وزارة التعليم العالي والبحث العلمي

BADJI MOKHTAR-ANNABA UNIVERSITY
UNIVERSITE BADJI MOKHTAR ANNABA

جامعة باجي مختار
-عنابة-

Faculty of Sciences

Year : 2023/2024

Department of Mathematics

THESIS

Presented with a view to obtaining the doctorate degree

CONTROL OF HIERARCHICAL MEAN FIELD TYPE GAMES

Branch
Applied Mathematics

Speciality
ACTUARIAL SCIENCE

By
Zahrate El Oula FRIHI

SUPERVISOR : Mohamed Riad REMITA Prof. ENSIA. ALGER

In front of the jury

PRESIDENT :	Fatima Zohra BENMOSTEFA	Prof.	U.B.M. ANNABA
EXAMINER :	Med Cherif BOURAS	Prof.	U.B.M. ANNABA
EXAMINER :	Halim ZEGHDOUDI	Prof.	U.B.M. ANNABA
EXAMINER :	Abdelali EZZEBSA	MCA.	U. GUELMA

وزارة التعليم العالي والبحث العلمي

BADJI MOKHTAR-ANNABA UNIVERSITY
UNIVERSITE BADJI MOKHTAR ANNABA

جامعة باجي مختار -عنابة-

Faculté des Sciences

Département de Mathématiques

Année : 2023/2024

THÈSE

Présentée en vue de l’obtention du diplôme de Doctorat

CONTRÔLE DES JEUX DE TYPE CHAMP MOYEN HIÉRARCHIQUE

Filière
Mathématiques Appliquées

Spécialité
ACTUARIAT

Par
FRIHI Zahrate El Oula

DIRECTEUR DE THÈSE : REMITA Mohamed Riad Prof. ENSIA. ALGER

Devant le jury

PRESIDENT :	BENMOSTEFA Fatima Zohra	Prof.	U.B.M. ANNABA
EXAMINATEUR :	BOURAS Med Cherif	Prof.	U.B.M. ANNABA
EXAMINATEUR :	ZEGHDOUDI Halim	Prof.	U.B.M. ANNABA
EXAMINATEUR :	EZZEBSA Abdelali	MCA.	U. GUELMA

To my dear mother Menidjel Zakia
and my brother Frihi Alla Eddine.

Abstract

This thesis focuses on solving hierarchical leadership mean-field-type games with jump diffusion state equations and two different performing functionals using two different methods, concluding with a numerical analysis of the impact of the hierarchy on the final solution. In Paper A, we provide a solution for a hierarchical mean-field type optimal control problem with polynomial performing functional in three forms using the dynamic programming principle method. The first form is a one-level mean-field type game when the player acts simultaneously without information differences. The second form is a Stackelberg mean field type game, where the players are divided into two groups in two levels, one acting first, then the other reacting. The last case is a fully hierarchical mean-field type game when the number of hierarchy levels is equal to the number of players. Numerical investigation proves the existence of a significant effect of the hierarchy form on the optimal performing functional. In paper B, we are interested in the hierarchical mean-field type optimal control problem of electricity production in a smart grid energy market composed of $n-1$ prosumers with $n \geq 2$ and one major producer, the market leader. These n agents impose their production strategies sequentially in a hierarchal order. The market leader is at the top of the hierarchy, and the other prosumers follow him at separate levels according to their production capacity. We model the problem by a jump-diffusion stochastic differential state equation and a quadratic cost functional. The solution is established using the direct method, and two numerical scenarios are considered.

Keywords— Mean-field-type hierarchical control, Stackelberg mean field type game, polynomial objective functional, electricity production control, energy market.

ملخص

تركز هذه الأطروحة على حل و دراسة الألعاب الهرمية من نوع المتوسط الميداني حيث حالة اللعبة ممثلة بمعادلة تفاضلية عشوائية بقفزات و دالة الأداء ممثلة بدالتين مختلفين وهذا بإستخدام طريقتين مختلفتين وتطبيق أمثلة و تحليلات عددية لدراسة تأثير التسلسل الهرمي على الحل النهائي للعبة. في الورقة البحثية أ، قدمنا حلاً لمشكلة التحكم الأمثل من نوع المتوسط الميداني الهرمي مع دالة أداء كثيرة الحدود لثلاثة حالات بإستخدام طريقة مبدأ البرمجة الديناميكية. الحالة الأولى عبارة عن لعبة من نوع متوسط ميداني ذات مستوى واحد حيث يحدد كل اللاعبين إستراتيجيتهم في آن واحد دون وجود إختلافات في المعلومة. الحالة الثانية هي لعبة ستاكلبيرغ من نوع المتوسط الميداني حيث يتم تقسيم اللاعبين إلى مجموعتين على مستويين، أحدهما تقرر أولاً، ثم الأخرى تتفاعل و تقوم بالَرد. الحالة الأخيرة هي لعبة من نوع متوسط ميدان ذات ترتيب هرمي كامل أي عندما يكون عدد مستويات الهرم مساويًا لعدد اللاعبين. أثبت التطبيق العددي وجود تأثير مهم لشكل التسلسل الهرمي على دالة الأداء المثلى. في الورقة البحثية ب، نهتم بمشكلة التحكم الأمثل من نوع المتوسط الميداني الهرمي لكمية الكهرباء المنتجة في شبكة ذكية لسوق الطاقة الذي يتكون من $n-1$ منتجين و مستهلكين في آن واحد مع $n \geq 2$ ومنتج رئيسي واحد، الشركة الرائدة في السوق. يفرض n وكيل استراتيجيات الإنتاج الخاصة بهم بالتسلسل في ترتيب هرمي. يقع قائد السوق في أعلى التسلسل الهرمي، ويتبعه المنتجون الآخرون في مستويات منفصلة من الهرم وفقًا لقدرتهم الإنتاجية. نقوم بتمثيل المشكلة بمعادلة حالة تفاضلية عشوائية ذات قفزات و بدالة تكلفة تربيعية. تم إيجاد الإنتاج الأمثل باستخدام الطريقة المباشرة و دراسة تأثير التسلسل الهرمي بتطبيق سيناريوهين عدديين.

كلمات مفتاحية-- تحكم هرمي من نوع متوسط ميداني، لعبة ستاكلبيرغ من نوع متوسط ميداني، دالة أداء كثيرة الحدود، التحكم في إنتاج الكهرباء، سوق الطاقة.

Resumé

Cette thèse se focalise sur la résolution des jeux de leadership hiérarchiques de type champ moyen avec des équations d'état différentielles stochastique avec saut et deux fonctionnelles de performance différentes en utilisant deux méthodes de résolution différentes, afin de conclure par des analyses numériques sur l'impact de la hiérarchie sur la solution finale. Dans l'article A, nous fournissons une solution pour un problème de contrôle optimal de type champ moyen hiérarchique avec une fonctionnelle de performance polynomiale sous trois formes en utilisant la méthode de programmation dynamique. La première forme est un jeu de type champ moyen à un seul niveau lorsque les joueurs agissent simultanément sans différences d'information. La deuxième forme est un jeu de Stackelberg de type champ moyen, où les joueurs sont divisés en deux groupes à deux niveaux, l'un agissant en premier, puis l'autre réagissant. Le dernier cas est un jeu de type champ moyen entièrement hiérarchique où le nombre de niveaux de la hiérarchie est égal au nombre de joueurs. L'étude numérique prouve l'existence d'un effet significatif de la forme de l'hiérarchie sur la fonction de performance optimale. Dans l'article B, nous nous intéressons au problème de contrôle optimal de type champ moyen hiérarchique de la production d'électricité dans un marché énergétique de réseau intelligent composé de $n-1$ avec $n \geq 2$ prosommateurs et d'un producteur majeur, le leader du marché. Ces n agents imposent leurs stratégies de production de manière séquentielle dans un ordre hiérarchique. Le leader du marché est au sommet de la hiérarchie, et les autres prosommateurs le suivent à des niveaux séparés en fonction de leur capacité de production. Nous modélisons le problème par une équation d'état différentielle stochastique avec saut et par une fonctionnelle de coût quadratique. La solution est établie à l'aide de la méthode directe et deux scénarios numériques sont envisagés.

Mots clés— Contrôle hiérarchique de type champ moyen, jeu de Stackelberg de type champ moyen, fonction objectif polynomiale, contrôle de la production d'électricité, marché de l'énergie.

Acknowledgments

First, I thank God for giving me the power, faith, and courage to achieve this thesis.

I want to express my gratitude to my supervisor, Pr. Mohamed Riad Remita, for his advice and understanding during my time as a Ph.D. student.

Words cannot express my appreciation to Pr. Boualem Djehiche, for the time and the interest he provided to this work, his guidance, and valuable feedback, have been instrumental in shaping the direction and quality of this work.

I am also profoundly thankful to Pr. Hamidou Tembine, Dr. Salah Eddine Choutri, and Dr. Julian Barreiro-Gomez for their cooperation and expertise sharing.

Lastly, my heartfelt appreciation goes to my family especially my mother, brother, and uncles, Hocine Frihi and Nour Eddine Menidjel, for all kinds of support and encouragement they provided during my university pursuits.

Frihi Zahrate El Oula
Annaba, Algeria, 2023.

Table of Contents

Abstract i

Resumé iii

Acknowledgments iv

Table of Contents v

Table of Contents v

List of Figures vii

List of Tables viii

I Introduction 1

A short overview of game theory 3
1 Game Representations . . . 3
1.1 The Strategic Form . . . 3
1.2 The Extensive Form . . . 5
2 Games Classification . . . 6
3 Game solution concept . . . 8
3.1 Nash equilibrium . . . 8
3.2 Subgame-Perfect Nash Equilibrium . . . 12
4 Stochastic Differential Game . . . 15
4.1 Information Structures . . . 17
4.2 Nash equilibrium . . . 18
4.3 Dynamic Programming Principle . . . 18
5 Mean Field Game . . . 24
6 Mean Field Type Game . . . 27
6.1 Mean-Field-Type Game with Jump . . . 27

6.2 Solution Approaches 28
References 36

II Research Papers in Hierarchical Mean Field Type Games and Application **38**

A Hierarchical structures and leadership design in mean-field-type games with polynomial cost **40**
1 Introduction 41
2 The Setup 43
2.1 Games with polynomial cost 44
2.2 Hierarchical Leader Design and Algorithmic Approach 45
3 Nash Mean-Field-Type Equilibrium 49
4 Multiple Leaders and Multiple Followers 55
5 Fully hierarchical game 61
6 Numerical Investigation 69
6.1 Effect of the number of leaders on the total cost 69
6.2 Impact of the Hierarchical Structures 74
7 Conclusion 75
References 76

B Hierarchical Mean-Field Type Control of Price Dynamics for Electricity in Smart Grid **78**
1 Introduction 79
2 The Setup 80
3 Hierarchical mean-field type control problem 82
4 Main result 83
5 Numerical examples 93
6 Conclusion 97
References 98

Perspective **100**

List of Figures

1 The Entry Game Tree . . . 6

2 The Entry Subgames . . . 12

3 The Nash equilibrium of the last entry subgame . . . 13

4 The new tree form of the entry game . . . 14

5 The subgame-perfect Nash equilibrium of the entry game . . . 14

A.1 Different hierarchical designs and their solution concepts are considered in this paper. . . . 46

A.2 Possible combinations in the hierarchical leadership design for two decision-makers. Ordered Bell number $B(2) = 3$. . . . 48

A.3 Possible combinations in the hierarchical leadership design for three decision-makers. Ordered Bell number $B(3) = 13$. . . . 48

A.4 Number of possible hierarchical structures for a given set of decision-makers described by the ordered Bell number $B(I)$. . . . 49

A.5 Evolution of the differential equations $\dot{\bar{\alpha}}^{\text{leader/follower}}$, and the corresponding initial values for different numbers of leaders in the **homogeneous** scenario. 71

A.6 Evolution of the differential equations $\dot{\bar{\alpha}}^{\text{leader/follower}}$, and the corresponding initial values for different numbers of leaders in the **heterogeneous** scenario. 73

A.7 Evolution of the sum of differential equations and the corresponding total cost for the **heterogeneous** scenario. . . . 73

A.8 Evolution of the differential equations $\sum_{j\in\mathcal{I}} \dot{\bar{\alpha}}_j(t)$, and the corresponding initial values for different hierarchical structures in the **heterogeneous** scenario. . . . 75

B.1 Mean-field-type hierarchical performance corresponding to Scenario 1. . . . 95

B.2 Mean-field-type hierarchical performance corresponding to Scenario 2. . . . 96

List of Tables

1 The Prisoner's Dilemma Game Matrix 5

A.1 Summary of $\bar{\alpha}_0^{\text{leader}}$, $\bar{\alpha}_0^{\text{follower}}$, and $S(|\mathcal{I}_L|, \delta_{x_0})$ for the different number of leaders in the **homogeneous** scenario. 70

A.2 Summary of $\bar{\alpha}_0^{\text{leader}}$, $\bar{\alpha}_0^{\text{follower}}$, and $S(|\mathcal{I}_L|, \delta_{x_0})$ for the different number of leaders in the **heterogeneous** scenario. 72

A.3 Total cost for the different hierarchical orders in a three-player case in the **heterogeneous** scenario. 74

B.1 Parameters for scenario 1 and 2 94

B.2 The revenues of the agents for scenarios 1 and 2. 94

Part I

Introduction

A short overview of game theory

Game theory is a discipline that uses mathematical models to study the strategic interactions of multiple rational decision-makers, known as players. A player can have very different meanings; that can be a human, a plant, an animal, a robot, etc., which means that the notion of games dates back to the beginning of time. However, the theory of games was born in 1944 through the publication of the famous book "Theory of Games and Economic Behavior" by the mathematician John von Neumann and the economist Oskar Morgenstern (Neumann & Morgenstern, 1944). This foundation was strengthened by the work of John Nash in 1951, who introduced the notion of a solution for non-zero-sum games in (Nash, 1951). Since then, game theory has experienced significant mathematical development and many applications in various disciplines: social sciences, biology, computer science, economics sciences, politics, etc. The success of the field is particularly remarkable in economics, where several game theorists received the Nobel Prize.
In this chapter, we present a short tour of game theory by introducing the basic concepts that will be used in the second part of the thesis.

1 Game Representations

In game theory, the mathematical definition of a game takes two forms.

1.1 The Strategic Form

Definition 1.1. *A game in strategic or normal form is defined by three features.*

1. *A set of I players, $\mathcal{I} = \{1, ..., I\}$.*

2. *For each player $i, i \in \mathcal{I}$, a collection of all possible strategies*

$$U_i = \{u_i^1, u_i^2, ..., u_i^{k_i}\}.$$

If each player chooses a strategy u_i we can represent the outcome of the game as a vector

$$u = \{u_1, u_2, ..., u_I\}.$$

3. *For each player $i, i \in \mathcal{I}$, a payoff function l_i (named also a utility function), which gives player i a value for each game outcome*

$$l_i : U = \prod_{i \in \mathcal{I}} U_i \to R.$$

□

Remark 1.1. *In the normal form*

- *The players act simultaneously without knowing each other's actions.*
- *A matrix is used to represent the game.* □

Example 1.1 (The Prisoner's Dilemma by Merrill Flood and Melvin Dresher (1950))**.** *Two criminals are arrested by the police after an armed robbery and interrogated in separate rooms. Each is given the choice of denouncing his accomplice for using the gun (strategy D) and receiving a reduced sentence, or remaining silent (strategy N) and being convicted only of the robbery.*
This situation is a kind of game in the normal form, where

- *The players set $I = 2, \mathcal{I} = \{1, 2\} = \{Suspect1, Suspect2\}$.*
- *The strategies set of each player's $k_1 = K_2 = 2, U_1 = \{u_1^1 = D, u_1^2 = N\}$,* $U_2 = \{u_2^1 = D, u_2^2 = N\}$.
- *The outcome of the game $u = \{(N, N), (N, D), (D, N), (D, D)\}$.*
- *The payoff of each player (The years of imprisonment to which each is sentenced)*

$$\begin{aligned} l_1(u_1^1 = D, u_2^2 = N) = l_2(u_1^2 = N, u_2^1 = D) = 1, \\ l_1(u_1^2 = N, u_2^1 = D) = l_2(u_1^1 = D, u_2^2 = N) = 5, \\ l_1(N, N) = l_2(N, N) = 2, \\ l_1(D, D) = l_2(D, D) = 4. \end{aligned}$$

- *The matrix that represents the game*

		Suspect 2	
		D	N
Suspect 1	D	(4,4)	(1,5)
	N	(5,1)	(2,2)

Table 1: The Prisoner's Dilemma Game Matrix

1.2 The Extensive Form

Definition 1.2. *A game in extensive form is defined by*

1. *A set of I players, $\mathcal{I} = \{1, ..., I\}$.*
2. *The order of play for each player.*
3. *The information that each player has about the moves of the other players.*
4. *A set of actions for each player $i \in \mathcal{I}, U_i = \{u_i^1, u_i^2, ..., u_i^{k_i}\}$.*
5. *A payoff function $l_i(.)$ for each player $i \in \mathcal{I}$.* □

Remark 1.2. *In the extensive form*

- *The players act sequentially over time, they can play once or multiple times with the knowledge of the actions played before them.*
- *A tree used to represent the game, containing*
  - *initial vertex (node),*
  - *decision vertices,*
  - *terminal vertices,*
  - *branches connecting each node to its successors.* □

Example 1.2 (The Entry Model by Dixit's (1979))**.** *We have two firms, The first (Firm 1) is an incumbent monopolist making \$9 million in profits, and the second (Firm 2) is a new competitor who wants to enter the market (strategy E). If Firm 2 decides to enter, the monopolist can either fight (strategy F) by slashing prices or accommodate and create a common monopoly (strategy C).*

This economic example is a kind of game in the extensive form, where

- *The players set* $I = 2, \mathcal{I} = \{1, 2\} = \{Firm1, Firm2\}$.
- *The strategies set of each player's* $K_1 = K_2 = 2, U_1 = \{u_1^1 = F, u_1^2 = C\}$, $U_2 = \{u_2^1 = E, u_2^2 = O\}$.
- *The order of play Firm 2 acts first, then Firm 1 reacts.*
- *The information that Firm 1 contains Firm 2 played* u_2^1.
- *The outcome of the game*

$$U = \{(O), (E, F), (E, C)\}.$$

- *The payoff of each player (The profits of both companies)*

$$l_1(O) = 0, l_2(O) = 9,$$
$$l_1(E, F) = -2, l_2(E, F) = -4,$$
$$l_1(E, C) = 3, l_2(E, C) = 5.$$

- *The tree that represents the game*

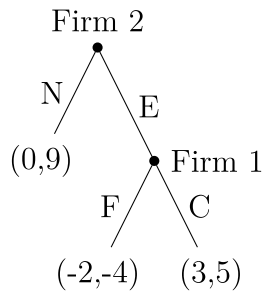


Fig. 1: The Entry Game Tree

2 Games Classification

According to the different contexts in which the players interact, games can be classified in game theory as

- **Cooperative and non-cooperative games**

  In cooperative games, players form a group or sub-groups and make enforceable agreements to apply certain strategies in order to reach certain payoffs.

While, in non-cooperative games, players prefer to play individually in a selfish manner, even when they have the opportunity to communicate, where each of them tries to maximize his profit using his best strategy. For more understanding, let us take the example of the prisoner's dilemma, which is a non-cooperative game, but, if the suspects can communicate, they will commit to keeping quiet, which transforms the game into a cooperative one.

- **Static and dynamic games**
  A static game is a game that can be played simultaneously or at different times, but in either case, the players have no information about the decisions of each other. Otherwise, we have the dynamic game, which is a sequential or repeated game, where each player has a set of information about the previous actions of the game. To deal with the static game we use the strategic form, and to deal with the dynamic game we use the extensive form. The relation between these two types of games is that the dynamic game is a generalization of the static game when the last one is repeated a finite or infinite number of times.

- **Complete and incomplete information games**
  A complete information game is a game where all players have equal access to all relevant information about the game. This means that every player knows the game rules, all the other players, the strategies available to each of them, and the associated payoffs. Various games with complete information can be cited, such as chess, tic-tac-toe, rock-paper-scissors, and the prisoner's dilemma. In contrast, incomplete information games or Bayesian games involve hidden or private information, where players may have partial or varying knowledge about the game. Poker is a well-known example of an incomplete information game. Each player holds private cards that are hidden from the other players. Also, the entry game can be a Bayesian game if firms don't know the payoffs of each other.

- **Perfect and imperfect information games**
  We say that a game is a perfect information game if all decisions taken during this game are visible to all players, and each player knows all strategies played before him. But, when the player ignores some or all of these strategies, the game is said to be an imperfect information game. A static game is a game with imperfect information,

while a dynamic game can be with perfect or imperfect information.

- **Symmetric and asymmetric games**

  For a game to be symmetric, all the players must have the same set of strategies, and there must be no inherent advantage or disadvantage associated with a particular player so that the players' identities do not influence the payoffs. If any of the previous conditions are not verified, then the game is said to be asymmetric. As a result, we will get symmetry payoffs like in the prisoner's dilemma example or asymmetry payoffs like in the entry example.

- **Zero-sum and non-zero-sum games**

  A zero-sum game is a particular case of the constant-sum game, where the sum of payoffs equals the same constant value each time the players change their decisions; in the zero-sum game case, this constant is zero. These types of games are based on the idea that if a party of players wins, the other party loses, and vice versa. In a non-zero-sum game, the value of the constant is not zero and changes as the players' decisions change. Also, in the non-zero-sum game, the notion of win-loss is not necessarily exciting, i.e., if one wins, it does not mean that the other loses so the game can be a win-win or a loss-loss. Returning to the prisoner's dilemma example, the sum of the two payoffs found in each cell of the table has different strictly positive values (4,6,6,8), indicating that the game is a non-zero-sum game.

3 Game solution concept

A solution concept is certain conditions imposed on a game to have a solution and on a strategy profile to be a solution. A game solution is a point of equilibrium among all players presenting the ideal outcome of a game. A game may have no solution, one solution, or multiple solutions. In game theory, there are many solution concepts, depending on the type of game. For the interest of our papers in this section, we define the Nash equilibrium solution concept and the Stagelberg equilibrium solution concept.

3.1 Nash equilibrium

The most famous solution concept in game theory is the Nash equilibrium, developed by the American mathematician John Forbes Nash in 1950 (Nash, 1951) and named after him. Its importance resides in predicting the solution of a wide large of games. The Nash equilibrium was first interpreted for a static, non-cooperative game with complete information before being extended to other types of games.

Definition 3.1. ***Best Response*** *A strategy $u_i \in U_i$ of player i is the best response (BR) to his opponents' strategies $u_{-i} = (u_1, ..., u_{i-1}, u_{i+1}, ..., u_I)$ if*

$$\forall i \in \mathcal{I}, \forall u'_i \in U_i; \quad l_i(u_i, u_{-i}) \geq l_i(u'_{-i}, u_{-i}),$$

with $U_i = u'_1, ..., u'_{i-1}, u_i, u'_{i+1}, ..., u'_I$. □

Definition 3.2. ***Best Response Correspondence*** *The best response correspondence $BR_i(u_{-i})$ of player i is the application defined by*

$$BR_i : U_{-i} = \underset{j \in \mathcal{I} \setminus \{i\}}{\times} U_j \to 2^{U_i}$$
$$u_{-i} \to BR_i(u_{-i}) = arg \max_{u'_i \in U_i} l_i(u'_{-i}, u_{-i}).$$

2^{U_i} is the power set that represents all the possible subsets of U_i. □

Remark 3.1.

- *$BR_i(u_{-i})$ is the player i set of all strategies that are the best response to u_{-i}. If u_i is the unique best response to u_{-i} then $BR_i(u_{-i}) = u_i$.*
- *The best response correspondence $BR(u)$ of the game is the application defined by*

$$BR : U = \underset{i \in \mathcal{I}}{\times} U_i \to 2^U$$
$$u \to BR(u) = \prod_{i \in \mathcal{I}} BR_i(u_{-i}).$$

- *u is a fixed point of $BR(u)$ if $u \in BR(u)$.* □

Definition 3.3. ***Nash equilibrium*** *A strategy profile $u^* = (u^*_1, u^*_2, ..., u^*_I)$ is a Nash equilibrium if no player is interested in unilaterally deviating from his strategy u^*_i, considering the expected best decisions of the other players u^*_{-i}. Formally, u^* is a Nash equilibrium if*

$$\forall i \in \mathcal{I}, \forall u_i \in U_i; \quad l_i(u^*_i, u^*_{-i}) \geq l_i(u_i, u^*_{-i}). \tag{1}$$

□

Definition 3.4. *A strategy profile u^* is a Nash equilibrium if*

$$u^* \in BR(u^*), \tag{2}$$

otherwise, a strategy profile u^* is a Nash equilibrium if

$$\forall i \in \mathcal{I}, u^*_i \in BR(u^*_{-i}).$$

□

Remark 3.2. *The last characterization defines the Nash equilibrium as a solution to a fixed point problem. To prove the existence of this point, Nash used the Kakutani (in 1941) and the Brouwer (in 1912) fixed point theorems.* □

Example 3.1. *To define the Nash equilibriums for the Prisoner's Dilemma game, we need to express the intersection points of the best responses of all players, which verifies (2.2) as follows*

- *The suspect 1 best responses* $BR_1(D) = \{D\}$, $BR_1(N) = \{D\}$.
- *The suspect 2 best responses* $BR_2(D) = \{D\}$, $BR_2(N) = \{D\}$.
- *The best responses correspondences of the game*
  - $BR(D, D) = BR_1(D) \times BR_2(D) = \{D\} \times \{D\} = \{(D, D)\}$,
  - $BR(D, N) = BR_1(D) \times BR_2(N) = \{(D, D)\}$,
  - $BR(N, D) = BR_1(N) \times BR_2(D) = \{(D, D)\}$,
  - $BR(N, N) = BR_1(N) \times BR_2(N) = \{(D, D)\}$.
- *The strategy profile that verifies (2.2) is* $BR(D, D) = (D, D)$.

⇒ *This game has only one Nash equilibrium* $u^* = (D, D)$.

Example 3.2. *(The Cournot Model by Antoine Augustin Cournot (1838)) In the Cournot oligopoly model (Cournot, 1838), the market is composed of n firms that compete by choosing simultaneously their production quantities* q_i *of the same single good. The market price is given by the linear inverse demand function* $P = a - \sum_{i=1}^{I} q_i = a - Q$ *and the profit of each producer i is equal to its revenue minus its cost*

$$l_i(q_1, ..., q_n) = q_i P - C_i(q_i).$$

The strategic form of this model when $I = 2$ is known as Cournot's duopoly game defined as

- *The players set* $I = 2, \mathcal{I} = \{1, 2\} = \{Producer1, Producer2\}$.
- *The strategies set of each player's* $U_i = [0, +\infty[, i = \{1, 2\}$.
- *The outcome of the game* $u = (q_1, q_2)$, *for* $q_i \in U_i, i = \{1, 2\}$.
- *The payoff of each player (the profit of each firm)*

$$l_1(q_1, q_2) = q_1(a - (q_1 + q_2)) - C_1(q_1),$$
$$l_2(q_1, q_2) = q_2(a - (q_1 + q_2)) - C_2(q_2).$$

To find the Nash equilibrium for this game, we will take the following steps

- ***Step 1.*** *Write the profit maximization problem for each firm*

$$\max_{q_1} l_1(q_1, q_2) = \max_{q_1} (q_1(a - (q_1 + q_2)) - cq_1)),$$
$$\max_{q_2} l_2(q_1, q_2) = \max_{q_2} (q_2(a - (q_1 + q_2)) - cq_2)).$$

- ***Step 2.*** *Find the best response function for each firm by solving the previous problems*

$$BR_1(q_2) = q_1^* = \underset{q_1}{argmax}\ (q_1(a - (q_1 + q_2)) - cq_1)),$$
$$BR_2(q_1) = q_2^* = \underset{q_2}{argmax}\ (q_2(a - (q_1 + q_2)) - cq_2)),$$

  which is given by the first-order condition,

$$\frac{\partial l_1}{\partial q_1} = -2q_1 - q_2 - c + a = 0 \Rightarrow BR_1(q_2) = q_1^* = \frac{-q_2 - c + a}{2},$$
$$\frac{\partial l_2}{\partial q_2} = -2q_2 - q_1 - c + a = 0 \Rightarrow BR_2(q_1) = q_2^* = \frac{-q_1 - c + a}{2}.$$

- ***Step 3.*** *Find the Nash equilibrium* (q_1^*, q_2^*) *by solving the two best response equations simultaneously*

$$\begin{cases} BR_1(q_2^*) = q_1^* \\ BR_2(q_1^*) = q_2^* \end{cases} \Longrightarrow \begin{cases} q_1^* = \dfrac{-q_2^* - c + a}{2} \\ q_2^* = \dfrac{-q_1^* - c + a}{2} \end{cases} \Longrightarrow \begin{cases} q_1^* = q_2^* = \dfrac{a - c}{3}. \end{cases}$$

  The game solution is $u^* = (\frac{a - c}{3}, \frac{a - c}{3})$.

3.2 Subgame-Perfect Nash Equilibrium

The subgame-perfect Nash equilibrium, introduced by Reinhard Selten in 1975 (Selten, 1974), is a solution concept for dynamic games with complete information. The concept is to decompose the game into sequential subgames, where the Nash equilibrium of each subgame is a perfect subgame equilibrium for the original game. To solve these subgames, consequently, the full game, we use the backward induction algorithm.

Definition 3.5. *A subgame of an extensive form game is a single decision node with all of its successor's nodes and branches.* $\square$

Remark 3.3.

- *Every game in extensive form has at least one subgame: the game itself.*
- *The entry game has two subgames.*

(a) The last subgame   (b) The first subgame

Fig. 2: The Entry Subgames

$\square$

Definition 3.6. ***Subgame-Perfect Nash Equilibrium*** *A strategy profile $u^* = (u_1^*, ..., u_I^*)$ is a subgame-perfect Nash equilibrium of an extensive form game if it constitutes a Nash equilibrium for every subgame.* $\square$

Backward induction

The backward induction method was discovered by Arthur Cayley in 1875 while solving the secretary problem. In game theory, the first use of backward induction was in the solution of zero-sum two-player games by John von Neumann and Oskar Morgenstern in their book (Neumann & Morgenstern, 1944). For the determination of subgame-perfect equilibria of a dynamic game with complete and perfect information, we use the following backward induction algorithms

- **Step 1.** Compute a Nash equilibrium for the terminal subgame of the game tree, given all previous players' decisions.
- **Step 2.** Delete this subgame and replace its initial node with the payoff vector corresponding to its equilibrium.
- **Step 3.** Compute a Nash equilibrium for the new terminal subgame of the new game tree, given all the decisions of the players who moved previously and the optimal action for the player who moves in the deleted subgame.
- **Step 4.** Repeat steps 2 and 3 until you have only the subgame that starts at the first node of the game.
- **Step 5.** Compute a Nash equilibrium for this first subgame, given all the optimal actions for the players who move in the deleted subgames. This Nash equilibrium is a subgame-perfect Nash equilibrium of the whole game.

Example 3.3. *As an example of the backward induction algorithm, we solve the entry game*

- ***Step 1.*** *The Nash equilibrium of the last subgame is* (E, C) *where,* $C = B_1(E) = \underset{u_i}{argmax}\ l_1(u_i, E)$.

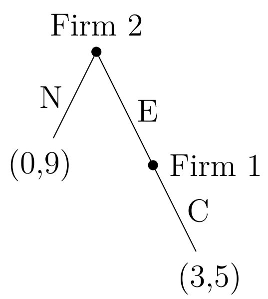


Fig. 3: The Nash equilibrium of the last entry subgame

- ***Step 2.*** *The new game tree form is*

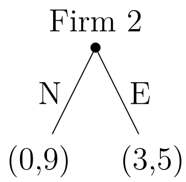


Fig. 4: The new tree form of the entry game

- ***Step 3.*** *The subgame-perfect Nash equilibrium of the entry game is* (E, C) *where,* $E = B_2(u_1^*) = \underset{u_i}{argmax}\ l_2(u_i, u_1^*)$.

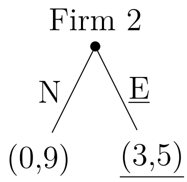


Fig. 5: The subgame-perfect Nash equilibrium of the entry game

Example 3.4. *(The Stackelberg Model by Heinrich Freiherr von Stackelberg (1934)) The Stackelberg leadership model (Stackelberg, 1948) is an extensive form of the Cournot model. In the Stackelberg duopoly model, firms move sequentially. The firm that moves first is called the leader, and the firm that reacts second is called the follower. We apply the backward induction principle to derive the Stackelberg subgame-perfect Nash equilibrium. That is, we start by determining the optimal strategy of the follower and work backward to determine the leader's optimal strategy given that of the follower.*

- ***Step 1.*** *Solve the follower maximization problem given the leader's decision using the first-order condition as in the Cournot model*

$$\frac{\partial l_2}{\partial q_2} = -2q_2 - q_1 - c + a = 0 \Rightarrow BR_2(q_1) = q_2^* = \frac{-q_1 - c + a}{2}.$$

- ***Step 2.*** *Solve the leader maximization problem given the follower optimal strategy using the first-order condition*

$$\begin{aligned}\frac{\partial l_1}{\partial q_1} &= -2q_1 - BR_2(q_1) - q_1 BR_2'(q_1) - c + a = 0\\ &\Leftrightarrow -2q_1 + \frac{q_1+c-a}{2} + \frac{q_1}{2} - c + a = 0\\ &\Leftrightarrow -2q_1 - \frac{q_1+c-a}{2} - \frac{q_1}{2} - c + a = 0\\ &\Leftrightarrow BR_1(q_2) = q_1^* = \frac{a-c}{2}.\end{aligned}$$

⇒ From steps 1 and 2, the subgame-perfect Nash equilibrium of the Stackelberg duopoly game is $u^* = (q_1^*, q_2^*)$, where

$$\begin{aligned} q_1^* &= \frac{a-c}{2},\\ q_2^* &= \frac{-q_1^* - c + a}{2} = \frac{a-c}{2} - \frac{a-c}{4}\\ \Leftrightarrow q_2^* &= \frac{a-c}{4}.\end{aligned}$$

Comparing the output and benefit of each firm in the Stackelberg and Cournot models, we find that

$$\begin{aligned} q_1^{*S} > q_1^{*C} \quad &\Rightarrow \quad l_1^S(q_1^{*S}, q_2^{*S}) > l_1^C(q_1^{*C}, q_2^{*C}),\\ q_2^{*S} < q_2^{*C} \quad &\Rightarrow \quad l_2^S(q_1^{*S}, q_2^{*S}) < l_2^C(q_1^{*C}, q_2^{*C}).\end{aligned}$$

In the Stackelberg subgame perfect equilibrium, the leader makes the highest revenue compared to the follower and the first firm in the Cournot Nach equilibrium. This extra profit of the leader is due to the advantage of the first move. Therefore, we can conclude that sequential moves and the order of these moves have a significant impact on the solution of the game.

4 Stochastic Differential Game

Stochastic differential games emerged in the 1960s when researchers began using stochastic calculus and stochastic optimal control to model and solve dynamic games that unfold over a continuous time horizon. A stochastic differential game with I players is defined on probability space $(\Omega, \mathbb{F}, P)$ and composed of

- I natural filtrations $\mathbb{F} := \left(\mathbb{F}^1, ..., \mathbb{F}^I\right)$, where for all $1 \leq i \leq I, \mathbb{F}^i := \left(\mathcal{F}_s^i\right)_{t \leq s \leq T}$ formalize the information available to each player i at the time s to choose his action.

- I state equations as Itô stochastic differential equations (SDEs) describing the evolution of the players in the game (the game state of each player).

$$\begin{cases} dx_1(s) = b_1(s, x_1(s), ..., x_I(s), u_1(s), ..., u_I(s))ds + \sigma_1(s, x_1(s), ..., x_I(s), u_1(s), ..., u_I(s))dB_1(s), \\ dx_2(s) = b_2(s, x_1(s), ..., x_I(s), u_1(s), ..., u_I(s))ds + \sigma_2(s, x_1(s), ..., x_I(s), u_1(s), ..., u_I(s))dB_2(s), \\ \quad \vdots \\ dx_I(s) = b_I(s, x_1(s), ..., x_I(s), u_1(s), ..., u_I(s))ds + \sigma_I(s, x_1(s), ..., x_I(s), u_1(s), ..., u_I(s))dB_I(s), \end{cases} \tag{3}$$

  where for each player i

  - $x_i(t) := x_i^t, 0 \leq t \leq s \leq T$ is a given initial state,
  - $b_i \ : [t, T] \times \mathbb{R}^I \times \prod_{i=1}^{I} U_i \to \mathbb{R}$ is the drift function,
  - $\sigma_i : \ [t, T] \times \mathbb{R}^I \times \prod_{i=1}^{I} U_i \to \mathbb{R}$ is the diffusion coefficient function,
  - $u_i(s) \in U_i = \mathbb{R}$ is the action (control),
  - $B_i(s)$ is the Brownian motion.
  - $x_i(s) \equiv x_i^{t,x_i^t}(s)$ is the unique solution to his state equation under the initial time t and the initial state x_i^t.

- I expected objective functions, which can be benefits or utilities in the case of maximization, and costs or losses in the case of minimization.

$$\begin{cases} L_1(x_1(t), ..., x_I(t), u_1(t), ..., u_I(t)) = \mathbb{E}\left(\int_t^T l_1(s, x_1(s), ..., x_I(s), u_1(s), ..., u_I(s))ds\right) \\ \qquad\qquad + \mathbb{E}\left(h_1(x_1(T), ..., x_I(T))\right), \\ L_2(x_1(t), ..., x_I(t), u_1(t), ..., u_I(t)) = \mathbb{E}\left(\int_t^T l_2(s, x_1(s), ..., x_I(s), u_1(s), ..., u_I(s))ds\right) \\ \qquad\qquad + \mathbb{E}\left(h_2(x_1(T), ..., x_I(T))\right), \\ \qquad \vdots \\ L_I(x_1(t), ..., x_I(t), u_1(t), ..., u_I(t)) = \mathbb{E}\left(\int_t^T l_I(s, x_1(s), ..., x_I(s), u_1(s), ..., u_I(s))ds\right) \\ \qquad\qquad + \mathbb{E}\left(h_I(x_1(T), ..., x_I(T))\right), \end{cases} \tag{4}$$

where for each player i

- $l_i: \ [t,T] \times \mathbb{R}^I \times \prod_{i=1}^{I} U_i \to \mathbb{R}$ is the instantaneous benefit or loss function,
- $h_i : \mathbb{R}^I \to \mathbb{R}$ is the terminal benefit or loss function.

4.1 Information Structures

In a stochastic differential game, there are many different information structures available to each player concerning the state of the game and the previous actions of the other players at the time s when he chooses his control. Let us define the most common ones.

Definition 4.1. ***Open-loop strategy*** *A control action $u_i(s)$ of player i is selected according to an open-loop strategy if*

$$u_i(s) = \gamma_i(x_i^t, s),$$

where $\gamma_i(.,.) : \mathbb{R}^I \times [t,T] \to U_i$ a strategy function. □

Definition 4.2. ***Closed-loop strategy*** *A control action $u_i(s)$ of player i is selected according to a closed-loop strategy if*

$$u_i(s) = \gamma_i(x_i^{[t,s]}, s),$$

where $x_i^{[t,s]}$ is the state of player i from time t to time s, $t < s$. □

Definition 4.3. ***State-feedback strategy*** *A control action $u_i(s)$ of player i is selected according to a state-feedback strategy if*

$$u_i(s) = \gamma_i(x_i(s), s).$$

□

4.2 Nash equilibrium

An I-tuple $(u_1^*(s), u_2^*(s), ..., u_I^*(s))$ of controls is a Nash equilibrium of a stochastic differential game under one of the information structures if they solve the following I stochastic optimal control problems

$$\begin{cases} L_1\left(x_1(t), ..., x_I(t), u_1^*(t), u_2^*(t), ..., u_I^*(t)\right) = \max\limits_{u_1(s)} \left(\min\limits_{u_1(s)}\right) L_1\left(x_1(t), ..., x_I(t), u_1(t), u_2^*(t), ..., u_I^*(t)\right), \\ L_2\left(x_1(t), ..., x_I(t), u_1^*(t), u_2^*(t), ..., u_I^*(t)\right) = \max\limits_{u_2(s)} \left(\min\limits_{u_2(s)}\right) L_2\left(x_1(t), ..., x_I(t), u_1^*(t), u_2(t), ..., u_I^*(t)\right), \\ \qquad \vdots \\ L_I\left(x_1(t), ..., x_I(t), u_1^*(t), u_2^*(t), ..., u_I^*(t)\right) = \max\limits_{u_I(s)} \left(\min\limits_{u_I(s)}\right) L_I\left(x_1(t), ..., x_I(t), u_1^*(t), u_2^*(t), ..., u_I(t)\right). \end{cases} \tag{5}$$

Remark 4.1. *In systems (3) and (4), players interact among states and controls, but we can have other systems in which the players interact only among states or only among controls.* □

4.3 Dynamic Programming Principle

Originated by the American mathematician Richard Bellman in 1950, the dynamic programming principle (DPP) is one of the powerful methods that can be used to solve the system (5). The basic idea of this mathematical technique is to divide the principal optimal control problem into subproblems with different initial times and states, where the optimal control for these subproblems is the optimal control for the global problem, known as the Bellman principle of optimality. To solve the stochastic optimal control, one must solve a nonlinear second-order partial differential equation named the Hamilton-Jacobi-Bellma equation (HJB), which relates all the subproblems.

Definition 4.4. *The value functions or the cost-to-go functions of I players at initial time t and initial state x_i^t is given by*

$$\begin{cases} V_1\left(t, x_1^t, ..., x_I^t\right) = \max\limits_{u_1(s)} \left(\min\limits_{u_1(s)}\right) \ L_1\left(x_1(t), ..., x_I(t), u_1(t), u_2^*(t), ..., u_I^*(t)\right), \\ V_2\left(t, x_1^t, ..., x_I^t\right) = \max\limits_{u_2(s)} \left(\min\limits_{u_2(s)}\right) \ L_2\left(x_1(t), ..., x_I(t), u_1^*(t), u_2(t), ..., u_I^*(t)\right), \\ \qquad\qquad \vdots \\ V_I\left(t, x_1^t, ..., x_I^t\right) = \max\limits_{u_I(s)} \left(\min\limits_{u_I(s)}\right) \ L_I\left(x_1(t), ..., x_I(t), u_1^*(t), u_2^*(t), ..., u_I(t)\right), \end{cases} \tag{6}$$

with $V_i\left(T, x_1^t, ..., x_I^t\right) = h_i(x_1^t, ..., x_I^t)$ is the value function of the player i at terminal time T.

□

Theorem 4.1. *For any $(\tau, x_i^\tau) \in [t, T] \times \mathbb{R}$ where τ satisfies $t \leq \tau \leq T$, the value functions defined in (6) can be computed backward from the value functions $\left(V_1\left(\tau, x_1^\tau, ..., x_I^\tau\right), ..., V_I\left(\tau, x_1^\tau, ..., x_I^\tau\right)\right)$ as follow*

$$\begin{cases} V_1\left(t, x_1^t, ..., x_I^t\right) = \max\limits_{u_1(s)} \left(\min\limits_{u_1(s)}\right) \ \mathbb{E}\left[\int_t^\tau l_1(s, x_1(s), ..., x_N(s), u_1(s), u_2^*(s), ..., u_I^*(s))ds + V_1\left(\tau, x_1^\tau, ..., x_I^\tau\right)\right], \\ V_2\left(t, x_1^t, ..., x_I^t\right) = \max\limits_{u_2(s)} \left(\min\limits_{u_2(s)}\right) \ \mathbb{E}\left[\int_t^\tau l_2(s, x_1(s), ..., x_I(s), u_1^*(s), u_2(s), ..., u_I^*(s))ds + V_2\left(\tau, x_1^\tau, ..., x_I^\tau\right)\right], \\ \qquad \vdots \\ V_I\left(t, x_1^t, ..., x_I^t\right) = \max\limits_{u_I(s)} \left(\min\limits_{u_I(s)}\right) \ \mathbb{E}\left[\int_t^\tau l_I(t, x_1(s), ..., x_I(s), u_1^*(s), u_2^*(s), ..., u_I(s))ds + V_I\left(\tau, x_1^\tau, ..., x_I^\tau\right)\right]. \end{cases} \tag{7}$$

These I equations are called the Bellman equations or the dynamic programming equations.
□

Proof. From (4) we have for all $1 \leq i \leq I$

$$\begin{aligned} L_i(x_1(t), ..., x_i(t), ..., x_I(t), u_1(t), ..., u_i(t), ..., u_I(t)) =& \mathbb{E}\left(\int_t^T l_i(s, x_1^{t,x_1^t}(s), ..., x_I^{t,x_I^t}(s), u_1(s), ..., u_I(s))ds \right. \\ &\left. + h_i(x_i^{t,x_i^t}(T), ..., x_I^{t,x_I^t}(T))\right) \\ =& \mathbb{E}\left(\int_t^\tau l_i(s, x_1^{t,x_1^t}(s), ...x_I^{t,x_I^t}(s), u_1(s), ..., u_I(s))ds\right) \\ +& \mathbb{E}\left(\int_\tau^T l_i(s, x_1^{t,x_1^t}(s), ..., x_I^{t,x_I^t}(s), u_1(s), ..., u_I(s))ds \right. \\ &\left. + h_i(x_i^{t,x_i^t}(T), ..., x_I^{t,x_I^t}(T))\right), \end{aligned}$$

Applying the iterated conditioning and the strong Markov property that consists if τ: $0 \leq t \leq \tau \leq T$ is a stopping time with respect to $\mathbb{F}_i$, then

$$\mathbb{E}\big[f(x^{t,x^t}(s=\tau+\delta)) \mid \mathcal{F}^i_\tau\big] = \mathbb{E}\big[f(x^{\tau,x^\tau}(s=\delta))\big],$$

where $\delta = s - \tau$, see page 117 in (Øksendal, 2003) for the full definition.

We get:

$$\begin{aligned}
&\mathbb{E}\bigg(\int_\tau^T l_i(s, x_1^{t,x_1^t}(s), ..., x_I^{t,x_I^t}(s), u_1(s), ..., u_I(s))ds + h_i(x_i^{t,x_i^t}(T), ..., x_I^{t,x_I^t}(T))\bigg) \\
&\qquad = \mathbb{E}\Bigg(\mathbb{E}\Bigg(\bigg(\int_\tau^T l_i(s, x_1^{t,x_1^t}(s), ..., x_I^{t,x_I^t}(s), u_1(s), ..., u_I(s))ds + h_i(x_i^{t,x_i^t}(T), ..., x_I^{t,x_I^t}(T))\bigg) \bigg| \mathcal{F}^i_\tau\Bigg)\Bigg) \\
&\qquad = \mathbb{E}\Bigg(\mathbb{E}\bigg(\int_\tau^T l_i(s, x_1^{\tau,x_1^\tau}(s), ..., x_I^{\tau,x_I^\tau}(s), u_1(s), ..., u_I(s))ds + h_i(x_i^{\tau,x_i^\tau}(T), ..., x_I^{\tau,x_I^\tau}(T))\bigg)\Bigg) \\
&\qquad = \mathbb{E}\bigg(L_i(x_1(\tau), ..., x_i(\tau), ..., x_I(\tau), u_1(\tau), ..., u_i(\tau), ..., u_I(\tau))\bigg),
\end{aligned}$$

$$\begin{aligned}
\Rightarrow L_i(x_1(t), ..., x_i(t), ..., x_I(t), u_1(t), ..., u_i(t), ..., u_I(t)) &= \mathbb{E}\bigg(\int_t^\tau l_i(s, x_1^{t,x_1^t}(s), ...x_I^{t,x_I^t}(s), u_1(s), ..., u_I(s))ds\bigg) \\
&+ \mathbb{E}\bigg(L_i(x_1(\tau), ..., x_i(\tau), ..., x_I(\tau), u_1(\tau), ..., u_i(\tau), ..., u_I(\tau))\bigg),
\end{aligned}$$

In order to complete the proof of Theorem 4.1 and have:

$$\begin{aligned}
V_i\left(t, x_1^t, ..., x_I^t\right) &= \min_{u_i(t)} L_i\left(x_1(t), ..., x_I(t), u_1^*(t), ..., u_i(t), ..., u_I^*(t)\right), \\
&= \min_{u_i(t)} \mathbb{E}\left[\int_t^\tau l_i(s, x_1(s), ..., x_I(s), u_1^*(s), ..., u_i(s), ..., u_I^*(s))ds + V_i\left(\tau, x_1^\tau, ..., x_I^\tau\right)\right],
\end{aligned}$$

You can consult the pages 180 and 181 in (Yong & Zhou, 1999).

■

Hamilton-Jacobi-Bellman Equation

To solve each value function $V_i\left(t, x_1^t, ..., x_I^t\right)$ in (7), we define a partial differential equation (PDE) for each of them.

Theorem 4.2. *Assume that each value function* $V_i\left(t, x_1^t, ..., x_I^t\right) \in C^2\left([t, T] \times \mathbb{R}^I\right)$. *Then* $\left(V_1\left(t, x_1^t, ..., x_I^t\right), ..., V_I\left(t, x_1^t, ..., x_I^t\right)\right)$ *is a solution of the following nonlinear second-order partial differential equations*

$$
\begin{cases}
\dfrac{\partial V_1}{\partial t}(t, x_1^t, ..., x_I^t) + \min\limits_{u_1(t)}\Big[l_1(t, x_1(t), ..., x_I(t), u_1(t), u_2^*(t), ..., u_I^*(t)) \\
\qquad + b_1(t, x_1(t), ..., x_I(t), u_1(t), u_2^*(t), ..., u_I^*(t)) \displaystyle\sum_{j=1}^{I} \frac{\partial V_1}{\partial x_j^t}\left(t, x_1^t, ..., x_I^t\right) \\
\qquad + \dfrac{1}{2}(\sigma_1(t, x_1(t), ..., x_I(t), u_1(t), u_2^*(t), ..., u_I^*(t)))^2 Tr(\mathbb{H}_x(V_1\left(t, x_1^t, ..., x_I^t\right)))\Big] = 0\ , \\
\dfrac{\partial V_2}{\partial t}(t, x_1^t, ..., x_I^t) + \min\limits_{u_2(t)}\Big[l_2(t, x_1(t), ..., x_I(t), u_1^*(t), u_2(t), ..., u_I^*(t)) \\
\qquad + b_2(t, x_1(t), ..., x_I(t), u_1^*(t), u_2(t), ..., u_I^*(t)) \displaystyle\sum_{j=1}^{I} \frac{\partial V_2}{\partial x_j^t}\left(t, x_1^t, ..., x_I^I\right) \\
\qquad + \dfrac{1}{2}(\sigma_2(t, x_1(t), ..., x_I(t), u_1^*(t), u_2(t), ..., u_I^*(t)))^2 Tr(\mathbb{H}_x(V_2\left(t, x_1^t, ..., x_I^t\right)))\Big] = 0\ , \\
\qquad \vdots \\
\dfrac{\partial V_I}{\partial t}(t, x_1^t, ..., x_I^t) + \min\limits_{u_I(t)}\Big[l_I(t, x_1(t), ..., x_I(t), u_1^*(t), u_2^*(t), ..., u_I(t)) \\
\qquad + b_I(t, x_1(t), ..., x_I(t), u_1^*(t), u_2^*(t), ..., u_I(t)) \displaystyle\sum_{j=1}^{I} \frac{\partial V_I}{\partial x_j^t}\left(t, x_1^t, ..., x_I^t\right) \\
\qquad + \dfrac{1}{2}(\sigma_I(t, x_1(t), ..., x_I(t), u_1^*(t), u_2^*(t), ..., u_I(t)))^2 Tr(\mathbb{H}_x(V_I\left(t, x_1^t, ..., x_I^t\right)))\Big] = 0\ ,
\end{cases}
\tag{8}
$$

where $\mathbb{H}_x(V_i\left(t, x_1^t, ..., x_I^t\right))$ *the Hessian matrix of* V_i *w.r.t* x *and* Tr *the trace operator.*

□

Proof. First, we start by setting $\tau = t + h$, for $0 \leq h \leq \tau - t$ in the Bellman principle of optimality

$$
V_i\left(t, x_1^t, ..., x_I^t\right) = \min_{u_i(t)} \mathbb{E}\left[\int_t^{t+h} l_i(s, x_1(s), ..., x_N(s), u_1(s), u_2^*(s), ..., u_I^*(s)) ds + V_i\left(t+h, x_1^{t+h}, ..., x_I^{t+h}\right)\right].
$$

After that, we apply the Ito stochastic differentiation rule to the function V_i

$$V_i\left(t+h, x_1^{t+h}, ..., x_I^{t+h}\right) = V_i\left(t, x_1^t, ..., x_I^t\right) + \int_t^{t+h} \left(\frac{\partial V_i}{\partial s}\left(s, x_1^s, ..., x_I^s\right) + b_i \sum_{j=1}^{I} \frac{\partial V_i}{\partial x_j^s}\left(t, x_1^t, ..., x_I^t\right) \right) ds,$$

$$+ \int_t^{t+h} \frac{1}{2}(\sigma_i)^2 Tr(\mathbb{H}_x(V_i\left(s, x_1^s, ..., x_I^s\right)))ds + \int_t^{t+h} \sigma_i \sum_{j=1}^{I} \frac{\partial V_i}{\partial x_j^s}\left(t, x_1^s, ..., x_I^s\right) dB_i(s),$$

such that $b_i = b_i(t, x_1(t), ..., x_I(t), u_1(t), ..., u_I(t))$,
$\sigma_i = \sigma_i(t, x_1(t), ..., x_I(t), u_1(t), ..., u_I(t))$.

The next step is to replace this Ito formula in the ith equation of (7)

$$\begin{aligned} V_i\left(t, x_1^t, ..., x_I^t\right) =& \min_{u_i(t)} \mathbb{E}\Bigg[\int_t^{t+h} l_i(s, x_1(s), ..., x_I(s), u_1^*(s), ..., u_i(s), ..., u_I^*(s))ds + V_i\left(t, x_1^t, ..., x_I^t\right) \\ &+ \int_t^{t+h} \left(\frac{\partial V_i}{\partial s}\left(s, x_1^s, ..., x_I^s\right) + b_i^* \sum_{j=1}^{I} \frac{\partial V_i}{\partial x_j^s}\left(t, x_1^t, ..., x_I^t\right) \right) ds, \\ &+ \int_t^{t+h} \frac{1}{2}(\sigma_i^*)^2 Tr(\mathbb{H}_x(V_i\left(s, x_1^s, ..., x_I^s\right)))ds + \int_t^{t+h} \sigma_i^* \sum_{j=1}^{I} \frac{\partial V_i}{\partial x_j^s}\left(t, x_1^s, ..., x_I^s\right) dB_i(s) \Bigg], \end{aligned}$$

where $b_i^* = b_i(t, x_1(t), ..., x_I(t), u_1(t), ..., u_i^*(t), ..., u_I(t))$,
$\sigma_i^* = \sigma_i(t, x_1(t), ..., x_I(t), u_1^*(t), ..., u_i(t), ..., u_I^*(t))$.

Taken outside of the minimum the terms that do not depend on the action u_i and dividing the equation by h leads to

$$\begin{aligned} &\mathbb{E}\left[\frac{\int_t^{t+h} \frac{\partial V_i}{\partial s}\left(s, x_1^s, ..., x_I^s\right) ds}{h} \right] + \min_{u_i(t)} \mathbb{E}\Bigg[\frac{\int_t^{t+h} l_i(s, x_1(s), ..., x_I(s), u_1^*(s), ..., u_i(s), ..., u_I^*(s))ds}{h} \\ &+ \frac{\int_t^{t+h} b_i^* \sum_{j=1}^{I} \frac{\partial V_i}{\partial x_j^s}\left(t, x_1^s, ..., x_I^s\right) ds}{h} + \frac{\int_t^{t+h} \frac{1}{2}(\sigma_i^*)^2 Tr(\mathbb{H}_x(V_i\left(s, x_1^s, ..., x_I^s\right)))ds}{h} \\ &+ \frac{\int_t^{t+h} \sigma_i^* \sum_{j=1}^{I} \frac{\partial V_i}{\partial x_j^s}\left(t, x_1^s, ..., x_I^s\right) dB_i(s)}{h} \Bigg] = 0. \end{aligned} \tag{9}$$

We know that if f_i is a continuous, nonnegative random variable, the two following properties hold

- $\mathbb{E}(f_i) = 0 \Rightarrow f_i = 0$,
- $\lim_{h \to 0} \frac{\int_t^{t+h} f_i\left(s, x_1^s, ..., x_I^s\right) ds}{h} = f_i\left(t, x_1^t, ..., x_I^t\right).$

Using these results in (9) gives us the HJB equation

$$\frac{\partial V_i}{\partial t}\left(t, x_1^t, ..., x_I^t\right) + \min_{u_i(t)}\Big[l_i(t, x_1(t), ..., x_I(t), u_1^*(t), ..., u_i(t), ..., u_I(t))$$
$$+ b_i(t, x_1(t), ..., x_I(t), u_1^*(t), ..., u_i(t), ..., u_I^*(t)) \sum_{j=1}^{I} \frac{\partial V_i}{\partial x_j^t}\left(t, x_1^t, ..., x_I^t\right)$$
$$+ \frac{1}{2}(\sigma_i(t, x_1(t), ..., x_I(t), u_1^*(t), ..., u_i(t), ..., u_I^*(t)))^2 Tr(\mathbb{H}_x(V_i\left(t, x_1^t, ..., x_I^t\right))) = 0.$$

■

Remark 4.2.

- *The second term in each equation in the (8) system is known as the Hamiltonian H_i for all $1 \le i \le I$*

$$\begin{aligned} H_i(t, x_1(t), ..., x_I(t), p_i, q_i) = \min_{u_i(t)}\Big[& l_i(t, x_1(t), ..., x_I(t), u_1^*(t), u_2^*(t), ..., u_i(t), ..., u_I^*(t)) \\ & + b_i(t, x_1(t), ..., x_I(t), u_1^*(t), u_2^*(t), ..., u_i(t), ..., u_I^*(t)) p_i(t) \\ & + \frac{1}{2}(\sigma_i(t, x_1(t), ..., x_I(t), u_1^*(t), u_2^*(t), ..., u_i(t), ..., u_I^*(t)))^2 q_i(t)\Big] = 0, \end{aligned} \tag{10}$$

  where $p_i(t) = \sum_{j=1}^{I} \frac{\partial V_i}{\partial x_j^t}\left(t, x_1^t, ..., x_I^t\right)$, *and* $q_i(t) = Tr(\mathbb{H}_x(V_i\left(t, x_1^t, ..., x_I^t\right)))$.

- *The PDEs in the (8) system are called the Hamilton-Jacobi-Bellman equations.*

- *To have a solution for the I stochastic optimal control problems in (5), we must apply the verification technique that involves the following steps.*

  - ***Step 1.*** *Solve the Hamilton-Jacobi-Bellman equations in (8) to find the value functions* $(V_1\left(t, x_1^t, ..., x_I^t\right), ..., V_I\left(t, x_1^t, ..., x_I^t\right))$.
  - ***Step 2.*** *Find the Nash equilibrium (the I optimal actions)* $(u_1^*(t), u_2^*(t), ..., u_I^*(t))$ *through (9).*
  - ***Step 3.*** *Replace the I optimal actions in the I state equations (3) and solve these I SDEs (with the initial time and state* $(0, x_i^0)$*) to obtain the optimal states* $(x_1^*(t), x_2^*(t), ..., x_I^*(t))$.

  $\Longrightarrow$ *As a Final result we get for each player i on optimal pair* $(x_i^*(t), u_i^*(t))$.

□

5 Mean Field Game

Computing a Nash equilibrium of a stochastic differential game with $I \geq 2$ players involves solving the I-coupled HJB equations in (8) simultaneously. As the number of players I increases, the resolution of the system becomes more and more complicated, if not impossible, due to the increased interactions and couplings among the players. Jean-Michel Lasry and Pierre-Louis Lions proposed the mean-field game (MFG) in (Lasry & Lions, 2007) to deal with this large population game problem. The principal idea of the authors is to transfer the concept of the mean field from the physics area, which considers the interaction of large numbers of particles where each particle has a negligible effect on the whole system, to the game theory area. To reduce complexity, each player in MFG plays against the mass of the population rather than against each other individually (the interaction is among the player states). In this case, the systems (3) and (4) become as follows respectively

$$\begin{cases} dx_1(s) = b_1\Big(s, x_1(s), \dfrac{1}{I}\displaystyle\sum_{j=1}^{I}\delta_{x_j(s)}, u_1(s)\Big)ds + \sigma_1\Big(s, x_1(s), \dfrac{1}{I}\displaystyle\sum_{j=1}^{I}\delta_{x_j(s)}, u_1(s)\Big)dB_1(s), \\ dx_2(s) = b_2\Big(s, x_2(s), \dfrac{1}{I}\displaystyle\sum_{j=1}^{I}\delta_{x_j(s)}, u_2(s)\Big)ds + \sigma_2\Big(s, x_2(s), \dfrac{1}{I}\displaystyle\sum_{j=1}^{I}\delta_{x_j(s)}, u_2(s)\Big)dB_2(s), \\ \quad\vdots \\ dx_I(s) = b_I\Big(s, x_I(s), \dfrac{1}{I}\displaystyle\sum_{j=1}^{I}\delta_{x_j(s)}, u_I(s)\Big)ds + \sigma_I\Big(s, x_I(s), \dfrac{1}{I}\displaystyle\sum_{j=1}^{I}\delta_{x_j(s)}, u_I(s)\Big)dB_I(s), \end{cases} \tag{11}$$

subject to,

$$\begin{cases} L_1\Big(x_1(t), \dfrac{1}{I}\displaystyle\sum_{j=1}^{I}\delta_{x_j(t)}, u_1(t)\Big) = \mathbb{E}\left(\displaystyle\int_t^T l_1\Big(s, x_1(s), \dfrac{1}{I}\sum_{j=1}^{I}\delta_{x_j(s)}, u_1(s)\Big)ds + h_1\Big(x_1(T), \dfrac{1}{I}\sum_{j=1}^{I}\delta_{x_j(s)}\Big)\right), \\ L_2\Big(x_2(t), \dfrac{1}{I}\displaystyle\sum_{j=1}^{I}\delta_{x_j(t)}, u_2(t)\Big) = \mathbb{E}\left(\displaystyle\int_t^T l_2\Big(s, x_2(s), \dfrac{1}{I}\sum_{j=1}^{I}\delta_{x_j(s)}, u_2(s)\Big)ds + h_2\Big(x_2(T), \dfrac{1}{I}\sum_{j=1}^{I}\delta_{x_j(T)}\Big)\right), \\ \qquad\quad\vdots \\ L_I\Big(x_I(t), \dfrac{1}{I}\displaystyle\sum_{j=1}^{I}\delta_{x_j(t)}, u_I(t)\Big) = \mathbb{E}\left(\displaystyle\int_t^T l_I\Big(s, x_I(s), \dfrac{1}{I}\sum_{j=1}^{I}\delta_{x_j(s)}, u_I(s)\Big)ds + h_I\Big(x_I(T), \dfrac{1}{I}\sum_{j=1}^{I}\delta_{x_j(T)}\Big)\right), \end{cases} \tag{12}$$

where $\dfrac{1}{I}\displaystyle\sum_{j=1}^{I}\delta_{x_j(.)}$ is the empirical measure of all the player's states and $\delta_{x_j(.)}$ is the Dirac measure (the unit mass at the state $x_j(.)$).

Also in MFG, we assume the following assumptions

- The number of players I is infinitely large.
- The players are indistinguishable and symmetric.
- The players are nonatomic, meaning that the individual action of one player does not affect the global state of the game.

When $I \to \infty$ and the three conditions are satisfied, the asymptotic formulations of the last two systems are expected to not depend on i (i.e., nonparameterized). To simplify the resolution of the equilibrium problem in this large game limit, we restrict the system (11) composed of I identical equations in the limit to one dynamics equation, knowing as a controlled McKean-Vlasov equation of one representative player who tries to minimize his objective function given the aggregate behavior of the other players. Thus, the mean-field game problem is formulated as follows

$$\min_{u(s)} \qquad L(x(t), u(t)) = \mathbb{E}\left(\int_t^T l(s, x(s), m(s), u(s))ds + h(x(T), m(T)) \right),$$

subject to

$$\begin{cases} dx(s) = b(s, x(s), m(s), u(s))ds + \sigma(s, x(s), m(s), u(s))dB(s), \quad s \in [t, T], \\ x(t) = x. \end{cases} \tag{13}$$

Where the mean field term $m(s) \in \mathcal{P}(\mathbb{R})$ represents the probability distribution of a generic player with state x at time s. According to the law of large numbers, $m(s)$ is defined as

$$m(s) = \lim_{I \to \infty} \frac{1}{I} \left(\sum_{j=1}^{I} \delta_{x_j(s)=x(s)} \right),$$

such that $\delta = 1$ when $x_j(s) = x(s)$ and $\delta = 0$ when $x_j(s) \neq x(s)$.

Solving the MFG problem is equivalent to solving the next pair of an HJB equation describing the evolution of the value function of a representative agent and a Fokker-Planck-Kolmogorov (FPK) equation describing the evolution of the probability distribution $m(t)$,

$$\begin{cases} \dfrac{\partial V}{\partial t}(t,x) + \min\limits_{u(t)} \Big[l(t, x, m(t), u(t)) + b(t, x, m(t), u(t)) \dfrac{\partial V}{\partial x}(t,x) \\ \qquad\qquad + \dfrac{1}{2}\sigma(t, x, m(t), u(t))^2 \dfrac{\partial^2 V}{\partial x \partial x}(t,x) \Big] = 0, \; t \in [0, T] \\ V(T, x) = h(x, m(T)), \\ \dfrac{\partial m}{\partial t}(t) + \dfrac{\partial (b(t, x, m(t), u^*(t)) m(t))}{\partial x} - \dfrac{\partial^2 (\sigma(t, x, m(t), u^*(t))^2 m(t))}{2 \partial x \partial x} = 0, \; t \in [0, T], \\ m(0) = m_0. \end{cases} \tag{14}$$

with

$$H(t,x,m,p,q) = \min_{u(t)} \Big[l(t,x,m(t),u(t)) + b(t,x,m(t),u(t))p(t) + \frac{1}{2}\sigma(t,x,m(t),u(t))^2 q(t) \Big] = 0, \tag{15}$$

$p(t) = \dfrac{\partial V}{\partial x}(t,x)$ and $q(t) = \dfrac{\partial^2 V}{\partial x \partial x}(t,x)$.

Proof.
- The proof of the HJB equation in the MFG system (14) is taken in the same way as the proof of the theorem 2.21, where the mean field term $m(t)$ is a fixed deterministic function acting as a parameter.
- For the FPK equation, see the proof 6.2 without the jump part. ■

The mean-field game equilibrium is the solution pair $(u^*(t), m(t))$ of (14) system, where $u^*(t)$ satisfies

$$L(x, u^*(t)) \leq L(x, u(t)),$$

and $m(t) = \mathcal{L}(x^*(t))$, with $x^*(t)$ the optimal state.

To obtain this pair, we must solve this fixed point problem by following the next steps.

- **Step 1.** Solving the HJB equation in (14) by freezing the mean field term $m(t)$ (supposed as an input) to get the value function $V(t,x) = \min\limits_{u(s)} L(x, u(t))$.
- **Step 2.** Determine the optimal action $u^*(t) = \gamma^*(t,x,m,p,q)$ by solving the Hamiltonian H in (15) and then the optimal state $x^*(t)$ by solving

$$\begin{cases} dx^*(t) = b(t,x^*(t),m(t),u^*(t))ds + \sigma(t,x^*(t),m(t),u^*(t))dB(t), & t \in [0,T], \\ x^*(0) = x_0^*. \end{cases}$$

- **Step 3.** Identify the probability law $m(t)$ of the optimal state $x^*(t)$ $(m(t) = \mathcal{L}(x^*(t)))$ by solving the FPK equation in the MFG system (14).

6 Mean Field Type Game

The mean-field game assumptions exclude many real-world modeling problems. For example, in engineering, economics, and various other domains, the agents are finite, not anonymous, and each player takes actions that significantly affect the state of the game. A relaxed version of the MFG called the mean-field type game (MFTG) was suggested by Djehiche et al. in (Andersson & Djehiche, 2011; Buckdahn R. Djehiche & Li, 2011) for a one-player game and extended to a multiplayer game in (Tembine, 2015; Tcheukam & Tembine, 2016; Djehiche B. Tcheukam & Tembine, 2017; Tembine, 2017; Barreiro-Gomez J. Duncan & Tembine, 2019c; T. Duncan & Tembine, 2018; Barreiro-Gomez J. Duncan & Tembine, 2019b). To give the player more weight in the game, the authors incorporate the first moments of his state and action into the state dynamics equation and the performance functional (cost functional).

A $I \geq 2$ finite stochastic MFTG problem is defined as follows

$$
\begin{aligned}
&\min_{u_i(s)} L_i(x, u_i(t), u_{-i}(t)) \\
&= \min_{u_i(s)} \mathbb{E}\left(\int_t^T l_i(s, x(s), \mathbb{E}[x(s)], u_i(s), \mathbb{E}[u_i(s)], u_{-i}(s), \mathbb{E}[u_{-i}(s)])ds + h_i(x(T), \mathbb{E}[x(T)])\right),
\end{aligned}
$$

subject to

$$
\begin{cases}
dx(s) = b(s, x(s), \mathbb{E}[x(s)], u_i(s), \mathbb{E}[u_i(s)], u_{-i}(s), \mathbb{E}[u_{-i}(s)])ds \\
\qquad\qquad +\sigma(s, x(s), \mathbb{E}[x(s)], u_i(s), \mathbb{E}[u_i(s)], u_{-i}(s), \mathbb{E}[u_{-i}(s)])dB(s), \quad s \in [t, T], \\
x(t) = x,
\end{cases} \tag{16}
$$

where $u_{-i}(s) = (u_1(s), ..., u_{i-1}(s), u_{i+1}(s), ..., u_I(s))$,
$\mathbb{E}[x(s)] = \int_t^T y(s)m(s, y)dy$,
$\mathbb{E}[u_i(s)] = \int_t^T u_i(s, y, m)m(s, y)dy$,
and $m(t, x)$ is the probability density of the state x at time t.

A Nash mean field type equilibrium is a strategies profile $((u_1^*(t), ..., u_i^*(t), ..., u_I^*(t)))$, that satisfies for each player $i \in \mathcal{I}$

$$
L_i(x, u_i^*(t), u_{-i}^*(t)) \leq L_i(x, u_i(t), u_{-i}^*(t)).
$$

6.1 Mean-Field-Type Game with Jump

When mathematics is applied to other scientific branches such as finance, economics, computer science, engineering, biology, or meteorology, a natural question arises: Do the obtained results fit those of these branches? Empirical studies of different trajectories in different fields prove the existence of sudden variations with large amplitude (jumps) due to various external factors. For example, in finance, jumps in market prices are caused by political decisions, epidemics, natural disasters, inflation, wars, etc., such as the infamous stock market crash of 19 October 1987, when the Dow Jones Industrial Average plunged 23% in a single day. Other examples include jumps in the number of viewers trajectory in social media and TV channels or jumps in the amount of rainfall trajectory resulting from climate change and pollution. These empirical facts motivated us to model the state of the MFTG in the second part of the thesis by a jump-diffusion process instead of a purely continuous process as follows

$$\begin{cases} dx(s) = b(s, x(s), \mathbb{E}[x(s)], u_i(s), \mathbb{E}[u_i(s)], u_{-i}(s), \mathbb{E}[u_{-i}(s)])ds \\ \qquad\qquad +\sigma(s, x(s), \mathbb{E}[x(s)], u_i(s), \mathbb{E}[u_i(s)], u_{-i}(s), \mathbb{E}[u_{-i}(s)])dB(s), \\ \qquad\qquad + \int_\Theta \mu(s, x(s), \mathbb{E}[x(s)], u_i(s), \mathbb{E}[u_i(s)], u_{-i}(s), \mathbb{E}[u_{-i}(s)], \theta)\tilde{N}(ds, d\theta),\ s \in [t, T], \\ x(t) = x,\ t \in [0, T], \end{cases} \tag{17}$$

where: $\Theta = \mathbb{R}_+ \backslash \{0\}$ is the set of jump size,

$\mu:\ [t, T] \times \mathbb{R} \times \mathcal{P}(\mathbb{R}) \times \prod_{i=1}^{I} U_i \times \Theta \to \mathbb{R}$ is the jump rate,

$\tilde{N}(dt, d\theta) = N(dt, d\theta) - \nu(d\theta)ds$ is the compensated Poisson measure,

$N(ds, d\theta)$ is the Poisson measure with intensity measure $\nu(d\theta)ds$,

the B and N processes are independent of each other.

B is used to capture smaller disturbances and N is used for larger system jumps.

For complete definitions and stochastic calculus based on jump processes, please refer to (Cont & Tankov, 1st Edition, 2003; Øksendal & Sulem-Bialobroda, 2st Edition, 2009).

6.2 Solution Approaches

For solving an MFTG problem with jumps, we present two different methods

Dynamic Programming Principle

The state SDEs of the MFTG problems 16 and 17 are not appropriate to the dynamic programming principle due to the existence of the density $m(t, x)$, which renders the state non-Markovian, and the iterated expectation no longer holds. The way out of this issue is to identify an augmented state where the DPP can be applied. A satisfying augmented state is the density $m(t, x)$ that solves an FPK equation. We only need to rewrite the cost-functional L_i as a function of $m(t, x)$ to obtain a classical deterministic differential game problem given by

$$\begin{aligned}
&\min_{u_i(s)} L_i(m_t(x), u_i(t), u_{-i}^*(t)) \\
&\quad = \min_{u_i(s)} \int_t^T \int_{\mathbb{R}} l_i(s, x, m(s,x), u_i(s), u_{-i}^*(s)) m(s,x) dx ds + \int_{\mathbb{R}} h_i(x, m(T,x)) m(T,x) dx,
\end{aligned}$$

subject to

$$\begin{cases} \dfrac{\partial m}{\partial s}(s,x) + \dfrac{\partial (bm(s,x))}{\partial x} - \dfrac{\partial^2(\sigma^2 m(s,x))}{2\partial x \partial x} - J^*[m], \; s \in [t,T], \\ m(t,x) = m_t(x). \end{cases} \tag{18}$$

where $b := b(s, x, m(s,x), u_i(s), u_{-i}(s))$,
$\sigma := \sigma(s, x, m(s,x), u_i(s), u_{-i}(s))$,
$\mu(s,\theta) := \mu(s, x, m(s,x), u_i(s), u_{-i}(s), \theta)$,
and $J^*[m]$ the adjoint operator of the jump operator J given by

$$J[m] = \int_\Theta \Big[m(s, x + \mu(s,\theta)) - m(s,x) - \mu(s,\theta) \frac{\partial m}{dx}(s,x) \Big] \nu(d\theta).$$

Proof. **The FPK equation with jump**
We start by applying the Ito differentiation for an arbitrary function noted $g(x)$, where $x := x(s)$ is the jump-diffusion process in 17 (see Lemma 2 in (Bensoussan A. Djehiche B. Tembine & Yam, 2019))

$$\begin{aligned}
dg(x) = &\Big(b \frac{\partial g(x)}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2 g(x)}{\partial x \partial x} + J[g] \Big) ds + \sigma \frac{\partial g(x)}{\partial x} dB(s) \\
&+ \int_\Theta [g(x + \mu(\theta)) - g(x)] \tilde{N}(ds, d\theta),
\end{aligned}$$

where $J[g]$ is the jump operator of $g(x)$ defined by

$$J[g] = \int_\Theta \Big[g(x + \mu(\theta)) - g(x) - \mu(\theta) \frac{dg(x)}{dx} \Big] \nu(d\theta),$$

After that, we take the expectations of both sides

$$\begin{aligned}
\mathbb{E}[dg(x)] &= \mathbb{E}\Big[b \frac{\partial g(x)}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2 g(x)}{\partial x \partial x} + J[g] \Big] ds, \\
\Rightarrow \; \frac{d}{ds} \mathbb{E}[g(x)] &= \mathbb{E}\Big[b \frac{\partial g(x)}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2 g(x)}{\partial x \partial x} + J[g] \Big],
\end{aligned}$$

Then, we rewrite the last equation in terms of the probability density $m(s,x)$

$$\begin{aligned}\frac{d}{ds}\int_{\mathbb{R}} g(x)m(s,x)dx &= \int_{\mathbb{R}} \Big[b\frac{\partial g(x)}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2 g(x)}{\partial x \partial x} + J[g]\Big]m(s,x)dx \\ &= \int_{\mathbb{R}} b\frac{\partial g(x)}{\partial x}m(s,x)dx + \int_{\mathbb{R}} \frac{\sigma^2}{2}\frac{\partial^2 g(x)}{\partial x \partial x}m(s,x)dx + \int_{\mathbb{R}} J[g]m(s,x)dx, \\ \Rightarrow \int_{\mathbb{R}} g(x)\frac{\partial m(s,x)}{\partial s}dx &= \int_{\mathbb{R}} b\frac{\partial g(x)}{\partial x}m(s,x)dx + \int_{\mathbb{R}} \frac{\sigma^2}{2}\frac{\partial^2 g(x)}{\partial x \partial x}m(s,x)dx + \langle J[g], m(s,x)\rangle,\end{aligned}$$

We know that: $\langle J[g], m(s,x)\rangle = \langle g(x), J^*[m]\rangle$, where $\langle , \rangle$ is the inner product, replacing that in the last equation and by integration by parts, we get

$$\int_{\mathbb{R}} g(x)\frac{\partial m(s,x)}{\partial s}dx = -\int_{\mathbb{R}} g(x)\frac{\partial (bm(s,x)}{\partial x}dx + \int_{\mathbb{R}} g(x)\frac{\partial^2(\sigma^2 m(s,x))}{2\partial x \partial x}dx + \langle g(x), J^*[m]\rangle,$$

Now we combine the integrals

$$\begin{aligned}\int_{\mathbb{R}} g(x)\Big[\frac{\partial m(s,x)}{\partial s} + \frac{\partial (bm(s,x)}{\partial x} - \frac{\partial^2(\sigma^2 m(s,x))}{2\partial x \partial x} - J^*[m]\Big]dx = 0, \\ \Rightarrow \frac{\partial m(s,x)}{\partial s} + \frac{\partial (bm(s,x)}{\partial x} - \frac{\partial^2(\sigma^2 m(s,x))}{2\partial x \partial x} - J^*[m] = 0.\end{aligned}$$

The last equation is the FPK equation. ■

Definition 6.1. *The cost-to-go function of player i is defined by*

$$\begin{cases} V_i(t, m_t(x)) = \min\limits_{u_i(s)} \ L_i(m_t(x), u_i(t), u^*_{-i}(t)) \\ V_i(T, m_T(x)) = \displaystyle\int_{\mathbb{R}} h_i(x, m_t(x))m_t(x)dx, t \in [0,T]. \end{cases} \tag{19}$$

Proposition 6.1. *For any $(\tau, m(\tau,x)) \in [t,T] \times \mathcal{P}(\mathbb{R})$ with $t \leq \tau \leq T$, we have*

$$V_i(t, m_t(x)) = \min_{u_i(s)} \Big[\int_t^{t+h}\int_{\mathbb{R}} l_i(s,x,m(s,x),u_i(s),u^*_{-i}(s))m(s,x)dxds + V_i(t+h, m(t+h,x))\Big] \tag{20}$$

□

Proof. Let the right side of the equality be denoted by $\bar{V}_i(t, m_t(x))$. For any $\varrho > 0$, there exists $u_i(s) \in U_i[t,T]$ so that

$$
\begin{aligned}
V_i(t, m_t(x)) + \varrho > L_i(m_t(x), u_i(t), u^*_{-i}(t)) = & \int_t^T \int_{\mathbb{R}} l_i(s, x, m(s,x), u_i(s), u^*_{-i}(s)) m(s,x) dx ds \\
& + \int_{\mathbb{R}} h_i(x, m(T,x)) m(T,x) dx, \\
& = \int_t^{t+h} \int_{\mathbb{R}} l_i(s, x, m(s,x), u_i(s), u^*_{-i}(s)) m(s,x) dx ds \\
& + \int_{t+h}^T \int_{\mathbb{R}} l_i(s, x, m(s,x), u_i(s), u^*_{-i}(s)) m(s,x) dx ds \\
& + \int_{\mathbb{R}} h_i(x, m(T,x)) m(T,x) dx, \\
& \geq \int_t^{t+h} \int_{\mathbb{R}} l_i(s, x, m(s,x), u_i(s), u^*_{-i}(s)) m(s,x) dx ds \\
& + V_i(t+h, m(t+h,x)) \geq \bar{V}_i(t, m_t(x)).
\end{aligned}
$$

Inversely, for any $\varrho > 0$ and a given $u_i(s) \in U_i[t,T]$, there exists $\tilde{u}_i(s) \in U_i[t,T]$ that matches with $u_i(s)$ on $[t, t+h]$, so that

$$
L_i(\tilde{m}(t+h,x), \tilde{u}_i(t+h), \tilde{u}^*_{-i}(t+h)) \leq V_i(t+h, \tilde{m}(t+h,x)) + \varrho,
$$

with $\tilde{m}(s,x)$ the solution of 18 at time s with control $\tilde{u}(s)$, where $\tilde{m}_t(x)$ is the initial state. Therefore

$$
\begin{aligned}
V_i(t, m_t(x)) = V_i(t, \tilde{m}_t(x)) \leq & L_i(\tilde{m}_t(x), \tilde{u}_i(t), \tilde{u}^*_{-i}(t)) \\
& = \int_t^T \int_{\mathbb{R}} l_i(s, x, \tilde{m}(s,x), \tilde{u}_i(s), \tilde{u}^*_{-i}(s)) \tilde{m}(s,x) dx ds \\
& + \int_{\mathbb{R}} h_i(x, \tilde{m}(T,x)) \tilde{m}(T,x) dx, \\
& = \int_t^{t+h} \int_{\mathbb{R}} l_i(s, x, \tilde{m}(s,x), u_i(s), u^*_{-i}(s)) \tilde{m}(s,x) dx ds \\
& + \int_{t+h}^T \int_{\mathbb{R}} l_i(s, x, \tilde{m}(s,x), \tilde{u}_i(s), \tilde{u}^*_{-i}(s)) \tilde{m}(s,x) dx ds \\
& + \int_{\mathbb{R}} h_i(x, \tilde{m}(T,x)) \tilde{m}(T,x) dx, \\
& = \int_t^{t+h} \int_{\mathbb{R}} l_i(s, x, \tilde{m}(s,x), u_i(s), u^*_{-i}(s)) \tilde{m}(s,x) dx ds \\
& + L_i(\tilde{m}(t+h,x), \tilde{u}_i(t+h), \tilde{u}^*_{-i}(t+h)) \\
& \leq \int_t^{t+h} \int_{\mathbb{R}} l_i(s, x, \tilde{m}(s,x), u_i(s), u^*_{-i}(s)) \tilde{m}(s,x) dx ds \\
& + V_i(t+h, \tilde{m}(t+h,x)) + \varrho, \leq \bar{V}_i(t, m_t(x)).
\end{aligned}
$$

We make $\varrho \longrightarrow 0$ and we take the infimum, we get

$$\begin{aligned}\bar{V}_i(t, m_t(x)) \leq V_i(t, m_t(x)) &\leq \bar{V}_i(t, m_t(x)),\\ \Rightarrow V_i(t, m_t(x)) &= \bar{V}_i(t, m_t(x)),\end{aligned}$$

which concludes the proof. $\blacksquare$

Proposition 6.2. *Suppose $V_i(t, m_t(x)) \in C^2\left([t, T] \times \mathcal{P}(\mathbb{R})\right)$, then $V_i(t, m_t(x))$ is a solution of the following HJB equation*

$$\frac{\partial V_i}{dt}(t, m_t(x)) + \int_{\mathbb{R}} \min_{u_i(s)} H_i(t, x, m_t(x), V_{im}, V_{ixm}, V_{ixxm}) m(t, x) dx = 0, \tag{21}$$

where the Hamiltonian H_i has the form

$$\begin{aligned}H_i(t, x, m_t(x), V_{im}, V_{ixm}, V_{ixxm}) = l_i(s, x, m(s, x), u_i(s), u^*_{-i}(s)) + bV_{ixm} + \frac{\sigma^2}{2} V_{ixxm}\\ + \int_{\Theta} \Big[V_{im}(s, x + \mu(s, \theta)) - V_{im} - \mu(s, \theta) V_{ixm}\Big] \nu(d\theta),\end{aligned} \tag{22}$$

such that $V_{im}(t, x, m_t(x)) := \dfrac{\partial V_i}{\partial m}(t, m_t(x))$,

$$V_{ixm} := \frac{\partial V_{im}}{\partial x}\left(t, m_t(x)\right),$$

$$V_{ixxm} := \frac{\partial^2 V_{im}}{\partial x \partial x}\left(t, m_t(x)\right). \qquad \square$$

Proof. Applying the Taylor series to $V_i(t + h, m(t + h, x))$, we get

$$\begin{aligned}V_i(t + h, m(t + h, x)) =& V_i(t, m_t(x)) + h \frac{\partial V_i}{\partial t}\left(t, m_t(x)\right) + (m(t + h, x) - m(t, x)) \frac{\partial V_i}{\partial m}\left(t, m_t(x)\right)\\ &+ o(h),\end{aligned}$$

from 18 we can rewrite $V_i(t + h, m(t + h, x))$ as

$$\begin{aligned}V_i(t + h, m(t + h, x)) = V_i(t, m_t(x)) + h \frac{\partial V_i}{\partial t}\left(t, m_t(x)\right) + \int_t^{t+h} \Bigg\{ - \frac{\partial(bm(s, x))}{\partial x} + \frac{\partial^2(\sigma^2 m\left(s, x\right))}{2\partial x \partial x}\\ + J^*[m]\Bigg\} ds \frac{\partial V_i}{\partial m}\left(t, m_t(x)\right) + o(h),\end{aligned}$$

replacing $V_i(t+h, m(t+h,x))$ in 20 yields

$$V_i(t, m_t(x)) = \min_{u_i(s)} \Big[\int_t^{t+h} \int_{\mathbb{R}} l_i(s,x,m(s,x),u_i(s),u^*_{-i}(s))m(s,x)dxds + V_i(t,m_t(x)) \\ +h\frac{\partial V_i}{\partial t}(t,m_t(x)) + \int_t^{t+h} \Big\{ -\frac{\partial(bm(s,x))}{\partial x} + \frac{\partial^2(\sigma^2 m(s,x))}{2\partial x \partial x} \\ + J^*[m]\Big\} ds \frac{\partial V_i}{\partial m}(t,m_t(x)) + o(h)\Big],$$

get the terms $V_i(t,m_t(x))$ and $h\frac{\partial V_i}{\partial t}(t,m_t(x))$ outside the minimum. Next, We dived by h, as follows

$$\frac{\partial V_i}{\partial t}(t,m_t(x)) + \min_{u_i(s)} \Big[\frac{\int_t^{t+h}\int_{\mathbb{R}} l_i(s,x,m(s,x),u_i(s),u^*_{-i}(s))m(s,x)dxds}{h} \\ + \frac{\int_t^{t+h} \Big\{ -\frac{\partial(bm(s,x))}{\partial x} + \frac{\partial^2(\sigma^2 m(s,x))}{2\partial x \partial x} + J^*[m]\Big\} ds \frac{\partial V_i}{\partial m}(t,m_t(x))}{h} + \frac{o(h)}{h}\Big],$$

letting $h \to 0$, leads to

$$\frac{\partial V_i}{\partial t}(t,m_t(x)) + \min_{u_i(s)} \Big[\int_{\mathbb{R}} l_i(s,x,m(s,x),u_i(s),u^*_{-i}(s))m(s,x)dx \\ + \Big\{ -\frac{\partial(bm(s,x))}{\partial x} + \frac{\partial^2(\sigma^2 m(s,x))}{2\partial x \partial x} + J^*[m]\Big\} \frac{\partial V_i}{\partial m}(t,m_t(x))\Big],$$

Applying the integration by part, we obtain

$$\frac{\partial V_i}{\partial t}(t,m_t(x)) + \min_{u_i(s)} \Big[\int_{\mathbb{R}} \Big\{ l_i(s,x,m(s,x),u_i(s),u^*_{-i}(s)) \\ + bV_{ixm} + \frac{\sigma^2}{2} V_{ixxm} + J[V_{im}]\Big\} m(s,x)dx\Big],$$

where,

$$J[V_{im}] = \int_{\Theta} \Big[V_{im}(s, x+\mu(s,\theta)) - V_{im} - \mu(s,\theta)V_{ixm}\Big]\nu(d\theta).$$

Thus, the Proposition 6.2 is proved. ■

For a better understanding of the dynamic programming principle method in the resolution of an MFTG optimal control problem, see Paper A.1, where we used the method for solving a hierarchical MFTG problem with polynomial payoff.

Direct Method

Another interesting approach to solving the MFTG problems is the direct method, proposed first by Duncan and al. (e.g., (T. E. Duncan & Pasik-Duncan, 2012, 2013; T. E. Duncan, 2014, 2016)), after being applied to the MFTG in (Barreiro-Gomez J. Duncan & Tembine, 2019c; T. Duncan & Tembine, 2018; Barreiro-Gomez J. Duncan & Tembine, 2019b; Barreiro-Gomez J. Duncan T. E. Pasik-Duncan & Tembine, 2020; Barreiro-Gomez J. Duncan & Tembine, 2019a, 2020; Tian R. Yu & Zhang, 2020). It is called direct or verification because it provides semi-explicit solutions without the need to solve the HJB-Kolmogorov equations for the dynamic programming method, or backward-forward stochastic differential equations of the Pontryagin's type for the stochastic maximum principle method. These semi-explicit solutions allow us to make numerical applications and analyses, which are of great importance in mathematical modeling. The direct method is a simple procedure based on the following five main steps

- **Step 1.** Define the statement of the MFTG problem (state equation and performance functional) that you are interested in solving.
- **Step 2.** From the structure of the terminal performance functional $h_i(x(T), \mathbb{E}[x(T)])$, identify a suitable guess functional $f(t, x)$ for the optimal performance. This guess functional is composed of unknown deterministic coefficients that need to be determined.
- **Step 3.** Apply Itô's formula to the guess functional $f(t, x)$. In the case of the jump-diffusion processes, Itô's formula has the form

$$\begin{aligned}
f_i(T, x(T)) - f_i(t, x(t)) &= \int_t^T \Big[\frac{\partial f_i}{ds}(s, x(s)) + b\frac{\partial f_i}{dx}(s, x(s)) + \frac{\sigma^2}{2}\frac{\partial f_i}{dxdx}(s, x(s))\Big] ds \\
&+ \int_t^T \sigma \frac{\partial f_i}{dx}(s, x(s)) dB(s) \\
&+ \int_t^T \int_\Theta \Big[f_i(s, x(s) + \mu(s, \theta)) - f_i(s, x(s)) - \mu(s, \theta)\frac{\partial f_i}{dx}(s, x(s))\Big]\nu(d\theta) ds, \\
&+ \int_t^T \int_\Theta [f_i(s_-, x(s) + \mu(s_-, \theta)) - f_i(s_-, x(s))]\tilde{N}(ds, d\theta),\ s \in [t, T],
\end{aligned} \tag{23}$$

where $b := b(s, x(s), \mathbb{E}[x(s)], u_i(s), \mathbb{E}[u_i(s)], u_{-i}(s), \mathbb{E}[u_{-i}(s)])$,
$\sigma := \sigma(s, x(s), \mathbb{E}[x(s)], u_i(s), \mathbb{E}[u_i(s)], u_{-i}(s), \mathbb{E}[u_{-i}(s)])$,
$\mu(s, \theta) := \mu(s, x(s), \mathbb{E}[x(s)], u_i(s), \mathbb{E}[u_i(s)], u_{-i}(s), \mathbb{E}[u_{-i}(s)], \theta)$.

After that, compute the gap $L_i(x, u_i(t), u_{-i}(t)) - \mathbb{E}[f_i(t, x)]$ and group terms by common factors.

- **Step 4.** Complete the squares for both the control and the expected value of the control to get them into a quadratic expression, and then rearrange the gap expression by grouping terms that have the state and the expected value of the state as common factors.

- **Step 5.** Apply an identification process for the optimal control, the expected value of the optimal control, and ordinary differential equations, such that the gap is minimized.

$\Longrightarrow$ At the end of the five steps, one gets

- The optimal actions $(u_i^*(s), u_{-i}^*(s))$ and their expected values $(\mathbb{E}[u_1^*(s)], ..., \mathbb{E}[u_I^*(s)])$ in semi-explicit form.

- A system of ordinary differential equations (Riccati equations) in which the solution defines the deterministic coefficients of the guess functional and gives an explicit form to the optimal controls and their expectations.

- The optimal state $x^*(s)$ by replacing the optimal controls and the expected values of the optimal controls in the state equation 17.

- The expected value of the optimal state $\mathbb{E}(x^*(s))$.

For a better understanding of the direct method, see the proof 4 in Paper B, where we used the method for solving a hierarchical MFTG problem in the electricity market.

References

Andersson, D., & Djehiche, B. (2011). A maximum principle for sdes of mean-field type. *Applied Mathematics & Optimization*, *63*(2011), 341-356.

Barreiro-Gomez J. Duncan, T. E., & Tembine, H. (2019a). Co-opetitive linear-quadratic mean-field-type games. *IEEE transactions on cybernetics*, *50*(12), 5089-5098.

Barreiro-Gomez J. Duncan, T. E., & Tembine, H. (2019b). Linear-quadratic mean-field-type games: Jump-diffusion process with regime switching. *IEEE Transactions on Automatic Control*, *64*(10), 4329-4336.

Barreiro-Gomez J. Duncan, T. E., & Tembine, H. (2019c). Linear-quadratic mean-field-type games with multiple input constraints. *IEEE Control Systems Letters*, *3*(3), 511-516.

Barreiro-Gomez J. Duncan, T. E., & Tembine, H. (2020). Discrete-time linear-quadratic mean-field-type repeated games: Perfect, incomplete, and imperfect information. *Automatica*(112), 0005-1098.

Barreiro-Gomez J. Duncan T. E. Pasik-Duncan, B., & Tembine, H. (2020). Semi-explicit solutions to some non-linear non-quadratic mean-field-type games: A direct method. *IEEE transactions on automatic control*, 1–14.

Bensoussan A. Djehiche B. Tembine, H., & Yam, S. C. P. (2019). Mean-field-type games with jump and regime switching. *Dynamic Games and Applications*, 1-39.

Buckdahn R. Djehiche, B., & Li, J. (2011). A general stochastic maximum principle for SDEs of mean-field type. *Applied Mathematics & Optimization*, *64*(2011), 197-216.

Cont, R., & Tankov, P. (1st Edition, 2003). *Financial modelling with jump processes.* Chapman and Hall/CRC.

Cournot, A. A. (1838). *Recherches sure les principes mathématiques de la théorie des richesses* (Vol. 213 pp).

Djehiche B. Tcheukam, A., & Tembine, H. (2017). Mean-field-type games in engineering. *AIMS Electronics and Electrical Engineering*, *1*(1), 18.

Duncan, T., & Tembine, H. (2018). Linear-quadratic mean-field-type games: A direct method. *Games*, *9*(7).

Duncan, T. E. (2014). Linear-quadratic stochastic differential games with general noise processes. *Models and Methods in Economics and Management Science: Essays in Honor of Charles S. Tapiero, (eds. F. El Ouardighi and K. Kogan), Operations Research and Management Series, Springer Intern. Publishing, Switzerland*, *198*, 17-26.

Duncan, T. E. (2016). Linear exponential quadratic stochastic differential games.

IEEE Transactions on Automatic Control, *61*, 2550-2552.

Duncan, T. E., & Pasik-Duncan, B. (2012). Linear-exponential-quadratic control for stochastic equations in a hilbert space. *Dynamic Systems and Applications*, *21*, 407–416.

Duncan, T. E., & Pasik-Duncan, B. (2013). Linear-quadratic fractional gaussian control. *SIAM Journal on Control and Optimization*, *51*, 4604–4619.

Lasry, J. M., & Lions, P. L. (2007). Mean field games. *Japanese Journal of Mathematics, 2 (1):229–260*.

Nash, J. (1951). Non-cooperative games. *Annals of Mathematics*, *54*(2), 286-295.

Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press, Princeton.

Selten, R. (1974). Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*(4), 25–55.

Stackelberg, H. V. (1948). The theory of the market economy, trans. by aj peacock, london, william hodge. *Originally published as Grundlagen der Theoretischen Volkswirtschaftlehre.*

Tcheukam, A., & Tembine, H. (2016). Mean-field-type games for distributed power networks in presence of prosumers. In *2016 28th chinese control and decision conference (ccdc)* (p. 446-451). 2016.

Tembine, H. (2015). Risk-sensitive mean-field-type games with l^p-norm drifts. *Automatica*, *59*(2015), 224-237.

Tembine, H. (2017). Mean-field-type games. *AIMS Math*, *2*(4), 706–735.

Tian R. Yu, Z., & Zhang, R. (2020). A closed-loop saddle point for zero-sum linear-quadratic stochastic differential games with mean-field type. *Systems Control Letters*, *1136*(1), 104624.

Yong, J., & Zhou, X. Y. (1999). *Stochastic controls: Hamiltonian systems and hjb equations* (Vol. 43). Springer.

Øksendal, B. (2003). *Stochastic differential equations, an introduction with applications*. Universitext. Springer-Verlag, Berlin, sixth edition.

Øksendal, B., & Sulem-Bialobroda, A. (2st Edition, 2009). *Applied stochastic control of jump diffusions*. Springer Berlin, Heidelberg.

Part II

Research Papers in:

Hierarchical Mean Field Type Games and Application

Paper A

Hierarchical structures and leadership design in mean-field-type games with polynomial cost

Zahrate El Oula Frihi[1], Julian Barreiro-Gomez [2], Salah Eddine Choutri [3] and Hamidou Tembine[4]

The paper has been published in
Games journal, 11(3), 30; 2020. https://doi.org/10.3390/g11030030

Abstract

This article presents a class of hierarchical mean-field-type games with multiple layers and non-quadratic polynomial costs. The decision-makers act in sequential order with informational differences. We first examine the single layer case where each decision-maker does not have the information about the other control strategies. We derive the Nash mean-field-type equilibrium and cost in linear state-and-mean-field feedback form by using a partial integro-differential system. Then, we examine the Stackelberg two-layer problem with multiple leaders and multiple followers. Numerical illustrations show that, in the symmetric case, having only one leader is not necessarily optimal for the total sum cost. Having too many leaders may be also suboptimal for the total sum cost. The methodology is extended to multi-level hierarchical systems. It is shown that the order of the play plays a key role in the total performance of the system. We also identify specific range of parameters for which the Nash equilibrium coincides with the hierarchical solution independently of the number of layers and the order of play. In the heterogeneous case, it is shown that the total cost is significantly affected by the design of the hierarchical structure of the problem.

1 Introduction

The idea of hierarchy dates back a least to 1934, when Stackelberg (Stackelberg, 1948) introduced a game that models markets where some firms have stronger influence on others. Stackelberg games consist of two players, a leader and a follower. The leader who moves first, decides an optimal strategy after anticipating the best response of the follower. Then, the follower eventually chooses the anticipated best response to optimize her cost or payoff. Therefore, this game is a game with two-level hierarchy. A dynamic Linear-Quadratic (LQ) Stackelberg differential game was studied by Samaan and Cruz (Simaan & Cruz, 1973a). A stochastic LQ Stackelberg differential game was investigated by Bagchi and Başar(Bagchi & Basar, 1981). Bensoussan et al. (Bensoussan, Chen, & Sethi, 2015) derive a maximum principle for the leader's Stackelberg solution under the adapted closed-loop memoryless information structure.

Two or more players, the Stackelberg game is called hierarchical game and it becomes more interesting and involved due to its multi-layer structure including various forms of information. The players act in sequential order such that each one of them is a leader for the previous and a follower of the next player in the hierarchy. For hierarchical mean-field-free differential games, see e.g. (Pan & Yong, 1991; Simaan & Cruz, 1973b; Cruz, 1978; Gardner & Cruz, 1978; Basar & Selbuz, 1979).

Only few works consider hierarchical structures in mean-field related games. Open-loop Stackelberg solutions are addressed in linear-quadratic setting in (Lin, Jiang, & Zhang, 2019; Du & Wu, 2019);and in the context of large populations, mean-field Stackelberg games are investigated in (Moon & Basar, 2015; Bensoussan, Chau, & Yam, 2015; Bensoussan, Chau, Lai, & Yam, 2017; Averboukh, 2018; Moon & Basar, 2018). Besides, the leader-follower configuration has been used in several problems and fields to illustrate and model a variety of hierarchical behaviors. For instance, in (Shi, Wang, & Xiong, 2016), a leader-follower stochastic differential game with asymmetric information is studied motivated by applications in finance, economics and management engineering. In (Nourian, Caines, Malhamé, & Huang, 2012), a large population leader-follower stochastic multi-agent systems is analyzed with coupled cost functions and by using a mean-field Linear-Quadratic-Gaussian (LQG) approach. Regarding control applications, (Cai & Hu, 2017) presents a tracking control design in a distributed manner in a multiagent system configured in a leader-follower fashion, and it is shown that the setup can be used to model the power sharing problem in microgrids. In (Li, Shi, & Chen, 2018), a security problem in networked control systems is studied by means of a Stackelberg approach, and in (Barreiro-Gomez, Ocampo-Martinez, & Quijano, 2017) a hierarchical control structure or sequential predictive control is designed for a large-scale water system. In (Sutter & Rivas, 2014), leadership is studied in the context of public goods games by means of the reward and punishment effects. *The works mentioned above do not consider hierarchical mean-field-type game setting where the payoff functionals are non-linear with respect to the probability measure of the state.*

Hierarchical mean-field-type game theory studies a class of hierarchical games in which the payoffs and/or state dynamics depend not only on the state-action pairs but also the distribution of them (Barreiro-Gomez, Duncan, & Tembine, 2019). This class of games offers several features:

- A single decision-maker can have a strong impact on the mean-field terms,
- The expected payoffs are not necessarily linear with respect to the state distribution,
- The number of decision-makers is not necessarily infinite.

Games with non-linear distribution-dependent quantity-of-interest are very attractive in terms of applications since the non-linear dependence of the payoff functions in terms of state distribution allows us to capture risk measures, which are functionals of variance, inverse quantile, and/or higher moments. In portfolio optimization, for instance, payoff functions may include the third and the fourth moments known as the kurtosis and skewness (e.g.(Beardsley, Field, & Xiao, 2012; Theodossiou & Savva, 2016)). Generally, equilibrium solutions to mean-field type games are presented as either open-loop or closed-loop solutions. The open-loop solutions are controls that do not explicitly depend on the state process at time t, i.e., they are rather adapted processes that depend only on time and the initial data. The stochastic maximum principle can be used as a methodology for finding such optimal control strategies. Closed-loop solutions (aka feedback solutions) are deterministic functions that depend on the state of the process at time t as well as its marginal distribution. The dual adjoint functions which are obtained from the Hamilton-Jacobi-Bellman (HJB) equations can be used for finding feedback optimal controls. We will use this approach throughout this paper. For linear quadratic stochastic differential games, Sun and Yong (Sun & Yong, 2014) established that the existence of open-loop optimal control strategies is equivalent to the solvability of the corresponding optimality system which is a forward-backward (Stochastic Differential Equation) SDE, and the existence of closed-loop optimal strategies is equivalent to the existence of a regular solution to the corresponding Riccati equation.

Our contribution can be summarized as follows. This work examines a class of hierarchical mean-field-type games with multiple layers, multiple leaders, and multiple followers. Based on infinite dimensional partial integro-differential equations (PIDEs) on the space of measures, we provide semi-explicit solutions in closed-loop form of a class of master systems with hierarchical structure and non-quadratic cost, which are not covered in the earlier works. Recall that the non-quadratic costs allow analyzing other classes of higher risk terms, e.g., kurtosis (Beardsley et al., 2012; Theodossiou & Savva, 2016). The novelty of this paper mainly lies in the analysis of the effect of hierarchy and leadership on the solutions.

The rest of this article is structured as follows. We present the model setup in Section 2. Section 3 investigates the Nash equilibrium (no leader). Section 4 presents Stackelberg solution. The multi-layer case is presented in Section 5. Numerical examples are presented in Section 6. Finally, concluding remarks are drawn in Section 7.

2 The Setup

There are $I \geq 2$ number of decision-makers interacting within the time horizon $[t_0, t_1]$, $t_0 < t_1$. The set of decision-makers is denoted by $\mathcal{I} = \{1, 2, \dots, I\}$. Decision-maker $i \in \mathcal{I}$ has a control action $u_i \in U_i = \mathbb{R}$. The state x is driven by a Drift-Jump-Diffusion process of mean-field type given by

$$dx = bdt + \sigma dB + \int_{\Theta} \mu(., \theta)\tilde{N}(dt, d\theta), \quad x(t_0) \sim m(t_0, .),$$

where

Drift: $b \ : [t_0, t_1] \times \mathbb{R} \times \prod_{j=1}^{I} U_j \times \mathcal{P}(\mathbb{R}) \to \mathbb{R}$,
Diffusion coefficient: $\sigma : \ [t_0, t_1] \times \mathbb{R} \times \prod_{j=1}^{I} U_j \times \mathcal{P}(\mathbb{R}) \to \mathbb{R}$,
Brownian motion B,

Set of jump size: $\Theta = \mathbb{R}_+ \backslash \{0\}$,
Jump: $N(dt, d\theta)$,
Compensated jump: $\tilde{N}(dt, d\theta) = N(dt, d\theta) - \nu(d\theta)dt$,
Jump rate: $\mu : \ [t_0, t_1] \times \mathbb{R} \times \prod_{j=1}^{I} U_j \times \mathcal{P}(\mathbb{R}) \times \Theta \to \mathbb{R}$,

where $\mathcal{P}(\mathbb{R})$ denotes the set of probability measures on $\mathbb{R}$. We assume that $x(t_0), B$ and N are mutually independent. The performance functional of decision-maker i is

$$L_i(u, m_0) = h_i(x(t_1), m(t_1)) + \int_{t_0}^{t_1} l_i(t, x, u, m)dt,$$

where $m(t,dy)=\mathbb{P}_{x(t)}(dy)$ is the probability measure of the state $x(t)$ at time t, and

$$\begin{array}{l} l_i:\ [t_0,t_1]\times\mathbb{R}\times\prod_{j=1}^{I}U_j\times\mathcal{P}(\mathbb{R})\to\mathbb{R},\\ h_i:\mathbb{R}\times\mathcal{P}(\mathbb{R})\to\mathbb{R}. \end{array}$$

In addition, each decision-maker is assumed to have a computational capability such as being able to compute an aggregative term of m from the model. Let $\mathcal{U}_i$ be the set of control strategies of decision-maker i that are progressively measurable with respect to the filtration generated by the unions of events in $\{B,N\}$.

2.1 Games with polynomial cost

We investigate the mean-field-type game with the following data:

$$\begin{aligned}
t_0 &= 0,\quad t_1=T>0, && \text{(A.1a)}\\
l_i(t,x,u,m) &= q_i\frac{(x-\bar{x})^{2k_i}}{2k_i}+r_i\frac{(u_i-\bar{u}_i)^{2k_i}}{2k_i}+c_i(x-\bar{x})^{2k_i-1}(u_i-\bar{u}_i)\\
&\quad+\sum_{j\in\mathcal{I}\backslash\{i\}}\epsilon_{ij}(x-\bar{x})^{2(k_i-1)}(u_i-\bar{u}_i)(u_j-\bar{u}_j)\\
&\quad+\bar{q}_i\frac{\bar{x}^{2\bar{k}_i}}{2\bar{k}_i}+\bar{r}_i\frac{\bar{u}_i^{2\bar{k}_i}}{2\bar{k}_i}+\bar{c}_i\bar{x}^{2\bar{k}_i-1}\bar{u}_i+\sum_{j\neq i}\bar{\epsilon}_{ij}\bar{x}^{2(\bar{k}_i-1)}\bar{u}_i\bar{u}_j, && \text{(A.1b)}\\
h_i(x,m) &= q_{iT}\frac{(x_T-\bar{x}_T)^{2k_i}}{2k_i}+\bar{q}_{iT}\frac{\bar{x}_T^{2\bar{k}_i}}{2\bar{k}_i}, && \text{(A.1c)}\\
b(t,x,u,m) &= b_1(x-\bar{x})+\bar{b}_1\bar{x}+\sum_{j\in\mathcal{I}}\left[b_{2j}(u_j-\bar{u}_j)+\bar{b}_{2j}\bar{u}_j\right], && \text{(A.1d)}\\
\sigma(t,x,u,m) &= (x-\bar{x})\tilde{\sigma}, && \text{(A.1e)}
\end{aligned}$$

$$\begin{aligned}
\mu(t,x,u,m,\theta) &= (x-\bar{x})\tilde{\mu}(.,\theta), && \text{(A.2a)}\\
\bar{x}(t) &= \int ym(t,dy), && \text{(A.2b)}\\
\bar{u}_i(t) &= \int u_i(t,y,m)m(t,dy),\ i\in\mathcal{I}, && \text{(A.2c)}
\end{aligned}$$

where $k_i\geq 1,\ \bar{k}_i\geq 1$, are natural numbers, and the coefficients are time-dependent. The coefficient functions $q_i,r_i,\ \bar{q}_i$ and $\bar{r}_i$ are nonnegative functions and

$$\int_{t_0}^{t_1}\left[\tilde{\sigma}^2(t)+\int_{\Theta}\Big((1+\tilde{\mu}(t,\theta))^{2k_i}-1-2k_i\tilde{\mu}(t,\theta)\Big)\nu(d\theta)\right]dt<+\infty.$$

2.2 Hierarchical Leader Design and Algorithmic Approach

The hierarchical leadership design consists of finding the optimal number of hierarchical layers h and the non-empty subsets of players $\mathcal{I}_1, \dots, \mathcal{I}_h$ partitioning the set of all players as

$$\mathcal{I} = \bigcup_{k=1}^{h} \mathcal{I}_k, \text{ and if } k \neq k',\ \mathcal{I}_k \cap \mathcal{I}_{k'} = \emptyset.$$

The performance functional for the hierarchical design is the sum cost at the chosen hierarchical solution, i.e.,

$$\inf_h \inf_{(\mathcal{I}_1,\dots,\mathcal{I}_h):\cup_{k=1}^h \mathcal{I}_k = \mathcal{I}} S(h, \mathcal{I}_1, \dots, \mathcal{I}_h).$$

Here, we take into consideration three main game scenarios described as follows. First, the game has a unique layer, i.e., a situation in which all the players select their strategies simultaneously. Second, the game is played in two layers (bi-level hierarchy). The players are grouped into two sets ($h = 2$) known as leaders, those who decide first and simultaneously, and followers, those who react against the decision of the leaders. Third, the game is structured to take into account as many layers as the number of players (fully hierarchical configuration with $h = I$), i.e., players select strategically in sequence one by one in I layers. For all configurations, let L_i^* denote the optimal cost of the player $i \in \mathcal{I}$ in the hierarchical mean-field-type game problem and $S(h, \mathcal{I}_1, \dots, \mathcal{I}_h) = \sum_{i\in\mathcal{I}} L_i^*$ denotes the total (social) cost at the hierarchical solution. The hierarchical leadership design consists of determining the optimal leaders, followers, and/or number of layers such that the total cost is minimized

Notice that, for both the bi-level and fully hierarchical cases, there are multiple combinations for the players. In the bi-level scenario, the set of all possible sets of leaders is given by the power set $2^{\mathcal{I}}$, and any set of leaders is denoted by $\mathcal{I}_L \subseteq 2^{\mathcal{I}}$ with the corresponding set of followers $\mathcal{I}_F = \mathcal{I} \setminus \mathcal{I}_L$. Regarding the fully-hierarchical game, there are as many possibilities in the strategic ordering as permutations of the set of players $\mathcal{I}$. All possible permutations of the players are considered.

For the bi-level case, the optimal set for leaders and followers is

$$\begin{aligned} \mathcal{I}_L^* &\in \arg\min_{2^{\mathcal{I}}} \ S(2, \mathcal{I}_1, \mathcal{I}_2), \\ \mathcal{I}_F^* &= \mathcal{I} \setminus \mathcal{I}_L^*. \end{aligned}$$

On the other hand, for the fully-hierarchical case, we have that the optimal permutation is

$$(\mathcal{I}_1^*, \dots, \mathcal{I}_I^*) \in \arg\min_{\mathcal{I}_1,\dots,\mathcal{I}_I} \ S(I, \mathcal{I}_1, \dots, \mathcal{I}_I).$$

In this paper, we study the three aforementioned scenarios involving one, two, and I layers as presented in Figure A.1. We also present under which conditions all the three configurations have the same solution, i.e., when the Nash solution coincides with the hierarchical solutions at different layers. Furthermore, we present numerical examples considering different levels of hierarchy. The problem addressed in this paper can be interpreted as a mechanism design that, instead of determining the appropriate cost functionals or utility functions to induce a desired output, we design the best hierarchical structure in order to reduce the overall social cost.

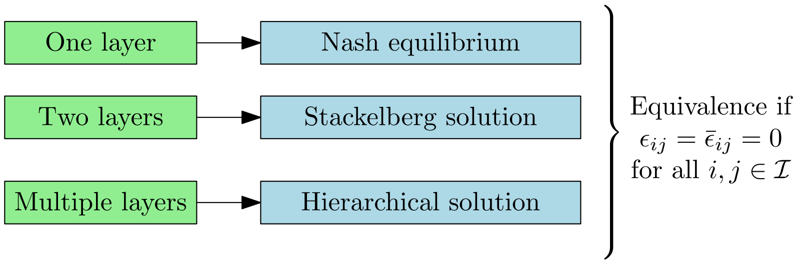


Fig. A.1: Different hierarchical designs and their solution concepts are considered in this paper.

Remark 2.1 (Feasibility and Existence). *The set of possible combinations for the layers/levels and players per level is non-empty and finite. Then, the optimal hierarchical leader design is feasible and there exists an optimal solution (combination) such that the social cost is minimized.* □

Since the feasible set of possible combinations for the hierarchical configurations is non-empty and finite, then it is possible to find the best hierarchical structure by means of Algorithm 1. The main results evoked in the Algorithm 1 given by Propositions 1, 2, and 3, are presented throughout the paper.

Algorithm 1: Finding the best hierarchical structure

Result: Leadership design in multi-level hierarchical games
initialization;
$\mathcal{I} \leftarrow \{1, \ldots, I\}$, set of decision-makers;
$I \leftarrow |\mathcal{I}|$, number of decision-makers;
$\mathcal{H}^* \leftarrow \mathcal{I}$, initialization for the partition ;
$S^* \leftarrow \infty$, initial arbitrary social cost ;
$tb \leftarrow B(I)$, number of possible leadership structures;
$\{\mathcal{H}(1), \ldots, \mathcal{H}(B(I))\} \leftarrow$ set of all $B(I)$ possible leadership structures;
$i \leftarrow 0$, initial index to test the leadership structure;
while $i \leq tb$ **do**
  $\ell \leftarrow$ number of levels in the structure $\mathcal{H}(i)$;
  $\{\mathcal{I}_1, \ldots, \mathcal{I}_\ell\} \leftarrow$ partition from $\mathcal{H}(i)$;
  switch ℓ **do**
    if $\ell = 1$**:** $S(I, \{\mathcal{I}_1, \ldots, \mathcal{I}_\ell\}) \leftarrow$ social cost for the Nash game Proposition 1 ;
    if $\ell = 2$**:** $S(I, \{\mathcal{I}_1, \ldots, \mathcal{I}_\ell\}) \leftarrow$ social cost for the Stackelberg game Proposition 2 ;
    if $\ell \geq 3$**:** $S(I, \{\mathcal{I}_1, \ldots, \mathcal{I}_\ell\}) \leftarrow$ social cost for the Hierarchical game Proposition 3 ;
  end
  if $S(I, \{\mathcal{I}_1, \ldots, \mathcal{I}_\ell\}) < S^*$ **then**
    $\mathcal{H}^* \leftarrow \mathcal{H}(i)$ update of the hierarchical structure;
    $S^* \leftarrow S(I, \{\mathcal{I}_1, \ldots, \mathcal{I}_\ell\})$ update of the social cost;
  else
    $\mathcal{H}(i)$ is a candidate optimal design;
  end
  $i \leftarrow i + 1$;
end
The optimal leadership design is $\mathcal{H}^*$ with social cost S^*;

According to the procedure in Algorithm 1, one of the main concerns in the leadership design problem is related to the dimensionality of the feasible set for the hierarchical structures (NP-hard problem). The total number of combinations is given by the total number of ordered partitions from a set, such total combinations are computed by means of the *ordered Bell number* $B : \mathbb{N} \to \mathbb{N}$, i.e., for I players we have:

$$B(I) = \sum_{k=0}^{I} \sum_{j=0}^{k} (-1)^{k-j} \left(\begin{array}{c} k \\ j \end{array} \right) j^n.$$

For instance, if $I = 2$, then there are $B(2) = 3$ possible leadership configurations as shown in Figure A.2; i $I = 3$, then there are $B(3) = 13$ possible leadership structures presented in Figure A.3, and $B(4) = 75$, $B(5) = 541$, and $B(6) = 4683$. Figure A.4 illustrates the rapid increment of the number of combinations as the decision-makers increase. Notice that it is not possible to have more levels than players in the hierarchical game ($h \leq I$). The following sections are devoted to the presentation of semi-explicit solutions for hierarchical mean-field-type games with different levels from one (Nash scenario) up to the number of players I (fully-hierarchical scenario).

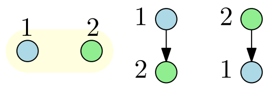


Fig. A.2: Possible combinations in the hierarchical leadership design for two decision-makers. Ordered Bell number $B(2) = 3$.

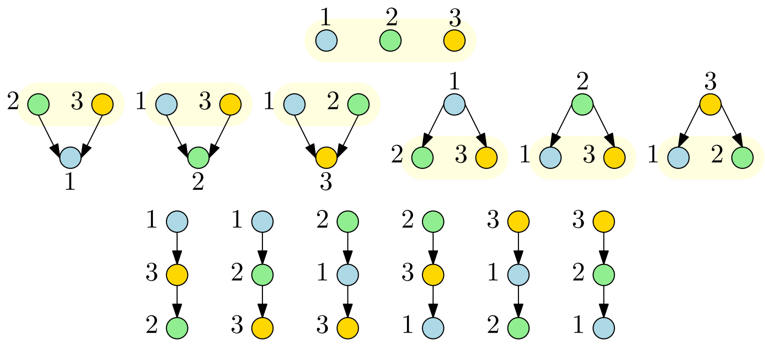


Fig. A.3: Possible combinations in the hierarchical leadership design for three decision-makers. Ordered Bell number $B(3) = 13$.

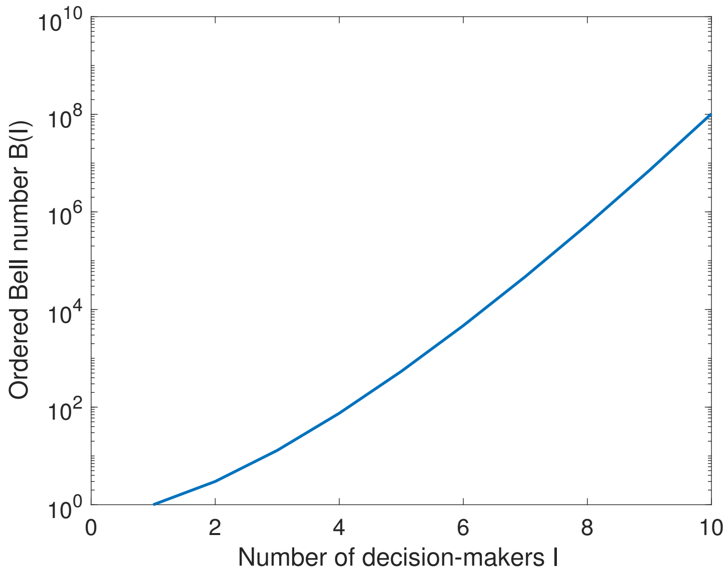


Fig. A.4: Number of possible hierarchical structures for a given set of decision-makers described by the ordered Bell number $B(I)$.

3 Nash Mean-Field-Type Equilibrium

The risk-neutral mean-field-type game is given by

$$(\mathcal{I}, U_i, \mathcal{U}_i, \mathbb{E}[L_i])_{i\in\mathcal{I}}.$$

A risk-neutral Nash Mean-Field-Type Equilibrium is a solution of the following fixed-point problem:

$$\begin{aligned}
&i \in \mathcal{I},\\
&\mathbb{E}[L_i(u^*)] = \inf_{u_i\in\mathcal{U}_i} \mathbb{E}[L_i(u_1^*, \dots, u_{i-1}^*, u_i, u_{i+1}^*, \dots, u_I^*)].
\end{aligned}$$

Let $\hat{V}_i(t, m)$ be the optimal cost-to-go from m at time $t \in (t_0, t_1)$ given the strategies of the others, i.e.,

$$\hat{V}_i(t, m) = \inf_{u_i} \mathbb{E}[h_i(x(t_1), m(t_1)) + \int_t^{t_1} l_i(t, x, u, m)dt' | m(t) = m].$$

We say that $\hat{V}_{i,m}(t, x, m) := \hat{V}_{i,m}(t, m)(x)$ is a Gâteaux derivative of $\hat{V}_i(t, m)$ with respect to the measure m if

$$\lim_{\tau\to 0} \frac{d}{d\tau} \hat{V}_i(t, m + \tau\tilde{m}) = \int \hat{V}_{i,m}(t, m)(x)\tilde{m}(dx). \quad \text{(A.3)}$$

If $\int \tilde{m}(dx) = 0$ then by adding a constant to $\hat{V}_{i,m}(t,x,m)$ does not change the value of the integral in (A.3). For any scalar λ and $m \in \mathcal{P}(\mathbb{R})$ one has, $\lambda = \lambda \int m(dx)$. Thus, λ is also a Gâteaux-derivative of the constant function λ. However in our problem, the term $\hat{V}_{i,xm}$ which is the gradient of $x \mapsto \hat{V}_{i,m}(t,x,m)$ will be used in the Hamiltonian, and $\hat{V}_{i,xm}$ does not have the constant ambiguity. Let us denote the jump operator J as

$$J[\phi_i] := \int_\Theta [\phi_{i,m}(t^-, x+\mu) - \phi_{i,m} - \mu\phi_{i,xm}]\nu(d\theta),$$

Let us introduce the integrand Hamiltonian as

$$\begin{aligned} &H_i(t,x,m,\hat{V}_m,\hat{V}_{xm},\hat{V}_{xxm}) \\ &\quad = \inf_{u_i \in U_i} \left\{ l_i + b\,\hat{V}_{i,xm} + \frac{\sigma^2}{2}\hat{V}_{i,xxm} + \int_\Theta [\hat{V}_{i,m}(t^-, x+\mu) - \hat{V}_{i,m} - \mu\hat{V}_{i,xm}]\nu(d\theta) \right\}. \end{aligned}$$

A sufficiency condition for a risk-neutral Nash equilibrium system is given by the following PIDE system:

$$0 = \hat{V}_{i,t}(t,m) + \int H_i(t,x,m,\hat{V}_m,\hat{V}_{xm},\hat{V}_{xxm})m(dx), \tag{A.4a}$$

$$\hat{V}_i(t_1,m) = \int m(dy)h_i(y,m), i \in \mathcal{I}. \tag{A.4b}$$

We refer the reader to Proposition 6.2 from (Bensoussan, Djehiche, Tembine, & Yam, 2020) for a derivation of this equilibrium system. The system (A.4) is an infinite dimensional PIDE system in m and it provides the Nash equilibrium values of the mean-field-type game. Notice that from (A.4) the equilibrium strategies, are the best response to the integrand Hamiltonian and can be expressed as functions of $t, x, m, \hat{V}_{i,m}, \hat{V}_{i,xm}, \hat{V}_{i,xxm}$.

Next, we provide semi-explicitly the Nash mean-field-type equilibrium in linear state-and-mean-field feedback strategies. To do so, we use (A.4).

Proposition 3.1. *A risk-neutral Nash mean-field-type equilibrium is given in a semi-explicit way as follows:*

$$u_i^{ne} = -\eta_i\left(x - \int ym(dy)\right) - \bar{\eta}_i \int ym(dy), \tag{A.5a}$$
$$0 = -r_i\eta_i^{2k_i-1} - \sum_{j\neq i}\epsilon_{ij}\eta_j + b_{2i}\alpha_i + c_i, \tag{A.5b}$$
$$0 = -\bar{r}_i\bar{\eta}_i^{2\bar{k}_i-1} - \sum_{j\neq i}\bar{\epsilon}_{ij}\bar{\eta}_j + \bar{b}_{2i}\bar{\alpha}_i + \bar{c}_i, \tag{A.5c}$$

$$\hat{V}_i(t,m) = \alpha_i \int_x \frac{(x-\int ym(dy))^{2k_i}}{2k_i} m(dx) + \bar{\alpha}_i \frac{(\int ym(dy))^{2\bar{k}_i}}{2\bar{k}_i}, \tag{A.6a}$$
$$0 = \dot{\alpha}_i + q_i + r_i\eta_i^{2k_i} - 2k_ic_i\eta_i + 2k_i\sum_{j\neq i}\epsilon_{ij}\eta_i\eta_j + 2k_i\alpha_i[b_1 - \sum_{j\in\mathcal{I}} b_{2j}\eta_j] \tag{A.6b}$$
$$+ 2k_i(2k_i-1)\alpha_i\frac{1}{2}\tilde{\sigma}^2 + \alpha_i\int_\Theta \left[(1+\tilde{\mu})^{2k_i} - 1 - 2k_i\tilde{\mu}\right]\nu(d\theta),$$
$$\alpha_i(T) = q_{iT}, \tag{A.6c}$$
$$0 = \dot{\bar{\alpha}}_i + \bar{q}_i + \bar{r}_i\bar{\eta}_i^{2\bar{k}_i} - 2\bar{k}_i\bar{c}_i\bar{\eta}_i + 2\bar{k}_i\sum_{j\neq i}\bar{\epsilon}_{ij}\bar{\eta}_i\bar{\eta}_j + 2\bar{k}_i\bar{\alpha}_i\left[\bar{b}_1 - \sum_j \bar{b}_{2j}\bar{\eta}_j\right], \tag{A.6d}$$
$$\bar{\alpha}_i(T) = \bar{q}_{iT}, \tag{A.6e}$$

for all $i \in \mathcal{I}$ with

$$\int ym(t,dy) = \left[\int ym(0,dy)\right] e^{\int_0^t \left[\bar{b}_1 - \sum_j \bar{b}_{2j}\bar{\eta}_j\right]dt'}, \tag{A.6f}$$

whenever the above coefficient system admits a solution that does not escape within $[t_0, t_1]$. □

Proof. Under the assumption of perfect state observation and perfect knowledge of the model, a sufficiency condition for equilibrium is given by the PIDE system (A.4). We aim to solve (A.4). To do so, we start with the following guess functional of decision-maker i as

$$\hat{V}_i(t,m) = \alpha_i(t) \int_x \frac{(x-\int ym(dy))^{2k_i}}{2k_i} m(dx) + \bar{\alpha}_i(t) \frac{(\int ym(dy))^{2\bar{k}_i}}{2\bar{k}_i},$$

where the coefficient functions α_i and $\bar{\alpha}_i$ need to be determined. Notice that, for $k_i = 1$, the functional $\hat{V}_i(t,m)$ becomes a mean-variance-dependent functional, and for an arbitrary parameter k_i, the functional may support higher order moments. We compute the key terms $\hat{V}_{i,m}(t,m)$, $\hat{V}_{i,xm}(t,m)$, $\hat{V}_{i,xxm}(t,m)$.

$$\begin{aligned}\hat{V}_{i,m}(t,m) = &-\alpha_i x \int \left(y - \int zm(dz)\right)^{2k_i-1} m(dy) + \alpha_i \frac{(x-\int ym(dy))^{2k_i}}{2k_i} \\ &+ \bar{\alpha}_i x \left(\int ym(dy)\right)^{2k_i-1}, \end{aligned} \tag{A.7a}$$

$$\begin{aligned}\hat{V}_{i,xm}(t,m) = &-\alpha_i \int \left(y - \int zm(dz)\right)^{2k_i-1} m(dy) + \alpha_i \left(x - \int ym(dy)\right)^{2k_i-1} \\ &+ \bar{\alpha}_i \left(\int ym(dy)\right)^{2k_i-1}, \end{aligned} \tag{A.8a}$$

$$\hat{V}_{i,xxm}(t,m) = (2k_i-1)\alpha_i \left(x - \int ym(dy)\right)^{2(k_i-1)}, \tag{A.8b}$$

$$\hat{V}_{i,m}(t,m)(x+\mu) - \hat{V}_{i,m}(t,m)(x) - \mu \hat{V}_{i,xm}(t,m)(x) \tag{A.8c}$$

$$= \alpha_i \frac{(x-\int ym(dy))^{2k_i}}{2k_i}\left[(1+\tilde{\mu})^{2k_i} - 1 - 2k_i\tilde{\mu}\right] + \tilde{\epsilon}, \tag{A.8d}$$

with $\int \tilde{\epsilon} m(dy) = 0$. The Integrand Hamiltonian is strictly convex in $(u_i - \bar{u}_i, \bar{u}_i)$. The optimal control strategy is the unique minimizer of

$$\begin{aligned} &r_i \frac{(u_i-\bar{u}_i)^{2k_i}}{2k_i} + c_i(x-\bar{x})^{2k_i-1}(u_i-\bar{u}_i) + \sum_{j\neq i} \epsilon_{ij}(x-\bar{x})^{2(k_i-1)}(u_i-\bar{u}_i)(u_j-\bar{u}_j) \\ &+ \left[\hat{V}_{i,xm}(t,m) - \int \hat{V}_{i,xm}(t,m)(x)m(dx)\right] \sum_{j\in\mathcal{I}} b_{2j}(u_j-\bar{u}_j) + \bar{r}_i \frac{\bar{u}_i^{2\bar{k}_i}}{2\bar{k}_i} + \bar{c}_i \bar{x}^{2\bar{k}_i-1}\bar{u}_i \\ &+ \sum_{j\neq i} \bar{\epsilon}_{ij} \bar{x}^{2(\bar{k}_i-1)} \bar{u}_i \bar{u}_j + \left[\int \hat{V}_{i,xm}(t,m)(x)m(dx)\right] \sum_j \bar{b}_{2j}\bar{u}_j. \end{aligned} \tag{A.9}$$

By strictly convexity and by orthogonality between $(u_i - \bar{u}_i)$ and $\bar{u}_i$ the following condition system holds:

$$\begin{aligned} &i \in \mathcal{I}, \\ &0 = r_i(u_i-\bar{u}_i)^{2k_i-1} + c_i(x-\bar{x})^{2k_i-1} + \sum_{j\neq i}\epsilon_{ij}(x-\bar{x})^{2(k_i-1)}(u_j-\bar{u}_j) \\ &\quad + \left[\hat{V}_{i,xm}(t,m) - \int \hat{V}_{i,xm}(t,m)(x)m(dx)\right] b_{2i}, \end{aligned} \tag{A.10a}$$

$$0 = \bar{r}_i \bar{u}_i^{2\bar{k}_i-1} + \bar{c}_i \bar{x}^{2\bar{k}_i-1} + \sum_{j\neq i} \bar{\epsilon}_{ij}\bar{x}^{2(\bar{k}_i-1)}\bar{u}_j + \left[\int \hat{V}_{i,xm}(t,m)(x)m(dx)\right]\bar{b}_{2i}. \tag{A.10b}$$

By solving the previously mentioned conditions, one obtains the optimal control input in a closed-loop form. The linear state-and-mean-field-type feedback strategy $u_i = -\eta_i(x - \int ym(dy)) - \bar{\eta}_i \int ym(dy),\ i \in \mathcal{I}$ solves the system if the coefficients satisfy

$$
\begin{aligned}
& i \in \mathcal{I}, \\
& 0 = -r_i \eta_i^{2k_i - 1} - \sum_{j \neq i} \epsilon_{ij} \eta_j + b_{2i} \alpha_i + c_i, && \text{(A.11a)} \\
& 0 = -\bar{r}_i \bar{\eta}_i^{2\bar{k}_i - 1} - \sum_{j \neq i} \bar{\epsilon}_{ij} \bar{\eta}_j + \bar{b}_{2i} \bar{\alpha}_i + \bar{c}_i, && \text{(A.11b)}
\end{aligned}
$$

The integrand Hamiltonian of i becomes

$$
\begin{aligned}
H_i = & \left[q_i + r_i \eta_i^{2k_i} - 2k_i c_i \eta_i + 2k_i \sum_{j \neq i} \epsilon_{ij} \eta_i \eta_j \right] \frac{(x - \int ym(dy))^{2k_i}}{2k_i} \\
& + 2k_i \alpha_i \left[b_1 - \sum_{j \in \mathcal{I}} b_{2j} \eta_j \right] \frac{(x - \int ym(dy))^{2k_i}}{2k_i} + 2k_i(2k_i - 1)\alpha_i \frac{1}{2} \tilde{\sigma}^2 \frac{(x - \int ym(dy))^{2k_i}}{2k_i} \\
& + \alpha_i \int_\Theta [(1 + \tilde{\mu})^{2k_i} - 1 - 2k_i \tilde{\mu}] \nu(d\theta) \frac{(x - \int ym(dy))^{2k_i}}{2k_i} \\
& + \left[\bar{q}_i + \bar{r}_i \bar{\eta}_i^{2\bar{k}_i} - 2\bar{k}_i \bar{c}_i \bar{\eta}_i \right] \frac{(\int ym(dy))^{2\bar{k}_i}}{2\bar{k}_i} + \left[2\bar{k}_i \sum_{j \neq i} \bar{\epsilon}_{ij} \bar{\eta}_i \bar{\eta}_j \right] \frac{(\int ym(dy))^{2\bar{k}_i}}{2\bar{k}_i} \\
& + 2\bar{k}_i \bar{\alpha}_i \left[\bar{b}_1 - \sum_j \bar{b}_{2j} \bar{\eta}_j \right] \frac{(\int ym(dy))^{2\bar{k}_i - 1}}{2\bar{k}_i} + \tilde{\epsilon}_2. && \text{(A.12)}
\end{aligned}
$$

By identification, the coefficients α_i solve the following ordinary differential equation:

$$
\begin{aligned}
0 = & \dot{\alpha}_i + q_i + r_i \eta_i^{2k_i} - 2k_i c_i \eta_i + 2k_i \sum_{j \neq i} \epsilon_{ij} \eta_i \eta_j + 2k_i \alpha_i [b_1 - \sum_{j \in \mathcal{I}} b_{2j} \eta_j] \\
& + 2k_i(2k_i - 1)\alpha_i \frac{1}{2} \tilde{\sigma}^2 + \alpha_i \int_\Theta [(1 + \tilde{\mu})^{2k_i} - 1 - 2k_i \tilde{\mu}] \nu(d\theta), && \text{(A.13a)} \\
\alpha_i(T) = & \ q_{iT}, && \text{(A.13b)} \\
0 = & \ \dot{\bar{\alpha}}_i + \bar{q}_i + \bar{r}_i \bar{\eta}_i^{2\bar{k}_i} - 2\bar{k}_i \bar{c}_i \bar{\eta}_i + 2\bar{k}_i \sum_{j \neq i} \bar{\epsilon}_{ij} \bar{\eta}_i \bar{\eta}_j + 2\bar{k}_i \bar{\alpha}_i [\bar{b}_1 - \sum_j \bar{b}_{2j} \bar{\eta}_j], && \text{(A.13c)} \\
\bar{\alpha}_i(T) = & \ \bar{q}_{iT}. && \text{(A.13d)}
\end{aligned}
$$

The aggregate mean-field term $\int ym(t, dy)$ can be derived in a semi-explicit way by taking the expected value of the state dynamics. It follows that

$$
\int ym(t, dy) = \left[\int ym(0, dy) \right] e^{\int_0^t \left[\bar{b}_1 - \sum_j \bar{b}_{2j} \bar{\eta}_j \right] dt}. \qquad \blacksquare
$$

The following Remark discusses the existence and uniqueness of the η terms in Proposition 3.1.

Remark 3.1. *The uniqueness of the coefficient system (A.5) in η requires a strong condition, i.e.,*

$$0 = -r_i\eta_i^{2k_i-1} - \sum_{j\neq i}\epsilon_{ij}\eta_j + b_{2i}\alpha_i + c_i.$$

- *Let I be an arbitrary integer and $k_i = k = 1$, the system in η becomes linear and has a unique solution if and only if the determinant of the matrix M is non-zero, with $M_{ii} = r_i$ and $M_{ij} = \epsilon_{ij}$, $i \neq j$. When the determinant is zero, the resulting control strategies become non-admissible and the costs become infinite.*

- *For $k_i = k = 2$, and $I = 2$ the system in η becomes a binary cubic polynomial given by*

  $$\begin{aligned} r_1\eta_1^3 + \epsilon_{12}\eta_2 - b_{21}\alpha_1 - c_1 = 0,\\ r_2\eta_2^3 + \epsilon_{21}\eta_1 - b_{22}\alpha_2 - c_2 = 0. \end{aligned}$$

  For $\epsilon_{12} = 0$ there is a unique solution given by

  $$\eta_1 = \left(\frac{b_{21}\alpha_1 + c_1}{r_1}\right)^{\frac{1}{3}},\ \eta_2 = \left(\frac{-\epsilon_{21}\eta_1 + b_{22}\alpha_2 + c_2}{r_2}\right)^{\frac{1}{3}}.$$

  For $\epsilon_{12} \neq 0$ we derive from the first equation that

  $$\eta_2 = \frac{-r_1\eta_1^3 + b_{21}\alpha_1 + c_1}{\epsilon_{12}}.$$

  By substituting it to the second equation we arrive at

  $$r_2\left(\frac{-r_1\eta_1^3 + b_{21}\alpha_1 + c_1}{\epsilon_{12}}\right)^3 + \epsilon_{21}\eta_1 - b_{21}\alpha_1 - c_1 = 0,$$

  The latter equation is a polynomial of odd degree "9". It has a unique real root in η_1 if its derivative has a constant sign. Its derivative is

  $$\epsilon_{21} - 9\frac{r_1r_2}{\epsilon_{12}}\eta_1^2\left(\frac{-r_1\eta_1^3 + b_{21}\alpha_1 + c_1}{\epsilon_{12}}\right)^2.$$

  It has a constant sign if ϵ_{21} and $\frac{r_1r_2}{\epsilon_{12}}$ have opposite signs. If r_1 and r_2 are positive then the condition is reduced to

  $$\epsilon_{21}\epsilon_{12} \leq 0.$$

- *$I = 2$ and arbitrary $k_i \geq 1$. Thus, a sufficiency condition is that ϵ_{ji} and $(2k_i - 1)(2k_j - 1)\frac{r_i r_j}{\epsilon_{ij}}$ have opposite signs. In particular if $k_i \geq 1, k_j \geq 1,\ r_i > 0, r_j > 0$, then the condition reduces to*

$$\epsilon_{ij}\epsilon_{ji} \leq 0.$$

- *The same reasoning applies to the system in $\bar{\eta}$ and has a unique real solution if*

$$\bar{\epsilon}_{ij}\bar{\epsilon}_{ji} \leq 0.$$

- *For $I \geq 3$ decision-makers and arbitrary $k_i \geq 1$ the system can be rewritten as a fixed-point equation which fulfills a contraction mapping condition if the norms of r and ϵ are sufficiently small. In this case, there is a unique solution.*

□

In the next section, we investigate the bi-level case with multiple leaders and multiple followers.

4 Multiple Leaders and Multiple Followers

We consider the description in (A.1) in a bi-level hierarchical game with two and more leaders, i.e., $|\mathcal{I}_L| \geq 2$, and two and more followers, i.e., $|\mathcal{I}_F| \geq 2$.
We restrict our attention to the admissible strategies which are Lipschitz in the state x. Given the strategies of the leaders $(u_i)_{i\in\mathcal{I}_L} \in \prod_{i\in\mathcal{I}_L}\mathcal{U}_i$, a risk-neutral best response strategy of follower j is a strategy that solves $\inf_{\mathcal{U}_j}\ \mathbb{E}[L_j]$. The set of risk-neutral best responses of j is denoted by $\mathrm{rnBR}_j((u_i)_{i\in\mathcal{I}_L}, (u_{j'})_{j'\in\mathcal{I}_F\setminus\{j\}})$.

A mean-field-type risk-neutral Nash equilibrium among the followers given the first movers' strategies $(u_i)_{i\in\mathcal{I}_L} \in \prod_{i\in\mathcal{I}_L}\mathcal{U}_i$, is a strategy profile $(u_j,\ j \in \mathcal{I}_F)$ of all followers such that for every decision-maker $j \in \mathcal{I}_F$,

$$u_j \in \mathrm{rnBR}_j((u_i)_{i\in\mathcal{I}_L}; (u^{\mathrm{rn}}_{j'})_{j'\in\mathcal{I}_F\setminus\{j\}}).$$

The followers solve the following Nash game given the strategy of the leaders

$(u_i)_{i\in\mathcal{I}_L}$, i.e.,

$$j \in \mathcal{I}_F :$$
$$0 = \hat{V}_{j,t}(t,m) + \int H_j^r(x,m,(\hat{V}_{j',m},\hat{V}_{j',xm},\hat{V}_{j',xxm})_{j'\in\mathcal{I}_F}|(u_i)_{i\in\mathcal{I}_L})m(dx), \tag{A.14a}$$
$$\hat{V}_j(t_1,m) = \int m(dy)h_j(y,m), \tag{A.14b}$$
$$H_j^r = \inf_{u_j\in U_j}\left\{ l_j + b\,\hat{V}_{j,xm} + \frac{\sigma^2}{2}\hat{V}_{j,xxm} + J[\hat{V}_{j,m}]|(u_i)_{i\in\mathcal{I}_L}\right\}. \tag{A.14c}$$

Then, the leaders solve the following PIDE system:

$$i \in \mathcal{I}_L :$$
$$0 = \hat{V}_{i,t}(t,m) + \int H_i^r(x,m,(\hat{V}_{i',m},\hat{V}_{i',xm},\hat{V}_{i,xxm})_{i'\in\mathcal{I}_L\cup\mathcal{I}_F})m(dx), \tag{A.15a}$$
$$\hat{V}_i(t_1,m) = \int m(dy)h_i(y,m), \tag{A.15b}$$
$$H_i^r = \inf_{u_i\in U_i}\left\{ l_i + b\,\hat{V}_{i,xm} + \frac{\sigma^2}{2}\hat{V}_{i,xxm} + J[\hat{V}_{i,m}]|\{u_j^*(.,(u_i)_{i\in\mathcal{I}_L})\}_{j\in\mathcal{I}_F}\right\}, \tag{A.15c}$$

A minimizer of the integrand Hamiltonian H_i^r, denoted by

$$u_i^{ss} = u_i^{ss}(t,x,m,(\hat{V}_{i',m},\hat{V}_{i',xm},\hat{V}_{i',xxm})_{i'\in\mathcal{I}_L\cup\mathcal{I}_F}),$$

provides a candidate Stackelberg strategy of the leader i. A mean-field-type risk-neutral Stackelberg solution between multiple leaders and multiple followers is a strategy $((u_i^{ss})_{i\in\mathcal{I}_L},(u_j^{ss})_{j\in\mathcal{I}_F})$ of all decision-makers such that

$$i \in \mathcal{I}_L,$$
$$u_i^{ss} \in \arg\min_{u_i\in\mathcal{U}_i}\left\{ \mathbb{E}L_i(x,u_i,(u_{i'}^{ss})_{i\in\mathcal{I}_L\setminus\{i\}},(u_j^{ss})_{j\in\mathcal{I}_F}) : \right.$$
$$\left. u_j^{ss} \in \mathrm{rnBR}_j((u_i^{ss})_{i\in\mathcal{I}_L};(u_{j'}^{ss})_{j'\in\mathcal{I}_F\setminus\{j\}}\right\},$$

and for every follower

$$j \in \mathcal{I}_F,\ u_j^{ss} \in \mathrm{rnBR}_j((u_i^{ss})_{i\in\mathcal{I}_L};(u_{j'}^{ss})_{j'\in\mathcal{I}_F\setminus\{j\}}).$$

The next result presents the Stackelberg mean-field-type solution involving several leaders and followers in a semi-explicit manner.

Proposition 4.1. *The risk-neutral Stackelberg mean-field-type solution with multiple leaders and multiple followers is given in a semi-explicit way as follows:*

$$
\begin{aligned}
& u_j^{ss} = -\eta_j \left(x - \int ym(dy)\right) - \bar{\eta}_j \int ym(dy), j \in \mathcal{I}_F, && \text{(A.16a)}\\
& j \in \mathcal{I}_F: \\
& 0 = -r_j \eta_j^{2k_j-1} - \sum_{j' \in \mathcal{I}_F \backslash \{j\}} \epsilon_{jj'} \eta_{j'} - \sum_{i \in \mathcal{I}_L} \epsilon_{ji} \eta_i + b_{2j} \alpha_j + c_j, \\
& 0 = -\bar{r}_j \bar{\eta}_j^{2\bar{k}_j-1} - \sum_{j' \in \mathcal{I}_F \backslash \{j\}} \bar{\epsilon}_{jj'} \bar{\eta}_{j'} - \sum_{i \in \mathcal{I}_L} \bar{\epsilon}_{ji} \bar{\eta}_i + \bar{b}_{2j} \bar{\alpha}_j + \bar{c}_j, \\
& i \in \mathcal{I}_L: \\
& 0 = -r_i \eta_i^{2k_i-1} - \sum_{i' \in \mathcal{I}_L \backslash \{i\}} \epsilon_{ii'} \eta_{i'} - \sum_{j \in \mathcal{I}_F} \epsilon_{ij} \eta_j + b_{2i} \alpha_i + \sum_{j \in \mathcal{I}_F} \epsilon_{ij} \eta_i \frac{\epsilon_{ji}}{(2k_j - 1) r_j \eta_j^{2k_j-2}} \\
& \quad - \sum_{j \in \mathcal{I}_F} b_{2j} \frac{\epsilon_{ji}}{(2k_j - 1) r_j \eta_j^{2k_j-2}} \alpha_i + c_i, \\
& 0 = -\bar{r}_i \bar{\eta}_i^{2\bar{k}_i-1} - \sum_{i' \in \mathcal{I}_L \backslash \{i\}} \bar{\epsilon}_{ii'} \bar{\eta}_{i'} - \sum_{j \in \mathcal{I}_F} \bar{\epsilon}_{ij} \bar{\eta}_j + \bar{b}_{2i} \bar{\alpha}_i + \sum_{j \in \mathcal{I}_F} \bar{\epsilon}_{ij} \bar{\eta}_i \frac{\bar{\epsilon}_{ji}}{(2\bar{k}_j - 1) \bar{r}_j \bar{\eta}_j^{2\bar{k}_j-2}} \\
& \quad - \sum_{j \in \mathcal{I}_F} \bar{b}_{2j} \frac{\bar{\epsilon}_{ji}}{(2\bar{k}_j - 1) \bar{r}_j \bar{\eta}_j^{2\bar{k}_j-2}} \bar{\alpha}_i + \bar{c}_i,
\end{aligned}
$$

and

$$
\begin{aligned}
& \hat{V}_i(0, m) = \alpha_i(0) \int_x \frac{(x - \int ym_0(dy))^{2k_i}}{2k_i} m_0(dx) + \bar{\alpha}_i(0) \frac{(\int ym_0(dy))^{2\bar{k}_i}}{2\bar{k}_i}, && \text{(A.16b)}\\
& 0 = \dot{\alpha}_i + q_i + r_i \eta_i^{2k_i} - 2k_i c_i \eta_i + 2k_i \sum_{i' \in \mathcal{I}_L \backslash \{i\}} \epsilon_{ii'} \eta_i \eta_{i'} + 2k_i \sum_{j \in \mathcal{I}_F} \epsilon_{ij} \eta_i \eta_j \\
& \quad + 2k_i \left[b_1 - \sum_{i' \in \mathcal{I}_L} b_{2i'} \eta_{i'} - \sum_{j \in \mathcal{I}_F} b_{2j} \eta_j \right] \alpha_i + 2k_i (2k_i - 1) \alpha_i \frac{1}{2} \tilde{\sigma}^2 \\
& \quad + \alpha_i \int_\Theta [(1 + \tilde{\mu})^{2k_i} - 1 - 2k_i \tilde{\mu}] \nu(d\theta), && \text{(A.16c)}\\
& \alpha_i(T) = q_{iT}, && \text{(A.16d)}\\
& 0 = \dot{\bar{\alpha}}_i + \bar{q}_i + \bar{r}_i \bar{\eta}_i^{2\bar{k}_i} - 2\bar{k}_i \bar{c}_i \bar{\eta}_i + 2\bar{k}_i \sum_{i' \in \mathcal{I}_L \backslash \{i\}} \bar{\epsilon}_{ii'} \bar{\eta}_i \bar{\eta}_{i'} + 2\bar{k}_i \sum_{j \in \mathcal{I}_F} \bar{\epsilon}_{ij} \bar{\eta}_i \bar{\eta}_j \\
& \quad + 2\bar{k}_i \left[\bar{b}_1 - \sum_{i' \in \mathcal{I}_L} \bar{b}_{2i'} \bar{\eta}_{i'} - \sum_{j \in \mathcal{I}_F} \bar{b}_{2j} \bar{\eta}_j \right] \bar{\alpha}_i, && \text{(A.16e)}\\
& \bar{\alpha}_i(T) = \bar{q}_{iT}, && \text{(A.16f)}
\end{aligned}
$$

with

$$\int ym(t,dy) = \left[\int ym(0,dy)\right] e^{\int_0^t [\bar{b}_1 - \sum_j \bar{b}_{2j}\bar{\eta}_j] dt'}, \tag{A.16g}$$

whenever the above coefficient system admits a unique solution. □

Proof. For the data in (A.1), the integrand Hamiltonian H_j^r has a unique minimizer, denoted by

$$u_j^* = u_j^*(t,x,m,(\hat{V}_{j',m},\hat{V}_{j',xm},\hat{V}_{j',xxm})_{j'\in\mathcal{I}_F},(u_i)_{i\in\mathcal{I}_L}),$$

which provides the reaction strategies of the follower decision-makers. Following (A.1) with leaders in $\mathcal{I}_L$ and followers in $\mathcal{I}_F$, the first order optimality condition yields

$$\begin{aligned}
& j \in \mathcal{I}_F, \\
& 0 = r_j(u_j-\bar{u}_j)^{2k_j-1} + c_i(x-\bar{x})^{2k_i-1} + \sum_{j'\in\mathcal{I}_F\setminus\{j\}} \epsilon_{jj'}(x-\bar{x})^{2(k_j-1)}(u_{j'}-\bar{u}_{j'}) \\
& \quad + \sum_{i\in\mathcal{I}_L} \epsilon_{ji}(x-\bar{x})^{2(k_j-1)}(u_i-\bar{u}_i) + \left[\hat{V}_{j,xm}(t,m) - \int \hat{V}_{j,xm}(t,m)(x)m(dx)\right] b_{2j},
\end{aligned} \tag{A.17a}$$

$$\begin{aligned}
& 0 = \bar{r}_j\bar{u}_j^{2\bar{k}_j-1} + \bar{c}_i\bar{x}^{2\bar{k}_i-1} + \sum_{j'\in\mathcal{I}_F\setminus\{j\}} \bar{\epsilon}_{jj'}\bar{x}^{2(\bar{k}_j-1)}\bar{u}_{j'} + \sum_{i\in\mathcal{I}_L} \bar{\epsilon}_{ji}\bar{x}^{2(\bar{k}_j-1)}\bar{u}_i \\
& \quad + \left[\int \hat{V}_{j,xm}(t,m)(x)m(dx)\right]\bar{b}_{2j},
\end{aligned} \tag{A.17b}$$

and

$$j \in \mathcal{I}_F,$$

$$\sum_{i\in\mathcal{I}_L} \epsilon_{ji}\eta_i = -r_j\eta_j^{2k_j-1} - \sum_{j'\in\mathcal{I}_F\setminus\{j\}} \epsilon_{jj'}\eta_{j'} + b_{2j}\alpha_j + c_j, \tag{A.18a}$$

$$\sum_{i\in\mathcal{I}_L} \bar{\epsilon}_{ji}\bar{\eta}_i = -\bar{r}_j\bar{\eta}_j^{2\bar{k}_j-1} - \sum_{j'\in\mathcal{I}_F\setminus\{j\}} \bar{\epsilon}_{jj'}\bar{\eta}_{j'} + \bar{b}_{2j}\bar{\alpha}_j + \bar{c}_j, \tag{A.18b}$$

which provides $\{\eta_j,\bar{\eta}_j\}_{j\in\mathcal{I}_F}$ as function of $\{\eta_i,\bar{\eta}_i\}_{i\in\mathcal{I}_L}$ and $\alpha,\bar{\alpha}$. Following (A.1) with leaders in $\mathcal{I}_L$ and followers in $\mathcal{I}_F$, the leaders' integrand Hamiltonian can be rewritten as follows

$$
\begin{aligned}
H_i^r &= \inf_{u_i \in U_i} \{l_i + b\, \hat{V}_{i,xm}\} + \frac{\sigma^2}{2}\hat{V}_{i,xxm} + J[\hat{V}_{i,m}], \\
&= \inf_{u_i \in U_i} q_i \frac{(x-\bar{x})^{2k_i}}{2k_i} + r_i \frac{(u_i - \bar{u}_i)^{2k_i}}{2k_i} + c_i (x-\bar{x})^{2k_i-1}(u_i - \bar{u}_i) \\
&+ \sum_{i' \in \mathcal{I}_L \setminus \{i\}} \epsilon_{ii'} (x-\bar{x})^{2(k_i-1)} (u_i - \bar{u}_i)(u_{i'} - \bar{u}_{i'}) \\
&+ \sum_{j \in \mathcal{I}_F} \epsilon_{ij} (x-\bar{x})^{2(k_i-1)} (u_i - \bar{u}_i)(u_j^* - \bar{u}_j^*) + \bar{q}_i \frac{\bar{x}^{2\bar{k}_i}}{2\bar{k}_i} + \bar{r}_i \frac{\bar{u}_i^{2\bar{k}_i}}{2\bar{k}_i} + \bar{c}_i \bar{x}^{2\bar{k}_i-1}\bar{u}_i \\
&+ \sum_{i' \in \mathcal{I}_L \setminus \{i\}} \bar{\epsilon}_{ii'} \bar{x}^{2(\bar{k}_i-1)} \bar{u}_i \bar{u}_{i'} + \sum_{j \in \mathcal{I}_F} \bar{\epsilon}_{ij} \bar{x}^{2(\bar{k}_i-1)} \bar{u}_i \bar{u}_j^* \\
&+ \left\{ b_1(x-\bar{x}) + \sum_{i' \in \mathcal{I}_L} b_{2i'}(u_{i'} - \bar{u}_{i'}) + \sum_{j \in \mathcal{I}_F} b_{2j}(u_j^* - \bar{u}_j^*) \right\} \hat{V}_{i,xm} \\
&+ \{\bar{b}_1 \bar{x} + \sum_{i' \in \mathcal{I}_L} \bar{b}_{2i'} \bar{u}_{i'} + \sum_{j \in \mathcal{I}_F} \bar{b}_{2j} \bar{u}_j^* \} \hat{V}_{i,xm} + \frac{\sigma^2}{2}\hat{V}_{i,xxm} + J[\hat{V}_{i,m}]
\end{aligned}
$$

In view of (A.17),

$$
\begin{cases}
\dfrac{\partial(u_j^* - \bar{u}_j^*)}{\partial(u_i - \bar{u}_i)} = -\dfrac{\epsilon_{ji}}{(2k_j-1) r_j \eta_j^{2k_j-2}}, \\
\dfrac{\partial \bar{u}_j^*}{\partial \bar{u}_i} = -\dfrac{\bar{\epsilon}_{ji}}{(2\bar{k}_j-1) \bar{r}_j \bar{\eta}_j^{2\bar{k}_j-2}},
\end{cases}
$$

The optimal Stackelberg strategies of the leaders satisfy the following system:

$$
\begin{aligned}
0 &= r_i (u_i - \bar{u}_i)^{2k_i-1} + c_i (x-\bar{x})^{2k_i-1} + \sum_{i' \in \mathcal{I}_L \setminus \{i\}} \epsilon_{ii'} (x-\bar{x})^{2(k_i-1)} (u_{i'} - \bar{u}_{i'}) \\
&+ \sum_{j \in \mathcal{I}_F} \epsilon_{ij} (x-\bar{x})^{2(k_i-1)} (u_j^* - \bar{u}_j^*) - \sum_{j \in \mathcal{I}_F} \epsilon_{ij} (x-\bar{x})^{2(k_i-1)} (u_i - \bar{u}_i) \frac{\epsilon_{ji}}{(2k_j-1) r_j \eta_j^{2k_j-2}} \\
&+ \left[b_{2i} - \sum_{j \in \mathcal{I}_F} b_{2j} \frac{\epsilon_{ji}}{(2k_j-1) r_j \eta_j^{2k_j-2}} \right] \alpha_i (x-\bar{x})^{2k_i-1}, \\
0 &= \bar{r}_i \bar{u}_i^{2\bar{k}_i-1} + \bar{c}_i \bar{x}^{2\bar{k}_i-1} + \sum_{i' \in \mathcal{I}_L \setminus \{i\}} \bar{\epsilon}_{ii'} \bar{x}^{2(\bar{k}_i-1)} \bar{u}_{i'} + \sum_{j \in \mathcal{I}_F} \bar{\epsilon}_{ij} \bar{x}^{2(\bar{k}_i-1)} \bar{u}_j^* \\
&- \sum_{j \in \mathcal{I}_F} \bar{\epsilon}_{ij} \bar{x}^{2(\bar{k}_i-1)} \bar{u}_i \frac{\bar{\epsilon}_{ji}}{(2\bar{k}_j-1) \bar{r}_j \bar{\eta}_j^{2\bar{k}_j-2}} + \left[\bar{b}_{2i} - \sum_{j \in \mathcal{I}_F} \bar{b}_{2j} \frac{\bar{\epsilon}_{ji}}{(2\bar{k}_j-1) \bar{r}_j \bar{\eta}_j^{2\bar{k}_j-2}} \right] \bar{\alpha}_i \bar{x}^{2\bar{k}_i-1},
\end{aligned}
$$

whose solution provides the coefficients $(\eta_i^{ss}, \bar{\eta}_i^{ss})_{i \in \mathcal{L}}$. ■

Remark 4.1. *Clearly, the mean-field-type Nash equilibrium in (A.5) differs from the Stackelberg solution in (A.16) when the ϵ_{ij} are non-zero.* □

No control-coupling within classes

It follows from (A.16) that, for $\epsilon_{jj'} = 0 = \bar{\epsilon}_{jj'}$ for $(j, j') \in \mathcal{I}_F^2$, the term η_j is explicitly given by

$$\eta_j = \left\{ \frac{-\sum_{i\in\mathcal{I}_L} \epsilon_{ji}\eta_i + b_{2j}\alpha_j + c_j}{r_j} \right\}^{\frac{1}{2k_j-1}},$$

and

$$\bar{\eta}_j = \left\{ \frac{-\sum_{i\in\mathcal{I}_L} \bar{\epsilon}_{ji}\bar{\eta}_i + \bar{b}_{2j}\bar{\alpha}_j + \bar{c}_j}{\bar{r}_j} \right\}^{\frac{1}{2\bar{k}_j-1}}.$$

No Leader and All Followers

In this case, there is no leader. All decision-makers are followers. This case is similar to the model proposed in the Nash game above. The solution is given by (A.5).

One Leader and Multiple Followers

There is a unique leader in $\mathcal{I}_L$, and the remaining decision-makers in $\mathcal{I}_F$ are followers. $\mathcal{I} = \mathcal{I}_L \cup \mathcal{I}_F$. We assume that, the leader (decision-maker $1 \in \mathcal{I}_L$) uses a state-and-mean-field type feedback strategy $u_1(t, x, m)$ and each of the followers (decision-maker $j \in \mathcal{I}_F$) finds state-and-mean-field type feedback strategy $u_j(t, x, m, u_1)$ given u_1. The followers solve a Nash game given the strategy of the leader u_1.

Multiple Leaders and One Follower

Since there is only one follower the reaction set of the follower will be computed given the strategies of the leaders.

All Leaders and No Follower

In this case, there is no follower. All decision-makers are leaders. In terms of information structure, this case is similar to the model proposed in the Nash game above. The solution is given by (A.5).

5 Fully hierarchical game

In the previous sections, we had only bi-level game problems. In this section we make as many levels as the number of decision-makers. There are $|\mathcal{I}|$ hierarchical levels. At each layer i, decision-maker i chooses a control strategy u_i knowing the control strategy of the preceding decision-makers i.e., $\{i-1, \dots, 1\}$. This becomes a sequential decision-making problem. We use a backward induction method to solve the hierarchical game problem. This means that, the decision-making problem at the last layer I, which is the reaction of decision-maker I, can be seen as a mean-field-type control problem. This is because at the $i-$th level, the strategies $(u_{i'})_{i' \in \{1,\dots,i-1\}}$ are already known by decision-maker i.

The Proposition 5.1 next, presents the multi-level hierarchical-structure solution in the context of mean-field-type games in a semi-explicit manner.

Proposition 5.1. *The risk-neutral $I-$level hierarchical mean-field-type solution is given in a semi-explicit way as follows:*

$$u_i^{hs} = -\eta_i \left(x - \int ym(dy) \right) - \bar{\eta}_i \int ym(dy), i \in \mathcal{I}, \tag{A.19a}$$

$$\hat{V}_i(0, m) = \alpha_i(0) \int_x \frac{(x - \int ym_0(dy))^{2k_i}}{2k_i} m_0(dx) + \bar{\alpha}_i(0) \frac{(\int ym_0(dy))^{2\bar{k}_i}}{2\bar{k}_i}, \tag{A.19b}$$

with

$$\int ym(t, dy) = \left[\int ym(0, dy) \right] e^{\int_0^t \left[\bar{b}_1 - \sum_j \bar{b}_{2j} \bar{\eta}_j \right] dt}, \tag{A.19c}$$

where the coefficient functions are given by

$$
\begin{aligned}
&\textit{Level } 1:\\
&\quad 0=-r_1\eta_1^{2k_1-1}+c_1-\sum_{j=2}^{I}\epsilon_{1,j}\eta_j+\sum_{j=2}^{I}\epsilon_{1,j}\eta_i\frac{\epsilon_{ji}}{(2k_j-1)r_j}\eta_j^{-2(k_j-1)}\\
&\qquad+\left[b_{2,1}-\sum_{j=2}^{I}b_{2j}\frac{\epsilon_{j1}}{(2k_j-1)r_j}\eta_j^{-2(k_j-1)}\right]\alpha_1,\\
&\quad 0=\dot{\alpha}_1+q_1+r_1\eta_1^{2k_1}-2k_1c_1\eta_1+2k_1\sum_{j=2}^{I}\epsilon_{1j}\eta_1\eta_j+2k_1\{b_1-b_{21}\eta_1-\sum_{j=2}^{I}b_{2j}\eta_j\}\alpha_1\\
&\qquad+2k_1(2k_1-1)\alpha_1\frac{1}{2}\tilde{\sigma}^2+\alpha_1\int_{\Theta}[(1+\tilde{\mu})^{2k_1}-1-2k_1\tilde{\mu}]\nu(d\theta),\\
&\alpha_1(T)=q_{1T},\\
&\quad 0=-\bar{r}_1\bar{\eta}_1^{2\bar{k}_1-1}+\bar{c}_1-\sum_{j=2}^{I}\bar{\epsilon}_{1,j}\bar{\eta}_j+\sum_{j=2}^{I}\bar{\epsilon}_{1,j}\bar{\eta}_1\frac{\bar{\epsilon}_{j1}}{(2\bar{k}_j-1)\bar{r}_j}\bar{\eta}_j^{-2(\bar{k}_j-1)}\\
&\qquad+\left[\bar{b}_{21}-\sum_{j=2}^{I}\bar{b}_{2j}\frac{\bar{\epsilon}_{j1}}{(2\bar{k}_j-1)\bar{r}_j}\bar{\eta}_j^{-2(\bar{k}_j-1)}\right]\bar{\alpha}_1,\\
&\quad 0=\dot{\bar{\alpha}}_1+\bar{q}_1+\bar{r}_1\bar{\eta}_1^{2\bar{k}_1}-2\bar{k}_1\bar{c}_1\bar{\eta}_1+2\bar{k}_1\sum_{j=2}^{I}\bar{\epsilon}_{1j}\bar{\eta}_i\bar{\eta}_j+2\bar{k}_1\{\bar{b}_1-\bar{b}_{21}\bar{\eta}_1-\sum_{j=2}^{I}\bar{b}_{2j}\bar{\eta}_j\}\bar{\alpha}_1,\\
&\bar{\alpha}_1(T)=\bar{q}_{1T}.
\end{aligned}
$$

Level i :

$$
\begin{aligned}
0 &= -r_i\eta_i^{2k_i-1} + c_i - \sum_{i'=1}^{i-1}\epsilon_{I-1,i'}\eta_{i'} - \sum_{j=i+1}^{I}\epsilon_{i,j}\eta_j + \sum_{j=i+1}^{I}\epsilon_{i,j}\eta_i\frac{\epsilon_{ji}}{(2k_j-1)r_j}\eta_j^{-2(k_j-1)} \\
&+ \left[b_{2i} - \sum_{j=i+1}^{I} b_{2j}\frac{\epsilon_{ji}}{(2k_j-1)r_j}\eta_j^{-2(k_j-1)}\right]\alpha_i, \\
0 &= \dot{\alpha}_i + q_i + r_i\eta_i^{2k_i} - 2k_ic_i\eta_i + 2k_i\sum_{i'=1}^{i-1}\epsilon_{ii'}\eta_i\eta_{i'} + 2k_i\sum_{j=i+1}^{I}\epsilon_{ij}\eta_i\eta_j \\
&+ 2k_i\{b_1 - \sum_{i'=1}^{i-1} b_{2i'}\eta_{i'} - b_{2i}\eta_i - \sum_{j=i+1}^{I} b_{2j}\eta_j\}\alpha_i + 2k_i(2k_i-1)\alpha_i\frac{1}{2}\tilde{\sigma}^2 \\
&+ \alpha_i\int_\Theta[(1+\tilde{\mu})^{2k_i} - 1 - 2k_i\tilde{\mu}]\nu(d\theta), \\
\alpha_i(T) &= q_{iT}, \\
0 &= -\bar{r}_i\bar{\eta}_i^{2\bar{k}_i-1} + \bar{c}_i - \sum_{i'=1}^{i-1}\bar{\epsilon}_{I-1,i'}\bar{\eta}_{i'} - \sum_{j=i+1}^{I}\bar{\epsilon}_{i,j}\bar{\eta}_j + \sum_{j=i+1}^{I}\bar{\epsilon}_{i,j}\bar{\eta}_i\frac{\bar{\epsilon}_{ji}}{(2\bar{k}_j-1)\bar{r}_j}\bar{\eta}_j^{-2(\bar{k}_j-1)} \\
&+ \left[\bar{b}_{2i} - \sum_{j=i+1}^{I}\bar{b}_{2j}\frac{\bar{\epsilon}_{ji}}{(2\bar{k}_j-1)\bar{r}_j}\bar{\eta}_j^{-2(\bar{k}_j-1)}\right]\bar{\alpha}_i, \\
0 &= \dot{\bar{\alpha}}_i + \bar{q}_i + \bar{r}_i\bar{\eta}_i^{2\bar{k}_i} - 2\bar{k}_i\bar{c}_i\bar{\eta}_i + 2\bar{k}_i\sum_{i'=1}^{i-1}\bar{\epsilon}_{ii'}\bar{\eta}_i\bar{\eta}_{i'} + 2\bar{k}_i\sum_{j=i+1}^{I}\bar{\epsilon}_{ij}\bar{\eta}_i\bar{\eta}_j \\
&+ 2\bar{k}_i\left[\bar{b}_1 - \sum_{i'=1}^{i-1}\bar{b}_{2i'}\bar{\eta}_{i'} - \bar{b}_{2i}\bar{\eta}_i - \sum_{j=i+1}^{I}\bar{b}_{2j}\bar{\eta}_j\right]\bar{\alpha}_i, \\
\bar{\alpha}_i(T) &= \bar{q}_{iT}.
\end{aligned}
$$

Level I :

$$\eta_I = \left(\frac{-\sum_{j=1}^{I-1}\epsilon_{I,j}\eta_j + b_{2I}\alpha_I + c_I}{r_I}\right)^{\frac{1}{2k_I-1}},$$

$$0 = \dot{\alpha}_I + q_I + r_I\eta_I^{2k_I} - 2k_Ic_I\eta_I + 2k_I\sum_{i'=1}^{I-1}\epsilon_{Ii'}\eta_I\eta_{i'} + 2k_I\{b_1 - \sum_{i'=1}^{I-1} b_{2i'}\eta_{i'} - b_{2I}\eta_I\}\alpha_I$$

$$+ 2k_I(2k_I-1)\alpha_I\frac{1}{2}\tilde{\sigma}^2 + \alpha_I\int_\Theta[(1+\tilde{\mu})^{2k_I} - 1 - 2k_I\tilde{\mu}]\nu(d\theta),$$

$$\alpha_I(T) = q_{IT},$$

$$\bar{\eta}_I = \left(\frac{-\sum_{j=1}^{I-1}\bar{\epsilon}_{I,j}\bar{\eta}_j + \bar{b}_{2I}\bar{\alpha}_I + \bar{c}_I}{\bar{r}_I}\right)^{\frac{1}{2\bar{k}_I-1}},$$

$$0 = \dot{\bar{\alpha}}_I + \bar{q}_I + \bar{r}_I\bar{\eta}_I^{2\bar{k}_I} - 2\bar{k}_I\bar{c}_I\bar{\eta}_I + 2\bar{k}_I\sum_{i'=1}^{I-1}\bar{\epsilon}_{Ii'}\bar{\eta}_I\bar{\eta}_{i'} + 2\bar{k}_I\{\bar{b}_1 - \sum_{i'=1}^{i-1}\bar{b}_{2i'}\bar{\eta}_{i'} - \bar{b}_{2I}\bar{\eta}_I\}\bar{\alpha}_I,$$

$$\bar{\alpha}_I(T) = \bar{q}_{IT},$$

whenever these equations admit a solution. □

Proof. We use a backward induction procedure to prove the statement.

I-th hierarchical level:

When decision-maker I optimizes the preceding decision-makers have already chosen their strategy and that is known by I. Hence, integrand Hamiltonian of I is

$$\begin{aligned}
H_I &= \inf_{u_I\in U_I}\{l_I + b\,\hat{V}_{I,xm}\} + \frac{\sigma^2}{2}\hat{V}_{I,xxm} + J[\hat{V}_{I,m}] \\
&= \inf_{u_I\in U_I} q_I\frac{(x-\bar{x})^{2k_I}}{2k_I} + r_I\frac{(u_I-\bar{u}_I)^{2k_I}}{2k_I} + c_I(x-\bar{x})^{2k_I-1}(u_I-\bar{u}_I) \\
&+ \sum_{i'=1}^{I-1}\epsilon_{I,i'}(x-\bar{x})^{2(k_I-1)}(u_I-\bar{u}_I)(u_{i'}-\bar{u}_{i'}) + \left[b_1(x-\bar{x}) + \sum_{i'=1}^{I-1}b_{2i'}(u_{i'}-\bar{u}_{i'})\right]\hat{V}_{I,xm} \\
&+ b_{2,I}(u_I-\bar{u}_I)\hat{V}_{I,xm} + \bar{q}_I\frac{\bar{x}^{2\bar{k}_I}}{2\bar{k}_I} + \bar{r}_I\frac{\bar{u}_I^{2\bar{k}_I}}{2\bar{k}_I} + \bar{c}_I\bar{x}^{2\bar{k}_I-1}\bar{u}_I + \sum_{i'=1}^{I-1}\bar{\epsilon}_{I,i'}\bar{x}^{2(\bar{k}_I-1)}\bar{u}_I\bar{u}_{i'} \\
&+ \left[\bar{b}_1\bar{x} + \sum_{i'=1}^{I-1}\bar{b}_{2i'}\bar{u}_{i'} + \bar{b}_{2,I}\bar{u}_I\right]\hat{V}_{I-1,xm} + \frac{\sigma^2}{2}\hat{V}_{I,xxm} + J[\hat{V}_{I,m}].
\end{aligned}$$

It follows from strictly convex optimization above that the best response strategy can be expressed as:

$$
\begin{aligned}
u_I^* - \bar{u}_I^* &= -\xi_1^{\frac{1}{2k_I-1}},\\
\bar{u}_I^* &= -\xi_2^{\frac{1}{2\bar{k}_I-1}},
\end{aligned}
$$

where

$$
\begin{aligned}
\xi_1 &= \frac{1}{r_I}\left(\sum_{i'=1}^{I-1} \epsilon_{I,i'}(x-\bar{x})^{2(k_I-1)}(u_{i'}-\bar{u}_{i'}) + b_{2,I}\hat{V}_{I,xm} + c_I(x-\bar{x})^{2k_I-1}\right),\\
\xi_2 &= \frac{1}{\bar{r}_I}\left(\sum_{i'=1}^{I-1} \bar{\epsilon}_{I,i'}\bar{x}^{2(\bar{k}_I-1)}\bar{u}_{i'} + \bar{b}_{2,I}\hat{V}_{I,xm} + \bar{c}_I\bar{x}^{2\bar{k}_I-1}\right).
\end{aligned}
$$

In particular,

$$
\begin{aligned}
& i \leq I-1 :\\
\frac{\partial(u_I^*-\bar{u}_I^*)}{\partial(u_i-\bar{u}_i)} &= \frac{\epsilon_{I,i}}{(2k_I-1)r_I}(x-\bar{x})^{2(k_I-1)}(u_I^*-\bar{u}_I^*)^{-2(k_I-1)}, \qquad \text{(A.23a)}\\
\frac{\partial\bar{u}_I^*}{\partial\bar{u}_i} &= \frac{\bar{\epsilon}_{I,i}}{(2\bar{k}_I-1)\bar{r}_I}\bar{x}^{2(\bar{k}_I-1)}(\bar{u}_I^*)^{-2(\bar{k}_I-1)}. \qquad \text{(A.23b)}
\end{aligned}
$$

If the preceding decision-makers $\{1,2,\ldots,I-1\}$ have all used linear state-and-mean-field feedback strategies then the reaction of the I-th decision-maker who is at I-th level of hierarchy can be rewritten as

$$
\begin{aligned}
u_I^{seq} &= -\eta_I\left(x - \int ym(dy)\right) - \bar{\eta}_I\int ym(dy),\\
\eta_I &= \left(\frac{-\sum_{j=1}^{I-1}\epsilon_{I,j}\eta_j + b_{2I}\alpha_I + c_I}{r_I}\right)^{\frac{1}{2k_I-1}},\\
\bar{\eta}_I &= \left(\frac{-\sum_{j=1}^{I-1}\bar{\epsilon}_{I,j}\bar{\eta}_j + \bar{b}_{2I}\bar{\alpha}_I + \bar{c}_I}{\bar{r}_I}\right)^{\frac{1}{2\bar{k}_I-1}}.
\end{aligned}
$$

$(I-1)$-th hierarchical level

At the hierarchical level $I-1$, the preceding levels are $\{1,2,\ldots,I-2\}$ and the succeeding level is I. Having the expression of the optimal control strategies of the last layer I we can move to the preceding layer, i.e., $I-1$. Decision-maker $I-1$

has $u_1, \ldots, u_{I-2}$ and the reaction u_I^* of decision-maker I. Therefore, the integrand Hamiltonian of $I-1$ is given by

$$
\begin{aligned}
H_{I-1}^r &= \inf_{u_{I-1}\in U_{I-1}} \{l_i + b\,\hat{V}_{I-1,xm}\} + \frac{\sigma^2}{2}\hat{V}_{I-1,xxm} + J[\hat{V}_{I-1,m}] \\
&= \inf_{u_{I-1}\in U_{I-1}} r_{I-1}\frac{(u_{I-1}-\bar{u}_{I-1})^{2k_{I-1}}}{2k_{I-1}} + q_{I-1}\frac{(x-\bar{x})^{2k_{I-1}}}{2k_{I-1}} \\
&+ c_{I-1}(x-\bar{x})^{2k_{I-1}-1}(u_{I-1}-\bar{u}_{I-1}) \\
&+ \sum_{i'=1}^{I-2}\epsilon_{I-1,i'}(x-\bar{x})^{2(k_{I-1}-1)}(u_{I-1}-\bar{u}_{I-1})(u_{i'}-\bar{u}_{i'}) \\
&+ \epsilon_{I-1,I}(x-\bar{x})^{2(k_{I-1}-1)}(u_{I-1}-\bar{u}_{I-1})(u_I^*-\bar{u}_I^*) \\
&+ \left[b_1(x-\bar{x}) + \sum_{i'=1}^{I-2} b_{2i'}(u_{i'}-\bar{u}_{i'}) + b_{2,I-1}(u_{I-1}-\bar{u}_{I-1}) + b_{2I}(u_I^*-\bar{u}_I^*)\right]\hat{V}_{I-1,xm} \\
&+ \bar{q}_{I-1}\frac{\bar{x}^{2\bar{k}_{I-1}}}{2\bar{k}_{I-1}} + \bar{r}_{I-1}\frac{\bar{u}_{I-1}^{2\bar{k}_{I-1}}}{2\bar{k}_{I-1}} + \bar{c}_{I-1}\bar{x}^{2\bar{k}_{I-1}-1}\bar{u}_{I-1} + \sum_{i'=1}^{I-2}\bar{\epsilon}_{I-1,i'}\bar{x}^{2(\bar{k}_{I-1}-1)}\bar{u}_{I-1}\bar{u}_{i'} \\
&+ \bar{\epsilon}_{I-1,I}\bar{x}^{2(\bar{k}_{I-1}-1)}\bar{u}_{I-1}\bar{u}_I^* + \{\bar{b}_1\bar{x} + \sum_{i'=1}^{I-2}\bar{b}_{2i'}\bar{u}_{i'} + \bar{b}_{2,I-1}\bar{u}_{I-1} + \bar{b}_{2I}\bar{u}_I^*\}\hat{V}_{I-1,xm} \\
&+ \frac{\sigma^2}{2}\hat{V}_{I-1,xxm} + J[\hat{V}_{I-1,m}]
\end{aligned}
$$

In view of (A.23), the terms with $\bar{u}_I^*$ depend on $\bar{u}_{I-1}$, $\bar{u}_{I-2}$, …, $\bar{u}_1$. The first-order optimality condition for u_{I-1}^* yields

$$
\begin{aligned}
0 &= -r_{I-1}\eta_{I-1}^{2k_{I-1}-1} + c_{I-1} - \sum_{i'=1}^{I-2}\epsilon_{I-1,i'}\eta_{i'} - \epsilon_{I-1,I}\eta_I \\
&\quad + \epsilon_{I-1,I}\eta_{I-1}\frac{\epsilon_{I,I-1}}{(2k_I-1)r_I}\eta_I^{-2(k_I-1)} + \left\{b_{2,I-1} - b_{2I}\frac{\epsilon_{I,I-1}}{(2k_I-1)r_I}\eta_I^{-2(k_I-1)}\right\}\alpha_{I-1}, \\
0 &= -\bar{r}_{I-1}\bar{\eta}_{I-1}^{2\bar{k}_{I-1}-1} + \bar{c}_{I-1} - \sum_{i'=1}^{I-2}\bar{\epsilon}_{I-1,i'}\bar{\eta}_{i'} - \bar{\epsilon}_{I-1,I}\bar{\eta}_I \\
&\quad + \bar{\epsilon}_{I-1,I}\bar{\eta}_{I-1}\frac{\bar{\epsilon}_{I,I-1}}{(2\bar{k}_I-1)\bar{r}_I}\bar{\eta}_I^{-2(\bar{k}_I-1)} + \left\{\bar{b}_{2,I-1} - \bar{b}_{2,I}\frac{\bar{\epsilon}_{I,I-1}}{(2\bar{k}_I-1)\bar{r}_I}\bar{\eta}_I^{-2(\bar{k}_I-1)}\right\}\bar{\alpha}_{I-1},
\end{aligned}
$$

where we have used (A.23) for $i = I-1$.

$$
u_{I-1}^{seq} = -\eta_{I-1}\left(x - \int ym(dy)\right) - \bar{\eta}_{I-1}\int ym(dy). \tag{A.24}
$$

i-th hierarchical level

For $i \in \{2, \dots, I-2\}$,

$$
\begin{aligned}
H_i^r = &\inf_{u_i \in U_i} q_i \frac{(x-\bar{x})^{2k_i}}{2k_i} + r_i \frac{(u_i - \bar{u}_i)^{2k_i}}{2k_i} + c_i (x-\bar{x})^{2k_i - 1}(u_i - \bar{u}_i) \\
&+ \sum_{i'=1}^{i-1} \epsilon_{ii'} (x-\bar{x})^{2(k_i-1)} (u_i - \bar{u}_i)(u_{i'} - \bar{u}_{i'}) + \sum_{j=i+1}^{I} \epsilon_{ij} (x-\bar{x})^{2(k_i-1)} (u_i - \bar{u}_i)(u_j^* - \bar{u}_j^*) \\
&+ \left[b_1 (x - \bar{x}) + \sum_{i'=1}^{i-1} b_{2i'} (u_{i'} - \bar{u}_{i'}) + b_{2i}(u_i - \bar{u}_i) + \sum_{j=i+1}^{I} b_{2j} (u_j^* - \bar{u}_j^*) \right] \hat{V}_{i,xm} \\
&+ \bar{q}_i \frac{\bar{x}^{2\bar{k}_i}}{2\bar{k}_i} + \bar{r}_i \frac{\bar{u}_i^{2\bar{k}_i}}{2\bar{k}_i} + \bar{c}_i \bar{x}^{2\bar{k}_i - 1} \bar{u}_i + \sum_{i'=1}^{i-1} \bar{\epsilon}_{ii'} \bar{x}^{2(\bar{k}_i - 1)} \bar{u}_i \bar{u}_{i'} + \sum_{j=i+1}^{I} \bar{\epsilon}_{ij} \bar{x}^{2(\bar{k}_i-1)} \bar{u}_i \bar{u}_j^* \\
&+ \left\{ \bar{b}_1 \bar{x} + \sum_{i'=1}^{i-1} \bar{b}_{2i'} \bar{u}_{i'} + \bar{b}_{2i} \bar{u}_i + \sum_{j=i+1}^{I} \bar{b}_{2j} \bar{u}_j^* \right\} \hat{V}_{i,xm} + \frac{\sigma^2}{2} \hat{V}_{i,xxm} + J[\hat{V}_{i,m}].
\end{aligned}
$$

By identification from the first-order optimality condition the coefficient functions $\eta_i, \bar{\eta}_i$ satisfy the following equations

$$
\begin{aligned}
0 = &-r_i \eta_i^{2k_i - 1} + c_i - \sum_{i'=1}^{i-1} \epsilon_{I-1,i'} \eta_{i'} - \sum_{j=i+1}^{I} \epsilon_{i,j} \eta_j + \sum_{j=i+1}^{I} \epsilon_{i,j} \eta_i \frac{\epsilon_{ji}}{(2k_j - 1) r_j} \eta_j^{-2(k_j - 1)} \\
&+ \left\{ b_{2i} - \sum_{j=i+1}^{I} b_{2j} \frac{\epsilon_{ji}}{(2k_j - 1) r_j} \eta_j^{-2(k_j-1)} \right\} \alpha_i, \\
0 = &-\bar{r}_i \bar{\eta}_i^{2\bar{k}_i - 1} + \bar{c}_i - \sum_{i'=1}^{i-1} \bar{\epsilon}_{I-1,i'} \bar{\eta}_{i'} - \sum_{j=i+1}^{I} \bar{\epsilon}_{i,j} \bar{\eta}_j + \sum_{j=i+1}^{I} \bar{\epsilon}_{i,j} \eta_i \frac{\bar{\epsilon}_{ji}}{(2\bar{k}_j - 1) \bar{r}_j} \bar{\eta}_j^{-2(\bar{k}_j - 1)} \\
&+ \left\{ \bar{b}_{2i} - \sum_{j=i+1}^{I} \bar{b}_{2j} \frac{\bar{\epsilon}_{ji}}{(2\bar{k}_j - 1) \bar{r}_j} \bar{\eta}_j^{-2(\bar{k}_j-1)} \right\} \bar{\alpha}_i,
\end{aligned}
$$

1-st hierarchical level We now examine first level of the hierarchy. The integrand

Hamiltonian of decision-maker 1 is

$$
\begin{aligned}
H_1^r = & \inf_{u_1\in U_1} q_1\frac{(x-\bar{x})^{2k_1}}{2k_1} + r_1\frac{(u_1-\bar{u}_1)^{2k_1}}{2k_1} + c_1(x-\bar{x})^{2k_1-1}(u_1-\bar{u}_1) \\
& + \sum_{j=2}^{I}\epsilon_{1j}(x-\bar{x})^{2(k_1-1)}(u_1-\bar{u}_1)(u_j^*-\bar{u}_j^*) + \left\{ b_1(x-\bar{x}) + b_{21}(u_1-\bar{u}_1)\right\}\hat{V}_{1,xm} \\
& + \left\{\sum_{j=2}^{I} b_{2j}(u_j^*-\bar{u}_j^*)\right\}\hat{V}_{1,xm} + \bar{q}_1\frac{\bar{x}^{2\bar{k}_1}}{2\bar{k}_1} + \bar{r}_1\frac{\bar{u}_1^{2\bar{k}_1}}{2\bar{k}_1} + \bar{c}_1\bar{x}^{2\bar{k}_1-1}\bar{u}_1 + \sum_{j=2}^{I}\bar{\epsilon}_{1j}\bar{x}^{2(\bar{k}_1-1)}\bar{u}_1\bar{u}_j^* \\
& + \left\{\bar{b}_1\bar{x} + \bar{b}_{21}\bar{u}_1 + \sum_{j=2}^{I}\bar{b}_{2j}\bar{u}_j^*\right\}\hat{V}_{1,xm} + \frac{\sigma^2}{2}\hat{V}_{1,xxm} + J[\hat{V}_{1,m}].
\end{aligned}
$$

By identification from the first-order optimality condition the coefficient functions $\eta_1, \bar{\eta}_1$ satisfy the following equations

$$
\begin{aligned}
0 = & -r_1\eta_1^{2k_1-1} + c_1 - \sum_{j=2}^{I}\epsilon_{1j}\eta_j + \sum_{j=2}^{I}\epsilon_{1j}\eta_1\frac{\epsilon_{j1}}{(2k_j-1)r_j}\eta_j^{-2(k_j-1)} \\
& + \left\{ b_{21} - \sum_{j=2}^{I} b_{2j}\frac{\epsilon_{j1}}{(2k_j-1)r_j}\eta_j^{-2(k_j-1)}\right\}\alpha_1, \\
0 = & -\bar{r}_1\bar{\eta}_1^{2\bar{k}_1-1} + \bar{c}_1 - \sum_{j=2}^{I}\bar{\epsilon}_{1j}\bar{\eta}_j + \sum_{j=2}^{I}\bar{\epsilon}_{1j}\bar{\eta}_1\frac{\bar{\epsilon}_{j1}}{(2\bar{k}_j-1)\bar{r}_j}\bar{\eta}_j^{-2(\bar{k}_j-1)} \\
& + \left\{ \bar{b}_{21} - \sum_{j=2}^{I} \bar{b}_{2j}\frac{\bar{\epsilon}_{j1}}{(2\bar{k}_j-1)\bar{r}_j}\bar{\eta}_j^{-2(\bar{k}_j-1)}\right\}\bar{\alpha}_1.
\end{aligned}
$$

Putting all together we arrive at the announced statement. ■

From the analysis above the following remarks are in order:

- For $\epsilon_{ij}\neq 0, \bar{\epsilon}_{ij}\neq 0$, the order of the play matters because of the informational difference between the decision-makers at different levels of hierarchy in (A.19). One open question that we leave for future investigation is: How to determine the optimal ordering among all permutations of heterogenous decision-makers?
- When all the ϵ_{ij} and $\bar{\epsilon}_{ij}$ are zero, the Nash equilibrium coincides with the bi-level solution, which coincides with any level hierarchical solution. The order of the play and the informational difference do not generate an extra advantage for the first mover in this particular case. Consequently, the hierarchical leader design is only performed when the parameters $\epsilon_{ij}\neq 0, \bar{\epsilon}_{ij}\neq 0$.

6 Numerical Investigation

In this section, we perform some numerical examples in order to analyze two main scenarios. We study the effect of the number of leaders on the total cost for both homogeneous and heterogeneous scenarios, and we investigate the effect of the hierarchical structure considering a heterogeneous scenario.

6.1 Effect of the number of leaders on the total cost

We investigate the effect of the number of leaders on the total performance of the system. The total cost at the Stackelberg solution is

$$S(\mathcal{I}_L, m_0) = \sum_{i\in\mathcal{I}_L} \hat{V}_i(0,m_0) + \sum_{j\in\mathcal{I}_F} \hat{V}_j(0,m_0).$$

For $m_0 = \delta_{x_0}$, and $\bar{k}_i = \bar{k} \geq 1$, the total cost is

$$S(\mathcal{I}_L, m_0) = \left(\sum_{i\in\mathcal{I}_L} \bar{\alpha}_i(0) + \sum_{j\in\mathcal{I}_F} \bar{\alpha}_j(0)\right) \frac{x_0^{2\bar{k}}}{2\bar{k}}.$$

Uniform coupling and homogeneous players

When all other parameters are identical across the players except their role, $S(\mathcal{I}_L, m_0)$ can be expressed as a function $|\mathcal{I}_L|$. It follows from (A.16) that

$$\begin{aligned}
&\chi := |\mathcal{I}_L|,\\
&0 = -\bar{r}(\bar{\eta}^{fo})^{2\bar{k}-1} - (|\mathcal{I}| - \chi - 1)\bar{\epsilon}\bar{\eta}^{fo} - \chi\bar{\epsilon}\bar{\eta}^{lead} + \bar{b}_2\bar{\alpha}^{fo} + c,\\
0 = &- \bar{r}(\bar{\eta}^{lead})^{2\bar{k}-1} - (\chi - 1)\bar{\epsilon}\bar{\eta}^{lead} - (|\mathcal{I}| - \chi)\bar{\epsilon}\bar{\eta}^{fo} + \bar{b}_2\bar{\alpha}^{lead}\\
&+ \bar{c} + \frac{\bar{\epsilon}(|\mathcal{I}| - \chi)(\bar{\epsilon}\bar{\eta}^{lead} - \bar{\alpha}^{lead}\bar{b}_2)}{(2\bar{k}-1)\bar{r}(\bar{\eta}^{fo})^{2\bar{k}-2}},\\
\bar{\alpha}^{lead}(t_0) = \bar{q}_{t_1} + &\int_{t_0}^{t_1} \Big\{ \bar{q} + \bar{r}(\bar{\eta}^{lead})^{2\bar{k}} - 2\bar{k}\bar{c}\bar{\eta}^{lead} + 2\bar{k}\bar{\epsilon}\bar{\eta}^{lead}[(\chi - 1)\bar{\eta}^{lead}\\
&+ (|\mathcal{I}| - \chi)\bar{\eta}^{fo}] + 2\bar{k}\bar{\alpha}^{lead}[\bar{b}_1 - \bar{b}_2\bar{\eta}^{lead}\chi - \bar{b}_2\bar{\eta}^{fo}(|\mathcal{I}| - \chi)] \Big\} dt\\
\bar{\alpha}^{fo}(t_0) = \bar{q}_{t_1} + &\int_{t_0}^{t_1} \Big\{ \bar{q} + \bar{r}(\bar{\eta}^{fo})^{2\bar{k}} - 2\bar{k}\bar{c}\bar{\eta}^{fo} + 2\bar{k}\bar{\epsilon}\bar{\eta}^{fo}[(|\mathcal{I}| - \chi - 1)\bar{\eta}^{fo} + \chi\bar{\eta}^{lead}]\\
&+ 2\bar{k}\bar{\alpha}^{fo}[\bar{b}_1 - \bar{b}_2\bar{\eta}^{lead}\chi - \bar{b}_2\bar{\eta}^{fo}(|\mathcal{I}| - \chi)] \Big\} dt.
\end{aligned}$$

The optimal number of leaders is given by

$$|\mathcal{I}_L| \in \arg\min_{\chi}[\chi\bar{\alpha}^{lead}(0) + (|\mathcal{I}| - \chi)\bar{\alpha}^{fo}(0)],$$

where $\bar{\alpha}$ depends on χ as well. We observe that the latter function is not necessarily monotone in $\chi = |\mathcal{I}_L|$. This means that increasing the number of leaders in the interaction does not necessarily improve the total performance of the system.

We numerically investigate $S(|\mathcal{I}_L|, \delta_{x_0})$ as a function of $\chi = |\mathcal{I}_L|$ for $|\mathcal{I}| = 6$. Let us consider a symmetric six-player game problem involving the parameters presented here:

$$\begin{aligned}
\bar{c}_i &= \bar{c} = 0, \forall i \in \mathcal{I}, & \bar{k}_i &= \bar{k} = 1, \forall i \in \mathcal{I},\\
\bar{\epsilon}_i &= \bar{\epsilon} = 1, \forall i \in \mathcal{I}, & b_{2i} &= b_2 = 0.1, \forall i \in \mathcal{I},\\
\bar{b}_{2i} &= \bar{b}_2 = 0.5, \forall i \in \mathcal{I}, & \bar{r}_i &= \bar{r} = 2, \forall i \in \mathcal{I},\\
\bar{q}_i &= \bar{q} = 1, \forall i \in \mathcal{I}, & \bar{q}_{iT} &= \bar{q}_T = 2, \forall i \in \mathcal{I},\\
T &= 0.1.
\end{aligned}$$

Table A.1: Summary of $\bar{\alpha}_0^{\text{leader}}$, $\bar{\alpha}_0^{\text{follower}}$, and $S(|\mathcal{I}_L|, \delta_{x_0})$ for the different number of leaders in the **homogeneous** scenario.

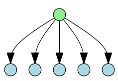
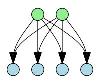
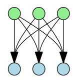
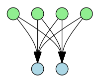
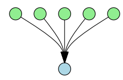
Leader(s)-Follower(s) Structure					
Individual leader cost	3.132	3.37	9.772	3.107	**2.968**
Individual follower cost	1.217	**0.2931**	0.3481	2.933	3.562
Total cost	9.219	**7.911**	30.36	18.29	18.4

Figure A.5 presents the evolution of both $\dot{\bar{\alpha}}^{\text{leader}}$ and $\dot{\bar{\alpha}}^{\text{follower}}$ for different number of leaders $|\mathcal{I}_L|$. Notice that, the initial values $\bar{\alpha}_0^{\text{leader}}$ and $\bar{\alpha}_0^{\text{follower}}$ determine the optimal cost considering that

$$\bar{x}_0 = \bar{x}(0) = \int ym(0, dy) = \int y\delta_{x_0}(dy) = x_0. \tag{A.25}$$

Figure A.5 and Table A.1 also show that, under the considered parameters, the lowest total cost is obtained when $|\mathcal{I}_L| = 2$, corresponding to a cost $S(|\mathcal{I}_L|, \delta_{x_0}) = 7.911$. These results offer insight into the structure-game design for the sake of either individual or total costs. We observe that having only a leader is suboptimal for the

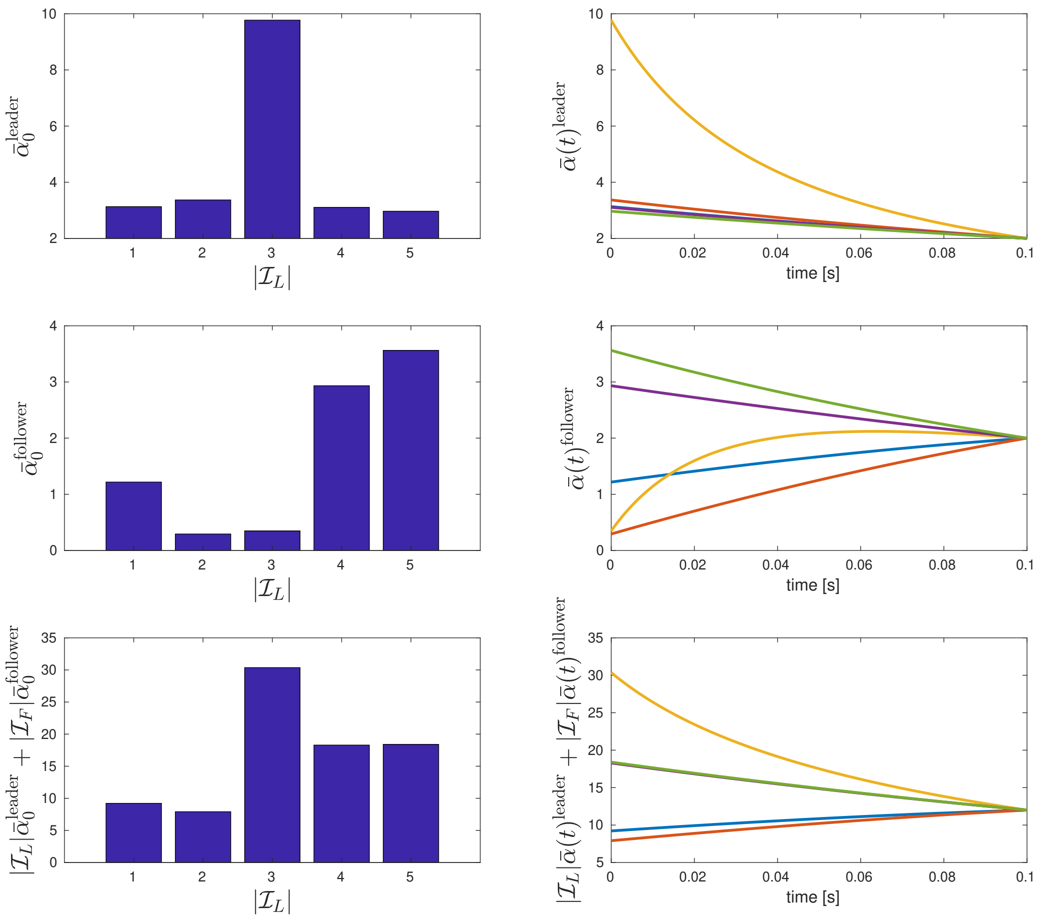


Fig. A.5: Evolution of the differential equations $\dot{\bar{\alpha}}^{\text{leader/follower}}$, and the corresponding initial values for different numbers of leaders in the **homogeneous** scenario.

total cost. Having too many leaders (the majority of the decision-makers as leaders) is not suboptimal for the total cost. In this setting, there is a tradeoff between leaders and followers so that the system's cost gets balanced.

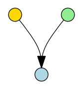
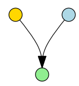
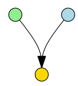
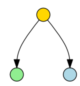
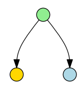
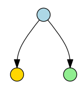
Uniform coupling and heterogeneous players

Now we investigate the two-layer case with uniform coupling, i.e., $\bar{\epsilon}_{ij} = 0.1$, for all combinations $i, j \in \mathcal{I}$ and for the heterogeneous case with $|\mathcal{I}| = 3$.
We consider the following parameters:

$$\begin{aligned}
b_{21} &= 0.1, & b_{22} &= 0.2, & b_{23} &= 0.3,\\
\bar{b}_{21} &= 0.5, & \bar{b}_{22} &= 0.6, & \bar{b}_{23} &= 0.7,\\
\bar{r}_1 &= 2, & \bar{r}_2 &= 2.1, & \bar{r}_3 &= 2.2,\\
\bar{q}_1 &= 1, & \bar{q}_2 &= 2, & \bar{q}_3 &= 3,\\
\bar{q}_{1T} &= 4, & \bar{q}_{2T} &= 6, & \bar{q}_{3T} &= 8,\\
\bar{b}_1 &= 2, & T &= 1, & \bar{k}_i &= \bar{k} = 1,\ \forall i \in \mathcal{I},
\end{aligned}$$

Figure A.6 shows the evolution of $\bar{\alpha}_1, \bar{\alpha}_2$, and $\bar{\alpha}_3$ for the different topologies presented in Table A.2. It can be seen in Figure A.7 that all the structures return a close value for the total cost. However, Table A.2 shows that the best topology is the last one where the third player acts as the unique leader assuming an initial condition such that (A.25) holds.

Table A.2: Summary of $\bar{\alpha}_0^{\text{leader}}$, $\bar{\alpha}_0^{\text{follower}}$, and $S(|\mathcal{I}_L|, \delta_{x_0})$ for the different number of leaders in the **heterogeneous** scenario.

Leader(s)-Follower(s) Structure						
Leaders	{1}{2}	{1}{3}	{2}{3}	{1}	{2}	{3}
Followers	{3}	{2}	{1}	{2}{3}	{1}{3}	{1}{2}
Total cost	17.14	16.96	16.99	17.04	17.13	**16.92**

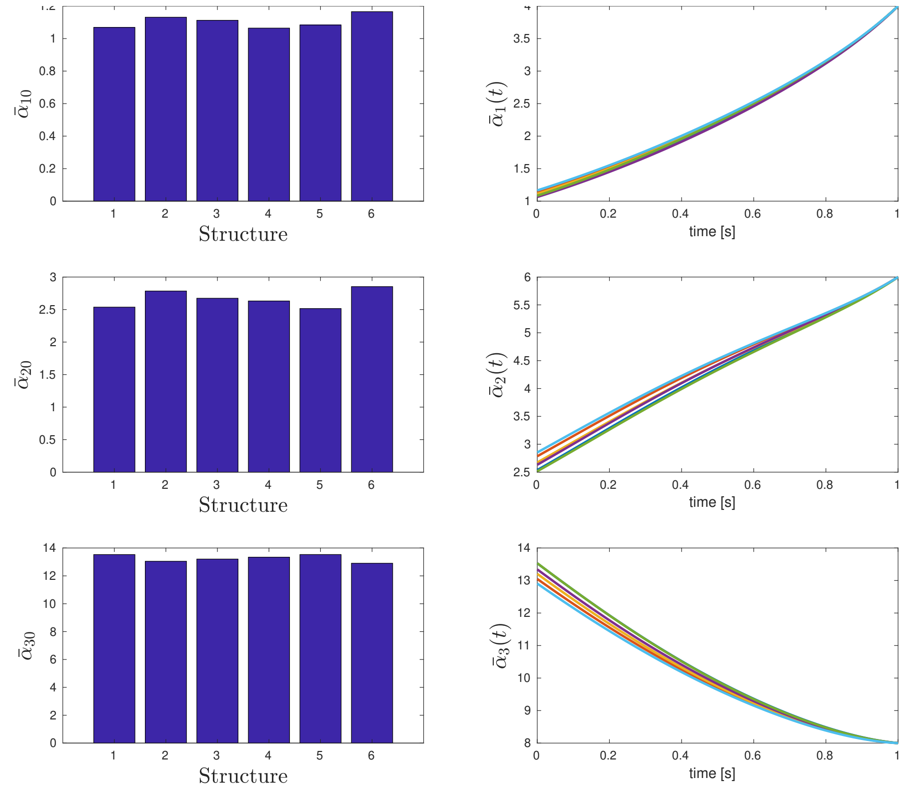


Fig. A.6: Evolution of the differential equations $\dot{\bar{\alpha}}^{\text{leader/follower}}$, and the corresponding initial values for different numbers of leaders in the **heterogeneous** scenario.

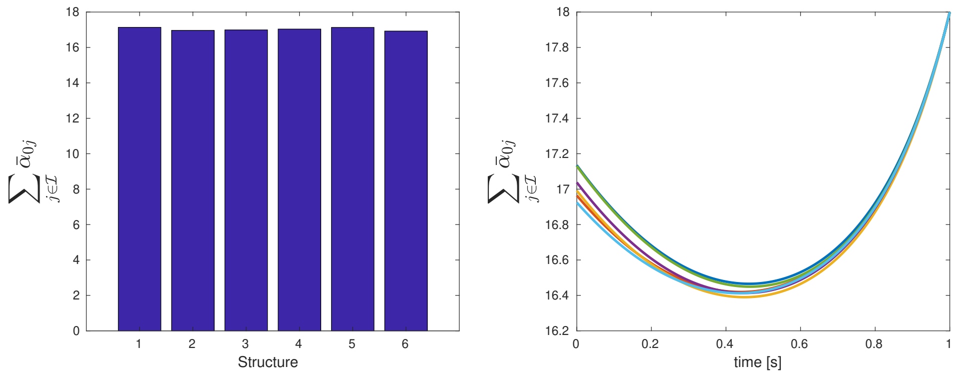


Fig. A.7: Evolution of the sum of differential equations and the corresponding total cost for the **heterogeneous** scenario.

6.2 Impact of the Hierarchical Structures

Here, we analyze the impact on the order of the strategic selection, i.e., the hierarchical order on the heterogeneous case with $|\mathcal{I}| = 3$. We consider the following heterogeneous parameters:

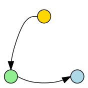
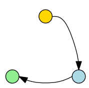
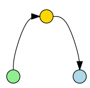
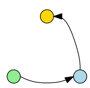
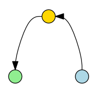
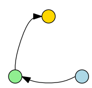
$$
\begin{aligned}
b_{21} &= 0.1, & b_{22} &= 0.2, & b_{23} &= 0.3,\\
\bar{b}_{21} &= 0.4, & \bar{b}_{22} &= 0.5, & \bar{b}_{23} &= 0.6,\\
\bar{r}_1 &= 1, & \bar{r}_2 &= 2, & \bar{r}_3 &= 3,\\
\bar{q}_1 &= 1.1, & \bar{q}_2 &= 1.2, & \bar{q}_3 &= 1.3,\\
\bar{q}_{1T} &= 2.1, & \bar{q}_{2T} &= 2.2, & \bar{q}_{3T} &= 2.3,\\
\bar{b}_1 &= 2, & T &= 0.1, & \bar{k}_i &= \bar{k} = 1, \ \forall i \in \mathcal{I},
\end{aligned}
$$

and

$$
\bar{\epsilon} = \begin{pmatrix} 1 & 1.2 & 1.1 \\ 1.5 & 1 & 1.6 \\ 1.3 & 1.4 & 1 \end{pmatrix}.
$$

Table A.3 shows the summary of the total costs for the six different possible hierarchical orders assuming an initial condition such that (A.25) holds. It can be seen that the third configuration is the best to minimize the total cost. Moreover, Figure A.8 presents the evolution of the equations $\sum_{j\in\mathcal{I}} \dot{\alpha}_j(t)$ for all the possible structures.

Table A.3: Total cost for the different hierarchical orders in a three-player case in the **heterogeneous** scenario.

Hierarchical Structure						
Combination label	1	2	3	4	5	6
Hierarchical order	{1}{2}{3}	{1}{3}{2}	{2}{1}{3}	{2}{3}{1}	{3}{1}{2}	{3}{2}{1}
Total cost	6.124	7.464	**5.864**	8.757	6.894	8.433

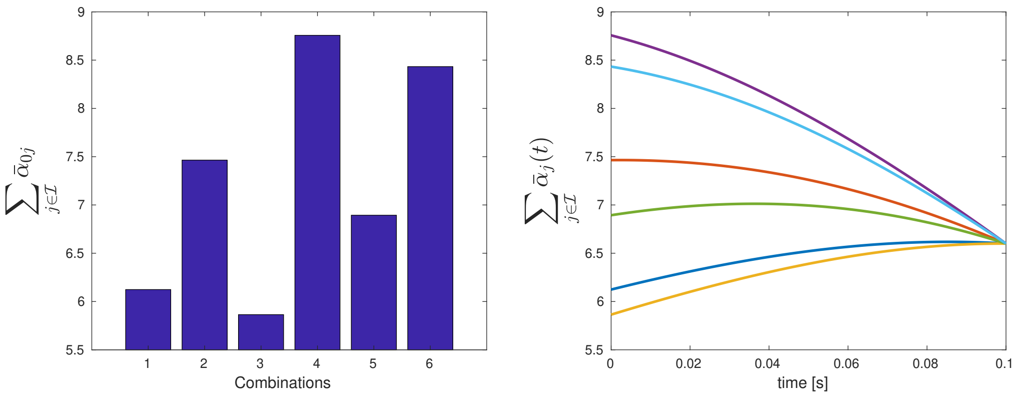


Fig. A.8: Evolution of the differential equations $\sum_{j\in\mathcal{I}} \dot{\bar{\alpha}}_j(t)$, and the corresponding initial values for different hierarchical structures in the **heterogeneous** scenario.

7 Conclusion

We have examined multi-layer hierarchical mean-field-type games with non-quadratic polynomial costs. We derived hierarchical mean-field-type solutions in the linear state-and-mean-field feedback form by using a partial integro-differential system, and we have established the relationship between the Nash and the hierarchical solutions.

Furthermore, we have studied the impact of the number of leaders on a bi-level Stackelberg problem for both symmetric and non-symmetric scenarios. In addition, we have shown that the number of layers, permutations of the decision-makers per layer, and their identity affect significantly the total cost of the system. We have also shown numerically that the ordering among all permutations of heterogenous decision-makers may reduce the cost by a significant proportion depending on the horizon.

References

Averboukh, A. Y. (2018). Stackelberg solution for first-order mean-field game with a major player. *Izv. Inst. Mat. Inform. Udmurt.*, *52*.

Bagchi, A., & Basar, T. (1981). Stackelberg strategies in linear-quadratic stochastic differential games. *Journal of Optimization Theory and Applications*, *35*(3), 443-464.

Barreiro-Gomez, J., Duncan, T. E., & Tembine, H. (2019). Linear-quadratic mean-field-type games: Jump-diffusion process with regime switching. *IEEE Transactions on Automatic Control*, *64*(10), 4329-4336.

Barreiro-Gomez, J., Ocampo-Martinez, C., & Quijano, N. (2017). Partitioning for large-scale systems: A sequential distributed MPC design. In *20th ifac world congress* (p. 8838-8843). Toulouse, France.

Basar, T., & Selbuz, H. (1979). Closed-loop stackelberg strategies with applications in the optimal control of multilevel systems. *IEEE Transactions on Automatic Control*, *24*(2), 166–179.

Beardsley, X. W., Field, B., & Xiao, M. (2012). Mean-variance-skewness-kurtosis portfolio optimization with return and liquidity. *Communications in Mathematical Finance*, *1*(1), 13–49.

Bensoussan, A., Chau, M., Lai, Y., & Yam, S. (2017). Linear-quadratic mean field Stackelberg games with state and control delays. *SIAM Journal on Control and Optimization*, *55*(4), 2748-2781.

Bensoussan, A., Chau, M. H. M., & Yam, S. C. P. (2015). Mean-field Stackelberg games: Aggregation of delayed instructions. *SIAM J. Control Optim.*, *53*(4), 2237-2266.

Bensoussan, A., Chen, S., & 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.

Bensoussan, A., Djehiche, B., Tembine, H., & Yam, S. C. P. (2020). Mean-field-type games with jump and regime switching. *Dynamic Games and Applications https://doi.org/10.1007/s13235-019-00306-2*.

Cai, H., & Hu, G. (2017). Distributed tracking control of an interconnected leader-follower multiagent system. *IEEE Transactions on Automatic Control*, *62*(7), 3494-3501.

Cruz, J. (1978). Leader-follower strategies for multilevel systems. *IEEE Transactions on Automatic Control*, *23*(2), 244–255.

Du, K., & Wu, Z. (2019). Linear-quadratic Stackelberg game for mean-field backward stochastic differential system and application. *Mathematical Problems in Engineering*, *Article ID 1798585, 17 pages.*

Gardner, B., & Cruz, J. (1978). Feedback stackelberg strategy for m-level hierarchical games. *IEEE Transactions on Automatic Control*, *23*(3), 489–491.

Li, Y., Shi, D., & Chen, T. (2018). False data injection attacks on networked control systems: A Stackelberg game analysis. *IEEE Transactions on Automatic Control*, *63*(10), 3503-3509.

Lin, Y., Jiang, X., & Zhang, W. (2019). An open-loop Stackelberg strategy for the linear quadratic mean-field stochastic differential game. *IEEE Transactions on Automatic Control*, *64*(1), 97-110.

Moon, J., & Basar, T. (2015). Linear-quadratic stochastic differential Stackelberg games with a high population of followers. *in Proc. 54th IEEE Conf. Decision Control, Osaka, Japan*, 2270-2275.

Moon, J., & Basar, T. (2018). Linear quadratic mean field stackelberg differential games. *Automatica*, *97*(2018), 200-213.

Nourian, M., Caines, P., Malhamé, R. P., & Huang, M. (2012). Mean field LQG control in leader-follower stochastic multi-agent systems: Likelihood ratio based adaptation. *IEEE Transactions on Automatic Control*, *57*(11), 2801-2816.

Pan, L., & Yong, J. (1991). A differential game with multi-level of hierarchy. *Journal of mathematical analysis and applications*, *161*(2), 522–544.

Shi, J., Wang, G., & Xiong, J. (2016). Leader-follower stochastic differential game with asymmetric information and applications. *Automatica*, *63*(2016), 60-73.

Simaan, M., & Cruz, J. (1973b). A stackelberg solution for games with many players. *IEEE Transactions on Automatic Control*, *18*(3), 322–324.

Simaan, M., & Cruz, J. B. (1973a). On the stackelberg strategy in nonzero-sum games. *Journal of Optimization Theory and Applications*, *11*(5), 533–555.

Stackelberg, H. V. (1948). The theory of the market economy, trans. by aj peacock, london, william hodge. *Originally published as Grundlagen der Theoretischen Volkswirtschaftlehre*.

Sun, J., & Yong, J. (2014). Linear quadratic stochastic differential games: Open-loop and closed-loop saddle points. *SIAM Journal on Control and Optimization*, *52*(6), 4082-4121. Retrieved from https://doi.org/10.1137/140953642 doi: 10.1137/140953642

Sutter, M., & Rivas, M. F. (2014). Leadership, reward and punishment in sequential public goods experiments. In P. A. V. Lange, B. Rockenbach, & T. Yamagishi (Eds.), *Reward and punishment in social dilemmas* (p. 1-39). Oxford, Reino Unido: Oxford University Press.

Theodossiou, P., & Savva, C. S. (2016). Skewness and the relation between risk and return. *Management Science*, *62*(6), 1598-1609.

Paper B

Hierarchical Mean-Field Type Control of Price Dynamics for Electricity in Smart Grid

Zahrate El Oula Frihi[1], Salah Eddine Choutri [2], Julian Barreiro-Gomez [3] and Hamidou Tembine[4]

The paper has been published in
Journal of Systems Science and Complexity, Vol. 35, pp.1–17, 2022.
https://doi.org/10.1007/s11424-021-0176-3

Abstract

This paper solves a mean-field type hierarchical optimal control problem in the electricity market. We consider $n-1$ prosumers and one producer. The ith prosumer, for $1 < i < n$, is a leader of the $(i-1)$th prosumer and is a follower of the $(i+1)$th prosumer. The first player (agent) is the follower at the bottom whereas the nth is the leader at the top. The problem is described by a linear jump-diffusion system of conditional mean-field type, where the conditioning is with respect to common noise, and a quadratic cost functional involving the second moment, the square of the conditional expectation of the controls of the agents. We provide a semi-explicit solution to the corresponding mean-field-type hierarchical control problem with common noise. Finally, we illustrate the obtained result via a numerical example with two different scenarios.

1 Introduction

Leader-follower games were first introduced by Stackelberg (Stackelberg, 1948) in 1934, to model markets where some firms have a stronger influence on others. Stackelberg games are non-zero-sum static games with a two-level hierarchy as they consist of two players, a major player (the leader) and a minor player (the follower). The minor player chooses a response strategy (assumed rational), for any announced strategy from the major player, such that her own performance criterion is optimized. The major player predicts the best response of the minor player and chooses a strategy to optimize her performance criterion (assuming that she knows the performance criterion of the minor player). A dynamic LQ Stackelberg differential game was studied by samaan and cruz (Simaan & Cruz, 1973a). The stochastic LQ Stackelberg differential game was investigated by Bagchi and Basar in (Bagchi & Basar, 1981).

For $n > 2$ the leader-follower problem is called multi-hierarchical, each player is a leader for the previous one and a follower of the next player in the hierarchy. The first and the last players are the leader at the top and the follower at the bottom, respectively. For multi-hierarchical differential games, see e.g. (Pan & Yong, 1991; Li & Yu, 2018; Simaan & Cruz, 1973b; Cruz, 1978; Gardner & Cruz, 1978; Basar & Selbuz, 1979).

In the present paper we study a stochastic mean-field type Cournot hierarchical control problem of $n-1$ electricity prosumers (followers-leaders at different levels) and one electricity producer (leader at the top). The considered electricity price model is the same as in (Djehiche, Barreiro-Gomez, & Tembine, 2020), where the authors formulated a stochastic mean-field-type dynamic Cournot game between elec-

tricity producers, based on similar deterministic models (e.g. (Evans, 1922; Ross, 1925)). We refer the reader to (B. F. Hobbs & Pang, 2007; Metzler, Hobbs, & Pang, 2003; B. E. Hobbs, 2001; Willems, 2002; Hogan, 1997), for research involving Cournot-based models for the electricity market. There are other works developing an application of mean field control involving price dynamics. For instance, in (Wanga & Huang, 2019), mean field control is designed to optimize over dynamic production considering sticky price and risk-free profit.
Our problem is described by a linear jump-diffusion system of conditional mean-field type and a quadratic cost functional. The goal of this work is to, first, find the optimal solutions to the mean-field type hierarchical control problem and, second, to investigate the effect of the hierarchy on the solutions. We summarize our contribution as follows. We formulate a mean-field-type hierarchical control problem with common noise and jump-diffusion. We provide semi-explicit solution with a structure in state-and-mean-field feedback form, using a direct method consisting of square completion technique. Finally, we present a numerical example with 3 agents under two different scenarios (a homogeneous case as well as a heterogeneous case), to validate our theoretical results and investigate the effect of the hierarchy on the optimal solutions and the revenues for each agent. The rest of this paper is organized as follows. The next section introduces the setup for the model and the key quantities. In section 3 we present the hierarchical mean-field type problem. Section 4 presents the main results. In section 5, we present a numerical example with different scenarios. Section 6 concludes the paper.

2 The Setup

We consider a hierarchical mean-field type control problem described by the following settings. Let $\mathcal{T} := [0, T]$ be the time horizon with $T > 0$. The energy market in our model is described by $n \geq 2$ agents, one producer, and $n-1$ prosumers following a hierarchical structure. Each agent $i \in \{1, \dots, n\}$, at time $t \in \mathcal{T}$, has an output $u_i(t) \geq 0$. The log-price dynamics, as modeled in (Djehiche et al., 2020), are given by $p(0) = p_0$ and

$$dp(t) = s[a - D(t) - p(t)]dt + \left(\sigma dB(t) + \int_{\theta\in\Theta} \mu(\theta)\widetilde{N}(dt, d\theta)\right) + \sigma_0 dB_0(t), \quad \text{(B.1)}$$

where

$$D(t) := \sum_{i=1}^{n} u_i(t),$$

is the supply at time t and

$$\widetilde{N}(dt, d\theta) = N(dt, d\theta) - v(d\theta)dt,$$

is a compensated martingale and v is assumed to be a radon measure over Θ. The process N is a jump process with Levy measure $v(d\theta)$ such that

$$\int_{\Theta} \mu^2(\theta)v(d\theta) < +\infty.$$

The processes B and B_0 are standard Brownian motions representing, respectively, a local and a global noise in the model. It is assumed that all these processes are mutually independent and the global noise is observed by all agents. Denote by $\mathcal{F}_t^{B_0}$ the filtration generated by the observed common noise up to time t. The number s is positive and the quantities a, σ, σ_0 are fixed parameters. We assume that initial distribution of p_0 is square integrable.
The conditional price $\bar{p}(t) := E[p(t)|\mathcal{F}_t^{B_0}],\ 0 \leq t \leq \mathcal{T})$, which is driven by the common noise B_0, solves the following stochastic differential equation

$$\begin{aligned} d\bar{p}(t) &= s[a - \bar{D}(t) - \bar{p}(t)]dt + \sigma_0 dB_0(t), \\ \bar{p}(0) &= \bar{p}_0, \end{aligned}$$

where

$$\bar{D}(t) := \sum_{i=1}^{n} \bar{u}_i(t).$$

At time $t \in \mathcal{T}$, agent i gains $\bar{p}(t)u_i(t) - C_i(u_i(t))$ where $C_i : \mathbb{R} \to \mathbb{R}$, represents her instant cost given by

$$C_i(u_i) = c_i u_i(t) + \frac{r_i u_i^2(t)}{2} + \frac{\bar{r}_i \bar{u}_i^2(t)}{2}.$$

The term $\bar{u}_i(t) = E[u_i(t)|\mathcal{F}_t^{B_0}]$ is the conditional expectation of agent i's output given the common noise B_0 (the global uncertainty). The payoff functional (or the long-term revenue) of each agent i is

$$\mathcal{R}_i(p_0, u_1(t), \ldots, u_n(t)) = -\frac{q}{2}e^{-\lambda_i T}(p(T) - \bar{p}(T))^2 + \int_0^T e^{-\lambda_i t}[\bar{p}(t)u_i(t) - C_i(u_i(t))]dt,$$

where $c_i, r_i, \bar{r}_i$ and q are non-negative parameters and λ_i is a discount factor for the agent i.
Note that all agents are coupled through the price functional. Furthermore, the payoff functional is of mean-field type since it involves two conditional mean-field terms: $\bar{p}(t)$ and $\bar{u}_i^2(t)$ based on the observations of the common noise B_0 up to t. For ease of notation, we dropped the dependence on t for: $p, \bar{p}, u, \bar{u}, D, \bar{D}, B$ and B_0.

3 Hierarchical mean-field type control problem

The hierarchical control problem we consider consists of the following leader-follower order: given $n-1$ prosumers and one producer, the ith prosumer is a leader of the $(i-1)$th prosumer and is a follower of the $(i+1)$th prosumer, for $1 < i < n$. When considering only two levels in the strategic interaction, i.e., a unique producer and a unique prosumer, the hierarchical mean-field-type control problem becomes a Stackelberg problem as in (Li & Yu, 2018). Here, we address the solution for n levels including the bi-level Stackelberg problem. The nth agent (the producer) is the leader at the top level and the 1st prosumer is the follower at the bottom level. More precisely, denote

$$\mathcal{U}_i = L^2_{\mathcal{F}^{p_0,p,B_0}}([0,T]\times\mathbb{R},\mathbb{R})$$

the set of square integrable, $\mathcal{F}^{p_0,p,B_0}$-progressively measurable feedback controls, all the agents aim to maximize their revenues as follows:

For given $u_i \in \mathcal{U}_i,\ i\in\{2,\dots,n\}$, Prosumer 1 aims to find her optimal control

$$u_1^* = u_1^*(u_2,\dots,u_n)\in\mathcal{U}_1,$$

such that

$$E[\mathcal{R}_1(p_0,u_1^*,\dots,u_n)] = \sup_{u_1\in\mathcal{U}_1} E[\mathcal{R}_1(p_0,u_1,\dots,u_n)],$$

under the dynamics in (B.1). Then, given the optimal control of the 1st prosumer, Prosumer 2 finds her optimal control

$$u_2^* = u_2^*(u_3,\dots,u_n)\in\mathcal{U}_2,$$

such that

$$E[\mathcal{R}_2(p_0,u_1^*,u_2^*,\dots,u_n)] = \sup_{u_2\in\mathcal{U}_2} E[\mathcal{R}_2(p_0,u_1^*(u_2,\dots,u_n),u_2,\dots,u_n)],$$

under the log-price dynamics

$$dp = s[a - D^* - p]dt + \left(\sigma dB + \int_{\theta\in\Theta}\mu(\theta)\widetilde{N}(dt,d\theta)\right) + \sigma_0 dB_0, \tag{B.2}$$

where

$$D^*(t) := u_1^*(t) + \sum_{i=2}^n u_i(t).$$

Thereby, we can define $u_1^*, u_2^*, \dots, u_n^*$ inductively and the producer's optimal control u_n^* will not depend on the controls of the $n-1$ prosumers.

The same model remarks discussed in (Djehiche et al., 2020) are applicable to the price dynamics considered in this hierarchical game problem as presented next.

- This log-price dynamics model can be re-interpreted as an error to the standard inverse demand model.
- The constant s allows us to navigate between several regimes.
- The jump term $\tilde{N}$ captures some of the big changes in the market that may happen randomly, e.g., regulation law change.
- The global uncertainty B_o captures common noise in the market, for example, weather conditions and temperature field in specific season.
- The conditional log-price is calculated based on the common noise that is observed. The revenue is computed from the conditional log-price.
- This revenue model is similar to the one considered by Jovanovic (see page 652 in (Jovanovic, 1982)) who studied discrete-time mean-field games for selection and evolution industry. Therein the conditional state appears as well. However, (Jovanovic, 1982) considered that firms are too small to affect the log-price. Here, each of the n firms can influence the price and cannot be neglected. Global uncertainty was not considered in (Jovanovic, 1982).

4 Main result

In this section, we present the main result of the paper in the form of the following proposition.

Proposition 4.1. *The optimal controls for the n agents are in state-and-conditional mean-field feedback form:*

$$u_i^*(t) = -\frac{s((p(t)-\bar{p}(t))\alpha_i(t)+\xi_i(t)}{e^{-\lambda_i t} r_i} + \frac{e^{-\lambda_i t}(\bar{p}(t)-c_i) - \beta_i(t)s\bar{p}(t) - s\gamma_i(t)}{e^{-\lambda_i t}(r_i+\bar{r}_i)}.$$

The conditional optimal price:

$$\begin{aligned} d\bar{p}(t) =& s\left[a - \sum_{i=1}^{n}\left(\frac{e^{-\lambda_i t}(\bar{p}(t)-c_i) - \beta_i(t)s\bar{p}(t) - s\gamma_i(t)}{e^{-\lambda_i t}(r_i+\bar{r}_i)}\right) - \bar{p}(t)\right]dt + \sigma_0 dB_0(t), \\ \bar{p}(0) =& \bar{p}_0. \end{aligned}$$

The stochastic functions α_i, β_i, γ_i, δ_i and ξ_i are $\mathcal{F}_t^{B_0}-$measurable and solve the following stochastic Riccati system: for $1 \le i \le n$,

$$
\begin{aligned}
d\alpha_i &= \left(2s\alpha_i - \frac{s^2\alpha_i^2}{e^{-\lambda_i t}r_i} - 2s^2\alpha_i \sum_{j\neq i} \frac{\alpha_j}{e^{-\lambda_j t}r_j}\right) dt + \alpha_{i,0} dB_0, \\
\alpha_i(T) &= -qe^{-\lambda_i T}, \\
d\beta_i &= \left(-\frac{(e^{-\lambda_i t} - \beta_i s)^2}{e^{-\lambda_i t}(r_i + \bar{r}_i)} + 2s\beta_i \sum_{j\neq i} \frac{(e^{-\lambda_j t} - \beta_j s)}{e^{-\lambda_j t}(r_j + \bar{r}_j)} + 2\beta_i s\right) dt + \beta_{i,0} dB_0, \\
\beta_i(T) &= 0, \\
d\gamma_i &= -\left(\beta_i s(a + \sum_{j\neq i} \frac{e^{-\lambda_j t}c_j + s\gamma_j}{e^{-\lambda_j t}(r_j + \bar{r}_j)}) + \sigma_0\beta_{i,0} - s\gamma_i \sum_{j\neq i} \frac{(e^{-\lambda_j t} - s\beta_j)}{e^{-\lambda_j t}(r_j + \bar{r}_j)}\right. \\
&\quad \left. + \sigma_0\beta_{i,0} - s\gamma_i - \frac{(e^{-\lambda_i t} - \beta_i s)(s\gamma_i + e^{-\lambda_i t}c_i)}{(r_i + \bar{r}_i)e^{-\lambda_i t}} \right) dt - \beta_i\sigma_0 dB_0, \\
\gamma_i(T) &= 0, \\
d\xi_i &= s\left(\xi_i - s\alpha_i \sum_{j\neq i} \frac{\xi_j}{e^{-\lambda_j t}r_j} - \frac{s\alpha_i\xi_i}{e^{-\lambda_i t}r_i} - s\xi_i \sum_{j\neq i} \frac{\alpha_j}{e^{-\lambda_j t}r_j}\right) dt + \xi_{i,0} dB_0, \\
\xi_i(T) &= 0, \\
d\delta_i &= -\left(\frac{1}{2}\frac{(s\gamma_i + e^{-\lambda_i t}c_i)^2}{e^{-\lambda_i t}(r_i + \bar{r}_i)} + s^2\xi_i \sum_{j\neq i} \frac{\xi_j}{e^{-\lambda_j t}r_j} + \frac{1}{2}\frac{s^2\xi_i^2}{e^{-\lambda_i t}r_i} + \frac{1}{2}\beta_i\sigma_0^2\right. \\
&\quad \left. + \frac{\alpha_i}{2}(\sigma^2 + \int_\Theta \mu^2(\theta)v(d\theta)) + s\gamma_i a + s\gamma_i \sum_{j\neq i} \frac{e^{-\lambda_j t}c_j + s\gamma_j}{e^{-\lambda_j t}(r_j + \bar{r}_j)} + \gamma_{i,0}\sigma_0 \right) dt \\
&\quad - \sigma_0\gamma_i dB_0, \\
\delta_i(T) &= 0.
\end{aligned}
$$

□

Remark 4.1. *The optimal control input u_i^* is independent from the Riccati equation δ_i, for all i. Nevertheless, the differential equation δ_i affects the optimal revenue, which is given by the initial value of the ansatz, i.e., for all i*

$$
\begin{aligned}
E[\mathcal{R}_i^* - f_i(0, p(0)|\mathcal{F}_T^{B_0})] &= E[\mathcal{R}_i^* - \frac{1}{2}\alpha_i(0)(p_0 - \bar{p}_0)^2 - \frac{1}{2}\beta_i(0)\bar{p}_0^2 \\
&\quad - \xi_i(0)(p_0 - \bar{p}_0) - \gamma_i(0)\bar{p}_0 - \delta_i(0)|\mathcal{F}_T^{B_0})] \\
&= 0.
\end{aligned}
$$

Proof. We prove it by a backward inductive argument described in the first part 3.2 of the thesis. We start with Prosumer 1 and we use a direct method 6.2 consisting of the following guess functional for the quadratic discounted revenue functional,

$$f_1(t,p) = \frac{1}{2}\alpha_1(t)(p-\bar{p})^2 + \frac{1}{2}\beta_1(t)\bar{p}^2 + \xi_1(t)(p-\bar{p}) + \gamma_1(t)\bar{p} + \delta_1(t), \tag{B.3}$$

where $\alpha_1, \beta_1, \gamma_1, \delta_1$ and ξ_1 are random functions of time t, that are $\mathcal{F}_t^{B_0}-$measurable such that

$$f_1(T, p(T)) = -\frac{q}{2}e^{-\lambda_1 T}\alpha_1(T)(p(T)-\bar{p}(T))^2.$$

We derive a formula for $df_1(t,p)$ using Itô's formula for the jump-diffusion process. We have

$$\mathrm{d}\bar{p}^2 = \left(2s\bar{p}(a-\bar{D}-\bar{p}) + \sigma_o^2 s\right)\mathrm{d}t + 2\bar{p}\sigma_o \mathrm{d}B_o, \tag{B.4}$$

$$\mathrm{d}\left[\frac{\beta_1\bar{p}^2}{2}\right] = \frac{\bar{p}^2\mathrm{d}\beta_1}{2} + \frac{\beta_1\mathrm{d}[\bar{p}^2]}{2} + \beta_{1,o}\bar{p}\sigma_o\mathrm{d}t, \tag{B.5}$$

therefore, replacing (B.4) in (B.5) yields

$$\mathrm{d}\left[\frac{\beta_1\bar{p}^2}{2}\right] = \frac{1}{2}\bar{p}^2\mathrm{d}\beta_1 + \frac{1}{2}\beta_1\left(2s\bar{p}\left(a-\bar{D}-\bar{p}\right) + \sigma_o^2\right)\mathrm{d}t + \beta_i\bar{p}\sigma_o\mathrm{d}B_o + \beta_{1,o}\bar{p}\sigma_o\mathrm{d}t.$$

We compute the difference between p and $\bar{p}$.

$$\mathrm{d}\left[p-\bar{p}\right] = -s\left(D-\bar{D}+p-\bar{p}\right)\mathrm{d}t + \left(\sigma\mathrm{d}B + \int_\Theta \mu(\theta)\tilde{N}(\mathrm{d}t,\mathrm{d}\theta)\right).$$

Then,

$$\mathrm{d}[\xi_1(p-\bar{p})] = \xi_1\left(-s(D-\bar{D}+p-\bar{p})dt + (\sigma dB + \int_\Theta \mu(\theta)\widetilde{N}(dt,d\theta))\right) + (p-\bar{p})d\xi_1,$$

and

$$\begin{aligned}\mathrm{d}\left[p-\bar{p}\right]^2 = {} & 2(p-\bar{p})\left(\sigma\mathrm{d}B + \int_\Theta \mu(\theta)\tilde{N}(\mathrm{d}t,\mathrm{d}\theta)\right) - 2s(p-\bar{p})(D-\bar{D}+p-\bar{p})\mathrm{d}t \\ & + \left(\sigma^2 + \int_\Theta \mu^2(\theta)\nu(\mathrm{d}\theta)\right)\mathrm{d}t.\end{aligned} \tag{B.6}$$

Moreover,

$$\mathrm{d}\left[\frac{\alpha_1(p-\bar{p})^2}{2}\right] = \frac{(p-\bar{p})^2}{2}\mathrm{d}\alpha_1 + \frac{1}{2}\alpha_1\mathrm{d}\left[(p-\bar{p})^2\right] + 0, \tag{B.7}$$

and replacing (B.6) in (B.7) yields

$$\begin{aligned}
\mathrm{d}\left[\frac{\alpha_1(p-\bar{p})^2}{2}\right] = &\frac{(p-\bar{p})^2}{2}\mathrm{d}\alpha_1 + \frac{1}{2}\alpha_1\left(\sigma^2 + \int_\Theta \mu^2(\theta)\nu(\mathrm{d}\theta)\right)\mathrm{d}t \\
&- s(p-\bar{p})\alpha_1(D-\bar{D}+p-\bar{p})\mathrm{d}t + (p-\bar{p})\alpha_1\sigma\mathrm{d}B \\
&+ (p-\bar{p})\alpha_1\int_\Theta \mu(\theta)\tilde{N}(\mathrm{d}t,\mathrm{d}\theta).
\end{aligned}$$

Finally,

$$\mathrm{d}[\gamma_1\bar{p}] = \bar{p}\mathrm{d}\gamma_1 + \left(s\gamma_1\left(a-\bar{D}-\bar{p}\right) + \gamma_{1,o}\sigma_o\right)\mathrm{d}t + \sigma_o\gamma_1\mathrm{d}B_o.$$

Thus,

$$\begin{aligned}
df_1(t,p) = &\frac{(p-\bar{p})^2 d\alpha_1}{2} + \frac{\alpha_1}{2}\left(\sigma^2 + \int_\Theta \mu^2(\theta)v(d\theta)\right)dt \\
&- (s(p-\bar{p})\alpha_1 + s\xi_1)(D-\bar{D}+p-\bar{p})dt + \frac{1}{2}\bar{p}^2 d\beta_1 \\
&+ \frac{1}{2}\beta_1(2s\bar{p}(a-\bar{D}-\bar{p}) + \sigma_0^2)dt + \beta_1\bar{p}\sigma_0 dB_0 + \beta_{1,0}\bar{p}\sigma_0 dt \\
&+ \bar{p}d\gamma_1 + \left(s\alpha_1(a-\bar{D}-\bar{p}) + \gamma_{1,0}\sigma_0\right)dt + \sigma_0\gamma_1 dB_0 + d\delta_1 \\
&+ ((p-\bar{p})\alpha_1 + \xi_1)(\sigma dB + \int_\Theta \mu(\theta)\widetilde{N}(dt,d\theta)) + (p-\bar{p})d\xi_1.
\end{aligned}$$

Note that the expected revenue can be written as

$$\begin{aligned}
e^{-\lambda_1 t}E[\bar{p}u_1 - C_1(u_1)|\mathcal{F}_T^{B_0}] = &e^{-\lambda_1 t}\ E[(\bar{p}-c_1)\bar{u}_1|\mathcal{F}_T^{B_0}] \\
&- e^{-\lambda_1 t}\ E[\frac{1}{2}r_1(u_1-\bar{u}_1)^2 - \frac{1}{2}(r_1+\bar{r}_1)\bar{u}_1^2|\mathcal{F}_T^{B_0}].
\end{aligned}$$

Next, we integrate $df_1(t,p)$ from 0 to T, then take the conditional expectation with respect to the filtration $\mathcal{F}_T^{B_0}$ and we, finally, express the difference between the revenue and the guess functional evaluated at 0.

$$
\begin{aligned}
E[\mathcal{R}_1 - f_1(0,p(0))|\mathcal{F}_T^{B_0}] = &\int_0^T e^{-\lambda_1 t} E[(\bar{p}-c_1)\bar{u}_1 \\
&- \frac{1}{2} r_1(u_1-\bar{u}_1)^2 - \frac{1}{2}(r_1+\bar{r}_1)\bar{u}_1^2|\mathcal{F}_T^{B_0}]dt \\
&+ \frac{\alpha_1(T) - qe^{-\lambda_1 T}}{2} E[(p(T)-\bar{p}(T))^2|\mathcal{F}_T^{B_0}] \\
&+ E[\int_0^T \frac{(p-\bar{p})^2 d\alpha_1}{2}|\mathcal{F}_T^{B_0}] \\
&+ E[\int_0^T \frac{\alpha_1}{2}\left(\sigma^2 + \int_\Theta \mu^2(\theta)v(d\theta)\right) dt|\mathcal{F}_T^{B_0}] \\
&- E[\int_0^T (s(p-\bar{p})\alpha_1 + s\xi_1)(D-\bar{D}+p-\bar{p})dt|\mathcal{F}_T^{B_0}] \\
&+ \frac{1}{2}E[\int_0^T \beta_1(2s\bar{p}(a-\bar{D}-\bar{p}) + \sigma_0^2)dt|\mathcal{F}_T^{B_0}] \\
&+ E[\int_0^T \beta_1\bar{p}\sigma_0 dB_0|\mathcal{F}_T^{B_0}] + E[\int_0^T \beta_{1,0}\bar{p}\sigma_0 dt|\mathcal{F}_T^{B_0}] \\
&+ E[\int_0^T \bar{p}d\gamma_1|\mathcal{F}_T^{B_0}] + E[\int_0^T \sigma_0\gamma_1 dB_0|\mathcal{F}_T^{B_0}] \\
&+ E[\int_0^T d\delta_1|\mathcal{F}_T^{B_0}] + E[\int_0^T (p-\bar{p})d\xi_1|\mathcal{F}_T^{B_0}] \\
&+ E[\int_0^T \left(s\gamma_1(a-\bar{D}-\bar{p}) + \gamma_{1,0}\sigma_0\right) dt|\mathcal{F}_T^{B_0}] \\
&+ \frac{1}{2}E[\int_0^T \bar{p}^2 d\beta_1|\mathcal{F}_T^{B_0}].
\end{aligned}
$$

Now, we rearrange and complete the square for the terms

$$
\begin{aligned}
e^{-\lambda_1 t}\ (\ 2(\bar{p}-c_1)\bar{u}_1 - r_1(u_1-\bar{u}_1)^2 - (r_1+\bar{r}_1)\bar{u}_1^2\ ) \\
-2s((p-\bar{p})\alpha_1 + \xi_1)(D-\bar{D}) - 2\beta_1 s\bar{p}\bar{D} - 2s\gamma_1\bar{D}
\end{aligned}
$$

where

$$
\bar{D} = \sum_{j=1}^n \bar{u_j} \quad \text{and} \quad D = \sum_{j=1}^n u_j.
$$

We have

$$e^{-\lambda_1 t}\ \big(\ 2(\bar{p}-c_1)\bar{u}_1 - r_1(u_1-\bar{u}_1)^2 - (r_1+\bar{r}_1)\bar{u}_1^2\ \big) - 2s((p-\bar{p})\alpha_1+\xi_1)(u_1-\bar{u}_1)$$
$$-2s((p-\bar{p})\alpha_1+\xi_1)\sum_{j=2}^{n}(u_j-\bar{u}_j) - 2\beta_1 s\bar{p}\ \bar{u}_1 - 2s\gamma_1\bar{u}_1 - 2s\beta_1\bar{p}\sum_{j=2}^{n}\bar{u}_j$$
$$-2s\gamma_1\sum_{j=2}^{n}\bar{u}_j$$

is equal to

$$\begin{aligned}
&-(r_1+\bar{r}_1)e^{-\lambda_1 t}\left(\bar{u}_1 - \frac{e^{-\lambda_1 t}(\bar{p}-c_1)-\beta_1 s\bar{p}-s\gamma_1}{e^{-\lambda_1 t}(r_1+\bar{r}_1)}\right)^2\\
&-r_1 e^{-\lambda_1 t}\left(u_1-\bar{u}_1+\frac{s((p-\bar{p})\alpha_1+\xi_1)}{e^{-\lambda_1 t}r_1}\right)^2\\
&+\frac{s^2(p-\bar{p})^2\alpha_1^2+2s^2(p-\bar{p})\alpha_1\xi_1+s^2\xi_1^2}{e^{-\lambda_1 t}r_1}\\
&+\frac{(e^{-\lambda_1 t}(\bar{p}-c_1)-\beta_1 s\bar{p}-s\gamma_1)^2}{e^{-\lambda_1 t}(r_1+\bar{r}_1)} - 2s\beta_1\bar{p}\sum_{j=2}^{n}\bar{u}_j\\
&-2s((p-\bar{p})\alpha_1+\xi_1)\sum_{j=2}^{n}(u_j-\bar{u}_j) - 2s\gamma_1\sum_{j=2}^{n}\bar{u}_j.
\end{aligned}$$

Then,

$$
\begin{aligned}
&E[\mathcal{R}_1 - f_1(0,p(0))|\mathcal{F}_T^{B_0}]\\
&= + \frac{\alpha_1(T) - qe^{-\lambda_1 T}}{2} E[(p(T)-\bar{p}(T))^2|\mathcal{F}_T^{B_0}]\\
&\quad - E\left[\int_0^T \frac{r_1+\bar{r}_1}{2} e^{-\lambda_1 t}\left(\bar{u}_1 - \frac{e^{-\lambda_1 t}(\bar{p}-c_1) - \beta_1 s\bar{p} - s\gamma_1}{e^{-\lambda_1 t}(r_1+\bar{r}_1)}\right)^2 dt|\mathcal{F}_T^{B_0}\right]\\
&\quad - E\left[\int_0^T \frac{r_1}{2} e^{-\lambda_1 t}\left(u_1 - \bar{u}_1 + \frac{s((p-\bar{p})\alpha_1+\xi_1)}{e^{-\lambda_1 t} r_1}\right)^2 dt|\mathcal{F}_T^{B_0}\right]\\
&\quad + E\left[\int_0^T \frac{(p-\bar{p})^2}{2}\left\{d\alpha_1 - (2s\alpha_1 - \frac{s^2\alpha_1^2}{e^{-\lambda_1 t} r_1})dt\right\}|\mathcal{F}_T^{B_0}\right]\\
&\quad + E\left[\int_0^T \frac{(p-\bar{p})}{2}\left\{2d\xi_1 - \left(2s\xi_1 + 2s\alpha_1\sum_{j=2}^n (u_j - \bar{u}_j)\right) dt\right.\right.\\
&\quad \left.\left. - \frac{2s^2\alpha_1\xi_1}{e^{-\lambda_1 t} r_1} dt\right\}|\mathcal{F}_T^{B_0}\right] + E\left[\int_0^T \frac{\bar{p}}{2}\left\{2d\gamma_1 + (2\beta_1 s(a - \sum_{j=2}^n \bar{u}_j)\right.\right.\\
&\quad + 2\sigma_0\beta_{1,0} - 2s\gamma_1 - \frac{2(e^{-\lambda_1 t} - \beta_1 s)(s\gamma_1 + e^{-\lambda_1 t}c_1)}{(r_1+\bar{r}_1)e^{-\lambda_1 t}})dt\\
&\quad \left.\left. +2\beta_0\sigma_0 dB_0\right\}|\mathcal{F}_T^{B_0}\right]\\
&\quad + E\left[\int_0^T \frac{\bar{p}^2}{2}\left\{d\beta_1 - (2\beta_1 s - \frac{(e^{-\lambda_1 t} - \beta_1 s)^2}{e^{-\lambda_1 t}(r_1+\bar{r}_1)})dt\right\}|\mathcal{F}_T^{B_0}\right]\\
&\quad + \frac{1}{2}E\left[\int_0^T \left\{2d\delta_1 + \frac{(s\gamma_1 + e^{-\lambda_1}c_1)^2}{e^{-\lambda_1}(r_1+\bar{r}_1)}dt + \frac{\xi_1^2 s^2}{e^{-\lambda_1 t} r_1}dt + \beta_1\sigma_0^2 dt\right.\right.\\
&\quad - 2s\xi_1\sum_{j=2}^n (u_j - \bar{u}_j)dt + \alpha_1(\sigma^2 + \int_\Theta \mu^2(\theta)v(d\theta))dt\\
&\quad + 2sa\gamma_1 dt - 2s\gamma_1\sum_{j=2}^n \bar{u}_j dt + 2\gamma_{1,0}\sigma_0 dt\\
&\quad \left.\left. +2\sigma_0\gamma_1 dB_0\right\}|\mathcal{F}_T^{B_0}\right].
\end{aligned}
$$

We deduce that $E[\mathcal{R}_1 - f_1(0,p(0))|\mathcal{F}_T^{B_0}] \leq 0$ and the equality occurs when

$$\bar{u}_1^* = \frac{e^{-\lambda_1 t}(\bar{p}-c_1) - \beta_1 s\bar{p} - s\gamma_1}{e^{-\lambda_1 t}(r_1+\bar{r}_1)},$$

and

$$u_1^* = -\frac{s((p-\bar{p})\alpha_1+\xi_1)}{e^{-\lambda_1 t} r_1} + \bar{u}_1^*.$$

Moreover,

$$
\begin{aligned}
d\alpha_1 &= \left(2s\alpha_1 - \frac{s^2\alpha_1^2}{e^{-\lambda_1 t}r_1}\right)dt + \alpha_{1,0}dB_0,\\
\alpha_1(T) &= -qe^{-\lambda_1 T},\\
d\beta_1 &= \left(2\beta_1 s - \frac{(e^{-\lambda_1 t}-\beta_1 s)^2}{e^{-\lambda_1 t}(r_1+\bar{r}_1)}\right)dt + \beta_{1,0}dB_0,\\
\beta_1(T) &= 0,\\
d\gamma_1 &= -\Bigg(\beta_1 s(a - \sum_{j=2}^{n}\bar{u}_j) + \sigma_0\beta_{1,0} - s\gamma_1\\
&\quad - \frac{(e^{-\lambda_1 t}-\beta_1 s)(s\gamma_1 + e^{-\lambda_1 t}c_1)}{(r_1+\bar{r}_1)e^{-\lambda_1 t}}\Bigg)dt - \beta_1\sigma_0 dB_0,\\
\gamma_1(T) &= 0,\\
d\xi_1 &= s\left(\xi_1 + \alpha_1\sum_{j=2}^{n}(u_j - \bar{u}_j) - \frac{s\alpha_1\xi_1}{e^{-\lambda_1 t}r_1}\right)dt\\
&\quad + \xi_{1,0}dB_0,\\
\xi_1(T) &= 0,\\
d\delta_1 &= -\Bigg(\frac{1}{2}\frac{(s\gamma_1 + e^{-\lambda_1 t}c_1)^2}{e^{-\lambda_1 t}(r_1+\bar{r}_1)} + \frac{1}{2}\frac{\xi_1^2 s^2}{e^{-\lambda_1 t}r_1} + \frac{\beta_1}{2}\sigma_0^2 - s\xi_1\sum_{j=2}^{n}(u_j-\bar{u}_j)\\
&\quad + \frac{\alpha_1}{2}(\sigma^2 + \int_\Theta \mu^2(\theta)v(d\theta)) + s\gamma_1(a - \sum_{j=2}^{n}\bar{u}_j) + \gamma_{1,0}\sigma_0\Bigg)dt\\
&\quad - \sigma_0\gamma_1 dB_0,\\
\alpha(T) &= 0.
\end{aligned}
$$

Similarly, we can find $u_2^*, \ldots, u_{n-1}^*$. Next, we find the optimal control for u_n^*. Giving $\bar{u}_2^*, \ldots, \bar{u}_{n-1}^*$, the conditional log-price dynamics becomes

$$
\begin{aligned}
d\bar{p} &= s\left[a - \sum_{j=1}^{n-1}\frac{e^{-\lambda_j t}(\bar{p}-c_j) - \beta_j s\bar{p} - s\gamma_j}{e^{-\lambda_j t}(r_j+\bar{r}_j)} - \bar{u}_n - \bar{p}\right]dt\\
&\quad + \sigma_0 dB_0, \quad \bar{p}(0) = \bar{p}_0.
\end{aligned}
$$

Now, we use the guess functional

$$
f_n(t,p) = \frac{1}{2}\alpha_n(t)(p-\bar{p})^2 + \frac{1}{2}\beta_n(t)\bar{p}^2 + \xi_n(t)(p-\bar{p}) + \gamma_n(t)\bar{p} + \delta_n(t), \tag{B.8}
$$

where $\alpha_n, \beta_n, \gamma_n$ and δ_n and ξ_n are random functions of time t, that are $\mathcal{F}_t^{B_0}-$measurable such that

$$f_n(T, p(T)) = -\frac{q}{2}e^{-\lambda_n T}\alpha_n(t)(p(T) - \bar{p}(T))^2.$$

Using the Ito formula we obtain

$$\begin{aligned}
E[\mathcal{R}_n - f_n(0, p(0))|\mathcal{F}_T^{B_0}] = &\int_0^T e^{-\lambda_n t} E[(\bar{p} - c_n)\bar{u}_n \\
&- \frac{1}{2}r_n(u_n - \bar{u}_n)^2 - \frac{1}{2}(r_n + \bar{r}_n)\bar{u}_n^2|\mathcal{F}_T^{B_0}]dt \\
&+ \frac{\alpha_n(T) - qe^{-\lambda_n T}}{2}E[(p(T) - \bar{p}(T))^2|\mathcal{F}_T^{B_0}] \\
&+ E[\int_0^T \frac{(p - \bar{p})^2 d\alpha_n}{2}|\mathcal{F}_T^{B_0}] \\
&+ E[\int_0^T \frac{\alpha_n}{2}\left(\sigma^2 + \int_\Theta \mu^2(\theta)v(d\theta)\right)dt|\mathcal{F}_T^{B_0}] \\
&- E[\int_0^T (s(p - \bar{p})\alpha_n + s\xi_n)(D^* - \bar{D}^* + p - \bar{p})dt|\mathcal{F}_T^{B_0}] \\
&+ \frac{1}{2}E[\int_0^T \beta_n(2s\bar{p}(a - \bar{D}^* - \bar{p}) + \sigma_0^2)dt|\mathcal{F}_T^{B_0}] \\
&+ E[\int_0^T \beta_n\bar{p}\sigma_0 dB_0|\mathcal{F}_T^{B_0}] + E[\int_0^T \beta_{n,0}\bar{p}\sigma_0 dt|\mathcal{F}_T^{B_0}] \\
&+ E[\int_0^T \bar{p}d\gamma_n|\mathcal{F}_T^{B_0}] + E[\int_0^T \sigma_0\gamma_n dB_0|\mathcal{F}_T^{B_0}] \\
&+ E[\int_0^T d\delta_n|\mathcal{F}_T^{B_0}] + E[\int_0^T (p - \bar{p})d\xi_n|\mathcal{F}_T^{B_0}] \\
&+ E[\int_0^T \left(s\gamma_n(a - \bar{D}^* - \bar{p}) + \gamma_{n,0}\sigma_0\right)dt|\mathcal{F}_T^{B_0}] \\
&+ \frac{1}{2}E[\int_0^T \bar{p}^2 d\beta_n|\mathcal{F}_T^{B_0}],
\end{aligned}$$

where

$$D^* = \sum_{j=1}^{n-1} u_j^* + u_n \quad \text{and} \quad \bar{D}^* = \sum_{j=1}^{n-1} \bar{u}_j^* + \bar{u}_n.$$

As for player 1, we complete the squares to obtain

$$
\begin{aligned}
&E[\mathcal{R}_n - f_n(0,p(0))|\mathcal{F}_T^{B_0}] \\
&\quad = -E\left[\int_0^T \frac{r_n+\bar{r}_n}{2} e^{-\lambda_n t}\left(\bar{u}_n - \frac{e^{-\lambda_n t}(\bar{p}-c_n) - \beta_n s\bar{p} - s\gamma_n}{e^{-\lambda_n t}(r_n+\bar{r}_n)}\right)^2 dt|\mathcal{F}_T^{B_0}\right] \\
&\quad - E\left[\int_0^T \frac{r_n}{2} e^{-\lambda_n t}\left(u_n - \bar{u}_n + \frac{s((p-\bar{p})\alpha_n + \xi_n)}{e^{-\lambda_n t} r_n}\right)^2 dt|\mathcal{F}_T^{B_0}\right] \\
&\quad + E\left[\int_0^T \frac{(p-\bar{p})^2}{2}\left\{d\alpha_n - (2s\alpha_n - \frac{s^2\alpha_n^2}{e^{-\lambda_n t} r_n})dt - 2s^2\alpha_n \sum_{j=1}^{n-1}\frac{\alpha_j}{e^{-\lambda_j t} r_j}dt\right\}|\mathcal{F}_T^{B_0}\right] \\
&\quad + E\left[\int_0^T (p-\bar{p})\left\{d\xi_n - s\xi_n dt + \left(s^2\alpha_n \sum_{j=1}^{n-1}\frac{\xi_j}{e^{-\lambda_j t} r_j}\right.\right.\right. \\
&\quad \left.\left.\left. + s^2\xi_n \sum_{j=1}^{n-1}\frac{\alpha_j}{e^{-\lambda_j t} r_j} + \frac{s^2\alpha_n\xi_n}{e^{-\lambda_n t} r_n}\right) dt\right\}|\mathcal{F}_T^{B_0}\right] \\
&\quad + E\left[\int_0^T \bar{p}\left\{d\gamma_n + s\beta_n a dt + s\beta_n \sum_{j=1}^{n-1}\frac{e^{-\lambda_j t}c_j + s\gamma_j}{e^{-\lambda_j t}(r_j+\bar{r}_j)}dt - s\gamma_n \sum_{j=1}^{n-1}\frac{e^{-\lambda_j t} - s\beta_j}{e^{-\lambda_j t}(r_j+\bar{r}_j)}dt\right.\right. \\
&\quad \left.\left. + \sigma_0\beta_{n,0}dt - s\gamma_n dt - \frac{(e^{-\lambda_n t} - \beta_n s)(s\gamma_n + e^{-\lambda_n t}c_n)}{(r_n+\bar{r}_n)e^{-\lambda_n t}}dt + \beta_n\sigma_0 dB_0\right\}|\mathcal{F}_T^{B_0}\right] \\
&\quad + E\left[\int_0^T \frac{\bar{p}^2}{2}\left\{d\beta_n - \left(2\beta_n s - \frac{(e^{-\lambda_n t} - \beta_n s)^2}{e^{-\lambda_n t}(r_n+\bar{r}_n)} + 2s\beta_n \sum_{j=1}^{n-1}\frac{(e^{-\lambda_j t} - \beta_j s)}{e^{-\lambda_j t}(r_j+\bar{r}_j)}\right) dt\right\}|\mathcal{F}_T^{B_0}\right] \\
&\quad + E\left[\int_0^T \left\{d\delta_n + \frac{1}{2}\frac{(s\gamma_n + e^{-\lambda_n}c_2)^2}{e^{-\lambda_n}(r_n+\bar{r}_n)}dt + s^2\xi_n \sum_{j=1}^{n-1}\frac{\xi_j}{e^{-\lambda_j t} r_j}dt\right.\right. \\
&\quad + \frac{1}{2}\frac{s^2\xi_n^2}{e^{-\lambda_n} r_n}dt + \frac{1}{2}\beta_n\sigma_0^2 dt + \frac{\alpha_n}{2}(\sigma^2 + \int_\Theta \mu^2(\theta)v(d\theta))dt + s\gamma_n a dt \\
&\quad \left.\left. + s\gamma_n \sum_{j=1}^{n-1}\frac{e^{-\lambda_j t}c_j + s\gamma_j}{e^{-\lambda_j t}(r_j+\bar{r}_i)}dt + \gamma_{n,0}\sigma_0 dt + \sigma_0\gamma_n dB_0\right\}|\mathcal{F}_T^{B_0}\right].
\end{aligned}
$$

We deduce that $E[\mathcal{R}_n - f_n(0,p(0))|\mathcal{F}_T^{B_0}] \leq 0$ and the equality occurs when

$$\bar{u}_n^* = \frac{e^{-\lambda_n t}(\bar{p}-c_n) - \beta_n s\bar{p} - s\gamma_n}{e^{-\lambda_n t}(r_n+\bar{r}_n)},$$

and

$$u_n^* = -\frac{s((p-\bar{p})\alpha_n + \xi_n)}{e^{-\lambda_n t} r_n} + \bar{u}_n^*.$$

Moreover,

$$
\begin{aligned}
d\alpha_n &= \left(2s\alpha_n - \frac{s^2\alpha_n^2}{e^{-\lambda_n t} r_n} - 2s^2\alpha_n \sum_{i=1}^{n-1} \frac{\alpha_i}{e^{-\lambda_i t} r_i}\right) dt + \alpha_{n,0} dB_0,\\
\alpha_n(T) &= -qe^{-\lambda_n T},\\
d\beta_n &= \left(-\frac{(e^{-\lambda_n t} - \beta_n s)^2}{e^{-\lambda_n t}(r_n + \bar{r}_n)} + 2s\beta_n \sum_{i=1}^{n-1} \frac{(e^{-\lambda_i t} - \beta_i s)}{e^{-\lambda_i t}(r_i + \bar{r}_i)} + 2\beta_n s\right) dt + \beta_{n,0} dB_0,\\
\beta_n(T) &= 0,\\
d\gamma_n &= -\Bigg(\beta_n s(a + \sum_{i=1}^{n-1} \frac{e^{-\lambda_i t} c_i + s\gamma_i}{e^{-\lambda_i t}(r_i + \bar{r}_i)}) + \sigma_0\beta_{n,0} - s\gamma_n \sum_{i=1}^{n-1} \frac{(e^{-\lambda_i t} - s\beta_i)}{e^{-\lambda_i t}(r_i + \bar{r}_i)} + \sigma_0\beta_{n,0}\\
&\quad - s\gamma_n - \frac{(e^{-\lambda_n t} - \beta_n s)(s\gamma_n + e^{-\lambda_n t} c_n)}{(r_n + \bar{r}_n)e^{-\lambda_n t}}\Bigg) dt - \beta_n\sigma_0 dB_0,\\
\gamma_n(T) &= 0,\\
d\xi_n &= s\left(\xi_n - s\alpha_n \sum_{i=1}^{n-1} \frac{\xi_i}{e^{-\lambda_i t} r_i} - \frac{s\alpha_n\xi_n}{e^{-\lambda_n t} r_n} - s\xi_n \sum_{i=1}^{n-1} \frac{\alpha_i}{e^{-\lambda_i t} r_i}\right) dt + \xi_{n,0} dB_0,\\
\xi_n(T) &= 0,\\
d\delta_n &= -\Bigg(\frac{1}{2} \frac{(s\gamma_n + e^{-\lambda_n} c_2)^2}{e^{-\lambda_n}(r_n + \bar{r}_n)} + s^2\xi_n \sum_{i=1}^{n-1} \frac{\xi_i}{e^{-\lambda_i t} r_i} + \frac{1}{2} \frac{s^2\xi_n^2}{e^{-\lambda_n} r_n} + \frac{1}{2}\beta_n\sigma_0^2\\
&\quad + \frac{\alpha_n}{2}\left(\sigma^2 + \int_\Theta \mu^2(\theta) v(d\theta)\right) + s\gamma_n a + s\gamma_n \sum_{i=1}^{n-1} \frac{e^{-\lambda_i t} c_i + s\gamma_i}{e^{-\lambda_i t}(r_i + \bar{r}_i)} + \gamma_{n,0}\sigma_0\Bigg) dt\\
&\quad - \sigma_0\gamma_n dB_0,\\
\delta_n(T) &= 0.
\end{aligned}
$$

The result follows by replacing backwardly the optimal solutions. ■

5 Numerical examples

The prosumers do not necessarily have the same equipment and technology. In order to produce electricity based on solar panels, solar cells are made from a semiconducting material that converts light into electricity. The most common material used as a semiconductor during the solar cell manufacturing process is silicon. Monocrystalline solar panels, Polycrystalline solar panels, Bifacial solar panels, and thin-film solar panels, each have their own advantages and disadvantages. It implies that each agent may have different production cost parameters. Therefore, we consider a scenario where all agents are heterogeneous with regard to the structure of their costs. We, furthermore, consider a homogeneous case, where all the agents have the same cost structure, to investigate the effect of hierarchy on the optimal solutions and revenues. Note that the production u_i needs to be bounded (by the production capacity). However, working with the constrained control set $[0, c]$ creates a solution (bang bang control) that is different from the unconstrained one. In order to capture it we use penalization term for big control.

Table B.1: Parameters for scenario 1 and 2

Parameter	**Scenario 1**	**Scenario 2**
T	15	15
p_0	10	10
c_1, c_2, c_3	1	1
λ_1	0.1	0.1
λ_2	0.15	0.1
λ_3	0.2	0.1
s	0.5	0.5
a	5	5
$r_1,\ \bar{r}_1$	10	20
$r_2,\ \bar{r}_2$	15	20
$r_3,\ \bar{r}_3$	20	20
q	30	30
σ	2	2
σ_0	0	0
μ	0.2	0.2

Table B.2: The revenues of the agents for scenarios 1 and 2.

Scenario	$\mathcal{R}_1^*$	$\mathcal{R}_2^*$	$\mathcal{R}_3^*$
1	2.6373e+05	1.4758e+05	9.5672e+04
2	1.3913e+05	1.3913e+05	1.3913e+05

Figures B.1 and B.2 present the behavior of the mean-field-type control problem corresponding to Scenarios 1 and 2, respectively. We notice different evolution for

$\bar{u}_i^*$ and $\mathcal{R}_i(u_i^*)$, for all i.

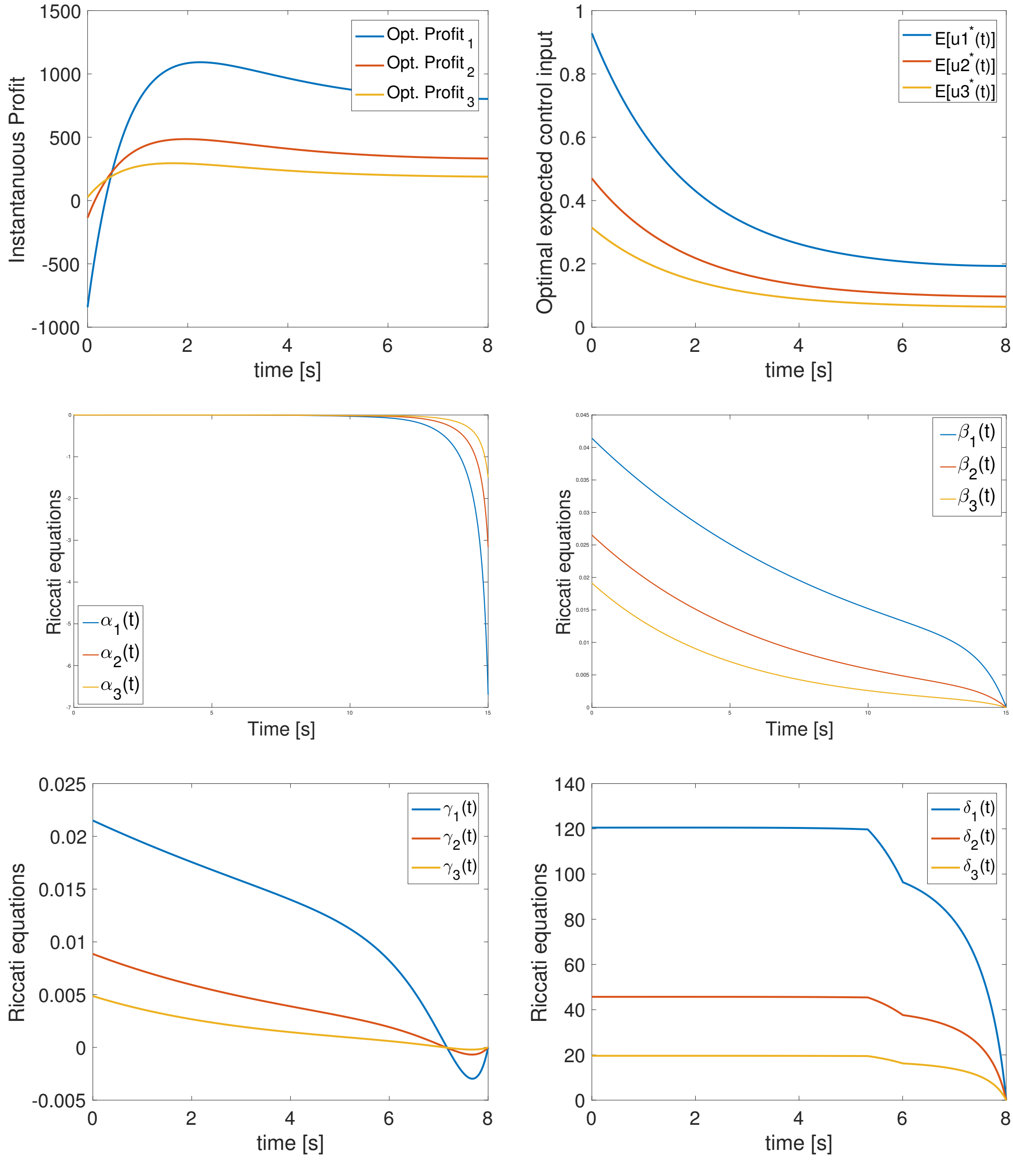


Fig. B.1: Mean-field-type hierarchical performance corresponding to Scenario 1.

Besides, it can be seen that the Riccati equations satisfy the boundary conditions e.g., $\alpha_i(T) = -q = -1$.

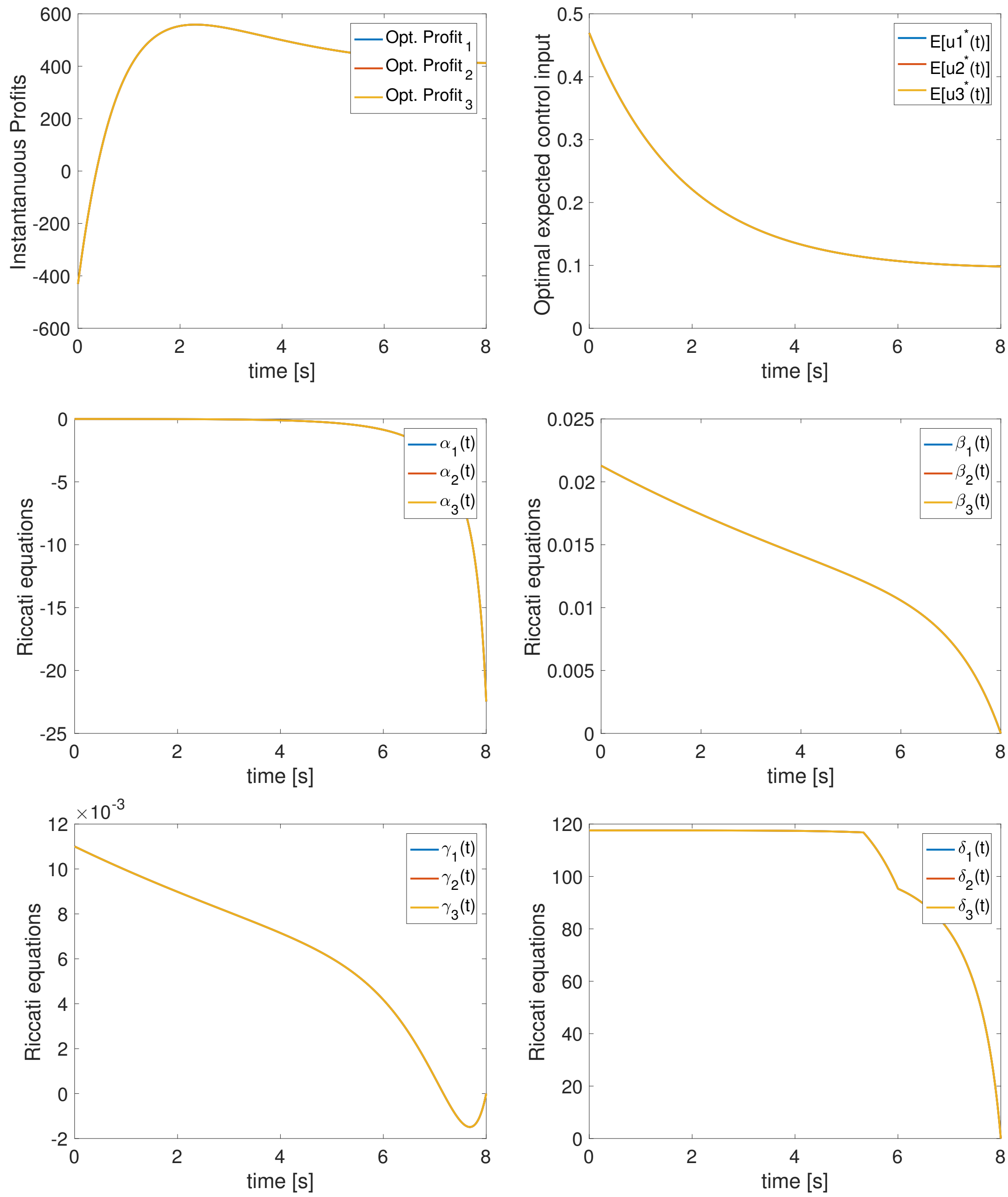


Fig. B.2: Mean-field-type hierarchical performance corresponding to Scenario 2.

6 Conclusion

We formulated a mean-field type hierarchical profit optimization problem that includes $n-1$ prosumers and a producer in an electricity market. All the agents are coupled through the conditional price dynamics that is given by a linear jump-diffusion system of conditional mean-field type. We considered a quadratic cost functional of mean-field-type and we provided a semi-explicit solution of the corresponding mean-field-type hierarchical control problem, using a direct method consisting of a square completion technique.

The optimal controls are in state-and-conditional mean-field feedback form. They coincide with Nash equilibrium solutions obtained in (Djehiche et al., 2020), where a mean-field-type game between electricity producers is considered. The coincidence is due to the fact that there are no cross terms $u_i u_j$ in the considered cost functional. This makes the expression of the optimal control of each agent does not depend explicitly on the optimal controls of the others. Therefore, being the first to decide on a strategy does not award any advantages.

The numerical examples validate our theoretical results. Table B.2 highlights the fact that in the homogeneous case, where all the agents have the same production cost parameters, the major player (the producer) does not make the highest profit despite of the precedence of action and thus, the hierarchy has no effect on the optimal controls, which is in line with Proposition 4.1.

References

Bagchi, A., & Basar, T. (1981). Stackelberg strategies in linear-quadratic stochastic differential games. *Journal of Optimization Theory and Applications*, *35*(3), 443-464.

Basar, T., & Selbuz, H. (1979). Closed-loop stackelberg strategies with applications in the optimal control of multilevel systems. *IEEE Transactions on Automatic Control*, *24*(2), 166–179.

Cruz, J. (1978). Leader-follower strategies for multilevel systems. *IEEE Transactions on Automatic Control*, *23*(2), 244–255.

Djehiche, B., Barreiro-Gomez, J., & Tembine, H. (2020). Price dynamics for electricity in smart grid via mean-field-type games. *Dynamic Games and Applications.* (DOI:https://doi.org/10.1007/s13235-020)

Evans, G. C. (1922). A simple theory of competition. *American Mathematical Monthly*, *29*(10), 371–380.

Gardner, B., & Cruz, J. (1978). Feedback stackelberg strategy for m-level hierarchical games. *IEEE Transactions on Automatic Control*, *23*(3), 489–491.

Hobbs, B. E. (2001). Linear complementarity models of nash-cournot competition in bilateral and poolco power markets. *IEEE Transactions on power systems*, *16*(2), 194–202.

Hobbs, B. F., & Pang, J. (2007). Nash-cournot equilibria in electric power markets with piecewise linear demand functions and joint constraints. *Operations Research*, *55*(1), 113–127.

Hogan, W. (1997). A market power model with strategic interaction in electricity networks. *The Energy Journal*, *18*(4).

Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica*, *50*(3), 649-670.

Li, N., & Yu, Z. (2018). Forward-backward stochastic differential equations and linear-quadratic generalized stackelberg games. *SIAM Journal on Control and Optimization*, *56*(6), 4148–4180.

Metzler, C., Hobbs, B. F., & Pang, J. (2003). Nash-cournot equilibria in power markets on a linearized dc network with arbitrage: Formulations and properties. *Networks and Spatial Economics*, *3*(2), 123–150.

Pan, L., & Yong, J. (1991). A differential game with multi-level of hierarchy. *Journal of mathematical analysis and applications*, *161*(2), 522–544.

Ross, C. F. (1925). A mathematical theory of competition. *American journal of mathematics*, *47*(3), 163–175.

Simaan, M., & Cruz, J. (1973b). A stackelberg solution for games with many players. *IEEE Transactions on Automatic Control*, *18*(3), 322–324.

Simaan, M., & Cruz, J. B. (1973a). On the stackelberg strategy in nonzero-sum games. *Journal of Optimization Theory and Applications*, *11*(5), 533–555.

Stackelberg, H. V. (1948). The theory of the market economy, trans. by aj peacock, london, william hodge. *Originally published as Grundlagen der Theoretischen Volkswirtschaftlehre*.

Wanga, B., & Huang, M. (2019). Mean field production output control with sticky prices: Nash and social solutions. *Automatica*, *100*(2019), 90-98.

Willems, B. (2002). Modeling cournot competition in an electricity market with transmission constraints. *The Energy Journal*, *23*(3).

Perspective

Open questions that we leave for future investigations are:

- finds, theoretically, the optimal ordering among all permutations of heterogenous decision-makers, and examines the benefits/costs of structure design and leadership.
- Model other real-world problems that have a hierarchical structure, such as cyber-attacks, the spread of epidemics, and traffic flows, and use real data for the numerical results.
- Use reinforcement learning (q-learning) and rather artificial intelligence techniques to solve hierarchy MFTG and other types of games.