t) - y_1(t)]\pi_1(t)}{(\pi_0(t) + \pi_1(t))^2} \cdot \left[ \pi_1(t)(\theta_A^-)' \left( \frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)} \right) + \pi_0(t)(\theta_A^+)' \left( \frac{\pi_1(t)}{\pi_0(t) + \pi_1(t)} \right) \right]
$$
\[
\begin{align*}
\dot{y}_1(t) &= -\frac{[y_1(t) - y_0(t)]\pi_0(t)}{(\pi_0(t) + \pi_1(t))^2} \cdot \left[ \pi_1(t)\theta_A^+ \left( \frac{\pi_1(t)}{\pi_0(t) + \pi_1(t)} \right) - y_2(t)\nu_1\hat{A}_0(y(t)) - \sigma_A \exp(\sigma_A\pi_0(t) + \sigma_B\pi_2(t)) \right] \\
&\quad - \frac{1}{2} \gamma_I((\hat{a}_0(y(t)))^2 - \phi_A \left( \frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)} \right) - \phi_A' \left( \frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)} \right) \cdot \frac{\pi_1(t)}{\pi_0(t) + \pi_1(t)}, \\
\dot{y}_2(t) &= -\frac{[y_2(t) - y_3(t)]\pi_2(t)}{(\pi_2(t) + \pi_3(t))^2} \cdot \left[ \pi_2(t)\theta_B^+ \left( \frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right) - y_3(t)\nu_1\hat{A}_2(y(t)) - \sigma_B \exp(\sigma_A\pi_0(t) + \sigma_B\pi_2(t)) \right] \\
&\quad - \frac{1}{2} \gamma_I((\hat{a}_2(y(t)))^2 - \phi_B \left( \frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right) - \phi_B' \left( \frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right) \cdot \frac{\pi_3(t)}{\pi_2(t) + \pi_3(t)}, \\
\dot{y}_3(t) &= -\frac{[y_3(t) - y_2(t)]\pi_3(t)}{(\pi_2(t) + \pi_3(t))^2} \cdot \left[ \pi_3(t)\theta_B^+ \left( \frac{\pi_3(t)}{\pi_2(t) + \pi_3(t)} \right) - y_1(t)\nu_1\hat{A}_2(y(t)) - \sigma_B \exp(\sigma_A\pi_0(t) + \sigma_B\pi_2(t)) \right] \\
&\quad - \frac{1}{2} \gamma_I((\hat{a}_3(y(t)))^2 - \phi_B \left( \frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right) - \phi_B' \left( \frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right) \cdot \frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)}, \\
\pi_0(t) &= \pi_1(t)\theta_A^- \left( \frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)} \right) + \pi_2(t)\nu_1\hat{A}_2(y(t)) - \pi_0(t) \left[ \theta_A^- \left( \frac{\pi_1(t)}{\pi_0(t) + \pi_1(t)} \right) + \nu_1\hat{A}_0(y(t)) \right], \\
\pi_1(t) &= \pi_0(t)\theta_A^+ \left( \frac{\pi_1(t)}{\pi_0(t) + \pi_1(t)} \right) + \pi_3(t)\nu_1\hat{A}_3(y(t)) - \pi_1(t) \left[ \theta_A^- \left( \frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)} \right) + \nu_1\hat{A}_1(y(t)) \right], \\
\pi_2(t) &= \pi_3(t)\theta_B^- \left( \frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right) + \pi_0(t)\nu_1\hat{A}_0(y(t)) - \pi_2(t) \left[ \theta_B^- \left( \frac{\pi_3(t)}{\pi_2(t) + \pi_3(t)} \right) + \nu_1\hat{A}_2(y(t)) \right], \\
\pi_3(t) &= \pi_2(t)\theta_B^+ \left( \frac{\pi_3(t)}{\pi_2(t) + \pi_3(t)} \right) + \pi_1(t)\nu_1\hat{A}_1(y(t)) - \pi_3(t) \left[ \theta_B^- \left( \frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right) + \nu_1\hat{A}_3(y(t)) \right],
\end{align*}
\]
where the optimal control is defined by:

\[
\hat{a}_0(y) := b \left( \frac{\nu_I (y_0 - y_2)}{\gamma_I} \right), \quad \hat{a}_1(y) := b \left( \frac{\nu_H (y_1 - y_3)}{\gamma_H} \right), \\
\hat{a}_2(y) := b \left( \frac{\nu_I (y_2 - y_0)}{\gamma_I} \right), \quad \hat{a}_3(y) := b \left( \frac{\nu_H (y_3 - y_1)}{\gamma_H} \right),
\]

and the terminal conditions are \( \pi(0) = \pi^0 \) and \( y(T) = \nabla C_0(\pi(T)) \).

### 4.5.2 Mean field equilibrium in the absence of the authority

The system of ODEs characterizing the mean field game equilibrium consists of the Hamilton-Jacobi equation and the Kolmogorov equation.

\[
v_0(t) = [v_1(t) - v_0(t)] \theta^+_A \left( \frac{\pi_1(t)}{\pi_0(t) + \pi_1(t)} \right) + [v_2(t) - v_0(t)] \nu_I \hat{a}_0(v(t)) + \frac{1}{2} \gamma_I (\hat{a}_0(v(t)))^2 \\
+ \phi_A \left( \frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)} \right),
\]

\[
v_1(t) = [v_0(t) - v_1(t)] \theta^-_A \left( \frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)} \right) + [v_3(t) - v_1(t)] \nu_H \hat{a}_1(v(t)) + \frac{1}{2} \gamma_H (\hat{a}_1(v(t)))^2 \\
+ \phi_A \left( \frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)} \right),
\]

\[
v_2(t) = [v_3(t) - v_2(t)] \theta^+_B \left( \frac{\pi_3(t)}{\pi_2(t) + \pi_3(t)} \right) + [v_0(t) - v_2(t)] \nu_I \hat{a}_2(v(t)) + \frac{1}{2} \gamma_I (\hat{a}_2(v(t)))^2 \\
+ \phi_B \left( \frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right),
\]

\[
v_3(t) = [v_2(t) - v_3(t)] \theta^-_B \left( \frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right) + [v_1(t) - v_3(t)] \nu_H \hat{a}_3(v(t)) + \frac{1}{2} \gamma_H (\hat{a}_3(v(t)))^2 \\
+ \phi_B \left( \frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right),
\]
\[ \dot{\pi}_0(t) = \pi_1(t) \theta_A^- \left( \frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)} \right) + \pi_2(t) \nu_I \hat{a}_2(v(t)) - \pi_0(t) \left[ \theta_A^+ \left( \frac{\pi_1(t)}{\pi_0(t) + \pi_1(t)} \right) + \nu_I \hat{a}_0(v(t)) \right] , \]

\[ \dot{\pi}_1(t) = \pi_0(t) \theta_A^+ \left( \frac{\pi_1(t)}{\pi_0(t) + \pi_1(t)} \right) + \pi_3(t) \nu_H \hat{a}_3(v(t)) - \pi_1(t) \left[ \theta_A^- \left( \frac{\pi_0(t)}{\pi_0(t) + \pi_1(t)} \right) + \nu_H \hat{a}_1(v(t)) \right] , \]

\[ \dot{\pi}_2(t) = \pi_3(t) \theta_B^- \left( \frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right) + \pi_0(t) \nu_I \hat{a}_0(v(t)) - \pi_2(t) \left[ \theta_B^+ \left( \frac{\pi_3(t)}{\pi_2(t) + \pi_3(t)} \right) + \nu_I \hat{a}_2(v(t)) \right] , \]

\[ \dot{\pi}_3(t) = \pi_2(t) \theta_B^+ \left( \frac{\pi_3(t)}{\pi_2(t) + \pi_3(t)} \right) + \pi_1(t) \nu_H \hat{a}_1(v(t)) - \pi_3(t) \left[ \theta_B^- \left( \frac{\pi_2(t)}{\pi_2(t) + \pi_3(t)} \right) + \nu_H \hat{a}_3(v(t)) \right] , \]

where the optimal control is defined by:

\[ \hat{a}_0(v) := b \left( \frac{\nu_I (v_0 - v_2)}{\gamma_I} \right) , \quad \hat{a}_1(v) := b \left( \frac{\nu_H (v_1 - v_3)}{\gamma_H} \right) , \]

\[ \hat{a}_2(v) := b \left( \frac{\nu_I (v_2 - v_0)}{\gamma_I} \right) , \quad \hat{a}_3(v) := b \left( \frac{\nu_H (v_3 - v_1)}{\gamma_H} \right) , \]

and the terminal conditions are \( \pi(0) = p^o \) and \( v(T) = 0 \).
Chapter 5

Finite State Mean Field Game
with Major and Minor Players

Mean field games with major and minor players were introduced to accommodate the presence of one single player whose influence on the behavior of the remaining population does not vanish in the asymptotic regime of large games. In this chapter we develop the theory of these dynamic games when the states of the players belong to a finite space. In Chapter 2 and Chapter 3, we have given a brief account of the theory of finite state mean field games for a single homogeneous population of players via both the analytical and the probabilistic method. The present chapter is concerned with the extension of the analytical method to models with major and minor players. We search for closed-loop Nash equilibria and for this reason, we use the approach which was advocated in Carmona and Wang [2017], and called an alternative approach in Chapter 13 of Carmona and Delarue [2017].

Our interest in mean field games with major and minor players when the state space is finite was sparked by the four-state model by Kolokoltsov and Bensoussan [2016] for the behavior of computer owners facing cyber attacks. Even though the model was not introduced and treated as a game with major and minor players,
clearly, it is of this type if the behaviors of the attacker and the targets are strategic. Practical applications amenable to these models abound and a better theoretical understanding of their structures should lead to sorely needed numerical procedures to compute Nash equilibria.

Early forms of mean field games with major and minor players appeared in Huang [2010] in an infinite-horizon setting, in Nguyen and Huang [2012a] for finite time horizons, and Nourian and Caines [2013] offered a first generalization to non linear-quadratic cases. In these models, the state of the major player does not enter the dynamics of the states of the minor players: it only appears in their cost functionals. This was remedied in Nguyen and Huang [2012b] for linear quadratic models. The recent technical report Jaimungal and Nourian [2015] adds a major player to the particular case (without idiosyncratic random shocks) of an extended mean field game model of optimal execution introduced in Chapter 1 and solved in Chapter 4 of Carmona and Delarue [2017].

The asymmetry between major and minor players was emphasized in Bensoussan et al. [2016a] where the authors insist on the fact that the statistical distribution of the state of a generic minor player should be derived endogenously. Like Nourian et al. [2011], Bensoussan et al. [2016a] characterizes the limiting problem by a set of stochastic partial differential equations. However, Bensoussan et al. [2016a] seems to be solving a Stackelberg game, a particular type of equilibrium we treated in Chapter 4. Recall that in a Stackelberg game, the major player first chooses an action and the population of minor players then reach a Nash equilibrium according to the major player’s strategy. A Stackelberg equilibrium is reached when the major player optimizes its cost function which depends on the distribution of the minor players’ states at the Nash equilibrium. In some sense, a Stackelberg game can be formulated as an optimal control problem of Nash equilibria. This is in contrast with
Carmona and Zhu [2016] which also insists on the endogenous nature of the statistical distribution of the state of a generic minor player, but which formulates the search for a mean field equilibrium as the search for a Nash equilibrium in a two player game over the time evolutions of states, some of which being of McKean-Vlasov type.

In this chapter, we cast the search for Nash equilibria as a search for fixed points of the best response function constructed from the optimization problems of both types of players. Typically, in a mean field game with major and minor players, the dynamics of the state $X_t^0$ and the cost functions of the major player depend upon the statistical distribution $\mu_t$ of the state $X_t$ of a generic minor player. Meanwhile, the dynamics of the state $X_t$ and the cost of a generic minor player depend upon the values of the state $X_t^0$ and the control $\alpha_t^0$ of the major player as well as the statistical distribution $\mu_t$ which captures the mean field interactions between the minor players. In the rest of this chapter, we shall refer to this generic minor player as the deviating minor player, since our objective is to study the optimal response of any given minor player given the strategies of its peers and the major player.

As the first important result in this chapter, we prove that the processes $(X_t^0, \mu_t)$ and $(X_t^0, X_t, \mu_t)$ are Markovian and we characterize their laws by their infinitesimal generators. We start from the finite player version of the model and show convergence when the number of minor players goes to infinity. We rely on standard results from the theory of the convergence of Markov semigroups.

Note that the control of the major player implicitly influences $\mu_t$ through the major player’s state, so the major player’s optimization problem should be treated as an optimal control problem for McKean-Vlasov dynamics. On the other hand, for the deviating minor player’s problem, we are just dealing with a classical Markov decision problem in continuous time. This allows us to adapt to the finite state space the approach introduced in Carmona and Wang [2017] and reviewed in Chapter 13 of
Carmona and Delarue [2017], to define and construct Nash equilibria. We emphasize that these are Nash equilibria for the whole system major + minor players and not only for the minor players. This is fully justified by our results on the propagation of chaos and their applications to the proof that our mean field game equilibria provide approximate Nash equilibria for games with finitely many players, including both major and minor players.

The rest of this chapter is structured as follows. Games with finitely many minor players and a major player are introduced in Section 5.1 where we explain the conventions and notations we use to describe continuous-time controlled Markov processes in finite state spaces. We also identify the major and minor players by specifying the information structures available to them, the types of actions they can take, and the costs they incur. The short and non-technical Section 5.2 describes the mean field game strategy and emphasizes the steps needed in the search for Nash equilibria of the system. This is in contrast with some earlier works where the formulation of the problem lead to Stackelberg equilibria, in which only the minor players end up in an approximate Nash equilibrium. To keep with the intuition that the mean field game strategy is to implement a form of limit when the number of minor players grows to infinity, Section 5.3 considers the convergence of the Markov processes describing the states of the major + minor players system in this limit, and identifies the optimization problems which the major player and the deviating minor player need to solve in order to construct their best responses. This leads to the formalization of the search for a mean field equilibrium as the search of fixed points for the best response map constructed in this way. The optimization problems underpinning the definition of the best response map are studied in Section 5.4. There, we use dynamic programming to prove that the value functions of these optimization problems are viscosity solutions of HJB type Partial Integro-Differential Equations (PIDEs). Section 5.5 proves existence of the best response map and of Nash equilibria under reasonable
conditions. Next, Section 5.6 gives a verification theorem based on the existence of a classical solution to the master equation. The longer Section 5.7 proves that the solution of the mean field game problem provides approximate Nash equilibria for the finite player games. This vindicates our formulation as the right formulation of the problem if the goal is to find Nash equilibria for the system including both major and minor players. The proof is technical but standard in the literature on mean field games and it relies on the results of propagation of chaos. However, the latter are usually derived for stochastic differential systems with mean field interactions, and because we could not find the results we needed in the existing literature, we provide proofs of the main steps of the derivations of these results in the context of controlled Markov evolutions in finite state spaces. Finally, an appendix provides the proofs of some of the technical results we used in the text.

5.1 Game Model with Finitely Many Players

We consider a stochastic game in continuous time, involving a major player indexed by 0, and $N$ minor players indexed from 1 to $N$. The states of all the players $X_0^t, X_1^t, \ldots, X_N^t$ are described by a continuous-time finite-state Markov process. Let us denote $\{1, 2, \ldots, m_0\}$ the set of possible values taken by the state $X_0^t$ of the major player, and $\{1, 2, \ldots, m\}$ the set of possible values taken by the state $X_n^t$ of the minor players. We introduce the empirical distribution of the states of the minor players at time $t$:

$$
\mu_i^N = \left[ \frac{1}{N} \sum_{n=1}^{N} 1(X_n^t = 1), \frac{1}{N} \sum_{n=1}^{N} 1(X_n^t = 2), \ldots, \frac{1}{N} \sum_{n=1}^{N} 1(X_n^t = m - 1) \right].
$$

We denote by $\mathcal{S}$ the following set:

$$
\mathcal{S} := \{ x \in \mathbb{R}^{m-1} | x_i \geq 0, \sum x_i \leq 1 \}.
$$
This is the convex hull of the \((m - 1)\)-dimensional simplex, which can be identified with the space of probability measures on the state space of the minor players. We consider Markovian dynamics in continuous time, according to which the rates of jumps of the state of a generic minor player depend upon the value of its control, the empirical distribution of the states of all the minor players, as well as the major player’s control and state. We denote by \(A_0\) (resp. \(A\)) a convex set in which the major player (resp. all the minor players) can choose their controls. We introduce a function \(q\):

\[
[0, T] \times \{1, \ldots, m\}^2 \times A \times \{1, \ldots, m_0\} \times A_0 \times \bar{S} \rightarrow \mathbb{R}
\]

\[
(t, i, j, \alpha, i_0, \alpha_0, x) \rightarrow q(t, i, j, \alpha, i_0, \alpha_0, x),
\]

and we make the following assumption on \(q\):

**Assumption 5.1.1.** For all \((t, \alpha, i_0, \alpha_0, x) \in [0, T] \times A \times \{1, \ldots, m_0\} \times A_0 \times \bar{S}\), the matrix \([q(t, i, j, \alpha, i_0, \alpha_0, x)]_{1 \leq i, j \leq m}\) is a Q-matrix.

Then we assume that at time \(t\) if the state of minor player \(n\) is \(i\), the player will jump to state \(j\) at a rate given by:

\[
q(t, i, j, \alpha^n_t, X_0^t, \alpha_0^t, \mu^N_t),
\]

if \(X_0^t\) is the major player state, \(\alpha_0^t \in A_0\) is the major player control, \(\alpha^n_t \in A\) is the \(n\)-th minor player control and \(\mu^N_t \in \bar{S}\) is the empirical distribution of the minor player’s states. Our goal is to use these rates to completely specify the law of a continuous-time process for every player’s state in the following way: if at time \(t\) the \(n\)-th minor player is in state \(i\) and uses control \(\alpha^n_t\), if the major player is in state \(X_0^t\) and uses the control \(\alpha_0^t\), and if the empirical distribution of the states of the population of minor players is \(\mu^N_t\), then the probability of player \(n\) remaining in the same state during the
infinitesimal time interval \([t, t + \Delta t]\) is \([1 + q(t, i, i, \alpha_t^n, X_t^0, \alpha_t^0, \mu_t^N)\Delta t + o(\Delta t)]\), whereas the probability of this state changing to another state \(j\) during the same time interval is given by \([q(t, i, j, \alpha_t^n, X_t^0, \alpha_t^0, \mu_t^N)\Delta t + o(\Delta t)]\).

Similarly, to describe the evolution of the state of the major player we introduce a function \(q_0:\)

\[
[0, T] \times \{1, \ldots, m_0\}^2 \times A_0 \times \bar{S} \rightarrow \mathbb{R}
\]

\[
(t, i_0, j_0, \alpha_0, x) \rightarrow q_0(t, i_0, j_0, \alpha_0, x),
\]

which satisfies the following assumption:

**Assumption 5.1.2.** For each \((t, \alpha_0, x) \in [0, T] \times A_0 \times \bar{S}, [q_0(t, i_0, j_0, \alpha_0, x)]_{1 \leq i_0, j_0 \leq m_0}\) is a Q-matrix.

So if at time \(t\) the state of the major player is \(i_0\), its control is \(\alpha_0^0 \in A_0\), and the empirical distribution of the states of the minor players is \(\mu_t^N\), we assume the state of the major player will jump to state \(j_0\) at rate \(q_0(t, i_0, j_0, \alpha_0, x)\).

We now define the control strategies which are admissible to the major and minor players. In our model, we assume that the major player can only observe its own state and the empirical distribution of the states of the minor players, whereas each minor player can observe its own state, the state of the major player as well as the empirical distribution of the states of all the minor players. Furthermore, we only allow for Markovian strategies given by feedback functions. Therefore the major player’s control should be of the form \(\alpha_t^0 = \phi_0(t, X_t^0, \mu_t^N)\) for some feedback function:

\[
\phi_0 : [0, T] \times \{1, \ldots, m_0\} \times \bar{S} \rightarrow A_0
\]

\[(t, i_0, x) \rightarrow \phi_0(t, i_0, x),\]

147
and the control of minor player \( n \) should be of the form \( \alpha^n_t = \phi_n(t, X^n_t, X^0_t, \mu^N_t) \) for some feedback function:

\[
\phi_n : [0, T] \times \{1, \ldots, m\} \times \{1, \ldots, m_0\} \times \bar{S} \rightarrow A
\]

\[
(t, i, i_0, x) \rightarrow \phi_n(t, i, i_0, x).
\]

We denote the sets of admissible control strategies by \( A_0 \) and \( A_n \) respectively. In line with the notations used in Section 2.2.1 of Chapter 2, for each \( n \in \{0, 1, \ldots, N\} \), we denote an element of \( A_n \) by \( \alpha_n \leftrightarrow \phi_n \) to emphasize the role of the feedback function.

When we study the convergence of \( N \)-player game to the mean field game and identify the Nash equilibria for mean field game later in this chapter, we shall consider the admissible strategies of which the feedback function is Lipschitz in the measure. More specifically, we define the collection of Lipschitz strategies for the major and minor players as follow:

\[
L_0 := \{ \alpha_0 \leftrightarrow \phi_0 | \phi_0 \text{ is Lipschitz continuous in } x, \text{ uniformly in } t \text{ and } i_0. \}
\]

\[
L_n := \{ \alpha_n \leftrightarrow \phi_n | \phi_n \text{ is Lipschitz continuous in } x, \text{ uniformly in } t, i \text{ and } i_0. \}
\]

Let us define the joint dynamics of the states of all the players. We assume that conditioned on the current state of the system, the changes of states are independent for different players. This means that for all \( i_0, i_1, \ldots, i_N \) and \( j_0, j_1, \ldots, j_N \), where \( i_0, j_0 \in \{1, 2, \ldots, m_0\} \) and \( i_n, j_n \in \{1, 2, \ldots, m\} \) for \( n = 1, 2, \ldots, N \), we have:

\[
\mathbb{P}[X^0_{t+\Delta t} = j_0, X^1_{t+\Delta t} = j_1, \ldots, X^N_{t+\Delta t} = j_N | X^0_t = i_0, X^1_t = i_1, \ldots, X^N_t = i_N] \\
= [\mathbb{1}(i_0 = j_0) + q_0(t, i_0, j_0, \phi_0(t, i_0, \mu^N_t), \mu^N_t)\Delta t + o(\Delta t)] \\
\times \prod_{n=1}^N [\mathbb{1}(i_n = j_n) + q_n(t, i_n, j_n, \phi_n(t, i_n, i_0, \mu^N_t), i_0, \phi_0(t, i_0, \mu^N_t), \mu^N_t)\Delta t + o(\Delta t)].
\]

Formally, this statement is equivalent to the definition of the Q-matrix, say \( Q^{(N)} \), of the continuous-time Markov chain \( (X^0_t, X^1_t, X^2_t, \ldots X^N_t) \). The state space of this
Markov chain is the Cartesian product of each player’s state space. Therefore $Q^{(N)}$ is a square matrix of size $m_0 \cdot m^N$. The non-diagonal entry of $Q^{(N)}$ can be found by simply retaining the first order term in $\Delta t$ when expanding the above product of probabilities. Because we assume that the transitions of states are independent among the individual players, $Q^{(N)}$ is a sparse matrix.

Each individual player aims to minimize its expected cost in the game. We assume that these costs are given by:

$$J_{0,N}(\alpha_0, \alpha_1, \ldots, \alpha_N) := \mathbb{E} \left[ \int_0^T f_0(t, X^0_t, \phi_0(t, X^0_t, \mu^1_t, \mu^N_t)) dt + g_0(X^0_T, \mu^N_T) \right],$$

$$J_{n,N}(\alpha_0, \alpha_1, \ldots, \alpha_N) := \mathbb{E} \left[ \int_0^T f_n(t, X^n_n, \phi_n(t, X^n_n, X^0_t, \mu^1_t, \mu^N_t), X^0_t, \phi_0(t, X^0_t, \mu^N_t)) dt + g_n(X^n_n, X^0_T, \mu^N_T) \right],$$

where $f_0 : [0, T] \times \{1, \ldots, m_0\} \times A_0 \times \mathcal{S} \to \mathbb{R}$, $g_0 : \{1, \ldots, m_0\} \times \mathcal{S} \to \mathbb{R}$ are respectively the running cost and terminal cost of the major player, and $f_n : [0, T] \times \{1, \ldots, m\} \times A \times \{1, \ldots, m_0\} \times A_0 \times \mathcal{S} \to \mathbb{R}$, $g_n : \{1, \ldots, m\} \times \{1, \ldots, m_0\} \times \mathcal{S} \to \mathbb{R}$ are respectively the running cost and terminal cost of the minor player $n$.

In this chapter, we focus on the special case of symmetric games, for which all the minor players share the same transition rate function and cost function, i.e.: $q_n := q$, $f_n := f$, $g_n := g$, $J_{n,N} := J_N$, and we search for symmetric Nash equilibria. We say that a couple of feedback functions $(\hat{\phi}_0, \hat{\phi})$ form a symmetric Nash equilibrium if the controls $(\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_N)$ given by $\hat{\alpha}_0^0 = \hat{\phi}_0(t, X^0_t, \mu^1_t, \mu^N_t)$ and $\hat{\alpha}_n^0 = \hat{\phi}(t, X^n_n, X^0_t, \mu^1_t, \mu^N_t)$ for $n = 1, \ldots, N$, form a Nash equilibrium in the sense that:

$$J_{0,N}(\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_N) \leq J_{0,N}(\alpha_0, \alpha_1, \ldots, \alpha_N),$$

$$J_{N}(\alpha_0, \alpha_1, \ldots, \alpha_n, \ldots, \alpha_N) \leq J_{N}(\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \alpha_n, \ldots, \hat{\alpha}_N),$$

for any choices of alternative admissible controls $\alpha_0$ and $\alpha_n$ of the forms $\alpha_0^0 = \phi_0(t, X^0_t, \mu^N_t)$ and $\alpha_n^0 = \phi(t, X^n_n, X^0_t, \mu^N_t)$.
In order to simplify the notations, we will systematically use the following notations when there is no risk of possible confusion. When $\alpha_0 \in \mathcal{A}_0$ is given by a feedback function $\phi_0$ and $\alpha \in \mathcal{A}$ is given by a feedback function $\phi$, we denote by $q_0^{\phi_0}$, $q^{\phi_0,\phi}$, $f_0^{\phi_0}$ and $f^{\phi_0,\phi}$ the functions:

\[
q_0^{\phi_0}(t,i_0,j_0,x) := q_0(t,i_0,j_0,\phi_0(t,i_0,x),x),
\]
\[
q^{\phi_0,\phi}(t,i,j,i_0,x) := q(t,i,j,\phi(t,i,i_0,x),i_0,\phi_0(t,i_0,x),x),
\]
\[
f_0^{\phi_0}(t,i_0,x) := f_0(t,i_0,\phi_0(t,i_0,x),x),
\]
\[
f^{\phi_0,\phi}(t,i_0,i,x) := f(t,i,\phi(t,i,i_0,x),i_0,\phi_0(t,i_0,x),x).
\]

### 5.2 Mean Field Game Formulation

Solving for Nash equilibria when the number of players is finite is challenging. There are many reasons why the problem becomes quickly intractable. Among them is the fact that as the number of minor players increases, the dimension of the $Q$ - matrix of the system increases exponentially. The paradigm of mean field games consists in the analysis of the limiting case where the number $N$ of minor players tends to infinity. In this asymptotic regime, one expects that simplifications due to averaging effects will make it easier to find asymptotic solutions which could provide approximate equilibria for finite player games when the number $N$ of minor players is large enough. The rationale for such a belief is based on the intuition provided by classical results on the propagation of chaos for large particle systems with mean field interactions. We will develop these results later in the paper.

The advantage of considering the limit case is two-fold. First, when $N$ goes to infinity, the empirical distribution of the minor players’ states converges to a random measure $\mu_t$ which we expect to be the conditional distribution of any minor player’s
state, i.e.: 

\[ \mu_t^N \rightarrow \mu_t := \left( \mathbb{P}[X^n_t = 1|X^0_t], \mathbb{P}[X^n_t = 2|X^0_t], \ldots, \mathbb{P}[X^n_t = m-1|X^0_t] \right). \]

As we shall see later on in the next section, when considered together with the major player’s state and one of the minor player’s state, the resulting process is Markovian and its infinitesimal generator has a tractable form. Also, when the number of minor players goes to infinity, small perturbations of a single minor player’s strategy will have negligible influence on the distribution of the minor player’s states. This gives rise to a simple formulation of the typical minor player’s search for the best response to the control choices of the major player. In the limit \( N \rightarrow \infty \), we understand a Nash equilibrium as a situation in which neither the major player, nor a typical minor player could be better off by changing their control strategies. In order to formulate this limiting problem, we need to define the joint dynamics of the states of the major player and a deviating minor player, making sure that the dynamics of the state of the major player depend upon the statistical distribution of the states of the minor players, and that the dynamics of the state of the deviating minor player depend upon the values of the state and the control of the major player, their own state, and the statistical distribution of the states of all the minor players.

As argued in Carmona and Wang [2017], and echoed in Chapter 13 of Carmona and Delarue [2017], the best way to search for Nash equilibria in the mean field limit of games with major and minor players is first to identify the best response map of the major player and the deviating minor player by solving the optimization problems for the strategies of (1) the major player in response to the field of the minor players, and (2) the deviating minor player in response to the behavior of the major player and the other minor players. Solving these optimization problems separately provides a definition of the best response map for the system. One can then search for a fixed...
point for this best response map. So the search for Nash equilibria for the mean field game with major and minor players can be summarized in the following two steps.

**Step 1: Identifying the Best Response Map**

**Step 1.1: Major Player’s Problem**

Fix an admissible strategy $\alpha \in \mathbb{A}$ of the form $\alpha_t = \phi(t, X_t, X_0^0, \mu_t)$ for the minor players, and solve for the optimal control problem of the major player given that all the minor players use the feedback function $\phi$. We denote by $\phi^*_0(\phi)$ the feedback function giving the optimal strategy of this optimization problem.

Notice that, in order to formulate properly this optimization problem, we need to define Markovian dynamics for the couple $(X_0^0, X_t)$ where $X_t$ is interpreted as the state of a deviating minor player, and the (random) measure $\mu_t$ has to be defined clearly. This is done in the next section as the solution of the major player’s optimization problem, the Markovian dynamics being obtained from the limit of games with $N$ minor players.

**Step 1.2: Deviating Minor Player’s Problem**

We single out the deviating minor player and we search for its best response to the rest of the other players. So we fix an admissible strategy $\alpha_0 \in \mathbb{A}_0$ of the form $\alpha^0_t = \phi_0(t, X^0_t, \mu_t)$ for the major player, and an admissible strategy $\alpha \in \hat{\mathbb{A}}$ of the form $\alpha_t = \phi(t, X_t, X_0^0, \mu_t)$ for the remaining minor players. We then assume that the deviating minor player which we singled out responds to the other players by choosing an admissible strategy $\bar{\alpha} \in \hat{\mathbb{A}}$ of the form $\bar{\alpha}_t = \bar{\phi}(t, \bar{X}_t, X^0_t, \mu_t)$. Clearly, if we want to find the best response of the deviating minor player to the behavior of the major player and the field of the other minor players, we need to determine the dynamics for the triple $(X^0_t, \bar{X}_t, \mu_t)$, and define clearly what we mean by the (random) measure $\mu_t$. This is done in the next section as the solution of the deviating minor player optimization problem, the Markovian dynamics being obtained from the limit
of games with \( N \) minor players. We denote by \( \varphi^*(\phi_0, \phi) \) the feedback function giving the optimal strategy of this optimization problem.

**Step 2: Search for a Fixed Point of the Best Response Map**

A Nash equilibrium for the mean field game with major and minor players is a pair of Markov control strategies \((\hat{\phi}_0, \hat{\phi})\) which provides a fixed point of the best response mappings, i.e. \([\hat{\phi}_0, \hat{\phi}] = [\varphi^*_0(\hat{\phi}), \varphi^*(\hat{\phi}_0, \hat{\phi})]\).

Clearly, in order to carry out Step 1, we need to formulate properly the search for these two best responses, and study the limit \( N \to \infty \) of both optimization problems of interest.

### 5.3 Convergence of Large Finite Player Games

Throughout the rest of this chapter, we make the following assumptions on the regularity of the transition rate and cost functions:

**Assumption 5.3.1.** There exists a constant \( L > 0 \) such that for all \( i, j \in \{1, \ldots, m\} \), \( i_0, j_0 \in \{1, \ldots, m_0\} \) and all \( t, t' \in [0, T] \), \( x, x' \in \bar{S} \), \( \alpha_0, \alpha'_0 \in A_0 \) and \( \alpha, \alpha' \in A \), we have:

\[
| \left( f, f_0, g, g_0, q, q_0 \right)(i, j, i_0, j_0, t, x, \alpha_0, \alpha) - \left( f, f_0, g, g_0, q, q_0 \right)(i, j, i_0, j_0, t', x', \alpha_0', \alpha') | \\
\leq L\left( |t - t'| + \|x - x'| + \|\alpha_0 - \alpha'_0\| + \|\alpha - \alpha'\| \right).
\]

(5.4)

**Assumption 5.3.2.** There exists a constant \( C > 0 \) such that for all \( i, j \in \{1, \ldots, m\} \), \( i_0 \in \{1, \ldots, m_0\} \) and all \( t \in [0, T] \), \( x \in \bar{S} \), \( \alpha_0 \in A_0 \) and \( \alpha \in A \), we have:

\[
| q(t, i, j, \alpha, i_0, \alpha_0, x) | \leq C.
\]

(5.5)

Finally, we add a boundary condition on the Markovian dynamics of the minor players. Intuitively speaking, this assumption rules out the extinction: it says that
a minor player can no longer change its state, when the percentage of minor players who are in the same state falls below a certain threshold.

**Assumption 5.3.3.** There exists a constant \( \epsilon > 0 \) such that for all \( t \in [0,T] \), \( i, j \in \{1, \ldots, m-1\} \), \( i \neq j \) and \( \alpha_0 \in A_0 \) and \( \alpha \in A \), we have:

\[
x_i < \epsilon \iff q(t, i, j, \alpha, i_0, \alpha_0, x) = 0,
\]
\[1 - \sum_{k=1}^{m-1} x_k < \epsilon \iff q(t, m, i, \alpha, i_0, \alpha_0, x) = 0.
\]

The purpose of this section is to identify the state dynamics which should be posited in the formulation of the mean field game problem with major and minor players. In order to do so, we formulate the search for the best response of each player by first setting the game with finitely many minor players, and then letting the number of minor players go to infinity to identify the dynamics over which the best response should be computed in the limit.

### 5.3.1 Major Player’s Problem with Finitely Many Minor Players

For any integer \( N \) (fixed for the moment), we consider a game with \( N \) minor players, and we compute the best response of the major player when the minor players choose control strategies \( \alpha_n = (\alpha_t^n)_{0 \leq t \leq T} \) given by the same feedback function \( \phi \) so that \( \alpha_t^n = \phi(t, X_t^{n,N}, X_t^{0,N}, \mu_t^N) \) for \( n = 1, \cdots, N \). Here, \( X_t^{n,N} \) denotes the state of the \( n \)-th minor player at time \( t \), \( X_t^{0,N} \) the state of the major player, and \( \mu_t^N \) the empirical distribution of the states of the \( N \) minor players at time \( t \). The latter is a probability measure on the state space \( E = \{1, \cdots, m\} \), and we shall identify it with an element in \( \bar{S} \):

\[
\mu_t^N = \frac{1}{N} \left( \sum_{n=1}^{N} \mathbb{1}(X_t^{n,N} = 1), \sum_{n=1}^{N} \mathbb{1}(X_t^{n,N} = 2), \ldots, \sum_{n=1}^{N} \mathbb{1}(X_t^{n,N} = m-1) \right).
\]
So for each $i \in E$, $\mu_i^N(t)$ is the proportion of minor players whose state at time $t$ is equal to $i$. Consequently, for $N$ fixed, $\mu_i^N$ can be viewed as an element of the finite space $\{0, 1/N, \ldots, (N - 1)/N, 1\}^{m-1}$. We denote by $\mathcal{S}_N$ the set of possible values of $\mu_i^N$, in other words, we set:

$$\mathcal{S}_N := \left\{ \frac{1}{N}(n_1, n_2, \ldots, n_{m-1}); n_i \in \mathbb{N}, \sum n_i \leq N \right\}.$$  \hspace{1cm} (5.9)

Given the choice of control strategies made by the minor players, we denote by $\alpha_0 = (\alpha_t^0)_{0 \leq t \leq T}$ the control strategy of the major player, and we study the time evolution of the state of the system given these choices of control strategies. Later on, we shall find the optimal choice for the major player’s control $\alpha_0$ given by a feedback function $\phi_0$ in response to the choice of the feedback function $\phi$ of the minor players. While this optimization should be done over the dynamics of the whole state $(X^{0,N}_t, X^{1,N}_t, \ldots, X^{N,N}_t)$, we notice that the process $(X^{0,N}_t, \mu_t^N)_{0 \leq t \leq T}$ is sufficient to define the optimization problem of the major player, and that it is also a continuous-time Markov process in the finite state space $\{1, \ldots, m_0\} \times \mathcal{S}_N$.

Our goal is to show that as $N \to \infty$, the Markov process $(X^{0,N}_t, \mu_t^N)_{0 \leq t \leq T}$ converges in some sense to a Markov process $(X^0_t, \mu_t)_{0 \leq t \leq T}$. This will allow us to formulate the optimization problem of the major player in the mean field limit in terms of this limiting Markov process.

For each integer $N$, we denote by $\mathcal{G}^{\phi_0, \phi}_{0,N}$ the infinitesimal generator of the Markov process $(X^{0,N}_t, \mu_t^N)_{0 \leq t \leq T}$. Since the process is not time-homogeneous, when we say infinitesimal generator, we mean the infinitesimal generator of the space-time process $(t, X^{0,N}_t, \mu_t^N)_{0 \leq t \leq T}$. Except for the partial derivative with respect to time, this infinitesimal generator is given by the Q-matrix of the process, namely the instantaneous rates of jump in the state space $\{1, \ldots, m_0\} \times \mathcal{S}_N$. So if $F : [0, T] \times \{1, 2, \ldots, m_0\} \times \mathcal{S}_N$
\( S_N \rightarrow \mathbb{R} \) is \( C^1 \) in time,

\[
\begin{align*}
[G^\phi_{0,N} F](t, i_0, x) &= \partial_t F(t, i_0, x) + \sum_{j_0 \neq i_0} [F(t, j_0, x) - F(t, i_0, x)] q^\phi_0(t, i_0, j_0, x) \\
&\quad + \sum_{j \neq i} [F(t, i_0, x + \frac{1}{N} e_{ij}) - F(t, i_0, x)] N x_i q^\phi_0,\phi(t, i, j, i_0, x),
\end{align*}
\]

(5.10)

where the first summation in the right hand side corresponds to jumps in the state of the major player and the terms in the second summation account for the jumps of the state of one minor player from \( i \) to \( j \). Here we code the change in the empirical distribution \( x \) of the states of the minor players caused by the jump from \( i \in \{1, \ldots, m\} \) to \( j \in \{1, \ldots, m\} \) with \( j \neq i \), of the state of a single minor player as \((1/N)e_{ij}\) with the notation \( e_{ij} := e_j \mathbb{1}(j \neq m) - e_i \mathbb{1}(i \neq m) \) where \( e_i \) stands for the \( i \)-th vector in the canonical basis of the space \( \mathbb{R}^{m-1} \). We have also used the notation \( x_m = 1 - \sum_{i=1}^{m-1} x_i \) for the sake of simplicity.

Notice that the two summations appearing in (5.10) correspond to finite difference operators which are bounded. So the domain of the operator \( G^\phi_{0,N} \) is nothing else than the domain of the partial derivative with respect to time. Notice also that the sequence of generators \( G^\phi_{0,N} \) converges, at least formally, toward a limit which can easily be identified. Indeed, it is clear from the definition (5.10) that \( [G^\phi_{0,N} F](t, i_0, x) \) still makes sense if \( x \in \bar{S} \), where \( \bar{S} \) is the \( m - 1 \) dimensional simplex. Moreover, if \( F : [0, T] \times \{1, 2, \ldots, m_0\} \times \bar{S} \rightarrow \mathbb{R} \) is \( C^1 \) in both variables \( t \) and \( x \), we have

\[
[G^\phi_{0,N}^\phi F](t, i_0, x) \rightarrow [G^\phi_0 F](t, i_0, x),
\]

where the operator \( G^\phi_0 \) is defined by:

\[
\begin{align*}
[G^\phi_0 F](t, i_0, x) := \partial_t F(t, i_0, x) + \sum_{j_0 \neq i_0} [F(t, j_0, x) - F(t, i_0, x)] q^\phi_0(t, i_0, j_0, x) \\
&\quad + \sum_{i,j=1}^{m-1} \partial_{x_j} F(t, i_0, x) x_i q^\phi_0(t, i, j, i_0, x) \\
&\quad + (1 - \sum_{k=1}^{m-1} x_k) \sum_{j=1}^{m-1} \partial_{x_j} F(t, i_0, x) q^\phi_0(t, m, j, i_0, x).
\end{align*}
\]
So far, we have a sequence of time-inhomogeneous Markov processes \((X_{t}^{0,N}, \mu_{t}^{N})_{N \geq 0}\) characterized by their infinitesimal generators \(G_{0,N}^{\phi_{0},\phi}\) which converge to \(G_{0}^{\phi_{0},\phi}\). We now aim to show the existence of a limiting Markov process with infinitesimal generator \(G_{0}^{\phi_{0},\phi}\). The proof consists of first showing the existence of a Feller semigroup generated by the limiting generator \(G_{0}^{\phi_{0},\phi}\), and then applying an argument of convergence of semigroups.

**Remark 5.3.4.** The standard results in the theory of semigroups are tailor-made for time-homogeneous Markov processes. However, they can easily be adapted to the case of time-inhomogeneous Markov processes by simply considering the space-time expansion, specifically by augmenting the process \((X_{t}^{0,N}, \mu_{t}^{N})\) into \((t, X_{t}^{0,N}, \mu_{t}^{N})\) and considering the uniform convergence on all bounded time intervals.

Let us introduce some additional notations which are useful for the functional analysis of the infinitesimal generators and their corresponding semigroups. We set \(E_{N} = [0,T] \times \{1, \ldots, m_{0}\} \times \tilde{S}_{N}\) and \(E_{\infty} = [0,T] \times \{1, \ldots, m_{0}\} \times \tilde{S}\) for the state spaces, and we denote by \(C(E_{\infty})\) the Banach space for the norm \(\|F\|_{\infty} = \sup_{t,i_{0},x} |F(t,i_{0},x)|\), of the real valued continuous functions defined on \(E_{\infty}\). We also denote by \(C^{1}(E_{\infty})\) the collection of functions in \(C(E_{\infty})\) that are \(C^{1}\) in \(t\) and \(x\) for all \(i_{0} \in \{1, \ldots, m_{0}\}\).

Note that the Markov process \((t, X_{t}^{0,N}, \mu_{t}^{N})\) lives in \(E_{N}\) while the candidate limiting process \((t, X_{t}^{0,N}, \mu_{t}^{N})\) lives in \(E_{\infty}\). The difference is that \(\mu_{t}^{N}\) only takes values in \(\tilde{S}_{N}\), which is a finite subset of \(\tilde{S}\). Thus if we want to show the convergence, we need to reset all the processes on the same state space, and our first step should be to extend the definition of \((t, X_{t}^{0,N}, \mu_{t}^{N})\) to a Markov process taking value in \(E_{\infty}\). To do so, we extend the definition of the generator \(G_{0,N}^{\phi_{0},\phi}\) to accommodate functions \(F\) defined on the whole \(E_{\infty}\):

\[
[G_{0,N}^{\phi_{0},\phi}F](t, i_{0}, x) = \partial_{t}F(t, i_{0}, x) + \sum_{j_{0} \neq i_{0}} (F(t, j_{0}, x) - F(t, i_{0}, x)) q_{0,\phi_{0}}^{\phi_{0}}(t, i_{0}, j_{0}, x) \\
+ \sum_{j \neq i} (F(t, i_{0}, x + \frac{1}{N} e_{ij}) - F(t, i_{0}, x)) N x_{i} \mathbb{1}(x_{i} \geq \frac{1}{N}) q_{\phi_{0},\phi}^{\phi_{0}}(t, i, j, i_{0}, x).
\]
We claim that for $N$ large enough, $\mathcal{G}_{0,N}^{\phi,\phi}$ generates a Markov process with a Feller semigroup taking values in $E_{\infty}$. Indeed, when the initial distribution is a probability measure on $\{1, \ldots, m_0\} \times \hat{S}_N$, the process has exactly the same law as $(X^{0,N}_t, \mu^N_t)$. To see why this is true, let us denote for all $x \in \hat{S}$ the set $\hat{S}_N^x := (x + \frac{1}{N} \mathbb{Z}^{m-1}) \cap \hat{S}$. Then we can construct a Markov process starting from $(i, x)$ and living in the space of finite states $\{1, \ldots, m\} \times \hat{S}_N^x$, which has the same transition rates as those appearing in the definition of $\mathcal{G}_{0,N}^{\phi,\phi}$. In particular, the indicator function $1(x_i \geq \frac{1}{N})$ forbids the component $x$ to exit the domain $\hat{S}$. Assumption 5.3.3 implies that the transition function is continuous on $E$ when $N \geq 1/\epsilon$, where $\epsilon$ is the extinction threshold in the assumption. So this process is a continuous-time Markov process with continuous probability kernel in a compact space. By Proposition 4.4 in Swart and Winter [2013], it is a Feller process. In the following, we will still denote this extended version of the process as $(X^{0,N}_t, \mu^N_t)$.

**Proposition 5.3.5.** There exists a Feller semigroup $\mathcal{T} = (\mathcal{T}_t)_{t \geq 0}$ on the space $C(E_{\infty})$ such that the closure of $\mathcal{G}_{0,\phi}^{\phi}$ is the infinitesimal generator of $\mathcal{T}$.

**Proof.** We use the perturbation argument. Observe that $\mathcal{G}_{0,\phi}^{\phi}$ is the sum of two linear operators $\mathcal{H}$ and $\mathcal{K}$ on $C(E_{\infty})$:

$$\left[\mathcal{H}F\right](t, i_0, x) := \partial_t F(t, i_0, x) + \mathbf{v}(t, i_0, x) \cdot \nabla F(t, i_0, x),$$

$$\left[\mathcal{K}F\right](t, i_0, x) := \sum_{j_0 \neq i_0} \left[ F(t, j_0, x) - F(t, i_0, x) \right] q^{\phi,\phi}_0(t, i_0, j_0, x),$$

where we denote by $\nabla F(t, i_0, x)$ the gradient of $F$ with respect to $x$, and by $\mathbf{v}$ the vector field:

$$\mathbf{v}_j(t, i_0, x) := \sum_{i=1}^{m-1} x_i q^{\phi,\phi}(t, i, j, i_0, x) + \left(1 - \sum_{i=1}^{m-1} x_i \right) q^{\phi,\phi}(t, m, j, i_0, x).$$

Being a finite difference operator, $\mathcal{K}$ is a bounded operator on $C(E_{\infty})$ so the proof reduces to showing that $\mathcal{H}$ generates a Feller semigroup (see for example Theorem
7.1, Chapter 1 in Ethier and Kurtz [2009]). To show that the closure of $\mathcal{H}$ generates a strongly continuous semigroup on $C(E_\infty)$ we use the characteristics of the vector field $v$. For any $(t,i_0,x) \in E_\infty$, let $(Y^r_{u(t,i_0,x)})_{u \geq 0}$ be the solution of the following ODE:

$$dY^r_{u(t,i_0,x)} = v(t + u, i_0, Y^r_{u(t,i_0,x)})du, \quad Y^r_{0(t,i_0,x)} = x.$$  

The existence and uniqueness of the solution to this ODE are guaranteed by the Lipschitz continuity of the vector field $v$, which in turn is a consequence of the Lipschitz property of $q^{\phi_0,\phi}$. Notice that by Assumption 5.3.3, the process $Y^r_{u(t,i_0,x)}$ is confined to $\bar{S}$. So we can define the linear operator $T_s$ on $C(E_\infty)$:

$$[T_sF](t,i_0,x) := F(s + t, i_0, Y^s_{s(t,i_0,x)}).$$

Uniqueness of the solution implies that $(T_s)_{s \geq 0}$ is a semigroup. This semigroup is also strongly continuous. Indeed, by the boundedness of $v$, for a fixed $h > 0$, there exists a constant $C_0$ such that $|Y^r_{s(t,i_0,x)} - x| \leq C_0 s$ for all $s \leq h$ and $(t,i_0,x) \in E_\infty$. Combining this estimation with the fact that $F$ is uniformly continuous in $(t,x)$ for all $F \in C(E_\infty)$, we obtain that $\|T_sF - F\| \to 0, s \to 0$. Finally the semigroup $T$ is Feller, since the solution of ODE, $Y^r_{s(t,i_0,x)}$, depends continuously on the initial condition as a consequence of the Lipschitz property of the vector field $v$.

It is plain to check that $\mathcal{H}$ is the infinitesimal generator of the semigroup $T$, and the domain of $\mathcal{H}$ is $C^1(E_\infty)$. 

The following lemma is a simple adaptation of Theorem 6.1, Chapter 1 in Ethier and Kurtz [2009] and is an important ingredient in the proof of the convergence. It says that the convergence of the infinitesimal generators implies the convergence of the corresponding semigroups.

**Lemma 5.3.6.** For $N = 1,2,\ldots$, let $\{T_N(t)\}$ and $\{T(t)\}$ be strongly continuous contraction semigroups on a Banach space $L$ with infinitesimal generators $\mathcal{G}_N$ and $\mathcal{G}$,
respectively. Let $D$ be a core for $\mathcal{G}$ and assume that $D \subset \mathcal{D}(\mathcal{G}_N)$ for all $N \geq 1$. If $\lim_{N \to \infty} \mathcal{G}_N F = \mathcal{G} F$ for all $F \in D$, then for each $F \in L$, $\lim_{N \to \infty} T_N(t) F = T(t) F$ for all $t \geq 0$.

We are now ready to state and prove the main result of this section: the Markov process $(X_{0,N}^t, \mu_t^N)$ describing the dynamics of the state of the major player and the empirical distribution of the states of the $N$ minor players converges weakly to a Markov process with infinitesimal generator $\mathcal{G}^{\phi_0, \phi}$, when the players choose Lipschitz strategies.

**Theorem 5.3.7.** Assume that the major player chooses a control strategy $\alpha_0$ given by a Lipschitz feedback function $\phi_0$ and that all the minor players choose control strategies given by the same Lipschitz feedback function $\phi$. Let $i_0 \in \{1, \ldots, m_0\}$ and for each integer $N \geq 1$, let $x_N \in \bar{S}_N$ with limit $x \in \bar{S}$. Then the sequence of processes $(X_{t}^0, N, \mu_t^N)$ with initial conditions $X_{0,N}^0 = i_0$, $\mu_t^N = x_N$ converges weakly to a Markov process $(X_0^t, \mu_t)$ with initial condition $X_0^0 = i_0$, $\mu_0 = x$. The infinitesimal generator for $(X_t^0, \mu_t)$ is given by:

$$
\mathcal{G}^{\phi_0, \phi}(t, i_0, x) := \partial_t F(t, i_0, x) + \sum_{j_0 \neq i_0} [F(t, j_0, x) - F(t, i_0, x)] q^{\phi_0, \phi} \big(\mu_0, \mu_t^N, x\big) + \sum_{i,j=1}^{m-1} \partial x_j F(t, i_0, x) x_i q^{\phi_0, \phi} \big(\mu_0, \mu_t^N, x\big) + \sum_{k=1}^{m-1} \partial x_k F(t, i_0, x) (1 - \sum_{j=1}^{m-1} x_j) q^{\phi_0, \phi} \big(\mu_0, \mu_t^N, x\big).
$$

**Proof.** Let us denote by $(T_t^N)_{t \geq 0}$ the semigroup associated with the time-homogeneous Markov process $(t, X_{t}^0, N, \mu_t^N)$ and the infinitesimal generator $\mathcal{G}^{\phi_0, \phi}_{0,N}$. Recall that by the procedure of extension we described above, the process $(t, X_{t}^0, N, \mu_t^N)$ now lives in $E_\infty$ and the domain for $\mathcal{G}^{\phi_0, \phi}_{0,N}$ is $C(E_\infty)$. In light of Theorem 2.5, Chapter 4 in Ethier and Kurtz [2009] and Proposition 5.3.5 we just proved, it boils down to proving that for any $F \in E_\infty$ and $t \geq 0$, $T_t^N F$ converges to $T_t F$, where $(T_t)_{t \geq 0}$ is the strongly continuous semigroup generated by the closure of $\mathcal{G}^{\phi_0, \phi}_{0,N}$.
To show the convergence, we apply Lemma 5.3.6. It is easy to see that \( C^1(E\infty) \) is a core for \( \mathcal{G}_0^{\phi_0,\phi} \) and \( C^1(E\infty) \) is included in the domain of \( \mathcal{G}_0^{\phi_0,\phi}_N \). Therefore it only remains to show that for all \( F \in C^1(E\infty) \), \( \mathcal{G}_0^{\phi_0,\phi}_N F \) converges to \( \mathcal{G}_0^{\phi_0,\phi} F \) in the space \( (C(E\infty), \| \cdot \|) \). Using the notation \( x_m := 1 - \sum_{i=1}^{m-1} x_i \), we have:

\[
\left| \left[ \mathcal{G}_0^{\phi_0,\phi}_N F(t, i_0, x) - \mathcal{G}_0^{\phi_0,\phi} F(t, i_0, x) \right] \right| \\
= \sum_{j \neq i} \left| N(F(t, i_0, x + \frac{1}{N}e_{ij}) - F(t, i_0, x)) - (1(j \neq m)\partial_{x_j} F(t, i_0, x) - 1(i \neq m)\partial_{x_i} F(t, i_0, x)) \right| x_i q_N(t, i, j, i_0, x)
\]

where we applied the intermediate value theorem at the last inequality and \( \lambda_{i,j} \in [0, 1] \). Note that \( \lambda_{i,j} \) also depends on \( t, x, i_0 \), but we omit them for the sake of simplicity. We notice that \( F \in C^1(E\infty) \) and \( E\infty \) is compact, and therefore \( \partial_{x_i} F \) is uniformly continuous on \( E\infty \) for all \( i \), which immediately implies that \( \| \mathcal{G}_0^{\phi_0,\phi}_N F - \mathcal{G}_0^{\phi_0,\phi} F \| \rightarrow 0, N \rightarrow +\infty \). This completes the proof.

\[ \square \]

**Major Player’s Optimization Problem in the Mean Field Limit**

Given that all the minor players use control strategies based on the same feedback function \( \phi \), the best response of the major player is to use the strategy \( \hat{\alpha}_0 \) given by the feedback function \( \hat{\phi}_0 \) solving the optimal control problem:

\[
\inf_{\alpha_0 \leftrightarrow \phi_0 \in \mathcal{A}_0} \mathbb{E} \left[ \int_0^T f_0(t, X^0_t, \phi_0(t, X^0_t, \mu_0, \mu_t), \mu_t) dt + g_0(X^0_T, \mu_T) \right],
\]

where \( (X^0_t, \mu_t)_{0 \leq t \leq T} \) is the continuous-time Markov process with infinitesimal generator \( \mathcal{G}_0^{\phi_0,\phi}_0 \).

### 5.3.2 Deviating Minor Player’s Problem

We turn to the computation of the best response of the deviating minor player, which we assume to be player 1. We assume that the major player chooses a strategy \( \alpha_0 \in \mathcal{A}_0 \) of the form \( \alpha_0^t = \phi_0(t, X^0_{i,N} t, \mu_t^N) \) and that the minor players \( n \in \{2, \ldots, N\} \)

161
all use strategies $\alpha_n \in A$ of the form $\alpha^*_n = \phi(t, X^{n,N}_t, X^{0,N}_t, \mu^N_t)$, and that the first minor player uses strategy $\bar{\alpha} \in A$ of the form $\bar{\alpha}_t = \bar{\phi}(t, X^{1,N}_t, X^{0,N}_t, \mu^N_t)$. Clearly, by symmetry, whatever we are about to say after we singled the first minor player out as the deviating minor player, can be said if we single out any other minor player. As before, for each fixed $N$, the process $(X^{0,N}_t, X^{1,N}_t, \mu^N_t)$ is a finite-state continuous-time Markov process with state space $\{1, \ldots, m_0\} \times \{1, \ldots, m\} \times \mathcal{S}_N$ whose infinitesimal generator $G_N^{\phi_0, \phi, \bar{\phi}}$ is given, up to the time derivative, by the corresponding Q-matrix of infinitesimal jump rates. In the present situation, its value on any real valued function $F$ defined on $[0,T] \times \{1, \ldots, m_0\} \times \{1, \ldots, m\} \times \mathcal{S}_N$ such that $t \to F(t, i_0, i, x)$ is $C^1$ for any $i_0$, $i$ and $x$, is given by the formula:

$$
[G_N^{\phi_0, \phi, \bar{\phi}} F](t, i_0, i, x) := \partial_t F(t, i_0, i, x) + \sum_{j_0, j_0 \neq i_0} [F(t, j_0, i, x) - F(t, i_0, i, x)]q^{\phi_0}(t, i_0, j_0, x)
$$

$$
+ \sum_{j, j \neq i} [F(t, i_0, j, x + \frac{1}{N} e_{ij}) - F(t, i_0, i, x)]q^{\phi_0, \bar{\phi}}(t, i, j, i_0, x)
$$

$$
+ \sum_{(j, k), j \neq k} [F(t, i_0, i, x + \frac{1}{N} e_{kj}) - F(t, i_0, i, x)](N x_k - 1(k = i))q^{\phi_0, \phi}(t, k, j, i_0, x).
$$

(5.13)

As before, the summations appearing above correspond to single jumps when 1) only the state of the major player changes from state $i_0$ to $j_0$, 2) only the state of the singled out first minor player changes from state $i$ to $j$, and finally 3) the state of one of the last $N - 1$ minor players jumps from state $k$ to $j$.

Following the same treatment as in the major player’s problem, we have the convergence result for the process $(X^N_t, X^{0,N}_t, \mu^N_t)$:

**Theorem 5.3.8.** Assume that for each integer $N$, the major player chooses a control $\alpha_0$ given by a Lipschitz feedback function $\phi_0$, the first minor player chooses a control $\bar{\alpha}$ given by a Lipschitz feedback function $\bar{\phi}$, all the other minor players choose strategies given by the same Lipschitz feedback function $\phi$, and that these three feedback functions do not depend upon $N$. Let $i_0 \in \{1, \ldots, m_0\}$ and for each integer $N \geq 2$, let $x_N \in \mathcal{S}_N$ with limit $x \in \mathcal{S}$. Then the sequence of processes $(X^N_t, X^{0,N}_t, \mu^N_t)_{0 \leq t \leq T}$ with initial
conditions \( X_0^N = i, X_0^{0,N} = i_0 \) and \( \mu_0^N = x_N \) converges weakly to a Markov process \((X_t, X_t^0, \mu_t)\) with initial condition \( X_0 = i, X_0^0 = i_0 \) and \( \mu_0 = x \). Its infinitesimal generator is given by:

\[
\begin{align*}
\{G_{\phi_0, \phi, \bar{\phi}} F\}(t, i, i_0, x) := & \quad \partial_t F(t, i, i_0, x) + \sum_{j_0 \neq i_0} [F(t, j_0, i_0, x) - F(t, i_0, i_0, x)] q^\phi_0 (t, i_0, j_0, x) \\
& + \sum_{j \neq i} [F(t, j, i_0, x) - F(t, i, i_0, x)] q^{\phi_0, \bar{\phi}} (t, i, j, i_0, x) + \sum_{i, j=1}^{m-1} \partial_{x_j} F(t, i, i_0, x) q^{\phi_0, \bar{\phi}} (t, i, j, i_0, x) \\
& + (1 - \sum_{k=1}^m x_k) \sum_{j=1}^{m-1} \partial_{x_k} F(t, i, i_0, x) q^{\phi_0, \bar{\phi}} (t, m, j, i_0, x).
\end{align*}
\]

(5.14)

Deviating Minor Player’s Optimization Problem in the Mean Field Limit

Accordingly, in the mean field game limit, we define the search for the best response of the deviating minor player (i.e. the minor player we singled out) to the strategies adopted by the major player and the field of minor players as the following optimal control problem. Assuming that the major player uses a feedback function \( \phi_0 \) and all the other minor players use the feedback function \( \phi \), the best response of the deviating minor player is given by the solution of:

\[
\inf_{\bar{\phi} \in \mathcal{A}} \mathbb{E} \left[ \int_0^T f(t, X_t, \bar{\phi}(t, X_t, X_t^0, \mu_t), X_t^0, \phi_0(t, X_t^0, \mu_t), \mu_t) dt + g(X_T, X_T^0, \mu_T) \right],
\]

where \((X_t, X_t^0, \mu_t)_{0 \leq t \leq T}\) is a Markov process with infinitesimal generator \( \mathcal{G}_{\phi_0, \phi, \bar{\phi}} \). We shall denote by \( \bar{\phi} = \phi(\phi_0, \phi) \) the optimal feedback function providing the solution of this optimal control problem.

5.4 Optimization Problem for Individual Players

In this section, we use the dynamic programming principle to characterize the value functions of the major and minor players’ optimization problems as viscosity solutions of the corresponding Hamilton-Jacobi-Bellman (HJB for short) equations. We follow
the detailed arguments given in Chapter II of Fleming and Soner [2006]. For both the major and deviating minor player, we show that the value function solves a weakly coupled system of Partial Differential Equations (PDEs for short) in viscosity sense. We also prove an important uniqueness result for these solutions. This uniqueness result is important indeed because as the reader noticed, in defining the best response map, we implicitly assumed that these optimization problems could be solved and that their solutions were unique.

We first consider the value function of the major player’s optimization problem assuming that all the minor players use the feedback function $\phi$:

$$V_0^\phi(t, i_0, x) := \inf_{\alpha_0 \leftrightarrow \phi_0 \in A_0} \mathbb{E} \left[ \int_t^T f_0^\phi(s, X_s^0, \mu_s) ds + g_0(X_T^0, \mu_T) \mid X_t^0 = i_0, \mu_t = x \right]. \quad (5.15)$$

**Theorem 5.4.1.** Assume that for all $i_0 \in \{1, \ldots, m_0\}$, the mapping $(t, x) \rightarrow V_0^\phi(t, i_0, x)$ is continuous on $[0, T] \times \mathcal{S}$. Then $V_0^\phi$ is a viscosity solution to the system of $m_0$ PDEs on $[0, T] \times \mathcal{S}$:

$$0 = \partial_t v_0(t, i_0, x) + \inf_{\alpha_0 \in A_0} \left\{ f_0(t, i_0, \alpha_0, x) + \sum_{j_0 \neq i_0} [v_0(t, j_0, x) - v_0(t, i_0, x)] g_0(t, i_0, j_0, \alpha_0, x) 
+ (1 - \sum_{k=1}^{m-1} x_k) \sum_{k=1}^{m-1} \partial_{x_k} v_0(t, i_0, x) q(t, m, k, \phi(t, m, i_0, x), i_0, \alpha_0, x)
+ \sum_{i,j=1}^{m-1} \partial_{x_j} v_0(t, i_0, x) x_i q(t, i, j, \phi(t, i, i_0, x), i_0, \alpha_0, x) \right\},$$

$$0 = g_0(i_0, x) - v_0(T, i_0, x). \quad (5.16)$$

The notion of viscosity solution in the above result is specified by the following definition:

**Definition 5.4.2.** A real valued function $v_0$ defined on $[0, T] \times \{1, \ldots, m_0\} \times \mathcal{S}$ such that $v_0(\cdot, i_0, \cdot)$ is continuous on $[0, T] \times \{1, \ldots, m_0\} \times \mathcal{S}$ for all $i_0 \in \{1, \ldots, m_0\}$ is said to be a viscosity subsolution (resp. supersolution) if for any $(t, i_0, x) \in [0, T] \times \{1, \ldots, m_0\} \times \mathcal{S}$ and any $C^\infty$ function $\theta$ defined on $[0, T] \times \mathcal{S}$ such that the function
(v_0(\cdot, i_0, \cdot) - \theta) attains a maximum (resp. minimum) at (t, x) and v_0(t, i_0, x) = \theta(t, x), then the following inequalities hold:

\[
0 \leq (\text{resp. } \geq) \partial_t \theta(t, x) + \inf_{\alpha_0 \in A_0} \left\{ f_0(t, i_0, \alpha_0, x) + \sum_{j_0 \neq i_0} [v_0(t, j_0, x) - v_0(t, i_0, x)]q_0(t, i_0, j_0, \alpha_0, x) + (1 - \sum_{k=1}^{m-1} x_k) \sum_{k=1}^{m-1} \partial_{x_k} \theta(t, x)q(t, m, k, \phi(t, m, i_0, x), i_0, \alpha_0, x) \right. \\
\left. + \sum_{i, j=1}^{m-1} \partial_{x_j} \theta(t, x)q(t, i, j, \phi(t, i, i_0, x), i_0, \alpha_0, x) \right\}, \text{ if } t < T,
\]

\[
0 \leq (\text{resp. } \geq) g_0(i_0, x) - v_0(T, i_0, x).
\] (5.17)

If \( v_0 \) is both a viscosity subsolution and supersolution, we call it a viscosity solution.

Proof. Define \( C(\bar{S})^{m_0} \) the collection of mappings \((i_0, x) \to \theta(i_0, x)\) defined on \(\{1, \ldots, m_0\} \times \bar{S}\) such that \(\theta(i_0, \cdot)\) is continuous on \(\bar{S}\) for all \(i_0\). Define the dynamic programming operator \(T_{t,s}\) on \(C(\bar{S})^{m_0}\) by:

\[
[T_{t,s}\theta](i_0, x) := \inf_{\alpha_0 \leftrightarrow \phi_0} \mathbb{E} \left[ \int_t^s f_0^\phi(u, X_u^0, \mu_u)du + \theta(X_u^0, \mu_u)|X_t^0 = i_0, \mu_t = x \right].
\] (5.18)

where the Markov process \((X_u^0, \mu_u)_{0 \leq t \leq T}\) has infinitesimal generator \(G_0^{\phi_0, \phi}\). Then the value function can be expressed as:

\[
V_0^\phi(t, i_0, x) = [T_{t,T}g_0](i_0, x),
\]

and the dynamic programming principle says that:

\[
V_0^\phi(t, i_0, x) = [T_{t,s}V_0^\phi(s, \cdot, \cdot)](i_0, x), \quad (t, s, i_0, x) \in [0, T]^2 \times \{1, \ldots, m_0\} \times \bar{S}.
\]

We will use the following lemma for which the proof is given in the appendix.

Lemma 5.4.3. Let \( \Phi \) be a function on \([0, T] \times \{1, \ldots, m_0\} \times \bar{S} \) and \( i_0 \in \{1, \ldots, m_0\} \) such that \( \Phi(\cdot, i_0, \cdot) \) is \(C^1\) in \([0, T] \times \bar{S}\) and \( \Phi(\cdot, j_0, \cdot) \) is continuous in \([0, T] \times \bar{S}\) for all
Then we have:

\[
\lim_{h \to 0} \frac{1}{h} \left[ (T_{t+h} \Phi(t + h, \cdot, \cdot))(i_0, x) - \Phi(t, i_0, x) \right]
\]

\[
= \partial_t \Phi(t, i_0, x) + \inf_{\alpha_0 \in A_0} \left\{ f_0(t, i_0, \alpha_0, x) + (1 - \sum_{k=1}^{m-1} x_k) \sum_{k=1}^{m-1} \partial_{x_k} \Phi(t, i_0, x) q(t, m, k, \phi(t, m, i_0, x), i_0, \alpha_0, x) \right. \\
+ \left. \sum_{i,j=1}^{m-1} \partial_{x_i} \Phi(t, i, x) x_j q(t, i, j, \phi(t, i, i_0, x), i_0, \alpha_0, x) + \sum_{j_0 \neq i_0} [\Phi(t, j_0, x) - \Phi(t, i_0, x)] q_0(t, i_0, j_0, \alpha_0, x) \right\}.
\]

We now prove the subsolution property. Let \( \theta \) be a function defined on \([0, T] \times \bar{S} \) such that \((V_0(\cdot, i_0, \cdot) - \theta)\) attains a maximum at \((t, x)\) and \(V_0(t, i_0, x) = \theta(t, x)\). Define the function \( \Phi \) on \([0, T] \times \{1, \ldots, m_0\} \times \bar{S} \) by \( \Phi(\cdot, i_0, \cdot) := \theta \) and \( \Phi(\cdot, j_0, \cdot) := V_0(\cdot, j_0, \cdot) \) for \( j_0 \neq i_0 \). Then clearly \( \Phi \geq V_0 \), which implies:

\[
(T_{t,s} \Phi(s, \cdot, \cdot))(i_0, x) \geq (T_{t,s} V_0(s, \cdot, \cdot))(i_0, x).
\]

By the Dynamic Programming Principle and the fact that \( \Phi(t, i_0, x) = V_0(t, i_0, x) \), we have:

\[
\lim_{h \to 0} \frac{1}{h} \left[ (T_{t,s} \Phi(s, \cdot, \cdot))(i_0, x) - \Phi(t, i_0, x) \right] \geq 0.
\]

Then applying the lemma we obtain the desired inequality. The viscosity property for supersolution can be checked in exactly the same way. 

For later reference, we state the comparison principle for the HJB equation we just derived. Again, its proof is postponed to the appendix of this chapter.

**Theorem 5.4.4.** (Comparison Principle) Let us assume that the feedback function \( \phi \) is Lipschitz, and let \( w \) (resp. \( v \)) be a viscosity subsolution (resp. supersolution) of the equation (5.16). Then we have \( w \leq v \).

We now turn to the deviating minor agent’s optimization problem assuming that the major player uses the feedback function \( \phi_0 \) and all the other minor players use
the feedback function $\phi$. We define the value function:

$$V^{\phi_0,\phi}(t, i, i_0, x) := \inf_{\alpha \leftrightarrow \phi \in \mathcal{A}} \mathbb{E} \left[ \int_t^T f^{\phi_0,\phi}(s, X_s, X^0_s, \mu_s) ds + g(X_T, X^0_T, \mu_T) \mid X_t = i, X^0_t = i_0, \mu_t = x \right],$$

(5.19)

where the Markov process $(X_t, X^0_t, \mu_t)_{0 \leq t \leq T}$ has infinitesimal generator $G^{\phi_0,\phi,\phi}$. In line with the analysis of the major player’s problem, we can show that $V^{\phi_0,\phi}$ is the unique viscosity solution to a coupled system of PDEs.

**Theorem 5.4.5.** Assume that for all $i \in \{1, \ldots, m\}$ and $i_0 \in \{1, \ldots, m_0\}$, the mapping $(t, x) \rightarrow V^{\phi_0,\phi}(t, i, i_0, x)$ is continuous on $[0, T] \times \bar{S}$. Then $V^{\phi_0,\phi}$ is a viscosity solution to the system of PDEs:

$$0 = \partial_t v(t, i, i_0, x)$$

$$+ \inf_{\alpha \in \mathcal{A}} \left\{ f(t, i, \bar{\alpha}, i_0, \phi_0(t, i_0, x), x) + \sum_{j \neq i} [v(t, j, i_0, x) - v(t, i, i_0, x)] q(t, i, j, i_0, x) \phi_0(t, i_0, x) + \sum_{j \neq i_0} [v(t, i, j_0, x) - v(t, i, i_0, x)] q_{\phi_0}(t, i_0, j_0, x) + \sum_{i, j = 1}^{m-1} \partial_{x_j} v(t, i, i_0, x) q_{\phi}(t, i, j, i_0, x) \right\}$$

$$+ \sum_{j_0 \neq i_0}^{m-1} \sum_{k=1}^{m-1} \partial_{x_k} v(t, i, i_0, x) q_{\phi_0}(t, m, k, i_0, x),$$

$$0 = g(i, i_0, x) - v(T, i, i_0, x).$$

(5.20)

Moreover, if the feedback functions $\phi_0$ and $\phi$ are Lipschitz, then the above system of PDEs satisfies the comparison principle.

It turns out that the value functions $V^\phi_0$ and $V^{\phi_0,\phi}$ are Lipschitz in $(t, x)$. To establish this regularity property and estimate the Lipschitz constants, we need to first study the regularity of the value functions for the finite player games, and control the convergence in the regime of large games. We will state these results in Section 5.7, where we deal with the propagation of chaos and highlight more connections between finite player games and mean field games.

We conclude this section with a result which we will use frequently in the sequel. To state the result, we denote by $J^\phi_0$ the expected cost of the major player when it
uses the feedback function \( \phi_0 \) and the minor players all use the feedback function \( \phi \). Put differently:

\[
J_{\phi_0,\phi}(t, i_0, x) := \mathbb{E} \left[ \int_t^T f_{\phi_0}(s, X^0_s, \mu_s) ds + g_0(X^0_T, \mu_T) | X^0_t = i_0, \mu_t = x \right],
\]

(5.21)

where the Markov process \((X^0_s, \mu_s)_{t \leq s \leq T}\) has infinitesimal generator \( \mathcal{G}^{\phi_0,\phi} \). Then by definition, we have \( V^\phi(t, i_0, x) = \inf_{\alpha_0} J_{\phi_0,\phi}(t, i_0, x) \). Similarly, we denote by \( J_{\phi_0,\phi,\tilde{\phi}} \) the expected cost of the deviating minor player when it uses the feedback function \( \tilde{\phi} \), while the major player uses the feedback function \( \phi_0 \) and all the other minor players use the same feedback function \( \phi \):

\[
J_{\phi_0,\phi,\tilde{\phi}}(t, i, i_0, x) := \mathbb{E} \left[ \int_t^T f_{\phi_0,\tilde{\phi}}(s, X_s, X^0_s, \mu_s) ds + g(X_T, X^0_T, \mu_T) | X_t = i, X^0_t = i_0, \mu_t = x \right]
\]

(5.22)

where the Markov process \((X_s, X^0_s, \mu_s)_{t \leq s \leq T}\) has infinitesimal generator \( \mathcal{G}^{\phi_0,\phi,\tilde{\phi}} \).

**Proposition 5.4.6.** If the feedback functions \( \phi_0, \phi \) and \( \tilde{\phi} \) are Lipschitz, then \( J_{\phi_0,\phi} \) and \( J_{\phi_0,\phi,\tilde{\phi}} \) are respectively continuous viscosity solutions of the PDEs (5.23) and (5.24):

\[
\begin{cases}
0 = [\mathcal{G}^{\phi_0,\phi} v_0](t, i_0, x) + f_{\phi_0}(t, i_0, x), \\
0 = g_0(i_0, x) - v_0(T, i_0, x),
\end{cases}
\]

(5.23)

\[
\begin{cases}
0 = [\mathcal{G}^{\phi_0,\phi,\tilde{\phi}} v](t, i, i_0, x) + f_{\phi_0,\tilde{\phi}}(t, i, i_0, x), \\
0 = g(i, i_0, x) - v(T, i, i_0, x).
\end{cases}
\]

(5.24)

moreover, the PDEs (5.23) and (5.24) satisfy the comparison principle.

**Proof.** The continuity of \( J_{\phi_0,\phi} \) and \( J_{\phi_0,\phi,\tilde{\phi}} \) follows from the fact that \((X^0_t, \mu_t)\) and \((X_t, X^0_t, \mu_t)\) are Feller processes, which we have shown in Theorem 5.3.7 and Theorem 5.3.8. The viscosity property can be shown using the exact same technique as in the proof of Theorem 5.4.1. Finally, the comparison principle is a consequence of the
Lipschitz property of $f_0, f, q_0, q, \phi_0, \phi$ and $\bar{\phi}$, and it can be proved by slightly modifying the proof of Theorem 5.4.4.

5.5 Existence of Nash Equilibria

In this section, we prove existence of Nash equilibria when the minor player’s jump rates and cost functions do not depend upon the major player’s control. We work under the following assumption:

Assumption 5.5.1. Assumption 5.3.1 and Assumption 5.3.2 are in force. In addition, the transition rate function $q$ and the cost function $f$ for the minor player do not depend upon the major player’s control $\alpha_0 \in A_0$.

The following assumptions will guarantee the existence of optimal strategies for both the major player and the deviating minor player.

Assumption 5.5.2. For all $i_0 = 1, \ldots, m_0$, $(t, x) \in [0, T] \times \mathcal{P}$ and $v_0 \in \mathbb{R}^{m_0}$, the function $\alpha_0 \to f_0(t, i_0, \alpha_0, x) + \sum_{j_0 \neq i_0} (v_0[j_0] - v_0[i_0])q_0(t, i_0, j_0, \alpha_0, x)$ has a unique maximizer in $A_0$ denoted as $\hat{\alpha}_0(t, i_0, x, v_0)$. Additionally, $\hat{\alpha}_0$ is Lipschitz in $(t, x, v_0)$ for all $i_0 = 1, \ldots, m_0$ with common Lipschitz constant $L_{\alpha_0}$.

Assumption 5.5.3. For all $i = 1, \ldots, m$, $i_0 = 1, \ldots, m_0$, $(t, x) \in [0, T] \times \bar{\mathcal{S}}$ and $v \in \mathbb{R}^{m \times m_0}$, the function $\alpha \to f(t, \alpha, i, i_0, x) + \sum_{j \neq i} (v[j, i] - v[i, i_0])q(t, i, j, \alpha, i_0, x)$ has a unique minimizer in $A$ denoted as $\hat{\alpha}(t, i, i_0, x, v)$. Additionally, $\hat{\alpha}$ is Lipschitz in $(t, x, v)$ for all $i_0 = 1, \ldots, m_0$ and $i = 1, \ldots, m$ with common Lipschitz constant $L_{\alpha}$.

Remark 5.5.4. Assumptions 5.5.2 and 5.5.3 are essentially about the existence of a unique minimizer of the Hamiltonian for both the major and minor player’s optimization problem. It is very similar to Assumption 3.1.7 in Chapter 3, where we study finite state mean field games without the major player. In Remark 3.1.8 and
Assumption 3.1.9, we provided a set of sufficient conditions, namely the convexity of the cost functions and the linearity of the transition rate, which guarantees the existence of a unique Lipschitz-continuous minimizer of the Hamiltonian. We can easily adapt these conditions to the current context, so that Assumptions 5.5.2 and 5.5.3 hold.

**Proposition 5.5.5.** Under Assumptions 5.5.1, 5.5.2 & 5.5.3, we have:

(i) For any Lipschitz feedback function $\phi$ for the deviating minor player, the best response $\phi^*(\phi)$ of the major player exists and is given by:

$$
\phi^*(\phi)(t, i, i_0, x) = \hat{\alpha}(t, i, i_0, x, V_{\phi^0}(t, \cdot, x)),
$$

where $\hat{\alpha}$ is the minimizer defined in Assumption 5.5.2 and $V_{\phi^0}$ is the value function of the major player’s optimization problem.

(ii) For any Lipschitz feedback function $\phi_0$ for the major player and $\phi$ for the other minor players, the best response $\phi^*(\phi_0, \phi)$ of the deviating minor player exists and is given by:

$$
\phi^*(\phi_0, \phi)(t, i, i_0, x) = \hat{\alpha}(t, i, i_0, x, V_{\phi_0, \phi}(t, \cdot, \cdot, x)),
$$

where $\hat{\alpha}$ is the minimizer defined in Assumption 5.5.3 and $V_{\phi_0, \phi}$ is the value function of the deviating minor player’s optimization problem.

**Proof.** Consider the expected total cost $J_{\phi_0}^{\phi^0(\phi), \phi}$ of the major player when all of the minor players use the feedback function $\phi$ and the major player uses the strategy given by the feedback function $\phi_0^*(\phi)$ defined by (5.25). Also consider $V_{\phi}^0$, the value function of the major player’s optimization problem. By the definition of $\phi_0^*(\phi)$ and the PDE (5.16), we see that $V_{\phi}^0$ is a viscosity solution of the PDE (5.23) with $\phi_0 = \phi_0^*(\phi)$ and $\phi = \phi$ in Proposition 5.4.6. To be able to use the comparison principle, we need to show that $\phi_0 = \phi_0^*(\phi)$ and $\phi$ are Lipschitz. Indeed the Lipschitz property follows from Assumption 5.5.2 and Corollary 5.7.8 (see Section 5.7). Now since $J_{\phi_0}^{\phi_0^*(\phi), \phi}$ is another
viscosity solution for the same PDE, we conclude that $J^{\phi_0(\phi), \phi}_0 = V_0^{\phi} = \inf_{\alpha_0 \to \phi_0} J^{\phi_0, \phi}_0$ and hence the optimality of $\phi_0^*(\phi)$. Likewise we can show that $\phi^*(\phi_0, \phi)$ is the best response of the deviating minor player.

In order to show that the Nash equilibrium is actually given by a couple of Lipschitz feedback functions, we need an additional assumption on the regularity of the value functions.

**Assumption 5.5.6.** There exist two constants $L_{\phi_0}, L_\phi$, such that for all $L_{\phi_0}$-Lipschitz feedback function $\phi_0$ and $L_\phi$-Lipschitz feedback function $\phi$, $V_0^{\phi}$ is $(L_{\phi_0}/L_\alpha - 1)$-Lipschitz and $V^{\phi_0, \phi}$ is $(L_\phi/L_\alpha - 1)$-Lipschitz.

The above assumption holds, for example, when the horizon of the game is sufficiently small. We shall provide more details (see Remark 5.7.9 below) after we reveal important connections between finite player games and mean field games in Section 5.7. We now state and prove the result on the existence of Nash equilibria.

**Theorem 5.5.7.** Under Assumptions 5.5.1, 5.5.2, 5.5.3 & 5.5.6, there exists a Nash equilibrium in the sense that there exists Lipschitz feedback functions $\hat{\phi}_0$ and $\hat{\phi}$ such that:

$$[\hat{\phi}_0, \hat{\phi}] = [\phi_0^*(\hat{\phi}), \phi^*(\hat{\phi}_0, \hat{\phi})].$$

**Proof.** We apply Schauder’s fixed point theorem. To this end, we need to: (i) specify a Banach space $\mathbb{V}$ containing the admissible feedback functions $(\phi_0, \phi)$ as elements, and a relatively compact convex subset $\mathbb{K}$ of $\mathbb{V}$; (ii) show that the mapping $\mathcal{R} : [\phi_0, \phi] \to [\phi_0^*(\phi), \phi^*(\phi_0, \phi)]$ is continuous and leaves $\mathbb{K}$ invariant (i.e. $\mathcal{R}(\mathbb{K}) \subset \mathbb{K}$).

(i) Define $\mathbb{C}_0$ as the collection of $A_0$ - valued functions $\phi_0$ on $[0, T] \times \{1, \ldots, m_0\} \times \bar{S}$ such that $(t, x) \times \phi_0(t, i_0, x)$ is continuous for all $i_0$, $\mathbb{C}$ as the collection of $A$ - valued functions $\phi$ on $[0, T] \times \{1, \ldots, m\} \times \{1, \ldots, m_0\} \times \bar{S}$ such that $(t, x) \times \phi(t, i, i_0, x)$ is
continuous for all $i, i_0$, and set $\mathbb{V} := C_0 \times C$. For all $(\phi_0, \phi) \in \mathbb{V}$, we define the norm:

$$\| (\phi_0, \phi) \| := \max \left\{ \sup_{i_0, t \in [0,T], x \in S} |\phi_0(t, i_0, x)|, \sup_{i, i_0, t \in [0,T], x \in S} |\phi(t, i, i_0, x)| \right\}.$$

It is easy to check that $(\mathbb{V}, \| \cdot \|)$ is a Banach space. Next, we define $\mathbb{K}$ as the collection of elements in $\mathbb{V}$ such that the mappings $(t, x) \mapsto \phi_0(t, i_0, x)$ are $L^0$–Lipschitz and $(t, x) \mapsto \phi(t, i, i_0, x)$ are $L$–Lipschitz in $(t, x)$ for all $i_0 = 1, \ldots, m_0$ and $i = 1, \ldots, m$, where $L^0, L$ are specified in Assumption 5.5.6. Clearly $\mathbb{K}$ is convex. Now consider the family $(\phi_0(\cdot, i_0, \cdot), (\phi_0, \phi))_{(\phi_0, \phi) \in \mathbb{K}}$ of functions defined on $[0,T] \times \bar{S}$. Thanks to the Lipschitz property, we see immediately that the family is equicontinuous and pointwise bounded. Therefore by the Arzelà-Ascoli theorem, the family is compact with respect to the uniform norm. Repeating this argument for all $i, i_0$ we see that $\mathbb{K}$ is compact under the norm $\| \cdot \|$. moreover, thanks to Assumptions 5.5.2, 5.5.3 & 5.5.6, we obtain easily that $\mathbb{K}$ is stable by $\mathcal{R}$.

(ii) It remains to show that $\mathcal{R}$ is a continuous mapping. We use the following lemma:

**Lemma 5.5.8.** Let $(\phi^n_0, \phi^n)$ be a sequence in $\mathbb{K}$ converging to $(\phi_0, \phi)$ in $\| \cdot \|$, and denote by $V^n_0$ and $V^n$ the value functions of the major and the deviating minor players, respectively, associated with $(\phi^n_0, \phi^n)$. Then $V^n_0$ and $V^n$ converge uniformly to $V_0$ and $V$, respectively, where $V_0$ and $V$ are the value functions of the major player and the deviating minor player associated with $(\phi_0, \phi)$.

The proof of the lemma uses standard arguments from the theory of viscosity solutions. We give it in the appendix. Now, the continuity of the mapping $\mathcal{R}$ follows readily from Lemma 5.5.8, Proposition 5.5.5 and Assumption 5.5.2 & 5.5.3. This completes the proof.
5.6 The Master Equation for Nash Equilibria

If a Nash equilibrium exists and is given by feedback functions $\hat{\phi}_0$ for the major player and $\hat{\phi}$ for the minor players, these functions should also be equal to the respective minimizers of the Hamiltonians in the HJB equations of the optimization problems. This informal remark leads to a system of coupled PDEs with terminal conditions specified at $t = T$, which we expect to hold if the equilibrium exists. Now the natural question to ask is: if this system of PDEs has a solution, does this solution provide a Nash equilibrium? The following result provides a verification argument:

**Theorem 5.6.1.** (Verification Argument) Assume that there exists two functions $\hat{\phi}_0 : [0, T] \times \{1, \ldots, m_0\} \times S \ni (t, i_0, x) \to \hat{\phi}_0(t, i_0, x) \in \mathbb{R}$ and $\hat{\phi} : [0, T] \times \{1, \ldots, m\} \times \{1, \ldots, m_0\} \times S \ni (t, i, i_0, x) \to \hat{\phi}(t, i, i_0, x) \in \mathbb{R}$ such that the system of PDEs in $(v_0, v)$:

\[ 0 = [G_{\hat{\phi}_0, \hat{\phi}} v_0](t, i_0, x) + f_0(t, i_0, \hat{\phi}_0(t, i_0, x), x), \]
\[ v_0(T, i_0, x) = g_0(i_0, x), \]
\[ 0 = [G_{\hat{\phi}, \hat{\phi}} v](t, i, i_0, x) + f(t, i, \hat{\phi}(t, i, i_0, x), i_0, \hat{\phi}_0(t, i_0, x), x), \]
\[ v(T, i, i_0, x) = g(i, i_0, x), \]

admits a classical solution $(\hat{V}_0, \hat{V})$ (i.e. the solution are $C^1$ in $t$ and $x$). Assume in addition that:

\[ \hat{\phi}_0(t, i_0, x) = \hat{\alpha}_0(t, i_0, x, \hat{V}_0(t, \cdot, x)), \]
\[ \hat{\phi}(t, i, i_0, x) = \hat{\alpha}(t, i, i_0, x, \hat{V}(t, \cdot, \cdot, x)). \]  

Then $\hat{\phi}_0$ and $\hat{\phi}$ form a Nash equilibrium and $\hat{V}_0(0, X_0^0, \mu_0)$ and $\hat{V}(0, X_0, X_0^0, \mu_0)$ are the equilibrium expected costs of the major and minor players.

**Proof.** We show that $\hat{\phi}_0 = \phi_0^*(\hat{\phi})$ and $\hat{\phi} = \phi^*(\hat{\phi}_0, \hat{\phi})$. Notice first that $\hat{\phi}_0$ and $\hat{\phi}$ are Lipschitz strategies due to the regularity of $\hat{V}_0$ and $\hat{V}$, Assumption 5.5.2 and Assumption 5.5.3.
Consider the major player’s optimization problem where we let $\phi = \hat{\phi}$ and denote by $V_0^{\hat{\phi}}$ the corresponding value function. Then since $\hat{V}_0$ is a classical solution to (5.27) and because of (5.28), we deduce that $\hat{V}_0$ is a viscosity solution to the HJB equation (5.16) associated with the value function $V_0^{\hat{\phi}}$. By uniqueness of the viscosity solution, we conclude that $\hat{V}_0 = V_0^{\hat{\phi}}$.

On the other hand, if we denote by $J_0^{\hat{\phi}_0, \hat{\phi}}$ the expected cost function of the major player when it uses the feedback function $\hat{\phi}_0$ and all the minor players use strategy $\hat{\phi}$, then the fact that $\hat{V}_0$ is a classical solution to (5.27) implies that $\hat{V}_0$ is also a viscosity solution. Then by Proposition 5.4.6 we have $J_0^{\hat{\phi}_0, \hat{\phi}} = \hat{V}_0$ and therefore $J_0^{\hat{\phi}_0, \hat{\phi}} = V_0^{\hat{\phi}} = \inf_{\alpha_0 + \hat{\phi}_0} J_0^{\alpha_0, \hat{\phi}_0}$. This means that $\hat{\phi}_0$ is the best response of the major player to the minor players using the feedback function $\hat{\phi}$.

For the optimization problem of the deviating minor player, we use the same argument based on the uniqueness of solution of PDE to obtain $J_0^{\hat{\phi}_0, \hat{\phi}, \hat{\phi}} = \hat{V} = V^{\hat{\phi}_0, \hat{\phi}} = \inf_{\alpha + \hat{\phi}} J_0^{\hat{\phi}_0, \hat{\phi}, \hat{\phi}}$. This implies that $\hat{\phi}$ is the deviating minor player’s best response to the major player using feedback function $\hat{\phi}_0$ and the rest of the minor players using $\hat{\phi}$. We conclude that $\hat{\phi}_0$ and $\hat{\phi}$ form the desired fixed point for the best response map.

It is important to keep in mind that the above verification argument of the master equation does not speak to the problem of existence of Nash equilibria. However, it provides a convenient way to compute numerically the equilibrium via the solution of a coupled system of first-order PDEs.
5.7 Propagation of Chaos and Approximate Nash Equilibria

In this section we show that in the \((N+1)\)-player game (see the description in Section 5.1), when the major player and each minor player apply the respective equilibrium strategies in the mean field game, the system is in an approximate Nash equilibrium. To uncover this link, we first revisit the \((N+1)\)-player game. We show that for a certain strategy profile, the total expected cost of an individual player in the finite player game converges to that of the mean field game. Our argument is largely similar to the one used in proving the convergence of the numerical scheme for viscosity solutions. One crucial intermediate result we use here is the gradient estimate for the value functions of the \((N+1)\)-player game. Similar results were proved in Gomes et al. [2013] for discrete state mean field games without a major player. As a byproduct of the proof, we can also conclude that the value function of the mean field game is Lipschitz in the measure argument. In the rest of this section, we assume that Assumption 5.5.1 is in force.

5.7.1 Back to the \((N+1)\)-Player Game

In this section, we focus on the game with a major player and \(N\) minor players. We show that both the expected costs of individual players and the value functions of the players’ optimization problems can be characterized by coupled systems of ODEs, and their gradients are bounded by some constant independent of \(N\). Such a gradient estimate will be crucial in establishing results on propagation of chaos, as well as the regularity of the value functions for the limiting mean field game.

We start from the major player’s optimization problem. Consider a strategy profile where the major player chooses a Lipschitz feedback function \(\phi_0\) and all the \(N\) minor players choose the same Lipschitz feedback function \(\phi\). Recall that the process
comprising the major player’s state and the empirical distribution of the states of the minor players, i.e. \((X_t^{0,N}, \mu^N_t)\), is a finite state continuous-time Markov chain in the space \(\{1, \ldots, m_0\} \times \bar{S}_N\), where \(\bar{S}_N := \{\frac{1}{N}(k_1, \ldots, k_{m-1}) | \sum k_i \leq N, k_i \in \mathbb{N}\}\). Its infinitesimal generator \(G_{0,N}^{\phi_0,\phi}\) was given by (5.10). The total expected cost to the major player is given by:

\[
J_{0,N}^{\phi_0,\phi}(t, i_0, x) := \mathbb{E}\left[\int_t^T f_0(s, X_s^{0,N}, \mu^N_s) ds + g_0(X_T^{0,N}, \mu^N_T) | X_t^{0,N} = i_0, \mu^N_t = x\right],
\]

and the value function of the major player’s optimization problem by:

\[
V_{0,N}(t, i_0, x) := \inf_{\phi_0 \in \mathbb{A}_0} J_{0,N}^{\phi_0,\phi}(t, i_0, x).
\]

Despite the notation, \(J_{0,N}^{\phi_0,\phi}\) can be viewed as a function defined on \([0, T]\) with values given by vectors indexed by \((i_0, x)\). The following result shows that \(J_{0,N}^{\phi_0,\phi}\) is characterized by a coupled system of ODEs.

**Proposition 5.7.1.** Let \(\phi_0\) (resp. \(\phi\)) be a Lipschitz feedback function for the major (resp. minor) player. Then \(J_{0,N}^{\phi_0,\phi}\) is the unique classical solution of the system of ODEs:

\[
0 = \dot{\theta}(t, i_0, x) + f_0(t, i_0, x) + \sum_{j_0, j_0 \neq i_0} (\theta(t, j_0, x) - \theta(t, i_0, x)) q_0(t, i_0, j_0, x)
+ \sum_{i,j, j \neq i} (\theta(t, i_0, x + \frac{1}{N}e_{ij}) - \theta(t, i_0, x)) N x_i q_0(t, i, j, i_0, x),
\]

(5.29)

\[
0 = \theta(t, i_0, x) - g_0(i_0, x).
\]

**Proof.** The existence and uniqueness of the solution to (5.29) is an easy consequence of the Lipschitz property of the functions \(f_0, q, q_0, \phi_0, \phi\) and the Cauchy-Lipschitz Theorem. The fact that \(J_{0,N}^{\phi_0,\phi}\) is a solution to (5.29) follows from Dynkin’s formula. \(\square\)

We state without proof the similar result for \(V_{0,N}^{\phi}\).
Proposition 5.7.2. If Assumptions 5.5.1, 5.5.2 & 5.5.3 hold and $\phi$ is a Lipschitz strategy, then $V_{0,N}$ is the unique classical solution of the system of ODEs:

\[
0 = \dot{\theta}(t, i_0, x) + \inf_{\alpha_0 \in A_0} \left\{ f_0(t, i_0, \alpha_0, x) + \sum_{j_0, j_0 \neq i_0} (\theta(t, j_0, x) - \theta(t, i_0, x))q_0(t, i_0, j_0, \alpha_0, x) \right\}
+ \sum_{(i,j), j \neq i} \left( \theta(t, i_0, x + \frac{1}{N}e_{ij}) - \theta(t, i_0, x) \right)N_xq(t, i, j, \phi(t, i, i_0, x), i_0, x),
\]

\[
0 = \theta(t, i_0, x) - g_0(i_0, x).
\]

(5.30)

The following estimates for $J_{0,N}^{\phi_0}$ and $V_{0,N}$ will play a crucial role in proving convergence to the solution of the mean field game. Their proofs are postponed to the appendix.

Proposition 5.7.3. For any Lipschitz feedback function $\phi_0$ for the major player and Lipschitz feedback function $\phi$ for the minor player, there exists a constant $L$ only depending on $T$ and the Lipschitz constants and bounds of $\phi_0, \phi, q_0, q, f_0, g_0$ such that for all $N > 0$, $(t, i_0, x) \in [0, T] \times \{1, \ldots, m_0\} \times \bar{S}_N$ and $j, k \in \{1, \cdots, m\}, j \neq k$, we have:

\[
|J_{0,N}^{\phi_0,\phi}(t, i_0, x)| \leq \|g_0\|_\infty + T\|f_0\|_\infty, \tag{5.31}
\]

\[
|J_{0,N}^{\phi_0}(t, i_0, x + \frac{1}{N}e_{jk}) - J_{0,N}^{\phi_0}(t, i_0, x)| \leq \frac{L}{N}. \tag{5.32}
\]

Proposition 5.7.4. For each Lipschitz feedback function $\phi$ with Lipschitz constant $L_\phi$, there exist constants $C_0, C_1, C_2, C_3, C_4 > 0$ only depending on $m_0, m$, the Lipschitz constants and bounds of $g_0, q, f_0, g_0$, such that for all $N > 0$, $(t, i_0, x) \in [0, T] \times \{1, \ldots, m_0\} \times \bar{S}_N$ and $j, k \in \{1, \cdots, m\}, j \neq k$, we have:

\[
|V_{0,N}^{\phi}(t, i_0, x + \frac{1}{N}e_{jk}) - V_{0,N}^{\phi}(t, i_0, x)| \leq \frac{C_0 + C_1T + C_2T^2}{N} \exp[(C_3 + C_4L_\phi)T]. \tag{5.33}
\]

We now turn to the problem of the deviating minor player. We consider a strategy profile where the major player uses a feedback function $\phi_0$, the first $(N - 1)$ minor
players use a feedback function $\phi$ and the remaining (de facto the deviating) minor player uses the feedback function $\bar{\phi}$. We recall that $(X^N_t, X^0_{t}, \mu^N_t)$ is a Markov process with infinitesimal generator $G_{N}^{\phi, \bar{\phi}}$ defined as in (5.13). We are interested in the deviating minor player’s expected cost:

$$J_{N}^{\phi_0, \phi, \bar{\phi}}(t, i, i_0, x) := \mathbb{E}\left[\int_{t}^{T} f_{s}^{\phi_0, \bar{\phi}}(s, X^N_s, X^0_{s}, \mu^N_s) ds + g(X^N_T, X^0_{T}, \mu^N_T) \mid X^N_t = i, X^0_{t} = i_0, \mu^N_t = x\right],$$

as well as the value function of the deviating minor player’s optimization problem:

$$V_{N}^{\phi_0, \phi}(t, i, i_0, x) := \sup_{\alpha \leftrightarrow \bar{\phi} \in \mathcal{A}} J_{N}^{\phi_0, \phi, \bar{\phi}}(t, i, i_0, x).$$

In full analogy with Propositions 5.7.3 and Proposition 5.7.4, we state the following results without proof.

**Proposition 5.7.5.** For any Lipschitz feedback function $\phi_0$ for the major player and Lipschitz feedback function $\phi, \bar{\phi}$ for the minor player, there exists a constant $L$ only depending on $T$ and the Lipschitz constants and bounds of $\phi_0, \phi, \bar{\phi}, q_0, q, f, g$ such that for all $N > 0$, $(t, i_0, x) \in [0, T] \times \{1, \ldots, m_0\} \times \bar{\mathcal{S}}_N$ and $j, k \in \{1, \ldots, m\}, j \neq k$, we have:

$$|J_{N}^{\phi_0, \phi, \bar{\phi}}(t, i, i_0, x)| \leq \|g\|_{\infty} + T\|f\|_{\infty}, \quad (5.34)$$

$$|J_{N}^{\phi_0, \phi, \bar{\phi}}(t, i, i_0, x + \frac{1}{N} e_{jk}) - J_{N}^{\phi_0, \phi, \bar{\phi}}(t, i, i_0, x)| \leq \frac{L}{N}. \quad (5.35)$$

**Proposition 5.7.6.** There exist constants $D_0, D_1, D_2, D_3, D_4, D_5 > 0$ depending only on $m_0, m$, the Lipschitz constants and bounds of $q_0, q, f, g$ such that for all Lipschitz feedback functions $\phi_0$ and $\phi$ with Lipschitz constants $L_{\phi_0}$ and $L_{\phi}$ respectively, and for
all $N > 0$, $(t, i_0, x) \in [0, T] \times \{1, \ldots, m_0\} \times \bar{S}_N$, and $j, k \in \{1, \ldots, m\}, j \neq k$, we have:

\[
|V_{0, N}^{\phi_0, \phi}(t, i_0, x + \frac{1}{N} e_{jk}) - V_{0, N}^{\phi_0, \phi}(t, i_0, x)| \leq \frac{D_0 + D_1 T + D_2 T^2 + D_3 L_{\phi_0} T}{N} \exp[(D_4 + D_5 L_{\phi}) T].
\]

(5.36)

5.7.2 Propagation of Chaos

We now prove two important limiting results. They are related to the propagation of chaos in the sense that they identify the limiting behavior of an individual when interacting with the mean field. First, we prove uniform convergence of the value functions of the individual players’ optimization problems. Combined with the gradient estimates proven in the previous subsection, this establishes the Lipschitz property of the value functions in the mean field limit. Second, we prove that the expected costs of the individual players in the $(N + 1)$-player game converge to their mean field limits at the rate $(N^{-1/2})$. This will help us show that the Nash equilibrium of the mean field game provides approximate Nash equilibria for the finite player games.

**Theorem 5.7.7.** For all Lipschitz strategies $\phi_0, \phi$, we have:

\[
\sup_{t, i_0, x} |V_0^\phi(t, i_0, x) - V_{0, N}^\phi(t, i_0, x)| \to 0, \quad N \to +\infty, \tag{5.37}
\]

\[
\sup_{t, i, i_0, x} |V_{0, N}^{\phi_0, \phi}(t, i, i_0, x) - \hat{V}_{0, N}^{\phi_0, \phi}(t, i, i_0, x)| \to 0, \quad N \to +\infty. \tag{5.38}
\]

**Proof.** We only provide a proof for (5.37), as (5.38) can be shown in the exact the same way.

(i) Fix a Lipschitz strategy $\phi$ for the minor players. To simplify the notation, we set $v_N := V_0^{0, N}$. Notice that $(t, i_0, x) \to v_N(t, i_0, x)$ is only defined on $[0, T] \times \{1, \ldots, m_0\} \times \bar{S}_N$, so our first step is to extend the domain of $v_N(t, i_0, x)$ to $[0, T] \times \{1, \ldots, m_0\} \times \bar{S}$. This can be done by considering the linear interpolation of $v_N$. More specifically, for any $x \in \bar{S}$, we denote $x_k, k = 1, \ldots, 2^{m-1}$ the $2^{m-1}$ closest neighbors of $x$ in the set $\bar{S}_N$. There exists $\alpha_k, k = 1, \ldots, 2^{m-1}$ positive constants
such that $x = \sum_{k=1}^{2m-1} \alpha_k x_k$. We then define the extension, still denoted as $v_N$, to be $v_N(t, i_0, x) := \sum_{k=1}^{2N-1} \alpha_k v_N(t, i_0, x_k)$. It is straightforward to verify that $v_N$ is continuous in $(t, x)$, Lipschitz in $x$ uniformly in $(t, i_0)$, and $C^1$ in $t$. Using the boundedness and Lipschitz continuity of $v_N, f_0, q_0, q$ and $\phi$, we obtain a straightforward estimation:

$$\frac{L}{N} \geq \left| \frac{\dot{v}_N(t, i_0, x) + \inf_{\alpha_0 \in A_0} \{ f_0(t, i_0, \alpha_0, x) + \sum_{j_0, j_0 \neq i_0} (v_N(t, j_0, x) - v_N(t, i_0, x))q_0(t, i_0, j_0, \alpha_0, x) \}}{N} \right|,$$

(5.39)

$$\frac{L}{N} \geq \left| v_N(t, i_0, x) - g_0(i_0, x) \right|.$$

(5.40)

where the constant $L$ only depends on the bounds and Lipschitz constants of $f_0, q_0, q$ and $\phi$.

(ii) Now let us denote $\bar{v}(t, i_0, x) := \limsup v_N(t, i_0, x)$ and $\underline{v}(t, i_0, x) := \liminf v_N(t, i_0, x)$ (see Section 5.8.3 for definitions of the operators $\limsup$ and $\liminf$). We show that $\bar{v}$ and $\underline{v}$ are a viscosity subsolution and a viscosity supersolution of the HJB equation (5.16) of the major player, respectively. Recall that we assume now that $q$ does not depend on $\alpha$. Then since $V^0_\phi$ is also a viscosity solution to (5.16), the comparison principle allows us to conclude that $\bar{v}(t, i_0, x) = \underline{v}(t, i_0, x) = V^0_\phi$ and the uniform convergence follows by standard arguments.

(iii) It remains to show that $\bar{v}$ is a viscosity subsolution to the PDE (5.16). The proof of $\underline{v}$ being a viscosity supersolution can be done in exactly the same way.

Let $\theta$ be a smooth function and $(\bar{t}, i_0, \bar{x}) \in [0, T] \times \{1, \ldots, m_0\} \times \bar{S}$ be such that $(t, x) \rightarrow \bar{v}(t, i_0, x) - \theta(t, x)$ has a maximum at $(\bar{t}, \bar{x})$ and $\bar{v}(\bar{t}, i_0, \bar{x}) = \theta(\bar{t}, \bar{x})$. Then by Lemma 6.1 in Crandall et al. [1992], there exist sequences $N_n \rightarrow +\infty$, $t_n \rightarrow \bar{t}$, $x_n \rightarrow \bar{x}$ such that for each $n$, the mapping $(t, x) \rightarrow v_{N_n}(t, i_0, x) - \theta(t, x)$ attains a maximum at $t_n, x_n$ and $\delta_n := v_{N_n}(t_n, i_0, x_n) - \theta(t_n, x_n) \rightarrow 0$. Instead of extracting a
subsequence, we may assume that $v_{N_n}(t_n, \cdot, x_n) \to (r_1, \ldots, r_{m_0})$, where $r_{j_0} \leq \bar{v}(\bar{t}, j_0, \bar{x})$ and $r_{i_0} = \bar{v}(\bar{t}, i_0, \bar{x})$.

For $\bar{t} = T$, $\bar{v}(\bar{t}, i_0, \bar{x}) \leq g_0(i_0, \bar{x})$ follows easily from (5.40). Now assume that $\bar{t} < T$. Instead of extracting a subsequence, we may assume that $t_n < T$ for all $n$. Then by maximality we have $\partial_t \theta(t_n, x_n) = \partial_t v_{N_n}(t_n, x_n)$. Again by maximality, we have for all $i, j = 1, \ldots, m, i \neq j$:

$$v_{N_n}(t_n, i_0, x_n + \frac{1}{N_n} e_{i,j}) - v_{N_n}(t_n, i_0, x_n) \leq \theta(t_n, i_0, x_n + \frac{1}{N_n} e_{i,j}) - \theta(t_n, i_0, x_n).$$

Injecting the above inequalities into the estimation (5.39) and using the positivity of $q$, we obtain:

$$-\frac{L}{N_n} \leq \partial_t \theta(t_n, i_0, x_n) + \inf_{\alpha_0 \in A_0} \left\{ f_0(t_n, i_0, \alpha_0, x_n) + \sum_{j_0, j_0 \neq i_0} (v_{N_n}(t_n, j_0, x_n) - v_{N_n}(t_n, i_0, x_n))q_0(t_n, i_0, j_0, \alpha_0, x_n) \right\} + \sum_{(i,j), j \neq i} [\theta(t_n, i_0, x_n + \frac{1}{N_n} e_{i,j}) - \theta(t_n, i_0, x_n)]N_n(x_n)q(t_n, i, j, \phi(t_n, i, i_0, x_n), i_0, x_n)$$

Taking the limit in $n$ we obtain:

$$0 \leq \partial_t \theta(\bar{t}, i_0, \bar{x}) + \inf_{\alpha_0 \in A_0} \left\{ f_0(\bar{t}, i_0, \alpha_0, \bar{x}) + \sum_{j_0, j_0 \neq i_0} (r_{j_0} - r_{i_0})q_0(\bar{t}, i_0, j_0, \alpha_0, \bar{x}) \right\} + \sum_{(i,j), j \neq i} [1(j \neq m)\partial_{x_j} \theta(\bar{t}, i_0, \bar{x}) - 1(i \neq m)\partial_{x_i} \theta(\bar{t}, i_0, \bar{x})]x_i q(\bar{t}, i, j, \phi(\bar{t}, i, i_0, \bar{x}), i_0, \bar{x}).$$

Now since $q_0$ is positive and $r_{j_0} \leq \bar{v}(\bar{t}, j_0, \bar{x})$ for all $j_0 \neq i_0$ and $r_{i_0} = \bar{v}(\bar{t}, i_0, \bar{x})$ we have:

$$\inf_{\alpha_0 \in A_0} \left\{ f_0(\bar{t}, i_0, \alpha_0, \bar{x}) + \sum_{j_0, j_0 \neq i_0} (r_{j_0} - r_{i_0})q_0(\bar{t}, i_0, j_0, \alpha_0, \bar{x}) \right\}$$

$$\leq \inf_{\alpha_0 \in A_0} \left\{ f_0(\bar{t}, i_0, \alpha_0, \bar{x}) + \sum_{j_0, j_0 \neq i_0} (\bar{v}(\bar{t}, j_0, \bar{x}) - \bar{v}(\bar{t}, i_0, \bar{x}))q_0(\bar{t}, i_0, j_0, \alpha_0, \bar{x}) \right\}.$$

The desired inequality for the viscosity subsolution follows immediately. This completes the proof. \qed
As an immediate consequence of the uniform convergence and the gradient estimates for the value functions $V_{0,N}^{\phi}$ and $V_{N}^{\phi_0,\phi}$, we have:

**Corollary 5.7.8.** Under Assumption 5.5.1, 5.5.2 & 5.5.3, for all Lipschitz feedback functions $\phi_0$ and $\phi$ with Lipschitz constants $L_{\phi_0}$ and $L_{\phi}$, respectively, the value functions $V_{0}^{\phi}$ and $V_{N}^{\phi_0,\phi}$ are Lipschitz in $(t,x)$. More specifically, there exist strictly positive constants $B, C_i, i = 0, \ldots, 4, D_i, i = 0, \ldots, 5$ that only depend on the bounds and Lipschitz constants of $f, f_0, g_0, g, q$ and $q_0$ such that:

\[ |V_{0}^{\phi}(t,i_0,x) - V_{0}^{\phi}(s,i_0,y)| \leq B|t-s| + (C_0 + C_1T + C_2T^2)\exp((C_3 + C_4L_{\phi})T)||x-y||, \]  

(5.41)

\[ |V_{N}^{\phi_0,\phi}(t,i,i_0,x) - V_{N}^{\phi_0,\phi}(s,i,i_0,y)| \leq B|t-s| + (D_0 + D_1T + D_2T^2 + D_3L_{\phi_0}T)\exp((D_4 + D_5L_{\phi})T)||x-y||. \]  

(5.42)

**Proof.** The Lipschitz property in $x$ is an immediate consequence of Theorem 5.7.7, Proposition 5.7.4 and Proposition 5.7.6. To prove the Lipschitz property on $t$, we remark that for each $N$, $V_{0,N}^{\phi}$ and $V_{N}^{\phi_0,\phi}$ are Lipschitz in $t$, uniformly in $N$. Indeed $V_{0,N}^{\phi}$ is a classical solution of the system of ODEs (5.30). Then it is clear that $\frac{d}{dt}V_{0,N}^{\phi}$ is bounded by the bounds of $f_0, g_0, q_0$ and $q$. We deduce that $V_{0,N}^{\phi}$ is Lipschitz in $t$ with a Lipschitz constant that only depends on $m, m_0$ and the bounds of $f_0, g_0, q_0$ and $q$. By convergence of $V_{0,N}^{\phi}$, we conclude that $V_{0}^{\phi}$ is also Lipschitz in $t$ and shares the Lipschitz constant of $V_{0,N}^{\phi}$. The same argument applies to $V_{N}^{\phi_0,\phi}$.

**Remark 5.7.9.** From Corollary 5.7.8, we see that Assumption 5.5.6 holds when $T$ is sufficiently small. Indeed we can first choose $L_{\phi_0} > L_a(1 + \max\{B, C_0\})$ and $L_{\phi} > L_b(1 + \max\{B, D_0\})$ and then choose $T$ sufficiently small, so that the Lipschitz constant of $V_{0}^{\phi}$ is smaller than $(L_{\phi_0}/L_a - 1)$ and the Lipschitz constant of $V_{\phi_0,\phi}$ is smaller than $(L_{\phi}/L_b - 1)$.  

182
We now state our second result on the propagation of chaos: the expected cost of an individual player in the \((N + 1)\)-player game converges to the expected cost in the mean field game at a rate of \(N^{-1/2}\).

**Theorem 5.7.10.** There exists a constant \(L\) depending only on \(T\) and the Lipschitz constants of \(\phi_0, \phi, \phi_0, f, f_0, g, g_0, q\) and \(q_0\) such that for all \(N > 0, t \leq T, x \in \bar{S}_N, i = 1, \ldots, m\) and \(i_0 = 1, \ldots, m_0\), we have:

\[
|J_{0,N}^{\phi_0,\phi}(t, i_0, x) - J_{0,N}^{\phi_0,\phi}(t, i_0, x)| \leq L/\sqrt{N},
\]

\[
|J_{N}^{\phi_0,\bar{\phi}}(t, i, i_0, x) - J_{N}^{\phi_0,\phi,\bar{\phi}}(t, i, i_0, x)| \leq L/\sqrt{N}.
\]

Our proof is based on standard techniques from the convergence rate analysis of numerical schemes for viscosity solutions of PDEs, c.f. Briani et al. [2012] and Barles and Souganidis [1991] for example. The key step of the proof is the construction of a smooth subsolution and a smooth supersolution of the PDEs (5.23) and (5.24) that characterize \(J_{0,N}^{\phi_0,\phi}\) and \(J_{N}^{\phi_0,\phi,\bar{\phi}}\), respectively (see Proposition 5.4.6). We construct these solutions by mollifying an extended version of \(J_{0,N}^{\phi_0,\phi}\). Then we derive the bound by using the comparison principle. In the following, we detail the proof for the convergence rate of the major player’s expected cost. The case of the generic minor player can be dealt with exactly the same way.

Since \(J_{0,N}^{\phi_0,\phi}(t, i, x)\) is only defined for \(x \in \bar{S}_N\), in order to mollify \(J_{0,N}^{\phi_0,\phi}\), we need to first construct an extension of \(J_{0,N}^{\phi_0,\phi}\) defined for all \(x \in O\) for an open set \(O\) containing \(\bar{S}\). To this end, we consider the following system of ODEs:

\[
0 = \dot{\theta}(t, i_0, x) + \tilde{f}_0(t, i_0, \tilde{\phi}_0(t, i_0, x), x) + \sum_{j_0 \neq i_0} (\theta(t, j_0, x) - \theta(t, i_0, x))\tilde{q}_0(t, i_0, j_0, \tilde{\phi}_0(t, i_0, x), x)
\]

\[
+ \sum_{(i,j), j \neq i} (\theta(t, i_0, x + \frac{1}{N} e_{ij}) - \theta(t, i_0, x))N \max\{x_i, 0\}\tilde{q}(t, i, j, \tilde{\phi}(t, i, i_0, x), i_0, \tilde{\phi}_0(t, i_0, x), x),
\]

\[
0 = \theta(T, i_0, x) - \tilde{g}_0(i_0, x).
\]
Here $\tilde{\phi}_0$, $\tilde{\phi}$, $\tilde{f}_0$, $\tilde{g}_0$ and $\tilde{q}_0$ are, respectively, extensions of $\phi_0$, $\phi$, $f_0$, $g_0$ and $q_0$ from $x \in \tilde{S}$ to $x \in \mathbb{R}^{N-1}$, which are Lipschitz in $x$. The following is proved using the same arguments as for Proposition 5.7.1.

**Lemma 5.7.11.** The system of ODEs (5.29) admits a unique solution $v_N$ defined in $[0,T] \times \{1,\ldots,m_0\} \times \mathbb{R}^{m-1}$. Moreover we have:

(i) $v_N(t,i_0,x)$ is Lipschitz in $x$ uniformly in $t$ and $i_0$ and the Lipschitz constant only depends on $T$ and the Lipschitz constants of $\phi_0$, $\phi$, $f_0$, $g_0$ and $q_0$.

(ii) $v_N(t,i_0,x) = J^{\phi_0,\phi}_{0,N}(t,i_0,x)$ for all $x \in \tilde{S}_N$.

To construct smooth super and sub solutions, we use a family of mollifiers $\rho^\varepsilon$ defined by $\rho^\varepsilon(x) := \rho(x/\varepsilon)/\varepsilon^{N-1}$, where $\rho$ is a smooth and positive function with compact support in the unit ball of $\mathbb{R}^{m-1}$ and satisfying $\int_{\mathbb{R}^{m-1}} \rho(x)dx = 1$. For $\varepsilon > 0$, we define $v_N^\varepsilon$ as the mollification of $v_N$ on $[0,T] \times \{1,\ldots,m_0\} \times \tilde{S}$:

$$v_N^\varepsilon(t,i_0,x) := \int_{y \in \mathbb{R}^{m-1}} v_N(t,i_0,x-y)\rho^\varepsilon(y)dy.$$  

Using the Lipschitz property of $\phi_0$, $\phi$, $f_0$, $g_0$ and $q_0$ and straightforward estimates on the mollifier $\rho^\varepsilon$, we obtain the following properties on $v_N^\varepsilon$.

**Lemma 5.7.12.** $v_N^\varepsilon$ is $C^\infty$ in $x$ and $C^1$ in $t$. Moreover, there exists a constant $C$ that depends only on $T$ and the Lipschitz constants of $\phi_0$, $\phi$, $f_0$, $g_0$ and $q_0$ such that for all $i_0 = 1,\ldots,m_0$, $i = 1,\ldots,m$, $t \leq T$ and $x,y \in \tilde{S}$, the following estimations hold:

$$|[\mathcal{G}^{\phi_0,\phi}_{0,N}v_N^\varepsilon](t,i_0,x) + f_0(t,i_0,\phi_0(t,i_0,x),x)| \leq L\varepsilon,$$  

$$|v_N^\varepsilon(t,i_0,x) - g_0(i_0,x)| \leq L\varepsilon,$$  

$$|\partial_x v_N^\varepsilon(t,i_0,x) - \partial_x v_N^\varepsilon(t,i_0,y)| \leq \frac{L}{\varepsilon} \|x-y\|.$$  

184
We are now ready to prove Theorem 5.7.10. We construct viscosity super and sub solutions by adjusting $v_N^\epsilon$ with a linear function on $t$. Then the comparison principle allows us to conclude.

**Proof.** (of Theorem 5.7.10) We denote by $L$ a generic constant that only depends on $T$ and the Lipschitz constants of $\phi_0, \phi, f_0, g_0$ and $q_0$. Using (5.43), (5.44) and (5.45) in Lemma 5.7.12, we obtain:

$$|G_0^{\phi_0, \phi} v_N^\epsilon(t, i_0, x) + f_0(t, i_0, \phi_0(t, i_0, x), x)| \leq L \left( \epsilon + \frac{1}{N \epsilon} \right),$$

$$|v_N^\epsilon(T, i_0, x) - g_0(i_0, x)| \leq L \epsilon.$$

Next we define:

$$v_N^\pm(t, i_0, x) := v_N^\epsilon(t, i_0, x) \pm \left[ L(\epsilon + \frac{1}{N \epsilon})(T - t) + L \epsilon \right].$$

Since $v_N^-\epsilon$ and $v_N^+\epsilon$ are smooth, the above estimation immediately implies that $v_N^\epsilon$ and $v_N^\pm\epsilon$ are viscosity sub and super solutions of the PDE (5.23), respectively. By Proposition 5.4.6, $J_0^{\phi_0, \phi}$ is a continuous viscosity solution to the PDE (5.23). Then by the comparison principle we have $v_N^- \leq J_0^{\phi_0, \phi} \leq v_N^+\epsilon$, which implies:

$$|v_N^\epsilon(t, i_0, x) - J_0^{\phi_0, \phi}(t, i_0, x)| \leq L(\epsilon + \frac{1}{N \epsilon})(T - t) + L \epsilon.$$

Now using the property of the mollifier and Lemma 5.7.11, we have for all $t \leq T$, $i_0 = 1, \ldots, m_0$ and $x \in \bar{S}_N$:

$$|v_N^\epsilon(t, i_0, x) - J_0^{\phi_0, \phi}(t, i_0, x)| = |v_N^\epsilon(t, i_0, x) - v_N^N(t, i_0, x)| \leq L \epsilon.$$

The desired results follow by combining the above inequalities and choosing $\epsilon = 1/\sqrt{N}$. \qed
5.7.3 Approximate Nash Equilibria

The following is an immediate consequence of the above propagation of chaos results.

**Theorem 5.7.13.** Assume that the Mean Field Game attains Nash equilibrium when the major player uses a Lipschitz feedback function \( \hat{\phi}_0 \) and all the minor players uses a Lipschitz feedback function \( \hat{\phi} \). Denote \( L_0 \) the Lipschitz constant for \( \hat{\phi}_0 \) and \( \hat{\phi} \). Then for any \( L \geq L_0 \), the strategy profile \((\hat{\alpha}_0, \hat{\alpha}_1, \ldots, \hat{\alpha}_N)\) defined by

\[
\hat{\alpha}_0(t) := \hat{\phi}_0(t, X_t^{0,N}, \mu_t^N) \quad \text{and} \quad \hat{\alpha}_n(t) := \hat{\phi}(t, X_t^{n,N}, X_t^{0,N}, \mu_t^N)
\]

gives an approximative Nash equilibrium within all the \( L \)-Lipschitz strategies. More specifically, there exist constants \( C > 0 \) and \( N_0 \in \mathbb{N} \) depending on \( L_0 \) such that for any \( N \geq N_0 \), any \( L \)-Lipschitz feedback function \( \phi_0 \) for major player and any \( L \)-Lipschitz feedback function \( \phi \), we have:

\[
J_{0,N}(\hat{\phi}_0, \hat{\phi}, \ldots, \hat{\phi}) \leq J_{0,N}(\phi_0, \hat{\phi}, \ldots, \hat{\phi}) + C/\sqrt{N},
\]

\[
J_N(\hat{\alpha}_0, \hat{\alpha}, \ldots, \hat{\phi}) \leq J_N(\phi_0, \hat{\phi}, \ldots, \hat{\phi}) + C/\sqrt{N}.
\]

5.8 Appendix

5.8.1 Proof of Lemma 5.4.3

Recall the dynamic programming operator defined in (5.18). Let \( \Phi \) be a mapping on \([0, T] \times \{1, \ldots, m_0\} \times \mathcal{S} \) such that \( \Phi(\cdot, i_0, \cdot) \) is \( C^1 \) in \([0, T] \times \mathcal{S} \) and \( \Phi(\cdot, j_0, \cdot) \) is continuous in \([0, T] \times \mathcal{S} \) for \( j_0 \neq i_0 \). We are going to evaluate the following limit:

\[
\lim_{h \to 0} I_h := \lim_{h \to 0} \frac{1}{h} \{ [\mathcal{T}_{t,t+h}\Phi(t+h, \cdot, \cdot)](i_0, x) - \Phi(t, i_0, x) \}.
\]

(i) Let us first assume that \( \Phi(\cdot, j_0, \cdot) \) is \( C^1 \) in \([0, T] \times \mathcal{S} \) for all \( j_0 \). Consider a constant control \( \alpha_0 \). Then by the definition of the operator \( \mathcal{T}_{t,t+h} \) we have:

\[
I_h \leq \frac{1}{h} \left\{ \mathbb{E} \left[ \int_t^{t+h} f_0(u, X_u^0, \alpha_0, \mu_u)du + \Phi(t+h, X_{t+h}^0, \mu_{t+h})|X_t^0 = i_0, \mu = x \right] - \Phi(t, i_0, x) \right\}.
\]
Using the infinitesimal generator of the process \((u, X^0_u, \mu_u)\), the right hand side of the above inequality has a limit. Taking the limit and then the infimum over \(\alpha\), we obtain:

\[
\limsup_{h \to 0} I_h \leq \partial_t \Phi(t, i_0, x)
+ \inf_{\alpha_0 \in A_0} \left\{ f_0(t, i_0, \alpha_0, x) + \sum_{i,j=1}^{m-1} \partial_{x_j} \Phi(t, i_0, x) x_i q(t, i, j, \phi(t, i, i_0, x), i_0, \alpha_0, x) \right.

+ (1 - \sum_{k=1}^{m-1} x_k) \sum_{k=1}^{m-1} \partial_{x_k} \Phi(t, i_0, x) q(t, m, k, \phi(t, m, i_0, x), i_0, \alpha_0, x)

+ \left. \sum_{j_0 \neq i_0} [\Phi(t, j_0, x) - \Phi(t, i_0, x)] q_0(i_0, j_0, \alpha_0, x) \right\}.
\]

On the other hand, for all \(h > 0\), there exists a control \(\phi^h_{0} \) such that:

\[
\mathbb{E}\left[ \int_t^{t+h} f_0(u, X^0_u, \phi^h_{0}(u, X^0_u, \mu_u), \mu_u) du + \Phi(t + h, X^0_{t+h}, \mu_{t+h}) \bigg| X^0_t = i_0, \mu = x \right] \\
\leq T_{t,t+h} \Phi(t + h, \cdot, \cdot)(i_0, x) + h^2.
\]

This implies:

\[
I_h \geq \frac{1}{h} \left\{ \mathbb{E}\left[ \int_t^{t+h} f_0(u, X^0_u, \phi^h_{0}(u, X^0_u, \mu_u), \mu_u) du + \Phi(t + h, X^0_{t+h}, \mu_{t+h}) \bigg| X^0_t = i_0, \mu = x \right] \\
- \Phi(t, i_0, x) \right\} - h.
\]

Since \(\Phi(\cdot, j_0, \cdot)\) is \(C^1\) for all \(j_0\), this can be further written using the infinitesimal generator:

\[
I_h \geq \frac{1}{h} \mathbb{E}\left[ \int_t^{t+h} f_0(u, X^0_u, \phi^h_{0}(u, X^0_u, \mu_u), \mu_u) + [G^0_{\phi^h_{0}, \phi} \Phi](u, X^0_u, \mu_u) du \bigg| X^0_t = i_0, \mu = x \right] - h.
\]

Taking the supremum over the control and applying the dominated convergence theorem, we obtain:

\[
\liminf_{h \to 0} I_h
\]
\[
\geq \partial_t \Phi(t, i_0, x) + \inf_{\alpha_0 \in A_0} \left\{ f_0(t, i_0, \alpha_0, x) + \sum_{i,j=1}^{m-1} \partial_{x_j} \Phi(t, i_0, x)x_i q(t, i, j, \phi(t, i, i_0, x), i_0, \alpha_0, x) \right. \\
+ (1 - \sum_{k=1}^{m-1} x_k) \sum_{k=1}^{m-1} \partial_{x_k} \Phi(t, i_0, x)q(t, m, k, \phi(t, m, i_0, x), i_0, \alpha_0, x) \\
\left. + \sum_{j_0 \neq i_0} [\Phi(t, j_0, x) - \Phi(t, i_0, x)]q_0(i_0, j_0, \alpha_0, x) \right\}.
\]

This proves the lemma for \( \Phi \) such that \( \Phi(\cdot, j_0, \cdot) \) is \( C^1 \) for all \( j_0 \).

(ii) Now take a continuous mapping \( \Phi \) and only assume that \( \Phi(\cdot, i_0, \cdot) \) is \( C^1 \). Applying the Weierstrass approximation theorem, for any \( \epsilon > 0 \), there exists a \( C^1 \) function \( \phi_{j_0}^\epsilon \) on \([0, T] \times S\) for all \( j_0 \neq i_0 \) such that:

\[
\sup_{(t, x) \in [0, T] \times S} |\Phi(t, j_0, x) - \phi_{j_0}^\epsilon(t, x)| \leq \epsilon.
\]

Define \( \Phi'(t, j_0, x) := \phi_{j_0}^\epsilon(t, x) + \epsilon \) for \( j_0 \neq i_0 \) and \( \Phi'(t, i_0, x) := \Phi(t, i_0, x) \). Then we have \( \Phi^\epsilon \geq \Phi \). By the monotonicity of the operator \( T_{t,t+h} \) we have:

\[
\frac{1}{h} \{ [T_{t,t+h}\Phi'(t + h, \cdot, \cdot)](i_0, x) - \Phi'(t, i_0, x) \} \geq \frac{1}{h} \{ [T_{t,t+h}\Phi(t + h, \cdot, \cdot)](i_0, x) - \Phi(t, i_0, x) \} := I_h.
\]

Now applying the results from step (i), we obtain:

\[
\limsup_{h \to 0} I_h \\
\leq \partial_t \Phi'(t, i_0, x) + \inf_{\alpha_0 \in A_0} \left\{ f_0(t, i_0, \alpha_0, x) + \sum_{i,j=1}^{m-1} \partial_{x_j} \Phi'(t, i_0, x)x_i q(t, i, j, \phi(t, i, i_0, x), i_0, \alpha_0, x) \right.
\]
\[
+ (1 - \sum_{k=1}^{m-1} x_k) \sum_{k=1}^{m-1} \partial_{x_k} \Phi'(t, i_0, x)q(t, m, k, \phi(t, m, i_0, x), i_0, \alpha_0, x) \\
\left. + \sum_{j_0 \neq i_0} [\Phi'(t, j_0, x) - \Phi'(t, i_0, x)]q_0(t, i_0, j_0, \alpha_0, x) \right\}
\]

\[
= \partial_t \Phi(t, i_0, x) + \inf_{\alpha_0 \in A_0} \left\{ f_0(t, i_0, \alpha_0, x) + \sum_{i,j=1}^{m-1} \partial_{x_j} \Phi(t, i_0, x)x_i q(t, i, j, \phi(t, i, i_0, x), i_0, \alpha_0, x) \right.
\]
\[
+ (1 - \sum_{k=1}^{m-1} x_k) \sum_{k=1}^{m-1} \partial_{x_k} \Phi(t, i_0, x)q(t, m, k, \phi(t, m, i_0, x), i_0, \alpha_0, x) \\
\left. + \sum_{j_0 \neq i_0} [\Phi(t, j_0, x) - \Phi(t, i_0, x)]q_0(t, i_0, j_0, \alpha_0, x) \right\}.
\]
\[ + \sum_{j_0 \neq i_0} \{ \Phi(t, j_0, x) - \Phi(t, j_0, x) + \Phi(t, i_0, x) \} q_0(t, i_0, j_0, \alpha_0, x) \}\]

\[ \leq \partial_t \Phi(t, i_0, x) + \inf_{\alpha_0 \in A_0} \left\{ f_0(t, i_0, \alpha_0, x) + \sum_{i, j=1}^{m-1} \partial_{x_j} \Phi(t, i_0, x) q(t, i, j, \phi(t, i, i_0, x), i_0, \alpha_0, x) \right. \]

\[ + (1 - \sum_{k=1}^{m-1} x_k) \sum_{k=1}^{m-1} \partial_{x_k} \Phi(t, i_0, x) q(t, m, k, \phi(t, m, i_0, x), i_0, \alpha_0, x) \]

\[ + \sum_{j_0 \neq i_0} \{ \Phi(t, j_0, x) - \Phi(t, i_0, x) \} q_0(t, i_0, j_0, \alpha_0, x) \}\] + \epsilon L.

The last equality is due to the fact that \( q_0 \) is bounded and \( |\Phi^\epsilon(t, j_0, x) - \Phi(t, j_0, x)| \leq \epsilon \). We can write a similar inequality for \( \liminf_{h \to 0} I_h \). Then tending \( \epsilon \) to 0 yields the desired result.

### 5.8.2 Proof of Theorem 5.4.4

In this section, we present the proof of the comparison principle for the HJB equation associated with the major player’s optimization problem. The arguments used in this proof can be readily applied to prove the uniqueness of the solution to the deviating minor player’s HJB equation (5.20) (c.f. Theorem 5.4.5). The same argument can also be used to prove the uniqueness result for equations (5.23) and (5.24) (c.f. Proposition 5.4.6).

Let \( v \) and \( w \) be, respectively, a viscosity subsolution and a viscosity supersolution to equation (5.16). Our objective is to show \( v(t, i_0, x) \leq w(t, i_0, x) \) for all \( 1 \leq i_0 \leq m_0 \), \( x \in \bar{S} \) and \( t \in [0, T] \).
(i) Without loss of generality, we may assume that \( v \) is a viscosity subsolution of:

\[
0 = -\eta + \partial_t v_0(t, i_0, x) + \inf_{\alpha_0 \in A_0} \left\{ f_0(t, i_0, \alpha_0, x) + \sum_{j_0 \neq i_0} [v_0(t, j_0, x) - v_0(t, i_0, x)] g_0(t, i_0, j_0, \alpha_0, x) + \sum_{k=1}^{m-1} x_k v_0(t, i_0, x) q(t, m, k, \phi(t, m, i_0), i_0, \alpha_0, x) + \sum_{i,j=1}^{m-1} \partial_x^i v_0(t, i_0, x) x_i q(t, i, j, \phi(t, i, i_0), i_0, \alpha_0, x) \right\},
\]

\[
0 = v_0(T, i_0, x) - g_0(i_0, x), \quad \forall x \in \overline{\mathcal{S}},
\]

where \( \eta > 0 \) is a small parameter. Indeed, we may consider the function \( v_\eta(t, i, x) := v(t, i, x) - \eta(T - t) \). Then it is easy to see that \( v_\eta \) is a viscosity subsolution to the above equation. If we can prove \( v_\eta \leq w \), then tending \( \eta \) to 0 yields \( v \leq w \). In the following, we will only consider the subsolution \( v \) to equation (5.46) and the supersolution \( w \) to equation (5.16), and prove that \( v \leq w \).

(ii) For \( \epsilon > 0 \) and \( 1 \leq i_0 \leq m_0 \), consider the function \( \Gamma_{i_0, \epsilon} \) defined on \([0, T]^2 \times \overline{\mathcal{S}}^2\):

\[
\Gamma_{i_0, \epsilon}(t, s, x, y) := v(t, i_0, x) - w(s, i_0, y) - \frac{1}{\epsilon} |t - s|^2 - \frac{1}{\epsilon} \|x - y\|^2,
\]

where \( \| \cdot \| \) is the euclidian norm on \( \mathbb{R}^{(m_0 - 1)} \). Since \( \Gamma_{i_0, \epsilon} \) is a continuous function on a compact set, it attains the maximum denoted as \( N_{i_0, \epsilon} \). Denote by \((\hat{t}, \hat{s}, \hat{x}, \hat{y})\) the maximizer (which obviously depends on \( \epsilon \) and \( i_0 \), but for simplicity, we suppress this dependency in the notation). We show that for all \( 1 \leq i_0 \leq m_0 \), there exists a sequence \( \epsilon_n \to 0 \) and the corresponding maximizer \((\hat{t}_n, \hat{s}_n, \hat{x}_n, \hat{y}_n)\) such that

\[
(\hat{t}_n, \hat{s}_n, \hat{x}_n, \hat{y}_n) \to (\hat{t}, \hat{s}, \hat{x}, \hat{y}), \text{ where } (\hat{t}, \hat{x}) := \arg \sup_{(t, x) \in [0, T] \times \mathcal{S}} \{v(t, i_0, x) - w(t, i_0, x)\}, \quad (5.47a)
\]

\[
\frac{1}{\epsilon_n} |\hat{t}_n - \hat{s}_n|^2 + \frac{1}{\epsilon_n} \|\hat{x}_n - \hat{y}_n\|^2 \to 0, \quad (5.47b)
\]
\[ N_{i_0, \epsilon} \rightarrow N_{i_0} := \sup_{(t, x) \in [0, T] \times S} \{ v(t, i_0, x) - w(t, i_0, x) \}. \]  

(5.47c)

Indeed, for any \((t, x) \in [0, T] \times S\), we have \(v(t, i_0, x) - w(t, i_0, x) = \Gamma_{i_0, \epsilon}(t, t, x, x) \leq N_{i_0, \epsilon}\). Taking the supremum, we obtain \(N_{i_0} \leq N_{i_0, \epsilon}\) and therefore:

\[ \frac{1}{\epsilon} |\tilde{t} - \tilde{s}|^2 + \frac{1}{\epsilon} ||\tilde{x} - \tilde{y}||^2 \leq v(\tilde{t}, i_0, \tilde{x}) - w(\tilde{s}, i_0, \tilde{y}) - N_{i_0} \leq 2L - N_{i_0}. \]

The last inequality comes from the fact that \(v\) and \(w\) are bounded on the compact set \([0, T] \times S\). It follows that \(|\tilde{t} - \tilde{s}|^2 + ||\tilde{x} - \tilde{y}||^2 \rightarrow 0\). Now since the sequence \((\tilde{t}, \tilde{s}, \tilde{x}, \tilde{y})\) (indexed by \(\epsilon\)) is in a compact set, we can extract a subsequence \(\epsilon_n \rightarrow 0\) such that \((\tilde{t}_n, \tilde{s}_n, \tilde{x}_n, \tilde{y}_n) \rightarrow (\hat{t}, \hat{s}, \hat{x}, \hat{y})\). We have the following inequality:

\[
N_i \leq v(\tilde{t}_n, i_0, \tilde{x}_n) - w(\tilde{s}_n, i_0, \tilde{y}_n) - \frac{1}{\epsilon_n} |\tilde{t}_n - \tilde{s}_n|^2 - \frac{1}{\epsilon_n} ||\tilde{x}_n - \tilde{y}_n||^2 \\
= N_{i_0, \epsilon_n} \leq v(\tilde{t}_n, i_0, \tilde{x}_n) - w(\tilde{s}_n, i_0, \tilde{y}_n).
\]

Noticing that \(v(\tilde{t}_n, i_0, \tilde{x}_n) - w(\tilde{s}_n, i_0, \tilde{y}_n) \rightarrow v(\hat{t}, i_0, \hat{x}_n) - w(\hat{s}, i_0, \hat{y}_n) \leq N_{i_0}\), we deduce that \(N_{i_0} = v(\hat{t}, i_0, \hat{x}_n) - w(\hat{s}, i_0, \hat{y}_n)\) which implies (5.47a). (5.47b) and (5.47c) follows easily by taking the limit in \(n\) in the above inequality.

(iii) Now we prove the comparison principle. Using the notation introduced in step (ii), we need to prove \(N_{i_0} \leq 0\) for all \(1 \leq i_0 \leq m_0\). Assume that there exists \(1 \leq i_0 \leq m_0\) such that:

\[ N_{i_0} = \sup_{1 \leq j_0 \leq m_0} N_{j_0} > 0. \]

We now work towards a contradiction. Without loss of generality we assume that \(N_{i_0} > N_{j_0}\) for all \(j_0 \neq i_0\). We then consider the subsequence \((\epsilon_n, \tilde{t}_n, \tilde{s}_n, \tilde{x}_n, \tilde{y}_n) \rightarrow (0, \hat{t}, \hat{s}, \hat{x}, \hat{y})\) with regard to \(i_0\) constructed in step (ii), for which (5.47a), (5.47b) and (5.47c) are satisfied. Since \(v(\hat{t}, i_0, \hat{x}) - w(\hat{t}, i_0, \hat{x}) = N_{i_0} > 0\), we have \(\hat{t} \neq T\). Instead of extracting a subsequence, we may assume that \(\tilde{t}_n \neq T\) and \(\tilde{s}_n \neq T\) for all \(n \geq 0\).
moreover for any $j_0 \neq i_0$, we have:

$$v(\bar{t}, j_0, \bar{x}) - w(\bar{t}, j_0, \bar{x}) \leq N_j < N_i = v(\hat{t}, i_0, \bar{x}) - w(\hat{t}, i_0, \bar{x}).$$

Since $v(\bar{t}_n, j_0, \bar{x}_n) - w(s_n, j_0, \bar{y}_n) \to v(\bar{t}, j_0, \bar{x}) - w(\bar{t}, j_0, \bar{x})$, instead of extracting a subsequence, we can assume that for all $j_0 \neq i_0$ and $n \geq 0$:

$$v(\bar{t}_n, j_0, \bar{x}_n) - w(s_n, j_0, \bar{y}_n) \leq v(\bar{t}_n, i_0, \bar{x}_n) - w(s_n, i_0, \bar{y}_n).$$

(5.48)

In the following we suppress the index $n$ for the sequence $(\epsilon_n, \bar{t}_n, s_n, \bar{x}_n, \bar{y}_n)$ for sake of simplicity of notation. By definition of the maximizer, for any $(t, x) \in [0, T] \times \mathcal{S}$, we have:

$$v(t, i_0, x) - w(s, i_0, y) - \frac{1}{\epsilon}||x - y||^2 - \frac{1}{\epsilon}||t - s||^2 \leq v(\bar{t}, i_0, \bar{x}) - w(s, i_0, \bar{y}) - \frac{1}{\epsilon}||\bar{x} - \bar{y}||^2 - \frac{1}{\epsilon}||\bar{t} - \bar{s}||^2.

Therefore $v(\cdot, i_0, \cdot) - \phi$ attains a maximum at $(\bar{t}, \bar{x})$ where:

$$\phi(t, x) := \frac{1}{\epsilon}||x - y||^2 + \frac{1}{\epsilon}||t - s||^2, \quad \partial_x \phi(\bar{t}, \bar{x}) = \frac{2}{\epsilon}(\bar{t} - \bar{s}), \quad \nabla \phi(\bar{t}, \bar{x}) = \frac{2}{\epsilon}(\bar{x} - \bar{y}).$$

(5.49)

Similarly $w(\cdot, i_0, \cdot) - \psi$ attains a minimum at $(\bar{s}, \bar{y})$ where:

$$\psi(t, x) := -\frac{1}{\epsilon}||x - \bar{x}||^2 - \frac{1}{\epsilon}||t - \bar{t}||^2 \quad \partial_x \psi(\bar{s}, \bar{y}) = \frac{2}{\epsilon}(\bar{t} - \bar{s}) \quad \nabla \psi(\bar{s}, \bar{y}) = \frac{2}{\epsilon}(\bar{x} - \bar{y})$$

(5.50)

Since $\bar{s} \neq T$ and $\bar{t} \neq T$, the definition of the viscosity solution, along with equation (5.49) and (5.50) gives the following inequalities:

$$\eta \leq \frac{2}{\epsilon}(\bar{t} - \bar{s}) + \inf_{\alpha_0 \in \mathbb{A}_0} \left\{ f_0(\bar{t}, i_0, \alpha_0, \bar{x}) + \sum_{i,j=1}^{m-1} \frac{2}{\epsilon}(\bar{x}_j - \bar{y}_j)\bar{x}_i q(\bar{t}, i, j, \phi(\bar{t}, i, i_0, \bar{x}), i_0, \alpha_0, \bar{x}) \right.$$

$$+ \left. (1 - \sum_{k=1}^{m-1} \bar{x}_k) \sum_{k=1}^{m-1} \frac{2}{\epsilon}(\bar{x}_k - \bar{y}_k)q(\bar{t}, m, k, \phi(\bar{t}, m, i_0, \bar{x}), i_0, \alpha_0, \bar{x}) \right.$$  

$$+ \left. \sum_{j_0 \neq i_0} [v(\bar{t}, j_0, \bar{x}) - v(\bar{t}, i_0, \bar{x})]q_0(\bar{t}, i_0, j_0, \alpha_0, \bar{x}) \right\}.$$
\[0 \geq \frac{2}{\epsilon}(\bar{t} - \bar{s}) + \inf_{\alpha_0 \in A_0} \{ f_0(\bar{s}, i_0, \alpha_0, \bar{y}) + \sum_{i,j=1}^{m-1} \frac{2}{\epsilon}(x_j - y_j)\bar{y}_i q(\bar{s}, i, j, \phi(\bar{s}, i, i_0, \bar{y}), i_0, \alpha_0, \bar{y}) \]

\[+ (1 - \sum_{k=1}^{m-1} \bar{y}_k) \sum_{k=1}^{m-1} \frac{2}{\epsilon}(\bar{x}_k - \bar{y}_k)q(\bar{s}, m, k, \phi(\bar{s}, m, i_0, \bar{y}), i_0, \alpha_0, \bar{y})\]

\[+ \sum_{j_0 \neq i_0} [w(\bar{s}, j_0, \bar{y}) - w(\bar{s}, i_0, \bar{y})]q_0(\bar{s}, i_0, j_0, \alpha_0, \bar{y}) \}.

Subtracting the above two inequalities, we obtain:

\[0 < \eta \leq I_1 + I_2 + I_3, \quad (5.51)\]

where the three terms \(I_1, I_2\) and \(I_3\) will be dealt with in the following. For \(I_1\) we have:

\[I_1 := \sup_{\alpha_0 \in A_0} \{ f_0(\bar{t}, i_0, \alpha_0, \bar{x}) - f_0(\bar{s}, i_0, \alpha_0, \bar{y}) \}

\[+ \sup_{\alpha_0 \in A_0} \left\{ \sum_{j_0 \neq i_0} [v(\bar{t}, j_0, \bar{x}) - v(\bar{t}, i_0, \bar{x})]q_0(\bar{t}, i_0, j_0, \alpha_0, \bar{x}) - \sum_{j_0 \neq i_0} [w(\bar{s}, j_0, \bar{y}) - w(\bar{s}, i_0, \bar{y})]q_0(\bar{s}, i_0, j_0, \alpha_0, \bar{y}) \right\}

\[\leq \sup_{\alpha_0 \in A_0} \{ f_0(\bar{t}, i_0, \alpha_0, \bar{x}) - f_0(\bar{s}, i_0, \alpha_0, \bar{y}) \}

\[+ \sup_{\alpha_0 \in A_0} \left\{ \sum_{j_0 \neq i_0} [v(\bar{t}, j_0, \bar{x}) - v(\bar{t}, i_0, \bar{x})][q_0(\bar{t}, i_0, j_0, \alpha_0, \bar{x}) - q_0(\bar{s}, i_0, j_0, \alpha_0, \bar{y})] \right\}

\[+ \sup_{\alpha_0 \in A_0} \left\{ \sum_{j_0 \neq i_0} [v(\bar{t}, j_0, \bar{x}) - w(\bar{s}, j_0, \bar{y}) - v(\bar{t}, i_0, \bar{x}) + w(\bar{s}, i_0, \bar{y})]q_0(\bar{s}, i_0, j_0, \alpha_0, \bar{y}) \right\}

\[\leq L(\bar{t} - \bar{s}) + \|\bar{x} - \bar{y}\| + 2CL(\|\bar{t} - \bar{s}\| + \|\bar{x} - \bar{y}\|) + 0.\]

In the last inequality we use (5.48) and the fact that \(q_0(\bar{s}, i_0, j_0, \alpha_0, \bar{y}) \geq 0\) for \(j_0 \neq i_0\). We also use the Lipschitz property of \(f_0\) and \(q_0\). Now in light of (5.47a), we obtain \(I_1 \to 0\) as \(\epsilon \to 0\). Now turning to \(I_2\):

\[I_2 := \sup_{\alpha_0 \in A_0} \left\{ \sum_{i,j=1}^{m-1} \frac{2}{\epsilon}(\bar{x}_j - \bar{y}_j)[\bar{x}_i q(\bar{t}, i, j, \phi(\bar{t}, i, i_0, \bar{x}), i_0, \alpha_0, \bar{x}) - \bar{y}_i q(\bar{s}, i, j, \phi(\bar{s}, i, i_0, \bar{y}), i_0, \alpha_0, \bar{y})] \right\}

\[\leq \sup_{\alpha_0 \in A_0} \left\{ \sum_{i,j=1}^{m-1} \frac{2}{\epsilon}(\bar{x}_j - \bar{y}_j)(\bar{x}_i - \bar{y}_i)q(\bar{t}, i, j, \phi(\bar{t}, i, i_0, \bar{x}), i_0, \alpha, \bar{x}) \right\} \]
\[
+ \sup_{\alpha_0 \in A_0} \left\{ \frac{1}{m-1} \sum_{i,j=1}^{m-1} 2 \epsilon |\bar{y}_j (x_j - \bar{y}_j) (q(\bar{t}, i, j, \phi(\bar{t}, i, i_0, \alpha_0, \bar{x}), i_0, \alpha_0, \bar{x}) - q(\bar{s}, i, j, \phi(\bar{s}, i, i_0, \bar{y}), i_0, \alpha_0, \bar{y})| \right\}
\]

\leq \sum_{i,j=1}^{m-1} \frac{2}{\epsilon} C |x_j - \bar{y}_j| |x_i - \bar{y}_i| + \sum_{j=1}^{m-1} \frac{2}{\epsilon} |x_j - \bar{y}_j| L (m - 1) (|\bar{t} - \bar{s}| + \|\bar{x} - \bar{y}\|),

where in the last inequality we used the Lipschitz property of \(q\) and \(\phi\) uniformly in \(\alpha\), as well as the boundedness of the function \(q\). It follows that \(I_2 \leq C_1^\epsilon (|\bar{t} - \bar{s}|^2 + \|\bar{x} - \bar{y}\|^2)\) and by (5.47b) we see that \(I_2 \to 0\) as \(\epsilon \to 0\). Finally, we deal with \(I_3\), which is defined by:

\[
I_3 := \sup_{\alpha_0 \in A_0} \left\{ (1 - \sum_{k=1}^{m-1} \bar{x}_k) \sum_{k=1}^{m-1} \frac{2}{\epsilon} (\bar{x}_k - \bar{y}_k) q(\bar{t}, m, k, \phi(\bar{t}, m, i_0, \bar{x}), i_0, \alpha_0, \bar{x}) \right. \\
- \left. (1 - \sum_{k=1}^{m-1} \bar{y}_k) \sum_{k=1}^{m-1} \frac{2}{\epsilon} (\bar{x}_k - \bar{y}_k) q(\bar{s}, m, k, \phi(\bar{s}, m, i_0, \bar{y}), i_0, \alpha_0, \bar{y}) \right\}.
\]

Using a similar estimation as for \(I_2\), we obtain \(I_3 \leq C_1^\epsilon (|\bar{t} - \bar{s}|^2 + \|\bar{x} - \bar{y}\|^2)\). Therefore by tending \(\epsilon\) to 0 in the inequality (5.51), we obtain a contradiction. This completes the proof.

### 5.8.3 Proof of Lemma 5.5.8

The main tool we use for the proof is the theory of limit operation on viscosity solutions. We refer the reader to Crandall et al. [1992] for an introductory presentation of limit operation on viscosity solutions to non-linear second order PDEs. Here we adapt the results established therein to the case of a coupled system of non-linear first order PDEs, which is the HJB equation we are interested in throughout the paper. Let \(\mathcal{O}\) be some locally compact subset of \(\mathbb{R}^d\) and \((F_n)_{n \geq 0}\) be a sequence of functions defined on \(\{1, \ldots, m\} \times \mathcal{O}\). We define the limit operator \(\limsup_*\) and \(\liminf_*\) as follow:

\[
\limsup_* F_n(i, x) := \lim_{n \to +\infty, \epsilon \to 0} \sup \{ F_k(i, y) | k \geq n, \| y - x \| \leq \epsilon \},
\]

\[
\liminf_* F_n(i, x) := \lim_{n \to +\infty, \epsilon \to 0} \inf \{ F_k(i, y) | k \geq n, \| y - x \| \leq \epsilon \}.
\]

194
Intuitively, we expect that the limit of a sequence of viscosity solutions solves the limiting PDE. It turns out that this is the proper definition of the limit operation, as is stated in the following lemma:

**Lemma 5.8.1.** Let $\mathcal{O}$ be a locally compact subset of $\mathbb{R}^d$, $(u_n)_{n \geq 0}$ be a sequence of continuous functions defined on $\{1, \ldots, m\} \times \mathcal{O}$ and $H_n$ be a sequence of functions defined on $[0, T] \times \{1, \ldots, m\} \times \mathcal{O} \times \mathbb{R}^m \times \mathbb{R}^d$, such that for each $n$, $u_n$ is a viscosity solution to the system of PDEs $H_n(t, i, x, u, \partial_t u(\cdot, i, \cdot), \nabla u(\cdot, i, \cdot)) = 0, \forall 1 \leq i \leq m$, in the sense of Definition 5.4.2. Assume that for each $i, x, d, p$, the mapping $u \to H(t, i, x, u, d, p)$ is non-decreasing in the $j$-th component of $u$, for all $j \neq i$. Then $u^* := \lim \sup_n u_n$ (resp. $u_* := \lim \inf_n u_n$) is a viscosity subsolution (resp. supersolution) to $H^*(t, i, x, u, \partial_t u(i, \cdot), \nabla u(i, \cdot)) = 0, \forall 1 \leq i \leq m$ (resp. $H_*(t, i, x, u, \partial_t u(\cdot, i, \cdot), \nabla u(\cdot, i, \cdot)) = 0, \forall 1 \leq i \leq m$).

The proof requires the definition of viscosity subsolution and supersolution based on the notion of first-order subjet and superjet (see Definition 2.2 in Crandall et al. [1992]). Let $u$ be a continuous function defined on $[0, T] \times \{1, \ldots, m\} \times \mathcal{O}$. We define the first order superjet of $u$ on $(t, i, x)$ to be:

$$J^+ u(t, i, x) := \{(d, p) \in \mathbb{R} \times \mathbb{R}^d : u(s, i, y) \leq u(t, i, x) + d(s - t) + (y - x)p + o(|t - s| + \|y - x\|), (s, y) \to (t, x)\}$$

Then it can be showed that $u$ is a viscosity subsolution (resp. supersolution) to the equation:

$$H(t, i, x, u, \partial_t u(i, \cdot), \nabla u(i, \cdot)) = 0,$$

if only if for all $(t, i, x)$ and $(d, p) \in J^+ u(t, i, x)$ (resp. $(d, p) \in J^- u(t, i, x)$), we have $H(t, i, x, u, d, p) \geq 0$ (resp. $H(t, i, x, u, d, p) \leq 0$). We now give the detail of the proof of Lemma 5.8.1.

**Proof.** Fix $(t, i, x) \in [0, T] \times \{1, \ldots, m\} \times \mathcal{O}$ and let $(d, p) \in J^+ u^*(t, i, x)$. We want to show that:

$$H^*(t, i, x, u^*(\cdot, x), d, p) \geq 0$$
where \( u^*(t, \cdot, x) \) represents the \( m \)-dimensional vector \([u^*(t, k, x)]_{1 \leq k \leq m}\). By Lemma 6.1 in Crandall et al. [1992], there exists a sequence \( n_j \to +\infty, x_j \in \mathcal{O}, (d_j, p_j) \in \mathcal{J}^+ u_{n_j}(t, i, x_j) \) such that:

\[
(t_j, x_j, u_{n_j}(t_j, i, x_j), d_j, p_j) \to (t, x, u^*(t, i, x), d, p), \quad j \to +\infty.
\]

Since \( u_n \) is viscosity subsolution to \( H_n = 0 \), we have:

\[
H_n(t_j, i, x_j, u_{n_j}(t_j, i, x_j), d_j, p_j) \geq 0.
\]

Now let us denote:

\[
S_{n, \epsilon}^u(t, k, x) := \sup \{ u_j(s, k, y), \max(|s - t|, \|y - x\|) \leq \epsilon, j \geq n \}.
\]

Then we have \( u^*(t, i, x) = \lim_{n \to +\infty, \epsilon \to 0} S_{n, \epsilon}^u(t, i, x) \). Let us fix \( \delta > 0 \). Then there exists \( \epsilon^0 > 0 \) and \( N^0 > 0 \) such that for all \( \epsilon \leq \epsilon^0 \) and \( j \geq N^0 \), we have:

\[
|S_{n, \epsilon}^u(t, k, x) - u^*(t, k, x)| \leq \delta, \quad \forall k \neq i.
\]

Moreover, there exists \( N > 0 \) such that for all \( j \geq N \),

\[
\|(t_j, x_j, u_{n_j}(t_j, i, x_j), d_j, p_j) - (t, x, u^*(t, i, x), d, p)\| \leq \delta \land \epsilon^0.
\]

Then for any \( j \geq N^0 \lor N \), we have \( \|(t_j, x_j) - (t, x)\| \leq \epsilon^0 \), and by definition of \( S_{n, \epsilon}^u \) we deduce that \( u_{n_j}(t_j, k, x_j) \leq S_{n, \epsilon}^u(t, k, x) \) for all \( k \neq i \). By the monotonicity property of \( H_{n_j} \), we have:

\[
H_{n_j}(t_j, i, x_j, S_{n_j, \epsilon^0}^u(t, 1, x), \ldots, S_{n_j, \epsilon^0}^u(t, i-1, x), u_{n_j}(t_j, i, x_j), S_{n_j, \epsilon^0}^u(t, i+1, x), \ldots, S_{n_j, \epsilon^0}^u(t, m, x), d_j, p_j) \geq 0.
\]

Now in the above inequality, all the arguments of \( H_{n_j} \) except \( i \) are located in a ball of radius \( \delta \) centered on the point \( (t, x, u^*(t, \cdot, x), d, p) \). Thus, we have:

\[
S_{n_j, \delta}^H(t, i, x, u^*(t, \cdot, x), d, p) \geq 0,
\]
where we have defined:

\[ S_{n,\delta}^H(t, i, x, u, d, p) := \sup\{H_j(s, i, y, v, e, q), \| (s, y, v, e, q) - (t, x, u, d, p) \| \leq \epsilon, j \geq n \}. \]

We have just proved that for any \( \delta > 0 \), there exists \( m > 0 \) such that for any \( j \geq m \), we have \( S_{n, \delta}^H(t, i, x, u^*(t, \cdot, x), d, p) \geq 0 \). Since we have:

\[ H^*(t, i, x, u^*(t, \cdot, x), d, p) = \lim_{n \to +\infty, \epsilon \to 0} S_{n, \delta}^H(t, i, x, u^*(t, \cdot, x), d, p), \]

we deduce that \( H^*(t, i, x, u^*(t, \cdot, x), d, p) \geq 0. \] \( \square \)

Now going back to the proof of Lemma 5.5.8, we consider a converging sequence of elements \( (\phi_n^0, \phi_n) \to (\phi_0^0, \phi) \) in \( \mathbb{K} \), where we have defined \( \mathbb{K} \) to be the collection of major and minor player’s controls \( (\phi_0^0, \phi) \) that are \( L \)-Lipschitz in \( (t, x) \). We denote \( V_n^0 \) (resp. \( V_0^* \)) the value function of the major player’s control problem associated with the controls \( \phi_n \) (resp. \( \phi \)). We also use the notation \( V_n^* := \lim sup V_n^0 \) and \( V_0^* := \lim inf V_0^n \). For all \( n \geq 0 \), we define the operator \( H_n^0 \):

\[
H_n^0(t, i, x, u, d, p) := d + \inf_{\alpha_0 \in \mathcal{A}_0} \left\{ f_0(t, i, \alpha_0, x) + \sum_{j_0 \neq i_0} (u_{j_0} - u_{i_0})q_0(t, i_0, j_0, \alpha_0, x) \right. \\
+ \left. (1 - \sum_{k=1}^m x_k) \sum_{k=1}^{m-1} p_k q(t, m, k, \phi_0(t, m, i_0, x), i_0, \alpha_0, x) \right. \\
+ \left. \sum_{k,l=1}^{m-1} p_{k,l} x_k q(t, k, l, \phi_0(t, k, i_0, x), i_0, \alpha_0, x) \right\}, \text{ if } t < T,
\]

\[
H_n^0(T, i, x, u, d, p) := g_0(i_0, x) - u_{i_0}.
\]

Then \( V_n^0 \) is viscosity solution to the equation \( H_n^0 = 0 \). It is clear to see that the operator \( H_n^0 \) satisfies the monotonicity condition in Lemma 5.8.1. To evaluate \( H_0^* := \lim sup^* H_0^n \) and \( H_0^* := \lim inf^* H_0^n \), we remark that for each \( 1 \leq i_0 \leq m_0 \), the sequence of functions \( (t, x, u, d, p) \to H_n(t, i_0, x, u, d, p) \) is equicontinuous. Indeed, this is due to the fact that the sequence \( \phi_n \) is equicontinuous and the function \( q \) is Lipschitz. Therefore \( H_0^* \) and \( H_0^* \) are...
simply the limit in the pointwise sense when \( t < T \). When \( t = T \), the boundary condition needs to be taken into account. The following computation is straightforward to verify:

**Lemma 5.8.2.** Define the operator \( H^0 \) as:

\[
H_0(t, i_0, x, u, d, p) := d + \inf_{\alpha_0 \in A_0} \left\{ f_0(t, i_0, \alpha_0, x) + \sum_{j_0 \neq i_0} (u_{j_0} - u_{i_0}) q_0(t, i_0, j_0, \alpha_0, x) \right. \\
+ (1 - \sum_{k=1}^{m} x_k) \sum_{k=1}^{m-1} p_k q(t, m, k, \phi(t, m, i_0, x), i_0, \alpha_0, x) \\
+ \sum_{k,l=1}^{m-1} p_l x_k q(t, k, l, \phi(t, k, i_0, x), i_0, \alpha_0, x) \left. \right\}, \quad \forall t \leq T.
\]

Then we have:

\[
H_0^*(t, i_0, x, u, d, p) = H_0^*(t, i_0, x, u, d, p) = H_0(t, i_0, x, u, d, p) \quad \text{if} \quad t < T, \\
H_0^*(T, i_0, x, u, d, p) = \max\{ (g_0(i_0, x) - u_{i_0}), H_0(T, i_0, x, u, d, p) \}, \\
H_0^*(T, i_0, x, u, d, p) = \min\{ (g_0(i_0, x) - u_{i_0}), H_0(T, i_0, x, u, d, p) \}.
\]

From the proof of Theorem 5.4.1, we see immediately that \( V_0 \) is a viscosity subsolution (resp. supersolution) of \( H_0^* = 0 \) (resp. \( H_0^* = 0 \)). By Lemma 5.8.2, \( V_0^* \) (resp. \( V_0^* \)) is a viscosity subsolution of \( H_0^* = 0 \) (resp. supersolution of \( H_0^* = 0 \)). Indeed, following exactly the proof of Theorem 5.4.4, we can show that a viscosity supersolution of \( H_0^* = 0 \) is greater than a viscosity subsolution of \( H_0^* = 0 \). By definition of \( \limsup^* \) and \( \liminf^* \), we have \( V_0^* \geq V_0^* \). It follows that \( V_0 \leq V_0^* \leq V_0^* \), and therefore we have \( \limsup^* V_0^* = \liminf^* V_0^* = V_0 \). Then we obtain the uniform convergence of \( V_0^* \) to \( V_0 \) following Remark 6.4 in Crandall et al. [1992].

### 5.8.4 Proof of Propositions 5.7.3 & 5.7.4

We use similar techniques as in Gomes et al. [2013] where the author provides gradient estimates for N-player games without major player. Let us first remark that the system of
ODEs (5.29) can be written in the following form:

\[-\dot{\theta}_m(t) = f_m(t) + \sum_{m' \neq m} a_{m'm}(t)(\theta_{m'} - \theta_m), \quad \theta_m(T) = b_m,\]

where we denote the index \( m := (i_0, x) \in \{1, \ldots, m_0\} \times \tilde{S}_N \) and we notice that \( a_{m'm} \geq 0 \) for all \( m' \neq m \). This can be further written in the compact form:

\[-\dot{\theta}(t) = f(t) + m(t) \theta, \quad \theta(T) = b, \quad (5.52)\]

where \( m \) is a matrix indexed by \( m \), with all off-diagonal entries being positive and the sum of every row equals 0. Define \( \| \cdot \| \) to be the uniform norm of a vector: \( \| b \| := \max_m |b_m| \).

Instead of proving (i) in Proposition 5.7.3, we prove a more general result, which is a consequence of Lemma 4 and Lemma 5 in Gomes et al. [2013].

**Lemma 5.8.3.** Let \( \phi \) be a solution to the ODE (5.52). Assuming \( f \) is bounded, we have:

\[\|\theta(t)\| \leq \int_t^T \| f(s) \| ds + \| b \|.\]

**Proof.** For any \( t \leq s \leq T \) we denote the matrix \( K(t, s) \) as the solution of the following system:

\[\frac{dK(t, s)}{dt} = -m(t)K(t, s), \quad K(s, s) = I,\]

where \( I \) stands for the \( m \) by \( m \) identity matrix. Now for \( \theta \), we clearly have:

\[-\dot{\theta}(s) = f(s) + m(s)\theta(s) \leq \| f(s) \| e + m(s)\theta(s),\]

where we denote \( e \) to be the vector with all the components equal to 1. Then using Lemma 5 in Gomes et al. [2013], we have:

\[-K(t, s)\dot{\theta}(s) \leq K(t, s)m(t)\theta(s) + \| f(s) \| K(t, s)e.\]
Note that \( K(t, s) \dot{\theta}(s) + K(t, s)m(t)\theta(s) = \frac{d}{ds}K(t, s)\theta(s) \). We integrate the above inequality between \( t \) and \( T \) to obtain:

\[
\theta(t) \leq K(t, T)b + \int_t^T \|f(s)\|K(t, s)e \, ds.
\]

Now using Lemma 4 in Gomes et al. [2013], we have \( \|K(t, T)b\| \leq \|b\| \) and \( \|K(t, T)e\| \leq \|e\| = 1 \). This implies that:

\[
\max_m \theta_m(t) \leq \|b\| + \int_t^T \|f(s)\|ds.
\]

Indeed starting from the inequality \( \dot{\theta}(s) \geq -\|f(s)\|e + m(s)\phi(s) \) and going through the same steps we obtain:

\[
\min_m \theta_m(t) \geq -\|b\| - \int_t^T \|f(s)\|ds.
\]

The desired inequality follows. \( \square \)

Now we turn to the proof of Proposition 5.7.3. Let \( \theta \) be the unique solution to the system of ODEs (5.29). Recall the notation \( e_{ij} := \mathbb{1}(j \neq m)e_j - \mathbb{1}(i \neq m)e_i \). For any \( k \neq l \), define \( z(t, i_0, x, k, l) := \theta(t, i_0, x + \frac{1}{N}e_{kl}) - \theta(t, i_0, x) \). Then \( z(t, \cdot, \cdot, \cdot, \cdot) \) can be viewed as a vector indexed by \( i_0, x, k \) and \( l \). Substracting the ODEs satisfied by \( \theta(t, i_0, x + \frac{1}{N}e_{kl}) \) and \( \theta(t, i_0, x) \), we obtain that \( z \) solves the following system of ODEs:

\[
\begin{aligned}
-\dot{z}(t, i_0, x, k, l) &= F(t, i_0, x, k, l) + \sum_{j_0 \neq i_0} (z(t, j_0, x, k, l) - z(t, i_0, x, k, l))q^0_0(t, i_0, j_0, x) \\
&+ \sum_{(i,j), j \neq i} (z(t, i_0, x + \frac{1}{N}e_{ij}, k, l) - z(t, i_0, x, k, l))N_{ij}q^0(t, i, i_0, x), \\
z(T, i_0, x, k, l) &= g_0(i_0, x + \frac{1}{N}e_{kl}) - g_0(i_0, x),
\end{aligned}
\]

(5.53)

where we have defined the \( F \) to be:

\[
F(t, i_0, x, k, l)
\]

200
\[ f^0(t, i_0, x) - f^0_0(t, i_0, x) \]
\[ + \sum_{j_0 \neq i_0} \left[ \theta(t, j_0, x + \frac{1}{N} e_{kl}) - \theta(t, i_0, x + \frac{1}{N} e_{kl}) \right] q^0_0(t, i_0, j_0, x) \]
\[ + \sum_{i,j} \left[ \theta(t, i_0, x + \frac{1}{N} e_{ij} + \frac{1}{N} e_{kl}) - \theta(t, i_0, x + \frac{1}{N} e_{kl}) \right] \]
\[ \times N[(x + \frac{1}{N} e_{kl})_i q(t, i, j, i_0, x + \frac{1}{N} e_{kl}) - x_i q(t, i, j, i_0, x)] \].

Then by the uniform bound provided in Lemma 5.8.3, together with the Lipschitz property of \( f_0, g_0, q_0, q, \alpha, \) and \( \beta \), we have:

\[ \| F(t) \| \leq \frac{L}{N} + 2(T \| f_0 \|_\infty + \| g_0 \|_\infty) \frac{L}{N} + \| z(t) \| \cdot L, \quad \| z(T) \| \leq \frac{L}{N}, \]

where \( L \) is a generic Lipschitz constant. Now we are ready to apply Lemma 5.8.3 to (5.53).

We have for all \( t \leq T \):

\[ \| z(t) \| \leq \frac{L}{N} + \int_t^T \left( \frac{C}{N} + L \| z(s) \| \right) ds, \]

where \( C \) is a constant only depending on \( T, L, \| f_0 \|_\infty, \| g_0 \|_\infty \). Finally Gronwall’s inequality allows us to conclude.

The proof of Proposition 5.7.4 is similar. We consider the solution \( \theta \) to the ODE (5.30) and we keep the notation \( z(t, i_0, x, k, l) \) as before. For \( v \in \mathbb{R}^{m_0}, x \in \bar{S}, t \leq T \) and \( i_0 = 1, \ldots, m_0 \), denote:

\[ h_0(t, i_0, x, v) := \inf_{\alpha_0 \in A_0} \{ f_0(t, \alpha_0, i_0, x) + \sum_{j_0 \neq i_0} (v_{j_0} - v_{i_0}) q_0(t, i_0, j_0, \alpha_0, x) \}. \]

Then for all \( x, y \in \bar{S}, u, v \in \mathbb{R}^{m_0} \), using the Lipschitz property of \( f_0 \) and \( q_0 \) and the boundedness of \( q_0 \), we have:

\[ |h_0(t, i_0, x, v) - h_0(t, i_0, y, u)| \leq L|x-y| + 2(m_0-1) \max\{\|u\|, \|v\|\} |x-y| + C^2(m_0-1) \|v-u\|. \]

(5.54)
Subtracting the ODEs satisfied by \( \theta(t, i_0, x + \frac{1}{N} e_{kl}) \) and \( \theta(t, i_0, x) \), we obtain that \( z \) solves the following system of ODEs:

\[
-\dot{z}(t, i_0, x, k, l) = F(t, i_0, x, k, l) + \sum_{(i,j), j \neq i} (z(t, i_0, x + \frac{1}{N} e_{ij}, k, l) - z(t, i_0, x, k, l)) N x_i q^\phi(t, i, j, i_0, x),
\]

\[
z(T, i_0, x, k, l) = g_0(i_0, x + \frac{1}{N} e_{kl}) - g_0(i_0, x),
\]

where \( F \) is given by:

\[
F(t, i_0, x, k, l) := h_0(t, i_0, x + \frac{1}{N} e_{kl}, \theta(t, \cdot, x + \frac{1}{N} e_{kl})) - h_0(t, i_0, x, \theta(t, \cdot, x))
\]

\[
+ \sum_{(i,j), j \neq i} [\theta(t, i_0, x + \frac{1}{N} e_{ij} + \frac{1}{N} e_{kl}) - \theta(t, i_0, x + \frac{1}{N} e_{kl})] \times N[(x + \frac{1}{N} e_{kl}), q^\phi(t, i, j, i_0, x + \frac{1}{N} e_{kl})] - x_i q^\phi(t, i, j, i_0, x)].
\]

By the estimation in equation (5.54), the uniform bound provided in Lemma 5.8.3, together with the Lipschitz property of \( f_0, g_0, q_0, \phi_0, \) and \( \phi \), we have:

\[
\|F(t)\| \leq \frac{1}{N} [L + 2(m_0 - 1)(T\|f_0\|_\infty + \|g_0\|_\infty)] + 2C_q(m_0 - 1)\|z(t)\| + m(m - 1)(L_\phi L + C_q)\|z(t)\|.
\]

We apply Lemma 5.8.3 to obtain:

\[
\|z(t)\| \leq \frac{L}{N} + \int_t^T [m(m - 1)(L_\phi L + C_q) + 2C_q(m_0 - 1)]\|z(s)\|ds
\]

\[
+ \int_t^T \frac{1}{N}[L + 2(m_0 - 1)(T\|f_0\|_\infty + \|g_0\|_\infty)]ds
\]

\[
:= C_0 + C_1 T + C_2 T^2 + \int_t^T (C_3 + C_4 L_\phi)\|z(s)\|ds.
\]

We then use Gronwall’s inequality to conclude.
Chapter 6

Conclusion

In this final chapter, we provide a few directions for future research on finite state mean field games. We would like to warn the reader that these are some very preliminary ideas, and their implementations are potentially prone to insurmountable technical difficulties.

From Finite State Space to Countable State Space?

In the mean field game models we have presented in this thesis so far, we have assumed that the states of the players belong to a finite space. The first natural and meaningful extension we can consider is to formulate the games in an infinite and countable space. This is certainly the case in the information percolation model in Duffie et al. [2009] (see Section 1.2 for a detailed presentation of the model). If we use the analytical approach to study the Nash equilibria, we should still be able to derive the Hamilton-Jacobi-Bellman equation characterizing the optimality of the control and the Kolmogorov equation characterizing the flow of measures for players’ states. However, since we now have a coupled system of countably many ODEs, we need to reconsider the topological argument leading to the existence of the solutions. On the other hand, most of the arguments in the probabilistic approach will remain valid in the case of countable state space, provided that we can take care of issues regarding the integrability of the state process. This means that we need to carefully consider the assumptions on the transition rate functions.
If we managed to extend the mean field game to any discrete space, not only will we be able to handle a wider range of applications, we will be able to develop new numerical methods for diffusion-based mean field game. It is a well known numerical algorithm in stochastic optimal control of diffusion that we can approximate the Wiener processes by random walks, or use the quantification to approximate the diffusion process by a stochastic process living on a discrete grid. Therefore, one interesting question to ask here is whether the equilibrium of a diffusion-based mean field game can be approximated by that of a discrete state mean field game where the state process is the corresponding discrete approximation. If the answer is yes, can we also have the convergence of the master equations of the approximate finite state mean field games to the master equation of the stochastic differential mean field game?

**Dynamic Programming Principle for Finite State Principal Agent Problems?**

In Chapter 4, we have shown that the principal agent problem in a finite state space can be formulated as a McKean-Vlasov optimal control problem (see equation (4.28) and Theorem 4.2.5). We have also shown in Section 4.3 that with the additional assumptions of linearity of the transition rate, convex cost function and linear utility function of the minor players, the optimal control problem is Markovian and can be reduced to a deterministic control problem on the Kolmogorov equation. A natural question to ask is if the Markovian structure is still preserved with a more general set of assumptions. Another related question is: can we use the dynamic programming principle to tackle the optimization problem (4.28)? Indeed, dynamic programming principle for McKean-Vlasov control problems was developed for the diffusion controlled processes in Pham and Wei [2017]. However, the authors dealt with the optimal control of the strong solution of SDE in which the probability measure and the Wiener process are both fixed, whereas the situation we deal with here is an optimal control problem in the weak formulation, meaning that the control impacts both the probability measure and the martingale driving the process. If we can establish a specific version of dynamic programming principle applicable to the control problem (4.28), we can derive the HJB equation and compute the optimal contract in a more general setup.
Probabilistic Formulation for Finite State Mean Field Games with Major and Minor Players?

In Chapter 5, we applied the analytical method to study the finite state mean field game with major and minor players. In particular, we described the dynamics of the states via the infinitesimal generators and we tackled the optimization problems of the players by applying the dynamic programming principle. This allowed us to obtain the master equation characterizing the Nash equilibria. However, in order to obtain a well-defined dynamics for the players’ dynamics in the mean field limit, we had to make some strong assumptions on the information structure of the players. As a result, we were only able to establish the existence of Nash equilibria for a certain set of Markovian strategies and a small duration of the game. Indeed, as pointed out in Carmona and Wang [2017], for mean field games with major and minor players, different assumptions on the information structure and the admissible strategies lead to different Nash equilibria. This makes us wonder: can we improve the results of the existence of Nash equilibria by switching to a less restrictive information structure? One possible angle to attack the problem is to extend the weak formulation in Chapter 3 to the mean field game with a major player. Another possible way, as suggested by Cecchin and Fischer [2017], is to use SDEs driven by Poisson random measures to model the state processes of the players. No matter which approach we take, the success of these attempts hinges on whether we will be able to obtain a well-defined and tractable dynamics for the state processes of the players in the mean field game limit.
Bibliography


Josu Doncel, Nicolas Gast, and Bruno Gaujal. Mean-field games with explicit interactions. 2016.


