ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE
ENGINEERING AND TECHNOLOGY

MULTILAYER MEAN FIELD DIFFERENTIAL GAMES IN MULTI-AGENT SYSTEMS AND AN APPLICATION IN INTELLIGENT TRANSPORTATION

Ph.D. THESIS

Alper ÖNER

Department of Control and Automation Engineering

Control and Automation Engineering Doctorate Programme

JUNE 2018

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL OF SCIENCE
ENGINEERING AND TECHNOLOGY

MULTILAYER MEAN FIELD DIFFERENTIAL GAMES IN MULTI-AGENT SYSTEMS AND AN APPLICATION IN INTELLIGENT TRANSPORTATION

Ph.D. THESIS

Alper ÖNER
(504112101)

Department of Control and Automation Engineering

Control and Automation Engineering Doctorate Programme

Thesis Advisor: Assoc. Prof. Dr. Gülay Öke GÜNEL

JUNE 2018

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

ÇOKLU-KARAR VERİCİLİ SİSTEMLERDE ÇOKLU DÜZLEM ORTALAMA ALAN DİFERANSİYEL OYUNLARI VE AKILLI ULAŞIMDA BİR UYGULAMA

DOKTORA TEZİ

Alper ÖNER
(504112101)

Kontrol ve Otomasyon Mühendisliği Anabilim Dalı

Kontrol ve Otomasyon Mühendisliği Doktora Programı

Tez Danışmanı: Assoc. Prof. Dr. Gülay Öke GÜNEL

HAZİRAN 2018

Alper ÖNER, a Ph.D. student of ITU Graduate School of Science Engineering and Technology 504112101 successfully defended the thesis entitled "MULTILAYER MEAN FIELD DIFFERENTIAL GAMES IN MULTI-AGENT SYSTEMS AND AN APPLICATION IN INTELLIGENT TRANSPORTATION ", which he/she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Assoc. Prof. Dr. Gülay Öke GÜNEL** Istanbul Technical University	....
Jury Members :	**Prof. Dr. Hakan TEMELTAŞ** Istanbul Technical University	....
	Prof. Dr. Mehmet Turan SÖYLEMEZ Istanbul Technical University	....
	Assoc. Prof. Dr. Gülay Öke GÜNEL Istanbul Technical University	....
	Assist. Prof. Dr. Dilek TÜKEL Doğuş University	....
	Assist. Prof. Dr. Figen ÖZEN Haliç University	....

Date of Submission : 10 May 2018
Date of Defense : 20 June 2018

To Duygu and Deniz ÖNER,

FOREWORD

This thesis has emerged as a valuable work of an international collaboration. Most of the programming, preliminary works and ideas are developed in Istanbul Technical University, Turkey. Afterwards most of the theoretical works have been completed in Coordinated Science Lab. of University of Illinois at Urbana Champaign, USA. Therefore, firstly I want to thank to Prof. Tamer BAŞAR from UIUC. He always encouraged me for finding new ideas, and then he showed me right directions in my researches. Furthermore, postdoctoral researcher Dr. Naci SALDI had contributed to my mean field researches enormously via valuable countless discussions in UIUC. I could not continue my thesis without TUBITAK 2214-A International Research Scholarship for Ph.D. students and 2211-A Scholarship for the Ph.D. degree in USA and Turkey, respectively so I would like to thank TUBITAK for their valuable supports. Moreover, I also want to thank my advisor Assoc. Prof. Gülay Öke GÜNEL and my thesis committee members Prof. Hakan TEMELTAŞ and Assist. Prof. Dilek TÜKEL. They always gave me significant recommendations on the purpose of improving my researches. Prof. Ata MUĞAN, Prof. Metin GÖKAŞAN and Assist. Prof. Volkan SEZER provided me a peaceful and consciousness expanding working environment during my master and PhD education in Mechatronics Education and Research Center of ITU. Besides I want to thank my parents Dr. Adil ÖNER, Prof. Güneş Pernur ÖNER, who had always supported me financially and morally during my education life, and my sisters Prof. Yıldız Öner İYİDOĞAN and İdil ÖNER. Finally, my beloved wife Duygu ÖNER and son Deniz ÖNER, I could not achieve any thing without their supports and love.

Mayıs 2018

Alper ÖNER
(MSc. Mechatronic Engineer)

TABLE OF CONTENTS

Page

ABBREVIATIONS

App	: Appendix
a.s.	: almost surely
UIUC	: University of Illinois at Urbana Champaign
ITU	: Istanbul Technical University
TUBITAK	: Türkiye Bilimsel ve Teknolojik Araştırma Kurumu
MF	: Mean Field
MFG	: Mean Field Game
MSG	: Multi-level Stackelberg Game
LQ	: Linear Quadratic
SDE	: Stochastic Differential Equation
FBSDE	: Forward-Backward Stochastic Differential Equation
BSDE	: Backward Stochastic Differential Equation
SOC	: Stochastic Optimal Control
HJB	: Hamilton-Jacobi-Bellman
ITS	: Intelligent Transportation System
MF-MSG	: Mean Field Multilayer Stackelberg Game
MF-ITS	: Mean Field Intelligent Transportation System
ITUCiTSim	: Istanbul Technical University City Transportation Simulation
GGs	: Greenhouse Gases
AI	: Artificial Intelligence

SYMBOLS

$X > 0$	: A positive definite matrix
$X \geq 0$	: A positive semi-definite matrix
$\mathscr{R}^n$	: n-dimensional real-valued vectors
$\mathscr{R}^{m\times n}$	: $m \times n$-dimensional real-valued matrices
x^T	: The transpose of x
X^T	: The transpose of a matrix X
$\|x\|_S^2$	: $x^T Sx$, where $x \in \mathscr{R}^n$ and $S \geq 0$
$\|X\|$	: The induced 2-norm for $X \in \mathscr{R}^{n\times n}$
I_n	: The n $\times$ n-dimensional identity matrix
1_n	: The n-dimensional column vector whose elements are all 1
$0_{n\times m}$	: The n $\times$ m-dimensional zero matrix
0_n	: $0_{n\times 1}$
u_{-i}	: $\{u_1,\ldots,u_{i-1},u_{i+1},\ldots,u_N\}$
Ω, $\mathscr{F}_t, t \geq 0$	: Natural complete filtered probability space augmented by all the $\mathscr{P}$ null sets in $\mathscr{F}$

LIST OF TABLES

Page

LIST OF FIGURES

MULTILAYER MEAN FIELD DIFFERENTIAL GAMES IN MULTI-AGENT SYSTEMS AND AN APPLICATION IN INTELLIGENT TRANSPORTATION

SUMMARY

In this thesis, we describe the results we have obtained on linear-quadratic hierarchical Mean Field Stackelberg differential games with open-loop information structure. We name this new class of games as Mean Field (MF) Multilayer Stackelberg Games (MSG). The model that we have developed consists of three levels of decision making. There is a leader at the top, sub-leaders are at the intermediate level, and followers are at the lowest level. In this structure, Stackelberg Game is played between the leader, sub-leaders and followers in turn. Firstly the leader plays with the sub-leaders and then sub-leaders play with their followers. Consequently, the leader can control the followers not directly, but only through the sub-leaders. The followers are weakly coupled through a Mean Field term, affecting their individual costs. The main novelty in this thesis is to extend the Mean Field Game to multi layer framework. Theoretically, Mean Field Equilibrium arises in the infinite population limit. In our approach, instead of computing the exact Stackelberg equilibrium, the Mean Field equilibrium policies are obtained with the assumption of infinite population. When the population size is sufficiently large, the resulting policy can be regarded as an approximate Stackelberg equilibrium by the law of large numbers, therefore we apply this resulting policy to the original problem in the finite population case. The performance of the developed method is evaluated by a numerical example.

In the second part of the thesis, we adapt and apply the developed methodolgy to Intelligent Transportation Systems (ITS) for future Smart Cities. So, another important novelty of the thesis is to apply The Mean Field Game Theory to Automated Highway System (AHS), in order to minimize both total consumed energy and travel time. In this model, there is a Control Center which acts as a leader, Road Links in the intermediate level act as sub-leaders and finally Vehicles are followers. The Control Center, Road Links and Vehicles which are in consecutive layers play Mean Field (MF) Multilayer Stackelberg Games (MSG). Although, the Control Center cannot manipulate the Vehicles directly, it imposes its strategy through the Road Links in the intermediary layer. The Vehicles of each Road Link are weakly coupled through a Mean Field term, which affects their individual cost functions. In order to implement this algorithm to ITS, we have developed our highly realistic simulation environment, Istanbul Technical University City Transportation Simulation (ITUCiTSim) program. The effectiveness of the proposed method is justified by simulations performed in ITUCiTSim.

ÇOKLU-KARAR VERİCİLİ SİSTEMLERDE ÇOKLU DÜZLEM ORTALAMA ALAN DİFERANSİYEL OYUNLARI VE AKILLI ULAŞIMDA BİR UYGULAMA

ÖZET

Son yıllarda çoklu karar vericili sistemlerin kontrol uygulamaları hızla artmaktadır. Teknolojinin gelişmesi ile insanlık için bir çok uygulama alanı ortaya çıkmaktadır. Bilgisayarlardaki hesaplama kapasitelerinin her geçen gün artması, geçmişte hesaplama zorluğu olan çalışma alanında daha fazla çözüm üretilmesine olanak sağlamaktadır. Dünyadaki ekolojik düzenin hızla bozulduğu günümüzde akıllı şehirlerin, insanlığa getirdiği optimum çözümler ile ekolojik bozulmanın azalmasını sağlamaktadır. Bir çok parametrenin aynı anda değerlendirilmesi gereken uygulamalarda çoklu karar verici sistemlere şiddetle ihtiyaç vardır.

Akıllı enerji dağıtım sistemlerinde aynı anda bir çok şebekeye dağıtılması gereken enerjinin optimizasyonu amaçlanmaktadır. Enerji kaynaklarından yüzbinlerce kullanıcıya enerji kayıplarını minimuma indirecek şekilde bir dağıtım yapılması bir çoklu karar verici problemidir. Enerji problemini ele alırsak, günümüzde doğayı korumak için kullanılması hayati olan rüzgar ve güneş enerjisi gibi sistemler gerekmektedir. Bireysel kullanıcıların ürettikleri ihtiyaç fazlası enerjiyi şebekeye verirken aynı optimizasyonun gerçekleşmesi ve bu enerjinin değişken dinamik koşullar altında fiyatlanması da bir çoklu karar verici problemidir. Aynı şekilde son yıllarda çalışmaları hızla artan elektrikli araçların şarj istasyonlarına bağlanması durumunda ortaya çıkacan sistemin benzer şekilde bir çoklu kontrol problemi olarak çözülmesi gerekmektedir.

Çoklu karar vericili sistemlerin akıllı şehirlerde sıklıkla kullanılması gereken bir başka alan da akıllı ulaştırma sistemleridir. Şehir içindeki yollarda; bir akıllı kavşakta aynı anda sistemin göz önüne alması gereken yayalar, araçlar, kurallar ve ortam koşulları gibi bir çok parametre vardır. Daha büyük akıllı otobanlarda ise aynı anda bir çok aracın değerlendirilmesi gerektiği gibi zaman ve enerji ile ilgili bir çok parametre de göz önüne alınmalıdır. Karayollarının yanı sıra havalimanında uçakların zaman ve yakıt optimizasyonu yaparak koordine edilmesi bir çok araştırmacı tarafından son yıllarda fazlalıkla çalışılan bir çoklu karar verici sistem problemidir. Benzer problem deniz limanlarında da mevcuttur.

Akıllı sehirlerde insanları optimum şekilde gidecekleri yere ulaştıracak problemlerin çözümü dünyada en çok çalışılan konulardan biridir. Evinden işine gitmeyi amaçlayan yüzbinlerce kişinin yaşadığı bir ortamda karar vericileri enerji, zaman, yoğunluk veya farklı önceliklere göre mantıklı bir dağıtım yapmak önümüzdeki yıllarda çözülmesi gereken en önemli sorunlardan biri olacaktır.

Çoklu karar verici problemlerinin günlük hayatta doğrudan insanlığın karşısına çıktığı anlarda bu optimizasyonun yapılabilmesi için en önemli gereklilik veridir. Yüzbinlerce karar vericinin etkilendiği sistemlerde problemin hızla çözülmesi gerekirken kullanıcılardan verilerin uygun bir biçimde toplanması, daha sonra elde

edilen sonucun aynı şekilde zamanında dağıtılması önümüzdeki yıllarda çözülmesi gereken en önemli çoklu karar vericili sistem problemidir.

İnternet ağlarında büyük verinin toplanması, filtrelendirilmesi, sınıflandırılması ve anlamlandırılması temelde çözülmesi gereken en önemli çoklu karar verici problemlerinden biridir. Veri hızının fiziksel imkanlar ile her geçen yıl arttığı günümüzde matematiksel yöntemler ile bu fiziksel artış desteklenmeli, bu sayede optimum çözümler veri karmaşıklığına, anlamlandırılmasına ve hızına yeni çözümler üretilmelidir. Bu tezin öncelikli amaçlarından biri de bu veri karmaşıklığını azaltacak bir optimum çoklu kontrol yöntemi önermektir.

Önerilen çoklu kontrol yöntemi Ortalama Alan Oyun temelli bir oyun teorisidir. Oyun Teorisi akıllı karar vericilerin en uygun kararı vermesini sağlayan matematiksel modelleri oluşturan bir optimizasyon yöntemidir. Oyun Teorisi; ekonomi, politik bilimler, psikoloji, bilgisayar bilimleri ve biyoloji alanlarında sıklıkla kullanılmaktadır. Temel olarak bir oyuncunun kazandığı kadar diğer oyuncunun kaybettiği sıfır-toplam oyunundan geliştirilmiştir. Modern oyun teorisi John von Neumann'ın iki oyunculu sıfır-toplam karışık stratejili oyunun çözümünün var oluşunu ispatından sonra gelişmiştir. Von Neumann'ın sürekli kompakt konveks kümelerde Brouwer sabit-nokta (fixed-point) teorisini kullanarak gelistirdiği ispatı oyun teorisi ve ekonomi matematiği alanında standart bir yöntem olmuştur.

Oyun teorisinde, oyunlar eğer oyuncular arasında maliyeti düşürmek için bir yardımlaşma varsa iş birlikçi oyun; oyuncular arasında yardımlaşma yoksa ve hepsi kendi maliyetlerini göz önünde bulunduruyorsa rekabetçi oyun türü olarak adlandırılırlar. Nash Dengesi ise iki ya da daha çok oyuncunun bulunduğu rekabetçi oyunlarda her oyuncunun diğer rakiplerinin stratejilerini göz önünde bulundurarak kendi stratejilerini hesapladığı durumda, strateji değiştirmenin hiç bir oyuncuya daha fazla kazanç kazandırmadığı denge durumudur. Yani eğer her bir oyuncu bir strateji belirlediği durumda hiç bir oyuncu stratejisini değiştirerek herhangi bir kazanç saylayamadığı için diğer oyuncuların kararları sabit kalıyor ise Nash Dengesi sağlanmış demektir. Nash dengesi oyun teorisi içerisindeki en temel kavramlardan biridir. Nash dengesinin gerçekliği uygulamalı ekonomi yöntemleri test edilerek görülebilir.

Çoklu kontrol problemleri temel olarak iki şekilde ele alınır. Tek bir merkezde kontrol algoritmasının geliştirildiği, merkezi kontrol yöntemi bunlardan ilkidir. Merkezi kontrolde çok sayıda karar verici olması durumunda her bir karar vericinin verilerinin toplanması ciddi bir zaman problemine yok açmaktadır. Ayrıca çok sayıda toplanan verinin bir kontrol stretejisine dönüştürülmesi hesap karmaşıklığı büyük olan bir problemin çözümünü gerektirmektedir. Bu problemi çözmek için son yıllarda dağıtılmış kontrol yöntemi geliştirilmiştir. Dağıtılmış kontrol yönteminde bütün verilerin tek bir merkezde toplanmasına gerek yoktur. Her bu karar verici kendi verilerini göz önüne alarak merkezden ayrı kontrol stratejisini geliştirir. Çoklu kontrol yöntemi olarak Nash denklemi kullanılmak istendiğinde çoklu karar vericilerin Nash Dengesine ulaşması için birbirlerinin en doğru stratejilerini bilmeleri gerekmektedir. Bu sebeple çoklu karar vericili bir sistemde karar verici sayısı arttıkça Nash Dengesine ulaşmak hesaplanamayacak kadar karmaşık bir boyuta ulaşmaktadır.

Nash dengesinin çoklu karar vericilerde hesaplanmasının zorluklarından dolayı mantıklı bir yaklaşık Nash Oyunu geliştirme ihtiyacı ortaya çıkmıştır. Bu yaklaşık Nash Oyununa; Ortalama Alan Oyun yöntemi geliştirilerek bir çözüm getirilmiştir.

Ortalama Alan Oyun Yönteminde her bir karar verici ortalama kütle değeri tarafından etkilenmektedir. Bu etkiyi Ortalama Değer Terimi temsil etmektedir. Karar vericiler Nash Oyununu bu ortalama alan terimini kullanarak oynarlar. Diğer bir deyişle her karar verici ortalama alan değeri sayesinde birbirine zayıf olarak bağlıdır. Karar verici sayısı sonsuza gittiği zaman karar vericilerin durumları deneysel olarak belirli bir limite yakınsar. Bu sebeple yeterli sayıda fazla karar verici olduğu durumlarda optimal kontrol probleminde kullanılan ortalama alan değeri dışarıdan girilen bir değer olarak değerlendirilir. Bu şekilde bir yaklaşımla çoklu karar vericili dağıtılmış karmaşıklığı yüksek Nash Oyunu, tek karar vericili bir optimal kontrol problemine dönüştürülebilir. Bu optimal kontrol problemi ortalama alan değeri ile tutarlıdır ve problemin çözümü yaklaşık olarak Nash Oyunun çözümüne eşittir.

Optimal kontrol probleminin sonucu, yaklaşık olarak Nash oyunu sonucuna yakın olmasına rağmen ortalama alan oyunun çözümü matematiksel olarak farklı denklemlerin çözülmesini gerektirir. Tek karar vericili optimal kontrol probleminden farklı olarak ilerleyen zamanda Fokker-Planck kısmı diferensiyal denkleminin çözümü ve geri zamanda ise Hamilton-Jacobi-Bellman kısmı diferensiyal denkleminin çözümü gerekmektedir. Bu tip denklemlere ilişkili ileri-geri kısmı diferansiyel denklemler denmektedir.

Literatürde çalışılan Ortalama Alan Oyunları sıklıkla rekabetçi Nash türünden oyunlardır. Bu çalışmalarda çoklu hiyerarşik katmanlar kullanılmamıştır. Bizim bu tezde yaptığımız yeniliklerden biri de çoklu hiyerarşik bir yapıda Stackelberg Ortalama Alan Oyunu geliştirmektir. Stackelberg Oyununun ismi Heinrich Freiherr von Stackelberg'den gelmektedir. Oyun teorisinde, Stackelberg Oyunu hem lider hem de takip edici içeren bir oyundur. Lider bazı durumlarda pazar lideri olarak da nitelendirilebilir. Stackelberg oyununda lider takipçilere kendi stratejisini ilan eder ve takipçiler kendi kazançlarını liderin stratejisine göre optimize ederler. Lider de kendi stratejisini belirlerken takipçilerin kendi stratejisini göz önüne alacağını da hesaba katarak oyunu oynar ve karını maksimize eder. Lider takipçileri hakkında her veriyi bilirken, takipçilerin böyle bir bilgisi yoktur.

Bu tezde biz öncelikle açık çevrim lineer-ikinci dereceden (quadratic) çoklu hiyerarşik katmanlı ortalama alan Stackelberg oyunu geliştirip, daha sonra yeni teorik modeli, C++ ile geliştirdiğimiz simülasyon programında Akıllı Ulaşım Sistemlerine uyguladık. Geliştirdiğimiz model üç katmandan oluşmaktadır. En üst katman sistemin genel lideridir. Bu akıllı ulaşım sisteminde kontrol merkezi olarak nitelendirilebilir. Genel lider alt katmanlar ile Stackelberg Oyunu oynayarak kendi stratejisini sisteme empoze eder. Ara katman alt lider ya da akıllı ulaşım sistemi için yol bölümleri olarak düşünülebilir. En alt katmandaki takipçiler, akıllı ulaşım uygulamasında araçlar, ara katman ile Stakelberg Ortalama Alan Oyunu oynar. Bu şekilde genel lider en alt katmandaki takipçileri doğrudan kontrol edemez, ara katman sayesinde dolaylı olarak kontrol edebilir. Her takipçi ve ara lider sistemin genel ortalama alan değeri ile birbirleriyle ilişkili iken her ara liderin kontrol ettiği gruplar da kendi ortalama alan değerleri ile takipçilerini ilişkilendirir.

Bu yeni lineer-ikinci dereceden çoklu hiyerarşik katmanlı ortalama alan Stackelberg oyunu algoritması sayesinde genel lider sistemi ara liderlerden farklı maliyet koşullarında optimize edebilmektedir. Ara liderlerin farklı çevre koşullarından etkilendiği problemlerde ortalama alan oyununu katmanlara bölüp ara liderler kullanmak zor çevre koşullarına adaptasyonu sağlamaktadır. Bu yöntemin akıllı ulaşım sistemlerinde kullanılması ile gelecekte insanlık için çok önemli olan akıllı şehirlerde, trafik akışı optimize edilerek karbon salınımının azalması sağlayacaktır.

1. INTRODUCTION

Control and coordination of multi-agent systems is an important and challenging problem, and it has many diverse application areas such as mobile robotics [1], vehicle formation [2], flocking [3], consensus problems [4], and micro-economics [5], among others. In the literature, a significant amount of effort has been devoted to the theory of multi-agent systems in order to characterize the optimal decision rules. In these models, it is not feasible that a single agent has access to information available to all the other agents as in centralized control. Hence importance of decentralized decision, where each agent applies a strategy using only its local information, naturally arises.

In decentralized game problems, interaction between all agents should be incorporated into the solution process in order to obtain Nash equlibria, which is the optimality criterion used in noncooperative game theory [6–8]. However, solving Nash equlibria with a large number of agents creates serious complexity issues, and therefore, has to be abandoned. It is, therefore, reasonable to search for an approximate solution to the Nash equilibrium. Mean field game framework [9, 10] has been introduced recently in an attempt to overcome the difficulties that arise in solving for Nash equilibria of the multi-agent differential game problems. In the mean field model, each individual agent is affected by the average mass behavior of all agents which, is called the *Mean Field Term*. In other words,the mean field term couples each agents weakly. The law of large numbers dictates that a deterministic limit is converged by the empirical distribution of agents' state (i.e., mean field term), when the population size goes to infinity. Hence, for large enough population sizes, we can view the mean field term as an external variable to the optimal control problem that agents are faced with. Therefore, this way, we can transform the decentralized multi-agent game problem to a single agent decision problem under the condition that the state evaluation of an individual agent should be consistent with the mean field term, which is a result of the law of large numbers. This condition is called Nash certainty equivalence problem in technical literature.

In this thesis, we describe the results we have obtained on linear-quadratic hierarchical Mean Field Stackelberg differential games with open-loop information structure. In this game model, there are three levels. In the first level, there is a global leader which plays a Stackelberg game with sub-leaders which are in the second level and in the third level, there are followers which play a Stackelberg Mean Field game with their sub-leaders. Hence, the leader cannot manipulate the followers directly, however, sub-leaders link up followers to the leader as an intermediary layer. A *Mean Field* term couples followers of each sub-leader weakly. The main novelty in this thesis is to consider the Mean Field Game in accordance with a hierarchical multilayer structure that arises in the infinite population limit and compute the Mean Field equilibrium policies instead of computing exact Stackelberg equilibrium. Then the resulting policies are applied to the original problem. In case the population size is sufficiently large, the resulting policies will yield an approximate Stackelberg equilibrium, by the law of large numbers. A similar problem which combines *Stackelberg* game and the *Mean Field* approach was first searched in [11]. However, the model considered in [11] constitutes of only two levels which significantly simplifies the problem.

Researchers examined the leader-follower mean field games in the past decade but they usually solved the problem through Nash Games. Moon and Basar studied Stackelberg game in Mean Field Games [12] recently. The contribution of this thesis to Stackelberg differential mean field games is that the problem is solved in view of hierarchy layers, a sub-layer is introduced to the game and multi leaders are added to this sub-layer. The hierarchy of sub-leaders which play games on behalf of their followers makes it possible to manage the followers in case of different sub goals. In [13], Stackelberg equilibrium is solved from a multi layer point of view. In this thesis, *Stackelberg Mean Field* approach of [11] and *Stackelberg Multi Layer* solution of [13] are combined.

After developing the new type of multi agent control model as described above, it is applied to the Intelligent Transportation Systems in Smart City, which is an envisaged concept that is used for vast scope of application areas. There is no consistent definition for Smart City among academia but it simply stands for a green, safe and friendly city with improved life conditions for its inhabitants. These conditions are strongly related to economic, social, and environmental standards of the city in terms of city services for citizens, transportation, communication,health care systems, water, business, and energy infrastructure. In recent decades, the world population has increased significantly and by the year 2050 it is expected that people of the world will prevalently live in cities with a ratio of %70. Due to the fact that cities are currently generating %80 of greenhouse gases, also consume %75 of the world's resources and energy, a significant amount of research has to be performed to preserve world resources.

The development of ITS is a cornerstone in the design and implementation of smart cities. Traffic congestion is a common problem all around the world. Governments have to take precautions to assure efficient transportation systems. France, Belgium, the UK and Netherlands have incorporated in a major project called as Connecting Europe Facility, which aims to promote smart infrastructure using sensors; roads that generate energy; real-time visual insights into traffic flows, truck platooning and to decrease noise pollution and greenhouse emissions. Computer-controlled platooning trucks, which reduce fuel consumption as well as carbon emission, are now on the highways in Europe. They also help the traffic flow by reducing congestion. Although some researchers claim that autonomous vehicles will only travel in permitted roads because of their unpredictability, research on autonomous vehicles is increasing significantly in recent years. Traffic Management Systems have proliferated enormously throughout the years, since the first traffic lights were implemented in the second half of the 19th century. However future requirements of ITS needs smarter systems than traffic lights. Leading Information and Communication Technology (ICT) companies of the world are now starting to develop infrastructure systems for smart cities. Initially they tend to focus on utilities and traffic infrastructures specifically. Wasted hours and fuels are the rudiments of the problems in terms of ITS. Since there are a lot of ingredients in traffic assessments, day-to-day models do not yield efficient results, therefore connected dynamic models are needed. Apparently, it is not difficult to foresee that ITS will pave the way for the future smart cities. Yet, more research effort is required to merge cooperative control and traffic management strategies on individual vehicles. In this thesis instantaneous control of an AHS is elaborated to expand the standards with respect to stochastic multi agent games in smart cities.

1.1 Literature Review

The optimal solution of the single agent decision problem provides an approximation to Nash equilibria of games with large population sizes. However, unlike standard single agent optimal control problem, the characterization of this optimal solution leads to a Fokker-Planck partial differential equation evolving forward in time, and a Hamilton-Jacobi-Bellman partial differential equation evolving backward in time. These types of coupled forward-backward partial differential equations have been studied in [14]. Researchers have examined various types of mean field models recently such as mean field optimization and teams [15]; mean field stochastic control [9, 16, 17]; mean field stochastic games [15, 18–20]; mean field stochastic difference games [21], and mean field stochastic differential games [14, 22, 23]. In [11] Moon and Basar show that stochastic mean filed Stackelberg game constitutes an ε-Nash equilibrium for the followers, and then the authors prove that when N is sufficiently large ε could be picked arbitrarily close to zero.

Implementations of large-scale systems in real life became more applicable following a surge of computing power in the 1980s. Smart grids, cyber physical systems, autonomous robotics, internet networks, social networks are some application areas of large-scale systems. Aspects of successful large-scale systems require five factors. These are scalability, availability, manageability, security and development practice. Mean field games have enhanced scalability in large-scale systems in the past decade. For instance, the control of coupled oscillators has been examined in [24]. Furthermore mean field game theory has been implemented to a large number of plug-in electric vehicles in [25] to design an efficient charging profile, to a large number of electric water heating loads in [26] with the purpose of controlling them, and to a large-scale molecular biology network in [27] with the aim of analyzing its behavior. The application of mean field games in economics with a large number of firms is explained in reference [28], and wireless power control for a large number of users is discussed in [16].

Scenarios where there is one major agent and a large number of minor agents are considered in references [29] and [30]. ε-Nash equilibrium is obtained, and stochastic mean field approximation has been introduced in these papers. In [29] best stochastic mean field process is calculated through the state augmentation method in K distinct models. Fixed point analysis is carried out to the problem with the purpose of obtaining K distinct models in [30]. The nonlinear counterpart of mean field games with major and minor agents is examined in [31].

The strategy of each agent is determined in a non-cooperative way in mean field games that are discussed above, which are mainly known as *Nash* games. Furthermore hierarchical decision making between the agents does not occur. In Stackelberg games, however, the leader announces its optimum strategy beforehand based on follower's likely reactions. The leader knows everything that followers know but followers do not. Then each follower states its optimum strategy according to leader's strategy. Subsequently the leader implements its announced strategy. The equilibrium strategy in this game is called Stackelberg equilibrium [6]. In the literature dynamic games and Stackelberg differential games have been examined comprehensively, since the 1970s (see [6, 32–37] and references therein).

Networked models of smart cities connected via communication, to merge human life and technology have also been studied in technical literature. ICT is a key component for monitoring and analyzing the city because smart systems need instant environmental information to provide more efficient cities. Key global technology providers such as IBM, Cisco and Siemens invest in this area of constructing ICT infrastructure. Hence the Internet of Things and Big Data framework become more of an issue for smart cities. More detailed definitions for smart cities could be found in survey papers [38–42].

Clearly smart city is a complicated concept, which is composed of many parts of the human life. Hence for the purpose of sketching the concept of smart city, researchers should put forward innovative models for different subsystems. There are many components such as, road network, a rapid transit system, communication system, railway network, gas supply system, street light system, firefighting system, waste management system, economy system, apartment homes, power supply system, traffic light system, digital library, hospital system, law enforcement, water supply system, bridges, hotels, and office space. It is known that elder population proliferates on the world recently. Nowadays %14 of the world population is aging hence a care service model has been developed for smart city in [43]. The model provides smart shuttle service for elderly people as regards their demands. Different components of a smart city are considered as a multi criteria cost function in [44]. Then they use different architectures in order to construct a reasoning engine. Mobile crowd-sensing is a concept, which stems from implicit or explicit mobile user data. It reduces investment cost of infrastructures which are required for information in smart city. Therefore a trustworthiness model has been developed with reference to smart citizen in [45] with the intention of suggesting efficient crowd-sourcing algorithms. In [46] strategic frameworks for Dubai have been examined from perceptive of ICT. A game theory based hierarchic decision making algorithm for strategic energy management of a smart city has been proposed in [47]. The hierarchic structure is of service to fill an important gap, which is in accordance with subsystems of the energy manager.

Vast research has been conducted for improving traffic conditions such as flow, density, speed in the literature. Most of them are in accordance with machine learning methods on the purpose of prediction. Furthermore they cope with big data problem mostly. A new ITS architecture for parallel transportation management and control system has been developed in [48]. The architecture uses ITS cloud model, which composes of four layers, namely application, platform, unified resource and physical layers. Hereby, they estimate the traffic flow through Bayes expectation maximization in an artificial system. A big data solution for ITS has been discussed in [49]. There, sensing methods are presented with respect to different sensor types in traffic environment under big data challenging. Consequently, new methods have been put forward with the aim of handling big data. In [50], an ITS optimization method has been developed as an ant colony based mobile crowd-sensing algorithm, which optimizes time and path distances. A Stackelberg game type optimization technique is proposed for highways in [51], which is a game between the leader and each follower separately hence it ignores the interaction among followers. A traffic density based travel time function has been developed in comparison to flow based one in [52]. Furthermore the optimal time has been calculated under constraints for density, in this way a traffic assignment problem has been solved. A correlation between weather conditions and traffic flow has been studied by deep learning belief networks in [53]. Deep learning algorithm uses historic traffic data and weather conditions of ongoing moment so correlation value is computed. An intelligent trip modeling system has been proposed to predict traveling speed using machine learning methods in [54]. Researchers have developed an intelligent speeding prediction system in [55] to reduce intentional and unintentional speeding hence they prevent speeding-related traffic accidents and injuries. A vehicle traffic predictive cruise control system has been proposed to improve traffic operation and the fuel efficiency of vehicles based on the asymmetric traffic theory in [56]. Unnecessary deceleration and acceleration actions have been decreased under both uncongested and congested traffic conditions. The design, development, implementation, and testing of a cooperative adaptive cruise control system has been presented in [57]. It consists of two controllers, one to manage the approaching maneuver to the leading vehicle and the other to regulate car-following once the vehicle joins the platoon.

A large number of methods are used for traffic flow prediction such as autoregressive and moving average models, Kalman filter, support vector machine or neural networks in the literature. But rather than prediction we prefer to control the multi-agent system meanwhile each follower converges to the flow by mean field theory individually. In this way when vehicles travel on the highway, stop-and-go driving conditions will be prevented resulting in lower carbon emission. Here, some application examples are provided. Mean field game is applied to a routing network in terms of Wardrop equilibrium, which was characterized by an equal traffic density on all used paths, in [58] firstly. Production output adjustment in a large market has been investigated in [59]. In this model, large number of producers supply a certain product with sticky prices. A mean field game is played between vehicle owners and electric price for electric vehicles in smart grid [60]. So, electric consumption has been optimized for electric and hybrid vehicles through selling or buying energy while they connect to smart grids.

1.2 Purpose of The Thesis and Contributions

- A new linear quadratic (LQ) Mean Field (MF) Multi-level Stackelberg game (MSG) is developed [61]. This game consists of multi hierarchical levels. A global-leader imposes its strategy on sub-leaders. It plays a stochastic Stackelberg Game with its sub-leaders while considering mean field behavior of followers in upper level(1-2). The lower level (2-3) contains mean field followers of the game, each sub-leader plays Stackelberg mean field game among its followers.

- We combine mutli-layer Stackelberg games such as described in [13] with mean field games. So we have added more layers to the Stackelberg Mean Field game [11]. Combination of mutli-layer stochastic Stackelberg games and mean field constraints in optimal control problem yields more complex equations. However partitioning the system to sub-problems increases the performance.

- We carried out our (MF-MSG) algorithm to a large scale tracking problem which includes 1600 agents. Then we have examined the influence on the (MF-MSG) algorithm via comparison with systems which have smaller number of agents and one Stackelberg decision layer.

- A new type of Intelligent Transportation System (MF-ITS) is proposed for AHS in future smart cities. This method should be considered as a new Intelligent Speed Change algorithm in respect that each vehicle tends to converge its speed to the mean field term. Besides it is an optimal multi agent control algorithm since each vehicle estimates its optimal control with reference to its own and other agent's dynamics. Eventually, MF-ITS algorithm optimizes the system based on travel time, traffic flow and energy control. In this way it comes up with a method, which has positive effects on both environment and psychology of people.

- We have extended the research, new MF-MSG for ITS. We have used hierarchical structure of MF-MSG so we could use different cost functions for sub-leaders or global leader and different mean field term on the each part of the highway [62].

- There are mot many automated highway models for smart cities in the literature. Hereby, ITS is modeled as a mean field game for the first time. Then we have tested our algorithm on a realistic simulation program (ITUCiTSim), which is developed for implementing the real time traffic scenarios on a ITS highway. Most researches exploit past data of traffic conditions. Here, we come up with an instantaneous control model of the AHS, which is not affected adversely in case traffic conditions change day-to-day.

The rest of the thesis is organized as follows. Section 2 presents some mathematical background about stochastic differential equations, forward-backward differential equations, stochastic optimal control problem, and game theory. Section 3 deals with the approximate optimal control problem that each follower is faced with in the mean field limit, when sub-leaders impose their arbitrary strategy to the followers. Furthermore, approximate game problem of sub-leaders is solved subject to mean field of each group. Global leader problem is solved as a Stackelberg game in this section. In Section 4, we introduce the problem and explain the main difficulties to solve the problem exactly when the number of agents on automated highway system is large; we also translate the mean field problem to an ITS. Moreover we deal with the approximate optimal control problem that each vehicle has to solve in the mean field limit, when road-links impose their arbitrary strategies on the vehicles. Besides, road-links approximate the game problems which are solved subject to the mean field of each section of the AHS. The control center problem is also solved in this section, within a Stackelberg game framework. In Section 5, firstly numerical experiments, which consist of four different types of games that are related to different number of agents strategy and a control to one-layer hierarchical level, are given. Then we introduce our professional simulation program ITUCiTSim along with some examples on traffic simulations from the literature. Then, the results of numerical experiments are given for two different traffic scenarios. These are MFG control of vehicles under normal traffic and after an accident occurs. Here, each road-section imposes its strategy on a different number of vehicles and is managed by a control center in the Istanbul simulation model. Finally, section 6 concludes the thesis with a brief summary of the results and identifying directions for future research.

2. MATHEMATICAL PRELIMINARIES

2.1 Stochastic Differential Equations

A stochastic differential equation (SDE) is a differential equation which also includes stochastic terms. The solution of a SDE is a stochastic process. Some typical examples of SDEs are Brownian motion and the Wiener process. However, other types of random behavior are also possible (see further details [63]).

2.1.1 Probability spaces, random variables and stochastic processes

Some preliminary information about probability theory will be briefly given in this section.

Definition 2.1.1.1. Let a set Ω be nonempty, and let $\mathscr{F} \subseteq 2^{\Omega}$ (2^{Ω} is the set of all subsets in Ω), named as a class, be nonempty. We show $\mathscr{F}$ as

(i) a π-system if $A, B \in \mathscr{F} \Rightarrow A \cap B \in \mathscr{F}$,

(ii) a λ-system if

$$\begin{cases} \Omega \in \mathscr{F}; \\ A, B \in \mathscr{F},\ A \subseteq B \Rightarrow B \backslash A \in \mathscr{F}; \\ A_i \in \mathscr{F},\ A_i \uparrow A, i = 1, 2, \cdots \Rightarrow A \in \mathscr{F}, \end{cases}$$

(iii) a σ-system if

$$\begin{cases} \Omega \in \mathscr{F}; \\ A, B \in \mathscr{F},\ \Rightarrow B \backslash A \in \mathscr{F}; \\ A_i \in \mathscr{F},\ , i = 1, 2, \cdots \Rightarrow \bigcup\limits_{i=1}^{\infty} A_i \in \mathscr{F}. \end{cases}$$

Hereby the σ-field is generated by A, in case $\sigma(A)$ is the smallest σ-field containing A.

Definition 2.1.1.2. The *measurable space* is shown as $(\Omega, \mathscr{F})$. So a function $P : \mathscr{F} \rightarrow [0, 1]$ is a probability measure P on a measurable space $(\Omega, \mathscr{F})$ such that

(i) $P(\varnothing), P(\Omega) = 1$,

(ii) if $A_1, A_2, \cdots \in \mathscr{F}$ and $\{A_i\}_{i=1}^{\infty}$ is disjoint (i.e. $A_i \cup A_j = \varnothing$ if $i \neq j$) so

$$P\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} P(A_i).$$

Definition 2.1.1.3. The *triple* $(\Omega, \mathscr{F}, P)$ demonstrates a *probability space*. This is a complete probability space in case $\mathscr{F}$ consists all subsets $\mathscr{G}$ of Ω with P-outer measure zero, i.e. with

$$P^*(\mathscr{G}) := \inf\{P(\mathscr{F}); \mathscr{F} \in \mathscr{F}, \mathscr{G} \subset \mathscr{F}\} = 0.$$

If there is a smallest σ-algebra $\mathscr{H}_{\mathscr{U}}$ including $\mathscr{U}$ in any family $\mathscr{U}$ of subsets of Ω, namely

$$\bigcap \mathscr{H}_{\mathscr{U}} = \{\mathscr{H}; \mathscr{H}\, \sigma - \text{algebra of } \Omega, \mathscr{U} \subset \mathscr{H}\}.$$

$\mathscr{H}_{\mathscr{U}}$ is named as the σ-algebra constituted by $\mathscr{U}$.

The σ-algebra $\mathscr{H}_X$ constituted by X is the smallest σ-algebra on Ω including all the sets, if $X : \Omega \to R^n$ is any function, so that

$$X^{-1}(U); \quad U \subset R^n \text{ open.}$$

It could be shown that

$$\mathscr{H}_X = X^{-1}(B); \quad B \in \mathscr{B},$$

where $\mathscr{B}$ is the Borel σ-algebra on R^n. Overtly, $\mathscr{H}_X$ is the smallest σ-algebra with this feature, and X will then be $\mathscr{H}_X$-measurable.

Definition 2.1.1.4. If $X : \Omega \to R^n$, then an $\mathscr{F}$-measurable function is a *random variable* X. A probability measure μ_X on R^n is induced by every random variable , described by

$$\mu_X(B) = P(X^{-1}(B)),$$

the distribution of X is shown by μ_X.

Definition 2.1.1.5. A collection of random variables taken from R^n constitutes a *stochastic process*

$$\{X_t\}_{t\in T},$$

on the probability space $(\Omega, \mathscr{F}, P)$.

Definition 2.1.1.6 (Cauchy Sequence [64])**.** A sequence $\{x_n\}$ in a normed space where $\| x_n - x_m \| \to 0$ as $n,m \to \infty$ is a *Cauchy Sequence*; i.e., there is an integer N where $\| x_n - x_m \| < \varepsilon$ for all $n,m > N$, and given $\varepsilon > 0$.

In a normed space, every convergent sequence is a Cauchy Sequence since, if $x_n \to x$

$$\| x_n - x_m \| = \| x_n - x + x - x_m \| \leq \| x_n - x \| + \| x - x_m \| \to 0.$$

Definition 2.1.1.7 (Banach Space)**.** In case every Cauchy sequence from X has a limit in X, a normed linear vector space X is complete. Thus *Banach Space* stems from a complete normed vector space.

Definition 2.1.1.8 (The L^p Spaces)**.** Let the L^p-norm of X, $\| X \|_p$ in case of $p \in [1,\infty)$ be a constant and $X : \Omega \to R_n$ be a random variable, then

$$\| X \|_p = \| X \|_{L_p(P)} = \Big(|X(\omega)|^p dP(\omega) \Big)^{\frac{1}{p}}.$$

If $p = \infty$ we set

$$\| X \|_\infty = \| X \|_{L^\infty(P)} = \inf\{N \in R;\ |X(\omega)| \leq N \text{ a.s.}\}.$$

The related L^p-spaces are given as

$$L^p(P) = L^p(\Omega) = \{X : \Omega \to R^n;\ \| X \|_p < \infty\}.$$

The *Lp-spaces* are Banach spaces, and they are also normed linear spaces. A Hilbert space is a norm where $p = 2$, the space $L^2(P)$, i.e. a complete inner product space, with inner product

$$(X,Y)_{L^2(P)} := E[X \cdot Y];\ X,Y \in L^2(P).$$

Definition 2.1.1.9 (Brownian Motion)**.** There are some finite-dimensional distributions of B_t on Ω where $(\Omega, \mathscr{F}, P_x)$ is a probability space by Kolmogorov's theorem. $\{B_t\}_{t\geq 0}$ is a stochastic process and then

$$P^x(B_{t_1} \in F_1, \cdots, B_{t_k} \in F_k) = \int_{F_1 \times \cdots F_k} p(t_1, x, x_1) \ldots p(t_k - t_{k-1}, x_{k-1}, x_k) dx_1 \ldots dx_k,$$

is named as Brownian Motion.

2.1.2 Ito integrals

Let a mathematical solution of a SDE be computed by ito Integrals.

$$\frac{dX}{dt} = b(t, X_t) + \sigma(t, X_t) \cdot \text{"constant"},$$

where b and σ are given functions is a stochastic differential equation containing noise. Firstly 1-dimensional noise problem will be worked on the problem. Some stochastic process W_t will be examined, on the purpose of describing the noise term, such that

$$\frac{dX}{dt} = b(t, X_t) + \sigma(t, X_t) \cdot W_t. \tag{2.1}$$

In engineering problems that cover lots of circumstances; on the purpose of deriving a solution it should be assumed that W_t has, the following properties:

(i) $\{W_t\}$ is stationary, i.e. the (joint) distribution of $W_{t1+t}, \ldots, W_{tk+t}$ does not depend on t.

(ii) $E[W_t] = 0$ for all t.

(iii) $t_1 \neq t_2 \rightarrow W_{t1}$ and W_{t2} are independent.

Besides, the *white noise process* could be shown as W_t in stochastic processes generally. Let

$$X_{k+1} - X_k = b(t_k, X_k)\Delta t_k + \sigma(t_k, X_k) W_k \Delta t_k, \tag{2.2}$$

where

$$X_j = X_{t_j}, \quad W_k = W_{t_k}, \quad \Delta t_k = t_{k+1} - t_k.$$

In this equation the only such process with continuous paths is the Brownian motion B_t. Hence $V_t = B_t$ should hold and attained from (2.2)

$$X_k = X_0 + \sum_{j=0}^{k-1} b(t_j, X_j)\Delta t_j + \sum_{j=0}^{k-1} \sigma(t_j, X_j)\Delta B_j. \tag{2.3}$$

It could be shown easily that the limit of the right hand side of equation (2.3) exists, when $\Delta t_j \to 0$. Subsequently the usual integration notation could be denoted by using

$$X_k = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s, \tag{2.4}$$

and we can assume as a convention that equation (2.1) demonstrates that the equation (2.4) is satisfied by a stochastic process $X_t = X_t(\omega)$.

Now the *Ito' integral* could be shown for functions $f \in V$

$$I[f](\omega) = \int_S^T f(t, \omega)dB_t(\omega),$$

where B_t denotes the 1-dimensional Brownian motion.

The natural idea is: Firstly, $I[\phi]$ should be described for a basic class of functions ϕ. Then, we demonstrate that each $f \in V$ can be estimated by such ϕ's and we utilize these to describe $\int f dB$ as the limit of $\int \phi dB$ as $\phi \to f$.

Its elementary form describes a function $\phi \in V$, in case of

$$\phi(t, \omega) = e_j(\omega) \cdot X[t_j, t_{j+1})(t).$$

Finally since $\phi \in V$ each function e_j must be F_{t_j}-measurable.

The integral is given for elementary functions $\phi(t, \omega)$

$$\int_S^T \phi(t, \omega)dB_t(\omega) = \sum_{j \geq 0} e_j(\omega)[B_{t_{j+1}} - B_{t_j}](\omega).$$

Definition 2.1.2.1 (The Ito' integral)**.** Then the Ito' integral of f (from S to T) is described by

$$\int_S^T f(t, \omega)dB_t(\omega) = \lim_{n \to \infty} \int_S^T \phi_n(t, \omega)dB_t(\omega),$$

where a *sequence of elementary functions* is $\{\phi_n\}$ so that

$$E\Big[\int_S^T \big(f(t, \omega) - \phi_n(t, \omega)\big)^2\Big] \to 0 \qquad \text{as } n \to \infty.$$

2.1.3 The Ito formula

When calculating a given integral, the base definition of Ito' integrals cannot be used. The solution is derived by chain rule, where the case resembles ordinary Riemann integrals calculation but its basic definition is not used. The fundamental theorem of calculus is used instead of basic definitions of Riemann integrals.

Hereby, we have only integration theory instead of differentiation theory. Even though this kind of situation occurs, there is a solution. The Ito' formula is a version of the chain rule of an Ito' integral. It is quite practical for the purpose of solving Ito' integrals.

Definition 2.1.3.1 (1-dimensional Ito' processes)**.** A *stochastic process* X_t on $(\Omega, \mathscr{F}, P)$ of the equation (2.5) is a (1-dimensional) Ito' process (or stochastic integral) where B_t is a 1-dimensional *Brownian motion* on $(\Omega, \mathscr{F}, P)$.

$$X_t = X_0 + \int_0^t u(s,\omega)ds + \int_0^t v(s,\omega)dB_s, \tag{2.5}$$

where $v \in W_H$, such that

$$P\Big[\int_0^t v(s,\omega)^2 ds < \infty \text{ for all } t \geq 0\Big].$$

It is assumed that u is H_t-adapted and

$$P\Big[\int_0^t |u(s,\omega)| ds < \infty \text{ for all } t \geq 0\Big].$$

The equation (2.5) could occasionally be written in the shorter differential form as

$$dX_t = udt + vdB_t.$$

For instance, (2.5) could be presented by

$$d\Big(\frac{1}{2}B_t^2\Big) = \frac{1}{2}dt + B_t dB_t.$$

Now the first main result is given by:

Theorem 2.1.3.2 (The 1-dimensional Ito' formula))**.** *An Ito' process X_t could be denoted by*

$$dX_t = udt + vdB_t, \tag{2.6}$$

In case g is twice continuously differentiable on $[0,\infty) \times R$ and $g(t,x) \in C^2([0,\infty) \times R)$,

$$Y_t = g(t, X_t).$$

is an Ito' process again, and

$$dY_t = \frac{\partial g}{\partial t}(t,X_t)dt + \frac{\partial g}{\partial x}(t,X_t)dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t,X_t)\cdot(dX_t)^2,$$

where $(dX_t)^2 = (dX_t)\cdot(dX_t)$ *is calculated by the fact that the rules*

$$dt\cdot dt = dt\cdot dB_t = dB_t\cdot dt = 0 \quad dB_t\cdot dB_t = dt.$$

Theorem 2.1.3.3 (The general Ito' formula)**.** *Let*

$$dX_t = udt + vdB_t,$$

be an n-dimensional Ito' process as stated. Assume that $g(t,x) = (g_1(t,x),\ldots,g_p(t,x))$ *is a* C^2 *map from* $([0,\infty)\times R^n$ *into* $R^p)$. *Subsequently the process*

$$Y_{t,\omega} = g(t,X_t)$$

is an Ito' process again, whose component number k, Y_k, *is stated as*

$$dY_k = \frac{\partial g_k}{\partial t}(t,X)dt + \sum_i \frac{\partial g_k}{\partial x_i}(t,X)dX_i + \frac{1}{2}\sum_{i,j}\frac{\partial^2 g}{\partial x_i \partial x_j}(t,X)\cdot dX_i dX_j,$$

where

$$dt\cdot dt = dt\cdot dB_i = dB_i\cdot dt = 0 \quad dB_i\cdot dB_j = \delta_{ij}dt.$$

2.2 Forward-Backward Stochastic Differential Equations

In this section FBSDEs are given in finite time duration. A simple form of linear FBSDEs is searched, on the purpose of deriving necessary conditions for solvability. Subsequently we investigate necessary and sufficient conditions to assure solvability in FBSDEs. The state for one-dimensional Brownian motion will be shown (see further details [65]).

2.2.1 Solvability of linear FBSDEs

The system of coupled linear FBSDEs are taken into consideration:

$$\begin{aligned} dX(t) =& \{AX(t)+BY(t)+CZ(t)+Db(t)\}dt \\ &+\{A_1X(t)+B_1Y(t)+C_1Z(t)+D_1\sigma(t)\}dW(t), \\ dY(t) =& \{\hat{A}X(t)+\hat{B}Y(t)+\hat{C}Z(t)+\hat{D}\hat{b}(t)\}dt \\ &+\{\hat{A}_1X(t)+\hat{B}_1Y(t)+\hat{C}_1Z(t)+\hat{D}_1\hat{\sigma}(t)\}dW(s), \\ X(0) =& x, \qquad Y(T) = GX(T)+Fg \qquad t\in[0,T]. \end{aligned} \tag{2.7}$$

Here, g is a random variable, furthermore $b, \sigma, \hat{b}$ and $\hat{\sigma}$ are stochastic processes and A, B, C etc. are matrices of appropriate sizes. $\{F_t\}_{t \geq 0}$-adapted processes $X(\cdot), Y(\cdot)$ and $Z(\cdot)$ are searched for, valued in R_n, R_m and R_l, respectively.

Definition 2.2.1.1. If the sequent is satisfied for all $t \in [0,T]$, then a triple $(X,Y,Z) \in M[0,T]$ is named an adapted solution of (2.7), almost certain:

$$
\begin{aligned}
X(t) =& x + \int_0^t \{AX(s) + BY(s) + CZ(s) + Db(s)\} ds \\
& + \int_0^t \{A_1 X(s) + B_1 Y(s) + C_1 Z(s) + D_1 \sigma(s)\} dW(s), \\
Y(t) =& GX(T) + Fg - \int_t^T \{\hat{A}X(s) + \hat{B}Y(s) + \hat{C}Z(s) + \hat{D}\hat{b}(s)\} ds \\
& - \int_t^T \{\hat{A}_1 X(s) + \hat{B}_1 Y(s) + \hat{C}_1 Z(s) + \hat{D}_1 \hat{\sigma}(s)\} dW(s).
\end{aligned}
$$

(2.7) is accepted as solvable, in case (2.7) admits an adapted solution.

After some simplifications the following FBSDE is obtained:

$$
\begin{aligned}
dX(t) =& \{AX + BY + CZ\} dt \\
& + \{A_1 X + B_1 Y + C_1 Z\} dW(t), \\
dY(t) =& \{\hat{A}X + \hat{B}Y + \hat{C}Z\} dt + Z dW(s), \\
X(0) =& 0, \qquad Y(T) = g \qquad t \in [0,T].
\end{aligned}
\tag{2.8}
$$

By demonstrating

$$
\begin{aligned}
A = \begin{bmatrix} A & B \\ \hat{A} & \hat{B} \end{bmatrix}, \qquad C = \begin{bmatrix} C \\ \hat{C} \end{bmatrix} \\
A_1 = \begin{bmatrix} A_1 & B_1 \\ 0 & 0 \end{bmatrix}, \qquad C_1 = \begin{bmatrix} C_1 \\ I \end{bmatrix}.
\end{aligned}
$$

the equation (2.8) could be written as stated below:

$$
\begin{aligned}
d\left(\begin{bmatrix} X \\ Y \end{bmatrix} \right) =& \{A\left(\begin{bmatrix} X \\ Y \end{bmatrix} \right) + CZ\} dt + \{A_1 \left(\begin{bmatrix} X \\ Y \end{bmatrix} \right) C_1 Z\} dW(t), \\
X(0) =& 0, \qquad Y(T) = g \qquad t \in [0,T].
\end{aligned}
$$

(2.8) is named as a (linear) *stochastic control* system, where (X,Y) is the state and Z is the control.

2.2.2 A Riccati type equation

In this subsection, a sufficient condition for unique solvability of (2.8) will be obtained. We will attain a Riccati type equation and a BSDE in the form of (2.8).

Let an adapted solution of (2.8) be assumed as $(X,Y,Z) \in M[0,T]$. Then X and Y are related in accordance with the formula

$$Y(t) = P(t)X(t) + p(t), \ \forall t \in [0,T], \text{ a.s.} \tag{2.9}$$

where $p : [0,T] \times \Omega \to R^m$ is an $\{F_t\}_{t\geq 0}$-adapted process and $P : [0,T] \to R^{m\times n}$ is a deterministic matrix-valued function. Then we obtain the equations for $p(\cdot)$ and $P(\cdot)$. Initially, from the terminal condition and (2.9) in (2.8), we have

$$g = P(T)X(T) + p(T).$$

Let us assume

$$P(T) = 0, \qquad p(T) = g.$$

Due to fact that $p(\cdot)$ is necessitated to be $\{F_t\}_{t\geq 0}$-adapted and $g \in L^2_{F_T}(\Omega; R^m)$, a BSDE that $p(\cdot)$ satisfies is :

$$\begin{aligned} dp(t) =& \alpha(t)dt + q(t)W(t), \qquad t \in [0,T] \\ p(T) =& g, \end{aligned} \tag{2.10}$$

where $\alpha(\cdot), q(\cdot) \in L^2_F(0,T;R^m)$ are undetermined. Then, by Ito' formula:

$$\begin{aligned} dY(t) =& \{\dot{P}X + P[AX + BY + CZ] + \alpha\}dt \\ &+ \{P[A_1X + B_1Y + C_1Z] + q\}dW(t), \\ =& \{[\dot{P} + PA + PBP]X + PCZ + PBp + \alpha\}dt \\ &+ \{[PA_1 + PB_1P] + PC_1Z + PB_1p + q\}dW(t). \end{aligned} \tag{2.11}$$

A comparison of (2.11) with the second equation in (2.8) yields:

$$[\dot{P} + PA + PBP]X + PCZ + PBp + \alpha = [\hat{A} + \hat{B}P]X + \hat{C}Z + \hat{B}p. \tag{2.12}$$

It is possible to write (2.12) as

$$\begin{aligned} 0 =& [\dot{P} + PA + PBP - \hat{A} - \hat{B}P \\ &+ (PC - \hat{C})(I - PC_1)^{-1}(PA_1 + PB_1P)]X \\ &+ [PB - \hat{B} + (PC - \hat{C})(I - PC_1)^{-1}PB_1]p \\ &+ (PC - \hat{C})(I - PC_1)^{-1}q + \alpha. \end{aligned}$$

The differential equation of $R^{m\times n}$-valued function $P(\cdot)$ is given as:

$$
\begin{aligned}
&\dot{P}+PA+PBP-\hat{A}-\hat{B}P \\
&\quad +(PC-\hat{C})(I-PC_1)^{-1}(PA_1+PB_1P)=0, \qquad t\in[0,T], \\
&P(T)=0.
\end{aligned}
\tag{2.13}
$$

We call (2.13) as a *Riccati type equation*. The solution $P(\cdot)$ over $[0,T]$ is given by (2.13) so that

$$
[I-PC_1]^{-1} \text{ is bounded for} \qquad t\in[0,T]. \tag{2.14}
$$

Unifying this with (2.10), it is understood that the next BSDE should be :

$$
\begin{aligned}
dp(t)=&-\Big\{[PB-\hat{B}+(PC-\hat{C})(I-PC_1)^{-1}PB_1]p \\
&+(PC-\hat{C})(I-PC_1)^{-1}q\Big\}dt+qW(t), \qquad t\in[0,T] \\
p(T)=&g.
\end{aligned}
\tag{2.15}
$$

BSDE (2.15) gives a unique adapted solution $(p(\cdot),q(\cdot))\in N[0,T]$, in case there is a solution $P(\cdot)$ given by (2.13) so that (2.14) is satisfied. Subsequently it can be expressed as:

$$
\begin{aligned}
\tilde{A}=&A+BP+C(I-PC_1)^{-1}(PA_1+PB_1P), \\
\tilde{A_1}=&A_1+B_1P+C(I-PC_1)^{-1}(PA_1+PB_1P), \\
\tilde{b}=&Bp+C(I-PC_1)^{-1}(PB_1p+q), \\
\tilde{\sigma}=&B_1p+C_1(I-PC_1)^{-1}(PB_1p+q).
\end{aligned}
$$

It is understood that $\tilde{b}$,$\tilde{\sigma}$ are $\{F_t\}_{t\geq 0}$-adapted processes, also $\tilde{A}$ and $\tilde{A_1}$ are time-dependent matrix-valued functions. Then, under (2.14), a unique strong solution is given by the following SDE:

$$
\begin{aligned}
dX(t)=&(\tilde{A}+\tilde{b})dt+(\tilde{A_1}+\tilde{\sigma})dW, \qquad t\in[0,T], \\
X(0)=&x.
\end{aligned}
\tag{2.16}
$$

A representation of the adapted solution of FBSDE (2.7) is presented by the next theorem.

Theorem 2.2.2.1. *Let (2.13) give a solution $P(\cdot)$ so that (2.14) holds. A unique adapted solution $(X,Y,Z)\in M[0,T]$ is given by FBSDE (2.7) which is specified by (2.16), and (2.15).*

2.3 Stochastic Optimal Control Problems

In this subsection it is concentrated on *dynamic* systems. These systems are defined by Ito' stochastic differential equations and are occasionally named as *diffusion models*. *White noise* is the base source of uncertainty in these diffusion models. Because of the fact that the systems are dynamic, the concerned *decisions (controls)*, which are made established upon the most updated information convenient to the *decision makers (controllers)*, should also change in time. An optimal decision among all possible ones should be chosen by the decision makers, with the purpose of reaching the best anticipated result with respect to their aims. These types of optimization problems are known as *stochastic optimal control problems* (see further details [66]).

2.3.1 Formulations of stochastic optimal control problems

An m-dimensional standard Brownian motion $W(\cdot)$ is described, where the usual condition is satisfied by a filtered probability space $(\Omega, \mathscr{F}, \{F_t\}_{t\geq 0}, P)$. Then, take into consideration the *controlled* stochastic differential equation:

$$\begin{aligned} dx(t) =& b(t,x(t),u(t))dt + \sigma(t,x(t),u(t))dW(t), \\ x(0) =& x_0 \in R^n, \end{aligned} \tag{2.17}$$

where $b[0,T] \times R^n \times U \to R^n, \sigma : [0,T] \times R^n \times U \to R^{n\times m}$, with U being a specified separable metric space, and $T \in (0,\infty)$ being fixed. The control function $u(\cdot)$ represents the action, decision, or policy of the decision-makers (controllers). At any instant of time the controller knows some information (as given by the information field $\{F_t\}_{t\geq 0}$) of what has happened up to that moment, however may not predict what will happen subsequently because of the uncertainty of the system. This nonanticipative limitation in mathematical terms could be represented as "$u(\cdot)$ is $\{F_t\}_{t\geq 0}$-adapted". The control $u(\cdot)$ is taken from the set

$$\mathscr{U}[0,T] = \{u : [0,T] \times \Omega \to U |\; u(\cdot) \text{ is } \{F_t\}_{t\geq 0}\text{-adapted}\}.$$

Any $u(\cdot) \in \mathscr{U}[0,T]$ is named as a *feasible control*. Besides, it is possible that there may be some constraints on the states. Equation (2.17) has random coefficients. Given multifunction $S(t) : [0,T] \to 2^{R^n}$. The *state constraint* can be specified as

$$x(t) \in S(t), \quad \forall t \in [0,T].$$

It should be considered that other types of state constraints are also possible.

The cost functional is introduced as follows:

$$J(u(\cdot)) = E\Big\{ \int_0^T f(t,x(t),u(t))dt + h(x(T)) \Big\}. \tag{2.18}$$

Definition 2.3.1.1. Let $W(t)$ be a given m-dimensional standard motion, when $(\Omega, \mathscr{F}, \{F_t\}_{t\geq 0}, P)$ are specified to satisfy the usual conditions. A control $u(\cdot)$ is named an admissible control, and $(x(\cdot),u(\cdot))$ an admissible pair, if

(i) $x(\cdot)$ is the unique solution of equation (2.17),

(ii) $u(\cdot) \in u[0,T]$,

(iii) $f(\cdot,x(\cdot),u(\cdot)) \in L_{\mathscr{F}}(0,T;R)$ and $h(x(T)) \in L_{\mathscr{F}}(0,\Omega;R)$.

(iv) some prescribed state constraints are satisfied,

$\mathscr{U}[0,T]$ demonstrates the set of all admissible controls. Then the stochastic optimal control problem can be given as:

Problem. Minimize (2.18) over $\mathscr{U}[0,T]$. The main purpose of the problem is to find $\bar{u}(\cdot) \in \mathscr{U}[0,T]$ (if it ever exists), so that

$$J(\bar{u}(\cdot)) = \inf_{u(\cdot)\in\mathscr{U}[0,T]} J(u(\cdot)). \tag{2.19}$$

2.3.2 Maximum principle and stochastic Hamiltonian systems

The key idea of Pontryagin's maximum principle is derived from the classical calculus of variations. Duality is used in the derivation of the maximum principle. The main problem is that the Ito' stochastic integral $\int_t^{t+\varepsilon} \sigma dW$ is only of order $\sqrt{\varepsilon}$ (rather than ε as with the normal Lebesgue integral, hence the common first-order variation method fails. To deal with this problem, first-order and second-order terms in the Taylor expansion of the spike variation should be worked on. Subsequently a stochastic Hamiltonian system that consists of two forward-backward stochastic differential equations and a maximum condition with an additional term quadratic in the diffusion coefficient are obtained (see further details [66]).

Let us make the following assumptions:

(S0) $W(t)$ generates $\{\mathscr{F}_t\}_{t\geq 0}$ which is the natural filtration, augmented by all the P-null sets in $\mathscr{F}$.

(S1) $T > 0$, and (U,d) is a separable metric space.

(S2) $L > 0$ is a constant, the maps $b,\sigma,f,$ and h are measurable and there is a modulus of continuity $\bar{\omega} : [0,\infty) \to [0,\infty)$ so that for $\varphi(t,x,u) = b(t,x,u),\sigma(t,x,u),f(t,x,u),h(x)$, then

$$\begin{aligned} |\varphi(t,x,u)-\varphi(t,\hat{x},\hat{u})| &\leq L|x-\hat{x}|+\bar{\omega}(d(u,\hat{u})), \\ \forall t\in[0,T], \quad x,\hat{x} &\in R^n, \quad u,\hat{u}\in U, \\ |\varphi(t,0,u)|\leq L, \qquad \forall (t,u) &\in [0,T]\times U. \end{aligned}$$

(S3) The maps $b,\sigma,f,$ and h are C^2 in x. Furthermore, there is a modulus of continuity $\bar{\omega} : [0,\infty) \to [0,\infty)$, and a constant $L > 0$ so that for $\varphi(t,x,u) = b(t,x,u),\sigma(t,x,u),f(t,x,u),h(x)$,

$$\begin{aligned} |\varphi_x(t,x,u)-\varphi_x(t,\hat{x},\hat{u})| &\leq L|x-\hat{x}|+\bar{\omega}(d(u,\hat{u})), \\ |\varphi_{xx}(t,x,u)-\varphi_{xx}(t,\hat{x},\hat{u})| &\leq \bar{\omega}(|x-\hat{x}|+d(u,\hat{u})), \\ \forall t\in[0,T], \quad x,\hat{x} &\in R^n, \quad u,\hat{u}\in U. \end{aligned}$$

The adjoint variable $p(\cdot)$ has a central role in the maximum principle. The adjoint equation that $p(\cdot)$ (the terminal value is given) satisfies determines a backward ordinary differential equation. However, if the time is reversed, it is equivalent to a forward equation. Nevertheless the time can not be reversed in the stochastic case, as it might ruin the nonanticipativeness of the solutions. Instead, the following terminal value problem for a stochastic differential equation is introduced:

$$\begin{aligned} dp(t) =& \Big\{ b_x(t,\bar{x}(t),\bar{u}(t))^T p(t) + \sum_{j=1}^{m} \sigma_x^j(t,\bar{x}(t),\bar{u}(t))^T q_j \\ & -f_x(t,\bar{x}(t),\bar{u}(t))^T \Big\} dt + q(t)dW(t), \qquad t\in[0,T], \\ p(T) =& -h_x(\bar{x}(T)). \end{aligned} \tag{2.20}$$

Note that $(p(\cdot),q(\cdot))$ is a pair of $\{\mathscr{F}_t\}_{t\geq 0}$-adapted processes. Let the equation (2.20) be a *backward stochastic differential equation*. The solution $(p(\cdot),q(\cdot))$ is necessitated to be $\{\mathscr{F}_t\}_{t\geq 0}$-adapted, even though the equation is to be solved backwards (because the terminal value is specified). Any pair of processes $(p(\cdot),q(\cdot)) \in L_{\mathscr{F}}(0,T;R^n) \times (L_{\mathscr{F}}(0,T;R^n))^m$ satisfying (2.20) is named an adapted solution of (2.20). (2.20) admits a unique adapted solution $(p(\cdot),q(\cdot))$, when (S0)-(S3) are satisfied, for any $(\bar{x}(\cdot),\bar{u}(\cdot)) \in L_{\mathscr{F}}(0,T;R^n)\times U[0,T]$. Then (S0) could not be omitted.

The uncertainty or the risk factor in the system is introduced by an additional adjoint equation as follows:

$$\begin{aligned} dP(t) =&\Big\{ b_x(t,\bar{x}(t),\bar{u}(t))^T P(t) + P(t) b_x(t,\bar{x}(t),\bar{u}(t)) \\ &+ \sum_{j=1}^{m} \sigma_x^j(t,\bar{x}(t),\bar{u}(t))^T P(t) \sigma_x^j(t,\bar{x}(t),\bar{u}(t)) \\ &+ \sum_{j=1}^{m} \sigma_x^j(t,\bar{x}(t),\bar{u}(t))^T Q_j(t) + Q_j(t)\sigma_x^j(t,\bar{x}(t),\bar{u}(t)) \\ &- H_{xx}\big(t,\bar{x}(t),\bar{u}(t),p(t),q(t)\big)^T \Big\} dt + \sum_{j=1}^{m} Q_j(t) dW^j(t), \qquad t\in[0,T], \\ p(T) =& - h_{xx}(\bar{x}(T)), \end{aligned} \tag{2.21}$$

where the *Hamiltonian* H is given by

$$\begin{aligned} H(t,x,u,p,q) =& \langle p, b(t,x,u)\rangle + tr[q^T \sigma(t,x,u)] - f(t,x,u) \\ & (t,x,u,p,q) \in [0,T]\times R^n \times U \times R^n \times R^{n\times m}, \end{aligned} \tag{2.22}$$

and $(p(\cdot),q(\cdot))$ is the solution to (2.20). In the above (2.21), the unknown is again a pair of processes $(P(\cdot),Q(\cdot)) \in L^2_{\mathscr{F}}(0,T;S^n)\times(L^2_{\mathscr{F}}(0,T;S^n))^m$.

Theorem 2.3.2.1 (Stochastic Maximum Principle). *If (S0)-(S3) hold and $(\bar{x}(\cdot),\bar{u}(\cdot))$ are an optimal pair of the problem, then there are pairs of processes*

$$\begin{aligned} (p(\cdot),q(\cdot)) &\in L^2_{\mathscr{F}}(0,T;R^n)\times(L^2_{\mathscr{F}}(0,T;R^n))^m, \\ (P(\cdot),Q(\cdot)) &\in L^2_{\mathscr{F}}(0,T;S^n)\times(L^2_{\mathscr{F}}(0,T;S^n))^m, \end{aligned}$$

where

$$\begin{aligned} q(\cdot) = \big(q_1(\cdot),\dots,q_m(\cdot)\big), &\qquad Q(\cdot) = \big(Q_1(\cdot),\dots,Q_m(\cdot)\big), \\ q_j(\cdot) \in L^2_{\mathscr{F}}(0,T;R^n), \qquad Q_j(\cdot) &\in L^2_{\mathscr{F}}(0,T;S^n), \qquad 1\le j\le m, \end{aligned}$$

so the first-order and second-order adjoint equations (2.20) and (2.21) are satisfied, respectively, so that

$$\begin{aligned} H(t,\bar{x}(t),&\bar{u}(t),p(t),q(t)) - H(t,\bar{x}(t),u(t),p(t),q(t)) \\ &-\frac{1}{2} tr\Big(\{\sigma(t,\bar{x},\bar{u}) - \sigma(t,\bar{x},u)\}^T P(t) \\ &\cdot\{\sigma(t,\bar{x},\bar{u}) - \sigma(t,\bar{x},u)\}\Big) \ge 0, \\ &\forall u\in U \qquad t\in[0,T], \end{aligned} \tag{2.23}$$

or, equivalently,

$$H(t,\bar{x}(t),\bar{u}(t)) = \max_{u\in\mathscr{U}} H(t,\bar{x}(t),u(t)) \qquad t\in[0,T]. \tag{2.24}$$

The inequality (2.23) is named as the variational inequality, and (2.24) is named as the maximum condition.

The system (2.17) along with its first-order adjoint system can be written as:

$$\begin{aligned} dx(t) =& H_p(t,x(t),u(t),p(t),q(t))dt \\ & + H_q(t,x(t),u(t),p(t),q(t))dW(t), \\ dp(t) =& H_x(t,x(t),u(t),p(t),q(t))dt + q(t)dW(s), \\ x(0) =& x_0, \qquad p(T) = -h_x(x(T)). \end{aligned} \tag{2.25}$$

The combination of (2.25), (2.20), and (2.24) (or (2.24)) is named as an (extended) *stochastic Hamiltonian system*, with its solution being a 6-tuple $(x(\cdot),u(\cdot),p(\cdot),q(\cdot),P(\cdot),Q(\cdot)$. Therefore, Theorem 2.3.2.1 could be rephrased as the following.

Theorem 2.3.2.2. *If (S0)-(S3) hold and the problem admits an optimal pair* $(\bar{x}(\cdot),\bar{u}(\cdot)$, *then the stochastic Hamiltonian system (2.25), (2.20), and (2.24) (or (2.24)) are solved by the optimal 6-tuple* $(x(\cdot),u(\cdot),p(\cdot),q(\cdot),P(\cdot),Q(\cdot)$ *of the problem.*

It is understood from the above result that system (2.25) (with $u(\cdot)$ given) is also named as a *forward-backward stochastic differential* equation.

2.4 Game Theory

Game theory is a decision making method, which is played among players who may act in cooperative or competitive behavior. The individuals are named as decision makers or players. Main application areas of game theory are conflicting circumstances arising in economy, politics and wars. Furthermore it is also studied by control theory, dynamic programming and Pontryagin's "maximum principle". Let "**decisions**" are made by players, moreover there are "**actions**" constituting "**strategies**" on the other hand. (see further details [6]).

Definition 2.4.0.1 (Saddle-Point Equilibrium)**.** Let the players adopt row i^*, column j^* as a pair of strategies for a specified $(m\times n)$ matrix game $A=\{a_{ij}\}$. Then, if the inequality

$$a_{i^*j} \leq a_{i^*j^*} \leq a_{ij^*}$$

is satisfied for all $i = 1,\ldots,m$ and all $j = I,\ldots,n$, the strategy row i^*, column j^* constitutes a *saddle-point equilibrium*, and the matrix game has a *saddle point in pure strategies*. The related outcome $a_{i^*j^*}$ of the game is named as the *saddle-point value*, and is indicated by $V(A)$.

2.4.1 Continuous time infinite dynamic games (Differential games)

Some background and a general formulation of infinite dynamic games are provided in this subsection. "Differential game" is a game whose dynamics is decribed by differential equations. Since the term "differential game" is also used for other classes of games, it would be more suitable to use "dynamic game" as the general term. In infinite dynamic games, while the players obtain some dynamic information throughout the decision process, the action sets of the players consist of an infinite number of elements. Moreover, some games can be also described in discrete time. (see further details [6]).

In the literature *differential games* are also named as continuous-time infinite dynamic games, a differential equation specifies the state of decision problems in a time interval. Hence, these games could be formulated in case of prespecified fixed duration as follows:

Definition 2.4.1.1. A N-person dynamic game includes the following.

(i) The players' set is an index set $N = I,\ldots,N$.

(ii) The duration of the change of the game is specified by a time interval $[0,T]$which is denoted a priori.

(iii) The *trajectory space* of the game is an infinite set S_o with some topological structure. $\{x(t),\ 0 \leq t \leq T\}$ creates the permissible state trajectories of the game and specifies its elements. Moreover, S^0 is a subset of a finite- dimensional vector space for each fixed $t \in [0,T]$, $x(t) \in S^0$.

(iv) The *control (action) space* of P_i, whose elements $u^i(t)$, $0 \leq t \leq T$ are the control functions or simply the controls of P_i, for each $i \in N$ is described by an infinite set U^i

with some topological structure. Moreover, for each fixed $t \in [0,T],\ u^i(t) \in S^i$, there is a set $S^i \subseteq R^m, (i \in N)$.

(v) A differential equation is introduced

$$\frac{dx(t)}{dt} = f(t, x(t), u^1(t), \ldots, u^n(t)),\ x(0) = x_0, \tag{2.26}$$

where the state trajectory of the game is denoted by x_t. In the above equation the specified initial state is x_0 and the N-tuple of control functions are $\{u^i(t),\ 0 \leq t \leq T\}(i \in N)$.

(vi) For each $i \in N$ a set-valued function $\eta^i(\cdot)$ is identified as

$$\eta^i(t) = \{x(s),\ 0 \leq s \leq \varepsilon_t^i\},\ 0 \leq \varepsilon_t^i \leq t,$$

where the state information gained is indicated by $\eta^i(t)$, ε_t^i is nondecreasing in t, and recalled by P_i at time $t \in [0,n]$. The information pattern of P_i is specified by features of $\eta^i(\cdot)$ (actually, ε_t^i), and the information structure of the game is the collection (over $i \in N$) of these information structures.

(vii) The cylinder sets $x \in S_o, x(s) \in B$ where B is a Borel set in S^0 and $0 \leq s \leq \varepsilon_t^i$ constitutes a sigma-field N_t^i; in S_o. The *information field* of P_i is $N_t^i,\ t \leq t_0$.

(viii) The strategy space of P_i is a prespecified class Γ^i of mappings $\gamma^i : [0,T] \times S_o \to S^i$, with the property that $u^i(t) = \gamma^i(t,x)$ is N_t^i-measurable. A permissible strategy for P_i is Γ^i, and each of its elements is γ^i.

(ix) For each $i \in N$ two functionals $q^i : S^0 \to R,\ g^i : [0,T] \times S^0 \times S^1 \times \cdots \times S^N \to R$ are described, such that the composite functional

$$L^i(u^1, \ldots, u^N) = \int_0^T g^i(t, x(t), u^1(t), \ldots, u^N(t)),$$

is well-defined for every $u^j(t) = \gamma^j(t,x),\ \gamma^j \in \Gamma^j(j \in N)$,and the cost functional of P_i in the game is of fixed duration for each $i \in N$. L^i.

Definition 2.4.1.2. P_i's information structure in an N-person continuous-time deterministic dynamic game (differential game) of prespecified fixed duration $[0,T]$ is

(i) open-loop (OL) pattern $\varepsilon_t^i = \{x_0\},\ t \in [0,T]$,

(ii) closed-loop perfect state (CLPS) pattern if

$$\varepsilon_t^i = \{x_s,\ 0 \le s \le t\}, \qquad t \in [0,T],$$

(iii) ε-delayed closed-loop perfect state (ε (DCLPS) pattern if

$$\begin{cases} \varepsilon_t^i = \{x_0,\ 0 \le s \le \varepsilon\}, \\ \varepsilon_t^i = \{x_s,\ 0 \le s \le t-\varepsilon\}, \qquad \varepsilon < t, \end{cases}$$

where $\varepsilon > 0$ is fixed,

(iv) memoryless (perfect state) (MPS) pattern if $\varepsilon_t^i = \{x_0, x(t)\}$, $t \in [0,T]$,

(v) feedback (perfect state) (FB) pattern if $\varepsilon_t^i = \{x(t)\}$, $t \in [0,T]$.

Theorem 2.4.1.3. *Let us assume $S_o = C^n \in [0,T]$ within the framework of description 2.4.1.1, and the information patterns of description 2.4.1.2. If*

(i) for each $x \in S^0, u^i \in S^i, i \in N$ $f(t,x,u^1,\ldots,u^N)$ is continuous in $t \in [0,T]$,

(ii) for some $k > 0$ $f(t,x,u^1,\ldots,u^N)$ is uniformly Lipschitz in $x, u^1,\ldots,u^N$,

$$\begin{aligned} &\| f(t,x,u^1,\ldots,u^N) - f(t,\bar{x},\bar{u}^1,\ldots,\bar{u}^N) \| \\ &\le k \max_{0\le t\le T} \{\| x(t)-\bar{x}(t) \| + \sum_{i\in N} \| u^i(t)-\bar{u}^i(t) \|\}, \\ &x(\cdot),\ \bar{x}(\cdot) \in C^n[0,T]; \quad u^i(\cdot),\ \bar{u}^i(\cdot) \in U^i \quad (i \in N) \end{aligned}$$

(iii) for each $x(\cdot) \in C^n[0,T]$ $\gamma^i \in \Gamma^i(i \in N), \gamma^i(t,x)$ is continuous in t and uniformly Lipschitz in $x(\cdot) \in C^n[0,T]$, a unique solution is admitted by the differential equation (2.26), so that $u^i(t) = \gamma^i(t,x)$), and moreover this unique trajectory is continuous.

2.4.2 Hierarchical infinite dynamic games (Stackelberg games)

The Nash equilibrium solution concept is given heretofore, where the decision process is not dominated by a single player. Now we introduce a hierarchical equilibrium solution concept, which includes one player who is the leader of the game hence it imposes his strategy on the other player(s). The original work is given by H. Von Stackelberg (1934) [67], the leader has ability to impose its strategy in such a decision problem, and the followers react to the leader's strategy (see further details [6]).

Definition 2.4.2.1. For a Stackelberg Game, let us assume that $\gamma^{1^*} \in \Gamma^1$ is a Stackelberg strategy for the leader (P_1). An optimal strategy for the follower (P_2) includes any element $\gamma^{2^*} \in R^2(\gamma^{1^*})$ that consists of an equilibrium with γ^{1^*}. A Stackelberg solution for the game with P_1 as the leader is the pair $\{\gamma^{1^*}, \gamma^{2^*}\}$, and the cost pair $J^1(\gamma^{1^*}, \gamma^{2^*})), J^2(\gamma^{1^*}, \gamma^{2^*}))$ is the corresponding *Stackelberg equilibrium outcome*.

Assume that in infinite dynamic games with fixed duration the Stackelberg solution is derived. The number of players is limited to two in discrete time.

In case the information structure is open-loop and P_1 behaves as the leader, the Stackelberg solution for the class of two-person deterministic discrete-time infinite dynamic games of prescribed fixed duration is given. So the state progresses according to

$$x_{k+1} = f_k\big(x_k(t), u_k^1(t), u_k^2(t)\big), \quad k \in K, \tag{2.27}$$

where $x_k \in X = R^n, x_1$ is denoted a priori, $u_k^i \in U_k^i \subseteq R^{m_i} i = 1,2$; and the stage-additive cost functional for P_i is given as

$$J^i(\gamma^1, \gamma^2) \cong L^i(u^1, u^2) = \sum_{k=1}^{K} g_k^i(x_{k+1}, u_k^1, u_k^2, x_k), \qquad i = 1,2. \tag{2.28}$$

Firstly note that, the control vectors u^1, u^2 (which are of dimensions m_1K and m_2K, respectively) could lead to the cost functional $J^i(u^1, u^2)$ by recursive substitution of (2.27) into (2.28) assuming that the initial state x_1, of the global leader is known a priori. Then, if

(i) J^i is continuous on $U^1 \times U^2 (i = 1,2)$,

(ii) $J^2(u^1, \cdot)$ is strictly convex on U^2 for all $u^1 \in U^1$,

(iii) In case U^i is a closed and bounded subset of $R^{m_i}, (i = 1,2)$, a Stackelberg equilibrium solution could be obtained for the open-loop infinite game. The unique reaction curve of the follower by minimizing $J^2(u^1, u^2)$ over $u^2 \in U^2$ for every fixed $u^1 \in U^1$ is defined by a brute-force method to attain the related solution, which is a revealing optimization problem due to assumptions (i)-(iii) above. $T^2 : U^1 \to U^2$ demonstrates this unique mapping and minimization of $J^1(u^1, T^2u^1)$ over U^1 is the optimization problem faced by P_1 (the leader), hence in this open-loop game a Stackelberg strategy for the leader is provided .

Every announced strategy $\gamma^1 \in \Gamma^1$ of the leader defines the unique optimal response of the follower. Because of the fact that the information pattern is open-loop, the follower has to solve an optimization problem, defined as (for each fixed $u^1 \in U^1$)

$$\min_{u^2 \in U^2} L^2(u^1, u^2)$$

subject to

$$x_{k+1} = f_k(x_k, u_k^1, u_k^2) \qquad x_1 \text{ given.}$$

If a standard optimal control problem is unique under conditions (i)-(iii) then its solution exists.

Lemma 2.4.2.2. *Assume that additional to conditions (i)-(iii)*

(iv) $f_k(\cdot, u_k^1, u_k^2)$ *is continuously differentiable on* $R^n, (k \in K)$,

(v) $g_k(\cdot, u_k^1, u_k^2, \cdot)$ *is continuously differentiable on* $R^n \times R^n, (k \in K)$.

A unique optimal response of the follower (to be demonstrated by $\bar{\gamma}^2(x_1) = \bar{u}^2$*) exists under any explained strategy* $u^1 = \gamma^1 \in \Gamma^1$ *of the leader. Then the following relations hold:*

$$\bar{x}_{k+1} = f_k(\bar{x}_k, u_k^1, \bar{u}_k^2) \qquad \bar{x}_1 = x_1, \tag{2.29}$$

$$\bar{u}_k^2 = \arg_{u_k^2 \in U_k^2} \min H_k^2(p_{k+1}, u_k^1, u_k^2, \bar{x}_k), \tag{2.30}$$

$$\begin{aligned} p_k = \frac{\partial}{\partial x_k} f_k(\bar{x}_k, u_k^1, \bar{u}_k^2)' \Big[p_{k+1} + \frac{\partial}{\partial x_{k+1}} g_k^2\big(\bar{x}_{k+1}, u_k^1, \bar{u}_k^2, \bar{x}_k\big)' \Big] + \\ + \Big[\frac{\partial}{\partial x_k} g_k^2\big(\bar{x}_{k+1}, u_k^1, \bar{u}_k^2, \bar{x}_k\big) \Big]', \\ p_{K+1} = 0 \end{aligned} \tag{2.31}$$

$$H_k^2(p_{k+1}, u_k^1, u_k^2, \bar{x}_k) \cong g_k^2\big(f_k(x_k, u_k^1, u_k^2), u_k^1, u_k^2, x_k\big) + p_{k+1} f_k(x_k, u_k^1, u_k^2)', \quad k \in K. \tag{2.32}$$

Now, this optimal control problem associates with a sequence of n-dimensional co-state vectors $p_k, \ldots, p_{K+1}$.

Then assume that

(vi) $f_k(x_k, u_k^1, \cdot)$ *is continuously differentiable on* $U_k^2, (k \in K)$,

(vii) $g_k^2(x_{k+1}, u_k^1, \cdot, x_k, \cdot)$ *is continuously differentiable* $U_k^2, (k \in K)$,

(viii) An inner-point solution for every $u^1 \in U^1$ *is* $\bar{u}^2$ *as in Lemma 2.4.2.2, then (2.30) could in an equal manner be given as*

$$\nabla_{u_k^2} H_k^2(p_{k+1}, u_k^1, \bar{u}_k^2, \bar{x}_k) = 0, \qquad k \in K. \tag{2.33}$$

If the Stackelberg strategy of the leader is obtained, $L^1(u^1, u^2)$ *should be minimized, because of fact that (2.33), (2.29) and (2.31) describe the unique optimal response of the follower. Hence, the following optimal control problem deals with* P_1

$$\min_{u^1 \in U^1} L^1(u^1, u^2)$$

subject to

$$x_{k+1} = f_k(x_k, u_k^1, u_k^2), \qquad x_1 \text{ given.}$$

$$p_k = F_k(x_k, u_k^1, u_k^2, p_{k+1}), \qquad p_{K+1} = 0$$

$$\nabla_{u_k^2} H_k^2(p_{k+1}, u_k^1, u_k^2, x_k) = 0, \qquad k \in K. \tag{2.34}$$

where (2.32) defines the Hamiltonian and

$$F_k = \frac{\partial}{\partial x_k} f_k(x_k, u_k^1, u_k^2)' p_{k+1} + \Big[p_{k+1} + \Big[\frac{\partial}{\partial x_k} g_k^2\big(x_{k+1}, u_k^1, u_k^2, x_k\big)\Big]'. \tag{2.35}$$

Theorem 2.4.2.3. *Furthermore these conditions could be (i)-(viii) stated, so that*

(vi) $f_k(x_k, \cdot, u_k^2)$, $g_k^2(x_{k+1}, \cdot, u_k^2, x_k)$ *are continuously differentiable on* $U_k^1, (k \in K)$,

(vii) $g_k^1(\cdot, \cdot, \cdot, \cdot)$ *is continuously differentiable* $R^n \times U_k^1 \times U_k^2 \times R^n, (k \in K)$,

(viii) $f_k(\cdot, u_k^1, \cdot)$ *is twice continuously differentiable on* $R^n \times U_k^2$, *and* $g_k^2(\cdot, u_k^1, \cdot)$ *is twice continuously differentiable on* $R^n \times U_k^2 \times R^n, (k \in K)$.

If an open-loop Stackelberg equilibrium strategy for the leader is defined by $\gamma_k^{1^*}(x_1) = u_k^{1^*} \in \check{U}_k^1,\ k \in K$ in the dynamic game formulated, then the following relations are satisfied by finite vector sequences $\{\lambda_1, \dots, \lambda_K\}$, $\{\mu_1, \dots, \mu_K\}$, $\{v_1, \dots, v_K\}$:

$$x_{k+1}^* = f_k(x_k^*, u_k^{1^*}, u_k^{2^*}), \qquad x_1^* = x_1,$$

$$\nabla_{u_k^1} H_k^1\big(\lambda_k, \mu_k, v_k, p_{k+1}^*, u_k^{1^*}, u_k^{2^*}, x_k^*\big) = 0,$$

$$\nabla_{u_k^2} H_k^1\big(\lambda_k, \mu_k, v_k, p_{k+1}^*, u_k^{1^*}, u_k^{2^*}, x_k^*\big) = 0,$$

$$\lambda_{k-1}^{'} \frac{\partial}{\partial x_k} H_k^1\big(\lambda_k, \mu_k, v_k, p_{k+1}^*, u_k^{1^*}, u_k^{2^*}, x_k^*\big), \quad \lambda_K = 0,$$

$$\mu_{k+1}^{'} \frac{\partial}{\partial p_{k+1}} H_k^1\big(\lambda_k, \mu_k, v_k, p_{k+1}^*, u_k^{1^*}, u_k^{2^*}, x_k^*\big), \quad \mu_1 = 0,$$

$$\nabla_{u_k^2} H_k^2\big(p_{k+1}^*, u_k^{1^*}, u_k^{2^*}, x_k^*\big) = 0,$$

$$p_k^* = F_k(x_k^*, u_k^{1^*}, u_k^{2^*}, p_{k+1}^*), \qquad p_{k+1}^* = 0,$$

where

$$\begin{aligned} H_k^1 = g_k^1\big(f_k(x_k, u_k^1, u_k^2), u_k^1, u_k^2, x_k\big) + \lambda_k^{'} f_k(x_k, u_k^1, u_k^2) + \mu_k^{'} F_k(x_k, u_k^1, u_k^2, p_{k+1}) + \\ + v_k^{'} \nabla_{u_k^2} H_k^2\big(p_{k+1}, u_k^1, u_k^2, x_k\big)^{'}. \end{aligned} \tag{2.36}$$

3. MEAN FIELD MULTILAYER STACKELBERG DIFFERENTIAL GAMES IN MULTI-AGENT SYSTEMS

In this section, we consider a three layer version of the model of Stackelberg MFG. Hence, in our model, followers can directly be affected by their sub-leaders instead of the global leader. In addition, global leader and sub-leaders are still playing the original Stackelberg game in the infinite population limit. This model is motivated by the paper [13] in which multi-layer Stackelberg games have been studied. In order to solve the problem, we use *Stackelberg mean field* approach of [11] and *Stackelberg multi layer* solution of [13]. The section is organized as follows. In Section 3.1, we introduce the problem and explain main challenges to establish true Stackelberg equilibrium. Section 3.1.1 deals with the mean-field game of the followers. In Section 3.1.2, sub-leaders' problem is considered. Global leader problem is solved as a Stackelberg game in Section 3.1.3.

3.1 Problem Statement

An exact formulation of the multilayer Stackelberg mean field game problem is given in this section. Three hierarchical levels constitute the algorithm structure of the decision making procedure. The global leader $\mathscr{G}_0$ is in the top level, the sub-leaders $\mathscr{S}\mathscr{L}_i$ follow in the next step, which is the second. Finally N followers are in the third level. Followers are placed into groups that are connected to a specific sub-leader (see Fig. 3.1). A mean field term couples cost functions to be minimized, which belong to each player. Players at levels two and three play Nash games with the players in the same level, whereas across different levels Stackelberg game as in [13] is played.

The following controlled stochastic differential equation (SDE) denotes the state equation of the global leader,

$$dx_0(t) = \big\{A_0x_0(t) + B_0u_0(t)\big\}dt + D_0dW_0(t), \tag{3.1}$$

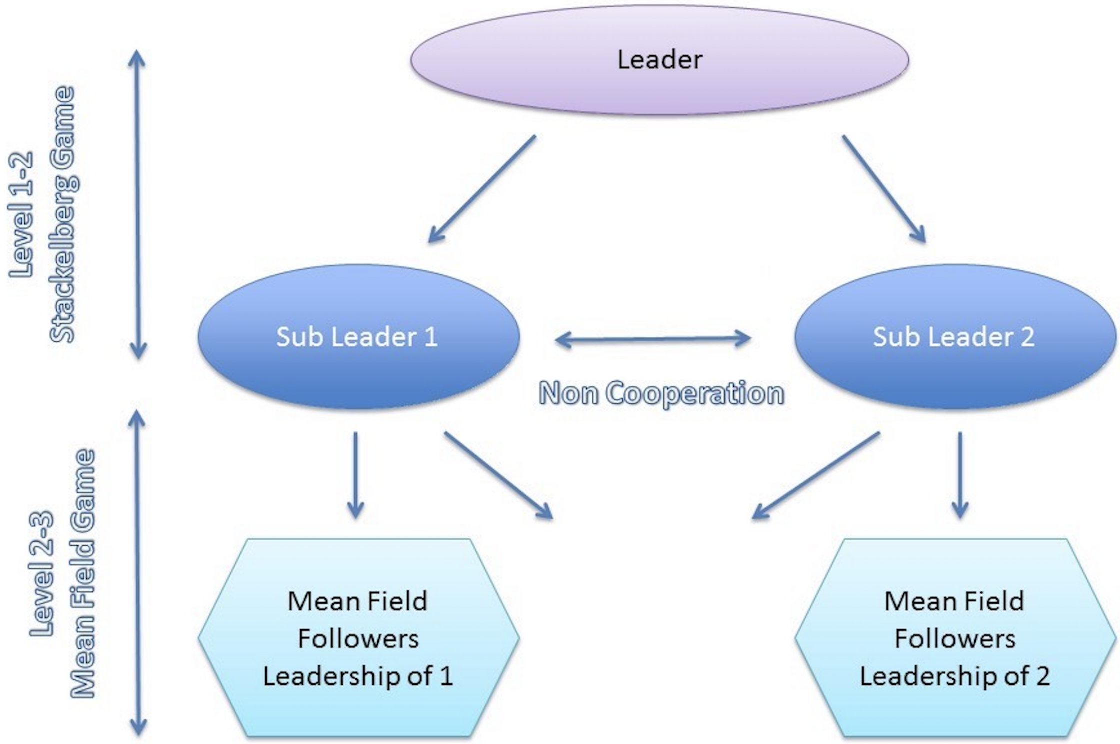

Figure 3.1 : Flow of the Stackelberg Game.

where, for each $t \geq 0$,the state of the global leader is given by $x_0(t) \in \mathscr{R}^n$, $\{W_0(t),\ t \geq 0\}$ is an n-dimensional Brownian motion, and its control is given by $u_0(t) \in \mathscr{R}^n$. Time-invariant matrices $\mathscr{G}_0$, A_0, B_0, and D_0, have appropriate dimensions. Here, $\mathscr{S}\mathscr{L}$ denotes the number of sub-leaders. Let N_j be the number of followers connected with sub-leader j ($j = 1,\ldots,\mathscr{S}\mathscr{L}$), the following SDE for $t \geq 0$ specifies the state equation of each sub-leader $\mathscr{S}\mathscr{L}_j$, with initial states $x_{lj}(0)$:

$$dx_{lj}(t) = \Big\{A_l x_{lj}(t) + B_l u_{lj}(t) + \sum_{\substack{j=1 \\ i \neq j}}^{\mathscr{S}\mathscr{L}} B_l u_{lj}(t) + F_0 u_0(t)\Big\}dt + D_l dW_{lj}(t), \qquad (3.2)$$

where, for each $t \geq 0$, the state of $\mathscr{S}\mathscr{L}_j$ is denoted by $x_{lj}(t) \in \mathscr{R}^n$, its control is given by $u_{lj}(t) \in \mathscr{R}^n$, and $\{W_{lj}(t),\ t \geq 0\}$ is an n-dimensional Brownian motion. Here, A_l, B_l, F_0, and $D_l$are the time-invariant matrices of $\mathscr{S}\mathscr{L}_j$ with appropriate dimensions.

The number of followers $\mathscr{F}_i^j$ ($1 \leq i \leq N_j$) for each sub-leader $\mathscr{SL}_j$ is denoted by N_j and their state equations defined with initial conditions $x_{fi}^{lj}(0)$, for $t \geq 0$ are given by:

$$dx_{fi}^{lj}(t) = \{A_f^{lj} x_{fi}^{lj}(t) + B_f u_{fi}^{lj}(t) + G_f u_{lj}(t)\}dt + D_f dW_{fi}^{lj}(t), \tag{3.3}$$

where, for each $t \geq 0$, the state of the follower $\mathscr{F}_{fi}^{lj}$ is denoted by $x_{fi}^{lj}(t) \in \mathscr{R}^n$, its control is $u_{fi}^{lj}(t) \in \mathscr{R}^n$, and $\{W_{fi}^{lj}(t),\ t \geq 0\}$ is an n-dimensional Brownian motion. Here, A_f^{lj}, B_f, G_f, and D_f are time-invariant matrices of $\mathscr{F}_i^j$ with appropriate dimensions. Followers may have different dynamics in each group as A_f^{lj} is indexed by j in this model. Let $\sum_{j=1}^{\mathscr{SL}} N_j = N$, where N is the sum of the followers in the game. As a consequence, if N is a variable we can accept N_j as a function of N (i.e., $N_j(N)$). We assume that for each $j = 1, \ldots, \mathscr{SL}$, $\lim_{N\to\infty} \frac{N}{N_j(N)} > 1$. Here, we are assuming that there is only a finite number of follower groups, however when the total number of followers is large, the number of followers related with each sub-leader is also large.

Let $\mathscr{F}_t$ be the σ-algebra generated by $\{x_{fi}^{lj}(0), x_{lj}(0), x_0(0);\ 1 \leq i \leq N,\ 1 \leq j \leq \mathscr{SL}\}$ and $\{W_{fi}^{lj}(\tau),\ W_{lj}(\tau),\ W_0(\tau);\ \tau \leq t,\ 1 \leq i \leq N,\ 1 \leq j \leq \mathscr{SL}\}$; that is,

$$\mathscr{F}_t = \sigma\Big(x_{fi}^{lj}(0),\ x_{lj}(0),\ x_0(0),\ W_{fi}^{lj}(\tau), W_{lj}(\tau),\ W_0(\tau); \tau \leq t,\ 1 \leq i \leq N,\ 1 \leq j \leq \mathscr{SL}\Big).$$

Let the centralized adapted open-loop information be $\mathscr{F}_t$ in the game model. We further state $\mathscr{F}_t^{i,j} = \sigma(x_{fi}^{lj}(0),\ W_{fi}^{lj}(\tau);\ \tau \leq t)$ as the local information of the i^{th} follower of the j^{th} sub-leader, which is also adapted open-loop, but decentralized.

Assumption

(a) $x_0(0)$, $\{x_{lj}(0);\ 1 \leq j \leq \mathscr{SL}\}$, and $\{x_{fi}^{lj}(0);\ 1 \leq i \leq N_j,\ 1 \leq j \leq \mathscr{SL}\}$ are independent of each other and have bounded second moments.

(b) $\{W_0(t); t \geq 0\}$, $\{W_{lj}(t);\ 1 \leq j \leq \mathscr{SL}, t \geq 0\}$, and $\{W_{fi}^{lj}(t);\ 1 \leq i \leq N_j,\ 1 \leq j \leq \mathscr{SL}, t \geq 0\}$ are independent of each other, which are also independent of $x_0(0)$, $\{x_{lj}(0);\ 1 \leq j \leq \mathscr{SL}\}$, and $\{x_{fi}^{lj}(0);\ 1 \leq i \leq N_j,\ 1 \leq j \leq \mathscr{SL}\}$.

The cost function for $\mathscr{G}_0$, to be minimized, is denoted by:

$$J_0^N(u_0, u^N) = E\int_0^T \Big\{\|x_0(t) - H_0 x^N(t)\|_{Q_0}^2 + \|u_0(t)\|_{R_0}^2\Big\}dt, \tag{3.4}$$

where $Q_0 > 0$, $H_0 > 0$, $R_0 > 0$, and $\|\cdot\|_Q$ specifies the weighted Euclidean norm. Here, $x^N(t)$ is the mean-field term denoted by

$$x^N(t) = \frac{1}{N}\sum_{j=1}^{\mathscr{SL}}\sum_{i=1}^{N_j} x_{fi}^{lj}(t).$$

Let $x^N(t)$ demonstrate the mass behavior of *all* followers. Here, the collection of policies of all sub-leaders and followers is given by u^N in the game; that is, $u^N = \{u_{lj}, u_{fi}^{lj}; 1 \leq i \leq N, 1 \leq j \leq \mathscr{SL}\}$. The penalty on tracking error is denoted by the first term in the cost function J_0^N, and the penalty on the control effort is given by the second term.

The cost function for $\mathscr{SL}_j$, to be minimized, is denoted by:

$$\begin{aligned} J_{lj}^N(u_{lj}, u_0, u_{-lj}^N) = E\int_0^T \Big\{ &\|x_{lj}(t) - H_l x_{lj}^N(t)\|_{Q_l}^2 + \|u_{lj}(t)\|_{R_l}^2 + \\ &2\sum_{\substack{k=1\\k\neq j}}^{\mathscr{SL}} u_{lk}^T(t) R_l u_{lk}(t) + 2u_{lj}^T(t) L_0 u_0(t)\Big\} dt, \end{aligned}$$

where $Q_l > 0, H_l > 0$, $R_l > 0$, $L_0 > 0$, and $u_{-lj}^N(t) := u^N(t) \setminus \{u_{lj}(t)\}$. Here, the mean field term related with sub-leader $\mathscr{SL}_j$ is x_{lj}^N and is denoted by

$$x_{lj}^N(t) = \frac{1}{N_j}\sum_{i=1}^{N_j} x_{fi}^{lj}(t).$$

Let the mass behavior of followers of $\mathscr{SL}_j$ be demonstrated by $x_{lj}^N(t)$. Here, the penalty on the control effort of other sub-leaders and the global-leader are given by the last two terms in the cost function J_{lj}^N. Assume that the mean field term couples weakly the sub-leaders with all the followers, and they are strongly coupled with each other and the global leader through their controls.

The cost function for $\mathscr{F}_{fi}^{lj}$, to be minimized, is denoted by:

$$J_{fi,lj}^N(u_{fi}^{lj}, u_0, u_{-(fi,lj)}^N) = E\int_0^T \Big\{ \|x_{fi}^{lj}(t) - H_f x_{lj}^N(t)\|_{Q_f}^2 + \|u_{fi}^{lj}(t)\|_{R_f}^2 + 2u_{fi}^{lj^T}(t) L_l u_{lj}(t)\Big\} dt, \tag{3.5}$$

where $Q_f > 0$, $H_f > 0$, $R_f > 0$, and $L_l > 0$. Assume that the followers, that are in the same group are coupled weakly with each other through their mean field term $x_{lj}^N(t)$ and they are strongly coupled with their sub-leader through its control u_{lj}.

As we addressed earlier, the adapted open-loop information structure considered in this game model is first centralized. Therefore, the classes of admissible controls for

$\mathscr{G}_0$, $\mathscr{SL}_i$ and $\mathscr{F}_{fi}^{lj}$ are $L^2_{\mathscr{F}}(0,T;R^n)$; that is, the set of all $\mathscr{F}_t$ adapted processes with finite second moments.

Searching for the precise Stackelberg solution for a large number of agents, we encounter two challenges. The first one is that when the number of followers N is large (i.e., *curse of dimensionality*), Nash equilibria between followers is hard to calculate for any arbitrary strategy of the leaders. Second one is that the sub-leaders encounter a large number of constraints in their game problem when the number of followers increases. As a result, it would be impossible to identify the precise solution and this leads us to seek for an approximated equilibrium. Consequently, mean field approach will be devised to compute the approximated equilibrium; that is, the mass behavior of the followers will be incorporated into the limiting mean field term, and so, all of the followers in each group can be represented by a generic agent which will decrease the complexity of the solution exceedingly. In this way, instead of using centralized information, only local information can be used for the players.

3.1.1 Stochastic mean field approximation for followers

In this part, the mean field approximation of Nash equilibrium for the game played among N_j followers of the j^{th} sub-leader $\mathscr{SL}_j$ is considered. Followers are linked with the players in their group through the mean field term $x_{lj}^N(t)$ and they are linked with followers in other groups via controls of their sub-leaders. Under the mean field approach, given the rules of global leader and sub-leaders, a generic agent of each group is faced with a stochastic control problem under a constraint on the expectation of its state; that is, the mean field term $x_{lj}^N(t)$ is replaced by a deterministic quantity $z_{lj}(t)$ that can be accepted as the limit of the $x_{lj}^N(t)$ by the law of large numbers. Consider that when a generic agent acts optimally, the expectation of its state should be the same with $z_{lj}(t)$ due to the Nash certainty equivalence (NCE) principle. The optimal decision rule for the non-standard stochastic control problem is identified by using the stochastic minimum principle.

Let the mean field term be denoted by $z_{lj}(t)$ $(1 \leq j \leq \mathscr{SL})$ in the infinite population limit $N_j \to \infty$. Hence, in the infinite population limit, the control u_{lj} and the mean field term $z_{lj}(t)$ of the sub-leader can be considered as exogenous signals in the cost function of a generic follower. Consequently, a generic follower has to solve the following stochastic control problem:

Minimize

$$J_{fi,lj}(u_{fi}^{lj},u_{lj}) = E\int_0^T \Big\{\|x_{fi}^{lj}(t)-H_f z_{lj}(t)\|_{Q_f}^2 + \|u_{fi}^{lj}(t)\|_{R_f}^2 + 2u_{fi}^{lj^T}(t)L_l u_{lj}(t)\Big\}dt, \tag{3.6}$$

subject to

$$dx_{fi}^{lj}(t) = \Big\{A_f^{lj}x_{fi}^{lj}(t) + B_f u_{fi}^{lj}(t) + G_f u_{lj}(t)\Big\}dt + D_f dW_{fi}^{lj}(t).$$

The stochastic minimum principle could be utilized to characterize the optimal policy (see [66, Chapter 3, Theorem 3.2]) due to fact that this is a classical linear-quadratic optimal stochastic control problem (with adapted open-loop information). To this end, Hamiltonian for this problem is denoted by:

$$\begin{aligned} H_{fi}^{lj}(x_{fi}^{lj},u_{fi}^{lj},p_{fi}) = {}& \frac{1}{2}\|x_{fi}^{lj}(t)-H_f z_{lj}(t)\|_{Q_f}^2 + \frac{1}{2}\|u_{fi}^{lj}(t)\|_{R_f}^2 + u_{fi}^{lj^T}(t)L_l u_{lj}(t) + Tr[r_{fi}^T(t)D_f] \\ & + p_{fi}^T(t)\big[A_f^{lj}x_{fi}^{lj}(t) + B_f u_{fi}^{lj}(t) + G_f u_{lj}(t)\big]. \end{aligned} \tag{3.7}$$

The corresponding state and adjoint equations and the (unique) optimal control are denoted by:

$$u_{fi}^{lj^*}(t) = -R_f^{-1}B_f^T p_{fi}(t) - R_f^{-1}L_l u_{lj}(t), \tag{3.8}$$

$$\begin{aligned} dx_{fi}^{lj^*}(t) &= \Big\{A_f^{lj}x_{fi}^{lj^*}(t) - B_f R_f^{-1}B_f^T p_{fi}(t) - \big(B_f R_f^{-1}L_l - G_f\big)u_{lj}(t)\Big\}dt + D_f dW_{fi}^{lj}(t), \\ dp_{fi}(t) &= \Big\{-A_f^{lj^T}p_{fi}(t) - Q_f\big[x_{fi}^{lj^*}(t) - H_f z_{lj}(t)\big]\Big\}dt + r_{fi}(t)dW_{fi}^{lj}(t) + r_{li}(t)dW_{li}(t), \end{aligned} \tag{3.9}$$

subject to the terminal condition on the adjoint variable, and the given initial condition on x_{fi}^{lj}

$$x_{fi}^{lj}(t)\Big|_{t=0} = x_{fi}^{lj}(0) \qquad p_{fi}(T) = 0.$$

Equation (3.9) is a forward-backward stochastic differential equation (FBSDE). Some simplifications may be appropriate because of the fact that $x_{fi}^{lj^*}(t)$ and $p_{fi}(t)$ satisfy stochastic coupled differential equations. (3.9) can be changed into a different form involving a Riccati type differential equation (RDE) by utilizing Ito's Lemma. In this way, $dx_{fi}^{lj*}(t)$ and $dp_{fi}(t)$ are given in (3.9) could be transformed into a different form and are decoupled. The equation could be transformed into a different form. Solving this Riccati equation is sufficient to obtain the solution of (3.9).

Theorem 3.1.1.1. *Given $z_{lj}(t)$ and $u_{lj}(t) \in L^2_{\mathscr{F}}(0, \mathscr{T}; \mathscr{R}^n)$, consider the local optimal control problem for $\mathscr{F}^{lj}_{fi}$. There exists a unique optimal controller u_{fi}^{lj*}. Moreover, $(x_{fi}^{lj*},\ u_{fi}^{lj*}) \in L^2_{\mathscr{F}}(0, \mathscr{T}; \mathscr{R}^n) \times L^2_{\mathscr{F}}(0, \mathscr{T}; \mathscr{R}^n)$ is the corresponding optimal control solution if and only if*

$$u_{fi}^{lj*}(t) = -R_f^{-1}B_f^T Z_{fi}(t)x_{fi}^{lj*}(t) - R_f^{-1}B_f^T\Phi_{fi}(t) - R_f^{-1}L_l u_{lj}(t), \tag{3.10}$$

where

$$\begin{aligned} dx_{fi}^{lj*}(t) = &\Big\{ \Big[A_f^{lj} - B_f R_f^{-1}B_f^T Z_{fi}(t)\Big] x_{fi}^{lj*}(t) - B_f R_f^{-1}B_f^T\Phi_{fi}(t) - \big(B_f R_f^{-1}L_l - G_f\big)u_{lj}(t)\Big\}dt \\ &+ D_f dW_{fi}^{lj}(t), \end{aligned} \tag{3.11}$$

$$\begin{aligned} d\Phi_{fi}(t) = &\Big\{ \Big[-A_f^{lj} + Z_{fi}(t)B_f R_f^{-1}B_f^T\Big]\Phi_{fi}(t) + Q_f H_f z_{lj}(t) - Z_{fi}(t)\big(B_f R_f^{-1}L_l - G_f\big)u_{lj}(t)\Big\}dt \\ &+ \{r_{fi}(t) - Z_{fi}(t)D_f\}dW_{fi}^{lj}(t) + r_{li}(t)dW_{li}(t), \end{aligned} \tag{3.12}$$

$$x_{fi}^{lj}(t)\Big|_{t=0} = x_{fi}^{lj}(0), \qquad \Phi_{fi}(T) = 0,$$

and the solution of the RDE is $Z_{fi}(t)$

$$\begin{aligned} \dot{Z}_{fi}(t) + A_f^{lj^T} Z_{fi}(t) + Z_{fi}(t)A_f^{lj} - Z_{fi}(t)B_f R_f^{-1}B_f^T Z_{fi}(t) + Q_f = 0, \\ Z_{fi}(T) = 0. \end{aligned} \tag{3.13}$$

Proof. The proof is given in Appendix A. □

Note that $\Phi_{fi}(t)$ is now decoupled from $x_{fi}^{lj^*}(t)$ and also there always exists a solution $Z_{fi}(t)$ for the RDE in (3.13). Therefore, one can compute the optimal policies by solving the decoupled SDE's in (3.11) and (3.12). Finally, for given z_{lj} and u_{lj}, $\left(x_{fi}^{lj^*}, p_{fi}, r_{fi}, r_{li}\right)$ has a unique solution in $\mathscr{L}_{\mathscr{F}}^{2}\left(0,T;\mathscr{R}^{2n},\mathscr{R}^{n\times q},\mathscr{R}^{n\times q}\right)$.

Consider that in the optimal control problem for a generic agent given above, the mean field term $z_{lj}(t)$ is accepted as an external variable. According to the Nash certainty equivalence principle (NCE), the state evolution of a generic agent should be consistent with the mean field term under the optimal control rule (i.e., the total population behavior of the followers). This means that the expectation of the state of a generic agent must be equal to the mean field term $z_{lj}(t)$. As a consequence, we denote the expectation of the state under optimal rule and the corresponding co-state as $z_{lj}(t)$ and $p_{mi}(t)$, respectively. With this notation, the coupled differential equations describing the evolution of the mean field term can be obtained by evaluating the expectations of the coupled SDE's in (3.9); that is,

$$\begin{aligned} dz_{lj}(t) &= \Big\{A_f^{lj} z_{lj}^*(t) - B_f R_f^{-1} B_f^T p_{mi}(t) - \left(B_f R_f^{-1} L_l - G_f\right) u_{lj}(t)\Big\} dt, \\ dp_{mi}(t) &= \Big\{-A_f^{lj^T} p_{mi}(t) - Q_f\left[z_{lj}(t) - H_f z_{lj}(t)\right]\Big\} dt, \end{aligned} \tag{3.14}$$

subject to the boundary conditions:

$$z_{lj}(t)\Big|_{t=0} = \lim_{N_j\to\infty} \frac{1}{N_j}\sum_{i=1}^{N_j} x_{fi}^{lj}(0), \qquad p_{mi}(T) = 0. \tag{3.15}$$

It must be mentioned that due to fact that the mean field term affects the cost functions of sub-leaders and global leader, the coupled differential equation given in (3.15) is a constraint between sub-leaders and the global leader in the game problem. Nevertheless, concatenating the state of the leaders and $\left(z_{lj}(t), p_{mi}(t)\right)$ could eliminate this constraint, and the Hamiltonian can be rewritten by this new state representation.

3.1.2 Stochastic mean field approximation for sub leaders

In this section we consider the Nash game played among sub-leaders given the limiting mean field term and the policy of the global leader. Recall that each sub-leader plays a Stackelberg mean field game with its followers.

We let $z_{li}(t)$ denote the limiting value of the mean field term in the i^{th} group for the infinite population case. Note that the mean field term $z_{li}(t)$, controls of other sub-leaders $\{u_{lj}(t)\}_{j\neq i}$, and control of global leader $u_0(t)$ can be taken as exogenous terms in i^{th} sub-leader's local control problem. Hence, i^{th} sub-leader is faced with the following control problem:

Minimize

$$J_{li}(u_{li}, u_0, u_{-li}^{\mathscr{S}\mathscr{L}}) = E\int_0^T \Big\{\|x_{li}(t) - H_l z_{li}(t)\|_{Q_l}^2 + \|u_{li}(t)\|_{R_l}^2 + 2\sum_{\substack{j=1\\ i\neq j}}^{\mathscr{S}\mathscr{L}} u_{lj}^T(t) R_l u_{li}(t)$$
$$+ 2u_{li}^T(t) L_0 u_0(t)\Big\} dt, \tag{3.16}$$

where $u_{-li}^{\mathscr{S}\mathscr{L}} = (u_{lj})_{j\neq i}$, subject to state dynamics of the sub-leader, and dynamics of the mean field term and its co-state (3.14), where the former is rewritten as:

$$dx_{li}(t) = \{A_l x_{li}(t) + B_l u_{li}(t) + \sum_{\substack{j=1\\ i\neq j}}^{\mathscr{S}\mathscr{L}} B_l u_{lj}(t) + F_0 u_0(t)\} dt + D_l dW_{li}(t).$$

The above optimal stochastic control problem is again solved using the stochastic minimum principle [66, Chapter 3, Theorem 3.2] (see also [68]). To this end, Hamiltonian of the control problem is given by:

$$H_{li}(x_{li}, u_{li}, p_{li}, \gamma_{li}, \beta_{li}) = \frac{1}{2}\|x_{li}(t) - H_l z_{li}(t)\|_{Q_l}^2 + \frac{1}{2}\|u_{li}(t)\|_{R_l}^2 + \sum_{\substack{j=1\\ i\neq j}}^{\mathscr{S}\mathscr{L}} u_{lj}^T(t) R_l u_{li}(t) +$$
$$u_{li}^T(t) L_0 u_0(t) + Tr[r_{li}^T(t) D_l] + p_{li}^T(t)\Big[A_l x_{li}(t) + B_l u_{li}(t) + \sum_{j=1 i\neq j}^{\mathscr{S}\mathscr{L}} B_l u_{lj}(t) + F_0 u_0(t)\Big] +$$
$$\gamma_{li}^T(t)\Big[A_f^{li} z_{li}(t) - B_f R_f^{-1} B_f^T p_{mi}(t) - \big(B_f R_f^{-1} L_l - G_f\big) u_{li}(t)\Big] + \beta_{li}^T(t)\Big[-A_f^{lj^T} p_{mi}(t) -$$
$$Q_f\left[z_{li}(t) - H_f z_{li}(t)\right]\Big].$$

Here, $\gamma_{li}(t)$ and β_{li} are adjoint functions, which correspond to the DEs satisfied by $z_{li}(t)$ and $p_{mi}(t)$, in (3.14), respectively. The (unique) optimal control and the corresponding state and adjoint equations are given as follows:

$$u_{li}^*(t) = -R_l^{-1}B_l^T p_{li}(t) - \sum_{\substack{j=1\\ i\neq j}}^{\mathscr{S}\mathscr{L}} R_l^{-1}R_l^T u_{lj}(t) - R_l^{-1}L_0 u_0^*(t) + R_l^{-1}\big(B_f R_f^{-1}L_l - G_f\big)^T \gamma_{li}(t), \tag{3.17}$$

$$\begin{aligned} dx_{li}^*(t) =& \Big\{ A_l x_{li}^*(t) - B_l R_l^{-1}B_l^T p_{li}(t) + \sum_{\substack{j=1\\ i\neq j}}^{\mathscr{S}\mathscr{L}} B_l\big(I_n - R_l^{-1}R_l^T\big)u_{lj}(t) + \Big(F_0 - B_l R_l^{-1}L_0\Big)u_0(t) \\ & + B_l R_l^{-1}\big(B_f R_f^{-1}L_l - G_f\big)^T \gamma_{li}(t)\Big\}dt + D_l dW_l(t), \\ dp_{li}(t) =& \Big\{ -A_l^T p_{li}(t) - Q_l[x_{li}^*(t) - H_l z_{li}^*(t)]\Big\}dt + r_{li}(t)dW_l(t) + r_0(t)dW_0(t), \end{aligned} \tag{3.18}$$

subject to the boundary conditions:

$$x_{li}(t)\Big|_{t=0} = x_{li}(0), \qquad p_{li}(T) = 0.$$

Equations describing the dynamics of state and co-state for the mean field term are given by:

$$\begin{aligned} dz_{li}(t) =& \Big\{ A_f^{li} z_{li}(t) - B_f R_f^{-1}B_f^T p_{mi}(t) + \Big(B_f R_f^{-1}L_l R_l^{-1}B_l^T - G_f R_l^{-1}B_l^T\Big)p_{li}(t) \\ & + \sum_{\substack{j=1\\ i\neq j}}^{\mathscr{S}\mathscr{L}} \Big(B_f R_f^{-1}L_l R_l^{-1}R_l^T - G_f R_l^{-1}R_l^T\Big)u_{lj}(t) + \Big(B_f R_f^{-1}L_l R_l^{-1}L_0 - G_f R_l^{-1}L_0\Big)u_0(t) \\ & - \big(B_f R_f^{-1}L_l - G_f\big)R_l^{-1}\big(B_f R_f^{-1}L_l - G_f\big)^T \gamma_{li}(t)\Big\}dt, \\ d\gamma_{li}(t) =& \Big\{ -A_f^{li^T}\gamma_{li}(t) - H_l^T Q_l[x_{li}^*(t) - H_l z_{li}^*(t)] + (I_n - H_f)^T Q_f^T \beta_{li}(t)\Big\}dt, \end{aligned} \tag{3.19}$$

subject to the boundary conditions:

$$z_{li}(0) = \lim_{N_i\to\infty} \frac{1}{N_i}\sum_{j=1}^{N_i} x_{fj}^{li}(0), \qquad \gamma_{li}(T) = 0.$$

The Hamiltonian also contains mean field adjoint function p_{mi} which has its own dynamics, so we also need the associated adjoint function dynamics to complete the necessary conditions. These dynamics are:

$$dp_{mi}(t) = \Big\{ -A_f^{li} p_{mi}(t) - Q_f \big[z_{li}^*(t) - H_f z_{li}^*(t) \big] \Big\} dt,$$
$$d\beta_{li}(t) = \Big\{ A_f^{li^T} \beta_{li}(t) + B_f R_f^{-1} B_f^T \gamma_{li}(t) \Big\} dt, \tag{3.20}$$

subject to the boundary conditions:

$$\beta_{li}(t)\Big|_{t=0} = 0, \qquad p_{mi}(T) = 0.$$

These equations will yield the optimal solution for the i^{th} sub-leader given controls of other sub-leaders $\{u_{lj}(t)\}_{j \neq i}$, and control of global leader $u_0(t)$. To find the equilibrium solution for the sub-leaders, these equations must be expressed in matrix form. We first write equations (3.18), (3.19) and (3.20) for all sub-leaders as a coupled matrix FBSDE (see (3.21) below) to ease the notation. Then, the equation (3.21) is transformed to a different form which involves a Riccati type equation using Ito's Lemma. Solving this Riccati equation is sufficient for obtaining the solution of FBSDE (3.21), and for characterizing the equilibrium solution.

Let us set

$$\mathscr{X}_l(t) = [x_{li}^*(t), \ldots, x_{l\mathscr{S}\mathscr{L}}^*(t), z_{li}^*(t), \ldots, z_{l\mathscr{S}\mathscr{L}}^*(t), \beta_{li}(t), \ldots, \beta_{l\mathscr{S}\mathscr{L}}(t)]^T,$$
$$\mathscr{Y}_l(t) = [p_{li}(t), \ldots, p_{l\mathscr{S}\mathscr{L}}(t), p_{mi}(t), \ldots, p_{m\mathscr{S}\mathscr{L}}(t), \gamma_{li}(t), \ldots, \gamma_{l\mathscr{S}\mathscr{L}}(t)]^T.$$

The overall coupled FBSDE that yields the equilibrium solution for sub-leaders can be compactly written as (see Appendix A for the construction of the matrices in the equation below):

$$d\mathscr{X}_l(t) = [\mathscr{A}_{l1}\mathscr{X}_l(t) + \mathscr{B}_{l1}\mathscr{Y}_l(t) + \mathscr{C}_{l1}]dt + \mathscr{D}_{l1} d\mathscr{W}_l(t) + \mathscr{D}_{l3} d\mathscr{W}_0(t),$$
$$d\mathscr{Y}_l(t) = [\mathscr{A}_{l2}\mathscr{X}_l(t) + \mathscr{B}_{l2}\mathscr{Y}_l(t)]dt + \mathscr{D}_{l2} d\mathscr{W}_l(t) + \mathscr{D}_{l4} d\mathscr{W}_0(t). \tag{3.21}$$

subject to the boundary conditions:

$$\mathscr{X}_l(t)\Big|_{t=0} = \mathscr{X}_l(0), \qquad \mathscr{Y}_l(T) = 0.$$

Theorem 3.1.2.1. *Given the control of global leader* $u_0(t)$*, consider the game problem for sub-leaders* $\mathscr{SL}_i$*,* $1 \le i \le \mathscr{SL}$*. There exists a unique equilibrium solution* $\{u^*_{li}\}_{i=1}^{\mathscr{SL}}$*. Moreover* $\{(x^*_{li},\ u^*_{li})\}_{i=1}^{\mathscr{SL}}$ *is the corresponding equilibrium solution if and only if for all* $i = 1, \ldots, \mathscr{SL}$*,*

$$u^*_{li}(t) = -R_l^{-1} B_l^T p_{li}(t) - \sum_{\substack{j=1 \\ i \neq j}}^{\mathscr{SL}} R_l^{-1} R_l^T u_{lj}(t) - R_l^{-1} L_0 u^*_0(t) + \left(R_l^{-1} B_f R_f^{-1} L_l - R_l^{-1} G_f\right)^T \gamma_{li}(t),$$

where

$$d\mathscr{X}_l(t) = \Big\{\mathscr{A}_{l1}\mathscr{X}_l(t) + \mathscr{B}_{l1}\Lambda_l(t)\mathscr{X}_l(t) + \mathscr{B}_{l1}\mathscr{V}_l(t) + \mathscr{C}_{l1}\Big\}dt - d\mathscr{W}_l(t), \tag{3.22}$$

$$d\mathscr{V}_l(t) = \Big\{\mathscr{B}_{l2}\mathscr{V}_l(t) - \Lambda_l^T(t)\mathscr{B}_{l1}\mathscr{V}_l(t) - \Lambda_l^T(t)\mathscr{C}_{l1}\Big\}dt + \big\{\mathscr{D}_{l2} - \Lambda_l^T\big\}d\mathscr{W}_l(t) + \mathscr{D}_{l4}d\mathscr{W}_0(t), \tag{3.23}$$

$$\mathscr{X}_l(t)\Big|_{t=0} = \mathscr{X}_l(0), \qquad \mathscr{V}_l(T) = 0,$$

and $\Lambda_l(t)$ *is the solution of the following RDE*

$$\begin{aligned} \dot{\Lambda}_l(t) + \Lambda_l^T(t)\mathscr{A}_{l1} - \mathscr{B}_{l2}\Lambda_l(t) - \Lambda_l^T(t)\mathscr{B}_{l1}\Lambda_l(t) - \mathscr{A}_{l2} = 0, \\ \Lambda_l(T) = 0, \end{aligned} \tag{3.24}$$

under the assumption that

$$\det\left\{ \begin{bmatrix} 0 & I_n \end{bmatrix} e^{\mathscr{A}_l t} \begin{bmatrix} I_n \\ 0 \end{bmatrix} \right\} > 0, \quad \forall t \in (0, T),$$

where the square matrix $\mathscr{A}_l$ *is given by*

$$\mathscr{A}_l = \begin{bmatrix} \mathscr{A}_{l1} & \mathscr{B}_{l1} \\ \mathscr{A}_{l2} & \mathscr{B}_{l2} \end{bmatrix}.$$

Proof. The proof is given in Appendix A. □

First (3.24) must be solved for Λ_l, then u_{li}^* must be computed by solving decoupled equations (3.22) and (3.23). Further details and solution of RDE is available in [65].

Remark 3.1.2.2. RDE is a non symmetric equation, so $\mathscr{A}_l$ must satisfy the above condition. Radon's Lemma is used for solving $\Lambda_l(t)$, see [65, (Chapter 2, Theorem 4.3)].

3.1.3 Stochastic mean field approximation for the global leader

In this section, the solution for the Stackelberg game between the global leader and the sub-leaders is derived. Sub-leaders apply the equilibrium solution obtained in the previous section given the control of global leader as an exogenous signal. The best control for the global leader which will eventually yield the (approximate) Stackelberg equilibrium is computed using the equilibrium solution that is derived in the previous section. It must be mentioned that the global leader can not directly affect the followers but can reach them only through sub-leaders.

Given the mean field term $z(t)$, which is the limit of $x^N(t)$ for infinite population, the global leader has to solve the following problem:

Minimize

$$J_0(u_0) = E\int_0^T \left\{\|x_0(t) - H_0 z(t)\|^2_{Q_0} + \|u_0(t)\|^2_{R_0}\right\}dt, \tag{3.25}$$

subject to

$$dx_0(t) = \left\{A_0 x_0(t) + B_0 u_0(t)\right\}dt + D_0 dW_0(t).$$

Global leader problem is a stochastic Stackelberg game with one leader and many followers (i.e. sub-leaders). The mean field term affects the cost function of the global leader, therefore the corresponding Hamiltonian equation should contain the dynamics of the global-leader as well as the dynamic equations of each sub-leader and mean field term. The optimal control of the global leader is obtained by the stochastic minimum principle [66, Chapter 3, Theorem 3.2] (see also [68]). The Hamiltonian of the local control problem for the global leader is obtained as:

$$H_0\big(x_0,u_0\big) = \frac{1}{2}\big\|x_0(t)-H_0z(t)\big\|_{Q_0}^2 + \frac{1}{2}\|u_0(t)\|_{R_0}^2 + Tr[r_0^T(t)D_0] + p_0^T(t)[A_0x_0(t)+B_0u_0(t)] + \gamma_0^T(t)$$
$$\Big[A_fz(t) - B_fR_f^{-1}B_f^Tp_{m0}(t) - \big(B_fR_f^{-1}L_l - G_f\big)u_0(t)\Big] + \beta_0^T(t)\Big[-A_f^Tp_{m0}(t) - Q_f\big[z(t)-H_fz(t)\big]\Big]$$
$$+\sum_{i=1}^{\mathscr{S}\mathscr{L}}\alpha_{0li}^T(t)\Big[A_lx_{li}(t) - B_lR_l^{-1}B_l^Tp_{li}(t) + \sum_{\substack{j=1\\ i\neq j}}^{\mathscr{S}\mathscr{L}}B_l\big(I_n - R_l^{-1}R_l^T\big)u_{lj}(t) + \Big(F_0 - B_lR_l^{-1}L_0\Big)u_0(t)$$
$$+B_lR_l^{-1}\big(B_fR_f^{-1}L_l - G_f\big)^T\gamma_{li}(t)\Big] + \sum_{i=1}^{\mathscr{S}\mathscr{L}}\delta_{0li}^T(t)\big[-A_l^Tp_{li}(t) - Q_l[x_{li}(t)-H_lz_{li}(t)]\big] + \sum_{i=1}^{\mathscr{S}\mathscr{L}}\beta_{0li}^T(t)$$
$$\Big[-A_f^{li^T}p_{mi}(t) - Q_f\big[z_{li}(t)-H_fz_{li}(t)\big]\Big] + \sum_{i=1}^{\mathscr{S}\mathscr{L}}\gamma_{0li}^T(t)\Big[A_f^{li}z_{li}(t) - B_fR_f^{-1}B_f^Tp_{mi}(t) + \Big(B_fR_f^{-1}L_lR_l^{-1}B_l^T$$
$$-G_fR_l^{-1}B_l^T\Big)p_{li}(t) + \sum_{\substack{j=1\\ i\neq j}}^{\mathscr{S}\mathscr{L}}\Big(B_fR_f^{-1}L_lR_l^{-1}R_l^T - G_fR_l^{-1}R_l^T\Big)u_{lj}(t) + \Big(B_fR_f^{-1}L_lR_l^{-1}L_0 - G_fR_l^{-1}L_0\Big)$$
$$u_0(t) - \big(B_fR_f^{-1}L_l - G_f\big)R_l^{-1}\big(B_fR_f^{-1}L_l - G_f\big)^T\gamma_{li}(t)\Big] + \sum_{i=1}^{\mathscr{S}\mathscr{L}}\eta_{0li}^T(t)\Big[A_f^{li^T}\beta_{li}(t) + B_fR_f^{-1}B_f^T\gamma_{li}(t)\Big]$$
$$+\sum_{i=1}^{\mathscr{S}\mathscr{L}}\zeta_{0li}^T(t)\Big[-A_f^{li^T}\gamma_{li}(t) - H_l^TQ_l[x_{li}(t)-H_lz_{li}(t)] + (I_n-H_f)^TQ_f^T\beta_{li}(t)\Big]. \tag{3.26}$$

Here, $\gamma_0(t)$, β_0, $\alpha_{0li}(t)$, $\delta_{0li}(t)$, β_{0li}, $\gamma_{0li}(t)$, $\eta_{0li}(t)$ and $\zeta_{0li}(t)$ are adjoint functions, which are used for finding extremal trajectories of $z(t)$, $p_{m0}(t)$, $x_{li}(t)$, $p_{li}(t)$, $p_{mi}(t)$, $z_{li}(t)$, $\beta_{li}(t)$ and $\gamma_{li}(t)$, respectively. The state transition matrix of all followers is given by $A_f = \dfrac{1}{\mathscr{S}\mathscr{L}}\sum_{j=1}^{\mathscr{S}\mathscr{L}}A_f^{lj}$. The (unique) optimal control and the corresponding state and adjoint equations are obtained as follows:

$$u_0^*(t) = -R_0^{-1}B_0^Tp_0(t) + R_0^{-1}\big(B_fR_f^{-1}L_l - G_f\big)^T\gamma_0(t) - \sum_{i=1}^{\mathscr{S}\mathscr{L}}R_0^{-1}\big(B_fR_f^{-1}L_lR_l^{-1}L_0 - G_fR_l^{-1}L_0\big)^T\gamma_{0li}(t)$$
$$-\sum_{i=1}^{\mathscr{S}\mathscr{L}}R_0^{-1}\big(F_0 - B_lR_l^{-1}L_0\big)^T\alpha_{0li}(t),$$

$$\begin{aligned}dx_0^*(t) =&\Big\{A_0x_0^*(t) - B_0R_0^{-1}B_0^Tp_0(t) - \sum_{i=1}^{\mathscr{S}\mathscr{L}}B_0R_0^{-1}\big(B_fR_f^{-1}L_lR_l^{-1}L_0 - G_fR_l^{-1}L_0\big)^T\gamma_{0li}(t)\\ &+B_0R_0^{-1}\big(B_fR_f^{-1}L_l - G_f\big)^T\gamma_0(t) - \sum_{i=1}^{\mathscr{S}\mathscr{L}}B_0R_0^{-1}\big(F_0 - B_lR_l^{-1}L_0\big)^T\alpha_{0li}(t)\Big\}dt + D_0dW_0(t),\\ dp_0(t) =&\Big\{-A_0^Tp_0(t) - Q_0[x_0^*(t) - H_0z^*(t)]\Big\}dt + r_0(t)dW_0(t).\end{aligned} \tag{3.27}$$

These stochastic coupled differential equations are valid for the sub-leaders. The boundary conditions are:

$$x_0(t)\Big|_{t=0} = x_0(0), \qquad p_0(T) = 0.$$

Using (3.26), the states and adjoint dynamics of sub-leaders are obtained as follows:

$$\begin{aligned}
dx_{li}^*(t) = \Big\{ & A_l x_{li}^*(t) - \Big(F_0 - B_l R_l^{-1} L_0\Big) R_0^{-1} B_0^T p_0(t) + \sum_{\substack{j=1 \\ i \neq j}}^{\mathscr{S}\mathscr{L}} B_l \big(I_n - R_l^{-1} R_l^T\big) u_{lj}(t) \\
& + \big(F_0 - B_l R_l^{-1} L_0\big) R_0^{-1} \big(B_f R_f^{-1} L_l - G_f\big)^T \gamma_0(t) + B_l R_l^{-1} \big(B_f R_f^{-1} L_l - G_f\big)^T \gamma_{li}(t) \\
& - B_l R_l^{-1} B_l^T p_{li}(t) - \sum_{i=1}^{\mathscr{S}\mathscr{L}} \big(F_0 - B_l R_l^{-1} L_0\big) R_0^{-1} \big(B_f R_f^{-1} L_l R_l^{-1} L_0 - G_f R_l^{-1} L_0\big)^T \gamma_{0li}(t) \\
& - \sum_{i=1}^{\mathscr{S}\mathscr{L}} \big(F_0 - B_l R_l^{-1} L_0\big) R_0^{-1} \big(F_0 - B_l R_l^{-1} L_0\big)^T \alpha_{0li}(t) \Big\} dt + D_0 dW_0(t), \\
d\alpha_{0li}(t) = \Big\{ & -A_l^T \alpha_{0li}(t) + Q_l^T \delta_{0li}(t) - Q_l^T H_l \zeta_{0li}(t) \Big\} dt + r_0(t) dW_0(t),
\end{aligned} \tag{3.28}$$

subject to the boundary conditions:

$$x_{li}(t)\Big|_{t=0} = x_{li}(0), \qquad \alpha_{0li}(T) = 0.$$

Equations (3.29) below give the dynamics of co-states of the states of sub-leaders and their adjoints:

$$\begin{aligned}
dp_{li}(t) &= \Big\{ -A_l^T p_{li}(t) - Q_l \big[x_{li}^*(t) - H_l z_{li}^*(t)\big] \Big\} dt + r_{li}(t) dW_l(t) + r_0(t) dW_0(t), \\
d\delta_{0li}(t) &= \Big\{ A_l \delta_{0li}(t) + B_l R_l^{-1} B_l^T \alpha_{0li}(t) \Big\} dt + r_0(t) dW_0(t),
\end{aligned} \tag{3.29}$$

subject to the boundary conditions:

$$\delta_{0li}(t)\Big|_{t=0} = 0, \qquad p_{li}(T) = 0.$$

Recall that $z_{li}^*(t)$ is the mean field value for sub-leaders. Adjoint functions γ_{0li} are used to estimate the optimal values of $z_{li}^*(t)$. States and co-states for mean field dynamics of sub-leaders are given by

$$
\begin{aligned}
dz_{li}^*(t) = \Big\{ & A_f^{li} z_{li}^*(t) - B_f R_f^{-1} B_f^T p_{mi}(t) - \Big(B_f R_f^{-1} L_l R_l^{-1} L_0 - G_f R_l^{-1} L_0 \Big) R_0^{-1} B_0^T p_0(t) \\
& + \Big(B_f R_f^{-1} L_l R_l^{-1} L_0 - G_f R_l^{-1} L_0 \Big) R_0^{-1} \big(B_f R_f^{-1} L_l - G_f \big)^T \gamma_0(t) + \Big(B_f R_f^{-1} L_l R_l^{-1} B_l^T \\
& - G_f R_l^{-1} B_l^T \Big) p_{li}(t) + \sum_{\substack{j=1 \\ i \neq j}}^{\mathscr{S}\mathscr{L}} \Big(B_f R_f^{-1} L_l R_l^{-1} R_l^T - G_f R_l^{-1} R_l^T \Big) u_{lj}(t) - \sum_{i=1}^{\mathscr{S}\mathscr{L}} \Big(B_f R_f^{-1} L_l R_l^{-1} L_0 \\
& - G_f R_l^{-1} L_0 \Big) R_0^{-1} \big(B_f R_f^{-1} L_l R_l^{-1} L_0 - G_f R_l^{-1} L_0 \big)^T \gamma_{0li}(t) - \big(B_f R_f^{-1} L_l - G_f \big) R_l^{-1} \big(B_f R_f^{-1} L_l \\
& - G_f \big)^T \gamma_{li}(t) - \sum_{i=1}^{\mathscr{S}\mathscr{L}} \Big(B_f R_f^{-1} L_l R_l^{-1} L_0 - G_f R_l^{-1} L_0 \Big) R_0^{-1} \big(F_0 - B_l R_l^{-1} L_0 \big)^T \alpha_{0li}(t) \Big\} dt, \\
d\gamma_{0li}(t) = \Big\{ & -A_f^{li^T} \gamma_{0li}(t) - H_l^T Q_l^T \delta_{0li}(t) + -Q_l^T H_l^2 \zeta_{0li}(t) + (I_n - H_f)^T Q_f^T \beta_{0li}(t) \Big\} dt,
\end{aligned}
\tag{3.30}
$$

subject to the boundary conditions:

$$
z_{li}(0) = \lim_{N_i \to \infty} \frac{1}{N_i} \sum_{j=1}^{N_i} x_{fj}^{li}(0), \qquad \gamma_{0li}(T) = 0.
$$

The dynamics for p_{mi} and β_{0li} are given by

$$
\begin{aligned}
dp_{mi}(t) &= \big\{ -A_f^{li^T} p_{mi}(t) - Q_f \big[z_{li}^*(t) - H_f z_{li}^*(t) \big] \big\} dt, \\
d\beta_{0li}(t) &= \Big\{ A_f^{li} \beta_{0li}(t) + B_f R_f^{-1} B_f^T \gamma_{0li}(t) \Big\} dt,
\end{aligned}
\tag{3.31}
$$

subject to the boundary conditions:

$$
\beta_{0li}(t) \Big|_{t=0} = 0, \qquad p_{mi}(T) = 0.
$$

$z^*(t)$ denotes the mean field value and γ_0 symbolizes its adjoint function for the global leader. States and co-states for the mean field dynamics are given by

$$
\begin{aligned}
dz^*(t) = \Big\{ & A_f z(t) - B_f R_f^{-1} B_f^T p_{m0}(t) + \big(B_f R_f^{-1} L_l - G_f \big) R_0^{-1} B_0^T p_0(t) - \big(B_f R_f^{-1} L_l - G_f \big) R_0^{-1} \\
& \big(B_f R_f^{-1} L_l - G_f \big)^T \gamma_0(t) + \sum_{i=1}^{\mathscr{S}\mathscr{L}} \big(B_f R_f^{-1} L_l - G_f \big) R_0^{-1} \big(B_f R_f^{-1} L_l R_l^{-1} L_0 - G_f R_l^{-1} L_0 \big)^T \gamma_{0li}(t) \\
& + \sum_{i=1}^{\mathscr{S}\mathscr{L}} \big(B_f R_f^{-1} L_l - G_f \big) R_0^{-1} \big(F_0 - B_l R_l^{-1} L_0 \big)^T \alpha_{0li}(t) \Big\} dt, \\
d\gamma_0(t) = \Big\{ & -A_f^T \gamma_0(t) + H_0^T Q_0 [x_0^*(t) - H_0 z^*(t)] + (I_n - H_f)^T Q_f^T \beta_0(t) \Big\} dt,
\end{aligned}
\tag{3.32}
$$

subject to the boundary conditions:

$$
z(0) = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{\mathscr{S}\mathscr{L}} \sum_{j=1}^{N_i} x_{fj}^{li}(0), \qquad \gamma_0(T) = 0.
$$

Mean field co-state dynamics p_{m0} and its adjoint β_0 for global leader are

$$dp_{m0}(t) = \{-A_f^T p_{m0}(t) - Q_f[z(t)^* - H_f z^*(t)]\}dt,$$
$$d\beta_0(t) = \Big\{A_f\beta_0(t) + B_f R_f^{-1} B_f^T \gamma_0(t)\Big\}dt,$$

subject to the boundary conditions:

$$\beta_0(t)\Big|_{t=0} = 0, \qquad p_{m0}(T) = 0.$$

The dynamics of other co-states that are effective in the dynamics of sub-leaders and their corresponding adjoints are given by:

$$\begin{aligned}
d\gamma_{li}(t) =& \Big\{-A_f^{li^T}\gamma_{li}(t) - H_l^T Q_l[x_{li}^*(t) - H_l z_{li}^*(t)] + (I_n - H_f)^T Q_f^T \beta_{li}(t)\Big\}dt + D_0 dW_0(t),\\
d\zeta_{0li}(t) =& \Big\{A_f^{li^T}\zeta_{0li}(t) - B_f R_f^{-1} B_f^T \eta_{0li}(t) - \big(B_f R_f^{-1} L_l - G_f\big) R_l^{-1} B_l^T \alpha_{0li}(t)\\
&+ \big(B_f R_f^{-1} L_l - G_f\big) R_l^{-1}\big(B_f R_f^{-1} L_l - G_f\big)^T \gamma_{0li}(t)\Big\}dt + r_0(t) dW_0(t),
\end{aligned} \tag{3.33}$$

subject to the boundary conditions:

$$\zeta_{0li}(t)\Big|_{t=0} = 0, \qquad \gamma_{li}(T) = 0,$$

$$\begin{aligned}
d\beta_{li}(t) &= \Big\{A_f^{li^T}\beta_{li}(t) + B_f R_f^{-1} B_f^T \gamma_{li}(t)\Big\}dt,\\
d\eta_{0li}(t) &= \Big\{-A_f^{li}\eta_{0li}(t) + (I_n - H_f)^T Q_f^T \zeta_{li}(t)\Big\}dt,
\end{aligned} \tag{3.34}$$

$$\eta_{0li}(t)\Big|_{t=0} = 0, \qquad \beta_{li}(T) = 0.$$

In order to solve this FBSDE, first the problem is put in matrix form. Coupled FBSDE (3.27), (3.28), (3.29), (3.30), (3.30), (3.31), (3.32), (3.33) and (3.34) are used to construct a new form as coupled matrix FBSDE.

Let us set

$$\begin{aligned}
\mathscr{X}_0(t) = [&x_0^*(t), z^*(t), \beta_0(t), x_{li}^*(t), ..., x_{l\mathscr{S}\mathscr{L}}^*(t), z_{li}^*(t), ..., z_{l\mathscr{S}\mathscr{L}}^*(t), \beta_{li}(t), ..., \beta_{l\mathscr{S}\mathscr{L}}(t),\\
&\delta_{0li}(t), ..., \delta_{0l\mathscr{S}\mathscr{L}}(t), \beta_{0li}(t), ..., \beta_{0l\mathscr{S}\mathscr{L}}(t), \zeta_{0li}(t), ..., \zeta_{0l\mathscr{S}\mathscr{L}}(t)]^T,\\
\mathscr{Y}_0(t) = [&p_0(t), p_{m0}(t), \gamma_0(t), p_{li}(t), ..., p_{l\mathscr{S}\mathscr{L}}(t), p_{mi}(t), ..., p_{m\mathscr{S}\mathscr{L}}(t), \gamma_{li}(t), ..., \gamma_{l\mathscr{S}\mathscr{L}}(t),\\
&\alpha_{0li}(t), ..., \alpha_{0l\mathscr{S}\mathscr{L}}(t), \gamma_{0li}(t), ..., \gamma_{0l\mathscr{S}\mathscr{L}}(t), \eta_{0li}(t), ..., \eta_{0l\mathscr{S}\mathscr{L}}(t)]^T,
\end{aligned}$$

$$
\begin{aligned}
d\mathscr{X}_0(t) &= [\mathscr{A}_{01}\mathscr{X}_0(t)+\mathscr{B}_{01}\mathscr{Y}_0(t)+\mathscr{C}_{01}]dt+\mathscr{D}_{01}d\mathscr{W}_0(t),\\
d\mathscr{Y}_0(t) &= [\mathscr{A}_{02}\mathscr{X}_0(t)+\mathscr{B}_{02}\mathscr{Y}_0(t)]dt+\mathscr{D}_{02}d\mathscr{W}_0(t),
\end{aligned} \tag{3.35}
$$

subject to the boundary conditions:

$$
\mathscr{X}_0(t)\Big|_{t=0}=\mathscr{X}_0(0), \qquad \mathscr{Y}_0(T)=0.
$$

Theorem 3.1.3.1. *Given the equilibrium solution for sub-leaders and followers, and the corresponding mean field term, consider the local optimal control problem for the global leader. There exists a unique optimal controller $u_0^*\in L^2_{\mathscr{F}}(0,\mathscr{T};\mathscr{R}^n)$, given by*

$$
\begin{aligned}
u_0^*(t) = &-R_0^{-1}B_0^T p_0(t)+R_0^{-1}\big(B_fR_f^{-1}L_l-G_f\big)^T\gamma_0(t)-\sum_{i=1}^{\mathscr{S}\mathscr{L}}R_0^{-1}\big(B_fR_f^{-1}L_lR_l^{-1}L_0\\
&-G_fR_l^{-1}L_0\big)^T\gamma_{0li}(t)-\sum_{i=1}^{\mathscr{S}\mathscr{L}}R_0^{-1}\big(F_0-B_lR_l^{-1}L_0\big)^T\alpha_{0li}(t),
\end{aligned}
$$

where

$$
d\mathscr{X}_0(t)=\Big\{\mathscr{A}_{01}\mathscr{X}_0(t)+\mathscr{B}_{01}\Lambda_0(t)\mathscr{X}_0(t)+\mathscr{B}_{01}\mathscr{V}_0(t)+\mathscr{C}_{01}\Big\}dt+\mathscr{D}_{01}d\mathscr{W}_0(t), \tag{3.36}
$$

$$
d\mathscr{V}_0(t)=\Big\{\mathscr{B}_{02}\mathscr{V}_0(t)-\Lambda_0^T(t)\mathscr{B}_{01}\mathscr{V}_0(t)-\Lambda_0^T(t)\mathscr{C}_{01}\Big\}dt+\big\{\mathscr{D}_{02}-\Lambda_0^T\big\}d\mathscr{W}_0(t), \tag{3.37}
$$

$$
\mathscr{X}_0(t)\Big|_{t=0}=\mathscr{X}_0(0), \qquad \mathscr{V}_0(T)=0,
$$

and $\Lambda_0(t)$ is the solution of the following RDE

$$
\begin{aligned}
\dot{\Lambda}_0(t)+\Lambda_0^T(t)\mathscr{A}_{01}-\mathscr{B}_{02}\Lambda_0(t)-\Lambda_0^T(t)\mathscr{B}_{01}\Lambda_0(t)-\mathscr{A}_{02}=0,\\
\Lambda_0(T)=0,
\end{aligned} \tag{3.38}
$$

under the assumption that

$$
\det\left\{\begin{bmatrix}0 & I_n\end{bmatrix}e^{\mathscr{A}_0 t}\begin{bmatrix}I_n\\0\end{bmatrix}\right\}>0, \quad \forall t\in(0,T),
$$

where $\mathscr{A}_0$ is given by

$$
\mathscr{A}_0=\begin{bmatrix}\mathscr{A}_{01} & \mathscr{B}_{01}\\ \mathscr{A}_{02} & \mathscr{B}_{02}\end{bmatrix}.
$$

Proof. The proof is given in Appendix A. □

If Λ_0 can be solved in (3.38), then $\mathscr{X}_0(t)$, $\mathscr{V}_0(t)$, and $u_0^*(t)$ can be computed by solving the decoupled equations (3.36) and (3.37). Further details and solution of RDE can be found in [65]. As derived above, $u_0^*(t)$ is the best response of the global leader to the equilibrium solution of the sub-leaders and followers given its control. Hence, the overall solution of the FBSDE in Theorem 3.1.3.1 with the control policies for sub-leaders and followers are computed via Theorems 3.1.2.1 and 3.1.1.1 will yield the (approximate) Stackelberg equilibrium.

Remark 3.1.3.2. Radon's Lemma is used to solve $\Lambda_0(t)$ same as Lemma 3.1.2.1.

4. MEAN FIELD DIFFERENTIAL GAMES IN INTELLIGENT TRANSPORTATION SYSTEMS

In this section the game problem is introduced, which has been adapted to AHS in an ITS. The goal is to propose a new multi agent control algorithm based on mean field games to adjust traffic flow on an AHS according to the mean value of the ITS so that stop-and-go driving conditions on the ITSs are eliminated. There are three levels in the game problem as was implemented in chapter 3. AHS is separated into sub-sections so that optimization could be came out according to different characteristics of the traffic flow and environmental conditions. This kind of hierarchical control provides great advantages on AHS since it allows the Control Center and Road-Links to optimize different cost functions. Implementing the mean field game for ITS provides a smoother flow on all vehicles, and also each vehicle determines its control according to a dynamic energy cost. While each road link plays a game among vehicles in an attempt to achieve a smoother flow in its own section, the global center also optimizes travel time of the ITS. So, as fluctuation on each vehicle is reduced by the mean field value, total energy and travel time are also optimized simultaneously.

The rest of the section is organized as follows. We first introduce the problem and then explain the main difficulties to solve the problem exactly when the number of agents is large on AHS, furthermore we translate mean field problem to on a ITS. Mean Field Nash Game deals with the approximate optimal control problem that each vehicle is faced with in the mean field limit, when road-links impose their arbitrary strategy to the vehicles. Moreover, the approximate game problem of the road-links is solved subject to mean field of each section of the AHS. Control center problem is also solved according to Stackelberg game principles in this section.

4.1 Problem Statement

In the first level there is the global leader (control center), $\mathscr{G}_0$, who dominates sub-leaders (road-links), $\mathscr{RL}_i$, which are at the second level. Stackelberg game is played between the global leader and the sub-leaders. The global leader wants to minimize its own cost function, $\mathscr{J}_0^N$. It chooses its strategy and then announces it to the sub-leaders $\mathscr{RL}_i$ who play a non-cooperative game among each other. There is a total of N followers in the third level and each follower is assigned to a sub-leader. Let $\mathscr{RL}$ be the number of sub-leaders and let N_{lj} be the number of followers linked to sub-leader j, $j = 1 \ldots \mathscr{RL}$. Each sub-leader plays a Stackelberg mean field game with its followers, (see Fig. 4.1) and wants to minimize its cost, denoted as $\mathscr{J}_{li}^N$.

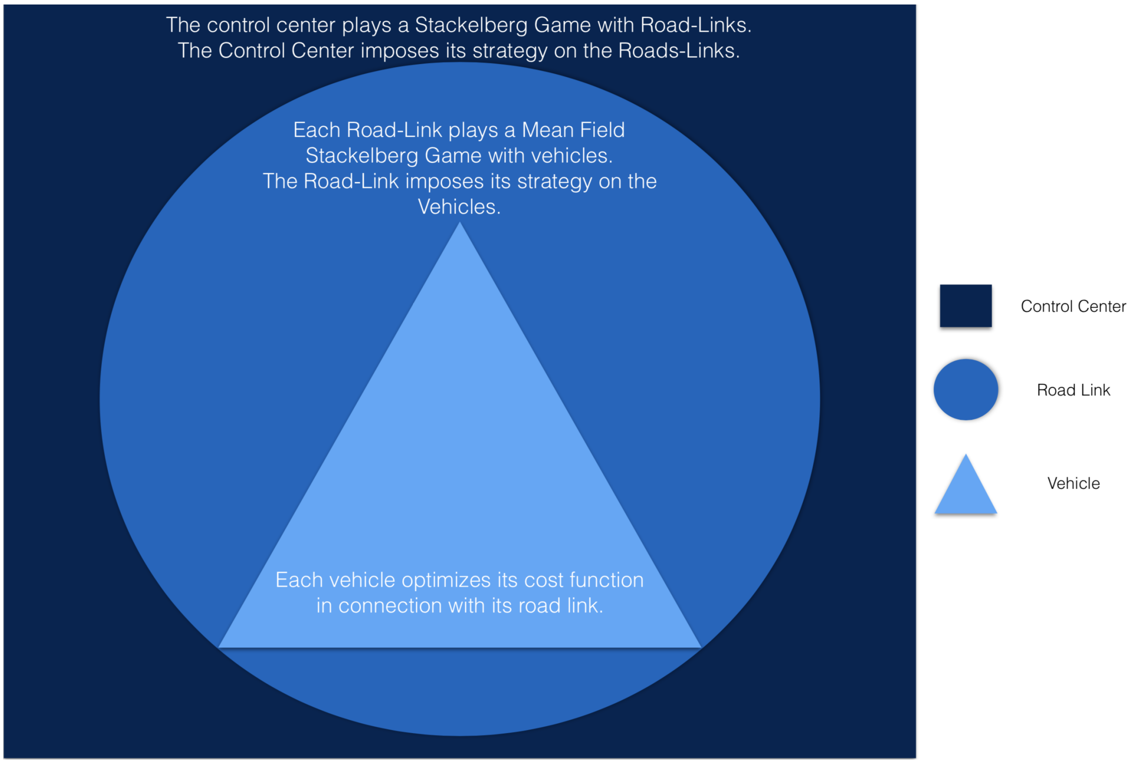

Figure 4.1 : Schema of the Stackelberg Game.

The stochastic differential equation (SDE) describing the state of the control center is given as:

$$dx_0(t) = \{a_0 x_0(t) + b_0 u_0(t)\}dt + d_0 dw_0(t), \tag{4.1}$$

where $x_0 \in \mathscr{R}$ is the state of the control center, $u_0 \in \mathscr{R}$ is the control computed by the control center, and $\{w_0(t),\ t \geq 0\}$ is a one-dimensional Brownian motion. The state equations of the road links, $\mathscr{RL}_i$, $1 \leq i \leq \mathscr{RL}$, are given by the following SDE's

$$dx_{li}(t) = \{a_l x_{li}(t) + b_l u_{li}(t) + f_0 u_0(t)\}dt + d_l dw_{li}(t), \tag{4.2}$$

where $x_{li} \in \mathscr{R}$ is the state of $\mathscr{RL}_i$, $u_{li} \in \mathscr{R}$ is its control, and $\{w_{li}(t),\ t \geq 0\}$ is a one-dimensional Brownian motion. Above, f_0 is a constant control of the global leader. For each road link $\mathscr{RL}_j$, $j = 1, \ldots, \mathscr{RL}$, the state equations of the followers, $\mathscr{F}_{fi}^{lj}$, $1 \leq i \leq N_{lj}$, dominated by $\mathscr{RL}_j$ are given by:

$$dx_{fi}^{lj}(t) = \{a_{fi}x_{fi}^{lj}(t) + b_{fi}u_{fi}^{lj}(t) + g_{fi}u_{lj}(t)\}dt + d_f dw_{fi}(t), \tag{4.3}$$

where $x_{fi}^{lj} \in \mathscr{R}$ is the state of the vehicle, $u_{fi}^{lj} \in \mathscr{R}$ is its control, and $\{w_{fi}(t),\ t \geq 0\}$ is a one-dimensional Brownian motion. N_{lj} is the number of vehicles governed by road link $\mathscr{RL}_j$ and $\sum_{j=1}^{\mathscr{RL}} N_{lj} = N$, where N is the total number of the vehicles in the game. Let $\lim_{N\to\infty} \frac{N_{lj}}{N} > 0$ for each $j = 1, \ldots, \mathscr{RL}$; that is, the number of vehicles for each road link is sufficiently large when the total number of followers is large. Each vehicle may have different dynamics in each road link in our model, therefore different $a_{fi},\ b_{fi},\ g_{fi}$ parameters can be assigned for each vehicle.

Assumption

(a) $x_0(0)$, $\{x_{fi}^{lj}(0),\ 1 \leq i \leq N\}$ and $\{x_{li}(0),\ 1 \leq i \leq \mathscr{RL}\}$ are independent of each other.

(b) $\{w_0(t), t \geq 0\}$, $\{w_{fi}(t),\ 1 \leq i \leq N, t \geq 0\}$, and $\{w_{li}(t),\ 1 \leq i \leq \mathscr{RL}, t \geq 0\}$ are independent of each other, which are also independent of $x_0(0)$, $\{x_{fi}^{lj}(0),\ 1 \leq i \leq N\}$ and $\{x_{li}(0),\ 1 \leq i \leq \mathscr{RL}\}$.

The performance index to be minimized for the control center, $\mathscr{G}_0$ is given by:

$$J_0^N(u_0, u^N) = E\int_0^T \Big\{q_0[\xi_0(t)x_0(t) - \nu^N(t)]^2 + r_{01}[u_0(t)]^2 + r_{02}[T_0(t)]^2\Big\}dt, \tag{4.4}$$

where $q_0 > 0$ and $r_{01} > 0,\ r_{02} > 0$, are the weighting values on tracking error, control and time respectively. The first term in J_0^N represents the penalty on tracking error of the flow, the second term stands for the penalty on the control effort, and third term is the penalty on the total travel time of ITS. It is possible to write:

$$\nu^N(t) = \xi_0(t) \times z^N(t), \tag{4.5}$$

where $\nu^N(t)$ represents the traffic flow of the system, ξ_0 is the density constant of the AHS and $z^N(t) = \dfrac{1}{N}\sum_{i=1}^N x_{fi}(t)$ is the *mean field term* that shows mass behavior of *all* vehicles. Here, it is assumed that mean field term gives the average velocity of the vehicles. Flow is affected directly by the first term, if the velocity of the followers converges to the mean field value, stop-and-go driving conditions are reduced throughout all AHS. u^N shows the controls of all other agents in the game; that is, $u^N = \{u_{lj}, u_{fi}^{lj}; 1 \leq i \leq N, 1 \leq j \leq \mathscr{RL}\}$.

In equations given below, flow based travel time function in ITS is represented by $T_0(t)$, C is the capacity of the AHS determined by the maximum flow possible on the highway, and t_f is the time taken to traverse in free flow conditions (when the density is zero and consequently the vehicles have free flow (maximum) speed). Bureau of Public Roads (BPR) (now FHWA) uses formula (4.6) below for $\tau_0(t)$ [69] which has a parameter α_τ with a typical value of 1 and β_τ which typically ranges between 2 and 12. In this thesis, we have assumed $\alpha_\tau = 1$ and $\beta_\tau = 2$ in our calculations.

$$\begin{aligned} T_0(t) =& t_f \tau_0(t), \\ \tau_0(t) =& \Big(1 + \beta_\tau \big(\frac{v^N}{C}\big)^{\alpha_\tau}\Big). \end{aligned} \tag{4.6}$$

The performance index to be minimized for road links $\mathscr{RL}_i$ is given by:

$$J_{li}^N(u_{li}, u_0, u_{-li}^N) = E \int_0^T \Big\{ q_{li}[\xi_{li}(t)x_{li}(t) - v_{li}^N(t)]^2 + r_{li}[u_{li}(t)]^2 + l_0 u_{li}(t) u_0(t) \Big\} dt, \tag{4.7}$$

$$v_{li}^N(t) = \xi_{li}(t) \times z_{li}^N(t). \tag{4.8}$$

where $q_{li} > 0$, $r_{li} > 0$, and $l_0 > 0$, are the weighting values on tracking error, control and leader control respectively. The mean field term that describes the mass behavior of the followers of $\mathscr{RL}_i$ is denoted by $z_{li}^N(t) = \frac{1}{N_{li}} \sum_{i=1}^{N_{li}} x_{fj}^{li}(t)$. The second and third terms in the performance index J_{li}^N represent the penalty on the control effort of the road links and the global-leader. Sub-leaders are weakly coupled with all the vehicles via the mean field term and they are strongly coupled with the global leader via their controls. Traffic flow of each road link, represented by $v_{li}^N(t)$ is given in (4.8) where ξ_{li} is the density of each road link. Each road link can reduce stop-and-go condition according to its own flow value due to the sub-mean field effect on the cost function (4.7).

The performance index to be minimized for vehicles $\mathscr{F}_{fi}^{lj}$ is given by:

$$\begin{aligned} J_{fi}^{N}(u_{fi}^{lj},u_{-fi}^{N}) =& E\int_{0}^{T}\Big\{q_{fi}[x_{fi}^{lj}(t)-z_{lj}^{N}(t)]^2+G(u_{fi}^{lj},u_{lj})+l_l u_{fi}^{lj}(t)u_{lj}(t)\Big\}dt,\\ G(u_{fi}^{lj},u_{lj}) =& r_{fi1}[u_{fi}^{lj}(t)]^2+r_{fi2}[x_{fi}^{lj}(t)]^2, \end{aligned} \tag{4.9}$$

where $q_{fi}>0$, $r_{fi1}>0$, $r_{fi2}>0$, and $l_l>0$. $G(u_{fi}^{lj},u_{lj})$ is the vehicle power related function of the performance index. The first term of $G(u_{fi}^{lj},u_{lj})$ is associated with the penalty on the control value of the vehicle. r_{fi1} involves the mass of the vehicle due to the fact that total energy is associated with the mass of the vehicle. The second term of $G(u_{fi}^{lj},u_{lj})$ represents the penalty on the velocity of the vehicle. When the vehicle goes faster on the highway, it faces more obstructive resistance effect such as rolling or speed correction. Therefore, coefficient r_{fi2} is related with approximated resistance effects and mass value of the vehicle besides the weight value of the cost function in our linear model (for detailed motor power models see [70]). Vehicles on the same road link are weakly coupled with each other via the mean field term z_{lj}^{N} and strongly coupled with their sub-leader via their controls u_{li}.

The classes of admissible controls for $\mathscr{G}_0$, $\mathscr{SL}_i$ and $\mathscr{F}_{fi}^{lj}$ is $L^2_{\mathscr{F}}(0,T;R)$; (the set of all adapted processes with finite second moments).

The most important difficulty when seeking exact Stackelberg solution within a large number of agents is that Nash equilibrium between followers is hard to obtain under arbitrary strategies of the leaders when the number of followers N is large (i.e., *curse of dimensionality*). The second major challenge is that when the number of the followers increases, the leaders face large numbers of constraints in their game problem. Hence, an exact solution would be impossible to obtain and mean field approach can be devised to obtain an approximate equilibrium. When the mass behavior of the followers are incorporated into the mean field term the totality of followers in each sub-team can be represented by a generic agent which will considerably decrease the complexity of the solution.

4.1.1 Stochastic mean field approximation for the vehicles

In this section, the mean field approximation of Nash equilibrium for the game played among N_j vehicles of the j^{th} road-link $\mathscr{SL}_j$ is developed. Vehicles are coupled through the mean field term $x_{lj}^{N}(t)$ with the players in their group and they are coupled with vehicles in other groups via controls of their road-links.

The dynamics and the cost functions of the vehicles separated into groups are given by (4.3) and (4.9). Each group is managed by its own sub-leader that may control its followers to achieve a different task.

$z_{lj}(t) \in L_F^2(0,T;R)$, $1 \leq j \leq \mathscr{S}\mathscr{L}$ represents the sub-mean field term in the infinite population limit $N_{lj} \to \infty$. The sub-mean field term $z_{lj}(t)$ and the control u_{li} of the sub-leader are exogenous terms in the cost function of a generic vehicle. A generic vehicle has to solve a stochastic control problem, given $z_{lj}(t)$ and u_{li}, to minimize its cost function (4.9) by using any admissible control $u_{fi}^{lj}(t)$. The (unique) optimal control and the corresponding state and adjoint equations are given as follows:

$$
\begin{aligned}
dx_{fi}^{lj*}(t) &= \left\{ a_f^{lj} x_{fi}^{lj*}(t) - \frac{b_f^2}{2r_{fi1}} p_{fi}(t) + \left(g_f - \frac{b_f l_l}{2r_{fi1}} \right) u_{lj}(t) \right\} dt + d_f dw_{fi}(t), \\
dp_{fi}(t) &= \left\{ -a_f^{lj} p_{fi}(t) - 2q_f [x_{fi}^{lj*}(t) - h_f z_{lj}(t)] - 2r_{fi2} x_{fi}^{lj*}(t) \right\} dt + \\
&\qquad e_{wfi}(t) dw_{fi}(t) + e_{wli}(t) dw_l(t),
\end{aligned}
\tag{4.10}
$$

where the state of a generic follower is calculated according to boundary conditions:

$$
x_{fi}^{lj}(t) = x_{fi}^{lj}(0) \qquad p_{fi}(T) = 0.
$$

(4.10) is a forward backward stochastic differential equation (FBSDE). It can be transformed into a different form which contains a Riccati differential equation (RDE) using Ito's Lemma and solving this Riccati equation is sufficient to obtain the solution of the FBSDE (4.10).

Theorem 4.1.1.1. *Given $z_{lj} \in L^2_{\mathscr{F}}\ (0,\mathscr{T};\mathscr{R})$ and $u_{lj}(t) \in L^2_{\mathscr{F}}\ (0,\mathscr{T};\mathscr{R})$, the local optimal control problem for F_{fi}^{lj}, $1 \leq i \leq N_{lj}$. admits a unique optimal controller u_{fi}^{lj*}. Moreover $(x_{fi}^{lj*},\ u_{fi}^{lj*}) \in L^2_{\mathscr{F}}\ (0,\mathscr{T};\mathscr{R}) \times L^2_{\mathscr{F}}\ (0,\mathscr{T};\mathscr{R})$ is the corresponding optimal control solution if and only if*

$$
u_{fi}^{lj*}(t) = -\frac{b_f}{2r_{fi1}} Z_{fi}(t) x_{fi}^{lj*}(t) - \frac{b_f}{2r_f} \Phi_{fi}(t) - \frac{l_l}{2r_{fi1}} u_{lj}(t), \tag{4.11}
$$

where

$$dx_{fi}^{lj*}(t) = \Big\{ \Big[a_f^{lj} - \frac{b_f^2}{2r_{fi1}} Z_{fi}(t) \Big] x_{fi}^{lj*}(t) - \frac{b_f^2}{2r_{fi1}} \Phi_{fi}(t) + \Big(g_f - \frac{l_l}{2r_{fi1}} \Big) u_{lj}(t) \Big\} dt + d_f dw_{fi}(t), \tag{4.12}$$

$$\begin{aligned} d\Phi_{fi}(t) = & \Big\{ \Big[-a_f^{lj} + \frac{b_f^2}{2r_{fi1}} Z_{fi}(t) \Big] \Phi_{fi}(t) + 2q_f h_f z_{lj}(t) - \frac{g_f - l_l}{2r_f} u_{lj}(t) Z_{fi}(t) \Big\} dt \\ & + \{ e_{wfi}(t) - Z_{fi}(t) \} dw_{fi}(t) + e_{wli}(t) dw_l(t), \end{aligned} \tag{4.13}$$

$$x_{fi}^{lj}(t) = x_{fi}^{lj}(0), \qquad \Phi_{fi}(T) = 0,$$

where $Z_{fi}(t)$ is the solution of the following RDE:

$$\begin{aligned} \dot{Z}_{fi}(t) + 2a_f^{lj} Z_{fi}(t) - \frac{b_f^2}{2r_{fi1}} Z_{fi}^2(t) + 2(q_f + r_{fi2}) = 0, \\ Z_{fi}(T) = 0. \end{aligned} \tag{4.14}$$

In the optimal control problem for a generic agent explained in detail above, the mean field term $z_{lj}(t)$ is the exogenous variable. According to the Nash certainty equivalence (NCE) principle the state evolution of a generic agent must be consistent with the mean field term (i.e., the total population behavior) which means that the expectation of the state of a generic agent must be equal to the mean field term $z_{lj}(t)$. In the following derivations, expectation of the state of the agent (which is equal to mean field term) and the corresponding co-state are as represented by $x_{mi}^*(t)$ and $p_{mi}(t)$, respectively and the coupled differential equations describing the evolution of the mean field term are obtained by taking the expectations of the coupled SDEs in (4.10).

4.1.2 Stochastic mean field approximation for the road links

In this section we consider the Nash game played among Road Links given the limiting mean field term and the policy of the control center. Each Road Links plays a Stackelberg mean field game with its vehicles.

The dynamics and the costs of road-links are given by equations (4.2) and (4.7) respectively. Let us consider the SDE for $\mathscr{RL}_i$, $1 \leq i \leq \mathscr{RL}$. The cost function of a generic road-link which uses approximated mean field value of its own group is described by (4.7) where $z_{li}(t) \in L_F^2(0,T;R)$ represents the infinite limit approximated mean field term. The mean field term $z_{li}(t)$, and control of global leader $u_0(t)$ are exogenous terms in i^{th} road-link's local control problem. The aim of each road-link is to minimize its cost function (4.7) using an admissible control $u_{li}(t)$. The (unique) optimal control and the corresponding state and adjoint equations are given as:

$$\begin{aligned} dx_{li}^*(t) = \Big\{a_l x_{li}^*(t) - \frac{b_l^2}{2r_l} p_{li}(t) + \Big(f_0 - \frac{b_l l_0}{2r_l}\Big) u_0^*(t) - \Big(\frac{b_l g_f}{2r_l} - \frac{b_l b_f l_l}{4r_f r_l}\Big)\gamma_{li}(t)\Big\} dt + \\ d_l dw_l(t), \\ dp_{li}(t) = \Big\{ - a_l p_{li}(t) - 2q_l \xi_{li}^2(t)[x_{li}^*(t) - h_l z_{li}^*(t)]\Big\} dt + e_{wli}(t) dw_l(t) + \\ e_{w0}(t) dw_0(t), \end{aligned} \tag{4.15}$$

where the state of each road-link can be obtained according to the boundary conditions:

$$x_{li}(t) = x_{li}(0), \qquad p_{li}(T) = 0.$$

Each road-link has its own mean field term denoted by $z_{li}^*(t)$. Adjoint functions γ_{li} are used to estimate the optimal values of $z_{li}^*(t)$. The dynamics of states and co-states for mean field dynamics are obtained as:

$$\begin{aligned} dz_{li}^*(t) = \Big\{a_f^{li} z_{li}^*(t) - \frac{b_f^2}{2r_f} p_{mi}(t) - \Big(\frac{g_f b_l}{2r_l} - \frac{b_l b_f l_l}{4r_f r_l}\Big) p_{li}(t) + \Big(\frac{b_f^2 l_l^2}{8r_f^2 r_l} + \frac{g_f^2}{2r_l}\Big)\gamma_{li}(t) - \\ \Big(\frac{g_f l_0}{2r_l} - \frac{l_0 b_f l_l}{4r_f r_l}\Big) u_0^*(t), \\ d\gamma_{li}(t) = \Big\{ - a_f^{li}\gamma_{li}(t) + 2h_l q_l \xi_{li}^2(t)[x_{li}^*(t) - h_l z_{li}^*(t)] + (2q_f - 2q_f h_f + 2r_{f2})\beta_{li}(t)\Big\} dt + \\ e_{wli}(t) dw_l(t) + e_{w0}(t) dw_0(t), \end{aligned} \tag{4.16}$$

subject to the boundary conditions:

$$z_{li}^*(0) = \lim_{N_i \to \infty} \frac{1}{N_i} \sum_{j=1}^{N_i} x_{fj}^{li}(0), \qquad \gamma_{li}(T) = 0.$$

The Hamiltonian includes mean field adjoint function p_{mi} as a constraint, so a new adjoint function β_{li} can be written to estimate the optimal values of the mean field adjoint function p_{mi}. Their dynamics are given by:

$$\begin{aligned} dp_{mi}(t) = \{ - a_f^{li} p_{mi}(t) - (2q_f - 2q_f h_f) z_{li}^*(t) - 2r_{f2} z_{li}^*(t)\} dt, \\ d\beta_{li}(t) = \Big\{a_f^{li}\beta_{li}(t) + \frac{b_f^2}{2r_f}\gamma_{li}(t) + e_{wli}(t) dw_l(t)\Big\} dt + e_{w0}(t) dw_0(t), \end{aligned} \tag{4.17}$$

subject to the boundary conditions:

$$\beta_{li}(t)\Big|_{t=0} = 0, \qquad p_{mi}(T) = 0.$$

First, the coupled equations (4.15), (4.16) and (4.17) for all sub-leaders are written as a coupled matrix FBSDE (see (4.18) below). Then, (4.18) is transformed into a different form which contains a RDE, using Ito's Lemma. Solving this Riccati equation is sufficient to obtain the solution of the FBSDE (4.18).

Let us set

$$\mathscr{X}_l(t) = [x^*_{li}(t), ..., x^*_{l\mathscr{S}\mathscr{L}}(t), z^*_{li}(t), ..., z^*_{l\mathscr{S}\mathscr{L}}(t), \beta_{li}(t), ..., \beta_{l\mathscr{S}\mathscr{L}}(t)]^T,$$
$$\mathscr{Y}_l(t) = [p_{li}(t), ..., p_{l\mathscr{S}\mathscr{L}}(t), p_{mi}(t), ..., p_{m\mathscr{S}\mathscr{L}}(t), \gamma_{li}(t), ..., \gamma_{l\mathscr{S}\mathscr{L}}(t)]^T.$$

The overall coupled FBSDE that yields the equilibrium solution for sub-leaders can be compactly written in the following form:

$$\begin{aligned} d\mathscr{X}_l(t) &= [\mathscr{A}_{l1}\mathscr{X}_l(t) + \mathscr{B}_{l1}\mathscr{Y}_l(t) + \mathscr{C}_{l1}]dt + \mathscr{D}_{l1}d\mathscr{W}_l(t) + \mathscr{D}_{l3}d\mathscr{W}_0(t), \\ d\mathscr{Y}_l(t) &= [\mathscr{A}_{l2}\mathscr{X}_l(t) + \mathscr{B}_{l2}\mathscr{Y}_l(t)]dt + \mathscr{D}_{l2}d\mathscr{W}_l(t) + \mathscr{D}_{l4}d\mathscr{W}_0(t), \end{aligned} \tag{4.18}$$

subject to the boundary conditions below:

$$\mathscr{X}_l(t)\Big|_{t=0} = \mathscr{X}_l(0), \qquad \mathscr{Y}_l(T) = 0.$$

Theorem 4.1.2.1. *Given the control of the global leader, $u_0(t)$, the game problem for sub-leaders $\mathscr{R}\mathscr{L}_i$, $1 \le i \le \mathscr{R}\mathscr{L}$, admits a unique equilibrium solution $\{u^*_{li}\}_{i=1}^{\mathscr{R}\mathscr{L}}$. Moreover $\{(x^*_{li},\ u^*_{li})\}_{i=1}^{\mathscr{R}\mathscr{L}}$ is the corresponding equilibrium solution if and only if for all $i = 1, \ldots, \mathscr{R}\mathscr{L}$,*

$$u^*_{li}(t) = -\frac{b_l}{2r_l}p_{li}(t) - \frac{l_0}{2r_l}u^*_0(t) - \Big(\frac{g_f}{2r_l} - \frac{b_f l_l}{4r_f r_l}\Big)\gamma_{li}(t),$$

where

$$d\mathscr{X}_l(t) = \Big\{\mathscr{A}_{l1}\mathscr{X}_l(t) + \mathscr{B}_{l1}\Lambda_l(t)\mathscr{X}_l(t) + \mathscr{B}_{l1}\mathscr{V}_l(t) + \mathscr{C}_{l1}\Big\}dt - d\mathscr{W}_l(t), \tag{4.19}$$

$$d\mathscr{V}_l(t) = \Big\{\mathscr{B}_{l2}\mathscr{V}_l(t) - \Lambda^T_l(t)\mathscr{B}_{l1}\mathscr{V}_l(t) - \Lambda^T_l(t)\mathscr{C}_{l1}\Big\}dt + \big\{\mathscr{D}_{l2} - \Lambda^T_l\big\}d\mathscr{W}_l(t) + \mathscr{D}_{l4}d\mathscr{W}_0(t), \tag{4.20}$$

$$\mathscr{X}_l(t) = \mathscr{X}_l(0), \qquad \mathscr{V}_l(T) = 0$$

Here, $\Lambda_l(t)$ is the solution of the following RDE:

$$\begin{aligned} \dot{\Lambda}_l(t) + \Lambda_l^T(t)\mathscr{A}_{l1} - \mathscr{B}_{l2}\Lambda_l(t) - \Lambda_l^T(t)\mathscr{B}_{l1}\Lambda_l(t) - \mathscr{A}_{l2} = 0, \\ \Lambda_l(T) = 0, \end{aligned} \tag{4.21}$$

where

$$\det\left\{ \begin{bmatrix} 0 & I \end{bmatrix} e^{\mathscr{A}_l t} \begin{bmatrix} I \\ 0 \end{bmatrix} \right\} > 0, \quad \forall t \in (0,T),$$

and the matrix $\mathscr{A}_l$ is given by

$$\mathscr{A}_l = \begin{bmatrix} \mathscr{A}_{l1} & \mathscr{B}_{l1} \\ \mathscr{A}_{l2} & \mathscr{B}_{l2} \end{bmatrix}.$$

4.1.3 Stochastic mean field approximation for the control center

Finally, a Stackelberg game is played between the control center and the road-links, where road-links apply the equilibrium solution computed in the previous section given the control of control center as an exogenous signal. The equilibrium solution that has been derived in the previous section depends on the control of the control center. Here, using these constraints the best result for the control center is obtained which will eventually lead to the (approximate) Stackelberg equilibrium. In these calculations, the control center can only reach the vehicles through road-links.

$\gamma_0(t)$, β_0, $\alpha_{0li}(t)$, $\delta_{0li}(t)$, β_{0li}, $\gamma_{0li}(t)$, $\eta_{0li}(t)$ and $\zeta_{0li}(t)$ are adjoint functions, which are used to obtain the extremal trajectory of $z(t)$, $p_{m0}(t)$, $x_{li}(t)$, $p_{li}(t)$, $p_{mi}(t)$, $z_{li}(t)$, $\beta_{li}(t)$ and $\gamma_{li}(t)$ respectively. These equations can be solved in a similar way as the road-links, as described in sub-section 4.1.2.

$$\begin{aligned} \mathscr{X}_0(t) = [&x_0^*(t), z^*(t), \beta_0(t), x_{li}^*(t), \ldots, x_{l\mathscr{S}\mathscr{L}}^*(t), z_{li}^*(t), \ldots, z_{l\mathscr{S}\mathscr{L}}^*(t), \beta_{li}(t), \ldots, \\ &\beta_{l\mathscr{S}\mathscr{L}}(t), \delta_{0li}(t), \ldots, \delta_{0l\mathscr{S}\mathscr{L}}(t), \beta_{0li}(t), \ldots, \beta_{0l\mathscr{S}\mathscr{L}}(t), \zeta_{0li}(t), \ldots, \zeta_{0l\mathscr{S}\mathscr{L}}(t)]^T, \\ \mathscr{Y}_0(t) = [&p_0(t), p_{m0}(t), \gamma_0(t), p_{li}(t), \ldots, p_{l\mathscr{S}\mathscr{L}}(t), p_{mi}(t), \ldots, p_{m\mathscr{S}\mathscr{L}}(t), \gamma_{li}(t), \ldots, \\ &\gamma_{l\mathscr{S}\mathscr{L}}(t), \alpha_{0li}(t), \ldots, \alpha_{0l\mathscr{S}\mathscr{L}}(t), \gamma_{0li}(t), \ldots, \gamma_{0l\mathscr{S}\mathscr{L}}(t), \eta_{0li}(t), \ldots, \eta_{0l\mathscr{S}\mathscr{L}}(t)]^T, \end{aligned}$$

$$\begin{aligned} d\mathscr{X}_0(t) &= [\mathscr{A}_{01}\mathscr{X}_0(t) + \mathscr{B}_{01}\mathscr{Y}_0(t)]dt + \mathscr{D}_{01}d\mathscr{W}_0(t), \\ d\mathscr{Y}_0(t) &= [\mathscr{A}_{02}\mathscr{X}_0(t) + \mathscr{B}_{02}\mathscr{Y}_0(t) + \mathscr{C}_{02}]dt + \mathscr{D}_{02}d\mathscr{W}_0(t), \end{aligned} \tag{4.22}$$

subject to the boundary conditions:

$$\mathscr{X}_0(t)\Big|_{t=0} = \mathscr{X}_0(0), \qquad \mathscr{Y}_0(T) = 0.$$

Theorem 4.1.3.1. *Given the equilibrium solution for sub-leaders and followers, and the corresponding mean field term, the local optimal control problem for $\mathscr{G}_0$, admits a unique optimal controller $u_0^* \in L^2_{\mathscr{F}}(0,\mathscr{T};\mathscr{R})$, given by*

$$u_0^*(t) = -\frac{b_0}{2r_0}p_0(t) - \Big(\frac{g_f}{2r_0} - \frac{b_f l_l}{4r_0 r_f}\Big)\gamma_0(t) + \sum_{i=1}^{\mathscr{S}\mathscr{L}}\Big(\frac{g_f l_0}{4r_0 r_l} - \frac{b_f l_l l_0}{8r_f r_l r_0}\Big)\gamma_{0li}(t) - \sum_{i=1}^{\mathscr{S}\mathscr{L}}\Big(\frac{f_0}{2r_0} - \frac{b_l l_0}{4r_0 r_l}\Big)\alpha_{0li}(t),$$

where

$$d\mathscr{X}_0(t) = \Big\{\mathscr{A}_{01}\mathscr{X}_0(t) + \mathscr{B}_{01}\Lambda_0(t)\mathscr{X}_0(t) + \mathscr{B}_{01}\mathscr{V}_0(t)\Big\}dt - d\mathscr{W}_0(t), \tag{4.23}$$

$$d\mathscr{V}_0(t) = \Big\{\mathscr{B}_{02}\mathscr{V}_0(t) - \Lambda_0^T(t)\mathscr{B}_{01}\mathscr{V}_0(t) + \mathscr{C}_{02}\Big\}dt + \big\{\mathscr{D}_{02} - \Lambda_0^T\big\}d\mathscr{W}_0(t), \tag{4.24}$$

$$\mathscr{X}_0(t) = \mathscr{X}_0(0), \mathscr{V}_0(T) = 0.$$

Here, $\Lambda_0(t)$ is the solution of the following RDE

$$\begin{aligned}\dot{\Lambda}_0(t) + \Lambda_0^T(t)\mathscr{A}_{01} - \mathscr{B}_{02}\Lambda_0(t) - \Lambda_0^T(t)\mathscr{B}_{01}\Lambda_0(t) - \mathscr{A}_{02} = 0,\\ \Lambda_0(T) = 0,\end{aligned} \tag{4.25}$$

where

$$\det\left\{\begin{bmatrix}0 & I\end{bmatrix} e^{\mathscr{A}_0 t}\begin{bmatrix}I\\0\end{bmatrix}\right\} > 0, \quad \forall t \in (0,T),$$

and the matrix $\mathscr{A}_0$ is given by

$$\mathscr{A}_0 = \begin{bmatrix}\mathscr{A}_{01} & \mathscr{B}_{01}\\ \mathscr{A}_{02} & \mathscr{B}_{02}\end{bmatrix}.$$

5. SIMULATION RESULTS

In this section two implementation works are presented. Firstly we have developed an innovative mean field game type in chapter 3, which is a multilayer hierarchical mean field game. Its theoretical analysis and mathematical equations are given extensively in that chapter. Here, the first result that is presented is related to MF-MSG. There is a global leader, which imposes its strategy to two sub-leaders. Two sub-leaders control different large number of followers. MF-MSG is examined according to 4 different test scenarios in MATLAB simulation environment.

The second implementation work given here has been developed for intelligent transportation systems. We have applied theory, which is obtained in chapter 3 to an ITS. This innovative work aims to control bunch of vehicles from a control center. Sub-leaders would be different road links and followers would be vehicles. It is hard to control lots of autonomous vehicles in a highway environment. Hence we have developed a new simulation environment, which has been implemented with C++ by QT editor. After the simulation program is explained in this section, scenarios with normal traffic flow and accident will be presented.

5.1 Numerical Examples for the MF-MSG

In this subsection we provide a numerical example to illustrate the main results. The following scalar dynamic equations for global leader, sub-leaders, and followers have been employed:

$$\begin{aligned}
dx_0(t) &= [0.02x_0(t) + 0.1u_0(t)]dt + 0.1dw_0(t),\\
dx_{l1}(t) &= [0.02x_{l1}(t) + 0.1u_{l1}(t) + 0.01u_{l2}(t) + 0.1u_0(t)]dt + 0.1dw_{l1}(t),\\
dx_{l2}(t) &= [0.02x_{l2}(t) + 0.1u_{l2}(t) + 0.01u_{l1}(t) + 0.1u_0(t)]dt + 0.1dw_{l2}(t),\\
dx_{fi}^1(t) &= [0.027x_{fi}^1(t) + 0.2u_{fi}^1(t) + 0.05u_{l1}(t)]dt + 0.1dw_{f1}(t),\\
dx_{fi}^2(t) &= [0.029x_{fi}^2(t) + 0.2u_{fi}^2(t) + 0.05u_{l2}(t)]dt + 0.1dw_{f2}(t),
\end{aligned}$$

There are only two sub-leaders in this example. The cost functions to be minimized are given as follows:

$$J_0^N(u_0) = E\int_0^{10}\Big\{0.1[x_0(t) - 1z^N(t)]^2 + [0.2(u_0(t))]^2\Big\}dt,$$

$$J_{l1}^N(u_{l1}, u_0, u_{-l1}^{\mathscr{S}\mathscr{L}}) = E\int_0^{10}\Big\{0.1[x_{l1}(t) - 1z_l1^N(t)]^2 + 0.2u_{l1}^2(t) + 0.02u_{l2}(t)u_{l1}(t) + 0.2u_{l1}(t)u_0(t)\Big\}dt,$$

$$J_{l2}^N(u_{l2}, u_0, u_{-l2}^{\mathscr{S}\mathscr{L}}) = E\int_0^{10}\Big\{0.1[x_{l2}(t) - 1z^l2^N(t)]^2 + 0.2u_{l2}^2(t) + 0.02u_{l1}(t)u_{l2}(t) + 0.2u_{l2}(t)u_0(t)\Big\}dt,$$

$$J_{fi}^N(u_{fi}^{lj}, u_0, u^{\mathscr{S}\mathscr{L}}) = E\int_0^{10}\Big\{0.015[x_{fi}^{lj}(t) - 1z_{lj}^N(t)]^2 + [0.2u_{fi}^{lj}(t)]^2 + 1.2u_{fi}^{lj}(t)u_{lj}(t)\Big\}dt.$$

The aim of the global leader is to force the followers to keep their state value at the mean field value. According to the simulation results the average mean field value of the followers is nearly 25, but the individual value of each follower is different from 25. The global leader tries to keep its own state value at 25, and also the followers keep their own values nearly at 25 through (MF)-(MSG). In the following results a comparison among four different games in a large scale tracking scenario is demonstrated. Different numbers of agents and hierarchical levels are employed in each game. Games start at the same initial conditions and use the same followers. The specification of each game is given below:

Game 1 consists of 1600 followers in 3 hierarchical levels (MF-MSG).

Game 2 consists of 800 followers in 3 hierarchical levels (MF-MSG).

Game 3 consists of 200 followers in 3 hierarchical levels (MF-MSG).

Game 4 consists of 1600 followers in 2 hierarchical levels (MF-SG).

5.1.1 Simulation results for the MF-MSG

The state of the global leader, $x_0(t)$ is depicted in Figure 5.1. The mean field value of the states converges to 25 in all the games. The global leader imposes its strategy on all the followers. It can be observed that at the beginning of the simulation the global leader is also affected by the mean field term, so it decreases its state value and

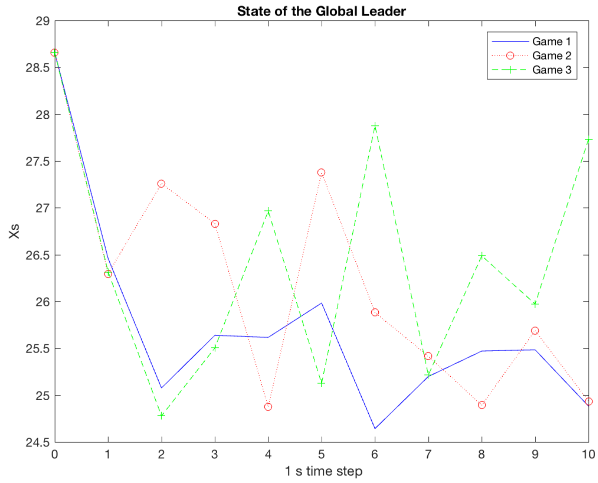

Figure 5.1 : State of the Global Leader.

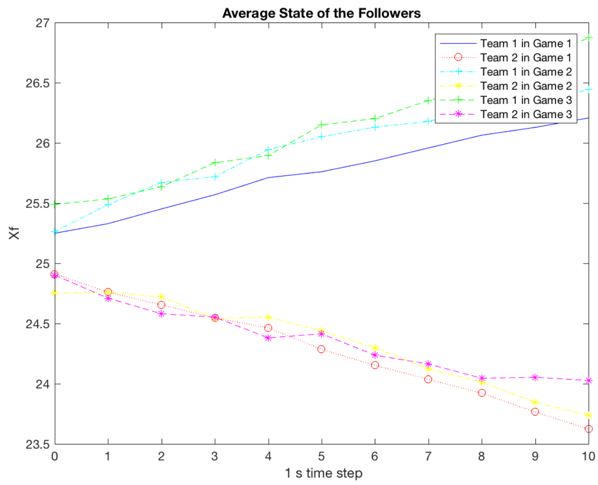

Figure 5.2 : Average State of Followers.

reaches the desired value nearly at 2nd simulation step. The global leader converges to desired value faster if the number of agents increases as in Game 1. Since the mean field error in Game 1 is smaller than in Game 2 and Game 3 the state of the global leader fluctuates less. It can be said that is applied a negative control.

The average state of all followers, $x_{fi}^{lj}(t)$ is illustrated in Figure 5.2 where it can be observed that the followers could nearly track the mean field term in all games. Average error increases as the simulation proceeds since adapted open loop control is applied. Each group performs differently due to their $a_f^{lj}(t)$ parameters. The number of followers higher in Game 1 compared to Game 2 and Game 3, so the followers in game 1 have smaller average state error. It can be concluded that as the mean field error decreases, the followers perform better control and the average follower state error decreases.

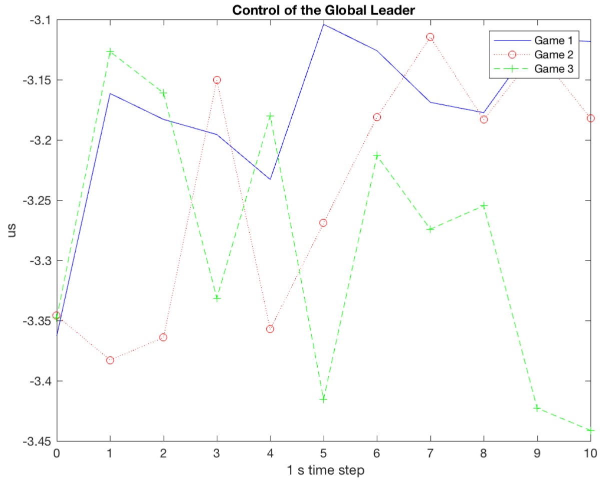

Figure 5.3 : Average Control of Global Leaders.

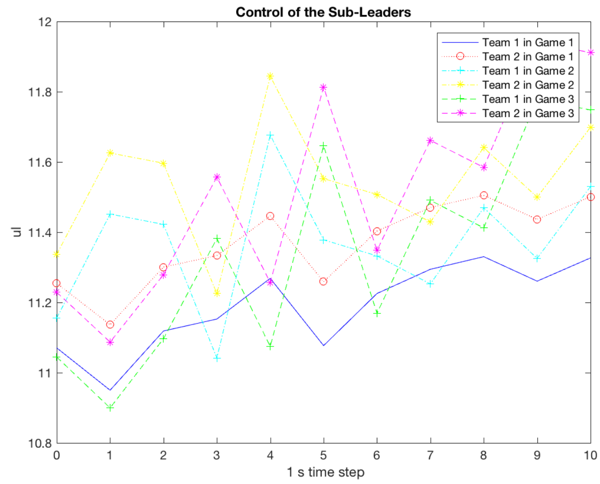

Figure 5.4 : Average Controls of Sub Leaders.

The average control of the global leader, $u_0(t)$ is shown in Figure 5.3. The global leader applies negative control on sub-leaders while sub-leaders apply positive control on the followers. Hence, followers are controlled indirectly by the global leader and smoother control is achieved. Also, sub-leaders may act differently compared to the leader.

The average control of sub-leaders, $u_{li}(t)$ is depicted in Figure 5.4. In our work, an additional layer of sub-leaders has been introduced compared to previous works which has crucial effects. Even though the global leader intends to impose a different control action, sub-leaders apply a specific control action to prevent unnecessary control caused by the mean field mass effect. So, error and cost decrease substantially. Hence, the additional layer of sub-leaders handles adverse effects of the environment more efficiently.

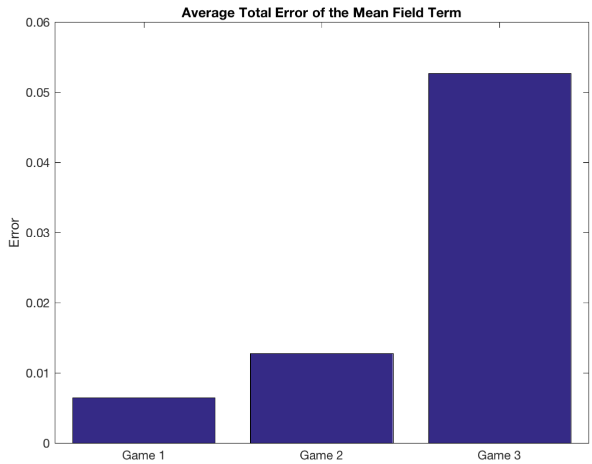

Figure 5.5 : Average Total Error of the Mean Field Term.

The average error of mean field terms is the difference between the exact mean field value and the approximated mean field value and it is illustrated in Figure 5.5. In Game 3, the number of agents is the lowest and the error has the highest value It can be observed that error decreases as the number of agents increases. Also, it is possible to predict that the error will converge to zero if the number of agents goes to infinity. In [11] Moon and Basar showed that.

A comparison of global leader states between Games 1 and 4, $x_0(t)$, in a different test environment than the previous results is provided in Figure 5.6. It can be observed that It can be observed that when initial state is higher than the desired value, both games could track the reference. There are more fluctuations in Game 1, since the Stackelberg game among sub-leaders and global-leader causes state trajectory variations. While the global leader determines its strategy it must take sub-leaders' dynamics into account. Although individual state trajectories show fluctuations, overall system performance is

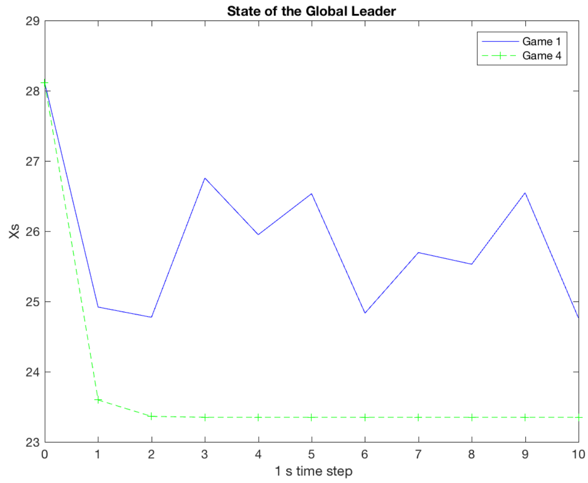

Figure 5.6 : State of the Global Leader in Games 1-4.

better. Also, it is possible to observe that the global leader state trajectory of Game 4 shows higher error. Moreover the state error has decreased in the multi-layer Game 1.

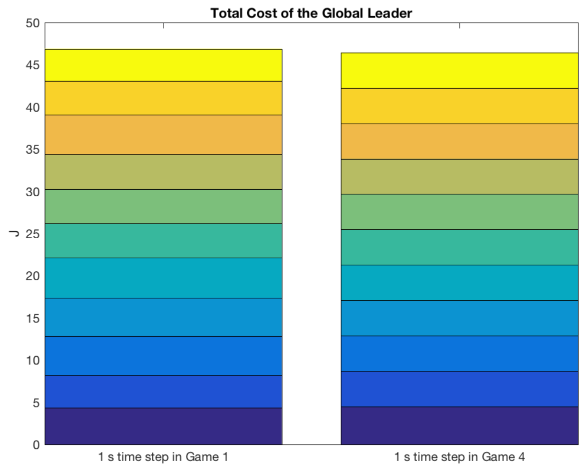

Figure 5.7 : The Global Leader Cost in Games 1-4.

A comparison of average cost of the global leader and followers between Games 1 and 4 respectively, $J_0(t)$, $J_{fi}(t)$ is illustrated in Figure 5.7 and Figure 5.8. While the cost of the global leader is approximately the same in both games, the additional sub-layer introduced in Game 1 reduces cost of the followers significantly. In both of the games there are 1600 agents which is the maximum number of agents in all the experiments. It can be observed that multi-layer approach performs better than the one layer approach due to specific control ability of sub-leaders' groups, which is shown in Figure 5.4.

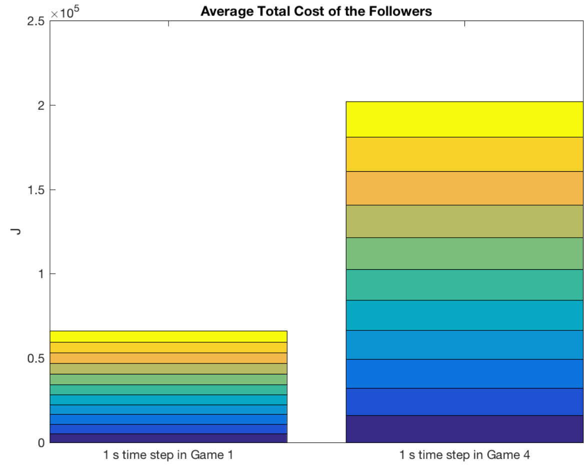

Figure 5.8 : Followers' Cost in Games 1-4.

In multi-layer game less cost is achieved compared to the one layer game. It can be concluded that Game 1 is more efficient with respect to Game 4 both from the perspective of cost and state error.

5.2 Numerical Examples for the MF-ITS

An important tool to test new ITS-oriented operational strategies and algorithms before they can be implemented on actual AHS is computer simulation. Some examples of professional Application Programming Interfaces (API) can be listed as Paramics [71], AIMSUN [72], METANET [73], CORSIM [74], and VISSIM [75]. Realistic conditions can be implemented before vehicles are on the road via an API which provides data exchange between interface and simulation environment. Models that have been developed have frequently been tested by already existing APIs. For instance a dynamic programming based adaptive signal processing algorithm has been proposed and tested using AIMSUN in [76]. In [77] authors used VISSIM to build a new hardware-in-the-loop traffic signal simulation framework to bridge traffic signals. On the other hand, in some works authors have developed their own simulation environments due to discrepancies between their methods and the ready to use simulation environments. For instance, in [78] authors needed an environment in continuous time, however most of ready-to-be-used APIs had been designed to be used in discrete time, so they built their own API through C++, to develop a simulation based iterative dynamic equilibrium traffic assignment model.

In this thesis, we needed to implement our complex algorithms on a new simulation environment to manage conditions via our own code, so we have used our own API, "ITUCiTSim", even though various useful ready-to-use simulation environments are still in use.

5.2.1 Real time simulation environment (ITUCiTSim)

Istanbul Technical University City Transportation Simulation (ITUCiTSim) is a professional program developed to test real time traffic conditions on AHSs [79]. It has been developed with QT C++, to create realistic traffic models taking into account road conditions, vehicle specifications, and traffic circumstances.

The monitoring screen of ITUCiTSim is depicted in Fig. 5.9. The desired road kilometer interval can be monitored and road sections can be selected from the left hand side of the screen. Each rectangle represents a different vehicle. Small and large rectangles represent cars and light trucks, which have different motor specifications and mass, respectively. Also, each color denotes different driver characteristics with regard to speed interval. Each vehicle possesses its own speed limit characteristic even if vehicles are included in the same color interval. The control panel of ITUCiTSim is illustrated in Fig. 5.10. As can be seen from the figure, environmental conditions can be adjusted as input parameters. Output parameters such as velocity, density, flow and carbon emission are shown on the right hand side of the panel. Furthermore, an accident can be created in the specific road kilometer with closed lines, which can be selected before the accident.

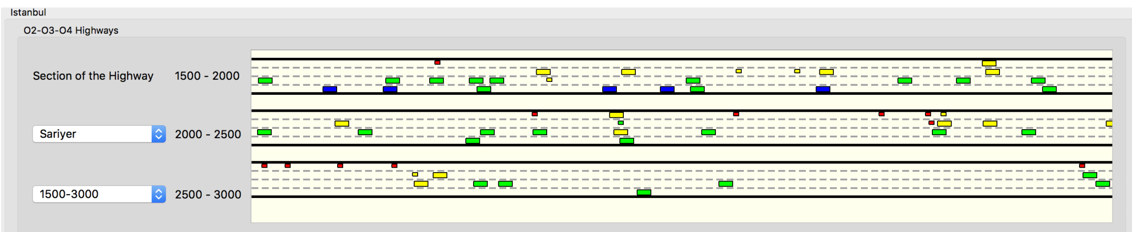

Figure 5.9 : ITUCiTSim Monitoring Screen.

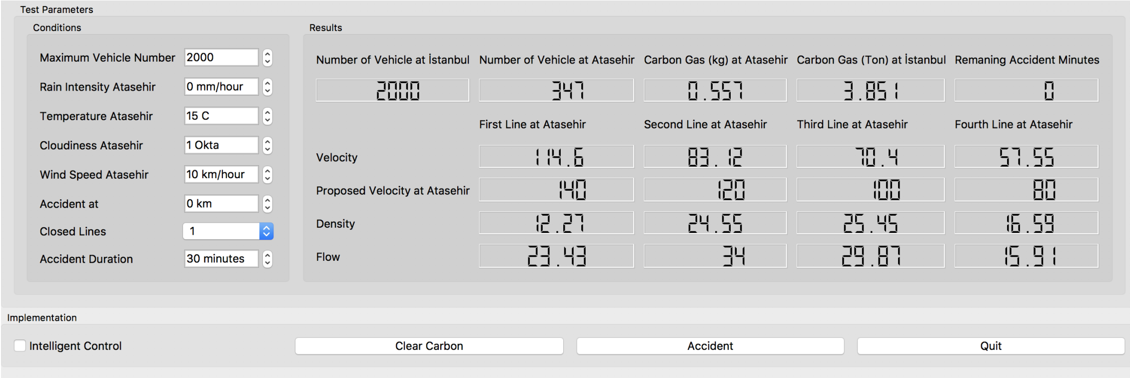

Figure 5.10 : ITUCiTSim - Control Panel.

A realistic case study model of Istanbul's E-80 highway has been employed to construct ITUCiTSim, and the data on the road links is given in Fig. 5.11. Vehicle specifications of ITUCiTSim are given in Table 5.1. Vehicle specifications must be considered carefully in a simulation environment since they affect numerical results directly.

Table 5.1 : ITUCiTSim Istanbul.

Vehicle Specifications		
Model	Car	Light Truck
Power (HP)	200	400
Speed(km/h)	80 - 140	60 – 80
Weight (kg)	2000	5000
Motor (liter)	1.5	5
Vehicle ratio on the road	%70	%30

5.2.2 Simulation results for the MF-ITS

In the thesis, simulations have been performed for the same road sections, vehicle parameters, and models as in [79]. A simulation of Istanbul E-80 highway model is developed using ITUCiTSim, the model is composed of 21 different road sections modeled as road-links in the MFG. The game includes 6 mean field road section and 1600 vehicles managed by a control center. The vehicles are distributed between the road links. The duration of the simulation is 2 minutes.

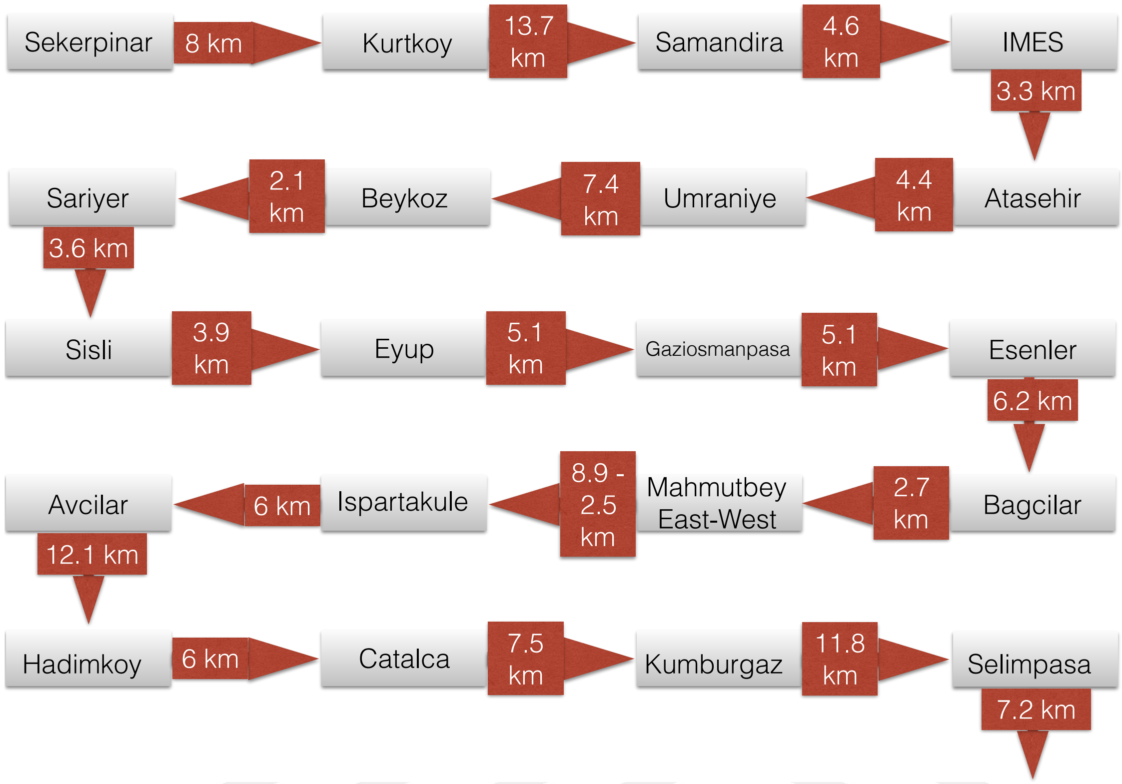

Figure 5.11 : Istanbul E-80 Highway Road Map.

1600 vehicles with their own characteristics are placed on random parts of six road links, initially. As the vehicles start to drive, mean field game is applied to manage them. The control center determines a cost function which is

$$J_0^N(u_0,u^N) = E\int_0^T \left\{ q_0[\xi_0(t)x_0(t) - \nu^N(t)]^2 + r_{01}[u_0(t)]^2 + r_{02}[T_0(t)]^2 \right\} dt.$$

Then it captures flow of the road , energy and travel time and imposes its strategy on the road links based on this cost function. Road links impose their strategies on their followers to minimize their cost functions with respect to their own traffic flow and energy characteristics. As a consequence, each vehicle specifies its strategy with reference to the mean field value while at the same time optimizing its cost function.

5.2.2.1 Normal traffic flow

Three different test scenarios with normal traffic flow conditions have been implemented. Figure 5.12 illustrates flow of the road-links. It can be noted that under

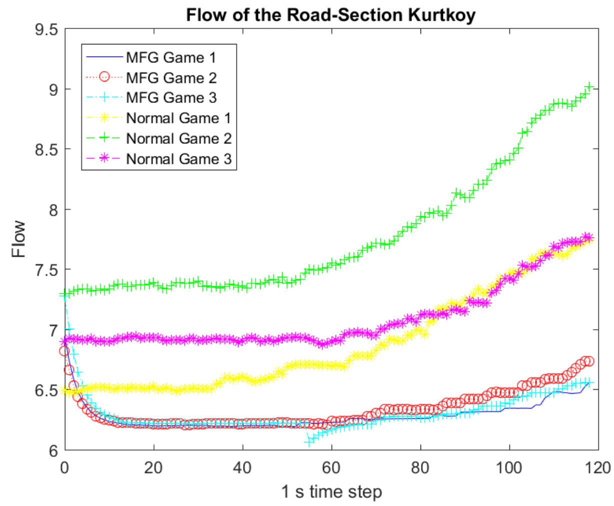

Figure 5.12 : Flow of the Kurtkoy Road Section.

mean field control a smoother flow is achieved. The flow does not change up to 60th second but after that traffic density increases which leads to a rise of the flow on the section of the road where accident has occurred

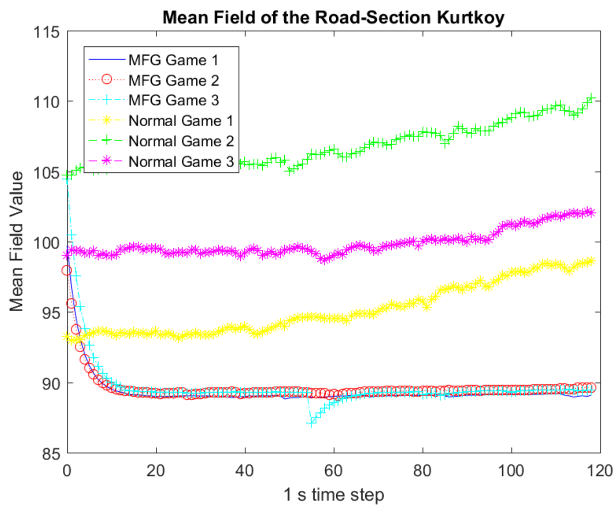

Figure 5.13 : Mean Field of the Kurtkoy Road Section.

During normal flow conditions the average speeds of the vehicles do not change as can be seen in figure 5.13. However, fluctuations on average speed is more than mean field test conditions. It van be observed that although the number of vehicles increases in this road-link, the MFG-ITS algorithm provides constant velocity on the highway.

The carbon emission of all vehicles in time has been depicted in Figure 5.14. In each test set different diversified car-truck ratio has been assigned which affects speed,

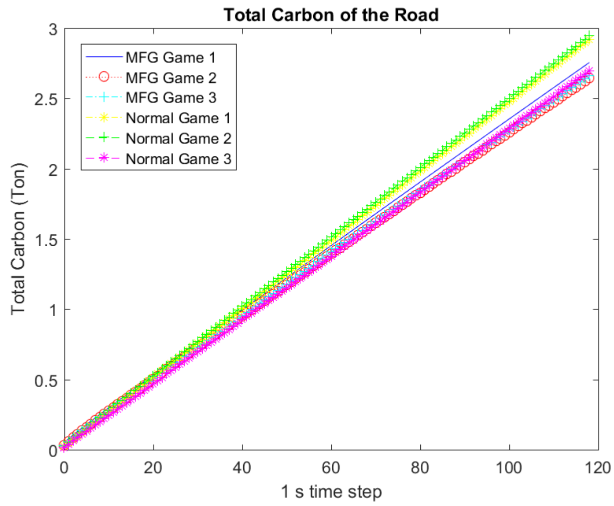

Figure 5.14 : Total Carbon Emission of the ITS.

fuel consumption, and stop-and-go number of vehicles. The total carbon emission is affected by vehicle characteristics in each different test set. It can be observed from varied test results that MFG can cope with total carbon emission on the ITS successfully. Total carbon emission is decreased approximately %6.35 at the end of the two minutes journey. In these tests it is assumed that there is no accident. In the case of an accident stop-and-go driving characteristic would increase, resulting in more fuel consumption and carbon emission.

5.2.2.2 Traffic flow after an accident

In these test scenarios, it is assumed that all road lanes are closed due to an accident. Then all lanes are re-opened and vehicles confront intense traffic. Hence the tests on traffic flow after an accident scenario analyzes a traffic environment involving highly intense stop-and-go driving conditions.

The flow of the road-link after an accident is depicted in Figure 5.15. The flows of normal-1 and normal-2 cases are less than the flow observed in the mean field game. Flow of the normal-3 scenario is faster than mean field games since average driving speed characteristics of vehicles are higher in this test scenario. It can be concluded that when vehicles are controlled via a mean field game, unnecessary stop-and-go is prevented, hence all vehicles can accelerate together faster than the uncontrolled case.

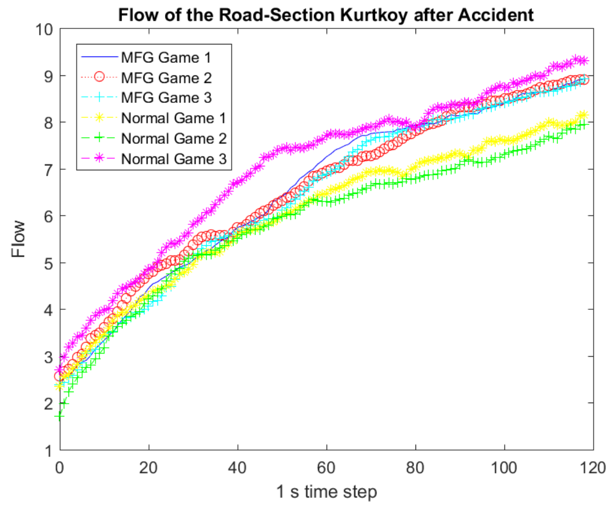

Figure 5.15 : Flow for Kurtkoy after Accident.

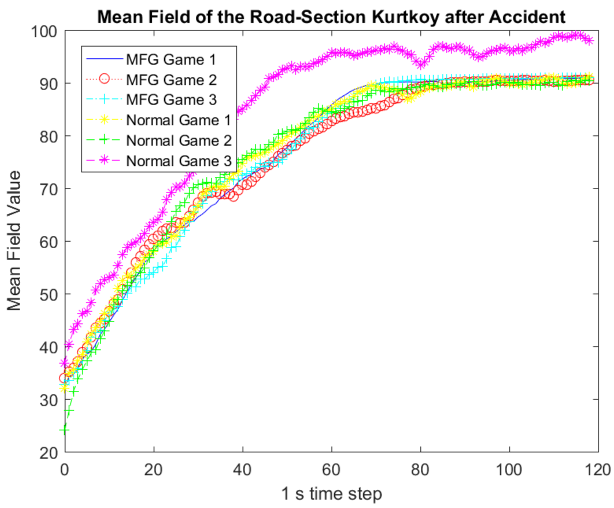

Figure 5.16 : Mean Field of Kurtkoy after Accident.

Average driving speed characteristics of vehicles in specified test scenarios are shown in Figure 5.16. The test results reveal that average velocities of all scenarios are approximately the same expect normal-3 because of its random environment conditions. In normal-1 and normal-2, vehicles try to get away from the congestion with a higher velocity but they fail to reach this higher velocity. Therefore they have approximately the same average velocity, however their flows are smaller values. As a result it can be concluded that mean field game helps to reduce the congestion.

The carbon emissions of all vehicles after an accident have been illustrated in Figure 5.17. It can be clearly seen that, in case of an accident MFG copes with total

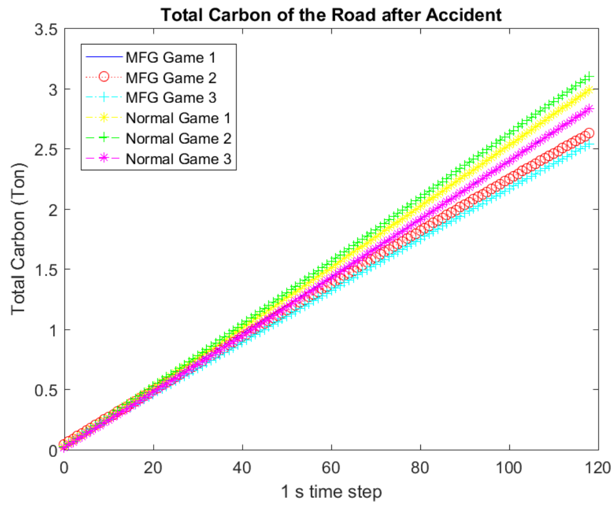

Figure 5.17 : Total Carbon Emission of the ITS after Accident.

carbon emission on the ITS successfully. MFG reduces the total carbon emission approximately %10.74 at end of the two-minutes journey.

6. CONCLUSIONS AND RECOMMENDATIONS

Growing body of research has been performed on mean field games recently due to fact that they provide an efficient solution for the multi agent control problems with large number of agents. Nash games can be applied through this new type of games. MFG involves large number of differential equations, which should be solved to derive a useful solution. This thesis has contributed to the mean field theory from a practical and useful point of view. Mean Field Multilayer Stackelberg Game can deal with complicated scenarios by dividing the main problem into sub-tasks. Divide and conquer technique is used in diverse research areas such as economics, politics, and computer sciences. Idea of our innovative model stems from divide and conquer technique. Hereby, complex problems could be divided to similar sub-systems, which have similar dynamics in the environment. Thus sub-problems with more precise models due to their analogous dynamics can be formulated and these sub-problems can be more efficiently solved.

In this thesis, we have applied MF-MSG algorithm to solve an important automated highway problem in smart cities. The importance of smart cities is increasing constantly among societies. The number of people living in cities is rising hence governments need more intelligent systems in order to optimize the management of time, energy, and wasted products. Smart cities should decrease time and energy consumption so emission of greenhouse gases must be decreased. One of the most important reasons of greenhouse gas emission is vehicles. Hence, it is expected that multi-agent control of non-autonomous vehicles on highways would lead to a significant decrease in this emission. There is an increasing research on autonomous vehicles recently and products of companies will be on the highway in near future. Autonomous vehicles could be easily directed through a control center to optimize energy and time. Therefore, a high necessity will arise for an algorithm which optimizes the large number of vehicles on an automated highway. This thesis has also contributed a solution to the important automated highway problem.

In this thesis, we have developed a new type of mean field game, which is a hierarchical layered game (MF-MSG). Then we have applied the new mean field game to the automated highway of the fictional smart city model of Istanbul. Hereby, we construct both a theoretical mean field model and give contributions to an environmental problem in a real life scenario.

6.1 Conclusion on MF-MSG

We have developed a new type of mean field approach for a class of multi-agent control problems. We call this new approach as Mean Field (MF) Multilayer Stackelberg Games (MSG). The approach enhances flexibility and applicability of the MFG. By the use of MF-MSG, agents can manage the system more efficiently when they encounter different kinds of restrictions. For example when some of the agents are affected by disturbances, they could be managed more robustly, since each group parameter could be adjusted separately in the physical system. Moreover sub-leaders may play different games among each other such as cooperative or rivalry games.

An optimal reference tracking problem is considered in this work. Each follower solves a local optimal control problem with its own group mean field term. Then sub-leaders solve their local optimal control problems, as nonstandard constrained optimization problems, with constraints imposed by the mean field process induced by all Nash followers. Finally the global leader solves its local optimal control problem as a Stackleberg Game while mean field process is arranged by sub-leaders. The global leader imposes its strategy on sub-leaders. Between levels 1 and 2, sub-leaders determine their strategies according to mean field of all the followers. Each sub-leader dominates its followers in level 3. Mean field game is played between level 2 and level 3. Each follower uses its group mean field value when sub-leaders determine their strategies.

Although the global leader can control the entire multi-agent system, it can not reach followers directly. The global leader controls sub-leaders through a Stackelberg Game and also sub-leaders manage their followers via a Stackelberg Mean Field Game. So sub-leaders can control their groups in spite of different environmental conditions. Furthermore they can also guide groups that have varied dynamics. In this way total cost and error decrease compared to previous approaches. Consequently we have shown four different experiments which demonstrate that cost efficiency increases and error converges to zero when the number of agent, goes to infinity.

We have implemented the case with 2 sub-leaders. If number of agents increases, more sub-leaders can be added to the system to increase efficiency. The case where a sub-leader controls a large number of agents is similar to a one layer case. Increasing the number of layers leads to a (simultaneous) increase in performance, but also aggravates computational load. So in the design this trade-off must be solved.

6.2 Conclusion on MF-ITS

In the second part of the thesis, we have applied the developed methodology of MF-MSG on an intelligent transportation system for future smart cities. In our approach, a control center will manage each vehicle in an automated highway in a multi-agent framework. There is a considerable body of technical literature on multi-agent control, however in large scale systems complexity is a major issue. Accordingly, we have introduced an innovative model for ITS whose analysis utilizes mean field game theory. In this model, complexity among vehicles is solved using MFGs in a Nash Game. In this model, the control center manages road links via its strategy, and also each vehicle determines its optimal action with respect to flow of its road link. The control center minimizes its cost function with respect to total travel time, energy and flow, and vehicles minimize their own cost functions with respect to only flow and their own energy functions.

Such a large scale ITS is not yet realizable in a real world scenario, therefore in this thesis we have developed an advanced simulation environment (ITUCiTSim) to perform our tests. In ITUCiTSim, a real road model of Istanbul's E-80 highway has been built, then the developed MF-ITS algorithm has been applied for the case of 1600 vehicles in 6 road-links, the duration of the tests have been chosen as 2 minutes. 2 different scenarios have been tested: Normal traffic flow and traffic flow after an accident. Test results show that, MF-ITS achieves smoother multi-agent control. It is also observed that each vehicle attains approximately the same velocity as computed by its cost function. Traffic congestion increases fuel consumption and carbon emission, this situation will not change even if the vehicles have low speed. When the whole system is controlled by MF-ITS each vehicle minimizes its own cost function, hence a significant reduction in total carbon emission is observed. The results show that as stop-and-go movements are prevented by controlling the vehicles, carbon emission is reduced under normal or congested flow conditions. Eventually, our MF-ITS algorithm could be used as an efficient control method with regard to environmental and economic issues in future smart cities.

6.3 Future Works

Complex AI problems can be effectively transformed into sub-problems using hierarchical systems. Mean field game theory is a useful tool for multi-agent decision making systems hence it will be required more complex differential equations in these types of games. If we choose stochastic differential equations in MFGs with the aim of modeling real world scenarios, then we could develop new solutions for SDEs. For example, in this thesis we dealt with the stochastic problem where noise terms did not contain any control parameter. If control parameter were affected by noise, then a different kind of solution would be required.

The mathematical part of this theses includes rich scientific problems, application of this mathematics to real world problems would be an important contributions to scientific literature. AI methods such as learning systems, natural language processing, computer vision, and decision-making systems will design the future life of humanity. All of these systems include huge amount of complex data so simplifying the problem of a decision making system to sub-problems would be crucial in finding the solution.

Nowadays data science is intensely used in most of the industry. Firstly user or environment data should be collected by a big data system. Big data is a concept, which manages huge amount of on-line data on a server. Thus intelligent systems could serve to a specific purpose, while they analyze exact requirements of a person through data science. Lots of implementation problems exist in big data and data science research area. Due to fact that both concepts include complex multi system problems, mean field games could be an efficient tool to obtain the solutions of these problems. If they can be modeled from the perspective of the multi agent game theory problem, mean field theory would optimize some of these implementation problems. Our innovative hierarchical model would also be an efficient tool.

Consequently, our approach could be implemented on large scale systems which consist of agents that have different dynamics. We could apply this theory in varied problems on AI, data science, big data, computer vision, machine learning, intelligent traffic management, power grids or biological networks.

As future work, we intend to improve our model taking into account more complicated traffic scenarios such as crossroads driving characteristics and more complex mathematical models as in semi-autonomous highways.

REFERENCES

[1] **Lin, Z.**, **Broucke, M. and Francis, B.** (2004). Local control strategies for groups of mobile autonomous agents, *IEEE Transaction on Automatic Control*, *49*(4), 622–629.

[2] **Fax, A.J. and Murray, R.M.** (2007). Information flow and cooperation control of vehicle formations, *IEEE Transaction on Automatic Control*, *49*(9), 1469–1479.

[3] **Perea, L.**, **Gomez, G. and Elosegui, P.** (2009). Extension of the Curker-Smale control law to space flight formations, *J. Guid, Control, Dyn.*, *32*(2), 527–537.

[4] **Ren, W.** (2007). Multi-vehicle consensus with a time-varying reference state, *Systems & Control Letters*, *56*, 474–483.

[5] **Lambson, V.** (1984). Self-enforcing collusion in large dynamic markets, *J. Econ. Theory*, *34*(2), 282–291.

[6] **Başar, T. and Olsder, G.J.** (1999). *Dynamic Noncooperative Game Theory*, SIAM.

[7] **Fudenberg, D. and Tirole, J.** (1991). *Game Theory*, The MIT Press.

[8] **Myerson, R.B.** (1991). *Game Theory: Analysis of Conflict*, Harvard University Press.

[9] **Huang, M.**, **Caines, P. and Malham, R.** (2007). Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ε-Nash equilibria, *IEEE Transaction on Automatic Control*, *52*(9), 1560–1571.

[10] **Li, T. and Zhang, J.F.** (2008). Asymptotically optimal decentralized control for large population stochastic multi-agent systems, *IEEE Transaction on Automatic Control*, *53*(7), 1643–1660.

[11] **Moon, M. and Başar, T.** (2016). Discrete-time mean field Stackelberg games with a large number of followers, *Proceedings of the 55th IEEE CDC*, Las Vegas, Nevada, USA, pp.4827–4832.

[12] **Moon, J. and Başar, T.** (2018). Linear Quadratic Mean Field Stackelberg Differantial Games, *SIAM Journal on Control and Optimization*, -(-), Under Review.

[13] **Başar, T.** (1983). Performance bounds for hierarchical systems under partial dynamic information, *J. Optimization Theory and Applications*, *39*(1), 67–87.

[14] **Lasry, J.M. and Lions, P.L.** (2007). Mean field games, *Jap. J. Math*, *2*(1), 229–260.

[15] **Tembine, H.**, **Le Boudec, J.Y.**, **ElAzouzi, R. and Altman, E.** (2009). Mean field asymptotic of Markov decision evolutionary games and teams, *Proceedings of Int. Conf. Game Theory for Networks(GameNets)*, Istanbul, Turkey, pp.140–150.

[16] **Huang, M.**, **Caines, P. and Malhamé, R.** (2003). Individual and mass behaviour in large population stochastic wireless power control problems: centralized and Nash equlibrium solutions, *Proceedings of the 42th IEEE CDC*, Maui, Hawaii, USA, pp.98–103.

[17] **Yin, H.**, **Mehta, P.**, **Meyn, S. and Shanblag, U.** (2010). Syncronization of coupled oscillators is a game, *Proceedings Amer. Control Conf. (ACC)*, Baltimore, MD, USA, pp.1783–1790.

[18] **Adlakha, S.**, **Johari, R.**, **Weintraub, G. and Goldsmith, A.** (2008). Oblivious equilbrium for large-scale stochastic games with unbounded costs, *Proceedings IEEE CDC*, Cancun, Mexico, pp.5531–5538.

[19] **Tembine, H.** (2011). Mean field stochastic games: Convergence, Q/H-learning and optimality, *Proceedings Amer. Control Conf. (ACC)*, San Francisco, CA, USA.

[20] **Weintraub, G.Y.**, **Benkard, C.L. and Van Roy, B.** (2005). Mean field games, *Adv. Neural Inf. Process. Syst.*, *18*.

[21] **Tembine, H. and Huang, M.** (2011). Mean field stochastic games: McKean-Vlasov dynamics, *Proceedings 50th IEEE Conf. Decision and Control and Eur. Control Conf. (CDC-ECC)*, Orlando, FL, USA.

[22] **Tembine, H.**, **Lasaulce, S. and Jungers, M.** (2010). Joint power control-allocation for green congnitive wireless networks using mean field theory, *Proceedings of the 5th IEEE Int. Conf. Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*.

[23] **Guéant, O.**, **Lasry, J.M. and Lions, P.L.** (2010). *Mean Field Games and Applications*, New York, NY, USA: Springer.

[24] **Yin, H.**, **Mehta, P.**, **Meyn, S. and Shanblag, U.** (2012). Synchronization of coupled oscillators is a game, *IEEE Transaction on Automatic Control*, *57*(4), 920–935.

[25] **Ma, Z.**, **Callaway, D. and Hiskens, I.** (2013). Decentralized charging control of large populations of plug-in electric vehicles, *IEEE Transaction on Control Systems Technology*, *21*(1), 67–78.

[26] **Kizilkale, A.C. and Malhamé, R.** (2014). Collective target tracking mean field control for Markovian jump-driven models of electric water heating loads, *Proceedings of the 19th IFAC World Congress*, Cape Town, South Africa, pp.1867–1872.

[27] **Zhu, Q.**, **Tembine, H. and Başar, T.** (2011). Hybrid risk-sensitive mean-field stochastic differential games with applications to molecular biology, *Proceedings of the 50th IEEE CDC and ECC*, Orlando, FL, USA, pp.4491–4497.

[28] **Weintraub, G.Y.**, **Benkard, C.L. and Van Roy, B.** (2008). Markov perfect industry dynamics with many firms, *Econometrica*, *76*(6), 1375–1411.

[29] **Huang, M.** (2010). Large-Population LQG games involving a major player: The Nash certainty equivalance principle, *SIAM Journal on Control and Optimization*, *48*(5), 3318–3353.

[30] **Nguyen, L. and Huang, M.** (2012). Linear-quadratic-Gaussian mixed games with continuum-parametrized minor players, *SIAM Journal on Control and Optimization*, *50*(5), 2907–2937.

[31] **Nguyen, L. and Caines, P.** (2013). ε-Nash mean field game theory for nonlinear stochastic dynamical systems with major and minor agents, *SIAM Journal on Control and Optimization*, *51*(4), 3302–3331.

[32] **Papavassilopoulos, G. and Cruz, J.B.** (1979). Nonclassical control problems and Stackelberg games, *IEEE Transaction on Automatic Control*, *24*(2), 155–166.

[33] **Simaan, M. and Cruz, J.B.** (1973). A Stackelberg solution for games with many players, *IEEE Transaction on Automatic Control*, *18*(3), 322–324.

[34] **Başar, T. and Selbuz, H.** (1979). Closed-Loop Stackelberg strategies with applications in optimal control of multilevel systems, *IEEE Transaction on Automatic Control*, *24*(2), 166–179.

[35] **Freiling, G.**, **Jank, G. and Lee, L.** (2001). Existence and Uniqueness of open-loop Stackelberg equlibria in linear-quadratic differential games, *J. Optimization Theory and Applications*, *110*(3), 515–544.

[36] **Yong, J.** (2002). A Leader-follower stochastic linear quadratic differantial game, *SIAM Journal on Control and Optimization*, *41*(4), 1015–1041.

[37] **Başar, T.**, **Bensoussan, A. and P., S.S.** (2010). Differentail games with mixed leadership: The open-loop solutionrmation, *Applied Mathematics and Computation*, *217*(3), 972–979, special Issue in Honor of George Leitman on his 85th Birth year.

[38] **Mohanty, S.P.**, **Choppali, B. and Kougianos, E.** (2016). Everything you wanted to know about smart cities: The Internet of things is the backbones, *IEEE Consumer Electronics Magazine*, *5*(3), 60–70.

[39] **Arroub, A.**, **Zahi, B.**, **Sabir, E. and Sadik, M.** (2016). A Literature Review on Smart Cities: Paradigms, Opportunities and Open Problems, *Intl. Conference on Wireless Networks and Mobile Communications (WINCOM)*, Fez, Morocco.

[40] **Dattakumar, A. and Sharmag, R.S.** (2016). Smart Cities and Knowledge Societies: Correlation, Causation or Distinct?, *IEEE Intl. Conference on Management of Innovation and Technology (ICMIT)*, Bangkok, Thailand.

[41] **Priano, F.H. and Guerra, C.F.** (2016). Fully Smart Cities: Analysis of the City to Smart City transformation process, *IEEE Intl. Conference on Smart Cities Conference (ISC2)*, Trento, Italy.

[42] **Batty, M.**, **Axhausen, K.W.**, **Giannotti, F.**, **Pozdnoukhov, A.**, **Bazzani, A.**, **Wachowicz, M.**, **Ouzounis, G. and Portugali, Y.** (2012). Smart cities of the future, *Eur. Phys. J. Spec. Top.*, *214*(1), 481—-518.

[43] **Wang, E. and Shi, J.** (2016). How Smart City Supports the Travel of Elderly Care Communities, *IEEE Intl. Conference on Smart Cities Conference (ISC2)*, Trento, Italy.

[44] **Přibyl, O. and Svítek, M.** (2015). System-oriented Approach to Smart Cities, *IEEE Intl. Conference on Smart Cities Conference (ISC2)*, Guadalajara, Mexico.

[45] **Pouryazdan, M. and Kantarci, K.W.** (2016). The Smart Citizen Factor in Trustworthy Smart City Crowdsensing, *IT Professional*, *18*(4), 26–33.

[46] **Hendy, M.E.**, **Miniaoui, S. and Fakhry, H.** (2015). Towards strategic information and communication technology (ICT) framework for smart cities decision-makers, *Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE)*, Shangri, Fiji.

[47] **Carli, R.**, **Dotoli, M. and Pellegrinoo, R.** (2016). A Hierarchical Decision-Making Strategy for the Energy Management of Smart Cities, *IEEE Transactions on Automation Science and Engineering*, *PP*(99), 1–19.

[48] **Zhu, F.**, **Li, Z.**, **Chen, S. and Xiong, G.** (2016). Parallel Transportation Management and Control System and Its Applications in Building Smart Cities, *IEEE Trans. Intell. Transp. Syst.*, *17*(6), 1576–1585.

[49] **Nandury, S.V. and Begum, B.A.** (2016). Strategies to Handle Big Data for Traffic Management in Smart Cities, *Intl. Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Jaipur, India.

[50] **Cerotti, D.**, **Distefano, S.**, **Merlino, G. and Puliafito, A.** (2016). A Crowd-Cooperative Approach for Intelligent Transportation Systems, *IEEE Trans. Intell. Transp. Syst.*, *PP*(99), 1–11.

[51] **Groot, N.**, **Schutter, B.D. and Hellendoorn, H.** (2015). Toward System-Optimal Routing in Traffic Networks: A Reverse Stackelberg Game Approach, *IEEE Trans. Intell. Transp. Syst.*, *16*(1), 29–40.

[52] **Kachroo, P. and Sastry, S.** (2016). Traffic Assignment Using a Density-Based Travel-Time Function for Intelligent Transportation Systems, *IEEE Trans. Intell. Transp. Syst.*, *17*(5), 1438–1447.

[53] **Koesdwiady, A., Soua, R. and Karray, F.** (2016). Improving Traffic Flow Prediction With Weather Information in Connected Cars: A Deep Learning Approach, *IEEE Transactions on Vehicular Technology*, *65*(2), 9508–9517.

[54] **Park, J., Murphey, Y.L., McGee, R., Kristinsson, J.G., Kuang, M.L. and Phillips, A.M.** (2014). Intelligent Trip Modeling for the Prediction of an Origin–Destination Traveling Speed Profile, *IEEE Trans. Intell. Transp. Syst.*, *15*(3), 1039–1053.

[55] **Zhao, G., Wu, C. and Qiao, C.** (2013). A Mathematical Model for the Prediction of Speeding with its Validation, *IEEE Trans. Intell. Transp. Syst.*, *14*(2), 828–836.

[56] **Tak, S., Kim, S. and Yeo, H.** (2016). A Study on the Traffic Predictive Cruise Control Strategy With Downstream Traffic Information, *IEEE Trans. Intell. Transp. Syst.*, *17*(7), 1932–1943.

[57] **Milanés, V., Shladover, S.E., Spring, J., Nowakowski, C., Kawazoe, H. and Nakamura, M.** (2014). Cooperative Adaptive Cruise Control in Real Traffic Situations, *IEEE Trans. Intell. Transp. Syst.*, *15*(1), 296–305.

[58] **Bauso, D., Zhang, X. and Papachristodoulou, A.** (2014). Density Flow over Networks: A Mean-field Game Theoretic Approach, *Proceedings 53th IEEE Conf. Decision and Control and Eur. Control Conf. (CDC-ECC)*, Los Angles, Ca, USA.

[59] **Wang, B. and Huang, M.** (2015). Dynamic Production Output Adjustment with Sticky Prices: A Mean Field Game Approach, *Proceedings 54th IEEE Conf. Decision and Control and Eur. Control Conf. (CDC-ECC)*, Osaka, Japan.

[60] **Couillet, R., Perlaza, S.M., Tembine, H. and Debbah, M.** (2012). Electrical Vehicles in the Smart Grid: A Mean Field Game Analysis, *IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS*, *30*(6).

[61] **Öner, A., Günel, G.O. and Saldi, N.** (2018). Mean Field Multilayer Stackelberg Differential Games in Multi-Agent Systems, *Appl. Comput. Math.*, *17*(1), 72–95.

[62] **Öner, A., T., B. and Günel, G.O.** (2017). Mean Field Differential Games in Intelligent Transportation Systems, *2017 Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, Napels, Italy.

[63] **Oksendal, B.** (2003). *Stochastic Differential Equations, An Introduction with Applications*, Springer-Verlag Berlin Heidelberg.

[64] **Luenberger, D.G.** (1969). *Optimization by Vector Space Methods*, John Wiley and Sons, Inc.

[65] **Ma, J. and Yong, J.** (1999). *Forward-Backward Stochastic Differential Equations and Their Applications*, Springer-Verlag Berlin Heidelberg.

[66] **Yong, J. and Zhou, X.Y.** (1999). *Stochastic Control: Hamiltonian Systems and HJB Equations*, Springer.

[67] **Stackelberg, V.H.**, (1934), Marktforrn und Gleichgewicht, Springer, Vienna.

[68] **Bensoussan, A.**, **Chen, S. and Sethi, S.P.** (2015). The Maximum Principle for Global Solutions of Stochastic Stackelberg Differential Games, *SIAM Journal on Control and Optimization*, *53*(4), 1956–1981.

[69] **Board, T.R.**, (1985), Highway capacity manual, Tech. Rep. 204, TRB, Washington, DC, USA.

[70] **Barth, M.**, **An, F.**, **Younglove, T.**, **Scora, G.**, **Levine, G.**, **Ross, M. and Wenzel, T.**, (2000), The Development of a Comprehensive Modal Emissions Model (Preliminary Draft Version 2), Final report to the National Cooperative Highway Research Program.

[71] **Ltd, Q.**, (2001), Paramics User Guide, U.K.

[72] **Syst, T.S.**, (2002), AIMSUN User Manual, Vers. 4.1, Barcelona, Spain.

[73] **Papageorgiou, M.**, **Papamichail, I.**, **Messmer, A. and Wang, Y.**, (2010), Traffic Simulation with METANET, Fundamentals of Traffic Simulation, 399 International Series in Operations Research and Management Science, Springer Science+Business Media.

[74] **Syst., F.H.A.**, (1997), Traffic Software Integrated System 97 User's Guide, Washington, DC.

[75] **AG, V.P.P.T.V.**, (2000), VISSIM 3.5 User Manual, Karlsruhe, Germany.

[76] **Fang, F.C. and Elefteriadou, L.** (2008). Capability-Enhanced Microscopic Simulation With Real-Time Traffic Signal Control, *IEEE Trans. Intell. Transp. Syst.*, *9*(4), 625–632.

[77] **Li, P. and Mirchandani, P.B.** (2016). A New Hardware-in-the-Loop Traffic Signal Simulation Framework to Bridge Traffic Signal Research and Practice, *IEEE Trans. Intell. Transp. Syst.*, *17*(9), 2430–2439.

[78] **Florian, M.**, **Mahut, M. and Tremblay, N.** (2008). Application of a simulation-based dynamic traffic assignment model, *European Journal of Operational Research*, *189*(3), 1381–1392.

[79] **Öner, A. and Günel, G.O.** (2015). A Case Study in Istanbul E-80: Intelligent Speed Change at the Highway for Reducing Gas Emission, *2015 Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, Budapest, Hungary.

APPENDICES

APPENDIXES

APPENDIX A

We will restrict our attention to the case $\mathscr{S}\mathscr{L} = 2$, for clarity of exposition.

For Sub-Leaders in Theorem 3.1.2.1

$$\mathscr{A}_{l1} = diag\{A_l, A_l, A_f^{l1}, A_f^{l2}, A_f^{l1}, A_f^{l2}\}, \qquad \mathscr{B}_{l2} = -\mathscr{A}_{l1},$$

$$\mathscr{A}_{l2} = \begin{bmatrix} -Q_l & 0 & Q_l H_l & 0 & 0 & 0 \\ 0 & -Q_l & 0 & Q_l H_l & 0 & 0 \\ 0 & 0 & -Q_f(1-H_f) & 0 & 0 & 0 \\ 0 & 0 & 0 & -Q_f(1-H_f) & 0 & 0 \\ -H_l^T Q_l & 0 & Q_l^T H_l^2 & 0 & (I_n - H_f)^T Q_f^T & 0 \\ 0 & -H_l^T Q_l & 0 & Q_l^T H_l^2 & 0 & (I_n - H_f)^T Q_f^T \end{bmatrix},$$

$$\mathscr{C}_{l1} = \begin{bmatrix} B_l\left(I_n - R_l^{-1} R_l^T\right) u_{l2}(t) + \left(F_0 - B_l R_l^{-1} L_0\right) u_0^*(t) \\ B_l\left(I_n - R_l^{-1} R_l^T\right) u_{l1}(t) + \left(F_0 - B_l R_l^{-1} L_0\right) u_0^*(t) \\ \left(B_f R_f^{-1} L_l R_l^{-1} R_l^T - G_f R_l^{-1} R_l^T\right) u_{l2}(t) + \left(B_f R_f^{-1} L_l R_l^{-1} L_0 - G_f R_l^{-1} L_0\right) u_0^*(t) \\ \left(B_f R_f^{-1} L_l R_l^{-1} R_l^T - G_f R_l^{-1} R_l^T\right) u_{l1}(t) + \left(B_f R_f^{-1} L_l R_l^{-1} L_0 - G_f R_l^{-1} L_0\right) u_0^*(t) \\ 0 \\ 0 \end{bmatrix},$$

$$\mathscr{D}_{l1} = \begin{bmatrix} D_l \\ D_l \\ 0_{n\times m} \\ 0_{n\times m} \\ r_{l1n\times m} \\ r_{l2n\times m} \end{bmatrix}, \qquad \mathscr{D}_{l2} = \begin{bmatrix} r_{l1n\times m} \\ r_{l2n\times m} \\ 0_{n\times m} \\ 0_{n\times m} \\ 0_{n\times m} \\ 0_{n\times m} \end{bmatrix}, \qquad \mathscr{D}_{l3} = \begin{bmatrix} 0_{n\times m} \\ 0_{n\times m} \\ 0_{n\times m} \\ 0_{n\times m} \\ r_{0n\times m} \\ r_{0n\times m} \end{bmatrix}, \qquad \mathscr{D}_{l4} = \begin{bmatrix} r_{0n\times m} \\ r_{0n\times m} \\ 0_{n\times m} \\ 0_{n\times m} \\ 0_{n\times m} \\ 0_{n\times m} \end{bmatrix},$$

$$\mathscr{B}_{lA} = \left(B_f R_f^{-1} L_l R_l^{-1} B_l^T - G_f R_l^{-1} B_l^T\right),$$
$$\mathscr{B}_{lB} = B_l R_l^{-1}\left(B_f R_f^{-1} L_l - G_f\right)^T,$$
$$\mathscr{B}_{lC} = -\left(B_f R_f^{-1} L_l - G_f\right) R_l^{-1}\left(B_f R_f^{-1} L_l - G_f\right)^T,$$

$$\mathscr{B}_{l1} = \begin{bmatrix} -B_l R_l^{-1} B_l^T & 0 & 0 & 0 & \mathscr{B}_{lB} & 0 \\ 0 & -B_l R_l^{-1} B_l^T & 0 & 0 & 0 & \mathscr{B}_{lB} \\ \mathscr{B}_{lA} & 0 & -B_f R_f^{-1} B_f^T & 0 & \mathscr{B}_{lC} & 0 \\ 0 & \mathscr{B}_{lA} & 0 & -B_f R_f^{-1} B_f^T & 0 & \mathscr{B}_{lC} \\ 0 & 0 & 0 & 0 & B_f R_f^{-1} B_f^T & 0 \\ 0 & 0 & 0 & 0 & 0 & B_f R_f^{-1} B_f^T \end{bmatrix}.$$

For Global-Leader in Theorem 3.1.3.1 The structure of the matrices of the global leader are similar to those of the sub-leaders.

APPENDIX B

Proof of Theorem 3.1.1.1. Equation (3.9) is derived using Hamilton equation (3.7), see [66, Chapter 3, Theorem 3.2]. Let us set

$$p_{fi}(t) = Z_{fi}(t)x_{fi}^{lj^*}(t) + \Phi_{fi}(t), \tag{A.1}$$

$$Z_{fi}(T) = 0 \qquad \Phi_{fi}(T) = G_f,\ G_f \in L^2_{\mathscr{F}}(\Omega, \mathscr{R}),$$

Equation (3.9) is a FBSDE, therefore Ito's formula is utilized to obtain the solution of the stochastic differential equation. The adjoint function is,

$$\begin{aligned} dp_{fi}(t) = &\Big\{\dot{Z}_{fi}(t)x_{fi}^{lj^*}(t) + Z_{fi}(t)\Big[\Big[A_f^{lj} - B_f R_f^{-1} B_f^T Z_{fi}(t)\Big]x_{fi}^{lj^*}(t) - B_f R_f^{-1} B_f^T \Phi_{fi}(t) - \\ &\big(B_f R_f^{-1} L_l - G_f\big)u_{lj}(t)\Big] + \dot{\Phi}_{fi}(t)\Big\}dt + Z_{fi}(t)dw_{fi}(t), \\ = &\Big\{\Big[-A_f^{lj} + Z_{fi}(t)B_f R_f^{-1} B_f^T\Big]\Phi_{fi}(t) + Q_f H_f z_{lj}(t) - Z_{fi}(t)\big(B_f R_f^{-1} L_l - G_f\big)u_{lj}(t)\Big\}dt \\ &+ \{r_{fi}(t) - Z_{fi}(t)D_f\}dW_{fi}^{lj}(t) + r_{li}(t)dW_{li}(t), \end{aligned}$$

Equilibrium of adjoint function of (3.9) is equal to above Ito's formula solution, then it can be written

$$\begin{aligned} 0 = &\Big\{\Big[\dot{Z}_{fi}(t) + A_f^{lj^T} Z_{fi}(t) + Z_{fi}(t)A_f^{lj} - Z_{fi}(t)B_f R_f^{-1} B_f^T Z_{fi}(t) + Q_f\Big]x_{fi}^{lj^*}(t) + \dot{\Phi}_{fi}(t) + \\ &\Big[A_f^{lj} - B_f R_f^{-1} B_f^T Z_{fi}(t)\Big]\Phi_{fi}(t) + \big(B_f R_f^{-1} L_l - G_f\big)Z_{fi}(t)u_{lj}(t) - Q_f H_f z_{lj}(t)\Big\}dt \\ &+ \Big\{Z_{fi}(t) + Z_{fi}(t)D_f - r_{fi}(t)\Big\}dW_{fi}(t) - r_{li}(t)dW_{li}(t). \end{aligned} \tag{A.2}$$

A RDE is obtained from the first part of equation (A.2. The solution of this RDE is employed in forward-backward differential equation (3.9).

$$0 = \dot{Z}_{fi}(t) + A_f^{lj^T} Z_{fi}(t) + Z_{fi}(t)A_f^{lj} - Z_{fi}(t)B_f R_f^{-1} B_f^T Z_{fi}(t) + Q_f.$$

The dynamic of $\Phi_{fi}(t)$ can be determined from second part of the equation (A.2):

$$\begin{aligned} d\Phi_{fi}(t) = &\Big\{\Big[-A_f^{lj} + Z_{fi}(t)B_f R_f^{-1} B_f^T\Big]\Phi_{fi}(t) + Q_f H_f z_{lj}(t) - Z_{fi}(t)\big(B_f R_f^{-1} L_l - G_f\big)u_{lj}(t)\Big\}dt \\ &+ \{r_{fi}(t) - Z_{fi}(t)D_f\}dW_{fi}^{lj}(t) + r_{li}(t)dW_{li}(t), \end{aligned}$$

The state dynamics can be written from equation (3.9) to find $x_{fi}^{lj^*}(t)$ as follows:

$$dx_{fi}^{lj^*}(t) = \Big\{A_f^{lj} x_{fi}^{lj^*}(t) - B_f R_f^{-1} B_f^T p_{fi}(t) - \big(B_f R_f^{-1} L_l - G_f\big)u_{lj}(t)\Big\}dt + D_f dW_{fi}^{lj}(t),$$

If equation (A.1 is substituted for $p_{fi}(t)$), $dx_{fi}^{lj^*}(t)$ is obtained as:

$$dx_{fi}^{lj^*}(t) = \Big\{ \Big[A_f^{lj} - B_f R_f^{-1} B_f^T Z_{fi}(t)\Big] x_{fi}^{lj^*}(t) - B_f R_f^{-1} B_f^T \Phi_{fi}(t) - \big(B_f R_f^{-1} L_l - G_f\big) u_{lj}(t) \Big\} dt \\ + D_f dW_{fi}^{lj}(t).$$

After finding the solution of the dynamic equations, its optimality must also be demonstrated. $J_{fi}(u_{fi}^{lj^*}, u_0, u^{\mathscr{S}\mathscr{L}})$ and $x_{fi}^{lj^*}(t)$ are convex and Lipschitz continuous in u_{fi}^{lj}. Hence, $(x_{fi}^{lj^*}, u_{fi}^{lj^*}(t))$ is optimal. See [66, Chapter 2, Section 5].

The Riccati equation has a unique solution for $Z_{fi}(t)$. So, forward backward equation (A) and (3.12) has a unique solution under control (3.10). We can conclude about the uniqueness of the solution to (3.9) from the existence of a unique solution to (3.13) and the uniqueness of the solution to (3.10) follows from the existence of a unique solution of (3.9). See [65, Chapter 2] .

This completes the proof. $\square$

APPENDIX C

Proof of Theorem 3.1.2.1 Equation (3.18) is derived using Hamilton equation (3.17), see [66, Chapter 3, Theorem 3.2]. Let us set

$$\mathscr{Y}_l(t) = \Lambda_l(t)\mathscr{X}_l(t) + \mathscr{V}_l(t), \tag{A.3}$$

$$\Lambda_l(T) = 0 \qquad \mathscr{Y}_l(T) = G_{li},\ G_{li} \in \mathscr{R}^n,$$

Equation (3.18) is a FBSDE, therefore Ito's formula is utilized to obtain the solution of the stochastic differential equation. The adjoint function is:

$$\begin{aligned} d\mathscr{Y}_l(t) =& \Big\{ \dot{\Lambda}_l(t)\mathscr{X}_l(t) + \Lambda_l^T(t)\left[\mathscr{A}_{l1}\mathscr{X}_l(t) + \mathscr{B}_{l1}\mathscr{Y}_l(t) + \mathscr{C}_{l1}\right] + \dot{\mathscr{V}}_l(t)\Big\} dt + \Lambda_l^T d\mathscr{W}_l(t), \\ =& [\mathscr{A}_{l2}\mathscr{X}_l(t) + \mathscr{B}_{l2}\mathscr{Y}_l(t)]dt + \mathscr{D}_{l2}d\mathscr{W}_l(t) + \mathscr{D}_{l4}d\mathscr{W}_0(t), \end{aligned} \tag{A.4}$$

Equilibrium of adjoint function of (3.18) is equal to above Ito's formula solution, then it can be written

$$\begin{aligned} 0 =& \Big\{ \left[\dot{\Lambda}_l(t) + \Lambda_l^T(t)\mathscr{A}_{l1} - \mathscr{B}_{l2}\Lambda_l(t) - \Lambda_l^T(t)\mathscr{B}_{l1}\Lambda_l(t) - \mathscr{A}_{l2}\right]\mathscr{X}_l(t) + \dot{\mathscr{V}}_l(t) + \Lambda_l^T(t)\mathscr{B}_{l1}\mathscr{V}_l(t) \\ &+ \Lambda_l^T(t)\mathscr{C}_{l1} - \mathscr{B}_{l2}\mathscr{V}_l(t)\Big\} dt + \left\{\Lambda_l^T - \mathscr{D}_{l2}\right\} dW_l(t) - \mathscr{D}_{l4}d\mathscr{W}_0(t), \end{aligned} \tag{A.5}$$

A Riccati equation is obtained from the first part of equation (A.5), this RDE is employed to solve the forward-backward differential equation (3.18).

$$0 = \dot{\Lambda}_l(t) + \Lambda_l^T(t)\mathscr{A}_{l1} - \mathscr{B}_{l2}\Lambda_l(t) - \Lambda_l^T(t)\mathscr{B}_{l1}\Lambda_l(t) - \mathscr{A}_{l2}$$

$d\mathscr{V}_l(t)$ follows from the second part of equation (A.5) as:

$$d\mathscr{V}_l(t) = \Big\{ \left[\mathscr{B}_{l2} - \Lambda_l^T(t)\mathscr{B}_{l1}\right]\mathscr{V}_l(t) - \Lambda_l^T(t)\mathscr{C}_{l1}\Big\} dt + \left\{\mathscr{D}_{l2} - \Lambda_l^T(t)\right\} dW_l(t) + \mathscr{D}_{l4}d\mathscr{W}_0(t).$$

We can rewrite equation (A.3) to find $\mathscr{X}_l(t)$ as:

$$\mathscr{V}_l(t) = \mathscr{Y}_l(t) - \Lambda_l(t)\mathscr{X}_l(t),$$

Using Ito's formula we obtain:

$$\dot{\mathscr{V}}_l(t) = \Big\{ \dot{\mathscr{Y}}_l(t) - \dot{\Lambda}_l(t)\mathscr{X}_l(t) - \Lambda_l^T(t)\dot{\mathscr{X}}_l(t)\Big\} dt - \Lambda_l^T(t)dW_l(t),$$

If equation (A.4) is put into the equation above, then it can be formulated as:

$$\begin{aligned} \dot{\mathscr{V}}_l(t) =& \Big\{ \dot{\Lambda}_l(t)\mathscr{X}_l(t) + \Lambda_l^T(t)\mathscr{A}_{l1}\mathscr{X}_l(t) + \Lambda_l^T(t)\mathscr{B}_{l1}\Lambda_l(t)\mathscr{X}_l(t) + \Lambda_l^T(t)\mathscr{B}_{l1}\mathscr{V}_l(t) + \Lambda_l^T(t)\mathscr{C}_{l1} \\ &+ \dot{\mathscr{V}}_l(t) - \dot{\Lambda}_l(t)\mathscr{X}_l(t) - \Lambda_l^T(t)\dot{\mathscr{X}}_l(t)\Big\} dt - \Lambda_l^T(t)dW_l(t), \end{aligned}$$

Finally $d\mathscr{X}_l(t)$ is derived as:

$$\dot{\mathscr{X}_l}(t) = \Big\{\mathscr{A}_{l1}\mathscr{X}_l(t) + \mathscr{B}_{l1}\Lambda_l(t)\mathscr{X}_l(t) + \mathscr{B}_{l1}\mathscr{V}_l(t) + \mathscr{C}_{l1}\Big\}dt - dW_l(t).$$

Equation (3.24) is a non symmetric Riccati equation and its solution can be obtained by Radon's Lemma. If the non symmetric Riccati Equation is solved using Radon's Lemma, we get:

$$\begin{bmatrix}\dot{\mathscr{M}_l}(t)\\ \dot{\mathscr{N}_l}(t)\end{bmatrix} = \begin{bmatrix}\mathscr{A}_{l1} & -\mathscr{B}_{l1}\\ -\mathscr{A}_{l2} & \mathscr{B}_{l2}\end{bmatrix}\begin{bmatrix}\mathscr{M}_l(t)\\ \mathscr{N}_l(t)\end{bmatrix}dt, \quad \begin{bmatrix}\mathscr{M}_l(T)\\ \mathscr{N}_l(T)\end{bmatrix} = \begin{bmatrix}I_n\\ 0\end{bmatrix},$$

$\mathscr{M}_l(t)$ is invertible, then

$$\Lambda_l(t) = \mathscr{N}_l(t)\mathscr{M}_l^{-1}(t),$$

The necessary and sufficient condition for uniqueness of the non symmetric Riccati Equation [65, (Chapter 2, Theorem 4.3)], is;

$$\det\left\{\begin{bmatrix}0 & I_n\end{bmatrix} e^{\mathscr{A}_l t}\begin{bmatrix}I_n\\ 0\end{bmatrix}\right\} > 0, \quad \forall t\,\varepsilon\,(0,T),$$

Here $\mathscr{A}_l$ matrix is denoted as,

$$\mathscr{A}_l = \begin{bmatrix}\mathscr{A}_{l1} & \mathscr{B}_{l1}\\ \mathscr{A}_{l2} & \mathscr{B}_{l2}\end{bmatrix}.$$

After finding the solution of the dynamic equations, its optimality must also be demonstrated. $J_{li}(u_{li}, u_0, u_{-li}^{\mathscr{S}\mathscr{L}})$ and $\mathscr{X}_l$ are convex and Lipschitz continuous in u_{li}. Hence, $(\mathscr{X}_l, u_{li}(t))$ is optimal. See [66, Chapter 2, Section 5].

The Riccati equation has a unique solution for $\Lambda_l(t)$. Hence, forward backward equation (3.22) and (3.23) have a unique solution under control (4.1.2.1). We can conclude about the uniqueness of the solution to (3.18) from the existence of a unique solution to (3.24) and the uniqueness of the solution to (4.1.2.1) follows from the existence of a unique solution of (3.18). See [65, Chapter 2].

This completes the proof.

APPENDIX D

Proof of Theorem 3.1.3.1 Let us set

$$\mathscr{Y}_0(t) = \Lambda_0(t)\mathscr{X}_0(t) + \mathscr{V}_0(t), \tag{A.6}$$

$$\Lambda_0(T) = 0 \qquad \mathscr{Y}_0(T) = G_0,\ G_0 \in \mathscr{R}^n.$$

The dynamics of $\mathscr{Y}_0$ can be expressed in two different ways by using Ito's formula.

$$\begin{aligned} d\mathscr{Y}_0(t) =& \Big\{ \dot{\Lambda}_0(t)\mathscr{X}_0(t) + \Lambda_0^T(t)\left[\mathscr{A}_{01}\mathscr{X}_0(t) + \mathscr{B}_{01}\mathscr{Y}_0(t) + C_{01}(t)\right] + \dot{\mathscr{V}}_0(t)\Big\}dt + \Lambda_0^T d\mathscr{W}_0(t), \\ =& [\mathscr{A}_{02}\mathscr{X}_0(t) + \mathscr{B}_{02}\mathscr{Y}_0(t)]dt + \mathscr{D}_{02}d\mathscr{W}_0(t). \end{aligned} \tag{A.7}$$

Subtracting the second equation from the first one, we obtain

$$\begin{aligned} 0 =& \Big\{ \left[\dot{\Lambda}_0(t) + \Lambda_l^T(t)\mathscr{A}_{01} - \mathscr{B}_{02}\Lambda_0(t) - \Lambda_0^T(t)\mathscr{B}_{01}\Lambda_0(t) - \mathscr{A}_{02}\right]\mathscr{X}_0(t) + \dot{\mathscr{V}}_0(t) + \Lambda_0^T(t)\mathscr{B}_{01}\mathscr{V}_0(t) \\ &+ \Lambda_0^T(t)\mathscr{C}_{01} - \mathscr{B}_{02}\mathscr{V}_0(t)\Big\}dt + \{\Lambda_0^T - \mathscr{D}_{02}\}dW_0(t). \end{aligned} \tag{A.8}$$

The following Riccati equation is derived from the first part of equation (A.8),

$$0 = \dot{\Lambda}_0(t) + \Lambda_0^T(t)\mathscr{A}_{01} - \mathscr{B}_{02}\Lambda_0(t) - \Lambda_0^T(t)\mathscr{B}_{01}\Lambda_0(t) - \mathscr{A}_{02} \tag{A.9}$$

$d\mathscr{V}_0(t)$ follows from the second part of equation (A.8), as:

$$d\mathscr{V}_0(t) = \Big\{\left[\mathscr{B}_{02} - \Lambda_0^T(t)\mathscr{B}_{01}\right]\mathscr{V}_0(t) - \Lambda_0^T(t)\mathscr{C}_{01}\Big\}dt + \left\{\mathscr{D}_{02} - \Lambda_0^T(t)\right\}dW_0(t).$$

We can rewrite equation (A.6) to find $\mathscr{X}_0(t)$ as:

$$\mathscr{V}_0(t) = \mathscr{Y}_0(t) - \Lambda_0(t)\mathscr{X}_0(t).$$

Using Ito's formula we obtain:

$$\dot{\mathscr{V}}_0(t) = \Big\{\dot{\mathscr{Y}}_0(t) - \dot{\Lambda}_0(t)\mathscr{X}_0(t) - \Lambda_0^T(t)\dot{\mathscr{X}}_0(t)\Big\}dt - \Lambda_0^T(t)dW_0(t).$$

If equation (A.7) is put into the equation above, then we obtain

$$\begin{aligned} \dot{\mathscr{V}}_0(t) =& \Big\{\dot{\Lambda}_0(t)\mathscr{X}_0(t) + \Lambda_0^T(t)\mathscr{A}_{01}\mathscr{X}_0(t) + \Lambda_0^T(t)\mathscr{B}_{01}\Lambda_0(t)\mathscr{X}_0(t) + \Lambda_0^T(t)\mathscr{B}_{01}\mathscr{V}_0(t) + \Lambda_0^T(t)\mathscr{C}_{01} \\ &+ \dot{\mathscr{V}}_0(t) - \dot{\Lambda}_0(t)\mathscr{X}_0(t) - \Lambda_0^T(t)\dot{\mathscr{X}}_0(t)\Big\}dt - \Lambda_0^T(t)dW_0(t). \end{aligned}$$

It is shown that $d\mathscr{X}_0(t)$ satisfies the following SDE

$$\dot{\mathscr{X}}_0(t) = \Big\{\mathscr{A}_{01}\mathscr{X}_0(t) + \mathscr{B}_{01}\Lambda_0(t)\mathscr{X}_0(t) + \mathscr{B}_{01}\mathscr{V}_0(t) + \mathscr{C}_{01}\Big\}dt - dW_0(t).$$

To show the existence of a unique equilibrium solution, it is sufficient to prove the existence of a unique solution to the RDE given in (A.9). Equation (A.9) is a non-symmetric Riccati equation and its solution can be found by Radon's Lemma; that is, if $(\mathscr{M}_0, \mathscr{N}_0)$ is a solution of the following equation

$$\begin{bmatrix} \dot{\mathscr{M}}_0(t) \\ \dot{\mathscr{N}}_0(t) \end{bmatrix} = \begin{bmatrix} \mathscr{A}_{01} & -\mathscr{B}_{01} \\ -\mathscr{A}_{02} & \mathscr{B}_{02} \end{bmatrix} \begin{bmatrix} \mathscr{M}_0(t) \\ \mathscr{N}_0(t) \end{bmatrix} dt, \quad \begin{bmatrix} \mathscr{M}_0(T) \\ \mathscr{N}_0(T) \end{bmatrix} = \begin{bmatrix} I_n \\ 0 \end{bmatrix},$$

where $\mathscr{M}_0(t)$ is invertible; then the solution of the Riccati equation in (A.9) is given by

$$\Lambda_0(t) = \mathscr{N}_0(t)\mathscr{M}_0^{-1}(t),$$

For existence and uniqueness of a solution of RDE [65, (Chapter 2, Theorem 4.3)], a necessary and sufficient condition is

$$\det\left\{ \begin{bmatrix} 0 & I_n \end{bmatrix} e^{\mathscr{A}_0 t} \begin{bmatrix} I_n \\ 0 \end{bmatrix} \right\} > 0, \quad \forall t\, \varepsilon\, (0, T),$$

where $\mathscr{A}_0$ can be denoted as

$$\mathscr{A}_0 = \begin{bmatrix} \mathscr{A}_{01} & \mathscr{B}_{01} \\ \mathscr{A}_{02} & \mathscr{B}_{02} \end{bmatrix}.$$

For our case, the solution is given by

$$\Lambda_0(t) = \left\{ \begin{bmatrix} 0 & I_n \end{bmatrix} e^{\mathscr{A}_0(t-T)} \begin{bmatrix} I_n \\ 0 \end{bmatrix} \right\} \left\{ \begin{bmatrix} 0_n & 0 \end{bmatrix} e^{\mathscr{A}_0(t-T)} \begin{bmatrix} I_n \\ 0 \end{bmatrix} \right\}^{-1} = \mathscr{N}_0(t)\mathscr{M}_0^{-1}(t).$$

This completes the proof.

CURRICULUM VITAE

Name Surname: Alper ÖNER

Place and Date of Birth: Şişli 25/03/1986

E-Mail: alperoner.itu@gmail.com

EDUCATION:

- **B.Sc.:** July 2011, Istanbul Technical University, Mechatronics Engineering
- **M.Sc.:** June 2009, Kocaeli University, Mechatronics Engineering

INDUSTRIAL & ACADEMIC EXPERIENCE:

Ata Technology Platform, Istanbul, Turkey,
Innovation and Transformation Department, August 2017 - ...
Job Description: *Software engineering on Machine Learning and IoT projects of franchiser of Burger King in Turkey and China,*
Position: Machine Learning Engineer (Python, Tensorflow, Optimization, Deep Learning, OpenCV, Ubuntu, Amazon Web Services(AWS), MySQL, GitHub and C, ZigBee, Arduino).

Yıldız Holding, Istanbul, Turkey,
PleiOne Lab, June 2017 - August 2017
Job Description: *Creating, developing and managing technology projects of 3rd largest biscuit producer in the world,*
Position: Creative Technologist (Managing the team, C/C++, IoT and Computer Vision).

Coordinated Science Lab, University of Illinois at Urbana-Champaign, IL, USA,
Decision and Control Laboratory, January 2016 - January 2017
Supervisor: Prof. Tamer Başar,
Project: *Stochastic Mean Field Games in Multi-Agent Systems,*
Position: Research Scientist.

Istanbul Technical University, Turkey,
MERC, Sep. 2014 - Dec. 2015, Jan. 2017 - June 2017
Supervisor: Assoc. Prof. Gülay Öke Günel,
Project: *ITUCiTSim Traffic Simulation GUI & Intelligent Transportation Project,*
Position: Software Engineer (QT Creator, C/C++, Game Theory, Multi-Agent

Control).

Istanbul Technical University, Turkey,
Mechatronics Education and Research Center, March 2010 - June 2017
Position: Teaching and Research Assistant.

Istanbul Technical University, Turkey,
Artificial Intelligence and Robotics Laboratory Jan. 2014 - Sep. 2014
Supervisor: Assoc. Prof. Sanem Sarıel,
Project: *Automated Planning and Learning Methods for Autonomous Mobile Robots,*
Position: Perception Engineer (Kinect, ROS, PCL, 3D-Camera, C/C++, LINUX).

Mechatronics Education and Research Center, Istanbul Technical University, Turkey,
Hybrid and Electrical Vehicles Control and Simulation Lab Oct. 2009 - Jan. 2014
Supervisor: Prof. Ata Muğan,
Project: *Unmanned Land Vehicles Project,*
Position: Perception Engineer (Laser, IBEO, SICK, Ultrasonic, 2D-Camera, C/C++, LabVIEW, Webots, OpenCV, ROS, LINUX).

Mechatronics Education and Research Center, Istanbul Technical University, Turkey,
Hybrid and Electrical Vehicles Control and Simulation Lab Jan. 2011 - Sep. 2011
Supervisor: Prof. Ata Muğan,
Project: *Convention of Traditional Car to Steer by Wire,*
Position: Research and Development Engineer (dSPACE, DC Motor, Microcontrollers, Motor Drivers, MATLAB, C/C++).

Robosistem Inc. June 2008 - July 2008
Position: Development Engineer (SolidWorks, PLC).

Kale Altinay Robotics July 2007 - August 2007
Position: Development Engineer (Robot Programming, PLC).

Kartes Inc. June 2007 - July 2007
Position: Production Engineer (AutoCAD).

REWARDS:

- AWARD FOR THE BEST PAPER PRESENTED BY A YOUNG RESEARCHER 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS 2017)
- The Scientific and Technological Research Council of Turkey (TUBITAK) Scholarship for the M.S. degree.
- TUBITAK Scholarship for the Ph.D. degree.
- TUBITAK International Research Scholarship for Ph.D. students.

- Graduating from Kocaeli University Mechatronics Engineering Department with a degree 2nd rank.
- Kocaeli University High Honor Student during all undergraduate terms.

PUBLICATIONS, AND PRESENTATIONS ON THE THESIS:

- **Öner A.**, Saldi N., Günel G. Ö., 2017. Mean Field Multilayer Stackelberg Differential Games in Multi-Agent Systems, ISSN 1683-3511 (print), ISSN 1683-6154 (online) *Appl. Comput. Math.,* Vol 17, No 1, 2018, 72-95.
- **Öner A.**, Başar T., Günel G. Ö., 2017. Mean Field Differential Games in Intelligent Transportation Systems. *IEEE Models and Technologies for Intelligent Transportation Systems (MT-ITS), 2017 International Conference on*, 978-1-5090-6484-7/17/ June 26-28, 2017 Napels, Italy.
- **Öner A.**, Günel G. Ö., 2015. A case study in Istanbul E-80: Intelligent speed change at the highway for reducing gas emission. *IEEE Models and Technologies for Intelligent Transportation Systems (MT-ITS), 2015 International Conference on*, 978-963-313-142-8 June 3-5, 2015 Budapest, Hungary.

OTHER PUBLICATIONS, AND PRESENTATIONS:

- Sezer V. ,Boyraz P. , Ercan Z., Dikilitaş Ç., Heceoğlu H. , **Öner A.**, Gökaşan M., 2014. Smart Mobile In-Vehicle Systems, Unmanned Ground Vehicle Otonobil: Design, Perception, and Decision Algorithms. *Springer New York*, 2014, pp 47-56.
- **Öner A.**, Günel G. Ö., 2014. Removing of the Laser Cluster Errors via Fisher Linear Discriminant Analysis in Unmanned Ground Vehicles. *Turkish National Comitee of Automatic Control 2014 National Meeting(TOK-14)*, Sept. 11-13, 2014, Kocaeli, Turkey.
- **Öner A.**, 2012. Mapping of Incremental Dynamic Environment Using Rao-Blackwellized Particle Filter. *12th International Conference Intelligent Autonomous Systems (IAS-12)*, June 26-29, 2012 Jeju Island, South Korea.
- **Öner A.**, Günel G. Ö., 2011. Mapping and Static Obstacle Detection in Unmanned Ground Vehicles. *Turkish National Comitee of Automatic Control 2011 National Meeting (TOK-11)*, Sept. 14-16, 2011, Izmir, Turkey.
- **Öner A.**, Günel G. Ö., 2011. Sensor Fusion with Occupancy Grid Map for Obstacle Segmentation in Intelligent Vehicles, *The 5th Biennial Workshop on Digital Signal Processing for In-Vehicle Systems* Sept. 4-7 , 2011, Kiel, Germany.
- Sezer V., Dikilitaş Ç., Ercan Z., Heceoğlu H., **Öner A**., Apak A., Gökaşan M. , Muğan A., 2011. Conversion of a conventional electric automobile into an unmanned ground vehicle (UGV). *IEEE International Conference on Mechatronics (ICM)* April 13-15, 2011, Istanbul, Turkey.