Titre: Title:	Collective Stochastic Discrete Choice Problems: A MIN-LQG Game Formulation
Auteur: Author:	Rabih Salhab
Date:	2018
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Salhab, R. (2018). Collective Stochastic Discrete Choice Problems: A MIN-LQG Game Formulation [Ph.D. thesis, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/3028/

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/3028/
Directeurs de recherche: Advisors:	Jérôme Le Ny, & Roland P. Malhamé
Programme: Program:	génie électrique

UNIVERSITÉ DE MONTRÉAL

COLLECTIVE STOCHASTIC DISCRETE CHOICE PROBLEMS: A MIN-LQG GAME FORMULATION

RABIH SALHAB
DÉPARTEMENT DE GÉNIE ÉLECTRIQUE
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

THÈSE PRÉSENTÉE EN VUE DE L'OBTENTION
DU DIPLÔME DE PHILOSOPHIÆ DOCTOR
(GÉNIE ÉLECTRIQUE)
JANVIER 2018

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Cette thèse intitulée:

COLLECTIVE STOCHASTIC DISCRETE CHOICE PROBLEMS: A MIN-LQG GAME FORMULATION

présentée par : SALHAB Rabih
en vue de l'obtention du diplôme de : Philosophiæ Doctor
a été dûment acceptée par le jury d'examen constitué de :

M. SAUSSIÉ David, Ph. D., président
M. MALHAMÉ Roland, Ph. D., membre et directeur de recherche
M. LE NY Jérôme, Ph. D., membre et codirecteur de recherche
M. CAINES Peter, Ph. D., membre
M. JADBABAIE Ali, Ph. D., membre externe

DEDICATION

To my family and my little Kinda

ACKNOWLEDGEMENTS

I'm very grateful to my supervisors Prof. Roland Malhamé and Prof. Jérôme Le Ny for giving me the opportunity to enter the research field. I thank them for their support and great advice in my research, and for their generous help in my future career.

I thank the members of my committee Prof. David Saussié, Prof. Peter Caines, Prof. Ali Jadbabaie and Prof. Mario Lefebvre for their time and effort to read my thesis. I'm grateful to Prof. Lahcen Saydy for supporting my accelerated transition from Master's to PhD.

I would like to thank all my professors and colleagues at Polytechnique, McGill and GERAD. I would like to thank my supportive cousins and friends Ibrahim Sbeity, Joumana Skeini, Youssef Sbeity, Hassan Sbeity Iman Dahwi, Sahar Komayha, Khalil Sbeity, Mohammed Badreddine, Issa Badreddine, Fatima Sbeity, Ali Mcheimech, Manal Komayha, Mohammad-Ali Mcheimech, Hadi El-Amine, Wissam El-Amine, Bob Komayha, Hassan Choueib, Ali Dirani, Ali Kassem, Nada Wehbe, Maryana Wazni, Mahmoud Ashour, Salwa Haidar, Talal Halabi, Mohammad Sbeity, Zahraa Sbeity, Ali Komayha, Najib Komayha, Soukayna Jaffal, Mohammad Khosravi, Meisam Mohammady, Gerard Degue, Saad Abobakr and Mohsen Ghafouri.

I'm deeply grateful to my mother, father, sisters, brother-in-law, my lovely niece Kinda and the rest of my family for their endless love, support and sacrifices.

RÉSUMÉ

On étudie dans cette thèse une classe de problèmes dynamiques concernant des choix collectifs discrets. Il s'agit de situations où un grand nombre d'agents exercent un effort pendant un horizon de temps fini pour se déplacer de leur position initiale et être à la fin de l'horizon aux alentours d'un nombre fini d'alternatives. En bougeant, ils doivent minimiser leurs efforts. De plus, le trajet de chaque agent est influencé par le comportement global de la population (l'effet social). Par exemple, un groupe de robots explorant un terrain doivent bouger de leurs positions initiales pour visiter plusieurs sites d'intérêt en minimisant leurs efforts. Le groupe peut se diviser en plusieurs sous-groupes, mais la taille de chacun doit rester suffisamment grande pour accomplir quelques tâches collectives, par exemple, des missions de secours. Ces problèmes sont aussi reliés aux élections, où les électeurs changent continuellement leurs opinions jusqu'au moment de décider pour qui ils veulent voter. Durant ce processus de choix d'un candidat, l'électeur optimise ses efforts pour ajuster son opinion et rester dans la zone de confort, c'est-à-dire, proche des opinions des autres. Le but de cette thèse est de comprendre les comportements individuels des agents et en déduire le comportement global de la population, et cela, lorsque les agents prennent leurs décisions de façon soit non-coopérative, soit coopérative. En d'autres termes, pour chacun des deux cas, on répond aux questions suivantes : comment un agent choisit-il sous l'influence de l'effet social? Peut-on prédire le comportement global de la population et la distribution des choix entre les alternatives?

On modélise les problèmes des choix discrets non-coopératifs par une nouvelle classe de jeux dynamiques non-coopératifs qu'on nomme “min-LQG”. Ce sont des versions modifiées des jeux LQG où les coûts individuels incluent un coût final qui est la distance minimale à l'ensemble des alternatives pour modéliser le phénomène des choix discrets. On étudie deux cas. Dans le premier (min-LQG dégénéré), on considère des agents décrits par des équations différentielles avec des conditions initiales aléatoires, alors que dans le deuxième (min-LQG non-dégénéré), leurs dynamiques sont des équations différentielles stochastiques. On utilise la méthodologie des jeux à champs moyen pour analyser les jeux. On développe des stratégies décentralisées, où chaque agent peut implémenter la sienne en observant son état et la distribution de probabilité des états initiaux des agents. Ces stratégies convergent vers un équilibre de Nash lorsque le nombre d'agents tend vers l'infini. Dans le cas dégénéré, un agent fait son choix d'alternative avant de bouger en se basant sur son état initial. Mais, dans le cas non-dégénéré, il ne peut plus s'engager à un choix dès le début. En effet, on montre que l'agent revisite son choix continuellement pour prendre en considération le risque

de faire une décision prématurée en faveur d'une alternative en se basant sur les conditions actuelles, alors que le bruit l'envoie à l'une des autres alternatives.

Pour les problèmes coopératifs des choix discrets, on considère le cas dégénéré seulement. On montre que le calcul d'un optimum social est complexe. De plus, son implémentation requiert beaucoup de communications entre les agents. Grâce à la méthodologie des jeux à champs moyen, des stratégies alternatives qui sont simples à calculer et implémenter sont proposées. Ces stratégies convergent vers un optimum social lorsque le nombre d'agents tend vers l'infini. On compare par un exemple numérique le comportement de la population dans les cas coopératif et non-coopératif. Les résultats montrent que les stratégies coopératives ont pour avantage de distribuer les agents d'une façon plus uniforme entre les alternatives. Ce comportement est préférable lorsqu'on considère l'exemple robotique pour améliorer la qualité des tâches collectives.

Alors que les choix finaux des alternatives dépendent des conditions initiales et du bruit, il est toujours possible de prédire la distribution de ces choix dans les cas non-coopératifs et coopératifs. En fait, on construit une bijection entre les solutions des problèmes non-coopératifs (équilibre de Nash approximatif) ou des problèmes coopératifs (optimum social approximatif) et les points fixes d'une fonction de dimension finie. Ces derniers sont les distributions potentielles de probabilités des choix sur l'ensemble des alternatives. Cette caractérisation des problèmes de choix discrets facilite la preuve d'existence de solutions et leur calcul pour ces problèmes de dimensions infinies.

Enfin, on étend le modèle des choix discrets dégénéré non-coopératif pour inclure un agent d'influence majeure dans le jeu. Cet agent (annonceur) fait des investissements publicitaires pour convaincre les autres agents (consommateurs) de choisir l'une des alternatives. La compétition considérée dans ce cas est de type Stackelberg, c.à.d., que l'annonceur investit, d'abord, puis les consommateurs choisissent une alternative sous l'influence de l'effet social et des investissements publicitaires. L'annonceur peut représenter, par exemple, l'équipe de compagne électorale de l'un des candidats dans l'exemple des élections. On considère au départ un continuum de consommateurs et l'on résout ce problème par la méthodologie des jeux à champs moyen. On montre que lorsque les opinions a priori des consommateurs vis-à-vis les alternatives sont suffisamment diversifiées, l'annonceur peut faire des investissements optimaux en connaissant uniquement la distribution de probabilité a priori de ces opinions. Quand ces investissements sont appliqués dans le cas réel d'un nombre fini de consommateurs, ils induisent une perte de performance dans le processus de publicité. Mais, cette perte devient négligeable quand le nombre de consommateurs croît suffisamment. On étudie le cas particulier où les opinions a priori ont une certaine distribution uniforme, et on développe des formes explicites des investissements publicitaires. De plus, on anticipe la distribution des choix sur l'ensemble des alternatives.

ABSTRACT

This thesis studies a class of dynamic collective discrete choice problems, whereby a large number of agents have to settle on one of a set of available alternatives within a finite time. Motion in choice space involves a control effort and it is socially influenced, in that the agents' paths are influenced by the aggregate behavior of the group. For example, a large collection of geographically dispersed robots exploring an unknown terrain must move with least energy expenditure from their initial randomly spread positions towards one of multiple potential sites of interest to visit. They do so, however, while trying to remain grouped to achieve some collective tasks, such as search and rescue. Another example is a mechanistic model of elections, where voters evolve continuously in opinion space until they form a final decision regarding who they should vote for. Until voters settle on a candidate, changes in their opinions will involve a mental strain, but at the same time, adhering to an individual opinion when it currently deviates from that of the majority will cause some level of discomfort. Even though both examples qualify as dynamic collective choice problems, they differ in the manner the agents make their choices. In the robotic example, a designer or social planner coordinates the robots' paths to optimize a global cost, while in the second example, the voters, left without any central coordinator, update their opinions selfishly irrespective of whether they make others better or worse-off. This suggests a classification of dynamic collective choice problems into cooperative and noncooperative types respectively. This work aims at modeling these problems and understanding how the population behaves on the microscopic (individual) and macroscopic (population) levels. In other words, we try to answer the following main questions: how does an agent make its choice under the social effect? Can one anticipate the evolution of the agents' state probability distribution (macroscopic behavior) along the time interval, and that of the agents' ultimate choices over the alternatives?

We formulate the non-cooperative problem as a dynamic non-cooperative game with a so-called "min-LQG" individual cost. This consists of a modified version of the linear quadratic Gaussian performance index that includes a choice-related minimization term in the final cost to capture the discrete choice aspect. We study two cases. In the first one, the degenerate case, the agents' states evolve according to a set of linear ordinary differential equations with the only source of randomness being their initial conditions. In the second one, the non-degenerate case, besides their random initial conditions, agents are subjected to random noise, thus evolving according to stochastic differential equations. On the methodological side, we use the mean field games theory to generate decentralized rational strategies, which

are shown to converge to a Nash equilibrium as the size of the agent population increases to infinity. We show that in the degenerate case an agent makes its choice of an alternative at the outset based on its initial condition and the initial probability distribution of the agent states. In the non-degenerate case, however, the agents can no longer commit to a choice from the beginning. Indeed, they continuously revise their decisions to account for the risk of making a premature "myopic" decision in favor of one alternative based on the current state, while being driven to one of the other alternatives by the noise process.

In the cooperative problem, we consider only the degenerate case. We show that the naïve approach to computing a social optimum becomes intractable as the size of the population increases. Furthermore, implementation of the resulting control law requires a significant amount of communication between the agents. Instead, we develop via the mean field games methodology a set of strategies that are simpler to compute and implement than the social optimum. While these strategies induce a priori a loss of optimality, the latter becomes negligible for a large number of agents. Indeed, they converge to a social optimum as the size of the population increases to infinity. As in the degenerate non-cooperative case, the agents make their choices at the outset based on their initial conditions and the agent states initial probability distribution. We provide a numerical example to compare the cooperative and non-cooperative formulations. The results show that the cooperative strategies have the advantage of much more evenly allocating the agents to the alternatives, a crucial feature required when considering the robotic example, for instance, to improve the quality of the collective tasks.

Although the individual paths and ultimate agent choices remain dictated by an individual's random initial state and the associated noise process history, the mean field games methodology allows one to anticipate both the probability distribution of the agents' choices over the alternatives and the evolution of the states' probability distribution in the cooperative as well as the non cooperative cases. In fact, we show that there exists a one-to-one map between the infinite population Nash equilibria in the non-cooperative case, or social optima in the cooperative one, and the fixed points of a particular finite dimensional map. These fixed points are the potential probability distributions of the agents' choices over the alternatives, and can in turn lead to a complete characterization of the corresponding macroscopic behaviors. As a result of the identification of this map, the infinite dimensional problem of identifying the candidate probabilistic evolutions of agent states in the dynamic non-cooperative and cooperative situations is reduced to that of computing candidate fixed points of a finite dimensional map. Subsequently, numerical schemes based on Quasi-Newton or the bisection methods are proposed to generate the infinite population Nash and social optima strategies.

Finally, in the presence of only two alternatives, the degenerate non-cooperative framework is further extended to include a single major agent called an advertiser, and a large collection of minor agents called consumers. The major agent spends money on advertisements aimed at convincing minor agents to choose a specific alternative among the two. The advertiser represents, for instance, the electoral campaign team of one of two opposing candidates in the elections example. He/She is “Stackelbergian”, in the sense that he/she invests at first, and then the consumers make their choices under the social and advertising effects. By going to the limit of a continuum of consumers, we solve the game via the mean field games methodology. In particular, when consumers have sufficiently diverse a priori opinions, we show that a unique Nash equilibrium exists at the consumers level. As a result, the advertiser can anticipate the consumers’ behavior and decide on optimal investments. These investments require that the advertiser know only the a priori probability distribution of the consumers’ opinions. When these investments are applied in real situations of a finite number of consumers, they induce a loss of performance in the advertising process. However, this loss becomes negligible in a large population. For the particular case of a uniform distribution of a priori opinions, we provide an explicit form of the advertiser’s optimal investment strategies, and consumers’ optimal choices. Moreover, we anticipate the probability distribution of the consumers’ choices over the alternatives.

TABLE OF CONTENTS

DEDICATION . . . iii

ACKNOWLEDGEMENTS . . . iv

RÉSUMÉ . . . v

ABSTRACT . . . viii

TABLE OF CONTENTS . . . xi

LIST OF TABLES . . . xiv

LIST OF FIGURES . . . xv

LIST OF ABBREVIATIONS AND NOTATIONS . . . xvi

LIST OF APPENDICES . . . xviii

CHAPTER 1 INTRODUCTION . . . 1
- 1.1 Research objectives . . . 2
- 1.2 Methodology and related works . . . 3
  - 1.2.1 Methodology - Mean field games theory . . . 3
  - 1.2.2 Relation to the discrete choice problems in microeconomics . . . 5
  - 1.2.3 Relation to opinion dynamics and elections . . . 6
  - 1.2.4 Relation to robotics . . . 7
  - 1.2.5 Relation to the advertising models . . . 7
- 1.3 Main contributions and structure of the thesis . . . 8
  - 1.3.1 First group of contributions . . . 8
  - 1.3.2 Second group of contributions . . . 11

CHAPTER 2 DEGENERATE NONCOOPERATIVE DYNAMIC COLLECTIVE DISCRETE CHOICE MODEL . . . 12
- 2.1 Mathematical model . . . 12
- 2.2 Exact Nash equilibria . . . 13
- 2.3 The mean field games methodology and the generic agent's best response . . . 17
- 2.4 Mean field equations . . . 20

2.5 Discussions 26
2.5.1 Commitment to the initial choice 26
2.5.2 Exact vs. approximate Nash equilibria 26
2.5.3 Full characterization of the limiting game 27
2.5.4 Simple numerical scheme 28
2.6 Nonuniform population with initial preferences 29
2.7 Approximate Nash equilibrium 34
2.8 Simulation results 36
2.9 Conclusion 37

CHAPTER 3 NON-DEGENERATE NONCOOPERATIVE DYNAMIC COLLECTIVE DISCRETE CHOICE MODEL 40
3.1 Mathematical model 40
3.2 The min-LQG optimal control problem and the generic agent's best response 42
3.2.1 Scalar binary choice case with A = Q = 0 42
3.2.2 General case 45
3.3 Mean field equations 55
3.4 Discussions 60
3.4.1 Indecisive agents 60
3.4.2 Numerical scheme 61
3.4.3 Relation to the discrete choice problems in microeconomics 63
3.5 Simulation results 64
3.5.1 Evolution of the probability distribution and sample paths 64
3.5.2 Influence of the social effect on the group's behavior 64
3.5.3 Confidence zone and influence of the noise intensity on the agents' individual behaviors 66
3.6 Conclusion 69

CHAPTER 4 DEGENERATE COOPERATIVE DYNAMIC COLLECTIVE DISCRETE CHOICE MODEL 71
4.1 Mathematical Model 71
4.2 Exact Social Optimum 72
4.3 Decentralized approximate Social Optimum 74
4.3.1 Person-by-person optimality 74
4.3.2 Asymptotic social optimum 76
4.4 Discussions 81
4.4.1 Uniform population 81

4.4.2 Exact vs. approximate social optima 82
4.4.3 Originality of the convergence proofs and technical assumptions . . . 83
4.5 Simulation Results 84
4.6 Conclusion 85

CHAPTER 5 DYNAMIC COLLECTIVE DISCRETE CHOICE MODEL WITH AN ADVERTISER 87
5.1 Mathematical model 87
5.2 Mean field Stackelberg competition 88
5.2.1 Consumers' game 89
5.2.2 Advertiser's game 92
5.2.3 Approximate Stackelberg solution 96
5.3 Case of uniform initial distribution 98
5.4 Simulation results 102
5.5 Conclusion 102

CHAPTER 6 CONCLUSION 104
6.1 Summary 104
6.2 Future directions 105

REFERENCES 107

APPENDICES 112

LIST OF TABLES

Table 2.1 Degenerate non-cooperative DCDC: exact and approximate Nash strategies . . . 27

Table 3.1 Degenerate and non-degenerate min-LQG optimal control problems . 69

Table 4.1 Cooperative DCDC: Exact and approximate social optima 82

LIST OF FIGURES

Figure 2.1 Degenerate non-cooperative DCDC: the choice process 26
Figure 2.2 Degenerate non-cooperative DCDC: influence of the social effect on the choices' distribution. 38
Figure 2.3 Degenerate non-cooperative DCDC: average state vs. sustainable mean for $Q = 8I_2$. 39
Figure 3.1 Non-degenerate non-cooperative DCDC: the choice process 61
Figure 3.2 Non-degenerate non-cooperative DCDC: distribution's evolution and sample paths. 65
Figure 3.3 Non-degenerate non-cooperative DCDC: influence of the social effect on the choices' distribution. 66
Figure 3.4 Non-degenerate non-cooperative DCDC: different splitting behaviors. 67
Figure 3.5 Non-degenerate non-cooperative DCDC: confidence zones (hatched areas). 68
Figure 3.6 Non-degenerate non-cooperative DCDC: evolution of the Gibbs distribution over a sample path for $\sigma = 6$. The red (resp. blue) downward (resp. upward) arrow is the element of the Gibbs distribution related to $p_1 = -10$ (resp. $p_2 = 10$). 68
Figure 4.1 Comparison of the cooperative and non-cooperative behaviors in the collective discrete choices problems. 85
Figure 4.2 Comparison of the choices' distributions and per agent costs in the cooperative and non-cooperative cases. 86
Figure 5.1 Evolution of the consumers' density in the absence and presence of an advertising effect. 103

LIST OF ABBREVIATIONS AND NOTATIONS

x_i^j	The superscript j refers to the alternative p_j, while the subscript i refers to agent i
$O(.)$	The Big-O in the Bachmann–Landau notation
$o(.)$	The Little-o in the Bachmann–Landau notation
u_{-i}	$(u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_n)$
$[x]_k$	The k-th element of the vector x
I_n	The $n \times n$ identity matrix
$1_{n\times m}$	The $n \times m$ matrix with entries equal to one
$\mathrm{diag}(M_1, \dots, M_l)$	The block-diagonal matrix with main diagonal blocks $M_1, \dots, M_l$
M'	The transpose of the matrix M
Tr(M)	The trace of the matrix M
$M \succ 0$	Positive definite matrix M
$M \succeq 0$	Positive semi-definite matrix M
$\lambda_{\max}(M)$	The largest eigenvalue of the matrix M
$\lambda_{\min}(M)$	The smallest eigenvalue of the matrix M
$\\|x\\|_M$	$\sqrt{\frac{1}{2}x'Mx}$, where x is a vector and $M \succeq 0$
$\otimes$	The Kronecker product
$\mathbb{P}(A)$	The probability of the event A
$\mathbb{E}(X)$	The expectation of the random variable X
$\mathbb{E}(X\|Y)$	The conditional expectation of the random variable X given the σ-algebra generated by the random variable Y
$\mathcal{N}(\mu, \Sigma)$	The Gaussian distribution with mean μ and variance Σ
$\sigma\{X_1, \dots, X_N\}$	The σ-algebra generated by the random variables X_1,...,X_N
$\mathbb{B}(X)$	The Borel σ-algebra on the topological set X
1_X	The indicator function of the set X
Y^X	The set of functions from X to Y
$C(X,Y)$	The set of continuous functions from X to Y
$L_2([0,T],\mathbb{R}^n)$	The Hilbert space of $\mathbb{R}^n$-valued Borel measurable functions f on $[0,T]$, such that $\int_0^T \\|f(t)\\|^2 dt < \infty$. The inner product of $f \in L_2([0,T],\mathbb{R}^n)$ and $g \in L_2([0,T],\mathbb{R}^n)$ is defined by $\langle f, g\rangle_{L_2} = \int_0^T f(t)'g(t)dt$, and the induced norm is denoted $\\|.\\|_{L_2}$

$C^{i,j}(X \times Y)$	The set of real-valued continuous functions $f(x, y)$ on $X \times Y$, such that the first $i-$th partial derivatives with respect to the x variable and the first $j-$th partial derivatives with respect to the y variable are continuous
$\|.\|$	The Euclidean norm
$\|.\|_\infty$	The supremum norm
$\overset{\circ}{X}$	The interior of the set X
$\overline{X}$	The closure of the set X
∂X	The boundary of the set X
X^c	The complement of the set X
$B(c, r)$	The ball of center c and radius r
$\frac{\partial h(t,x)}{\partial x}$	The gradient of the real function h w.r.t. $x \in \mathbb{R}^n$
$\frac{\partial^2 h(t,x)}{\partial x^2}$	The Hessian matrix of the real function h w.r.t. $x \in \mathbb{R}^n$.

DCDC	Dynamic Collective Discrete Choice
DCDCA	Dynamic Collective Discrete Choice with an Advertiser
FP	Fokker-Plank
HJB	Hamilton-Jacobi-Bellman
LQG	Linear Quadratic Gaussian
LQR	Linear Quadratic Regulator
MFG	Mean Field Games
PDE	Partial Differential Equation
ODE	Ordinary Differential Equation
SDE	Stochastic Differential Equation
i.i.d.	Independent and Identically Distributed
CDV	Choice Distribution Vector
CDM	Choice Distribution Matrix

LIST OF APPENDICES

Appendix A PUBLICATIONS 112

CHAPTER 1 INTRODUCTION

Discrete choice problems arise in situations where an individual makes a choice within a finite set of alternatives, such as modes of transportation (Koppelman and Sathi, 2005), entry and withdrawal from the labor market, or residential locations (Bhat and Guo, 2004). These choices are dictated by some personal factors, such as a person's financial status in the residential locations example. In some circumstances, the aggregate behavior of a group of individuals facing the same discrete choice problem contributes to the crystallization of the individual choices, and this is referred to as the social effect. For example, teenagers' decisions to smoke in schools are influenced by some personal factors, as well as by their peers' behaviors (Nakajima, 2007). These situations were studied by Brock and Durlauf under the name of discrete choice with social interaction (Brock and Durlauf, 2001). Our thesis proposes and develops two new classes of discrete choice problems entitled *dynamic collective discrete choice (DCDC)* and *dynamic collective discrete choice with an advertiser (DCDCA)*. The first is concerned with situations where a large number of agents exert an effort over a finite time horizon to move and ultimately settle on one of multiple predefined alternatives. They do so while influenced along the path by the agents population behavior. Throughout the thesis, the population behavior or macroscopic behavior will refer to the evolution of the agents' probability distribution. In the second problem, an additional agent/advertiser acting as a major agent makes some investments to persuade the other (minor) agents/consumers to move towards a specific alternative. In this case, the individual choices are shaped by some personal factors, and the social and advertising effects.

Applications of the DCDC model range from biology, politics, robotics to advertising. Indeed, the DCDC model helps understanding some successful biological collective decision making mechanisms. For example, a group of honey bees searching for a new nectar site has to move from its colony towards one of multiple sites of interests. Even though certain sites can be easier to reach and are more attractive for some bees, following the majority is still a priority to enhance the foraging ability (Seeley et al., 1991; Camazine et al., 1999). In politics, this model can be considered as a mechanistic representation of opinion crystallization in elections (Merrill and Grofman, 1999; Markus and Converse, 1979), where voters continuously update their opinions until forming a final decision regarding who they should vote for. Until voters settle on a candidate, changes in their opinions will involve a mental strain, but at the same time, voters adhering to an individual opinion when it currently deviates from that of the majority will experience some level of discomfort. The DCDCA model is relevant here when one of the candidates makes some investments in the form of electoral campaigns to influence the voters' opinions. The DCDC model is also connected to some navigation situations. For example, a large collection of robots explore an unknown terrain (Vail and Veloso, 2003; Halász et al., 2007; Hsieh et al., 2008), where they must move within a finite time horizon from their initial positions towards one of multiple sites of interest. While moving, each robot optimizes its efforts to remain close to the group and to arrive at the end of the time horizon in the vicinity of one of the sites. The group may split, but the size of the subgroups must remain large enough to perform some collective tasks, such as search, rescue, etc.

1.1 Research objectives

The purpose of the thesis is to develop abstract models for both the DCDC and DCDCA problems. The former is formulated as a dynamic game involving a large number of agents initially spread out in an Euclidean space. These agents must move within a finite time horizon from their initial positions towards one of multiple predefined alternatives. They do so while exerting as little effort as possible, and trying to remain grouped around their average. According to whether they act selfishly or cooperate, the DCDC problem is modeled as respectively a dynamic non-cooperative game or a social optimization problem. The DCDCA problem is, however, formulated by adding an extra "major" agent (Huang, 2010) (called advertiser in this case) to the non-cooperative DCDC game with two alternatives. The advertiser makes some investments to subsequently influence the paths of the other "minor" agents (called consumers in this case) and persuade them to choose a specific alternative among the two. Within this framework, our work aims at answering the following questions:

- What are the optimal strategies that lead the agents to settle with least effort on one of the alternatives under the social and advertising effects?
- What is the necessary information to compute and implement these strategies?
- Does an agent pick an alternative at the outset and commit to it? If not, how does it make its choice?
- Can the agents population macroscopic behavior and the probability distribution of their choices over the alternatives be anticipated?
- Does the population split between the alternatives in a unique or otherwise multiple ways?

1.2 Methodology and related works

1.2.1 Methodology - Mean field games theory

Game theory was formalized in 1944 by Von Neumann and Morgenstern (Von Neumann and Morgenstern, 2007). It studies situations where a group of interacting agents make simultaneous decisions to maximize their utility functions, or minimize their costs. An agent's utility function depends on its decision variable and those of the other agents. The main solution concept in non-cooperative games is the Nash equilibrium, formally defined by John Nash in 1951 (Nash, 1951). It is a set of decisions for all the agents (strategy profile), such that a unilateral deviant behavior is not profitable. Dynamic non-cooperative game theory was developed later. While in static games the decisions are made once, the agents update their strategies over time in dynamic games (Başar and Olsder, 1998). When the updating process is repeated continuously over a time interval, the Nash equilibrium is characterized by a set of coupled Hamilton-Jacobi-Bellman (HJB) equations, whose number is equal to that of the agents. Each equation describes an agent's best-response to the other agents' strategies (Başar and Olsder, 1998). As the size of the population grows, these equations become increasingly difficult, if not impossible, to solve.

In some situations, the agents become weakly coupled as their number increases sufficiently; more precisely, an isolated individual agent's decision has an increasingly negligible influence on the other agents' strategies, while their aggregate behavior considerably shapes the individual decisions. Ultimately, the agents interact anonymously through their *mean field*, i.e. their states' empirical distribution, as in the DCDC and DCDCA problems. The mean field games (MFG) methodology, which is used in this thesis, has proved to be a powerful technique to solve such games. It starts by assuming a continuum of agents, in which case, because of the law of large numbers, the mean field term becomes deterministic, although initially unknown. Unlike the finite population case discussed above, the infinite population Nash equilibrium is characterized by only two coupled partial differential equations, a backward HJB equation and a forward Fokker-Plank (FP) equation. The former characterizes a generic agent's best response to some assumed deterministic mean field term, while the latter propagates the would be mean field when all the agents implement their computed best responses. Consistency requires that sustainable mean field trajectories, if they exist, be replicated in the process. Thus the requirement that limiting equilibria satisfy a system of fixed point equations, herein given by the coupled HJB-FP equations. The corresponding best responses, when applied to the finite population, constitute approximate Nash equilibria (ϵ-Nash equilibria) (Huang et al., 2006, 2007).

Definition 1 ($\epsilon-$Nash Equilibrium). *Consider N agents, a set of strategy profiles $S = S_1 \times \cdots \times S_N$, and for each agent i, a cost function $J_i(u_1, \ldots, u_N)$, for $(u_1, \ldots, u_N) \in S$. A strategy profile $(u_1^*, \ldots, u_N^*) \in S$ is called an ϵ-Nash equilibrium with respect to the costs J_i, if there exists an $\epsilon > 0$, such that for any fixed $1 \leq i \leq N$, for all $u_i \in S_i$, we have $J_i(u_i, u_{-i}^*) \geq J_i(u_i^*, u_{-i}^*) - \epsilon$.*

The MFG theory was introduced in a series of papers by Huang, Caines and Malhamé (Huang et al., 2003, 2006, 2007), and independently by Lasry and Lions (Lasry and Lions, 2006a,b, 2007). Huang *et al.*'s 2003 and 2007 papers formulated the linear quadratic Gaussian (LQG) MFG theory. The general case was developed in (Huang et al., 2006; Lasry and Lions, 2006a,b, 2007). Though the results are similar, the approaches in these works are different. In fact, Huang *et al.* start by solving the game under the infinite size population assumption. Afterwards, they characterize the corresponding strategies when applied by the finite population as approximate Nash strategies ($\epsilon-$Nash). *We follow this approach in the thesis.* By contrast, Lasry and Lions solve the game for a finite number of agents, and claim that when the size of the population increases to infinity, the set of HJB-FP equations describing the agents' best responses and distributions, and whose number is equal to that of the agents, reduce to the aforementioned coupled HJB-FP equations. Cardaliaguet *et. al* show later (Cardaliaguet et al., 2015) that the finite Nash equilibrium converges in "average" to the unique solution of the "master equation". The latter is an equivalent formulation of HJB-FP equations.

The cooperative LQG MFG formulation was developed later in (Huang et al., 2012). Under an infinite population assumption, the authors investigate the structure of the LQG costs to develop decentralized person-by-person optimal strategies (Başar and Olsder, 1998), a weaker solution concept than the social optimum. They show that these strategies, when applied by a finite number of agents, converge to a social optimum as the size of the population increases to infinity.

Definition 2 (Person-by-person optimum). *Consider N agents, a set of strategy profiles $S = S_1 \times \cdots \times S_N$ and a social cost $J_{soc}(u_1, \ldots, u_N)$, $\forall (u_1, \ldots, u_N) \in S$. A strategy profile $(u_1^*, \ldots, u_N^*)$ is said to be person-by-person optimal with respect to the social cost J_{soc}, if for all $i \in \{1, \ldots, N\}$, for all $u_i \in S_i$, $J_{soc}(u_i, u_{-i}^*) \geq J_{soc}(u_i^*, u_{-i}^*)$.*

Definition 3 (Social optimum). *Consider N agents, a set of strategy profiles $S = S_1 \times \cdots \times S_N$ and a social cost $J_{soc}(u_1, \ldots, u_N)$, $\forall (u_1, \ldots, u_N) \in S$. A strategy profile $(u_1^*, \ldots, u_N^*)$ is said to be a social optimum with respect to the social cost J_{soc}, if $J_{soc}(u_i^*, u_{-i}^*) = \inf_{(u_i, u_{-i}) \in S} J_{soc}(u_i, u_{-i})$.*

In the above listed MFG models, each agent makes a tiny contribution to the mean field, which in its turn shapes the decisions of the agents. To account for influential agents, these models were extended in two directions. At first, Huang introduced the major-minor LQG MFG model (Huang, 2010). The general case was developed later in (Nourian and Caines, 2013). The major-minor models are non-cooperative MFGs where a major agent has a considerable impact on the other minor agents. The latter, however, influence each other and the major agent through their mean field. As in the standard MFG models, the minor and major agents make simultaneous decisions. The mean field is no longer deterministic for a continuum of minor agents, but a stochastic process adapted to the major agent's state. Subsequently, Bensoussan *et al.* have formulated a general Stackelberg MFG model (Bensoussan et al., 2014). Even though the influence relations between the agents in Huang and Bensoussan's models are similar, the major agent (or what Bensoussan et *al* call dominating agent) in the latter plays first, and then the minor agents make their decisions. The agents seek in this case a Stackelberg solution (Von Stackelberg, 1934; Başar and Olsder, 1998).

Definition 4 (Stackelberg Solution)**.** *Consider $N+1$ agents, a set of strategy profiles $S = S_0 \times \cdots \times S_N$, and for each agent k, a cost function $J_k(u_0, \ldots, u_N)$, $\forall (u_0, \ldots, u_N) \in S$. Suppose that agent 0 is the dominating agent. A strategy profile $(u_0^*, \ldots, u_N^*) \in S$ is called a Stackelberg solution w.r.t. the costs J_k, if there exists a map T from S_0 to $S_1 \times \cdots \times S_N$, such that for all $u_0 \in S_0$, $T(u_0)$ is a Nash Equilibrium w.r.t. J_k, $k = 1, \ldots, N$, and $u_0^* = \min_{u_0 \in S_0} J_0(u_0, T(u_0))$, with $T(u_0^*) = (u_1^*, \ldots, u_N^*)$.*

Although the MFG methodology constitutes the backbone of this thesis, our results are original with respect to the MFG literature. For a detailed discussion about the differences, we refer the reader to the main contributions in Section 1.3 below.

As discussed at the beginning of this chapter, the DCDC and DCDCA models are related to many problems in microeconomics, elections, robotics and advertising. We discuss this relation in the following paragraphs.

1.2.2 Relation to the discrete choice problems in microeconomics

Discrete choice problems were developed in microeconomics to analyze an individual's behavior facing a set of alternative choices. A first static model was formulated by McFadden in (McFadden, 1974), where an agent chooses among a finite set of alternatives the one that maximizes its utility. Later, Rust (Rust, 1994) introduced a dynamic discrete choice model involving a Markov decision process for each agent. While the social effect is absent in Rust's and McFadden's models, Brock and Durlauf (Brock and Durlauf, 2001) studied a discrete choice problem with *social interactions* modeled as a *static* non-cooperative game, where a large number of agents choose between two alternatives while being influenced by the average of their choices. The authors analyze the model using an approach similar to that of a static MFG, and inspired by statistical mechanics.

Even though Rust's model and the DCDC models of this thesis are dynamic discrete choice models, the agents make in the former repeated choices at each discrete time period, but under no social influence. In our non-degenerate DCDC model, however, they are continuously reassessing the adequacy of their would-be choices and current actions along their random state-space path, up until the end of the control horizon, at which point their ultimate choice of alternative becomes fully crystallized. Besides, the DCDC model includes peer influence. Thus, the DCDC formulation helps in modeling situations where the alternatives are identified as potential points in a suitable state space, for example, physical space in the robotics example, or opinion space in the election example. Moreover, implementation of a given choice involves movement under a social effect towards a final destination state, requiring control effort and constrained by specific dynamics.

Finally, the results in the non-degenerate DCDC case can be shown to have a direct relation with the economic discrete choice models. Indeed, we show in Section 3.4.3 that the DCDC model can be interpreted at each instant as a static discrete choice problem, where the individual costs include an additional term that penalizes myopic decisions.

1.2.3 Relation to opinion dynamics and elections

There have been two trends in modeling opinion dynamics during the past forty years. The first considers a group of agents evaluating a given subject. It tries to understand how they update their opinions, and how the update rules help predicting different behaviors, for instance, occurrence of a consensus, formation of opinion clusters, etc. In the early models (DeGroot, 1974; Lehrer, 1975), for example, each agent updates its opinion by averaging the opinions of the other agents. Under this update rule, it is shown that the agents' opinions converge to a common value. In other words, a consensus occurs. The averaging rule was modified later in many ways (Friedkin and Johnsen, 1990; Hegselmann and Krause, 2002; Jadbabaie et al., 2003; Acemoğlu et al., 2013). The second trend is the voter model (Liggett, 2012), whereby a group of agents choose continuously over a time horizon between two alternatives. At each instant, the probability that an agent changes its current choice of alternative from one to the other depends on the current neighbors' choices.

The first trend is related for example to contexts where the agents give their opinions on the quality of a product. The opinions take in this case values in a connected set. Even though in some cases (Hegselmann and Krause, 2002) they ultimately form clusters around a finite number of values, these values are not defined a priori, but result from the update rules. In the DCDC and DCDCA models, however, the opinions cluster around predefined alternatives that shape the form of the update rules. As for the second trend, the choices are restricted to a finite discrete set and are made at each instant. Although the voter model can model opinion dynamics in elections, it does not capture (as the DCDC and DCDCA do) the transition phase that a voter experiences before making a final decision regarding who he/she should vote for. During this phase, the voters are indecisive. Indeed, they optimize their efforts to steer their opinion state and stay in the comfort zone, i.e. close to the majority opinion state, and to ultimately make a definite choice of a candidate.

1.2.4 Relation to robotics

Far from social sciences, the DCDC model is also related to some mobile robot deployment problems (Halász et al., 2007; Hsieh et al., 2008). According to these models, a group of robots, initially located at some sites, needs to autonomously redistribute itself between the sites according to a desired distribution to achieve some collective tasks. Two main differences exist between these problems and the DCDC model. First, the deployment models do not consider scenarios where the robots are not initially at the sites, and need to travel to settle on these. Moreover, the deployment is enforced by a desired distribution irrespective of the energy expenditure. By contrast, in the case of the DCDC model, the robots move from random initial positions toward the alternatives with least cost. Although their final distribution over the sites is dictated by their dynamics and the location of the sites, the DCDC model provides parameters indicating for each robot the attractiveness of the sites, which can be calibrated to shape the final distribution. Second, unlike in (Halász et al., 2007; Hsieh et al., 2008), a robot's optimal strategy in the DCDC model takes into account a noise process, which represents the random forces that disturb the robot's path, for example.

1.2.5 Relation to the advertising models

Advertising models have been extensively studied in the context of differential games. Among these are the Vidale-Wolfe, the Excess, the Lanchester and the Diffusion models. For a detailed discussion about the differential advertising games, we refer the reader to (Erickson, 1995). Briefly, these games involve a finite number of producers that adjust their advertising rates to control their market shares and increase their profit. This approach does not take into account the consumers' behaviors and their influence on the market shares and profits. It cannot, for example, model situations where the consumers have different a priori preferences towards the products, or where they make their choices of a product under both the advertising and social effects. To account for such cases, the DCDCA model considers both the advertiser and consumers as decision makers in the game. In its current form, however, this model involves only a monopolistic advertising game, where two producers share the market, but only one of them is investing in advertising.

1.3 Main contributions and structure of the thesis

The main contributions of the thesis are divided into two groups. The first is related to the problem definition and research objectives. In other words, it contains the answers to the questions posed in Section 1.1. The second group is related to the techniques used in the proofs of some lemmas and theorems, which are intrinsic to the dynamic games in this work, and differ from those used in the classical MFG theory.

1.3.1 First group of contributions

As discussed in Section 1.1, the main objective of the thesis is to develop mathematical models for the DCDC and DCDCA problems. Accordingly, we introduce a new class of dynamic games and social optimization problems, called “min-LQG”, which involves a large number of agents with linear dynamics. The key feature of these problems is their individual cost, the “min-LQG” cost, a modified version of the LQG one, that includes a choice-related minimization term in its final cost to capture the discrete choice aspect.

For a large population, the discrete choices make the naïve approach to computing the exact solutions of the non-cooperative and cooperative problems intractable. Moreover, the implementation of these solutions requires a significant amount of communication between the agents. Indeed, we show in Sections 2.2 and 4.2 that a naïve method to compute the exact solutions involves solving an ordinary differential equation (ODE), whose dimension increases exponentially with the number of the agents. To implement these strategies, each agent needs to know the initial states of all the agents, and their parameters. Instead, we solve the DCDC and DCDCA problems via the MFG methodology. This allows to develop “simple” solutions, understand how the agents make their choices, and anticipate the aggregate agents’ behavior. More explicitly, we obtain the following results, which constitute the first group of contributions of this thesis.

Non-cooperative DCDC model

- We develop via the MFG methodology a set of strategies that are computationally tractable. Moreover, the implementation of these strategies requires that an agent know only its state, and the probability distributions of the initial states and parameters.

- We show that the mean field based strategies constitute an approximate Nash equilibrium ($\epsilon-$Nash equilibrium).

- In the degenerate case, i.e., in the absence of a noise process in the dynamics, we show that an agent makes its choice of an alternative prior starting to move, and based on its initial state.

- In the non-degenerate case, the agents can no longer commit to a choice from the beginning. Instead, they continuously revise their decisions to account for the risk of making a "myopic" decision in favor of one alternative based on the current state, while being driven to one of the other alternatives by the noise process. Accordingly, we interpret the min-LQG problem at each instant as a static discrete choice problem (McFadden, 1974), where an agent's cost of choosing one of the alternatives includes an additional term that penalizes myopic decisions.

- We show that there exists a one-to-one map between the infinite population Nash equilibria and the fixed points of a finite dimensional map. The latter are the potential probability distributions of the agents' choices over the alternatives. As a result, this one-to-one correspondence reduces the infinite dimensional problem of finding a Nash equilibrium for the limiting game to solving a fixed point for the finite dimensional map. Furthermore, it simplifies the existence proofs and numerical schemes to compute the equilibria.

For a detailed discussion of these contributions, we refer the reader to Sections 2.5 and 3.4. The degenerate non-cooperative DCDC model is introduced in Chapter 2, while Chapter 3 considers the non-degenerate case.

Degenerate cooperative DCDC model

- We develop via the MFG methodology a set of strategies that are computationally tractable. Moreover, the implementation of these strategies requires that an agent know only its state, and the probability distributions of the initial agent states and parameters.

- We show that the mean field based strategies converge to a social optimum as the number of agents increases to infinity.

- As in the degenerate non-cooperative case, we show that an agent makes a choice prior starting to move and based on its initial state.
- We show that there exists a one-to-one map between the infinite population social optima and the fixed points of a finite dimensional map. The latter are the potential probability distributions of the agents' choices over the alternatives. This characterization of the infinite population problem has consequences similar to those in the non-cooperative case.
- We provide a numerical example that illustrates the advantage of the cooperative strategies compared to the non-cooperative ones in evenly allocating the agents to the alternatives, a crucial feature required when considering the robotic example, for instance, to improve the quality of the collective tasks.

For a detailed discussion of these contributions, we refer the reader to Section 4.4. The degenerate cooperative DCDC model is studied in Chapter 4.

DCDCA model

To model the advertising effect, we add in Chapter 5 an advertiser to the degenerate non-cooperative DCDC game, which constitutes then a Stackelberg competition. The main contribution of the DCDCA model is to introduce an advertising model involving two competing alternatives and in which both the consumers and advertiser are part of the game. Our model describes the consumers' individual behaviors, from which it deduces the way the population of consumers splits along the alternatives under both the social and advertising effects. We start by assuming a continuum of consumers, and solve the game via the MFG methodology. When the consumers have sufficiently diverse a priori opinions, we show that a unique Nash equilibrium between the consumers exists. As a result, the advertiser can anticipate the consumers' behavior and make optimal investments. These investments require that the advertiser know only the probability distribution of the a priori opinions. When these investments are applied in real situations of a finite number of consumers, they induce a loss of performance in the advertising process. However, this loss becomes negligible in a large population. For a certain uniform distribution of a priori opinions, we provide an explicit form of the advertiser's optimal investment strategies, and consumers' optimal choices. Moreover, we anticipate the probability distribution of the consumers' choices over the alternatives.

1.3.2 Second group of contributions

The min-LQG final costs (2.2)-(3.2) change drastically the analysis of the non-cooperative DCDC game with respect to the standard non-cooperative MFG literature discussed in Subsection 1.2.1. Indeed, in the LQG MFG theory (Huang et al., 2003, 2007), the authors derive an explicit form of the generic agent's best response, which is a linear feedback policy. In this case, it is sufficient to know the initial mean of the population to solve the mean field equations (coupled HJB-FP equations). In the nonlinear MFG theory (Huang et al., 2006; Lasry and Lions, 2006b; Carmona and Delarue, 2013), however, an explicit solution is not generally possible, and the existence and uniqueness of a solution to the coupled HJB-FP is proved under strong assumptions. For example, in (Huang et al., 2006), the authors assume that the control laws are Liptschiz continuous with respect to the mean field term, and derive the result via Banach's fixed point theorem. In (Carmona and Delarue, 2013; Lasry and Lions, 2006b), the result is shown under a smooth and convex final cost assumption via Schauder's fixed point theorem (Conway, 2013). The non-cooperative min-LQG game lies somewhere between the LQG and the general MFG theories. Indeed, the linear dynamics and quadratic running costs allow one to derive an explicit form of the generic agent's best response, which is nonlinear in this case. Hence, one needs to know the complete distribution of the initial states in order to solve the mean field equations. Moreover, the best responses are not Lipschitz continuous and the final costs are not convex or smooth, which result in a multitude of solutions to the HJB-FP equations. But, thanks to the linear dynamics and quadratic running costs, we construct a one-to-one map between these solutions and the fixed points of a finite dimensional map. The existence of at least one fixed point, equivalently of a solution to the HJB-FP equations, is shown by Brouwer's fixed point theorem (Conway, 2013).

Similarly, the non-convexity and non-smoothness of the final costs in the cooperative DCDC problem require a proof machinery different from that of the LQG MFG (Huang et al., 2012). Indeed, the convergence proofs of the decentralized mean field based strategies to a social optimum are tightly related to the special form of the min-LQG final costs, see Lemma 6, Theorem 14, and Section 4.4.3. Finally, unlike in (Bensoussan et al., 2014, 2015), the advertiser's optimal control problem (5.12) in the DCDCA game involves a dynamic constraint that depends non-linearly on his/her state and the consumers' macroscopic behavior. An explicit Stackelberg solution is thus possible only in some special cases, for instance, when the consumers' a priori opinions have a certain uniform distribution.

CHAPTER 2 DEGENERATE NONCOOPERATIVE DYNAMIC COLLECTIVE DISCRETE CHOICE MODEL

We formulate the non-cooperative degenerate DCDC problem in Section 2.1 as a dynamic non-cooperative game, where an agent's state evolves according to a controlled linear ODE. The agents are cost coupled through their average state. In Section 2.2, we consider the problem of finding an exact Nash equilibrium for the game. We show that there exists a one-to-one map between the equilibria and the fixed points of a finite dimensional map. The latter is defined on a finite set of size increasing exponentially with the number of agents. A Nash equilibrium may not exist. Moreover, the naïve approach to computing an equilibrium, if it exists, becomes computationally intractable as the number of agents increases, and its implementation requires a significant amount of communication between the agents. Instead, we solve the game via the MFG methodology. We start by assuming a continuum of agents for which one can ascribe a deterministic, although initially unknown, mean trajectory. In order to compute its best response to this trajectory, a generic agent solves an optimal tracking problem, which constitutes the subject of Section 2.3. The problem of determining a sustainable mean trajectory, that is, a trajectory that is replicated by the mean trajectory of the agents when they optimally respond to it, is treated in Section 2.4. We show in that section that there exists a one-to-one map between the sustainable mean trajectories and the fixed points of a finite dimensional map. The latter are the potential probability distributions of the agents' choices over the alternatives. We derive conditions for the existence of a fixed point, equivalently of a sustainable mean trajectory, to hold. We discuss the results and contributions of the degenerate non-cooperative DCDC problem in Section 2.5. In Section 2.6, the model is further generalized to include initial preferences towards the alternatives, and nonuniform dynamics. In Section 2.7, we show that the mean field based strategies constitute an $\epsilon-$Nash equilibrium (See Definition 1) when applied by a finite number of agents. Finally, Section 2.8 provides some numerical simulation results, and Section 2.9 concludes this chapter.

2.1 Mathematical model

We formulate the degenerate non-cooperative DCDC problem as a dynamic non-cooperative game involving a large number N of agents with the following linear dynamics,

$$\frac{d}{dt}x_i = Ax_i + Bu_i, \qquad 1 \leq i \leq N, \tag{2.1}$$

where $x_i(t) \in \mathbb{R}^n$ is the state of agent i at time t and $u_i(t) \in \mathbb{R}^m$ its control input. In a large population, the knowledge of all the agents' initial states $x_i(0)$, $1 \leq i \leq N$, is not realistic and requires a complex communication network among the agents. It is thus convenient to think of the initial states as realizations of random variables resulting from a common probability distribution function in a collection of independent experiments. Formally, we assume that $x_i(0)$, $1 \leq i \leq N$, are independent and identically distributed (i.i.d.) random vectors on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with distribution equal to P_0 and finite second moment $\mathbb{E}\|x_i(0)\|^2 < \infty$. An agent i is associated with the following individual "min-LQG" cost functional,

$$J_i\left(u_i, \bar{x}\right) = \mathbb{E}\left(\int_0^T \left\{\|x_i - \bar{x}\|_Q^2 + \|u_i\|_R^2\right\} dt + \min_{1\leq j\leq l} \|x_i(T) - p_j\|_M^2\right), \tag{2.2}$$

where $Q \succeq 0$, $R \succ 0$, M is a large positive definite matrix in the sense that will become clear later (See Theorem 3 below), and $\|z\|_L = \sqrt{\frac{1}{2}z'Lz}$ for any vector z and matrix $L \succeq 0$. Along the path, the running cost forces the agents to exert as little control effort as possible, and to stay close to the population's average state $\bar{x} := 1/N \sum_{j=1}^N x_i$. At the end of the time horizon $T > 0$, each agent should reach the vicinity of one of the alternatives p_j, $1 \leq j \leq l$. Otherwise, it is strongly penalized by the final cost. Indeed, Theorem 3 below shows that the probability that a cost-minimizing agent is close to one of the alternatives can be arbitrary increased by making the final cost's coefficient M large enough. Hence, the overall individual cost captures the problem faced by each agent of deciding between a finite set of alternatives, while trying to remain close to the population's average trajectory. For each agent, we define the set of admissible control strategies,

$$\begin{aligned}\mathcal{U} =& \Big\{ u \in (\mathbb{R}^m)^{\Omega\times[0,T]} | u \text{ is } \sigma\left\{x_1(0), \ldots, x_N(0)\right\} \otimes \mathbb{B}([0,T]) - \text{measurable}, \\ & \text{and } \mathbb{E}\int_0^T \|u\|^2 dt < \infty \Big\}. \end{aligned} \tag{2.3}$$

2.2 Exact Nash equilibria

We consider in this section the problem of finding a Nash equilibrium for the non-cooperative DCDC game defined by (2.1) and (2.2). We assume here that each agent observes its state and the states of the other agents. In the following, we construct a one-to-one map between the Nash equilibria and the fixed points of a finite dimensional map. This allows us to analyze the existence and computation of an equilibrium.

Let us consider (u_i^*, u_{-i}^*) a Nash equilibrium with respect to the costs (2.2). We denote by

(x_i^*, x_{-i}^*) the corresponding states. At first, we show that the strategy u_i^* is the optimal control law of a linear quadratic regulator (LQR) problem. Moreover, the strategy profile (u_i^*, u_{-i}^*) is a Nash equilibrium for some LQR game. An agent i, $1 \leq i \leq N$, computes its best response to u_{-i}^* by solving the following degenerate min-LQG optimal control problem,

$$\inf_{u_i \in \mathcal{U}} \bar{J}_i(u_i, \bar{x}_{-i}^*, x_i(0)) = \inf_{u_i \in \mathcal{U}} \left(\int_0^T \left\{ \left\| x_i - \frac{1}{N} x_i - \bar{x}_{-i}^* \right\|_Q^2 + \|u_i\|_R^2 \right\} dt + \min_{1 \leq j \leq l} \|x_i(T) - p_j\|_M^2 \right), \tag{2.4}$$

subject to agent i's dynamics, where $\bar{x}_{-i}^* = \frac{1}{N} \sum_{k=1, k \neq i}^N x_k^*$.

The cost $\bar{J}_i$ is similar to (2.2), but we drop the expectation, since agent i observes the initial states of the other agents. Moreover, the others' states captured by $\bar{x}_{-i}^*$ are assumed known in this cost. In the following, we show that there exist l $\bar{x}_{-i}^*$-dependent basins of attraction, each corresponding to one of the alternatives. If agent i is initially inside one of the basins, then it moves under its best response to u_{-i}^* towards the corresponding alternative.

The cost function $\bar{J}_i$ defined in (2.4) can be written as the minimum of l LQR costs, each corresponding to one of the alternatives, as follows,

$$\bar{J}_i(u_i, \bar{x}_{-i}^*, x_i(0)) = \min_{1 \leq j \leq l} \bar{J}_i^j(u_i, \bar{x}_{-i}^*, x_i(0)), \tag{2.5}$$

$$\text{where } \bar{J}_i^j(u_i, \bar{x}_{-i}^*, x_i(0)) = \int_0^T \left\{ \left\| x_i - \frac{1}{N} x_i - \bar{x}_{-i}^* \right\|_Q^2 + \|u_i\|_R^2 \right\} dt + \|x_i(T) - p_j\|_M^2. \tag{2.6}$$

Moreover, we have that $\inf\limits_{u_i \in \mathcal{U}} \min\limits_{1 \leq j \leq l} \bar{J}_i^j(u_i, \bar{x}_{-i}^*, x_i(0)) = \min\limits_{1 \leq j \leq l} \inf\limits_{u_i \in \mathcal{U}} \bar{J}_i^j(u_i, \bar{x}_{-i}^*, x_i(0))$. Assuming a full state feedback, this equality implies that the optimal control law u_i^* of (2.5) is the optimal control law of the LQR problem with the least optimal cost. Explicitly, $u_i^* = u_i^k$ if $\bar{J}_i^k(u_i^k, \bar{x}_{-i}^*, x_i(0)) = \min\limits_{1 \leq j \leq l} \bar{J}_i^j(u_i^j, \bar{x}_{-i}^*, x_i(0))$, where u_i^j is the optimal control law of the LQR problem with cost functional $\bar{J}_i^j$.

The optimal value of $\bar{J}_i^j$, $1 \leq j \leq l$, is a function of $\bar{x}_{-i}^*$ and $x_i(0)$. This suggests to partition the space of initial states $\mathbb{R}^n$ into l $\bar{x}_{-i}^*$-dependent regions $D_i^j(\bar{x}_{-i}^*)$, $1 \leq j \leq l$, such that if agent i's initial state is inside $D_i^j(\bar{x}_{-i}^*)$, then the LQR problem corresponding to $\bar{J}_i^j$ is the less costly, and $u_i^* = u_i^j$. We recall the optimal value and the corresponding optimal control law (feedback policy) of $\bar{J}_i^j$ (Anderson and Moore, 2007),

$$\bar{J}_{i*}^j(\bar{x}_{-i}^*, x_i(0)) = \frac{1}{2} x_i(0)' \bar{\Gamma}(0) x_i(0) + \beta_i^j(0)' x_i(0) + \delta_i^j(0) \tag{2.7}$$

$$u_i^j(t, x, \bar{x}_{-i}^*) = -R^{-1} B' \left(\bar{\Gamma}(t) x + \beta_i^j(t) \right), \tag{2.8}$$

where $\bar{\Gamma}$, β_i^j and δ_i^j are respectively matrix-, vector-, and real-valued functions satisfying the following backward propagating differential equations:

$$\frac{d}{dt}\bar{\Gamma} = \bar{\Gamma}BR^{-1}B'\bar{\Gamma} - \bar{\Gamma}A - A'\bar{\Gamma} - Q\left(1-\frac{1}{N}\right)^2, \qquad \bar{\Gamma}(T) = M \tag{2.9}$$

$$\frac{d}{dt}\beta_i^j = \left(\bar{\Gamma}BR^{-1}B' - A'\right)\beta_i^j + Q\left(1-\frac{1}{N}\right)\bar{x}_{-i}^*, \qquad \beta_i^j(T) = -Mp_j \tag{2.10}$$

$$\frac{d}{dt}\delta_i^j = \frac{1}{2}(\beta_i^j)'BR^{-1}B'\beta_i^j - \left\|\bar{x}_{-i}^*\right\|_Q^2, \qquad \delta_i^j(T) = \|p_j\|_M^2. \tag{2.11}$$

The basins of attraction, defined in equation (2.12) below, are thus regions in $\mathbb{R}^n$ delimited by hyperplanes that depend on $\bar{x}_{-i}^*$.

$$\begin{aligned} D_i^j(\bar{x}_{-i}^*) &= \left\{x_0 \in \mathbb{R}^n | J_{i*}^j(\bar{x}_{-i}^*, x_0) \leq J_{i*}^k(\bar{x}_{-i}^*, x_0), \forall 1 \leq k \leq l\right\} \\ &= \left\{x_0 \in \mathbb{R}^n | \left(\beta_i^j(0) - \beta_i^k(0)\right)' x_0 + \delta_i^j(0) - \delta_i^k(0) \leq 0, \forall 1 \leq k \leq l\right\}. \end{aligned} \tag{2.12}$$

On the boundaries of the basins of attraction, two or more LQR problems have equal optimal costs. As a result, an agent initially on a boundary has multiple best responses. To avoid this situation, we make the following assumption.

Assumption 1. *Conventionally, we assume that if $x_i(0) \in \cap_{m=1}^k D_i^{j_m}(\bar{x}_{-i}^*)$, for some $j_1 < \cdots < j_k$, then agent i goes towards p_{j_1}.*

Under Assumption 1, agent i has a unique best response to u_{-i}^*,

$$u_i^*(t,x) = -R^{-1}B'\left(\bar{\Gamma}(t)x + \beta_i^j(t)\right), \qquad \text{if } x_i(0) \in D_i^j(\bar{x}_{-i}^*). \tag{2.13}$$

Given the initial conditions, equation (2.13) shows that an agent's best response is the optimal control law of the LQR problem that corresponds to the alternative chosen by this agent. Moreover, equations (2.13) and (2.9)-(2.11), for $1 \leq i \leq N$, imply that (u_i^*, u_{-i}^*) is a Nash equilibrium for the game defined by the N agents, but with the LQR costs that correspond to the chosen alternatives. Explicitly, if $d = (d_1, \ldots, d_N) \in \{p_1, \ldots, p_l\}^N$ are the chosen alternatives (i.e., agent i chooses d_i), then (u_i^*, u_{-i}^*) is a Nash equilibrium with respect to the following LQR costs, $1 \leq i \leq N$,

$$\bar{J}_i^{(d)}(u_i, u_{-i}) = \int_0^T \left\{\|x_i - \bar{x}\|_Q^2 + \|u_i\|_R^2\right\} dt + \|x_i(T) - d_i\|_M^2. \tag{2.14}$$

To construct the one-to-one map between the Nash equilibria and the fixed points of a finite dimensional map, let us define for each $d \in \{p_1, \ldots, p_l\}^N$ the non-cooperative LQR game with the individual costs $\bar{J}_i^{(d)}$, $1 \leq i \leq N$, and denote it by *Game*(d). We assume that for each d, *Game*(d) has a unique Nash equilibrium $Nash(d)$. We discuss this assumption in Remark 1 below. We define the finite dimensional map $F_{Nash} = F_2 \circ F_1$, where F_1 and F_2 are defined as follows. If we assign for each agent i an alternative d_i, then F_1 maps the assignment profile $d = (d_1, \dots, d_N)$ to the unique Nash equilibrium of the LQR game $Game(d)$ that corresponds to d. Formally, $F_1 : \{p_1, \dots, p_l\}^N \to \mathcal{U}^N$, such that $F_1(d) = Nash(d)$. On the other hand, F_2 maps a strategy profile $(u_1, \dots, u_N)$ to the choices made by the N agents, when each agent i optimally responds to the other agents' strategies u_{-i}. Formally, $F_2 : \mathcal{U}^N \to \{p_1, \dots, p_l\}^N$, such that $F_2(u_1, \dots, u_N) = (d_1, \dots, d_N)$, with d_i is the choice of agent i under its best response (2.13) to u_{-i}. We summarize the aforementioned functions in the following diagram,

$$\begin{array}{ccccc} \{p_1, \dots, p_l\}^N & \xrightarrow{F_1} & \mathcal{U}^N & \xrightarrow{F_2} & \{p_1, \dots, p_l\}^N \\ d & \longmapsto & F_1(d) = Nash(d) & \longmapsto & F_{Nash}(d) = F_2 \circ F_1(d). \end{array} \tag{2.15}$$

Theorem 1 (Exact Nash equilibrium)**.** *Assume that* $\forall d \in \{p_1, \dots, p_l\}^N$*,* $Game(d)$ *has a unique Nash equilibrium* $Nash(d)$*. Then,* (u_i^*, u_{-i}^*) *is a Nash equilibrium for the degenerate non-cooperative DCDC game if and only if* $(u_i^*, u_{-i}^*) = Nash(d)$*, with* d *is a fixed point of* F_{Nash}*.*

Proof. Let (u_i^*, u_{-i}^*) be a Nash equilibrium for the DCDC game, and $d = (d_1, \dots, d_N)$ be the chosen alternatives. Then, (u_i^*, u_{-i}^*) is the unique Nash equilibrium of the LQR game $Game(d)$. Furthermore, an agent i that optimally responds to u_{-i}^* will go towards d_i. Thus, d is a fixed point of F_{Nash}. Conversely, let $d = (d_1, \dots, d_N)$ be a fixed point of F_{Nash}. We prove that $(u_i^*, u_{-i}^*) := Nash(d)$ is a Nash equilibrium for the DCDC game. Fix $1 \le i \le N$. (u_i^*, u_{-i}^*) is a Nash equilibrium for $Game(d)$, therefore, $\inf_{u_i \in \mathcal{U}} \bar{J}_i^{(d)}(u_i, u_{-i}^*) = \bar{J}_i^{(d)}(u_i^*, u_{-i}^*)$. By the definition of $\bar{J}_i^{(d)}$, $\bar{J}_i^{(d)}(u_i^*, u_{-i}^*) \ge J_i(u_i^*, \bar{x}_i^*)$, where $\bar{x}_i^*$ is the population's average state under (u_i^*, u_{-i}^*). Moreover, $F_2(u_i^*, u_{-i}^*) = d$, since d is a fixed point of F_{Nash}. This means that the choice of agent i under its best response to u_{-i}^* is d_i. Hence, $\inf_{u_i \in \mathcal{U}} J_i(u_i, 1/N x_i + \bar{x}_{-i}^*) = \inf_{u_i \in \mathcal{U}} \bar{J}_i^{(d)}(u_i, u_{-i}^*)$. Thus, $\inf_{u_i \in \mathcal{U}} J_i(u_i, 1/N x_i + \bar{x}_{-i}^*) \ge J_i(u_i^*, \bar{x}_i^*)$, and (u_i^*, u_{-i}^*) is a Nash equilibrium for the DCDC game. □

According to Theorem 1, there exists a one-to-one map between the Nash equilibria of the degenerate non-cooperative DCDC game and the fixed points of a finite dimensional map F_{Nash}. But, this map is defined on the discrete set $\{p_1, \dots, p_l\}^N$. Hence, a fixed point of F_{Nash}, equivalently a Nash equilibrium, may not exist. A naïve method to compute an equilibrium, if it exists, is to run a search for a fixed point over the set $\{p_1, \dots, p_l\}^N$ of size l^N. This operation becomes quickly intractable as the size of the population becomes large. Furthermore, the implementation of the Nash strategies requires a significant amount of communication between the agents. Indeed, each agent needs to know its state and the initial states of all the other agents. Instead, we solve the game in the following sections via the MFG methodology.

Remark 1. *We assume in Theorem 1 that Game(d) has a unique Nash equilibrium for each d. Indeed, one can derive a sufficient condition for this assumption to hold as follows (Başar and Olsder, 1998, Theorem 6.12). Using Pontryagin's maximum principle, an agent's best response is described by its state and co-state equations. By concatenating these equations, we obtain a pair of coupled forward-backward linear ODE's. The existence and uniqueness of a solution to these equations is guarantied under a condition that does not depend on d. This condition involves the existence and uniqueness of a solution to a Riccati equation that does not depend on d.*

2.3 The mean field games methodology and the generic agent's best response

Following the MFG methodology (Huang et al., 2006), this section starts by assuming a continuum of agents. Accordingly, we assume that the mean field term $\bar{x}$ is deterministic and known throughout this section. The problem of determining $\bar{x}$ is treated in Section 2.4 below. The agents, which are cost coupled through $\bar{x}$, become decoupled under an infinite size population assumption. Hence, a generic (representative) agent of state x and control input u computes its best response to $\bar{x}$ by solving the following degenerate min-LQG optimal control problem,

$$\inf_{u \in \mathcal{U}} J\left(u, \bar{x}, x(0)\right) = \inf_{u \in \mathcal{U}} \left(\int_0^T \left\{ \|x - \bar{x}\|_Q^2 + \|u\|_R^2 \right\} dt + \min_{1 \le j \le l} \|x(T) - p_j\|_M^2 \right) \tag{2.16}$$

$$\text{s.t. } \frac{d}{dt} x = Ax + Bu, \tag{2.17}$$

where the generic agent's initial state $x(0)$ is a random variable with distribution equal to P_0. Similarly to Section 2.2, there exist l $\bar{x}$-dependent basins of attraction, each corresponding to one of the alternatives. If the generic agent is initially inside one of the basins, then it goes towards the corresponding alternative. We define the basins of attraction as follows,

$$\begin{aligned} D^j(\bar{x}) &= \left\{ x_0 \in \mathbb{R}^n | J_*^j(\bar{x}, x_0) \le J_*^k(\bar{x}, x_0), \forall 1 \le k \le l \right\} \\ &= \left\{ x_0 \in \mathbb{R}^n | \left(\beta^j(0) - \beta^k(0) \right)' x_0 + \delta^j(0) - \delta^k(0) \le 0, \forall 1 \le k \le l \right\}, \end{aligned} \tag{2.18}$$

$$\text{where } J_*^j(\bar{x}, x(0)) = \frac{1}{2}x(0)'\Gamma(0)x(0) + \beta^j(0)'x(0) + \delta^j(0), \tag{2.19}$$

$$\frac{d}{dt}\Gamma = \Gamma B R^{-1} B'\Gamma - \Gamma A - A'\Gamma - Q, \qquad \Gamma(T) = M \tag{2.20}$$

$$\frac{d}{dt}\beta^j = \left(\Gamma B R^{-1} B' - A'\right)\beta^j + Q\bar{x}, \qquad \beta^j(T) = -Mp_j \tag{2.21}$$

$$\frac{d}{dt}\delta^j = \frac{1}{2}(\beta^j)' B R^{-1} B'\beta^j - \|\bar{x}\|_Q^2, \qquad \delta^j(T) = \|p_j\|_M^2. \tag{2.22}$$

Assumption 2. *Conventionally, we assume that if $x(0) \in \cap_{m=1}^k D^{j_m}(\bar{x})$, for some $j_1 < \cdots < j_k$, then the generic agent goes towards p_{j_1}.*

Under Assumptions 3 and 5 in Sections 2.4 and 2.6, this convention does not affect the analysis in the remainder of this chapter. We summarize the above analysis in the following theorem.

Theorem 2 (The generic agent's best response)**.** *Under Assumption 2, the degenerate min-LQG optimal control problem* (2.16)-(2.17) *has a unique optimal control law (best response to $\bar{x}$)*

$$u_*(t, x, \bar{x}, x(0)) = u_*^j(t, x, \bar{x}, x(0)) := -R^{-1}B'\left(\Gamma(t)x + \beta^j(t)\right) \quad \text{if } x(0) \in D^j(\bar{x}), \tag{2.23}$$

where Γ, β^j, and δ^j are the unique solutions of (2.20)-(2.22).

As mentioned in Section 2.1, the final cost forces an agent to be close to one of the alternatives. Indeed, we show in the following theorem that the probability that an agent reaches an arbitrary small neighborhood of one of the alternatives increases with M. Furthermore, a generic agent cannot expect to reach exactly the alternatives, unless the optimal state at time T, $y_*(T)$, of the following optimal control problem is one of the alternatives,

$$\begin{aligned} J_0(u, \bar{x}, x(0)) &= \int_0^T \left\{\|y - \bar{x}\|_Q^2 + \|u\|_R^2\right\} dt \\ \text{s.t. } \frac{d}{dt}y &= Ay + Bu, \qquad y(0) = x(0). \end{aligned} \tag{2.24}$$

The result is proved for tracked paths $\bar{x}$ that are uniformly bounded with respect to M, a property that is shown to hold later in Theorem 4 for the desired tracked paths (sustainable mean trajectories).

Theorem 3 (Well-defined discrete choice problem)**.** *Suppose that the pair (A, B) is controllable and for each $M \succ 0$, the generic agent is optimally responding to a path $\bar{x}$, with x_**

is its optimal state. Assume that the path $\bar{x}$ is uniformly bounded with respect to M for the semi-norm $\left(\int_0^T \|.\|_Q^2 \mathrm{dt}\right)^{\frac{1}{2}}$. *Then, the following statements hold:*

(i) For any $\epsilon > 0$,

$$\mathbb{P}\left(\bigcap_{j=1}^{l}\{\|x_*(T)-p_j\|>\epsilon\}\right)=O\left(\frac{1}{\epsilon^2\lambda_{\min}(M)}\right). \tag{2.25}$$

(ii) The generic agent reaches at time T the alternative p_j if and only if $p_j = y_*(T)$.

Proof.

(i) It is sufficient to show that the expectation of the optimal cost of (2.16) $\mathbb{E}J_*(\bar{x}, x(0)) \leq K$, for some $K > 0$ independent of M. The result is then a direct consequence of Chebyshev's inequality (Durrett, 2010, Theorem 1.6.4)

$$\mathbb{P}\left(\min_{1\leq j\leq l}\|x_*(T)-p_j\|>\epsilon\right)\leq\frac{1}{\epsilon^2}\mathbb{E}\left(\min_{1\leq j\leq l}\|x_*(T)-p_j\|^2\right)\leq\frac{2}{\lambda_{\min}(M)\epsilon^2}\mathbb{E}J_*(\bar{x},x(0)). \tag{2.26}$$

The couple (A, B) is controllable, therefore, the following control law transfers the generic agent's state $\tilde{x}$ from $x(0)$ at time 0 to p_1 at time T (Rugh, 1996, Theorem 9.2),

$$\tilde{u}(t)=-B'e^{A't}\left(W^{-1}(0,T)\left(p_1-e^{AT}x(0)\right)\right)\in\mathcal{U}, \tag{2.27}$$

where W is the Gramian of (A, B). Following $\mathbb{E}\|x(0)\|^2 < \infty$, the term $\mathbb{E}\int_0^T \|\tilde{u}\|_R^2 dt$ is bounded by a constant $K_1 > 0$ that does not depend on M. By implementing $\tilde{u}$ in the generic agent's dynamics (2.17), and noting that the path $\bar{x}$ is uniformly bounded w.r.t. M, one can show that the generic agent's state $\tilde{x}$ satisfies $\mathbb{E}\int_0^T \|\tilde{x}-\bar{x}\|_Q^2 < K_2$, where $K_2 > 0$ does not depend on M. By the optimality of J_* and the definition of $\tilde{u}$, we obtain that

$$\mathbb{E}J_*(\bar{x},x(0))\leq\mathbb{E}J(\tilde{u},\bar{x},x(0))=\mathbb{E}\int_0^T\{\|\tilde{x}-\bar{x}\|_Q^2+\|\tilde{u}\|_Q^2\}dt\leq K_1+K_2. \tag{2.28}$$

This proves the first point.

(ii) Suppose that the generic agent reaches at time T the alternative p_j while minimizing its individual cost, that is, $x_*(T) = p_j$. The optimal cost is then equal to $J_*^j(\bar{x}, x(0))$, where J_*^j is defined in (2.19). We have, for all $u \in \mathcal{U}$, $J^j(u, \bar{x}, x(0)) \geq J_0(u, \bar{x}, x(0))$, with J_0 defined in (2.24). Under the optimal control law u_*^j of J^j, the generic agent reaches p_j. Therefore, $J^j(u_0, \bar{x}, x(0)) = J_0(u_0, \bar{x}, x(0)) = \min_u J_0(u, \bar{x}, x(0))$, where u_0

is the optimal control law of J_0. The uniqueness of an LQR optimal control law implies that $u_*^j = u_0$. Therefore, $p_j = y_*(T)$.

We now prove the sufficient condition. We have $p_j = y_*(T)$ for some $j \in \{1, \dots, l\}$. Moreover, $J^j(u, \bar{x}, x(0)) \geq J_0(u, \bar{x}, x(0))$ for all $u \in \mathcal{U}$. We have $\min_u J_0(u, \bar{x}, x(0)) = J_0(u_0, \bar{x}, x(0)) = J^j(u_0, \bar{x}, x(0))$. Therefore, the optimal control of J^j is u_0. Hence, the agent reaches p_j.

□

2.4 Mean field equations

Having computed a generic agent's best response to an arbitrary path $\bar{x}$, we consider in this section the problem of determining a sustainable $\bar{x}$, i.e., a trajectory that is replicated by the average state of an infinite number of agents when they optimally respond to it. A generic agent's optimal state under its best response (2.23) to a deterministic path $\bar{x}$ is a random variable adapted to its initial condition. Moreover, for any finite population, the agents' initial states are i.i.d.. Hence, the strong Law of Large Numbers (Durrett, 2010) suggests that the average state of an infinite number of agents is equal to the mathematical expectation of a generic agent's optimal state. Accordingly, a sustainable path $\bar{x}$ must satisfy the following mean field equations,

$$\frac{d}{dt}x_* = Ax_* + Bu_*(t, x_*, \bar{x}, x(0)), \qquad x_*(0) = x(0), \tag{2.29}$$

$$\mathbb{E}x_* = \bar{x}, \tag{2.30}$$

where x_* is the generic agent's optimal state under its best response u_* (2.23) to $\bar{x}$. Equations (2.29)-(2.30) constitute a nonlinear degenerate Mckean-Vlasov equation (Huang et al., 2006). Indeed, (2.29) is an ODE with random initial condition, where the right-hand side depends on the probability distribution of the solution x_* via the constraint (2.30).

Solving (2.29)-(2.30) directly is not an easy task. Alternatively, we derive in Lemma 1 below an equivalent representation of a solution $\bar{x}$ of (2.29)-(2.30). It consists of two forward-backward ODE's (2.31)-(2.32). These equations are coupled in the co-state boundary condition $\bar{q}(T)$ through what we call the "Choice distribution vector (CDV)" $\lambda(\bar{x})$. The CDV is a l dimensional probability vector, where the $k-$th element is the probability that the generic agent optimally responding to $\bar{x}$ is at time T closer to p_k than the other alternatives. Equations (2.31)-(2.32) are obtained by aggregating the state and co-state equations of a generic agent optimally responding to $\bar{x}$. The advantage of this new representation is that if one considers the CDV in the $\bar{q}(T)$ boundary condition as a parameter (say any l dimensional probability vector λ), then equations (2.31)-(2.32) become totally linear. As a result, they can be treated as the state and co-state equations of a LQR problem (2.47) that have a unique explicit solution (2.44) parametrized by λ. This implies that the solutions $\bar{x}$ of the mean field equations lies in the family of paths (2.44) parametrized by the l dimensional probability vectors. Conversely, in order to have a path $\bar{x}^\lambda$ parametrized by some candidate λ be a solution of (2.29)-(2.30), consistency requires that λ be equal to the associated CDV when the generic agent optimally responds to $\bar{x}^\lambda$. This is equivalent to requiring that λ be a fixed point of the finite dimensional map F defined by (2.45). Indeed, F maps a probability vector λ to the CDV $\lambda(\bar{x}^\lambda)$ when the generic agent optimally responds to $\bar{x}^\lambda$.

In effect, we establish a one-to-one map between the solutions of (2.29)-(2.30), equivalently the sustainable mean trajectories, and the fixed points of a finite dimensional map. Later, we show that there exists at least one fixed point, equivalently one sustainable path. We make the following technical assumption to exclude the case where the probability that a generic agent is initially on the boundaries of the basins of attraction is nonzero.

Assumption 3. *We assume that the P_0-measure of hyperplanes in $\mathbb{R}^n$ is zero.*

Assumption 3 is satisfied if the initial distribution P_0 is absolutely continuous with respect to the Lebesgue measure for example.

Lemma 1 (Equivalent representation)**.** *Under Assumption 3, a path $\bar{x}$ satisfies the mean field equations* (2.29)-(2.30) *if and only if it satisfies the following forward-backward ODE's,*

$$\frac{d}{dt}\bar{x} = A\bar{x} - BR^{-1}B'\bar{q}, \qquad \bar{x}(0) = \mathbb{E}x(0), \tag{2.31}$$

$$\frac{d}{dt}\bar{q} = -A'\bar{q}, \qquad \bar{q}(T) = M\left(\bar{x}(T) - \sum_{j=1}^{l}\lambda_j(\bar{x})p_j\right), \tag{2.32}$$

with $\lambda(\bar{x}) = (\lambda_1(\bar{x}), \dots, \lambda_l(\bar{x})) = \left(P_0\left(D^1(\bar{x})\right), \dots, P_0\left(D^l(\bar{x})\right)\right)$.

Proof. Consider $\bar{x}$ a solution of (2.29)-(2.30) and x_* the generic agent's optimal state corresponding to $\bar{x}$, and that satisfies (2.29). To prove that $\bar{x}$ satisfies (2.31)-(2.32), we derive at first a maximum principle for the degenerate min-LQG problem. If $x_*(0) \in D^j(\bar{x})$, then $u_* = u_*^j = -R^{-1}B'q_*^j$, where u_*^j and q_*^j are the optimal control law and co-state of the LQR problem that corresponds to p_j. The co-state satisfies the following backward ODE,

$$\frac{d}{dt}q_*^j = -A'q_*^j + Q(x_* - \bar{x}), \qquad q_*^j(T) = M(x_*(T) - p_j). \tag{2.33}$$

Hence, under Assumption 3, the optimal state x_* and co-state $q_* := \sum_{j=1}^{l} q_*^j 1_{D^j(\bar{x})}(x_*(0))$ of the degenerate min-LQG problem satisfy P_0-a.s the following forward-backward ODE's, which constitute the maximum principle of the degenerate min-LQG problem,

$$\frac{d}{dt}x_* = Ax_* - BR^{-1}R'q_*, \qquad x_*(0) = x(0) \tag{2.34}$$

$$\frac{d}{dt}q_* = -A'q_* + Q(x_* - \bar{x}), \qquad q_*(T) = M\left(x_*(T) - \sum_{j=1}^{l} p_j 1_{D^j(\bar{x})}(x_*(0))\right) \tag{2.35}$$

$$u_* = -R^{-1}B'q_*. \tag{2.36}$$

By taking the expectation of the right- and left-hand sides of (2.34)-(2.35), and noting that $\mathbb{E}x_* = \bar{x}$, we obtain that $\bar{x}$ satisfies (2.31)-(2.32), with $\bar{q} = \mathbb{E}q_*$.

Conversely, consider $(\bar{x}, \bar{q})$ satisfying (2.31)-(2.32), and x_* the generic agent's optimal state under its best response u_* to $\bar{x}$. In the following, we show that $\bar{x} = \mathbb{E}x_*$, or equivalently $\bar{x}$ is a solution of the mean field equations. Following the discussion above, x_* and its co-state q_* satisfy (2.34)-(2.35). By taking the expectation on both sides of equations (2.34)-(2.35), and subtracting the resulting equations from (2.31)-(2.32), we get,

$$\frac{d}{dt}e_x = Ae_x - BR^{-1}B'e_q, \qquad e_x(0) = 0, \tag{2.37}$$

$$\frac{d}{dt}e_q = -A'e_q + Qe_x \qquad e_q(T) = Me_x(T), \tag{2.38}$$

where $e_x = \mathbb{E}x_* - \bar{x}$ and $e_q = \mathbb{E}q_* - \bar{q}$. Equations (2.37)-(2.38) describe the optimal state e_x and co-state e_q of the following LQR problem,

$$\int_0^T \left\{ \|y - \bar{x}\|_Q^2 + \|v\|_R^2 \right\} dt + \|y(T)\|_M^2 \tag{2.39}$$

$$\text{s.t. } \frac{d}{dt}y = Ay + Bv, \qquad y(0) = 0, \tag{2.40}$$

which has a unique optimal solution $v_* = 0$. Thus, $e_x = e_q = 0$. This proves the result. $\square$

We define the following functions, which are used to compute the solution (2.44) of (2.31)-(2.32) when the CDV is considered as a parameter,

$$\frac{d}{dt}\gamma = -A'\gamma - \gamma A + \gamma B R^{-1} B'\gamma, \qquad \gamma(T) = M \tag{2.41}$$

$$\frac{d}{dt}R_1(t,s) = \left(A - BR^{-1}B'\gamma(t)\right) R_1(t,s), \qquad R_1(s,s) = I_n, \tag{2.42}$$

$$\frac{d}{dt}R_2(t) = \left(A - BR^{-1}B'\gamma(t)\right) R_2(t) + BR^{-1}B'R_1(T,t)'M, \qquad R_2(0) = 0, \tag{2.43}$$

$$\bar{x}^\lambda = R_1(t,0)\mathbb{E}x(0) + R_2(t)P\lambda, \tag{2.44}$$

where $\lambda \in S := \{\lambda = (\lambda_1, \dots, \lambda_l) \in \mathbb{R}^l | \lambda_j \geq 0 \text{ and } \sum_{j=1}^l \lambda_j = 1\}$ and $P = (p_1, \dots, p_l) \in \mathbb{R}^{n \times l}$ is the matrix of alternatives. Next, we define the finite dimensional map F from the simplex S into itself, such that, $\forall \lambda \in S$,

$$F(\lambda) = \left(P_0\left(D^1\left(\bar{x}^\lambda\right)\right), \dots, P_0\left(D^l\left(\bar{x}^\lambda\right)\right)\right). \tag{2.45}$$

We are now ready to state the main result of this chapter.

Theorem 4 (Sustainable mean trajectories)**.** *Under Assumption 3, the following statements hold:*

(i) $\bar{x}$ *satisfies* (2.29)-(2.30) *if and only if*

$$\bar{x} = \bar{x}^\lambda, \tag{2.46}$$

with $\lambda \in S$ *is a fixed point of* F*, where* $\bar{x}^\lambda$ *is defined in* (2.44).

(ii) F *has at least one fixed point, equivalently, there exists at least one* $\bar{x}$ *satisfying the mean field equations* (2.29)-(2.30).

(iii) The paths $\bar{x}^\lambda$ *are uniformly bounded with respect to* $M \succ 0$ *and* $\lambda \in [0,1]^l$*, for the semi-norm* $\left(\int_0^T \|.\|_Q^2 \mathrm{dt}\right)^{\frac{1}{2}}$.

Proof.

(i) Consider $\bar{x}$ a solution of the mean field equations (2.29)-(2.30). Following Lemma 1, $\bar{x}$ satisfies the equivalent representation (2.31)-(2.32). We define $\lambda := \lambda(\bar{x})$, where $\lambda(\bar{x})$ is the CDV of $\bar{x}$ defined in Lemma 1. Then, $\bar{x}$ and $\bar{q}$ are the optimal state and co-state of the following optimal LQR control problem,

$$\int_0^T \|v\|_R^2 dt + \|y(T) - P\lambda\|_M^2 \tag{2.47}$$
$$\text{s.t. } \frac{d}{dt}y = Ay + Bv, \qquad y(0) = \mathbb{E}x(0).$$

This, optimal control problem has a unique optimal state $\bar{x}^\lambda$ defined by (2.44). Hence, $\bar{x} = \bar{x}^\lambda$, where

$$\begin{aligned}\lambda = \lambda(\bar{x}) &= \left(P_0\left(D^1\left(\bar{x}\right)\right), \dots, P_0\left(D^l\left(\bar{x}\right)\right)\right)\\ &= \left(P_0\left(D^1\left(\bar{x}^\lambda\right)\right), \dots, P_0\left(D^l\left(\bar{x}^\lambda\right)\right)\right) = F(\lambda).\end{aligned} \tag{2.48}$$

The first equality follows from the definition of λ, the second from the definition of the CDV $\lambda(\bar{x})$, the third from $\bar{x} = \bar{x}^\lambda$, and the forth from the definition of F.

Conversely, let $\bar{x} = \bar{x}^\lambda$, for a fixed point $\lambda \in S$ of F. Following Lemma 1, it is sufficient to show that $\bar{x}$ satisfies the equivalent representation (2.31)-(2.32). But $\bar{x}^\lambda$ defined by (2.41)-(2.44) is the optimal state of (2.47). Hence, $\bar{x}^\lambda$ and its co-state $\bar{q} := \gamma\bar{x}^\lambda + \beta$ satisfy (2.31)-(2.32), but with $\bar{q}(T) = M(\bar{x}^\lambda - P\lambda)$. Here, β is the unique solution of $\frac{d}{dt}\beta = (\gamma BR^{-1}B' - A')\,\beta$, with $\beta(T) = -MP\lambda$. Thus, it remains to show that $\lambda = \lambda(\bar{x})$, where $\lambda(\bar{x})$ is defined in Lemma 1. Indeed,

$$\begin{aligned}\lambda = F(\lambda) &= \left(P_0\left(D^1\left(\bar{x}^\lambda\right)\right), \dots, P_0\left(D^l\left(\bar{x}^\lambda\right)\right)\right)\\ &= \left(P_0\left(D^1\left(\bar{x}\right)\right), \dots, P_0\left(D^l\left(\bar{x}\right)\right)\right) = \lambda(\bar{x}).\end{aligned} \tag{2.49}$$

The first equality follows from the hypothesis that λ is a fixed point of F, the second from the definition of F, the third from $\bar{x} = \bar{x}^\lambda$ and the fourth from the definition of $\lambda(\bar{x})$. This proves the first point.

(ii) F is a function from the convex compact set S into itself. It is sufficient to show that F is continuous. Brouwer's fixed point theorem (Conway, 2013, Section V.9) guaranties then the existence of at least one fixed point. The definition of F (2.45) includes the basins of attraction $H_\lambda^j := D^j(\bar{x}^\lambda)$, $1 \leq j \leq l$, parametrized by λ. Thus, we need at first to understand how these basins vary with λ. Indeed, by solving the linear ODE's (2.21)-(2.22) of β^j and δ^j, where $\bar{x}$ is replaced by $\bar{x}^\lambda$ defined in (2.44), and by replacing the explicit solutions β^j and δ^j in the expression of the basin of attraction, (2.18), we get the following expression for the basin of attraction H^j_λ,

$$H^j_\lambda = \left\{x_0 \in \mathbb{R}^n | \beta'_{jk} x_0 \leq \delta_{jk} + \theta'_{jk}\mathbb{E}x(0) + \xi'_{jk}P\lambda, \forall 1 \leq k \leq l\right\}, \tag{2.50}$$

where $\beta_{jk}, \theta_{jk}, \xi_{jk} \in \mathbb{R}^n$ and $\delta_{jk} \in \mathbb{R}$ depend on the alternatives p_j and p_k, and A, B, Q, R and M. More importantly, they do not depend on λ. Thus, if $\{\lambda_k\}_{k\in\mathbb{N}}$ is a sequence in S that converges to λ, then the boundaries of $H^j_{\lambda_k}$, which are hyperplanes in $\mathbb{R}^n$, are **translated** towards those of H^j_λ. Thus $1_{\overset{\circ}{H}{}^j_{\lambda_k}}(x_0)$ converges to $1_{\overset{\circ}{H}{}^j_{\lambda}}(x_0)$ for all $x_0 \in \mathbb{R}^n$. Under Assumption 3, this means that $1_{H^j_{\lambda_k}}(x_0)$ converges to $1_{H^j_\lambda}(x_0)$ P_0-a.s.. Lebesgue's dominated convergence theorem (Rudin, 1987) implies that $F(\lambda_k)$ converges to $F(\lambda)$. This completes the proof of the second point.

(iii) Following the proof of point (i), the paths $\bar{x}^\lambda$ are the optimal states of the control problem (2.47). Since (A, B) is controllable, the corresponding optimal control law v^λ satisfies

$$\int_0^T \|v^\lambda\|^2_R \mathrm{dt} \leq \int_0^T \|\tilde{v}^\lambda\|^2_R \mathrm{dt}, \tag{2.51}$$

where $\tilde{v}^\lambda = -B'e^{A't}\left(W^{-1}(0,T)\left(P\lambda - e^{AT}x(0)\right)\right)$ is the continuous control law that transfers the state y from $y(0)$ to $P\lambda$, and W the Gramian of (A, B). But $\tilde{v}^\lambda$ is independent of M and continuous with respect to λ. Hence,

$$\sup_{\lambda\in[0,1]^l} \int_0^T \|v^\lambda\|^2 \mathrm{dt} \leq \frac{1}{\lambda_{\min}(R)} \sup_{\lambda\in[0,1]^l} \int_0^T \|v^\lambda\|^2_R \mathrm{dt} \leq \frac{1}{\lambda_{\min}(R)} \max_{\lambda\in[0,1]^l} \int_0^T \|\tilde{v}^\lambda\|^2_R \mathrm{dt}. \tag{2.52}$$

Moreover,

$$\bar{x}^\lambda(t) = \exp(At)\mathbb{E}x(0) + \int_0^t \exp(A(t-\sigma))Bv^\lambda(\sigma)\mathrm{d}\sigma. \tag{2.53}$$

Therefore,

$$\int_0^T \|\bar{x}^\lambda\|^2_Q \mathrm{dt} \leq \lambda_{\max}(Q) \int_0^T \|\bar{x}_\lambda\|^2 \mathrm{dt} \leq K_1 + K_2 \int_0^T \|v^\lambda\|^2 \mathrm{dt}, \tag{2.54}$$

for some positive constants K_1 and K_2 that are independent of (M, λ). This proves the result.

□

Remark 2. *The third point of Theorem 4 is used to prove Theorem 3 in Section 2.3.*

2.5 Discussions

Theorems 2 and 4 include the following main contributions of the degenerate non-cooperative DCDC problem, which were listed in Section 1.3.1.

2.5.1 Commitment to the initial choice

Theorems 2 and 4 show that in the limit of an infinite number of agents, an agent makes its choice of an alternative based on its initial condition, and assuming that it knows the initial distribution P_0. Moreover, it commits to its choice along the path. Indeed, given the initial distribution P_0, the agents compute a fixed point λ of F defined in (2.45), and anticipate their mean trajectory $\bar{x} = \bar{x}^\lambda$ defined in (2.44). Subsequently, an agent i makes its choice of an alternative at time $t = 0$ as follows. It checks to which basin of attraction its initial state $x_i(0)$ belongs. If $x_i(0) \in D^j(\bar{x}^\lambda)$, then agent i goes toward the alternative p_j by implementing the optimal control law u_*^j of the LQR problem (2.19) that corresponds to p_j along the time horizon $[0, T]$. This choice process is summarized in Figure 2.1.

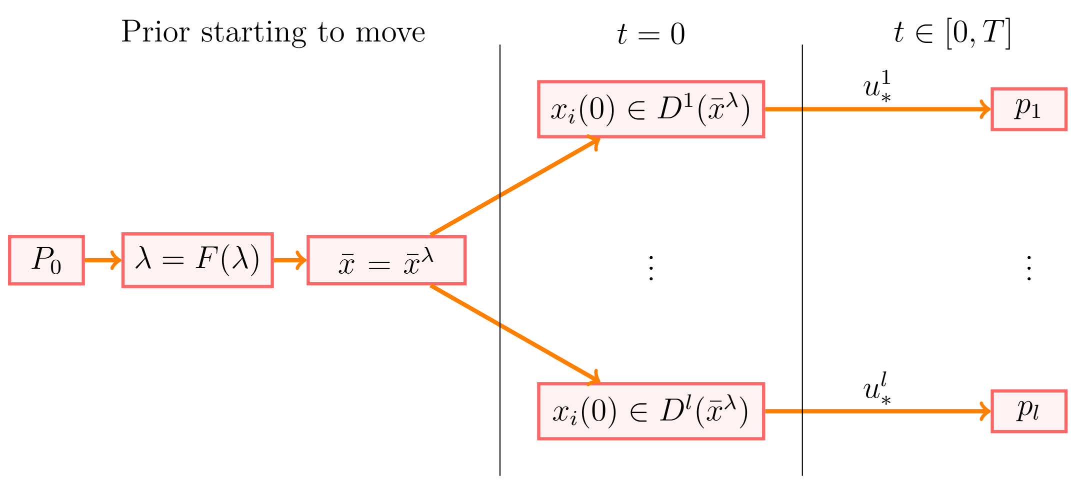

Figure 2.1 Degenerate non-cooperative DCDC: the choice process

2.5.2 Exact vs. approximate Nash equilibria

We show in Section 2.2 that the exact Nash equilibria are mapped one-to-one to the fixed points of F_{Nash}. A naïve approach to compute an equilibrium (if it exists) is to check for each $d \in \{p_1, \dots, p_l\}^N$ whether it is a fixed point of F_{Nash} or not. But, the computation of $F_{Nash}(d)$ includes solving the LQR game $Game(d)$, which involves a common Riccati equation of dimension n^2N, and a d-dependent linear ODE of dimension $nN+1$. In total, the naïve method requires solving an ODE of dimension $n^2N+l^N(nN+1)$. Furthermore, to implement an exact Nash equilibrium, each agent needs to know at least the exact initial states of all the other agents. Instead, we develop in Theorems 2 and 4 a set of decentralized strategies that are simple to compute, and where an agent is only required to know its state and the initial probability distribution P_0. In this case, we need only to solve one Riccati equation (2.20) of dimension n^2, l ODE's (2.21)-(2.22), each of dimension $n+1$, and a fixed point $\lambda \in \mathbb{R}^n$. We will show later in Section 2.7 that these strategies converge to an exact Nash equilibrium as the number of agents increases to infinity. We summarize the differences between the exact and approximate Nash equilibria in Table 2.1.

Table 2.1 Degenerate non-cooperative DCDC: exact and approximate Nash strategies

	Exact Nash	Approximate Nash
Computation	Solve an ODE of dimension $n^2N+l^N(nN+1)$.	Solve an ODE of dimension n^2+ln, and a fixed point $\lambda \in \mathbb{R}^n$.
Implementation	Each agent needs to know its state and the initial states of the others.	Each agent needs to know its state and P_0.

2.5.3 Full characterization of the limiting game

Theorem 4 constructs a one-to-one map between the infinite population Nash equilibria and the fixed points of F. This allows in some cases, for example the Gaussian binary choice case below, to derive sufficient conditions for the uniqueness of fixed points of F (equivalently, the uniqueness of infinite population Nash equilibria) to hold.

Gaussian binary choice case

We assume here that the number of alternatives is two ($l=2$) and the initial distribution P_0 is Gaussian with mean μ_0 and variance matrix Σ_0. For any $n\times n$ matrix Σ such that $\beta_{12}'\Sigma\beta_{12} < \left(\xi_{12}(p_1-p_2)\right)^2/2\pi$, we define

$$\begin{aligned}
a(\Sigma) &= \delta_{12}+\xi_{12}p_2-\sqrt{\log\xi_{12}(p_1-p_2)-\frac{1}{2}\log 2\pi\beta_{12}'\Sigma\beta_{12}}\sqrt{2\beta_{12}'\Sigma\beta_{12}}\\
b(\Sigma) &= \delta_{12}+\xi_{12}p_1+\sqrt{\log\xi_{12}(p_1-p_2)-\frac{1}{2}\log 2\pi\beta_{12}'\Sigma\beta_{12}}\sqrt{2\beta_{12}'\Sigma\beta_{12}}\\
S(\Sigma) &= \left\{\mu\in\mathbb{R}^n,\ \left(\beta_{12}'-\theta_{12}\right)\mu\in(a(\Sigma),b(\Sigma))\right\},
\end{aligned} \tag{2.55}$$

where β_{12}, δ_{12}, θ_{12} and ξ_{12} are defined in the proof of point (ii) of Theorem 4.

Corollary 1. *F has a unique fixed point if one of the following conditions is satisfied*

(i) $\beta_{12}'\Sigma_0\beta_{12} \geq \Big(\xi_{12}(p_1 - p_2)\Big)^2/2\pi$.

(ii) $\beta_{12}'\Sigma_0\beta_{12} < \Big(\xi_{12}(p_1 - p_2)\Big)^2/2\pi$ *and* $\mu_0 \notin S(\Sigma_0)$.

Proof. The generic agents' initial state $x(0)$ is distributed according to a Gaussian distribution $\mathcal{N}(\mu_0, \Sigma_0)$. Therefore, $\beta_{12}'x(0)$ has a normal distribution $\mathcal{N}\Big(\beta_{12}'\mu_0, \beta_{12}'\Sigma_0\beta_{12}\Big)$. Thus, one can analyze the dependence of $\Big[F(\alpha, 1-\alpha)\Big]_1 - \alpha$ on α to show that this function has a unique zero in $[0, 1]$ in case (i) or (ii) holds. Indeed, in both cases, the sign of the derivative with respect to α of $\Big[F(\alpha, 1-\alpha)\Big]_1 - \alpha$ does not change. As a result, this function is monotonic, and F has a unique fixed point. □

Corollary 1 states that in the Gaussian binary choice case, if the initial distribution of the agents has enough spread, then the agents make their choices in a unique way. On the other hand, if the uncertainty in their initial positions is low enough and the mean of population is inside the region $S(\Sigma_0)$ (a region delimited by two parallel hyperplanes), then the agents can split between the alternatives in multiple ways.

The map F characterizes also the way the agents split between the alternatives. Indeed, if $\lambda = (\lambda_1, \ldots, \lambda_l)$ is a fixed point of F, and the agents are optimally responding to $\bar{x}^\lambda$, then the Law of large numbers (Durrett, 2010) implies that $\mathbb{P}-$a.s. $\lim\limits_{N\to\infty} \sum_{i=1}^{N} 1_{D^j(\bar{x}^\lambda)}(x_i(0)) = \mathbb{E}1_{D^j(\bar{x}^\lambda)}(x(0)) = \lambda_j$. This means that, in the limit, $\lambda_j\%$ of the agents go towards p_j. In other words, the fixed points of F are the potential probability distributions of the agents' choices over the alternatives.

2.5.4 Simple numerical scheme

The computation of the mean field term in the general MFG theory involves solving a forward FP equation coupled with a backward HJB equation (Achdou and Capuzzo-Dolcetta, 2010). In our case, however, Theorem 4 reduces this infinite dimensional problem to the computation of a fixed point for the finite dimensional map F. This map is a vector of probabilities of some regions delimited by hyperplanes. Although a fixed point can be computed using Newton's method, this is computationally expensive as it requires the values of the inverse of the Jacobian matrix at the root estimates. Alternatively, we use a quasi Newton method such as Broyden's method (Broyden, 1965). According to this method, the inverse of the Jacobian is estimated recursively provided that F is continuously differentiable; this will be the case if the initial probability distribution has a continuous probability density function.

2.6 Nonuniform population with initial preferences

Hitherto, the agents' affinities towards different potential alternatives are dictated only by their a priori opinions (initial states) in $\mathbb{R}^n$. In this section, the model is further generalized by considering that the agents have also different tendencies towards the alternatives. When modeling smoking decision in schools for example (Nakajima, 2007), this could represent a teenager's tendency towards "Smoking" or "Not Smoking", which is the result of some endogenous factors such as parental pressure, financial condition, health, etc. In the elections example, it represents personal preferences that transcend party lines. Moreover, we assume that the agents have nonuniform dynamics and costs in this section.

We consider N agents with the following nonuniform dynamics,

$$\frac{d}{dt}x_i = A_i x_i + B_i u_i, \qquad 1 \leq i \leq N, \tag{2.56}$$

with random initial conditions as in Section 2.1. An agent i is associated with the following cost functional,

$$J_i\left(u_i, \bar{x}\right) = \mathbb{E}\left(\int_0^T \left\{\|x_i - \bar{x}\|^2_{Q_i} + \|u_i\|^2_{R_i}\right\} dt + \min_{1 \leq j \leq l} \|x_i(T) - p_j\|^2_{M_i^j}\right), \tag{2.57}$$

where $Q_i \succeq 0$, $R_i \succ 0$ and $M_i^j \succ 0$.

In contrast to the final cost coefficients in (2.2), the coefficients M_i^j, $1 \leq j \leq l$, in (2.57) depend on the alternatives, to model the initial tendency. By definition, we say that a population has a tendency to move towards p_k in a first scenario stronger than that in a second scenario, if in the absence of a social effect, the number of agents that go towards p_k in the first scenario is greater than that in the second scenario. We show in the following lines how the tendency, according to this definition, is controlled by M_i^j. Explicitly, we show that without a social effect ($Q_i = 0$), the number of agents that go towards an alternative p_j decreases as M_i^j increases (all other coefficients remaining constant). Hence, the tendency to move towards p_j decreases as M_i^j increases. To simplify the argument, we consider only the binary choice case $l = 2$. In the absence of a social effect, an agent i minimizes its individual cost (2.57). Following similar arguments to those of Theorem 2, agent i goes towards p_1 if and only if the optimal cost of the LQR problem corresponding to p_1 (i.e. with cost functional equal to (2.57), but where p_1 is the only available alternative) is less than that corresponding to p_2. Now, if the coefficient M_i^2 increases, then the optimal cost of the LQR problem corresponding to p_2 increases, and that corresponding to p_1 remains constant. Therefore, by increasing M_i^2, the number of agents that go towards p_2 cannot increase.

As N tends to infinity, it is convenient to represent the limiting sequence of parameters of $\{\theta_i\}_{i\in\{1,\dots,N\}} = \{(A_i, B_i, Q_i, R_i, M_i^1, \dots, M_i^l)\}_{i\in\{1,\dots,N\}}$ by a random vector θ, which belongs to a compact set Θ. We denote the empirical measure of the sequence $\{\theta_i\}_i$ as $P_\theta^N(\mathcal{A}) = 1/N\sum_{i=1}^N 1_{\{\theta_i\in\mathcal{A}\}}$ for all (Borel) measurable sets $\mathcal{A}$. We assume that P_θ^N converges weakly to the probability distribution P_θ of θ, that is, for all ϕ continuous, $\lim_{N\to\infty}\int_\Theta \phi(x)\mathrm{d}P_\theta^N(x) = \int_\Theta \phi(x)\mathrm{d}P_\theta(x)$. For further discussions about this assumption, one can refer to (Huang et al., 2012). The initial states $x_i(0)$, $1 \le i \le N$ and θ are assumed independent. Moreover, an agent i knows its initial position $x_i(0)$, its parameters θ_i, as well as the distributions P_0 and P_θ. The following analysis is developed for a generic agent with an initial position $x(0)$ and parameters $\theta = (A_\theta, B_\theta, Q_\theta, R_\theta, M_\theta^1, \dots, M_\theta^l)$. Assuming an infinite size population, a generic agent optimally responds to a posited deterministic although initially unknown continuous path $\bar{x}$. Following the analysis in Section 2.3, its best response is as follows,

$$u_*(t, x, \bar{x}, x(0), \theta) = -B_\theta' R_\theta^{-1}\left(\Gamma_\theta^j(t)x + \beta_\theta^j(t)\right), \qquad \text{if } x(0) \in D_\theta^j(\bar{x}), \tag{2.58}$$

with Γ_θ^j, β_θ^j and δ_θ^j are the unique solutions of the ODE's (2.20)-(2.22), where we replace the parameters (A, B, Q, R, M) by $(A_\theta, B_\theta, Q_\theta, R_\theta, M_\theta^j)$. The basins of attractions defined in the following equation are now regions in $\mathbb{R}^n$ delimited by quadric surfaces that depend on $\bar{x}$.

$$\begin{aligned} D_\theta^j(\bar{x}) = \Big\{ x_0 \in \mathbb{R}^n \Big| & \frac{1}{2}x_0'\left(\Gamma_\theta^j(0) - \Gamma_\theta^k(0)\right)x_0 + \left(\beta_\theta^j(0) - \beta_\theta^k(0)\right)' x_0 + \delta_\theta^j(0) - \delta_\theta^k(0) \le 0, \\ & \forall 1 \le k \le l \Big\}. \end{aligned} \tag{2.59}$$

The continuum of parameters, as represented by the random element θ, makes it impossible to reduce the search for a sustainable path $\bar{x}$ to a finite dimensional problem as in the uniform population case. The existence proof for such a path relies now on an abstract Banach space version of Brouwer's fixed point theorem, namely Schauder's fixed point theorem (Conway, 2013).

We define the operator $G : C([0,T], \mathbb{R}^n) \to C([0,T], \mathbb{R}^n)$ that maps a path $\bar{x}$ to the mean trajectory of a generic agent that responds optimally to $\bar{x}$. A sustainable path is thus a fixed point of G. We obtain the generic agent's optimal state by solving the linear ODE of β_θ^j, replacing the solution in (2.58) and the resulting optimal control law in (2.56). Thus, the optimal state is,

$$x_*\left(t,\bar{x},x(0),\theta\right)=\Phi_\theta^j(0,t)'x(0)+\int_0^t\Psi_\theta^j(\sigma,t,\sigma,T)M_\theta^jp_j\,\mathrm{d}\sigma$$
$$-\int_0^t\int_T^\sigma\Psi_\theta^j(\sigma,t,\sigma,\tau)Q_\theta\bar{x}(\tau)d\tau d\sigma,\qquad \text{if } x(0)\in D_\theta^j(\bar{x}),\qquad (2.60)$$

where $\Psi_\theta^j(\eta_1,\eta_2,\eta_3,\eta_4)=\Phi_\theta^j(\eta_1,\eta_2)'B_\theta R_\theta^{-1}B_\theta'\Phi_\theta^j(\eta_3,\eta_4)$, $\Pi_\theta^j(t)=\Gamma_\theta^j(t)B_\theta R_\theta^{-1}B_\theta'-A_\theta'$, and Φ_θ^j the unique solution of

$$\frac{d\Phi_\theta^j(t,\eta)}{dt}=\Pi_\theta^j(t)\Phi_\theta^j(t,\eta)\qquad \Phi_\theta^j(\eta,\eta)=I_n.\qquad (2.61)$$

When deriving (2.60), we use one main property of the state transition matrix Φ_θ^j, namely the state transition matrix $\tilde{\Phi}_\theta^j(\eta_1,\eta_2)$ of $-(\Pi_\theta^j)'$ (i.e. solution of (2.61), where Π_θ^j in the right hand side of (2.61) is replaced by $-(\Pi_\theta^j)'$) is equal to $\Phi_\theta^j(\eta_2,\eta_1)'$. For more details about the properties of the state transition matrix, we refer the reader to (Rugh, 1996). By taking the expectation on both sides of (2.60), we get the following expression of the operator G. For all $\bar{x}\in C([0,T],\mathbb{R}^n)$,

$$G(\bar{x})=\mathbb{E}x_*\left(t,\bar{x},x(0),\theta\right)=\sum_{j=1}^{l}\int_\Theta\int_{\mathbb{R}^n}1_{D_\theta^j(\bar{x})}(x_0)\Big\{\Phi_\theta^j(0,t)'x_0+\int_0^t\Psi_\theta^j(\sigma,t,\sigma,T)M_\theta^jp_j\,\mathrm{d}\sigma$$
$$-\int_0^t\int_T^\sigma\Psi_\theta^j(\sigma,t,\sigma,\tau)Q_\theta\bar{x}(\tau)d\tau d\sigma\Big\}dP_0(x_0)dP_\theta(\theta),\qquad (2.62)$$

where the second inequality follows from Fubini-Tonelli's theorem (Rudin, 1987). Theorem 5 below states that there exists a fixed point of G, i.e. a sustainable mean trajectory $\bar{x}$. We prove the result via Schauder's fixed point theorem, which requires the boundedness of G on bounded subsets of its domain, and its continuity. These are guaranteed by the following two assumptions respectively.

Assumption 4. *We assume that* $\sqrt{\max(k_1+k_2,k_3)}T<\pi/2$, *where*

$$k_1=\mathbb{E}\|x(0)\|\times\left(\sum_{j=1}^{l}\max_{(\theta,t)\in\Theta\times[0,T]}\|\Phi_\theta^j(0,t)\|\right)$$
$$k_2=\sum_{j=1}^{l}\max_{(\theta,t)\in\Theta\times[0,T]}\left\|\int_0^t\Psi_\theta^j(\sigma,t,\sigma,T)M_\theta^jp_jd\sigma\right\|\qquad (2.63)$$
$$k_3=\sum_{j=1}^{l}\max_{(\theta,t,\sigma,\tau)\in\Theta\times[0,T]^3}\|\Psi_\theta^j(\sigma,t,\sigma,\tau)Q_\theta\|.$$

Since Θ and $[0,T]$ are compact and Φ^j_θ is continuous with respect to time and parameter θ, k_1, k_2 and k_3 are well defined. Noting that the left hand side of the inequality tends to zero as T goes to zero, Assumption 4 can be satisfied for a short time horizon T for example.

Assumption 5. *We assume that the P_0-measure of quadric surfaces in $\mathbb{R}^n$ is zero.*

Similar to Assumption 3, Assumption 5 is satisfied when P_0 is absolutely continuous with respect to the Lebesgue measure for example.

Theorem 5 (Sustainable mean trajectories, nonuniform agents)**.** *Under Assumptions 4 and 5, G has at least one fixed point.*

Proof. We show the result by Schauder's fixed point theorem (Conway, 2013), which states that if G is a compact operator from the Banach space $\Big(C([0,T],\mathbb{R}^n), \|.\|_\infty\Big)$ into itself, and there exists a nonempty, bounded, closed and convex subset U of $\Big(C([0,T],\mathbb{R}^n), \|.\|_\infty\Big)$, such that $G(U) \subset U$, then G has at least one fixed point in U. G is a compact operator means that it is continuous and maps bounded sets to relatively compact sets. To show the continuity, let $\bar{x} \in C([0,T],\mathbb{R}^n)$ and $\{\bar{x}_k\}_{k\in\mathbb{N}}$ be a sequence converging to $\bar{x}$ in $\Big(C([0,T],\mathbb{R}^n), \|.\|_\infty\Big)$. Let

$$C_j > \max_{(\theta,t)\in\Theta\times[0,T]^2} \|\Phi^j_\theta(t)\| + \max_{(\theta,t)\in\Theta\times[0,T]^4} \|\Psi^j_\theta(t)\| + \max_{\theta\in\Theta} \|\theta\|. \tag{2.64}$$

Then,

$$\|G(\bar{x}_k) - G(\bar{x})\|_\infty \leq \sum_{j=1}^{l} C_j \Big\{ C_j \|\bar{x}_k - \bar{x}\|_\infty + V_{1j} + C_j\|p_j\|T + C_j\|\bar{x}\|_\infty T^2 V_{2j} \Big\}, \tag{2.65}$$

$$\begin{aligned}
\text{where } V_{1j} &= \int_\Theta \int_{\mathbb{R}^n} \left| 1_{D^j_\theta \bar{x}_k)}(x_0) - 1_{D^j_\theta(\bar{x})}(x_0) \right| \|x_0\| dP_0(x_0) dP_\theta(\theta) \\
V_{2j} &= \int_\Theta \int_{\mathbb{R}^n} \left| 1_{D^j_\theta(\bar{x}_k)}(x_0) - 1_{D^j_\theta(\bar{x})}(x_0) \right| dP_0(x_0) dP_\theta(\theta).
\end{aligned}$$

Under Assumption 5, we get that

$$V_{1j} = \int_\Theta \int_{\mathbb{R}^n} \left| 1_{\mathring{D}^j_\theta(\bar{x}_k)}(x_0) - 1_{\mathring{D}^j_\theta(\bar{x})}(x_0) \right| \|x_0\| dP_0(x_0) dP_\theta(\theta). \tag{2.66}$$

But, $\left| 1_{\mathring{D}^j_\theta(\bar{x}_k)}(x_0) - 1_{\mathring{D}^j_\theta(\bar{x})}(x_0) \right| \|x_0\| \leq 2\|x_0\|$ and converges to zero for all (x_0,θ) in $\mathbb{R}^n \times \Theta$. Moreover, $\mathbb{E}\|x(0)\| < \infty$. Therefore, by Lebesgue's dominated convergence theorem (Rudin, 1987), V_{1j} converges to zero. Similarly, V_{2j} converges to zero. Hence, G is continuous.

Let V be a bounded subset of $C([0,T],\mathbb{R}^n)$. We show via Arzela-Ascoli Theorem that the closure of $G(V)$ is compact, which along with the continuity of G imply that this operator is compact. Let $\{G(\bar{x}_k)\}_{k\in\mathbb{N}} \in G(V)$. By the continuity of $\Phi^j_\theta(\sigma,t)$ with respect to (σ,t,θ), of its derivative with respect to t and σ, and by the boundedness of $\bar{x}_k$, one can prove that for all (t,s) in $[0,T]^2$,

$$\|G(\bar{x}_k)(t) - G(\bar{x}_k)(s)\| \leq \Big(K_1\mathbb{E}\|x(0)\| + K_2\Big)|t-s|, \tag{2.67}$$

where K_1 and K_2 are positive constants. This inequality implies the uniform boundedness and equicontinuity of $\{G(\bar{x}_k)\}_{k\in\mathbb{N}}$. By Arzela-Ascoli Theorem (Conway, 2013), there exists a convergent subsequence of $\{G(\bar{x}_k)\}_{k\in\mathbb{N}}$. Hence, $G(V)$ and its closure are compact sets.

It remains to construct a nonempty, bounded, closed, convex subset $U \subset C([0,T],\mathbb{R}^n)$ such that $G(U) \subset U$. Let $C = \max(k_1+k_2, k_3)$, where k_1, k_2 and k_3 are defined in (2.63). Hence, $\|G(\bar{x})(t)\| \leq C + C\int_0^t \int_\sigma^T \|\bar{x}(\tau)\| d\tau \mathrm{d}\sigma$. We consider the following set

$$U = \Big\{\bar{x} \in C([0,T],\mathbb{R}^n) |\, \|\bar{x}(t)\| \leq R(t),\, \forall t \in [0,T]\Big\}, \tag{2.68}$$

where R is a continuous positive function on $[0,T]$ to be determined later. U is an nonempty, bounded, closed and convex subset of $C([0,T],\mathbb{R}^n)$. If we can find an R positive such that $R(t) = C + C\int_0^t\int_\sigma^T R(\tau)d\tau d\sigma$, for all $t\in[0,T]$, then for all $\bar{x}\in U$, we obtain that

$$\|G(\bar{x})(t)\| \leq C + C\int_0^t \int_\sigma^T R(\tau)d\tau d\sigma = R(t), \tag{2.69}$$

and $G(U)\subset U$. Hence, It remains to find such R. Note that the equality in (2.69) is equivalent to the second order differential equation $\frac{d^2R}{dt^2} = -CR$, with the boundary conditions, $R(0) = C$ and $\frac{dR}{dt}(T) = 0$. Thus, $R(t) = C/\cos(\sqrt{C}t)$, which is positive under Assumption 4. By finding R, U is well defined. This proves the result. □

Note that if T goes to zero, the costs become decoupled, and each agent will choose the "closest" alternative in the minimum energy sense. It is then expected that a Nash equilibrium exists in this case. Assumption 4 gives an upper bound on the time horizon T under which one can prove that such an equilibrium continues to exist.

Remark 3. *If the parameters take a finite number of values, i.e. Θ is a finite set, then we can relax Assumption 4 and construct a one to one map between the sustainable paths and the fixed points of a finite dimensional map. For more details about this case, we refer the reader to Chapter 3 below.*

2.7 Approximate Nash equilibrium

We develop in Sections 2.3, 2.4, and 2.6 a set of Nash strategies for a continuum of agents. These strategies are given by (2.23) in the uniform population case, and (2.58) in the nonuniform case. This section shows that these strategies, when applied by a finite number of agents, constitute an $\epsilon-$Nash equilibrium (see Chapter 1, Definition 1), where ϵ goes to zero as the number of agents increases to infinity. This equilibrium makes the group's behavior robust in the face of potential selfish behaviors. Indeed, unilateral deviations from the associated control policies are guaranteed to yield negligible cost reductions when the size of the group increases sufficiently.

Theorem 6 ($\epsilon-$Nash equilibrium). *When applied by a finite number N of agents, the strategies defined by* (2.23) *(resp.* (2.58)*) for a sustainable mean trajectory $\bar{x}$ constitute an ϵ_N-Nash equilibrium in the set $\mathcal{U}$* (2.3) *with respect to the costs* (2.2) *(resp.* (2.57)*), where ϵ_N goes to zero as N increases to infinity.*

Proof. Since the uniform population problem is a special case of the nonuniform one, we prove only the latter. Consider an arbitrary agent $i \in \{1, ..., N\}$ in a population of N agents. Suppose that agent i applies an arbitrary control law $u_i \in \mathcal{U}$, while the rest of the agents apply the mean field based control law defined by (2.58). That is, an agent $k \neq i$ of parameters $\theta_k = (A_k, B_k, Q_k, R_k, M_k^1, \dots, M_k^l)$ implements the control law $u_*(t, x, \bar{x}, x_k(0), \theta_k)$ where $\bar{x}$ is a fixed point of G defined in (2.62). $x_*(t, \bar{x}, x_k(0), \theta_k)$ defined in (2.60) is the corresponding optimal state. In the following, x_i denotes the state of agent i under its control input u_i, $\bar{x}_{-i}^{(N)}(t) = \frac{1}{N}\sum_{k=1,k\neq i}^{N} x_*(t, \bar{x}, x_k(0), \theta_k) + \frac{1}{N}x_i(t)$ the average of the population when agent i applies u_i, and $\bar{x}^{(N)}(t) = \frac{1}{N}\sum_{k=1}^{N} x_*(t, \bar{x}, x_k(0), \theta_k)$ the average of the population when all the agents apply (2.58).

Assume that agent i can profit by a unilateral deviation from the decentralized strategies (2.58). This means that

$$J_i\left(u_i, \bar{x}_{-i}^{(N)}\right) \leq J_i\left(u_*(x_*(t, \bar{x}, x_i(0), \theta_i), \bar{x}^{(N)}\right), \tag{2.70}$$

where J_i is defined by (2.57). We show in the following that this potential cost improvement is bounded by some ϵ_N that converges to zero as N increases to infinity. In view of (2.57), the compactness of Θ, the continuity of x_* with respect to θ, and $\mathbb{E}\|x_k(0)\|^2 < \infty$, the right hand side of (2.70) is bounded by $C_1 > 0$ independently of N. Hence, $\mathbb{E}\int_0^T\{\|x_i\|^2 + \|u_i\|^2\}dt < C_2$, where C_2 does not depend on N. We decompose the cost $J_i(u_i, \bar{x}_{-i}^{(N)})$ as follows,

$$J_i\left(u_i, \bar{x}_{-i}^{(N)}\right) = J_i(u_i, \bar{x}) + S_1 + S_2 + S_3 + S_4, \tag{2.71}$$

with

$$S_1 = \frac{1}{2N^2}\mathbb{E}\int_0^T \|x_i - x_*(t,\bar{x},x_i(0),\theta_i)\|_{Q_i}^2 dt + \frac{1}{N}\mathbb{E}\int_0^T (x_*(t,\bar{x},x_i(0),\theta_i) - x_i)' Q_i (x_i - \bar{x}) dt$$
$$S_2 = \frac{1}{N}\mathbb{E}\int_0^T (x_*(t,\bar{x},x_i(0),\theta_i) - x_i)' Q_i \left(\bar{x} - \bar{x}^{(N)}\right) dt \tag{2.72}$$
$$S_3 = \mathbb{E}\int_0^T \left(\bar{x} - \bar{x}^{(N)}\right)' Q_i (x_i - \bar{x}) dt + \mathbb{E}\int_0^T \left\|\bar{x} - \bar{x}^{(N)}\right\|_{Q_i}^2 dt,$$

where $\bar{x}$ is a fixed point of G. By Cauchy-Schwarz inequality,

$$|S_1| \leq \frac{\max\limits_{\theta_i\in\Theta}|Q_i|}{N}\left(\mathbb{E}\int_0^T \|x_*(t,\bar{x},x_i(0),\theta_i) - x_i\|^2 dt\right)^{\frac{1}{2}}\left(\mathbb{E}\int_0^T \|x_i - \bar{x}\|^2 dt\right)^{\frac{1}{2}}$$
$$+ \frac{\max\limits_{\theta_i\in\Theta}|Q_i|}{N^2}\mathbb{E}\int_0^T \|x_*(t,\bar{x},x_i(0),\theta_i) - x_i\|^2 dt. \tag{2.73}$$

In view of C_1 and C_2, $\left(\mathbb{E}\int_0^T \|x_*(t,\bar{x},x_i(0),\theta_i) - x_i\|^2 dt\right)^{\frac{1}{2}}$ and $\left(\mathbb{E}\int_0^T \|x_i - \bar{x}\|^2 dt\right)^{\frac{1}{2}}$ are bounded. Thus, $|S_1| \leq \eta_1/N$, where $\eta_1 > 0$. Similarly, $|S_2| \leq \eta_2/N$, where $\eta_2 > 0$. In the remaining of the proof, we show that $\mathbb{E}\int_0^T \left\|\bar{x} - \bar{x}^{(N)}\right\|^2 dt$ converges to zero as N goes to infinity. Cauchy-Schwarz inequality implies then that S_3 and

$$\left|J_i(u_*(t,x,\bar{x},x_i(0),\theta_i),\bar{x}) - J_i\left(u_*(t,x,\bar{x},x_i(0),\theta_i),\bar{x}^{(N)}\right)\right| =$$
$$\left|\mathbb{E}\int_0^T \left(\bar{x} - \bar{x}^{(N)}\right)' Q_i \left(2x_*(t,\bar{x},x_i(0),\theta_i) - \bar{x} - \bar{x}^{(N)}\right) dt\right| \tag{2.74}$$

converge to 0 as N increases to infinity. Moreover, the optimality implies that

$$J_i(u_*(t,\bar{x},x_i(0),\theta_i),\bar{x}) \leq J_i(u_i,\bar{x}). \tag{2.75}$$

Therefore,

$$J_i\left(u_i,\bar{x}_{-i}^{(N)}\right) \geq J_i\left(u_*(t,\bar{x},x_i(0),\theta_i),\bar{x}^{(N)}\right) + \epsilon_N, \tag{2.76}$$

where

$$\epsilon_N = J_i(u_*(t,\bar{x},x_i(0),\theta_i),\bar{x}) - J_i\left(u_*(t,\bar{x},x_i(0),\theta_i),\bar{x}^{(N)}\right) + S_1 + S_2 + S_3 \tag{2.77}$$

converges to 0 as N increases to infinity. Thus, it remains to show that $\mathbb{E}\int_0^T \left\|\bar{x} - \bar{x}^{(N)}\right\|^2 dt$ converges to zero as N goes to infinity. We define

$$\alpha_N = \left(\int_0^T \left\|\bar{x} - \mathbb{E}\bar{x}^{(N)}\right\|^2 dt\right)^{\frac{1}{2}} = \left(\int_0^T \left(\int_\Theta \bar{x}_\theta \,\mathrm{d}P_\theta(\theta) - \int_\Theta \bar{x}_\theta \,\mathrm{d}P_\theta^N(\theta)\right) dt\right)^{\frac{1}{2}}, \tag{2.78}$$

where $\bar{x}_\theta = \mathbb{E}(x_*|\theta)$, with the optimal state x_* is defined in (2.60). But,

$$\mathbb{E}\int_0^T \left\|\bar{x} - \bar{x}^{(N)}\right\|^2 dt \leq 2\alpha_N^2 + 2\mathbb{E}\int_0^T \left\|\frac{1}{N}\sum_{k=1}^N \Big(\mathbb{E}x_*(t,\bar{x},x_k(0),\theta_k) - x_*(t,\bar{x},x_k(0),\theta_k)\Big)\right\|^2 dt. \tag{2.79}$$

By the compactness of $[0,T]\times\Theta$, the family of functions $\bar{x}_\theta(t)$ defined on Θ and indexed by t is uniformly bounded and equicontinuous. Corollary 1.1.5 of (Stroock and Varadhan, 1979) implies that

$$\lim_{N\to+\infty}\sup_{t\in[0,T]}\left\|\bar{x}(t) - \frac{1}{N}\sum_{k=1}^N \mathbb{E}x_*(t,\bar{x},x_k(0),\theta_k)\right\| = 0. \tag{2.80}$$

Thus, α_N converges to 0 as N increases to infinity. By the independence of the initial conditions, and thus the independence of $x_*(t,\bar{x},x_k(0),\theta_k)$, $k\in\{1,\dots,N\}$, we obtain that

$$\mathbb{E}\int_0^T \Big\|\frac{1}{N}\sum_{k=1}^N \Big(\mathbb{E}x_*(t,\bar{x},x_k(0),\theta_k) - x_*(t,\bar{x},x_k(0),\theta_k)\Big)\Big\|^2 dt = O(1/N). \tag{2.81}$$

This proves the result. □

Remark 4 (Rate of convergence of the ϵ_N)**.** *Following the proof of Theorem 6, the profit margin ϵ_N of a unilateral deviant behavior converges to* 0 *with order $O(1/\sqrt{N}) + \alpha_N$, where α_N is defined by* (2.78)*. The rate of convergence of α_N depends on the rate of convergence of the finite population parameters distribution P_θ^N to the infinite population one P_θ. In case of a uniform population, the parameters set Θ is a singleton $\{\theta_0\}$, and the distributions P_θ^N and P_θ are equal to the point mass distribution at θ_0. Hence, α_N is identically zero in this case.*

Remark 5 (Deterministic initial conditions)**.** *One could assume that the initial conditions $x_i(0)$, $1\leq i\leq N$, are deterministic. In this case, the initial states' empirical measure is assumed to have a weak limit, known to all the agents. Noting that the optimal control law* (2.58) *is continuous with respect to the parameters, and discontinuous in the initial condition, considering the initial states as deterministic requires a special treatment to deal with the discontinuity and prove* (2.80)*. For more details about this case, we refer the reader to Chapter 4.*

2.8 Simulation results

We illustrate in this section the degenerate non-cooperative DCDC problem via a numerical example. We consider 300 agents with initial positions in $\mathbb{R}^2$ drawn from the Gaussian distribution $\mathcal{N}(0, 5I_2)$. These agents choose between the alternatives $p_1 = (-39.3, -10)$, $p_2 = (-27, 9.5)$ or $p_3 = (0, 40)$. They have uniform dynamics (2.1) and costs (2.2), where

$$A = \begin{bmatrix} 0 & 1 \\ 0.2 & -0.3 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 0.3 \end{bmatrix}, \tag{2.82}$$

$R = 10$ and $M = 6000I_2$. The game lasts for $T = 1.3$. To analyze the influence of the social effect on the choices' distribution, we consider four scenarios. Figure 2.2 shows the agents' initial and final positions, and the basins of attraction (2.18) of p_1, p_2 and p_3 for the different scenarios.

In the first scenario, an agent makes its choice of an alternative without a social effect, i.e. $Q = 0$. In this case, $\lambda = (0.24, 0.11, 0.65)$ is a fixed point of F defined in (2.45). We compute λ using the method described in Section 2.5. Without a social effect, the majority of the agents, 65%, choose p_3. These are the green stars in Figure 2.2-$Q = 0$. 24% choose p_1 (red crosses in Figure 2.2-$Q = 0$), and the rest p_2 (blue diamonds in Figure 2.2-$Q = 0$). As the social effect increases, it incites more agents to follow the majority. In the second scenario ($Q = 8I_2$), for example, the size of majority increases to 80%, while in the third one ($Q = 16I_2$), it reaches 86.3%. In the fourth scenario, as the social effect becomes strong enough ($Q = 24I_2$), a consensus to follow the majority occurs. The yellow balls are the agents that change their decisions with respect to those in the absence of a social effect. Finally, the relation between the size of the majority and the strength of the social effect is reflected in the shape of the basins of attraction. Indeed, Figure 2.2 shows that as the social effect increases, the basin of attraction of p_3 expends at the expenses of those of p_1 and p_2.

Finally, Figure 2.3 shows, for $Q = 8I_2$, that the sustainable mean trajectory $\bar{x}^\lambda$ defined in (2.44) for the fixed point $\lambda = (0.07, 0.13, 0.8)$ of F, and the average state of the 300 agents, when they optimally respond to $\bar{x}^\lambda$, are almost identical.

2.9 Conclusion

We formulate in this chapter the degenerate non-cooperative DCDC problem as a dynamic game with a novel class of individual costs, the min-LQG costs. To solve the game, we use the MFG methodology and develop a set of decentralized strategies that qualify as approximate Nash strategies as the size of the population increases sufficiently. According to these strategies, an agent chooses an alternative before it starts moving based on its initial condition. Moreover, it commits to its initial choice along the path. Afterwards, we construct a one-to-one map between the $\epsilon-$Nash equilibria and the fixed points of a finite dimensional

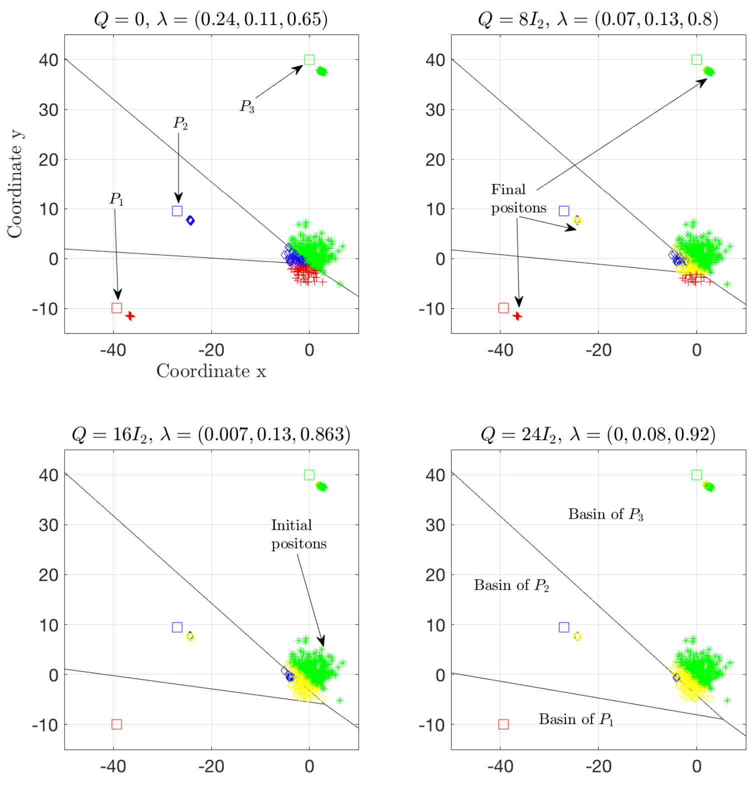

Figure 2.2 Degenerate non-cooperative DCDC: influence of the social effect on the choices' distribution.

map. The latter characterizes the game in terms of number of equilibria and the potential distributions of the agents' choices over the alternatives. Moreover, it reduces the infinite dimensional problem of solving the mean field equations to finding a fixed point of a finite dimensional map. The only randomness in the current formulation lies in the initial conditions. In the next chapter, we extend the current model to include noise processes in the agents' dynamics.

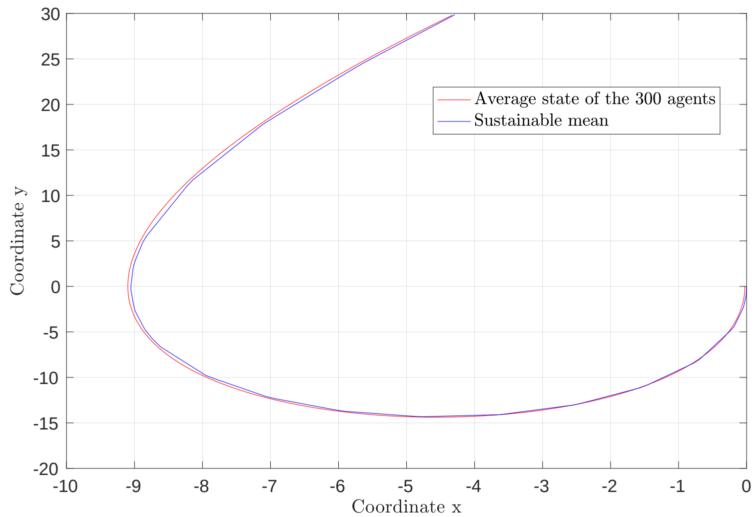

Figure 2.3 Degenerate non-cooperative DCDC: average state vs. sustainable mean for $Q = 8I_2$.

CHAPTER 3 NON-DEGENERATE NONCOOPERATIVE DYNAMIC COLLECTIVE DISCRETE CHOICE MODEL

We consider in this chapter the non-degenerate non-cooperative DCDC problem, where the agents are driven by a set of noise processes. In the robotic example of Section 1.2.4, this would model the forces that perturb a robot's trajectory while moving toward the potential sites. The noise processes can also represent the unexpected events that influence a voter's opinion during electoral campaigns in the elections example of Section 1.2.3.

We formulate in Section 3.1 the non-cooperative non-degenerate DCDC problem as a dynamic non-cooperative game. The agents' dynamics are now controlled diffusion processes, and their individual costs are similar to those in Chapter 2. Following the MFG methodology, Section 3.2 starts by considering a continuum of agents with assumed known mean trajectory. In order to compute its best response to this trajectory, a generic agent solves a non-degenerate min-LQG optimal tracking problem. We derive an explicit solution of this problem, and show that an agent faces a continuously revised discrete choice problem (McFadden, 1974), where the costs include a term penalizing myopic decisions. The problem of finding a sustainable mean trajectory is considered in Section 2.4. As in the degenerate case, we construct a one-to-one map between the sustainable mean trajectories and the fixed points of a finite dimensional map. The latter are the potential probability distributions of the agents' choices over the alternatives. We derive conditions for the existence of a sustainable mean trajectory to hold. In section 3.4, we discuss the results and contributions of this Chapter. Finally, Section 3.5 illustrates the non-degenerate DCDC problem by some numerical results, and Section 3.6 concludes this chapter.

3.1 Mathematical model

We model the non-degenerate non-cooperative DCDC problem as a stochastic dynamic non-cooperative game involving a large number N of agents with the following dynamics,

$$dx_i(t) = (A_i x_i(t) + B_i u_i(t))\, dt + \sigma_i dw_i(t), \tag{3.1}$$

for $1 \leq i \leq N$, where $A_i \in \mathbb{R}^{n\times n}$, $B_i \in \mathbb{R}^{n\times m}$, $\sigma_i \in \mathbb{R}^{n\times n}$, and $\{w_i, 1 \leq i \leq N\}$ are N independent Brownian motions in $\mathbb{R}^n$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\in[0,T]})$. $\{\mathcal{F}_t\}_{t\in[0,T]}$ is the augmented filtration of $\{\sigma(x_i(0), w_i(s), 0 \leq s \leq t, 1 \leq i \leq N)\}_{t\in[0,T]}$ (Karatzas and Shreve, 2012, Section 2.7), where $x_i(0)$, $1 \leq i \leq N$, are the initial states assumed i.i.d. with finite second moments $\mathbb{E}\|x_i(0)\|^2 < \infty$, and also independent of $\{w_i, 1 \le i \le N\}$. We assume that σ_i, $1 \le i \le N$, are invertible. In the remainder of this chapter, $\mathcal{M}([0,T], \mathbb{R}^n)$ denotes the set of progressively measurable $\mathbb{R}^n$-valued functions with respect to the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$. The vector $x_i(t) \in \mathbb{R}^n$ is the state of agent i at time t and $u_i(t) \in \mathbb{R}^m$ its control input. Each agent is associated with the following min-LQG individual cost functional:

$$J_i(u_i, \bar{x}) = \mathbb{E}\left(\int_0^T \left\{\|x_i - \bar{x}\|^2_{Q_i} + \|u_i\|^2_{R_i}\right\} dt + \min_{1\le j\le l} \|x_i(T) - p_j\|^2_{M_i}\right), \tag{3.2}$$

where $p_j \in \mathbb{R}^n$, $1 \le j \le l$, are the l alternatives, $T > 0$, $Q_i \succeq 0$, $R_i \succ 0$ and $M_i \succ 0$. The agents are cost coupled through their average state trajectory $\bar{x} = 1/N \sum_{i=1}^N x_i$. For a detailed discussion about the min-LQG cost, we refer the reader to Section 2.1.

We assume a population of k types of agents, that is, the vector of individual parameters $\theta_i := (A_i, B_i, \sigma_i, Q_i, R_i, M_i)$ takes values in a finite set $\{\Theta_1, \dots, \Theta_k\}$, which does not depend on the size of the population N. The empirical probability measure of the sequence $\{\theta_i\}_{i=1,\dots,N}$ is denoted by $P^N_\theta(\Theta_s) = 1/N \sum_{i=1}^N 1_{\{\theta_i = \Theta_s\}}$ for $s = 1, \dots, k$. We assume that $(P^N_\theta(\Theta_1), \dots, P^N_\theta(\Theta_k))$ converges to $P_\theta = (\alpha_1, \dots, \alpha_k)$, as $N \to \infty$, where $\alpha_s > 0$ for all $1 \le s \le k$.

For each agent, the set of admissible control laws is defined as follows:

$$\mathcal{U} = \left\{ u \in \mathcal{M}([0,T], \mathbb{R}^m) \,\middle|\, \mathbb{E}\int_0^T \|u(s)\|^2 ds < \infty \right\}. \tag{3.3}$$

The set of admissible Markov policies is

$$\begin{aligned} \mathcal{L} = \Big\{ & u \in (\mathbb{R}^m)^{[0,T]\times\mathbb{R}^n} | \exists L_1 > 0, \forall (t,x) \in [0,T]\times\mathbb{R}^n, \|u(t,x)\| \le L_1(1+\|x\|), \text{ and} \\ & \forall r > 0, \forall T' \in (0,T), \exists L_2 > 0, \forall \|(x,y)\| \le r, t \in [0,T'], \|u(t,x) - u(t,y)\| \le L_2\|x-y\| \Big\}. \end{aligned} \tag{3.4}$$

If $u \in \mathcal{L}$, then the stochastic differential equation (SDE) (3.1), with u_i equal to $u(t, x_i)$, has a unique strong solution (Karatzas and Shreve, 2012, Section 5.2). As shown below in Theorem 12, the mean field based strategies form $\epsilon-$Nash equilibria with respect to the space of admissible actions $\mathcal{U}$. Moreover, these strategies can be expressed as Markov policies (feedback policies), see (3.42).

3.2 The min-LQG optimal control problem and the generic agent's best response

Following the MFG methodology, we start by assuming a continuum of agents, with deterministic mean trajectory $\bar{x}$, which is supposed known in this section. The problem of determining $\bar{x}$ is treated in Section 3.3. In order to compute its best response to $\bar{x}$, a generic agent of state x, control input u and parameters $\theta = (A, B, \sigma, Q, R, M) \in \{\Theta_1, \ldots, \Theta_k\}$ solves the following non-degenerate min-LQG optimal control problem:

$$\begin{aligned} &\inf_{u \in \mathcal{U}} J\left(u, \bar{x}, x(0)\right) = \inf_{u \in \mathcal{U}} \mathbb{E}\left(\int_0^T \left\{ \|x - \bar{x}\|_Q^2 + \|u\|_R^2 \right\} dt + \min_{1 \le j \le l} \|x(T) - p_j\|_M^2 \middle| x(0) \right) \\ &\text{s.t. } dx(t) = \left(Ax(t) + Bu(t)\right) dt + \sigma dw(t), \end{aligned} \tag{3.5}$$

where w is a Brownian motion in $\mathbb{R}^n$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $x(0)$ is a random vector distributed according to the known distribution of the agents' initial states. The optimal cost-to-go function of (3.5) satisfies the following HJB equation (Fleming and Soner, 2006):

$$\begin{aligned} -\frac{\partial V}{\partial t} &= x' A' \frac{\partial V}{\partial x} - \frac{1}{2} \left(\frac{\partial V}{\partial x} \right)' B R^{-1} B' \frac{\partial V}{\partial x} + \frac{1}{2} \mathrm{Tr}\left(\sigma' \frac{\partial^2 V}{\partial x^2} \sigma \right) + \|x - \bar{x}\|_Q^2 \\ V(T, x) &= \min_{1 \le j \le l} \|x - p_j\|_M^2, \; \forall x \in \mathbb{R}^n. \end{aligned} \tag{3.6}$$

To solve (3.6) explicitly, we follow two strategies of Polya's second principle "devise a plan" (Pólya, 1957), namely, "consider special cases" and "guess and check". Indeed, we start in Section 3.2.1 by considering the scalar ($n = m = 1$) binary choice ($l = 2$) case with $A = Q = 0$. In this special case, the Hopf-Cole transformation transforms (3.6) into a heat equation, for which one can drive an explicit solution. Then, we propose in Section 3.2.2 a general form of this solution and show that it is indeed the unique solution of (3.6) in the general case.

3.2.1 Scalar binary choice case with $\mathbf{A} = \mathbf{Q} = \mathbf{0}$

We assume in this section that $n = m = 1$, $l = 2$ and $A = Q = 0$. In this case, the following generalized Hopf-Cole transformation (Evans, 1998, Chapter 4, Section 4.4)

$$\psi(t, x) = \exp\left(-\frac{B^2}{\sigma^2 R} V(t, x) \right) \tag{3.7}$$

transforms the HJB equation (3.6) to the following heat equation,

$$\begin{aligned} \frac{\partial \psi}{\partial t} &= -\frac{\sigma^2}{2}\frac{\partial^2 \psi}{\partial x^2} \\ \psi(T,x) &= \exp\left(-\frac{MB^2}{2\sigma^2 R}\min_{j=1,2}(x-p_j)^2\right), \quad \forall x \in \mathbb{R}. \end{aligned} \tag{3.8}$$

Without loss of generality, we assume that $p_1 < p_2$. This heat equation has a unique solution (Evans, 1998, Section 2.3, Theorem 1)

$$\begin{aligned} \psi(t,x) &= \frac{1}{\sigma\sqrt{2\pi(T-t)}}\int_{-\infty}^{\infty}\exp\left(-\frac{(x-y)^2}{2\sigma^2(T-t)}\right)\psi(T,y)dy \\ &= \frac{1}{\sigma\sqrt{2\pi(T-t)}}\left\{\int_{-\infty}^{c}\exp\left(-\frac{(x-y)^2}{2\sigma^2(T-t)}-\frac{MB^2}{2\sigma^2 R}(y-p_1)^2\right)dy\right. \\ &\qquad \left. + \int_{c}^{\infty}\exp\left(-\frac{(x-y)^2}{2\sigma^2(T-t)}-\frac{MB^2}{2\sigma^2 R}(y-p_2)^2\right)dy\right\}, \end{aligned} \tag{3.9}$$

where $c = (p_1+p_2)/2$. With the current form (3.9) of the solution, it is hard to guess what the general solution will look like. Hence, we derive in the remainder of this section another form of (3.9). In the degenerate case, the optimal control law of the min-LQG problem (2.23) is expressed in terms of the optimal policies and costs of the LQR problems (2.19) that a generic agent faces when one the alternative is available. This suggests to write the solution (3.9) in terms of the following LQG problems,

$$\begin{aligned} &\inf_{u\in\mathcal{U}} J^j(u,x(0)) = \inf_{u\in\mathcal{U}}\mathbb{E}\left(\int_0^T \frac{R}{2}u^2 dt + \frac{M}{2}(x(T)-p_j)^2 \Big| x(0)\right) \\ &\text{s.t. } dx(t) = Bu(t)dt + \sigma dw(t), \end{aligned} \tag{3.10}$$

for $j = 1,2$. Recall that the optimal cost-to-go V^j and optimal control law u_*^j of (3.10) are (Yong and Zhou, 1999, Chapter 6) $V^j(t,x) = \frac{1}{2}\Gamma(t)x^2 + \beta^j(t)x + \delta^j(t)$ and $u_*^j(t,x) = -\frac{B}{R}(\Gamma(t)x + \beta^j(t))$, where Γ, β^j and δ^j are equal to

$$\begin{aligned} \Gamma(t) &= \frac{MR}{R+MB^2(T-t)} \\ \beta^j(t) &= -\frac{MRp_j}{R+MB^2(T-t)} \\ \delta^j(t) &= \frac{MRp_j^2}{2(R+MB^2(T-t))} + \frac{\sigma^2 R}{2B^2}\left(\log\left(R+MB^2(T-t)\right) - \log R\right). \end{aligned} \tag{3.11}$$

By expanding the integrands in (3.9) and completing the squares, we get for $j = 1, 2$,

$$\frac{(x-y)^2}{2\sigma^2(T-t)} + \frac{MB^2}{2\sigma^2 R}(y-p_j)^2 = \frac{B^2}{2\sigma^2 R}\Gamma(t)x^2 + \frac{B^2}{\sigma^2 R}\beta_j(t)x + \frac{MB^2 p_j^2}{2\sigma^2(R+MB^2(T-t))} + \frac{R+MB^2(T-t)}{2\sigma^2 R(T-t)}\left(y - \frac{xR+MB^2 p_j(T-t)}{R+MB^2(T-t)}\right)^2. \quad (3.12)$$

Using these expressions in the integrands of (3.9) and making a change of variable

$$z_j = \sqrt{\frac{R+MB^2(T-t)}{R}}\left(y - \frac{xR+MB^2 p_j(T-t)}{R+MB^2(T-t)}\right), \quad (3.13)$$

for $j = 1, 2$, we get

$$\psi(t,x) = \exp\left(-\frac{B^2}{\sigma^2 R}V^1(t,x)\right)\mathbb{P}\left(\sigma w(T-t) \leq \frac{(c-x)r + MB^2(T-t)d}{\sqrt{R^2+MB^2R(T-t)}}\right) + \exp\left(-\frac{B^2}{\sigma^2 R}V^2(t,x)\right)\mathbb{P}\left(\sigma w(T-t) \geq \frac{(c-x)R - MB^2(T-t)d}{\sqrt{R^2+MB^2R(T-t)}}\right), \quad (3.14)$$

where $d = (p_2 - p_1)/2$.

In the following, we characterize the probabilities in (3.14). The optimal state x_*^j of (3.10) satisfies the linear SDE in (3.10), where u is replaced by the linear feedback u_*^j. After calculation, conditioning on $\{x_*^1(t) = x\}$, $x_*^1(T)$ is distributed according to the normal distribution (Karatzas and Shreve, 2012, page 354)

$$\mathcal{N}\left(\frac{xR+MB^2(T-t)p_1}{R+MB^2(T-t)}, \frac{\sigma^2 R(T-t)}{R+MB^2(T-t)}\right). \quad (3.15)$$

Thus,

$$\mathbb{P}\left(\sigma w(T-t) \leq \frac{(c-x)R + MB^2(T-t)d}{\sqrt{R^2+MB^2R(T-t)}}\right) = \mathbb{P}\left(x_*^1(T) \leq c \middle| x_*^1(t) = x\right). \quad (3.16)$$

Similarly, one can show that

$$\mathbb{P}\left(\sigma w(T-t) \geq \frac{(c-x)R - MB^2(T-t)d}{\sqrt{R^2+MB^2R(T-t)}}\right) = \mathbb{P}\left(x_*^2(T) \geq c \middle| x_*^2(t) = x\right). \quad (3.17)$$

We summarize the above discussion in the following Theorem.

Theorem 7. *In the scalar binary choice case with $A = Q = 0$, the HJB equation* (3.6) *has* *a unique solution*

$$V(t,x) = -\frac{\sigma^2 R}{B^2}\log\left[\exp\left(-\frac{B^2}{\sigma^2 R}V^1(t,x)\right)\mathbb{P}\left(x_*^1(T)\le c\middle|x_*^1(t)=x\right)\right.$$
$$\left.+\exp\left(-\frac{B^2}{\sigma^2 R}V^2(t,x)\right)\mathbb{P}\left(x_*^2(T)\ge c\middle|x_*^2(t)=x\right)\right], \tag{3.18}$$

where V^j, for $j = 1,2$, are the optimal cost-to-go functions of the standard LQG optimal control problems (3.10)*, and x_*^j, for $j = 1,2$, are the corresponding optimal states.*

3.2.2 General case

In this section, we propose an extension of (3.18), and show that it is indeed the unique solution of (3.6) in the general case. Moreover, we derive in this section an explicit formula for the min-LQG optimal control law.

The following notation is used in the remainder of this chapter. We define x_*^j, u_*^j and V^j to be the optimal state trajectory, optimal control law and optimal cost-to-go of the LQG tracking problem that the generic agent solves when p_j is the only available alternative, that is, (3.5) with $p_k = p_j$, for all $1 \le k \le l$. Recall that (Yong and Zhou, 1999, Chapter 6)

$$V^j(t,x) = \frac{1}{2}x'\Gamma(t)x + x'\beta^j(t) + \delta^j(t) \tag{3.19}$$
$$u_*^j(t,x) = -R^{-1}B'\left(\Gamma(t)x + \beta^j(t)\right) \tag{3.20}$$
$$dx_*^j(t) = \left(Ax_*^j(t) + Bu_*^j\left(t, x_*^j(t)\right)\right)dt + \sigma dw^j(t), \tag{3.21}$$

where Γ, β^j and δ^j are the unique solutions of

$$\begin{aligned}
&\frac{d}{dt}\Gamma(t) = \Gamma(t)BR^{-1}B'\Gamma(t) - A'\Gamma(t) - \Gamma(t)A - Q, && \Gamma(T) = M,\\
&\frac{d}{dt}\beta^j(t) = -\left(A - BR^{-1}B'\Gamma(t)\right)'\beta^j(t) + Q\bar{x}(t), && \beta^j(T) = -Mp_j,\\
&\frac{d}{dt}\delta^j(t) = \frac{1}{2}\beta^j(t)'BR^{-1}B'\beta^j(t) - \frac{1}{2}\mathrm{Tr}(\sigma'\Gamma(t)\sigma) - \|\bar{x}(t)\|_Q^2, && \delta^j(T) = \|p_j\|_M^2.
\end{aligned} \tag{3.22}$$

Remark 6. *The final cost in* (3.5) *is non-smooth. Hence, the corresponding HJB equation* (3.6) *and its transformed parabolic equation* (3.40) *below have non-smooth terminal conditions. However, as shown later in Lemma 3, these partial differential equations (PDE's) smooth out their solutions, i.e., the only non-smoothness occurs at the terminal time. Hence, all the PDE solutions in the remaining sections are to be understood in the strong sense.*

The term $\mathbb{P}(x_*^1(T) \leq c | x_*^1(t) = x)$ (resp. $\mathbb{P}(x_*^2(T) \geq c | x_*^2(t) = x)$) in (3.18) is the probability that an agent applying u_*^j is at time T closer to p_1 than p_2 (resp. p_2 than p_1), given that its current state is equal to x. To extend these conditional probabilities to the multidimensional multiple choice case, we define, for all $j \in \{1, \ldots, l\}$, $\mathcal{W}^j$ the Voronoi cell associated with p_j. That is, $\mathcal{W}^j = \{x \in \mathbb{R}^n,$ such that, $\|x - p_j\|_M \leq \|x - p_k\|_M, \forall 1 \leq k \leq l\}$. The generalized versions of these conditional probabilities are thus defined as follows,

$$g^j(t, x) := \mathbb{P}\left(x_*^j(T) \in \mathcal{W}^j \middle| x_*^j(t) = x\right), \tag{3.23}$$

where x_*^j defined in (3.21) is the optimal state of the LQG problem that corresponds to p_j.

As in the special case, we linearize the HJB equation (3.6) using a generalized Hopf-Cole transformation, see proof of Theorem 8 below. In the multidimensional case, however, we make the following assumption.

Assumption 6 (Control efficiency / noise intensity isotropy). *We assume that there exists a scalar $\eta > 0$, such that, $BR^{-1}B' = \eta\sigma\sigma'$.*

Remark 7. *Note the following:*

i. *Suppose that $n = m$, $A = Q = 0$, $B = diag(b_1, \ldots, b_n)$, $R = diag(r_1, \ldots, r_n)$ and $\sigma = diag(\sigma_1, \ldots, \sigma_n)$. σ_j is the intensity of the noise in the $j-th$ direction. The control variable acts on the state in the $j-th$ direction through the coefficient b_j, and the cost of this action in the same direction is evaluated through the coefficient r_j. Hence, the ratio b_j^2/r_j measures the efficiency of the control variable in the $j-th$ direction. Following these interpretations, Assumption 6 requires the ratio of the control efficiency to the noise intensity be identical in all the directions. In other words, it imposes a sort of isotropy on the ratio "control efficiency / noise intensity".*

ii. *Assumption 6 always holds in the scalar case ($n = m = 1$).*

iii. *If η exists, then it is strictly positive.*

iv. *If η exists, since we assumed that σ is invertible, so is $BR^{-1}B'$, and thus ker $B' = \{0\}$. In particular, the dimension of the control space is greater or equal to that of the state space ($m \geq n$). Then we must choose $R = \frac{1}{\eta}B'(\sigma\sigma')^{-1}B$, for some $\eta > 0$.*

v. *Assumption 6 is satisfied in particular if $B = R = \sigma = I_n$, a situation that has been studied previously in the context of other mean-field games (with $A = 0$) using the Hopf-Cole transformation, see (Gomes et al., 2016, Chapter 2) and the references therein.*

We now state the main result of this section.

Theorem 8 (Min-LQG optimal cost-to-go)**.** *Under Assumption 6, the HJB equation* (3.6) *has a unique solution* $(t,x) \mapsto V(t,x)$ *in* $C^{1,2}([0,T) \times \mathbb{R}^n) \cap C([0,T] \times \mathbb{R}^n)$*, defined as*

$$\begin{aligned} V(t,x) &= -\frac{1}{\eta} \log \left(\sum_{j=1}^{l} \exp\left(-\eta V^j(t,x)\right) g^j(t,x) \right), \forall (t,x) \in [0,T) \times \mathbb{R}^n \\ V(T,x) &= \min_{j=1,\dots,l} \|x - p_j\|_M^2, \ \forall x \in \mathbb{R}^n, \end{aligned} \tag{3.24}$$

where V^j *and* x_*^j *are defined in* (3.19) *and* (3.21)*.*

Before proving Theorem 8, we need some preliminary results. We start with a technical Lemma on the mean-square convergence of random variables.

Lemma 2. *Let* I *be a closed subset of* $\mathbb{R}^n$*. Let* $X_k \in \mathbb{R}^n$ *be a sequence of random variables with finite first and second moments. If* $\mathbb{E}[X_k] =: \mu_k \to \mu$ *for some vector* μ *not in* I*, and* $\mathbb{E}[\|X_k - \mu_k\|^2] \to 0$*, then* $\lim_{k\to\infty} \mathbb{P}(X_k \in I) = 0$*.*

Proof. $I \subset \mathbb{R}^n$ is a closed set and $\mu \notin I$, so the distance d between μ and I is strictly positive. Since, μ_k converges to μ, there exists $k_0 > 0$ such that for all $k \geq 0$ and for all x in I we have $\|x - \mu_k\| \geq d/2$. Hence, using Chebyshev's inequality (Durrett, 2010, Theorem 1.6.4),

$$\mathbb{P}(X_k \in I) \leq \mathbb{P}(\|X_k - \mu_k\| \geq d/2) \leq \frac{4}{d^2} \mathbb{E}[\|X_k - \mu_k\|^2], \tag{3.25}$$

for all $k \geq k_0$. The result follows since the right-hand side of the inequality is assumed to converge to 0. □

The following lemma concerns the regularity of the solution provided in Theorem 8.

Lemma 3 (Regularity of the solution)**.** V *defined in* (3.24) *is in* $C^{1,2}([0,T) \times \mathbb{R}^n) \cap C([0,T] \times \mathbb{R}^n)$*.*

Proof. Note that for g_j defined in (3.23), we can write

$$\begin{aligned} g^j(t,x) &= \mathbb{P}\left(x_*^j(T) \in \mathcal{W}^j \middle| x_*^j(t) = x\right) \\ &= \frac{1}{\sqrt{|2\pi\Sigma_t|}} \int_{\mathcal{W}^j} \exp\left(- \left\| y - \phi(T,t)x + \int_t^T \phi(T,\tau) B R^{-1} B' \beta^j(\tau) d\tau \right\|_{\Sigma_t^{-1}}^2 \right) dy \end{aligned} \tag{3.26}$$

for $\Sigma_t = \int_t^T \phi(T,\tau) \sigma \sigma' \phi(T,\tau)' d\tau$, where x_*^j, Γ and β^j are defined in (3.21) and (3.22), and the matrix-valued function $\phi(t,s)$ is the unique solution of

$$\frac{d}{dt}\phi(t,s) = \left(A - BR^{-1}B'\Gamma(t)\right)\phi(t,s), \tag{3.27}$$

with $\phi(s,s) = I_n$. The expression (3.26) follows from the fact that the solution of a linear SDE with deterministic initial condition has a normal distribution (Karatzas and Shreve, 2012, Section 2.5). In view of (3.24), (3.26), and $\Sigma_t \succ 0$ for all $t \in [0,T)$, V is in $C^{1,2}([0,T) \times \mathbb{R}^n)$. It remains to show the continuity on $\{T\} \times \mathbb{R}^n$. We start by considering $x \in \mathbb{R}^n \setminus \cup_{j=1}^l \partial\mathcal{W}^j$, and $(t_k, x_k) \in [0,T) \times \mathbb{R}^n$ converging to (T,x). We have $x \in \overset{\circ}{\mathcal{W}}{}^{j_0}$ for some $j_0 \in \{1,\dots,l\}$, and $x \notin \mathcal{W}^j$ for $j \neq j_0$. In view of (3.26), $g^j(t_k,x_k)$ is the probability that a Gaussian vector of mean $\phi(T,t_k)x_k - \int_{t_k}^T \phi(T,\tau)BR^{-1}B'\beta^j(\tau)d\tau$ (which converges to x with k) and covariance Σ_{t_k} (which converges to 0 with k) is in the closed set $\mathcal{W}^j$. In this way, each j defines a distinct sequence of random variables associated with the (t_k,x_k)'s. Now if one considers the closed set I of Lemma 2 to be any of the closed sets $\mathcal{W}^j$'s for $j \neq j_0$, one can conclude from this lemma that $g^j(t_k,x_k)$ must converge to 0 for $j \neq j_0$ and, as a consequence, to 1 for $j = j_0$ since the $\mathcal{W}^j$'s form a partition of the state space. Therefore, $V(t_k,x_k)$ converges to $V(T,x)$. Thus, V is continuous on $[0,T] \times \left(\mathbb{R}^n \setminus \cup_{j=1}^l \partial\mathcal{W}^j\right)$. Finally, consider a sequence $(t_k,x_k) \in [0,T) \times \mathbb{R}^n$ converging to (T,c), with $c \in \cup_{j=1}^l \partial\mathcal{W}^j$. We show that $V(t_k,x_k)$ converges to $V(T,c)$. Up to renumbering the Voronoi cells, we can assume without loss of generality that $c \in \partial\mathcal{W}^j$ for all $j \in \{1,\dots,i_0\}$ and $c \notin \cup_{j=i_0+1}^l \mathcal{W}^j$, for some $1 \leq i_0 \leq l$. We have,

$$\begin{aligned}
\xi_0 &= \sum_{j=1}^{l} \exp\left(-\eta V^j(t_k,x_k)\right) g^j(t_k,x_k) \\
&= \sum_{j=1}^{i_0} \exp\left(-\eta V^j(t_k,x_k)\right) g^j(t_k,x_k) + \sum_{j=i_0+1}^{l} \exp\left(-\eta V^j(t_k,x_k)\right) g^j(t_k,x_k).
\end{aligned} \tag{3.28}$$

Since $c \notin \cup_{j=i_0+1}^l \mathcal{W}^j$, one can use an argument similar to that above to show that the second term of the right-hand side of the second equality converges to 0.

Next, let $\epsilon > 0$ and fix $r > 0$ small enough so that $\overline{B(c,r)} \subset \left(\cap_{j=i_0+1}^l \mathcal{W}^j\right)^c$. The value of r will be determined later. The first term can be written

$$\sum_{j=1}^{i_0} \exp\left(-\eta V^j(t_k,x_k)\right) g^j(t_k,x_k) = \xi_1 + \xi_2, \tag{3.29}$$

where

$$\xi_1 = \sum_{j=1}^{i_0} \exp\left(-\eta V^j(t_k, x_k)\right) \mathbb{P}\left(x_*^j(T) \in \mathcal{W}^j \cap \overline{B(c,r)} \middle| x_*^j(t_k) = x_k\right) \tag{3.30}$$

$$\xi_2 = \sum_{j=1}^{i_0} \exp\left(-\eta V^j(t_k, x_k)\right) \mathbb{P}\left(x_*^j(T) \in \mathcal{W}^j \setminus \overline{B(c,r)} \middle| x_*^j(t_k) = x_k\right). \tag{3.31}$$

By Lemma 2, ξ_2 converges to zero. Next, by solving the linear ODE's of β^j and δ^j in (3.22) we get,

$$\beta^j(t) = -\phi(T,t)' M p_j + \int_T^t \phi(s,t)' Q \bar{x}(s) ds \tag{3.32}$$

$$\begin{aligned} \delta^j(t) = &-\frac{1}{2}\int_T^t \mathrm{Tr}(\sigma'\Gamma(\tau)\sigma)d\tau - \int_T^t \|\bar{x}(\tau)\|_Q^2 d\tau + \frac{1}{2} p_j' M \int_T^t \phi(T,\tau) B R^{-1} B' \phi(T,\tau)' d\tau M p_j \\ &- p_j' M \int_T^t \int_T^\tau \phi(T,\tau) B R^{-1} B' \phi(s,\tau)' Q \bar{x}(s) ds d\tau \\ &+ \frac{1}{2}\int_T^t \int_T^\tau \int_T^\tau \left(\phi(s,\tau)' Q \bar{x}(s)\right)' \phi(r,\tau)' Q \bar{x}(r) ds dr d\tau. \end{aligned} \tag{3.33}$$

By replacing (3.32)-(3.33) in the expressions (3.26) of g^j and (3.19) of V^j, one can show that under Assumption 6,

$$\xi_1 = \exp\left(-\eta V_0(t_k, x_k)\right) \sum_{j=1}^{i_0} \int_{\mathcal{W}^j \cap \overline{B(c,r)}} f_k(y) \exp\left(\eta(\|y\|_M^2 - \|y - p_j\|_M^2)\right) dy, \tag{3.34}$$

where $f_k(y)$ is the probability density function of the Gaussian distribution of mean $\phi(T, t_k) x_k - \int_{t_k}^T \phi(T,\tau) B R^{-1} B' \beta_0(\tau) d\tau$ and covariance matrix Σ_{t_k}, and V_0 and β_0 are equal to V^j and β^j defined in (3.19)-(3.22) but for $p_j = 0$. By the definition of c, $\|c - p_1\|_M^2 = \cdots = \|c - p_{i_0}\|_M^2$. Hence,

$$\begin{aligned} \xi_1 &= \exp\left(-\eta(V_0(t_k, x_k) - \|c\|_M^2 + \|c - p_j\|_M^2)\right) \sum_{j=1}^{i_0} \int_{\mathcal{W}^j \cap \overline{B(c,r)}} f_k(y) dy \\ &\quad + \exp\left(-\eta V_0(t_k, x_k)\right) \sum_{j=1}^{i_0} \int_{\mathcal{W}^j \cap \overline{B(c,r)}} f_k(y) f(y) dy \\ &\triangleq \xi_3 + \xi_4, \end{aligned} \tag{3.35}$$

where $f(y) = \exp\left(\eta(\|y\|_M^2 - \|y - p_j\|_M^2)\right) - \exp\left(\eta(\|c\|_M^2 - \|c - p_j\|_M^2)\right)$. $V_0(t_k, x_k)$ converges to $V_0(T, c) = \|c\|_M^2$. f_k converges in distribution to a point mass at c, and $\mathcal{W}^j \cap \overline{B(c,r)}$, $j = 1, \ldots, i_0$, is a partition of $\overline{B(c,r)}$. Therefore, ξ_3 converges to $\exp(-\eta \|c - p_j\|_M^2) = \exp(-\eta V(T, c))$. f is continuous, and $f(c) = 0$. Hence, one can choose r small enough so that $|f(y)| < \epsilon$ for all $y \in \overline{B(c,r)}$. Thus, $|\xi_4| \leq \epsilon$, and $\limsup_k |\xi_0 - \exp(-\eta V(T,c))| \leq \epsilon$. Since ϵ is arbitrary, ξ_0 converges to $\exp(-\eta V(T,c))$. This proves the result. □

Proof of Theorem 8. To finish the proof of Theorem 8, it remains to show that V satisfies the HJB equation (3.6). We define the transformations by a generalized Hopf-Cole transformation (Evans, 1998, Chapter 4-Section 4.4) of $V^j(t,x)$, $\psi^j(t,x) = \exp\left(-\eta V^j(t,x)\right)$, for $j = 1,\dots,l$. Recall (Yong and Zhou, 1999, Chapter 6) that the optimal cost-to-go V^j satisfies the HJB equation (3.6), but with the boundary condition equal to $V^j(T,x) = \|x - p_j\|_M^2$. By multiplying the right-hand and left-hand sides of (3.6) by $-\eta \exp\left(-\eta V^j(t,x)\right)$, one obtain that

$$\begin{aligned}-\frac{\partial \psi^j}{\partial t} =& x'A'\frac{\partial \psi^j}{\partial x} + \frac{1}{2}\mathrm{Tr}\left(\sigma'\frac{\partial^2 \psi^j}{\partial x^2}\sigma\right) - \eta\|x - \bar{x}\|_Q^2\psi^j \\ &+ \eta \exp\left(-\eta V^j(t,x)\right)\frac{1}{2}\left(\frac{\partial V^j}{\partial x}\right)'\left(BR^{-1}B' - \eta\sigma\sigma'\right)\frac{\partial V^j}{\partial x}.\end{aligned} \tag{3.36}$$

Thus, under Assumption 6, we get

$$\begin{aligned}-\frac{\partial \psi^j}{\partial t} &= x'A'\frac{\partial \psi^j}{\partial x} + \frac{1}{2}\mathrm{Tr}\left(\sigma'\frac{\partial^2 \psi^j}{\partial x^2}\sigma\right) - \eta\|x - \bar{x}\|_Q^2\psi^j \\ \psi^j(T,x) &= \exp\left(-\eta\|x - p_j\|_M^2\right), \quad \forall x \in \mathbb{R}^n.\end{aligned} \tag{3.37}$$

Define $\psi(t,x) = \exp\left(-\eta V(t,x)\right)$ the transformation of $V(t,x)$ defined in (3.24). Hence, we have $\psi(t,x) = \sum_{j=1}^{l}\psi^j(t,x)g^j(t,x)$. Equation (3.37), Assumption 6 and the identity $\frac{\partial \psi^j}{\partial x} = -\eta\left(\Gamma x + \beta^j\right)\psi^j$, where Γ and β^j are defined in (3.22), imply

$$\begin{aligned}&\frac{\partial \psi}{\partial t} + x'A'\frac{\partial \psi}{\partial x} + \frac{1}{2}\mathrm{Tr}\left(\sigma'\frac{\partial^2 \psi}{\partial x^2}\sigma\right) - \eta\|x - \bar{x}\|_Q^2\psi \\ &= \sum_{j=1}^{l}\left(\frac{\partial g^j}{\partial t} + \left(Ax - BR^{-1}B'\Gamma x - BR^{-1}B'\beta^j\right)'\frac{\partial g^j}{\partial x} + \frac{1}{2}\mathrm{Tr}\left(\sigma'\frac{\partial^2 g^j}{\partial x^2}\sigma\right)\right)\psi^j.\end{aligned} \tag{3.38}$$

The process x_*^j satisfies the SDE (3.21). Therefore, by Kolmogorov's backward equation (Karatzas and Shreve, 2012, Section 5.B),

$$\frac{\partial g^j}{\partial t} + \left(Ax - BR^{-1}B'\Gamma x - BR^{-1}B'\beta^j\right)'\frac{\partial g^j}{\partial x} + \frac{1}{2}\mathrm{Tr}\left(\sigma'\frac{\partial^2 g^j}{\partial x^2}\sigma\right) = 0. \tag{3.39}$$

Hence,

$$\frac{\partial \psi}{\partial t} + x'A'\frac{\partial \psi}{\partial x} + \frac{1}{2}\mathrm{Tr}\left(\sigma'\frac{\partial^2 \psi}{\partial x^2}\sigma\right) - \eta\|x - \bar{x}\|_Q^2\psi = 0. \tag{3.40}$$

By multiplying the right and left-hand sides of (3.40) by $\frac{1}{\eta}\exp(\eta V(t,x))$, we obtain that $V(t,x)$ satisfies (3.6). The uniqueness of the solution follows from the regularity of V (See Lemma 3) and the uniqueness of solutions to the uniform parabolic PDE (3.40) (Karatzas and Shreve, 2012, Theorem 7.6). □

Having solved the HJB equation related to the Min-LQG optimal control problem (3.5), we now prove the existence of a unique optimal control law and derive an explicit formula for this law. We define the following function:

$$\begin{aligned} u_*(t,x) &= -R^{-1}B'\frac{\partial V}{\partial x}(t,x), \qquad \forall (t,x)\in[0,T)\times\mathbb{R}^n, \\ u_*(T,x) &= 0, \qquad \forall x\in\mathbb{R}^n. \end{aligned} \tag{3.41}$$

Theorem 9 (Min-LQG optimal control law)**.** *Under Assumption 6, the following statements hold:*

(i) The function u_* defined in (3.41) has on $[0,T)\times\mathbb{R}^n$ the following form:

$$u_*(t,x) = \sum_{j=1}^{l} \frac{\exp\left(-\eta V^j(t,x)\right) g^j(t,x)}{\sum_{k=1}^{l}\exp\left(-\eta V^k(t,x)\right) g^k(t,x)} u_*^j(t,x), \tag{3.42}$$

with V^j, u_*^j and g^j respectively defined in (3.19), (3.20) and (3.23).

(ii) u_* is an admissible Markov policy, i.e. $u_*\in\mathcal{L}$ defined in (3.4).

(iii) $u_*\left(t,x_*(t,w)\right)$ is the unique optimal control law of (3.5), where $x_*(t,w)$ is the unique strong solution of the SDE in (3.5) with u equal to $u_*(t,x)$.

Proof.

(i) We have for all $(t,x)\in[0,T)\times\mathbb{R}^n$,

$$\begin{aligned} u_*(t,x) = -R^{-1}B'\frac{\partial V}{\partial x} &= \sum_{j=1}^{l}\frac{\psi^j(t,x)g^j(t,x)}{\sum_{k=1}^{l}\psi^k(t,x)g^k(t,x)}u_*^j(t,x) \\ &+ \frac{1}{\eta\sum_{k=1}^{l}\psi^k(t,x)g^k(t,x)}R^{-1}B'\sum_{j=1}^{l}\psi^j(t,x)\frac{\partial g^j}{\partial x}(t,x), \end{aligned} \tag{3.43}$$

where ψ^j defined in the proof of Theorem 8 is the transformation of V^j by the Hopf-Cole transformation. In the following, we show that the second summand is zero. By the change of variable $z = y - \phi(T,t)x + \int_t^T \phi(T,\tau)BR^{-1}B'\beta^j(\tau)d\tau$ in the expression of g^j (3.26) and Leibniz integral rule, we obtain that

$$\frac{\partial g^j}{\partial x}(t,x) = \frac{-\phi(T,t)}{\sqrt{|2\pi\Sigma_t|}} \int_{\partial\mathcal{W}^j - \phi(T,t)x + \int_t^T \phi(T,\tau)BR^{-1}B'\beta^j(\tau)d\tau} \exp\left(-\|z\|^2_{\Sigma_t^{-1}}\right) \vec{n}^j(z)ds(z) \tag{3.44}$$

$$= \frac{-\phi(T,t)}{\sqrt{|2\pi\Sigma_t|}} \int_{\partial\mathcal{W}^j} \exp\left(-\left\|y - \phi(T,t)x + \int_t^T \phi(T,\tau)BR^{-1}B'\beta_j(\tau)d\tau\right\|^2_{\Sigma_t^{-1}}\right) \vec{n}^j(y)ds(y),$$

where ϕ is defined in (3.27) and $\vec{n}^j(y)$ is the unit normal component of $\partial\mathcal{W}^j$ and its translation $\partial\mathcal{W}^j - \phi(T,t)x + \int_t^T \phi(T,\tau)BR^{-1}B'\beta^j(\tau)d\tau$. By replacing the expression of β^j and δ^j (3.32)-(3.33) in the expressions of the costs V^j defined in (3.19) and in the derivatives $\frac{\partial g^j}{\partial x}$, one can show that under Assumption 6,

$$\sum_{j=1}^{l} \psi^j(t,x)\frac{\partial g^j}{\partial x}(t,x)$$
$$= K_1(t,x)\sum_{j=1}^{l} \int_{\partial\mathcal{W}^j} \exp\left(K_2(t,x,y) + \eta\|y-p_j\|^2_M - \eta\|y\|^2_M\right) \vec{n}^j(y)ds(y), \tag{3.45}$$

where K_1 and K_2 are functions that do not depend on p_j, $\forall j \in \{1,\dots,l\}$. Note that $\partial\mathcal{W}^j = \cup_{i=1}^{k_j} O_i$, where the disjoint subsets (up to a subset of measure zero) $\{O_i\}_{i=1}^{k_j}$ are the common boundaries of $\mathcal{W}^j$ and the adjacent Voronoi cells. If O_i is the common boundary of $\mathcal{W}^j$ and some adjacent Voronoi Cell $\mathcal{W}^k$, then $\vec{n}^j(y) = -\vec{n}^k(y)$ for all $y \in O_i$. Moreover, by the definition of the Voronoi cells, $\|y-p_j\|_M = \|y-p_k\|_M$ for all $y \in O_i$. Therefore, the right-hand side of (3.45) is equal to zero. This proves the first point.

(ii) We show now that u_* is an admissible Markov policy. In view of (3.42), the function $\frac{\partial u_*}{\partial x}$ is continuous on $[0,T) \times \mathbb{R}^n$. Therefore, the local Lipschitz condition holds. Moreover, for all $(t,x) \in [0,T] \times \mathbb{R}^n$, we have

$$\|u_*(t,x)\| \le \sum_{j=1}^{l} \left\|u_*^j(t,x)\right\| \le \|R^{-1}B'\| \left(l\|\Gamma\|_\infty\|x\| + \sum_{j=1}^{l}\|\beta^j\|_\infty\right). \tag{3.46}$$

Hence, the linear growth condition is satisfied and this proves the second point. As a result, sufficient conditions are satisfied for the SDE defined in (3.5) and controlled by $u_*(t,x)$ to have a unique strong solution denoted x_* (Karatzas and Shreve, 2012, Section 5.2).

(iii) By the verification theorem (Fleming and Soner, 2006, Theorem 4.3.1), we know that u_* is the unique optimal control law of (3.5) if it is the unique minimizer (up to a set of measure 0) of the Hamiltonian $H(x, \frac{\partial V}{\partial x}, u, t) = (Ax + Bu)'\frac{\partial V}{\partial x} + \|x - \bar{x}\|_Q^2 + \|u\|_R^2$, and if the cost-to-go $V(t, x)$ has a polynomial growth in x and satisfies the HJB equation (3.6). For the first condition, we have for Lebesgue $\times$ $\mathbb{P}$-a.e $(t, \xi) \in [0, T] \times \Omega$,

$$u_* \left(t, x_*(t, \xi)\right) = -R^{-1}B'\frac{\partial V}{\partial x}\left(t, x_*(t, \xi)\right) = \operatorname*{argmin}_{u \in \mathbb{R}^n} H\left(x_*(t, \xi), \frac{\partial V}{\partial x}(t, x_*(t, \xi)), u, t\right). \tag{3.47}$$

In fact, the control law defined in (3.41) minimizes H except on the set $\{T\} \times \Omega$, which has a Lebesgue $\times$ $\mathbb{P}$ measure zero. Next, in view of (3.46) and the mean value theorem in several variables, we have for all $(t, x) \in [0, T) \times \mathbb{R}^n$, $\|V(t, x)\| \leq K_1(1 + \|x\|^2)$, for some $K_1 > 0$. Moreover, $\|V(T, x)\| \leq K_2(1 + \|x\|^2)$, for some $K_2 > 0$. Hence, for all $(t, x) \in [0, T] \times \mathbb{R}^n$, $\|V(t, x)\| \leq K\left(1 + \|x\|^2\right)$, for some $K > 0$. Moreover, as established in Theorem 3, $V \in C^{1,2}\left([0, T) \times \mathbb{R}^n\right) \cap C([0, T] \times \mathbb{R}^n)$ satisfies the HJB equation (3.6). This proves the result.

□

We show in the degenerate case (See Theorem 3) that the probability that an agent optimally responding to a sustainable mean trajectory reaches an arbitrary small neighborhood of one of the alternatives increases with M. In the following theorem, we extend this result to the non-degenerate scalar case. The proof relies on the analysis of the scalar version of the Riccati equation in (3.22). The extension of this analysis, and consequently the result of Theorem 10, to the multidimensional case is left for future work.

Theorem 10 (Well-defined discrete choice problem)**.** *In the scalar case* $(n = m = 1)$, *suppose that the paths* $\bar{x}$ *are uniformly bounded with respect to* M *for the norm* $\left(\int_0^T |.|^2 \mathrm{dt}\right)^{\frac{1}{2}}$. *Then, for any* $\epsilon > 0$,

$$\mathbb{P}\left(\bigcap_{j=1}^{l} \{|x_*(T) - p_j| > \epsilon\}\right) = O\left(\frac{1}{\epsilon^2 M} + \frac{\sigma^2}{2\epsilon^2}\frac{\log M}{M}\right). \tag{3.48}$$

Proof. To prove the result, it is sufficient to show that the expectation of the optimal cost $\mathbb{E}J_*(x(0)) \leq K + \frac{\sigma^2}{2}\log M$, for some $K > 0$ independent of M. The result is then a direct consequence of Chebyshev's inequality

$$\mathbb{P}\left(\min_{1 \leq j \leq l} |x_*(T) - p_j| > \epsilon\right) \leq \frac{1}{\epsilon^2}\mathbb{E}\min_{1 \leq j \leq l} |x_*(T) - p_j|^2 \leq \frac{2}{M\epsilon^2}\mathbb{E}J_*(x(0)). \tag{3.49}$$

To prove the boundedness of the cost, we start by the special case where $\bar{x} = 0$ and $p_j = 0$, for all $1 \leq j \leq l$, i.e. the LQG case. In view of the expression of V^j (3.19), the optimal cost is equal to $J_*(x(0)) = \frac{1}{2}\Gamma(0)x^2(0) + \frac{\sigma^2}{2}\int_0^T \Gamma(\tau)d\tau$. We now show that $\Gamma(0)$ is uniformly bounded with M and that $\int_0^T \Gamma(\tau)d\tau$ is of the order $\log M$ for large M. To prove the uniform boundness of $\Gamma(0)$, we consider the following LQR problem,

$$\begin{aligned} &\inf_v J_M(s,v) = \inf_v \int_s^T \left\{ \frac{Q}{2}y^2 + \frac{R}{2}v^2 \right\} dt + \frac{M}{2}|y(T)|^2 \\ &\text{s.t. } \frac{dy}{dt} = Ay + Bv, \qquad y(s) = 1. \end{aligned} \tag{3.50}$$

The optimal cost is $\inf_v J_M(0,v) = \frac{1}{2}\Gamma(0)$, where Γ is defined in (3.22). By the controllability of (A,B), one can find a continuous control law v_{10} that does not depend on M and, such that, the corresponding state y_{10} is at time T at 0. We have $\frac{1}{2}\Gamma(0) \leq J_M(0, v_{10})$. The right hand side of the inequality is finite and does not depend on M. Hence, $\Gamma(0)$ is uniformly bounded w.r.t. M. We now prove that $\int_0^T \Gamma(\tau)d\tau$ is of the order $\log M$ for large M. We have for all $M \geq 1$, $\frac{1}{2}\Gamma(s) = \inf_v J_M(s,v) \geq \inf_v J_1(s,v) = \frac{1}{2}\Gamma_1(s)$, where Γ_1 is equal to Γ in (3.22) but for $M = 1$. Hence, for all $M \geq 1$, $\min_{t\in[0,T]} \Gamma(t) \geq \min_{t\in[0,T]} \Gamma_1(t) := C > 0$. C does not depend on M. The last inequality follow from that fact that Γ_1 is continuous and strictly positive for all $t \in [0,T]$. By dividing by $\Gamma(t)$ on both sides of the Riccati equation in (3.22) and integrating on $[0,T]$ the right and left hand sides, we get

$$\log M - \log \Gamma(0) = \frac{B^2}{R}\int_0^T \Gamma(\tau)d\tau - 2AT - Q\int_0^T \frac{1}{\Gamma(\tau)}d\tau. \tag{3.51}$$

By the boundedness of $\Gamma(0)$ and $\left|\int_0^T \frac{1}{\Gamma(\tau)}d\tau\right| \leq \frac{1}{CT}$, we have $\int_0^T \Gamma(\tau)d\tau / \log M$ converges to 1 as M goes to infinity. Having shown the result for the special case, the case where $\bar{x} \neq 0$ and $p_1 = \cdots = p_l = p$ can be proved by making the change of variables $\tilde{x} = x - p$ and $\tilde{u} = u + \frac{A}{B}p$ and noting the uniform boundedness of $\bar{x}$ and that

$$\mathbb{E}J_*(x(0)) \leq \mathbb{E}\inf_{\tilde{u}} \mathbb{E}\left(\int_0^T \{Q\tilde{x}^2 + R\tilde{u}^2\}dt + \frac{M}{2}\tilde{x}^2(T)|x(0)\right) + \int_0^T Q(\bar{x}-p)^2 dt + \frac{RA^2T}{B^2}p^2. \tag{3.52}$$

Finally, we conclude the general case by the following inequality

$$\mathbb{E}J_*(x(0)) \leq \mathbb{E}\inf_u \mathbb{E}\left(\int_0^T \left\{\frac{Q}{2}(x-\bar{x})^2 + \frac{R}{2}u^2\right\}dt + \frac{M}{2}(x(T)-p_1)^2|x(0)\right). \tag{3.53}$$

□

In the degenerate case, the probability that a generic agent is at time T far from the alternatives converges to zero with order $1/M$ (See Theorem 3). In case $\sigma > 0$, however, the rate of convergence is $\sigma^2 \log M/2M$. Thus, to force the agents to be near the alternatives at time T, one needs to impose a penalty on the final distances to the alternatives greater in the presence of a noise than in its absence. This is due to the diffusive effect of the Brownian motion.

3.3 Mean field equations

This section considers the problem of finding a sustainable mean trajectory $\bar{x}$. In the remaining of this chapter, a subscript s refers to an agent with parameters $\Theta_s \in \{\Theta_1, \dots, \Theta_k\}$, where the set of parameters is defined in Section 3.1. We write $u_{s*}(t, x, \bar{x})$ instead of $u_{s*}(t, x)$ defined in (3.42) to emphasize the dependence on $\bar{x}$. A sustainable mean trajectory $\bar{x}$ must satisfy the following mean field equations,

$$\begin{aligned} \bar{x}(t) &= \sum_{s=1}^{k} \alpha_s \bar{x}_s(t), \text{ with } \bar{x}_s = \mathbb{E}[x_{s*}], \ 1 \le s \le k, \\ dx_{s*}(t) &= (A_s x_{s*}(t) + B_s u_{s*}(t, x_{s*}(t), \bar{x}))\, dt + \sigma_s dw_s(t), \end{aligned} \tag{3.54}$$

where $x_{s*}(0)$, $1 \le s \le k$, are i.i.d. according to the initial distribution of the agents, i.e., the distribution of $x_i(0)$, α_s are defined in Section 3.1 and $\{w_1, \dots, w_k\}$ are k independent Brownian motions, assumed independent of $\{x_{1*}(0), \dots, x_{k*}(0)\}$. The first equality follows from the fact that the average state when the agents optimally respond to $\bar{x}$ should converge to $\bar{x}$. But, the average state is equal to the weighted average of the averages of states in each type s, $1 \le s \le k$. The weight that corresponds to type s is equal to $P_\theta^N(\Theta_s)$, which is assumed to converge to α_s. Moreover, by the independence of the initial conditions and Brownian motions, the average state of the agents of type Θ_s converges by the strong law of large numbers to $\bar{x}_s(t)$.

To solve (3.54), we start as in the degenerate case by deriving an equivalent representation of the solution via the stochastic maximum principle. It consists of two forward-backward ODE's (3.60)-(3.61) coupled in the boundary condition $\bar{q}(T)$ through what we call a "Choice Distribution Matrix (CDM)" $\Lambda(\bar{x})$. A CDM $\Lambda(\bar{x})$ is a $k \times l$ row stochastic matrix with its (s, j) entry equal to the probability that a generic agent of type s is at time T closer (in the sense of the M-weighted l_2 norm) to p_j than any of the other alternatives when it optimally responds to $\bar{x}$. Afterwards, we construct a one-to-one map between the sustainable mean trajectories and the fixed point CDM's of a finite dimensional map F defined in (3.68) below. These fixed points are the potential distributions of the agents' choices over the alternatives.

In the remainder of this section, we adopt the following notations. Let $\bar{X} = (\bar{x}_1, \ldots, \bar{x}_k)$, $X_* = (x_{1*}, \ldots, x_{k*})$, $U_* = (u_{1*}, \ldots, u_{k*})$, $W = (w_1, \ldots, w_k)$, and $p = (p_1, \ldots, p_l)$. Let A, B, Q, R, M and σ be the block-diagonal matrices $\mathrm{diag}(A_1, \ldots, A_k)$, $\mathrm{diag}(B_1, \ldots, B_k)$, $\mathrm{diag}(Q_1, \ldots, Q_k)$, $\mathrm{diag}(R_1, \ldots, R_k)$, $\mathrm{diag}(M_1, \ldots, M_k)$ and $\mathrm{diag}(\sigma_1, \ldots, \sigma_k)$ respectively. Define $L = I_{nk} - 1_k \otimes P_1$, where 1_k is a column of k ones and $P_1 = P'_\theta \otimes I_n$ where P_θ is defined in Section 3.1 . The following assumption guarantees the existence and uniqueness of a solution for (3.60)-(3.61) whenever the CDM in the $\bar{q}(T)$ boundary condition is considered as a parameter.

Assumption 7. *We assume the existence of a solution on* $[0, T]$ *to the following (nonsymmetric) Riccati equation:*

$$\frac{d}{dt}\gamma = -A'\gamma - \gamma A + \gamma B R^{-1} B' \gamma + QL, \quad \gamma(T) = M. \tag{3.55}$$

Note that if Assumption 7 is satisfied, the solution of (3.55) is unique as a consequence of the smoothness of the right-hand side of (3.55) with respect to γ (Perko, 2013, Section 2.4, Lemma 1). For a uniform population, i.e., $k = 1$, we have $L = 0$ and hence Assumption 7 is always satisfied (Anderson and Moore, 2007, Section 2.3). For more details about Assumption 7, one can refer to (Freiling, 2002).

In the following lemma, we derive a stochastic maximum principle for the non-degenerate min-LQG problem.

Lemma 4 (Min-LQG stochastic maximum principle)**.** *The processes* $\left(q_{s*}(t), \frac{\partial^2 V_s}{\partial x^2}(t, x_{s*}(t))\right)$, $1 \leq s \leq k$, *with* $q_{s*}(t) = \frac{\partial V_s}{\partial x}(t, x_{s*}(t))$, *satisfy the following backward linear SDE:*

$$-\,dq_{s*}(t) = \left(A'_s q_{s*}(t) + Q_s(x_{s*}(t) - \bar{x}(t))\right) dt - \frac{\partial^2 V_s}{\partial x^2}\left((t, x_{s*}(t)\right) \sigma_s dw_s(t), \tag{3.56}$$

with $q_{s*}(T) = M_s\left(x_{s*}(T) - \sum_{j=1}^{l} 1_{\mathcal{W}^j}(x_{s*}(T)) p_j\right)$.

Proof. The function $\frac{\partial V_s}{\partial x}(t, x)$ is smooth on $[0, T) \times \mathbb{R}^n$. By applying Itô's formula (Karatzas and Shreve, 2012, Section 3.3.A) to $\frac{\partial V_s}{\partial x}\left(t, x_{s*}(t)\right)$, and by noting that V_s satisfies the HJB equation (3.6), we have

$$-\,dq_{s*}(t) = \left(A'_s q_{s*}(t) + Q_s(x_{s*}(t) - \bar{x}(t))\right) dt - \frac{\partial^2 V_s}{\partial x^2}\left((t, x_{s*}(t)\right) \sigma_s dw_s(t), \tag{3.57}$$

with $q_{s*}(0) = \frac{\partial V_s}{\partial x}\left(0, x_{s*}(0)\right)$. It remains to show that $\mathbb{P}$-a.s

$$\lim_{t \to T} \frac{\partial V_s}{\partial x}\left(t, x_{s*}(t)\right) = M_s\left(x_{s*}(T) - \sum_{j=1}^{l} 1_{\mathcal{W}^j}(x_{s*}(T)) p_j\right). \tag{3.58}$$

By Theorem 9, we have on $[0,T)\times\mathbb{R}^n$

$$\frac{\partial V_s}{\partial x}(t,x)=\sum_{j=1}^{l}\frac{\exp\left(-\eta_s V_s^j(t,x)\right)g_s^j(t,x)}{\sum_{k=1}^{l}\exp\left(-\eta_s V_s^k(t,x)\right)g_s^k(t,x)}\frac{\partial V_s^j}{\partial x}(t,x). \tag{3.59}$$

Fix $j\in\{1,\dots,l\}$. By Lemma 2, we have on $\{x_{s*}(T)\in\overline{\mathcal{W}}^j\}$, $\lim_{t\to T} g_s^j(t,x_{s*}(t))=1$ and $\lim_{t\to T} g_s^k(t,x_{s*}(t))=0$, for all $k\neq j$. Hence, on $\{x_{s*}(T)\in\overline{\mathcal{W}}^j\}$, we have $\lim_{t\to T}\frac{\partial V_s}{\partial x}(t,x_{s*}(t))=\lim_{t\to T}\frac{\partial V_s^j}{\partial x}(t,x_{s*}(t))=M_s(x_{s*}(T)-p_j)$. But, x_{s*} is the solution of an SDE with non-degenerate noise. Therefore, $\mathbb{P}\left(x_{s*}(T)\in\partial\mathcal{W}^j\right)=0$. This shows the result. □

In the following Lemma, we give an equivalent representation of a solution $\bar{x}$ of (3.54).

Lemma 5 (Equivalent representation)**.** *Under Assumption 7, $\bar{x}$ satisfies the mean field equations* (3.54) *if and only if it satisfies the following equations*

$$\frac{d}{dt}\bar{X}(t)=A\bar{X}(t)-BR^{-1}B'\bar{q}(t), \tag{3.60}$$

$$\frac{d}{dt}\bar{q}(t)=-A'\bar{q}(t)+QL\bar{X}(t), \tag{3.61}$$

$$\bar{x}(t)=P_1\bar{X}(t), \tag{3.62}$$

with $\bar{X}(0)=\mathbb{E}X(0)$ and $\bar{q}(T)=M\left(\bar{X}(T)-\Lambda(\bar{x})\otimes I_n p\right)$, where the CDM $\Lambda(\bar{x})$ is defined as follows:

$$\Lambda_{sj}(\bar{x})=\mathbb{P}(x_{s*}(T)\in\mathcal{W}^j),\qquad 1\le s\le k,\ \ 1\le j\le l \tag{3.63}$$

$$dX_*(t)=\left(AX_*(t)+BU_*\left(t,X_*(t),\bar{x}\right)\right)dt+\sigma dW(t). \tag{3.64}$$

$\Lambda_{sj}(\bar{x})$ is the probability that a generic agent of type s optimally responding to $\bar{x}$ is at time T closer to the alternative p_j than the other alternatives, and x_{s*} is its optimal state.

Proof. By taking the expectations on the right and the left hand sides of (3.56) and the SDE in (3.54), and in view of $\sum_{s=1}^{k}\alpha_s\bar{x}_s(t)=\bar{x}$, we get the necessary condition. To prove the sufficient condition, we consider $(\bar{X},\bar{x},\bar{q})$ satisfying (3.60)-(3.64). We define $(\hat{x}_s,\hat{q}_s)=(\mathbb{E}x_{s*},\mathbb{E}q_{s*})$, where (x_{s*},q_{s*}) are the s-type generic agent's optimal state and co-state when tracking $\bar{x}$. We define $e=(\hat{x}_1,\dots,\hat{x}_k)-\bar{X}$ and $\bar{q}_e=(\hat{q}_1,\dots,\hat{q}_k)-\bar{q}$. By taking expectations on the right and the left hand sides of (3.56) and the generic agent's dynamics, we obtain that

$$\begin{aligned} \frac{d}{dt}e(t) &= Ae(t) - BR^{-1}B'\bar{q}_e(t), \qquad e(0) = 0 \\ \frac{d}{dt}\bar{q}_e(t) &= -A'\bar{q}_e(t) + QLe(t), \qquad \bar{q}_e(T) = Me(T). \end{aligned} \tag{3.65}$$

Under Assumption 7, we define $q_e'(t) = \gamma(t)e(t)$, where $\gamma(t)$ is the unique solution of (3.55). We have $\frac{d(\bar{q}_e - q_e')}{dt} = -(A' - \gamma(t)BR^{-1}B')(\bar{q}_e - q_e')$, with $(\bar{q}_e(T) - q_e'(T)) = 0$. Hence, $\bar{q}_e(t) = \gamma(t)e(t)$. By replacing $\bar{q}_e(t) = \gamma(t)e(t)$ in the forward equation in (3.65), we obtain that $e = 0$. This proves the result. □

The following functions are used to compute the solution of (3.60)-(3.61), where the CDM is considered as a parameter Λ. Under Assumption 7, we define R_1 and R_2 such that, for all $s \geq 0$,

$$\begin{aligned} \frac{d}{dt}R_1(t,s) &= \left(A - BR^{-1}B'\gamma(t)\right)R_1(t,s), \\ \frac{d}{dt}R_2(t) &= \left(A - BR^{-1}B'\gamma(t)\right)R_2(t) + BR^{-1}B'R_1(T,t)'M, \end{aligned} \tag{3.66}$$

with $R_1(s,s) = I_{nk}$ and $R_2(0) = 0$, where γ is the unique solution of (3.55). We denote by S the set of $k \times l$ row stochastic matrices. For $\Lambda \in S$, define the function $\bar{x}^\Lambda : [0,T] \to \mathbb{R}^n$ by

$$\bar{x}^\Lambda(t) := P_1(R_1(t,0)\bar{X}(0) + R_2(t)\Lambda \otimes I_n\, p). \tag{3.67}$$

In the following theorem, we show that the sustainable mean trajectories belong to the family of trajectories $\bar{x}^\Lambda$, where $\Lambda \in S$. This family is generated by (3.60)-(3.62), where the CDM in the boundary condition is replaced by any $\Lambda \in S$. For $\bar{x}^\Lambda$ to be a sustainable mean trajectory, Λ should be the CDM, when the generic agents of the different types optimally respond to $\bar{x}^\Lambda$, i.e. $\Lambda(\bar{x}^\Lambda)$. This is equivalent to say that Λ is a fixed point of F defined below. Indeed, F maps $\lambda \in S$ to the CDM when the generic agents optimally respond to $\bar{x}^\Lambda$.

Thus, we define the finite dimensional map F from S into itself, such that for all $\Lambda \in S$,

$$F(\Lambda)_{sj} = \mathbb{P}(x_{s*}^\Lambda(T) \in \mathcal{W}^j), \tag{3.68}$$

where $X_*^\Lambda = (x_{1*}^\Lambda, \ldots, x_{k*}^\Lambda)$ is the unique strong solution of the following SDE parameterized by Λ

$$dX_*^\Lambda(t) = \left(AX_*^\Lambda(t) + BU_*\left(t, X_*^\Lambda(t), \bar{x}^\Lambda\right)\right)dt + \sigma dW(t), \quad \text{with } X_*^\Lambda(0) = X_*(0). \tag{3.69}$$

We state now the main result of this section.

Theorem 11 (Sustainable mean trajectories)**.** *Under Assumption 7, the following statements hold:*

(i) $\bar{x}$ satisfies the mean field equations (3.54) *if and only if*

$$\bar{x} = \bar{x}^{\Lambda} \tag{3.70}$$

where $\bar{x}^{\Lambda}$ is defined in (3.67) *and Λ is a fixed point of F.*

(ii) F is continuous and has at least one fixed point. Equivalently, (3.54) *has at least one solution $\bar{x}$.*

(iii) For a uniform population, i.e, k=1, the sustainable paths $\bar{x}$ are uniformly bounded with respect to M and $\Lambda \in S$ for the standard L_2 norm $\left(\int_0^T \|.\|^2 \mathrm{dt} \right)^{\frac{1}{2}}$.

Proof.

(i) Let $\bar{x}$ be a path satisfying the mean field equations (3.54). Then, by Lemma 5, $\bar{x}$ satisfies the equivalent representation (3.60)-(3.64). Under Assumption 7, using arguments similar to those used in Lemma 5, we obtain that (3.60) and (3.61) has a unique solution $(\bar{X}, \bar{q})$. Moreover, $\bar{q} = \gamma \bar{X} + \beta$, where γ is the unique solution of (3.55), and β is the unique solution of $\dot{\beta} = -(A - BR^{-1}B'\gamma)'\beta$ with $\beta(T) = -M\Lambda \otimes I_n p$. By replacing, $\bar{q} = \gamma \bar{X} + \beta$ in (3.60), we get that $\bar{x}$ is of the form (3.67). Next, by implementing this new form of $\bar{x}$ in the expression of (3.64) and by noting that Λ satisfies (3.63), Λ is a fixed point of F. Conversely, we consider Λ to be a fixed point of F, $\bar{X} = \Big(R_1(t,0)\bar{X}(0) + R_2(t)\Lambda \otimes I_n p \Big)$ and $\bar{x} = P_1 \bar{X}$. We define $\bar{q}(t) = -(BR^{-1}B')^{-1}(\frac{d}{dt}\bar{X}(t) - A\bar{X}(t))$. $(\bar{X}, \bar{q})$ satisfies (3.60)-(3.61). We have $\Lambda_{sj} = F(\Lambda)_{sj} = \mathbb{P}(x^{\Lambda}_{s*}(T) \in \mathcal{W}^j)$, where x^{Λ}_{s*} is defined in (3.69). But $\bar{x}$ is of the form (3.70), hence X^{Λ}_{*} is the unique strong solution of (3.64). Therefore, $\bar{x}$ satisfies (3.60)-(3.64), and by Lemma 5, it satisfies (3.54). This proves the first point.

(ii) To show the existence of a fixed point of F, it is sufficient to show that F is continuous, in which case Brouwer's fixed point theorem (Conway, 2013, Section V.9) ensures the existence of a fixed point. Equation (3.69) is a SDE depending on the parameter Λ. By (Skorokhod, 1981, Theorem 1), the joint distribution of X^{Λ}_{*} and the Brownian motion W is weakly continuous in Λ. Consider a sequence of stochastic matrices $\{\Lambda_n\}_{n \geq 0}$ converging to the stochastic matrix Λ. The distribution of $X^{\Lambda_n}_{*}(T)$ converges weakly to the distribution of $X_*^\lambda(T)$. Moreover, X_*^λ is the solution of a non-degenerate SDE. Hence, $\mathcal{W}^j$, $j = 1, \dots, l$, is a continuity set of the distribution of X_*^λ. Therefore, $\lim_n F(\Lambda_n)_{sj} = \lim_n \mathbb{P}(x_{s*}^{\Lambda_n}(T) \in \mathcal{W}^j) = \mathbb{P}(x_{s*}^{\Lambda}(T) \in \mathcal{W}^j) = F(\Lambda)_{sj}$, and so F is continuous. This proves the second point.

(iii) Finally, using arguments similar to those used in the proof of the third point of Theorem 4, one can show the third point.

□

Remark 8. *The third point of Theorem 11 is used to prove Theorem 10 in Section 3.2 for a uniform population.*

Hitherto, we assume a continuum of agents and find a set of Nash strategies (3.42). Using similar arguments to those used in Theorem 6, one can show the following result.

Theorem 12 (ϵ−Nash equilibrium)**.** *When applied by a finite number N of agents, the strategies defined by* (3.42) *for a sustainable mean trajectory $\bar{x}$ constitute an ϵ_N-Nash equilibrium in the set $\mathcal{U}$* (3.3) *with respect to the costs* (3.2)*, where ϵ_N goes to zero as N increases to infinity.*

3.4 Discussions

As in the degenerate case, the finite dimensional map F defined in (3.68) characterizes the non-degenerate game in terms of the number of approximate Nash equilibria and the potential distributions of the agents' choices over the alternatives. However, the presence of the noise process changes the way the agents make their choices and the numerical scheme to compute the Nash strategies.

3.4.1 Indecisive agents

While in the degenerate case an agent makes its choices of an alternative prior starting to move, Theorem 9 shows that in the non-degenerate case, an agent can no longer commit to a choice from the beginning. Indeed, the optimal strategy (3.42) is a convex combination of the optimal strategies to go to each alternatives u_*^j, $1 \le j \le l$. The weights constitute a spatio-temporal Gibbs distribution (Liggett, 2012) that puts more mass on the less costly and less "risky" alternative. An alternative p_j is considered riskier in state x at time t, if the Brownian motion has a higher chance of driving the state of an agent closer to another alternative at time T, when this agent implements u_*^j from (x,t) onwards. The Gibbs distribution $\big(Gibbs^1(t,x),\ldots,Gibbs^l(t,x)\big)$ at $(t,x)\in[0,T]\times\mathbb{R}^n$ is defined as follows,

$$Gibbs^j(t,x)=\frac{\exp\left(-\eta V^j(t,x)+\log g^j(t,x)\right)}{\sum_{k=1}^{l}\exp\left(-\eta V^k(t,x)+\log g^k(t,x)\right)},\qquad j=1,\ldots,l. \tag{3.71}$$

In this distribution, the cost of an alternative p_j is captured by V^j, while its risk by $-\log g^j$, where V^j and g^j are defined in (3.19) and (3.23).

To compute its optimal strategy, an agent needs to know its state and the probability distribution of the initial conditions P_0 and parameters P_θ. Indeed, given P_0 and P_θ, it computes a fixed point Λ of F defined in (3.68), and anticipates the mean trajectory $\bar{x}=\bar{x}^\Lambda$ defined in (3.67). Afterwards, the agent implements the feedback control law (3.42) by measuring instantaneously its state. This choice process is summarized in Figure 3.1.

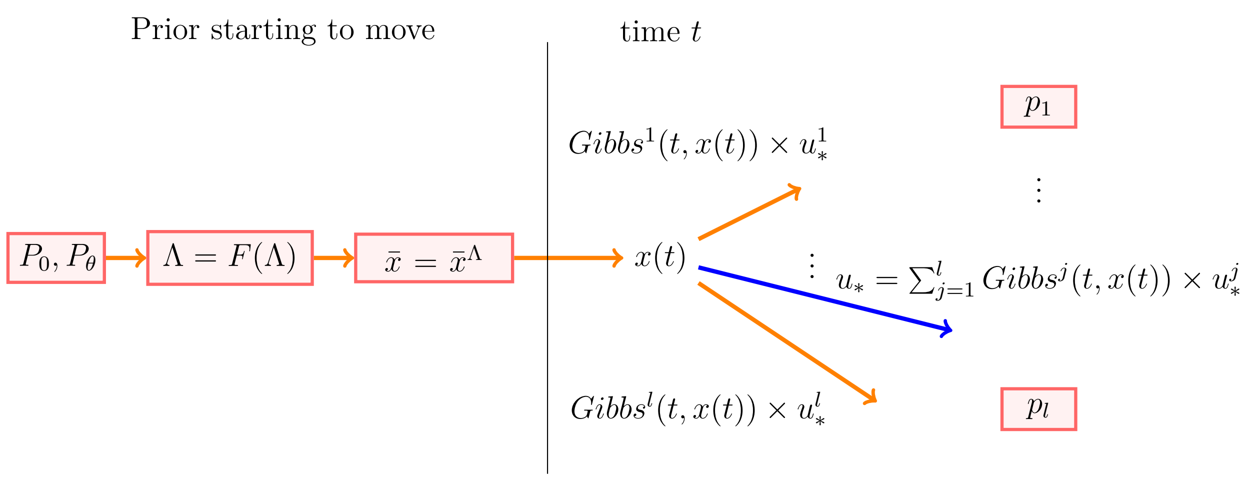

Figure 3.1 Non-degenerate non-cooperative DCDC: the choice process

3.4.2 Numerical scheme

As discussed above, to compute the optimal strategies (3.42), it is sufficient to find a fixed point Λ of F. But F is a map from the set S of $k\times l$ row stochastic matrices into itself. Hence, a fixed point can be computed by applying Broyden's method to $F(\Lambda)-\Lambda$. In the binary choice case with uniform population, S is the set of $(r,1-r)$, where $r\in[0,1]$, and $F(r,1-r)=([F(r,1-r)]_1,1-[F(r,1-r)]_1)$. Hence, one can apply the bisection method to the scalar function $G(r):=[F(r,1-r)]_1-r$ to find a fixed point of F. Both methods assume that one knows the value of F at any $\Lambda\in S$. But $F(\Lambda)$ involves the probability distribution of the random variable $X_*^\Lambda(T)$. This can be computed by solving the FP equation related to the diffusion process (3.69) via the finite difference implicit method (Pichler et al., 2013).

In the following, we develop a numerical scheme for the binary choice uniform population case. The function $G(r)$ defined in the previous paragraph is equal to $\int_{-\infty}^{c} p_r(T,x)dx$, where $c = (p_1 + p_2)/2$ and $p_r(t,x)$ is the probability density function of $X_*^{(r,1-r)}(t)$ defined by (3.69). It satisfies the following FP equation,

$$\frac{\partial p_r(t,x)}{\partial t} = -\frac{\partial \left(\mu(t,x,r)p_r(t,x)\right)}{\partial x} + \frac{\sigma^2}{2}\frac{\partial^2 p_r(t,x)}{\partial x^2}, \tag{3.72}$$

with $p_r(0,x) = p_0(x), \forall x \in \mathbb{R}$. Here $\mu(t,x,r) = Ax + Bu_*\left(t,x,\bar{x}^{(r,1-r)}\right)$ for $\bar{x}^{(r,1-r)}(t) = R_1(t,0)\bar{x}(0) + R_2(t)\left(rp_1 + (1-r)p_2\right)$, see (3.67).

We solve (3.72) via an implicit finite difference scheme (Pichler et al., 2013), with space domain $[x_{\min}, x_{\max}]$, time step Δt and space step Δx. We choose $x_{\min}$ and $x_{\max}$ such that the probability of $X_*^{(r,1-r)}$ to be outside $[x_{\min}, x_{\max}]$ is negligible. Since the final cost encourages the agents to move towards p_1 or p_2, one can heuristically set, in case of an initial Gaussian distribution $\mathcal{N}(\mu_0, \sigma_0^2)$, $x_{\min} = p_1 - |\mu_0| - 3\sqrt{T}\sigma - 3\sigma_0$ and $x_{\max} = p_2 + |\mu_0| + 3\sqrt{T}\sigma + 3\sigma_0$ (here we assume $p_1 < p_2$). These expressions reflect the fact that the agents initially centered around μ_0 with standard deviation σ_0 are spread in space by the Brownian motion, which has a maximum standard deviation $\sqrt{T}\sigma$, and forced by the optimal control laws to be centered at the end of the horizon around p_1 and p_2.

We denote by N_t the number of time nodes, by N_x the number of space nodes and by $p_r(k,i)$ and $\mu_r(k,i)$, for $0 \leq k \leq N_t$ and $0 \leq i \leq N_x$, the values of p_r and μ_r at $(t,x) = (k\Delta t, i(x_{\min} + \Delta x))$. The discretized FP equation (3.72) is then:

$$\begin{aligned}\frac{p_r(k+1,i) - p_r(k,i)}{\Delta t} = &-\frac{\mu_r(k+1,i+1)p_r(k+1,i+1) - \mu_r(k+1,i-1)p_r(k+1,i-1)}{2\Delta x} \\ &+ \frac{\sigma^2}{2}\frac{p_r(k+1,i+1) - 2p_r(k+1,i) + p_r(k+1,i-1)}{(\Delta x)^2},\end{aligned} \tag{3.73}$$

or in matrix form:

$$S_r^{k+1}P_r^{k+1} = P_r^k, \tag{3.74}$$

where $P_r^k = [p_r(k,0), ..., p_r(k,N_x)]'$ and $S_r^{k+1} = (s_{ij})_{i,j\in\{1,...,N_x\}}$ a tridiagonal matrix, with $s_{i(i-1)} = -\frac{\sigma^2}{2}\frac{\Delta t}{(\Delta x)^2} - \frac{\Delta t}{2\Delta x}\mu_r(k+1,i-1)$, $s_{i(i+1)} = -\frac{\sigma^2}{2}\frac{\Delta t}{(\Delta x)^2} + \frac{\Delta t}{2\Delta x}\mu_r(k+1,i+1)$ and $s_{ii} = 1 + \sigma^2\frac{\Delta t}{(\Delta x)^2}$. We summarize the numerical scheme in Algorithm 1.

Algorithm 1 Fixed Point Algorithm

procedure BISECTION METHOD
  $r_0 = 0$, $r_1 = 1$ and $err = 1$
  while $err > 0.01$ **do**
    $r = (r_0 + r_1)/2$
    procedure COMPUTATION OF $G(r)$
      for $k = 0$ to $N_t - 1$ **do**
        Find P_r^{k+1} s.t. $S_r^{k+1} P_r^{k+1} = P_r^k$
      end for
      $G(r) = \Delta x \sum_{i \le k_c} p_r(N_t, i)$, with $K_c = (c - x_{\min})/\Delta x$
      $err = |r_0 - r_1|$
      if $(G(r) - r)(G(r_0) - r_0) < 0$ **then**
        $r_1 = r$
      else
        $r_0 = r$
      end if
    end procedure
  end while
end procedure
Output: r, a fixed point of G

3.4.3 Relation to the discrete choice problems in microeconomics

We discuss in Section 1.2.2 the relation of the DCDC problems to the static discrete choice models in microeconomics. In this section, we show explicitly that our results are directly related to those in the microeconomic theory. Let us at first recall some facts about the static models (McFadden, 1974), whereby a generic person chooses between l alternatives. The cost he/she pays when choosing an alternative j is defined by $v^j = k(j) + \nu$, where $k(j)$ is a deterministic function that depends on personal publicly observable attributes, such as the person's financial status in the example of residential location choice (Bhat and Guo, 2004), and on alternative j. ν is a random variable accounting for personal idiosyncrasies unobservable by the social planner. The probability distribution of ν is specified so that the selection probabilities, i.e., the probabilities to choose each alternative given the observable attributes, satisfy three axioms given in (McFadden, 1974). McFadden showed that when ν has an extreme value distribution (McFadden, 1974), i.e. the cumulative distribution function of ν is equal to $F_\nu(x) = e^{-e^{-x}}$, then these axioms are satisfied and the probability that a cost-minimizing generic person chooses an alternative j is equal to $Pr^j = \frac{\exp(-k(j))}{\sum_{s=1}^{l}\exp(-k(s))}$. Now, the min-LQG optimal strategy at time t is $u_*(t,x,\bar{x}) = \sum_{j=1}^{l} Gibbs^j(t,x)u_*^j(t,x,\bar{x})$, where $Gibbs^j$ is defined in (3.71), V^j is the optimal cost if only the alternative p_j is available and u_*^j is the corresponding optimal strategy. The min-LQG optimal strategy can be interpreted as a mixed strategy between the pure strategies $u_*^j(t,x)$ (picking alternative p_j), $j = 1,\dots,l$. Within this framework, a generic agent at time t chooses the alternative p_j with probability $Gibbs^j(t,x)$. Thus, the min-LQG problem can be viewed at each time $t \in [0,T]$ as a static discrete choice problem, where the cost of choosing alternative p_j includes an additional term $-\frac{1}{\eta}\log\left(g^j(t,x)\right)$. This term increases with the probability that a generic agent making an early choice in favor of one alternative is driven by the Brownian motion toward one of the other alternatives. In other words, it measures the expected cost of making a premature choice of an alternative at time $t < T$, while ignoring the possibility of the future process noise upsetting the wisdom of such decision.

3.5 Simulation results

We illustrate in this section the non-degenerate non-cooperative DCDC model through some numerical examples.

3.5.1 Evolution of the probability distribution and sample paths

We consider a group of 50 agents with uniform dynamics, where the dynamics and costs parameters are $A = 0.1$, $B = 0.2$, $\sigma = 1.5$, $Q = 10$, $R = 5$, $M = 500$, $T = 2$ and $p_1 = -p_2 = -10$. The agents are initially drawn from the Gaussian distribution $\mathcal{N}(0.3, 1)$. Following the numerical scheme in Section 3.4, we find $\Lambda = (0.2, 0.8)$ a fixed point of F defined in (3.68). Figure 3.2 below shows the evolution of the agents' probability distribution and the sample paths of 10 agents. The distribution at time T is concentrated around the alternatives p_1 and p_2.

3.5.2 Influence of the social effect on the group's behavior

In this section, we analyze the influence of the social effect on the choices' distribution. We consider the example of Section 3.5.1 and compute the fixed points of F (or G defined in Section 3.4) for different values of Q. The results are shown in Figure 3.3. According to this figure, F has only one fixed point $(0.4, 0.6)$ without a social effect ($Q = 0$). Accordingly, 40% of the agents go towards p_1, and the rest towards p_2. As the strength of the social effect

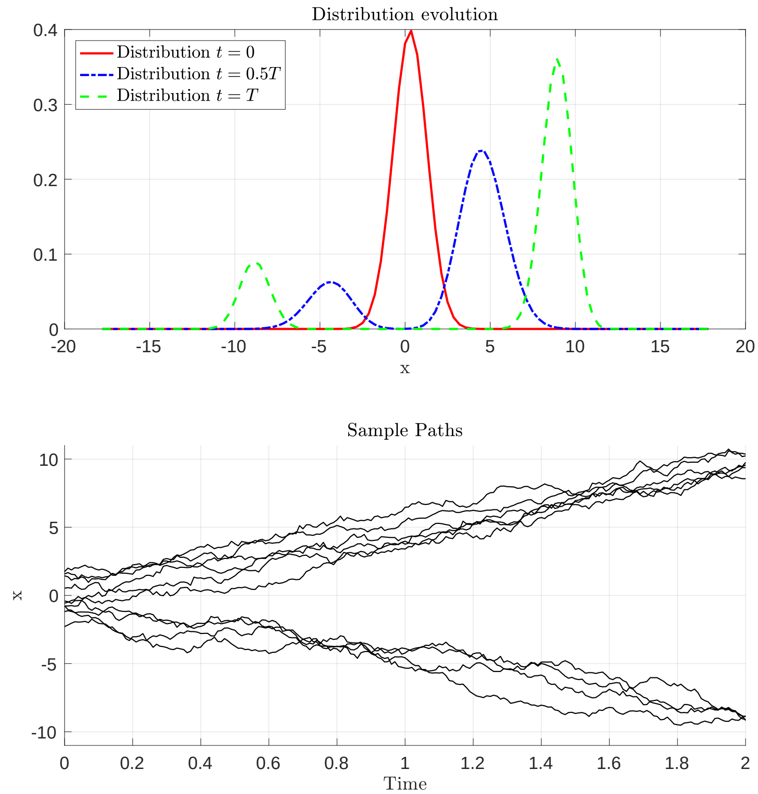

Figure 3.2 Non-degenerate non-cooperative DCDC: distribution's evolution and sample paths.

increases, the majority attracts more agents. For $Q = 20$, almost all the agents go towards p_2. When the social effect Q exceeds 21, F has three fixed points, where two of them correspond to consensus on one alternative. Indeed as Q increases arbitrarily, the agents essentially forget temporarily about the final cost, and the problem becomes a classical rendez-vous MFG where they tend to merge towards each other rapidly. If this occurs around the midpoint of the alternatives, then they stay grouped around this point, and before the end of the time horizon they split between the alternatives to minimize their final costs. Some large deviations are also possible, whereby a significant fraction of agents decides to choose one alternative, thus pulling everyone else towards it, which may help explain the non uniqueness of outcomes for $Q > 21$. Figure 3.4 illustrates the first behavior for a very strong social effect ($Q = 500$), i.e. when the population splits between the alternatives. Moreover, it compares it to the case where the social effect is absent. In both cases, the population splits almost equally between the alternatives. While the agents stay together as much as possible under a strong social effect, in the absence of a social effect, they split from the beginning into two groups, and each group moves to the less costly alternative. Although the frameworks are different, the results illustrated in Figure 3.3 resembles to the pitchfork bifurcation diagrams studied in (Gray et al., 2018) to model the influence of the social effect on the behavior of a population of honeybees choosing between two nectar sites. Indeed, the results in (Gray et al., 2018) show that for weak social effect, the dynamical system describing the evolution of the bees' opinions toward two different sites has a unique stable equilibrium that splits the population between the sites. For a strong social effect, however, the system has three equilibria; two stable ones that correspond to consensus to go to each site and one unstable equilibrium that splits the population between the sites.

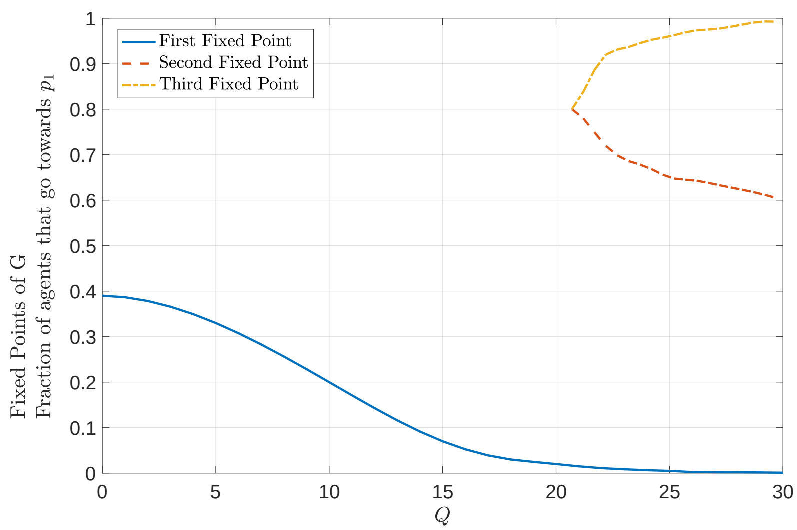

Figure 3.3 Non-degenerate non-cooperative DCDC: influence of the social effect on the choices' distribution.

3.5.3 Confidence zone and influence of the noise intensity on the agents' individual behaviors

Finally, we illustrate the influence of the noise intensity on the individual behaviors. Let us at first define what we call a "confidence zone". An agent is said to be in the confidence zone, if it chooses one of the alternatives with high probability. Formally, it is defined as follows,

$$CZ(e) = \{(t,x) \in [0,T] \times \mathbb{R}| \;\; Gibbs^1(t,x) > 1-e \text{ or } Gibbs^2(t,x) > 1-e\}, \tag{3.75}$$

where e is the confidence margin. Figure 3.5 shows the confidence zones (hatched areas) for $e = 0.1$ and for different values of σ. As expected, the area of $CZ(0.1)$ decreases as the

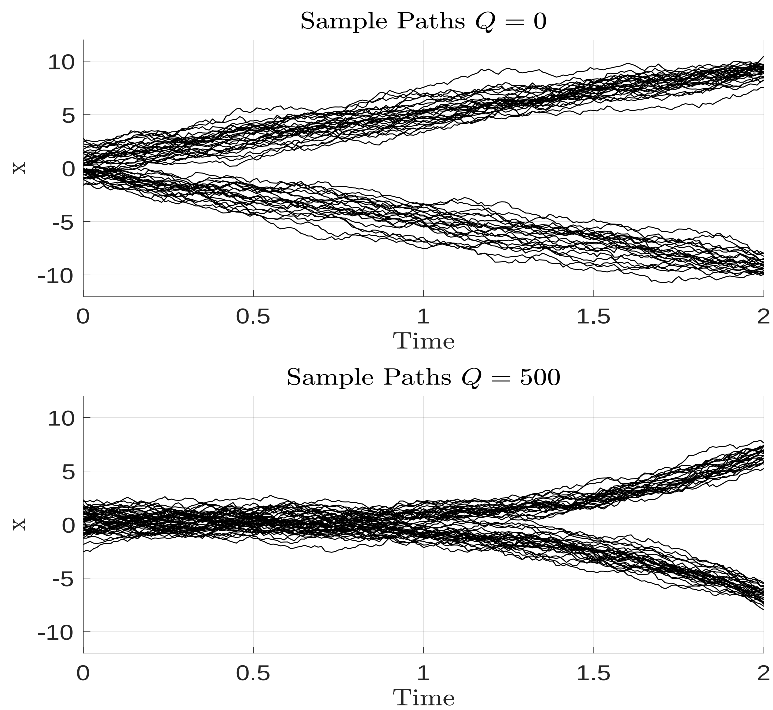

Figure 3.4 Non-degenerate non-cooperative DCDC: different splitting behaviors.

noise intensity increases. Indeed, the probability that an agent, making an early choice in favor of one alternative, is driven to the other one by the Brownian motion increases with the noise intensity, thus making the agent “indecisive”. Now, for a fixed σ, the confidence zone expands with time. Hence, the agents become more confident. This is due to the fact that, as time passes, the noise process has less time to drive the agent to other than the chosen alternative. Finally, we report in Figure 3.6 the Gibbs distribution for an agent at different instants of time. At time $t = 1$, the agent is closer to p_2 than at time $t = 1.75$. But, $Gibbs^2(1, x(1)) < Gibbs^2(1.75, x(1.75))$. This confirms that the agents become more confident with time.

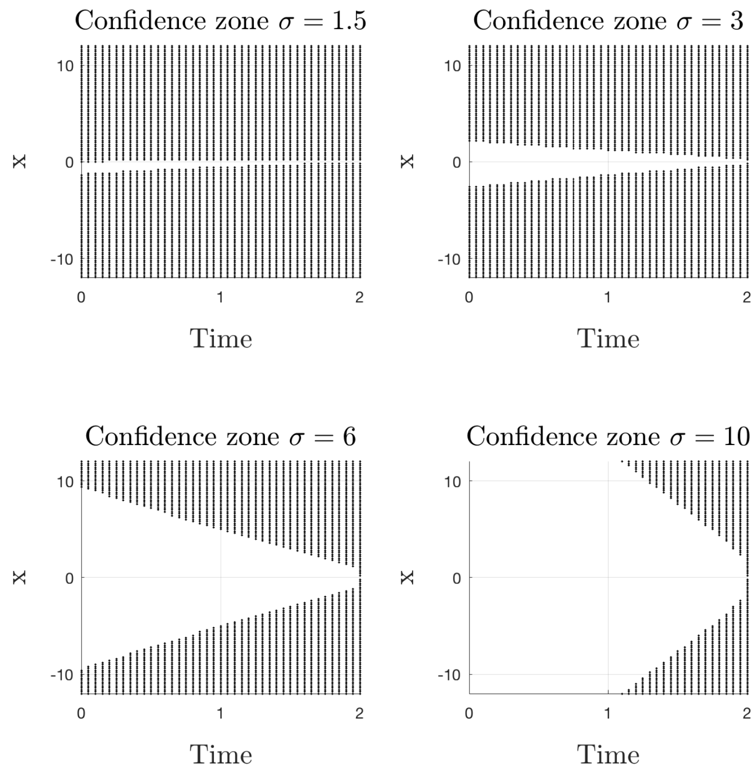

Figure 3.5 Non-degenerate non-cooperative DCDC: confidence zones (hatched areas).

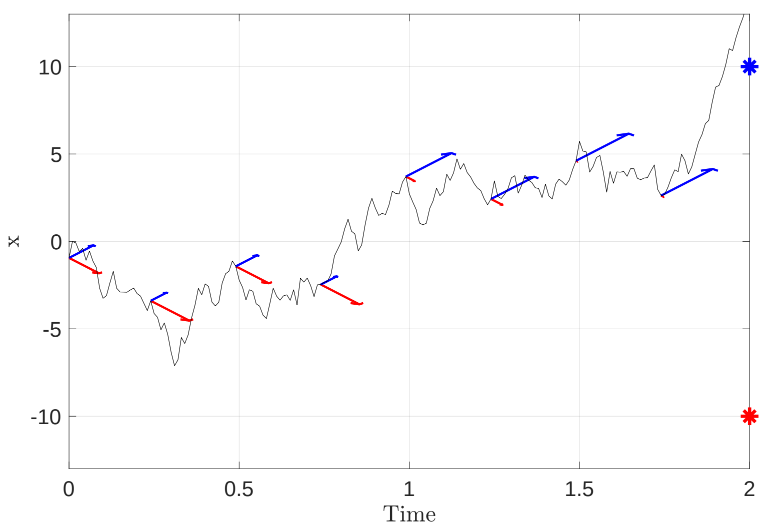

Figure 3.6 Non-degenerate non-cooperative DCDC: evolution of the Gibbs distribution over a sample path for $\sigma = 6$. The red (resp. blue) downward (resp. upward) arrow is the element of the Gibbs distribution related to $p_1 = -10$ (resp. $p_2 = 10$).

3.6 Conclusion

We formulate the non-degenerate non-cooperative DCDC problem as a stochastic dynamic game. The dynamics are controlled SDE's, and the agents have min-LQG individual costs. We solve explicitly the novel non-degenerate min-LQG optimal control problem, and compute a set of decentralized feedback strategies that qualify as approximate Nash as the size of population increases sufficiently. Following the strategies' form, we interpret the min-LQG optimal control problem at each instant as a static discrete choice problem. This includes an additional cost to penalize premature myopic decisions. The non-degenerate min-LQG optimal control law exhibits some other interesting properties that we summarize and compare to those of the degenerate case in Table 3.1.

Table 3.1 Degenerate and non-degenerate min-LQG optimal control problems

Degenerate min-LQG	Non-degenerate min-LQG
The optimal control law is equal to the less costly "pure" strategy u_*^j.	The optimal control law is a convex combination of the "pure" strategies, weighted by a Gibbs distribution that puts more mass on the less risky and less costly alternative.
The agents commit to their initial choices.	The agents reassess continuously the adequacy of their would-be choices along the path. As time passes, the agents become more confident about their choices.
The probability that an agent applying the optimal control law is far from the alternatives decays to zero faster than $1/M$.	The probability that an agent applying the optimal control law is far from the alternatives decays to zero faster than $\frac{1}{M} + \frac{\sigma^2 \log M}{M}$.

As in the degenerate case, we construct a one- to-one map between the $\epsilon-$Nash equilibria and the fixed points of a finite dimensional map. The latter characterize the game in terms of number of equilibria and the way the population splits between the alternatives. Furthermore, this one-to-one map allows us to propose a simple numerical scheme to solve the mean field equations.

The solution of the non-degenerate min-LQG optimal control problem is derived under Assumption 6, which is invoked at many places in the proofs. It is used in Theorem 8 to define the Hopf-Cole transformation and linearize the HJB equation (3.6), in Lemma 3 to show the continuity of the cost-to-go and in Theorem 9 to derive the min-LQG optimal control law. As discussed in Remark 7, this assumption imposes a kind of isotropy on the ratio "control efficiency / noise intensity". Following this interpretation of Assumption 6, a future direction to remove it would be to apply a continuous transformation (Lie transformation) on the set of solutions $((t, x), V(t, x))$, such that the transformed solution satisfies a LQG type (at least in the running term) HJB equation where the isotropy condition is satisfied.

Hitherto, the agents make their choices of an alternative in a non-cooperative fashion. The next chapter considers the cooperative DCDC problem.

CHAPTER 4 DEGENERATE COOPERATIVE DYNAMIC COLLECTIVE DISCRETE CHOICE MODEL

We develop in Chapters 2 and 3 the non-cooperative DCDC models, where an agent minimizes its cost irrespective of making the others better-off or worse-off. In some situations, this selfish behavior neglects the social context that imposes a sort of cooperation between the agents. Baker gave in (Baker, 1984) an example of such situations at the Chicago Options Exchange, where the relations among the traders, supposed to be non-cooperative, affect their trades. The aim of this chapter is to develop a cooperative DCDC model, which is formulated in Section 4.1. In Section 4.2, we solve for an exact social optimum (see Definition 3, Section 1.2.1). We propose a naïve method to compute an exact solution and show that it becomes quickly computationally intractable as the size of the population increases sufficiently. Moreover, its implementation requires a significant amount of communication between the agents. Instead, we develop in Section 4.3 via the MFG methodology a set of decentralized strategies that are simple to compute and implement. These strategies are shown to converge to a social optimum as the size of the population increases to infinity. We discuss the results in Section 4.4. Section 4.5 illustrates via a numerical example the advantage of the cooperative strategies with respect to the non-cooperative one in evenly allocating the agents to the alternatives. Finally, Section 4.6 concludes this chapter.

4.1 Mathematical Model

We formulate the degenerate cooperative DCDC problem as a social optimization problem involving N agents with dynamics,

$$\frac{d}{dt}x_i = A_i x_i + B_i u_i \qquad\qquad 1 \leq i \leq N, \tag{4.1}$$

where $A_i \in \mathbb{R}^{n\times n}$, $B_i \in \mathbb{R}^{n\times m}$, $x_i \in \mathbb{R}^n$ is the state of agent i, and $u_i \in \mathcal{U} = L_2([0,T],\mathbb{R}^m)$ its control input. The initial conditions $x_i(0)$, $1 \leq i \leq N$, are assumed deterministic. The agents cooperate to minimize the following common social cost,

$$J_{soc}\left(u_i, u_{-i}, \bar{x}, x_i(0), x_{-i}(0)\right) = \sum_{i=1}^{N} J_i\left(u_i, \bar{x}, x_i(0)\right), \tag{4.2}$$

where $\bar{x} = 1/N \sum_{i=1}^{N} x_i$. The individual costs are defined as follows,

$$J_i\left(u_i, \bar{x}, x_i(0)\right) = \int_0^T \left\{\|x_i - Z\bar{x}\|_Q^2 + \|u_i\|_{R_i}^2\right\} \mathrm{dt} + \min_{1 \le j \le l} \|x_i(T) - p_j\|_{M_i^j}^2, \tag{4.3}$$

where $Q \succeq 0$, $R_i \succ 0$, $M_i^j \succ 0$, $Z_i \in \mathbb{R}^{n \times n}$, and $p_j \in \mathbb{R}^n$, $1 \le j \le l$, are the alternatives. The coefficient R_i depends on the agent i. In the robotic swarm example, this reflects for instance the intention of the social planner to limit the mobility of some robots. For a detailed discussion about the individual costs, we refer the reader to Section 2.1.

When considering the limiting population ($N \to \infty$), it is convenient to represent the limiting sequence of $\{(x_i(0), \theta_i)\}_{i=1,\dots,N} := \{(x_i(0), A_i, B_i, R_i, M_i^1, \dots, M_i^l)\}_{i=1,\dots,N}$ by the random vector $(x(0), \theta)$ on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We assume that θ is in a compact set Θ. Let us denote the empirical measure of the sequence $\{(x_i(0), \theta_i)\}_i$ by $P_{0\theta}^N(\mathcal{A}) = \frac{1}{N} \sum_{i=1}^{N} 1_{\mathcal{A}}(x_i(0), \theta_i)$ for all (Borel) measurable sets $\mathcal{A}$. We assume that $P_{0\theta}^N$ has a weak limit the distribution of $(x(0), \theta)$, $P_{0\theta}$. We assume that $x(0)$ is independent of θ, that is $P_{0\theta} = P_0 \times P_\theta$, where P_0 and P_θ are respectively the marginal distributions of $x(0)$ and θ.

4.2 Exact Social Optimum

In this section, we consider the problem of finding a social optimum with respect to the social cost (4.2) (see Definition 3, Section 1.2.1). We start by showing that this problem is a degenerate min-LQG optimal control problem. Hence, we apply the results of Section 2.2 to compute a social optimum. We define the population's state $X = (x_1, \dots, x_N)$ and control profile $U = (u_1, \dots, u_N)$. A social optimum in the set $\mathcal{U}^N$ is the optimal control law $U_* \in \mathcal{U}^N$ that minimizes the social cost subject to the following population's dynamics,

$$\frac{d}{dt} X = AX + BU, \tag{4.4}$$

where $A = \mathrm{diag}(A_1, \dots, A_N)$ and $B = \mathrm{diag}(B_1, \dots, B_N)$. The individual costs can be written as follows,

$$J_i(u_i, \bar{x}, x_i(0)) = \min_{p_j \in \Delta} J_i^{p_j}(u_i, \bar{x}, x_i(0)), \tag{4.5}$$

where $\Delta = \{p_1, \dots, p_l\}$ is the set of alternatives, and

$$J_i^{p_j}(u_i, \bar{x}, x_i(0)) = \int_0^T \left\{\|x_i - Z\bar{x}\|_Q^2 + \|u_i\|_{R_i}^2\right\} \mathrm{dt} + \|x_i(T) - p_j\|_{M_i^j}^2. \tag{4.6}$$

Using the equality $c + \min(a, b) = \min(a + c, b + c)$, one can show by induction that,

$$J_{soc}(U, \bar{x}, X(0)) = \min_{d=(d_1,\dots,d_N)\in\Delta^N} J^d(U, X(0)), \tag{4.7}$$

where $J^d(U, X(0)) = \sum_{i=1}^{N} J_i^{d_i}(u_i, \bar{x}, x_i(0))$. J^d can be written as follows,

$$J^d(U, X(0)) = \int_0^T \left\{ \|X\|_{\hat{Q}}^2 + \|U\|_R^2 \right\} \mathrm{dt} + \|X(T) - d\|_{M^d}^2, \tag{4.8}$$

with $\hat{Q} = I_N \otimes Q + \frac{1}{N} 1_{N\times N} \otimes L$, $R = \mathrm{diag}(R_1, \dots, R_N)$, and $M^d = \mathrm{diag}(M_1^{d_1}, \dots, M_N^{d_N})$, and

$$L = Z'QZ - QZ - Z'Q. \tag{4.9}$$

According to equation (4.7), the social cost is a degenerate min-LQG cost, i.e. minimum of the l^N LQR costs (4.8). Let us denote by U_*^d and $J_*^d(X(0))$ the optimal control law and optimal cost of the LQR cost J^d, for $d \in \Delta^N$. Following the results of the degenerate min-LQG optimal control problem in Section 2.2, we obtain the following theorem.

Theorem 13 (Social optimum)**.** *The social cost* (4.2) *has a social optimum* $U_*^{d^*}$, *with* $d^* \in \underset{d\in\Delta^N}{\mathrm{argmin}}\, J_*^d(X(0))$.

The set Δ^N is the set of potential deployment configurations. For example, $d = (p_1, p_2, \dots, p_2)$ means that agent 1 chooses p_1, and the rest of the agents choose p_2. A deployment configuration $d \in \Delta^N$ costs the population $J^d(X(0))$. Theorem 13 states that a social optimum is the optimal control law of the optimal deployment configuration.

The exact solution of the cooperative DCDC problem presents the following difficulties:

- **Computational intractability:** To find a social optimum (naïvely), one needs to compute the l^N LQR costs of the different deployment configurations, and pick the less costly one. Moreover, each LQR problem involves solving a Riccati equation of dimension $nN \times nN$ and a linear ODE of dimension $nN + 1$. Hence, finding a social optimum becomes quickly intractable as the number of agents N increases.

- **Significant amount of communication:** According to U_*^d, the implementation of the social optimum requires that each agent know at least the exact initial conditions of all the agents $X(0)$, as well as the exact parameters θ_i, $1 \le i \le N$. As a result, the amount of communication between the agents increases drastically as the size of the population increases.

In the next section, we develop a set of strategies that are simpler to compute and implement than the exact social optimum. These strategies converge to a social optimum as the size of the population increases to infinity.

Remark 9. *Similarly to Theorem 3 in Chapter 2, we can force the cooperative agents to reach arbitrary small neighborhoods of the alternatives by increasing the coefficients M_i^j.*

4.3 Decentralized approximate Social Optimum

We develop in this section via the MFG methodology a set of decentralized strategies that are simple to compute and implement. These strategies converge to a social optimum as the size of the population increases to infinity. Following the methodology proposed in (Huang et al., 2012), we start by looking for a person-by-person solution for a continuum of agents (See Definition 2, Section 1.2.1). We show that this solution is a Nash equilibrium of a non-cooperative degenerate DCDC game. Here again, we apply the results of Chapter 2 to prove the existence of an equilibrium for a continuum of agents. Afterwards, we show that these strategies, when applied to a finite number of agents, converge to a social optimum as $N \to \infty$.

4.3.1 Person-by-person optimality

The person-by-person optimality is a weaker solution concept than the social optimum. These two concepts coincide, however, under some technical conditions, for instance, the convexity and smoothness of the costs in static games (Yüksel and Başar, 2013, Lemma 2.6.1). For a detailed discussion about these conditions in the LQG and min-LQG MFG cases, we refer the reader to Section 4.4.

Following the definition of the person-by-person solution (See Definition 2, Section 1.2.1), an agent i assumes that the other agents fixed their person-by-person strategies u_{-i}^*, and computes its strategy u_i^* by minimizing the social cost $J_{soc}(u_i, u_{-i}^*)$ over $u_i \in \mathcal{U}$. To simplify the presentation, we omit in this section the other arguments of J_{soc}. Similarly to (Huang et al., 2012), one can show that the social cost can be written $J_{soc}(u_i, u_{-i}^*) = J_{1,i}\left(u_i, \bar{x}_{-i}^*\right) +$

$J_{2,i}(u^*_{-i})$, where $\bar{x}^*_{-i} = 1/N\sum_{j=1,j\neq i}^N x^*_j$,

$$J_{1,i}\left(u_i, \bar{x}^*_{-i}\right) = \int_0^T \left\{\|x_i\|^2_{Q_N} + \left(\bar{x}^*_{-i}\right)' L_N x_i + \|u_i\|^2_{R_i}\right\} \mathrm{dt} + \min_{1\leq j\leq l} \|x_i(T) - p_j\|^2_{M_i^j}, \tag{4.10}$$

$$\text{with } Q_N = \left(I_n - \frac{1}{N}Z\right)' Q \left(I_n - \frac{1}{N}Z\right) + \frac{(N-1)}{N^2} Z'QZ,$$

$$L_N = -Z'Q\left(I_n - \frac{1}{N}Z\right) - QZ + \frac{q(N-1)}{N} Z'QZ.$$

The term $J_{2,i}(u^*_{-i})$ does not depend on the strategy u_i of agent i. Therefore, minimizing $J_{soc}(u_i, u^*_{-i})$ reduces to minimizing $J_{1,i}\left(u_i, \bar{x}^*_{-i}\right)$. Hence, a person-by-person solution (u^*_i, u^*_{-i}) with respect to the social cost is a fixed point of the system of equations $u^*_i \in \operatorname*{argmin}_{u_i\in\mathcal{U}} J_{1,i}\left(u_i, \bar{x}^*_{-i}\right)$, $1 \leq i \leq N$. Equivalently, (u^*_i, u^*_{-i}) is a Nash equilibrium with respect to the costs $J_{1,i}$, $1 \leq i \leq N$. But, these costs are of the degenerate min-LQG type with a weak coupling between the agents through $\bar{x}^*_{-i}$. Accordingly, we apply the results of Chapter 2 to compute a person-by-person solution for a continuum of agents.

Under the continuum of agents assumption, the mean field coupling term $\bar{x}^*_{-i}$ is assumed known and equal to $\bar{x}$. Moreover, the coefficients Q_N and L_N in the cost $J_{1,i}$ (4.10) converge to Q and L as $N \to \infty$. Hence, a generic agent of state x and parameters $\theta = (A_\theta, B_\theta, R_\theta, M^1_\theta, \ldots, M^l_\theta)$ minimizes the following min-LQG cost,

$$J(u, \bar{x}, x(0), \theta) = \int_0^T \left\{\|x\|^2_Q + \bar{x}'Lx + \|u\|^2_{R_\theta}\right\} \mathrm{dt} + \min_{1\leq j\leq l} \|x(T) - p_j\|^2_{M^j_\theta}, \tag{4.11}$$

where L is defined in (4.9). We define Γ^j_θ, β^j_θ and δ^j_θ to be the unique solutions of the ODE's (2.20)-(2.22), where we replace (A, B, R, M) by $(A_\theta, B_\theta, R_\theta, M^j_\theta)$, Q in (2.21) by $-L$ and Q in (2.22) by 0. Following the analysis of Section 2.6, the generic agent's best response to $\bar{x}$, and the corresponding state are on the set $\{x(0) \in D^j_\theta(\bar{x})\}$ as follows,

$$u_*(t, x, \bar{x}, x(0), \theta) = -B'_\theta R^{-1}_\theta \left(\Gamma^j_\theta(t)x + \beta^j_\theta(t)\right), \tag{4.12}$$

$$\begin{aligned} x_*(t, \bar{x}, x(0), \theta) = {} & \Phi^j_\theta(0, t)'x(0) + \int_0^t \Psi^j_\theta(\sigma, t, \sigma, T) M^j_\theta p_j \,\mathrm{d}\sigma \\ & + \int_0^t \int_T^\sigma \Psi^j_\theta(\sigma, t, \sigma, \tau) L\bar{x}(\tau) \,\mathrm{d}\tau \mathrm{d}\sigma, \end{aligned} \tag{4.13}$$

where Φ^j_θ, Ψ^j_θ and the basin of attraction $D^j_\theta(\bar{x})$, are defined as in Section 2.6. Moreover, there exists a sustainable mean trajectory, i.e., a fixed point of the map G that maps $\bar{x} \in C([0, T], \mathbb{R}^n)$ to $G(\bar{x}) = \mathbb{E}x_*(t, \bar{x}, x(0), \theta)$, under the following assumptions.

Assumption 8. *We assume that* $1/N\sum_{i=1}^N \|x_i(0)\|^2 < C$*, for all* $N > 0$.

Remark 10. *Using Portmanteau Theorem (Kallenberg, 2006, Chapter 3), one can show that Assumption 8 implies that* $\mathbb{E}\|x(0)\|^2 < \infty$*.*

Assumption 9. *Under Assumption 8, we assume that* $\sqrt{\max(k_1+k_2,k_3)}T < \pi/2$*, with* k_1*,* k_2 *and* k_3 *are defined as* (2.63)*, where* Q_θ *in the expression of* k_3 *is replaced by* L*.*

Assumption 10. *We assume that the* P_0-*measure of quadric surfaces in* $\mathbb{R}^n$ *is zero.*

According to Theorem 6 in Section 2.7, the strategies (4.12) constitute an $\epsilon-$Nash equilibrium with respect to the costs $J_{1,i}$ defined in (4.10), where ϵ converges to 0 as $N \to \infty$. Equivalently, these strategies are approximately person-by-person optimal. We show in the following section that they also converge to a social optimum under some conditions.

4.3.2 Asymptotic social optimum

We show in this section that the infinite population person-by-person strategies (4.12) are asymptotically socially optimal. Moreover, we give an explicit formula of the asymptotic social cost. We start by the following preliminary result, which states that the average state of N agents applying the strategies (4.12), for some fixed point of G, converges to this fixed point as $N \to \infty$. Given a fixed point path $\bar{x}$ of G, we define $u_*^{(N)}$ to be the strategy profile of the agents $1, \dots, N$ when they apply (4.12), and $\bar{x}^{(N)}(t)$ their average state under $u_*^{(N)}$, i.e.

$$\bar{x}^{(N)}(t) := \frac{1}{N}\sum_{i=1}^{N} x_*(t,\bar{x},x_i(0),\theta_i) = \int_{\mathbb{R}^n\times\Theta} x_*(t,\bar{x},x_0,\theta)\mathrm{d}P_{0\theta}^N(x_0,\theta). \tag{4.14}$$

Lemma 6. *Under Assumptions 8, 9 and 10,*

$$\lim_{N\to\infty}\int_0^T \left\|\bar{x}^{(N)}(t)-\bar{x}(t)\right\|^2 \mathrm{d}t = 0. \tag{4.15}$$

Proof. The functions Γ_θ^j, β_θ^j and δ_θ^j are continuous with respect to θ, which belongs to a compact set Θ. In view of the map G and (4.14), we have

$$\bar{x}^{(N)}(t)-\bar{x}(t) = \int_{\mathbb{R}^n\times\Theta} x_*(t,\bar{x},x_0,\theta)\mathrm{d}P_{0\theta}^N(x_0,\theta) - \int_{\mathbb{R}^n\times\Theta} x_*(t,\bar{x},x_0,\theta)\mathrm{d}P_{0\theta}(x_0,\theta). \tag{4.16}$$

If $x_*(t,\bar{x},x_0,\theta)$ was uniformly bounded and equicontinuous with respect to the initial conditions and parameters, then one could show the convergence by (Stroock and Varadhan, 1979, Corollary 1.1.5). But $x_*(t,\bar{x},x_0,\theta)$ is discontinuous. Alternatively, we show that the set of discontinuity points has a measure zero under Assumption 10. We then show that $\bar{x}^{(N)}$ converges pointwise to $\bar{x}$. Finally, we prove the uniform convergence, from which the result follows.

Pointwise convergence. $P_{0\theta}^N$ converges in distribution to $P_{0\theta}$. Therefore, there exist on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ a sequence of random variables (X_0^N, ξ_θ^N) of distribution $P_{0\theta}^N$ and a random variable (X_0, ξ_θ) of distribution $P_{0\theta}$, such that (X_0^N, ξ_θ^N) converges with probability one to (X_0, ξ_θ). Thus,

$$\bar{x}^{(N)}(t) - \bar{x}(t) = \int_\Omega \left(x_* \left(t, \bar{x}, X_0^N(\omega), \xi_\theta^N(\omega)\right) - x_* \left(t, \bar{x}, X_0(\omega), \xi_\theta(\omega)\right)\right) \mathrm{d}\mathbb{P}(\omega). \tag{4.17}$$

For a fixed t, the discontinuity points of $x_* (t, \bar{x}, x_0, \theta)$, considered now as a function of x_0 and θ, are included in the set $D = \{(x_0, \theta) \in \mathbb{R}^n \times \Theta \,|\, x_0 \in \partial D_\theta^j(\bar{x})\}$. Under Assumption 10 and the independence of $x(0)$ and θ, one can prove that $P_0 \times P_\theta(D) = 0$. Hence, $x_* \left(t, \bar{x}, X_0^N, \xi_\theta^N\right)$ converges with probability one to $x_* (t, \bar{x}, X_0, \xi_\theta)$. The compactness of $[0, T]$ and Θ, and the continuity of $\Pi_{\theta j}$ imply

$$\left\| x_* \left(t, \bar{x}, X_0^N, \xi_\theta^N\right) - x_* \left(t, \bar{x}, X_0, \xi_\theta\right) \right\| \leq K_1 \|X_0^N\| + K_2 \|X_0\| + K_3, \tag{4.18}$$

for some finite $K_1, K_2, K_3 > 0$ independent of N. $\bar{x}^{(N)}(t)$ converges pointwise to $\bar{x}(t)$ for all $t \in [0, T]$ as a consequence of Assumption 8, Remark 10 and Lebesgue's dominated convergence theorem.

Uniform convergence. As in the proof of Theorem 5, one can show that for all t_1, t_2, $\left\|\bar{x}^{(N)}(t_1) - \bar{x}^{(N)}(t_2)\right\| \leq K|t_1 - t_2|$ and $\|\bar{x}(t_1) - \bar{x}(t_2)\| \leq K|t_1 - t_2|$, where $K > 0$ is independent of N. We fix an $\epsilon > 0$ and consider a partition $0 = t_0 < t_1 < \cdots < t_j = T$ of $[0, T]$, such that for all $t, t' \in [t_k, t_{k+1}]$, for all $N \geq 1$, $\|\bar{x}^{(N)}(t) - \bar{x}^{(N)}(t')\| < \epsilon$ and $\|\bar{x}(t) - \bar{x}(t')\| < \epsilon$. By the pointwise convergence, there exists N_0 such that for all $N > N_0$, for all $1 \leq k \leq j$, $\|\bar{x}^{(N)}(t_k) - \bar{x}(t_k)\| < \epsilon$. We fix $N > N_0$. For an arbitrary $t \in [0, T]$, there exists k such that $t \in [t_k, t_{k+1}]$. We have

$$\|\bar{x}^{(N)}(t) - \bar{x}(t)\| \leq \|\bar{x}^{(N)}(t) - \bar{x}^{(N)}(t_k)\| + \|\bar{x}^{(N)}(t_k) - \bar{x}(t_k)\| + \|\bar{x}(t_k) - \bar{x}(t)\| \leq 3\epsilon. \tag{4.19}$$

This inequality holds for an arbitrary $t \in [0, T]$, therefore, $\lim_{N \to \infty} \sup_{t \in [0,T]} \|\bar{x}^{(N)}(t) - \bar{x}(t)\|^2 = 0$. This implies (4.15), and proves the result. □

The following assumption guaranties the asymptotic optimality of the mean-field based strategies (4.12) under a non-convex final cost. For more details about this assumption, we refer the reader to Section 4.4 below.

Assumption 11. *We assume that $L \succeq 0$, where L is defined in* (4.9).

Assumption 11 is satisfied, for example, when $Z = \alpha I_n$, with $\alpha < 0$ or $\alpha \geq 2$. *In the first* *case ($\alpha < 0$), the social effect is to drive the agents away from the mean of the population, while in the second case ($\alpha \geq 2$) the social effect is attraction.*

The following main result states that the per-agent social cost under the strategies (4.12) converges to the per-agent optimal social cost as the number of agents increases to infinity. In other words, the strategies (4.12) are asymptotically socially optimal.

Theorem 14 (Approximate social optimum)**.** *Under Assumptions 8, 10, and 11,*

$$\lim_{N\to\infty}\left|\inf_{U\in\mathcal{U}^N}\frac{1}{N}J_{soc}\left(U,x^{(N)},x_i(0),x_{-i}(0)\right)-\frac{1}{N}J_{soc}\left(u_*^{(N)},\bar{x}^{(N)},x_i(0),x_{-i}(0)\right)\right|=0, \tag{4.20}$$

where $u_*^{(N)}$ is the strategy profile of the N agents when they apply (4.12)*, $\bar{x}^{(N)}$ defined in* (4.14) *the corresponding average state, and $x^{(N)}=\frac{1}{N}\sum_{i=1}^N x_i$ is the average state of the N agents under the strategy profile $U\in\mathcal{U}^N$.*

Proof. Let $\bar{x}$ be a fixed point of G and $U=(u_1,\dots,u_N)\in\mathcal{U}^N$, such that,

$$\frac{1}{N}J_{soc}\left(U,x^{(N)},x_i(0),x_{-i}(0)\right)\leq\frac{1}{N}J_{soc}\left(u_*^{(N)},\bar{x}^{(N)},x_i(0),x_{-i}(0)\right). \tag{4.21}$$

In the following, we show that U can improve the social cost by at most ϵ_N, which converges to zero as $N\to\infty$. Noting (4.13), the compactness of Θ, the continuity of $\Gamma_\theta^j(t)$ with respect to t and θ and Assumption 8, one can prove that $(1/N)J_{soc}\left(u_*^{(N)},\bar{x}^{(N)},x_i(0),x_{-i}(0)\right)<c_0$, where c_0 is independent of N. Therefore, $(1/N)J_{soc}\left(U,x^{(N)},x_i(0),x_{-i}(0)\right)<c_0$ and

$$\frac{1}{N}\sum_{i=1}^N\int_0^T\left\{\|u_i(t)\|^2+\|u_i^*(t)\|^2+\|x_i(t)\|^2+\|x_i^*(t)\|^2\right\}\mathrm{dt}<c_1, \tag{4.22}$$

where $x_i^*=x_*(t,\bar{x},x_i(0),\theta_i)$, $u_i^*=u_*(t,x_i^*,\bar{x},x_i(0),\theta_i)$ and $c_1>0$ is independent of N. Let $\tilde{x}_i=x_i-x_i^*$, $\tilde{x}^{(N)}=(1/N)\sum_{i=1}^N\tilde{x}_i$ and $\tilde{u}_i=u_i-u_i^*$. We obtain that,

$$\begin{aligned}\frac{1}{N}J_{soc}\left(U,x^{(N)},x_i(0),x_{-i}(0)\right)&=\frac{1}{N}J_{soc}\left(u_*^{(N)},\bar{x}^{(N)},x_i(0),x_{-i}(0)\right)+\frac{1}{N}\sum_{i=1}^N\int_0^T\tilde{u}_i'R_iu_i^*\mathrm{dt}\\&+\frac{1}{N}\sum_{i=1}^N\int_0^T\left\{\left\|\tilde{x}_i-Z\tilde{x}^{(N)}\right\|_Q^2+\|\tilde{u}_i\|_{R_i}^2+\left(\tilde{x}_i-Z\tilde{x}^{(N)}\right)'Q\left(x_i^*-Z\bar{x}^{(N)}\right)\right\}\mathrm{dt}\\&+\frac{1}{N}\sum_{i=1}^N\min_{1\leq j\leq l}\|x_i(T)-p_j\|_{M_i^j}^2-\frac{1}{N}\sum_{i=1}^N\min_{1\leq j\leq l}\|x_i^*(T)-p_j\|_{M_i^j}^2.\end{aligned} \tag{4.23}$$

For a fixed point $\bar{x}$ of G, and recalling (4.11), we have

$$J(u_i, \bar{x}, x_i(0), \theta_i) = J(u_i^*, \bar{x}, x_i(0), \theta_i) + \int_0^T \left\{\|\tilde{x}_i\|_Q^2 + \|\tilde{u}_i\|_{R_i}^2 + \bar{x}' L\tilde{x}_i + \tilde{x}_i' Q x_i^* + \tilde{u}_i' R_i u_i^*\right\}\mathrm{dt} + \min_{1\le j\le l} \|x_i(T) - p_j\|_{M_i^j}^2 - \min_{1\le j\le l} \|x_i^*(T) - p_j\|_{M_i^j}^2. \tag{4.24}$$

Now (4.23) and (4.24) yield

$$\begin{aligned}\frac{1}{N} J_{soc}\left(U, x^{(N)}, x_i(0), x_{-i}(0)\right) &= \frac{1}{N} J_{soc}\left(u_*^{(N)}, \bar{x}^{(N)}, x_i(0), x_{-i}(0)\right) \\ &+ \frac{1}{N}\sum_{i=1}^{N} \left(J(u_i, \bar{x}, x_i(0), \theta_i) - J(u_i^*, \bar{x}, x_i(0), \theta_i)\right) \qquad (4.25)\\ &+ \int_0^T \left(\tilde{x}^{(N)}\right)' L\tilde{x}^{(N)}\mathrm{dt} + \int_0^T \left(\bar{x}^{(N)} - \bar{x}\right)' L\tilde{x}^{(N)}\mathrm{dt}.\end{aligned}$$

By the bounds c_0 and c_1, Cauchy-Schwarz inequality, and Lemma 6, we deduce that $\epsilon_N := \int_0^T (\bar{x}^{(N)} - \bar{x})' L\tilde{x}^{(N)}\mathrm{dt}$ converges to 0 as N goes to infinity. The optimality of u_i^* with respect to J and Assumption 11 imply

$$\frac{1}{N} J_{soc}\left(U, x^n(N), x_i(0), x_{-i}(0)\right) \ge \frac{1}{N} J_{soc}\left(u_*^{(N)}, \bar{x}^{(N)}, x_i(0), x_{-i}(0)\right) + \epsilon_N. \tag{4.26}$$

This proves the result. □

In the following theorem, we give an explicit form for the asymptotic per-agent optimal social cost. The obtained expression can be computed by knowing only the probability distributions P_0 and P_θ, and a fixed point path $\bar{x}$ of G.

Theorem 15 (Asymptotic per-agent optimal social cost)**.** *Under Assumptions 8, 10, and 11,*

$$\lim_{N\to\infty} \inf_{U\in\mathcal{U}^N} \frac{1}{N} J_{soc}\left(U, x^{(N)}, x_i(0), x_{-i}(0)\right) = -\frac{1}{2}\int_0^T \bar{x}' L\bar{x}\mathrm{dt} + \mathbb{E}\sum_{j=1}^{l} 1_{D_\theta^j(\bar{x})}(x(0))\left\{\frac{1}{2} x(0)'\Gamma_\theta^j(0)x(0) + \beta_\theta^j(0)' x(0) + \delta_\theta^j\right\}. \tag{4.27}$$

Before proving Theorem 15, we start by showing the following preliminary result to approximate the asymptotic per-agent cost.

Lemma 7. *Under Assumptions 8, 10, and 11 ,*

$$\lim_{N\to\infty} \left| \inf_{U\in\mathcal{U}^N} \frac{1}{N} J_{soc}\left(U, x^{(N)}, x_i(0), x_{-i}(0)\right) - \frac{1}{N} J_{soc}\left(u_*^{(N)}, \bar{x}, x_i(0), x_{-i}(0)\right)\right| = 0. \tag{4.28}$$

Proof. Using the same notations as in the proof of Theorem 14, we have obtain that,

$$\begin{aligned}
&\frac{1}{N} J_{soc}\left(u_*^{(N)}, \bar{x}^{(N)}, x_i(0), x_{-i}(0)\right) - \frac{1}{N} J_{soc}\left(u_*^{(N)}, \bar{x}, x_i(0), x_{-i}(0)\right) \\
&= \int_0^T \frac{1}{N} \sum_{i=1}^{N} \left(\left\| x_i^* - Z\bar{x}^{(N)} \right\|_Q^2 - \| x_i^* - Z\bar{x} \|_Q^2 \right) \mathrm{dt} \qquad (4.29) \\
&= \int_0^T \left\| Z\left(\bar{x}^{(N)} - \bar{x}\right) \right\|_Q^2 \mathrm{dt} + \int_0^T \left(\bar{x}^{(N)} - Z\bar{x}\right)' QZ\left(\bar{x} - \bar{x}^{(N)}\right) \mathrm{dt}.
\end{aligned}$$

Cauchy-Schwarz inequality and Lemma 6 imply

$$\lim_{N\to\infty} \left| \frac{1}{N} J_{soc}\left(u_*^{(N)}, \bar{x}^{(N)}, x_i(0), x_{-i}(0)\right) - \frac{1}{N} J_{soc}\left(u_*^{(N)}, \bar{x}, , x_i(0), x_{-i}(0)\right) \right| = 0. \qquad (4.30)$$

Therefore, we deduce by Theorem 14 the result. □

Proof of Theorem 15. It is sufficient to show that

$$\lim_{N\to\infty} \left| J_{soc}^{\infty}(\bar{x}) - \frac{1}{N} J_{soc}\left(u_*^{(N)}, \bar{x}, x_i(0), x_{-i}(0)\right) \right| = 0. \qquad (4.31)$$

where

$$\begin{aligned}
J_{soc}^{\infty}(\bar{x}) = \mathbb{E}\Bigg[\int_0^T \Big\{ &\| x_*(t, \bar{x}, x(0), \theta) - Z\bar{x} \|_Q^2 + \| u_*(t, x_*(t, \bar{x}, x(0), \theta), x(0), \theta) \|_{R_\theta}^2 \Big\} \mathrm{dt} \\
&+ \min_{1\le j\le l} \| x_*(T, \bar{x}, x(0), \theta) - p_j \|_{M_\theta^j}^2 \Bigg]. \qquad (4.32)
\end{aligned}$$

The result follows then from Lemma 7 and $J_{soc}^{\infty}(\bar{x}) = \psi_4 - \frac{1}{2}\int_0^T \bar{x}' L \bar{x} \mathrm{dt}$, where

$$\psi_4 = \mathbb{E} J_*(u, \bar{x}, x(0), \theta) = \sum_{j=1}^{l} \int_{\mathbb{R}^n \times \Theta} 1_{D_\theta^j(\bar{x})}(x_0) \Big\{ x_0' \Gamma_\theta^j(0) x_0 + \beta_\theta^j(0)' x_0 + \delta_\theta^j(0) \Big\} \mathrm{d}P_{0\theta}(x_0, \theta), \qquad (4.33)$$

with J_* is the optimal cost of (4.11). We use the same notation as in the proof of Lemma 6. We have

$$J_{soc}^{\infty}(\bar{x}) - \frac{1}{N} J_{soc}\left(u_*^{(N)}, \bar{x}, x_i(0), x_{-i}(0)\right) = \psi_1 + \psi_2 + \psi_3, \qquad (4.34)$$

where

$$\psi_1 = \int_0^T \int_\Omega \left\{ \|x_* (t, \bar{x}, X_0(\omega), \xi_\theta(\omega)) - Z\bar{x}\|_Q^2 - \left\|x_* \left(t, \bar{x}, X_0^N(\omega), \xi_\theta^N(\omega)\right) - Z\bar{x}\right\|_Q^2 \right\} \mathrm{d}\mathbb{P}(\omega)\mathrm{dt} \tag{4.35}$$

$$\psi_2 = \int_0^T \int_\Omega \left\{ \|u_* (t, x_*, \bar{x}, X_0(\omega), \xi_\theta(\omega))\|_{R_{\xi_\theta(\omega)}}^2 - \left\|u_* \left(t, x_*, \bar{x}, X_0^N(\omega), \xi_\theta^N(\omega)\right)\right\|_{R_{\xi_\theta^N(\omega)}}^2 \right\} \mathrm{d}\mathbb{P}(\omega)\mathrm{dt} \tag{4.36}$$

$$\begin{aligned}\psi_3 = &\int_\Omega \min_{1\le j\le l} \|x_* (T, \bar{x}, X_0(\omega), \xi_\theta(\omega)) - p_j\|_{M^j_{\xi_\theta(\omega)}}^2 \mathrm{d}\mathbb{P}(\omega)\\ &- \int_\Omega \min_{1\le j\le l} \left\|\hat{x} \left(T, \bar{x}, X_0^N(\omega), \xi_\theta^N(\omega)\right) - p_j\right\|_{M^j_{\xi_\theta^N(\omega)}}^2 \mathrm{d}\mathbb{P}(\omega).\end{aligned} \tag{4.37}$$

Noting that $a'Pa - b'Pb = (a+b)'P(a-b)$ and that the minimum of l continuous functions is continuous, one can prove by the same techniques used in the proof of Lemma 6 that ψ_1, ψ_2 and ψ_3 converge to zero as N goes to infinity. This proves the result. □

4.4 Discussions

We discuss in this section the results of this Chapter.

4.4.1 Uniform population

In case of a uniform population, that is, the set of parameters is $\Theta = \{(A, B, R, M, \ldots, M)\}$, one can use similar techniques to those in Section 2.4 to construct a one-to-one map between the sustainable mean trajectories (fixed points of G) and the fixed points of a finite dimensional map. The latter are the potential probability distributions of the agents' ultimate choices. The finite dimensional map is defined on the simplex $\{(\lambda_1, \ldots, \lambda_l) \in \mathbb{R}^l | \lambda_j \geq 0 \text{ and } \sum_{j=1}^l \lambda_j = 1\}$ as follows,

$$F(\lambda) = \left(P_0\left(D^1(\bar{x}^\lambda)\right), \ldots, P_0\left(D^l(\bar{x}^\lambda)\right)\right), \tag{4.38}$$

where $\bar{x}^\lambda$ is the optimal state of the following LQR problem,

$$\begin{aligned}&\int_0^T \left\{\|x\|_{L+Q}^2 + \|u\|_R^2\right\} dt + \|x(T) - p_\lambda\|_M^2\\ &\text{s.t. } \frac{d}{dt}x = Ax + Bu,\end{aligned} \tag{4.39}$$

with $p_\lambda = \sum_{j=1}^{l} \lambda_j p_j$, where $\lambda = (\lambda_1, \ldots, \lambda_l)$. This characterization of the sustainable mean trajectories has similar consequences to those in the non-cooperative case. For further discussions about these consequences, we refer the reader to Section 2.5.

4.4.2 Exact vs. approximate social optima

We develop in Section 4.3 an approximate social optimum (4.12). The corresponding strategies exhibit the following advantageous properties with respect to the exact solution developed in Section 4.2:

- **Tractable computations:** To compute a mean-field based strategy (4.12), an agent needs to compute a sustainable mean trajectory and the l LQR costs corresponding to the l alternatives, and pick the less costly one. Moreover, each LQR problem involves solving a Riccati equation of dimension $n \times n$ and a linear ODE of dimension $n+1$. In the uniform case, the computation of a sustainable mean trajectory reduces to finding a fixed point to the finite dimensional map F defined in (4.38).

- **Small amount of communication:** According to (4.12), the implementation of an approximate social optimum requires that each agent know only its own state and the probability distributions P_0 and P_θ.

We summarize and compare in Table 4.1 the main properties of the exact and approximate social optima.

Table 4.1 Cooperative DCDC: Exact and approximate social optima

	Exact social optimum	Approximate social optimum
Computation	Solve l^N LQR costs each involving an ODE of dim. $n^2N^2 + nN + 1$	Compute a sustainable mean and solve l LQR costs each involving an ODE of dim. $n^2 + n + 1$
Implementation	Each agent needs to know its state and parameters, and the states and parameters of all the other agents.	Each agent needs to know its state and parameters, and the distributions P_0 and P_θ

4.4.3 Originality of the convergence proofs and technical assumptions

Need for Assumption 11

A social optimum is necessarily person-by-person optimal. The inverse occurs under some conditions. For example, in static games, it is required that the cost be convex and smooth (Yüksel and Başar, 2013, Lemma 2.6.1). This condition, which is automatically satisfied in the standard LQG MFG's (Huang et al., 2012), is also implicitly used to show the convergence of the person-by-person solution to the social optimum. In our case, however, Assumption 11 is needed to deal with the non-convexity and non-smoothness of the final costs. Indeed, if we apply the techniques used in (Huang et al., 2012, Theorem 4.2) to the proof of Theorem 14, then by the convexity of the running cost, (4.23) implies that for all $U \in \mathcal{U}^N$,

$$\begin{aligned} J_{soc}\left(U, x^{(N)}, x_i(0), x_{-i}(0)\right) \geq\ & J_{soc}\left(u_*^{(N)}, \bar{x}^{(N)}, x_i(0), x_{-i}(0)\right) \\ & + \sum_{i=1}^{N} \int_0^T \left\{\tilde{u}_i' R_i u_i^* + \tilde{x}_i' \left(Q x_i^* + L\bar{x}\right)\right\} \mathrm{dt} + N\epsilon_N \\ & + \sum_{i=1}^{N} \min_{1\leq j\leq l} \|x_i(T) - p_j\|^2_{M_i^j} - \sum_{i=1}^{N} \min_{1\leq j\leq l} \|x_i^*(T) - p_j\|^2_{M_i^j}. \end{aligned} \tag{4.40}$$

We have $\frac{d}{dt}\tilde{x}_i'(\Gamma^j_{\theta_i} x_i^* + \beta^k_{\theta_i}) = -\tilde{u}_i' R_i u_i^* - \tilde{x}_i' \left(Q x_i^* + L\bar{x}\right)$. Hence,

$$\begin{aligned} \frac{1}{N} J_{soc}\left(U, x^{(N)}, x_i(0), x_{-i}(0)\right) \geq \frac{1}{N} J_{soc}\left(u_*^{(N)}, \bar{x}^{(N)}, x_i(0), x_{-i}(0)\right) \quad (4.41) \\ + \frac{1}{N} \sum_{i=1}^{N} \left(\phi_i\left(x_i(T)\right) - \phi_i\left(x_i^*(T)\right) - \tilde{x}_i'(T) \frac{d}{dx}\phi_i\left(x_i^*(T)\right)\right) + \epsilon_N, \end{aligned}$$

where ϕ_i is the final cost of agent i. If the final costs are smooth and convex (which is not the case), then (4.41) implies (4.20). To handle the non-convexity, steps (4.40) and (4.41) are replaced by (4.24), (4.25) and Assumption 11.

The nature of the initial conditions matters

In the LQG MFG's (Huang et al., 2012), the agents' optimal states are continuous with respect to the initial conditions. As a result, the authors could consider deterministic initial conditions, include them in the vector of parameters, and then "randomize" the limiting sequence $\{(x_i(0), \theta_i)\}_{i\geq 1}$ by assuming that the empirical measure of the finite sequence of $\{(x_i(0), \theta_i)\}_{1\leq i\leq N}$ converges weakly to a probability distribution (For more details about this randomization process, we refer the reader to (Huang et al., 2007)). This would not change the analysis (mainly the convergence proofs to a social optimum) with respect to the random initial conditions case. In our case, however, the optimal state (4.13) is discontinuous with respect to the initial condition. If we assume random initial conditions, then as in equation (2.80) in the degenerate DCDC game, Lemma 6 is a direct consequence of (Stroock and Varadhan, 1979, Corollary 1.1.5). But, for deterministic initial conditions, we need some constructions, the independence of P_0 and P_θ, and Assumption 11 to bypass the discontinuity (See the proof of Lemma 6).

4.5 Simulation Results

In this section, we compare via a numerical example the cooperative and non-cooperative behaviors in the DCDC problems. To this end, we draw 300 agents from the Gaussian distribution $\mathcal{N}([-5, 10], 15I_2)$. These agents have the following dynamics and cost parameters,

$$A_i = \begin{bmatrix} 0 & 1 \\ 0.02 & -0.3 \end{bmatrix} \qquad B_i = \begin{bmatrix} 0 \\ 0.3 \end{bmatrix}, \tag{4.42}$$

$R_i = 10I_2$, $M_i^j = 1200I_2$, $T = 2$, $p_1 = -p_2 = [-10 \quad 0]'$ and $Z = 3.5I_2$. We start by the situation where the agents make their choices in the absence of a social effect $Q = 0$. Afterwards, we increase gradually the social pressure, by increasing Q. Figure 4.1 illustrates the initial and final positions of the agents in the cooperative and non-cooperative cases, for $Q = 0$, $Q = 20I_2$ and $Q = 40I_2$. Moreover, it shows the basins of attraction in both cases. Figure 4.2 shows the choices' distribution and the corresponding per agent costs for different values of Q. For the computation of the agents' optimal strategies, we refer the reader to Section 4.4.1 for the cooperative case and Figure 2.1 for the non-cooperative one. In the absence of a social effect, the individual costs are decoupled, and each agent minimizes its own cost. Hence, the optimal strategies are identical in the cooperative and non-cooperative cases. Moreover, $\lambda = (0.17, 0.83)$ is a fixed point for the finite dimensional maps (4.38) and (2.45). Accordingly, the majority of the agents, 83%, choose p_2. As the social pressure increases, the cooperative and non-cooperative cases exhibit two opposite behaviors. The agents tend in the non-cooperative case to follow the majority, whereas in the cooperative context, the social pressure distributes the agents evenly between the alternatives. Indeed, in the non-cooperative case, the agents look at the mean as an exogenous trajectory. They try to beat it by moving in the same direction where the majority lies initially. As a result, the mean follows them, attracting more and more agents towards the majority. In the cooperative case, however, the agents cooperate to build a socially favorable mean, that is, a mean that doesn't make the majority better-off at the expenses of "harming" the minority. Hence, a fraction of the majority follows the minority to drift the mean towards the latter and ameliorate its cost. Consequently, this cooperation redistributes the agents between the alternatives.

Finally, it should be noted that in the cooperative case with $Q = 20I_2$, the 300 agents need in total 58 sec to compute their decentralized mean-field based strategies (4.12). On the other hand, the computation of one optimal value of (4.8) takes 56 sec. Hence, one needs (naïvely) $2^{300} \times 56$ sec to compute the exact social optimum[1].

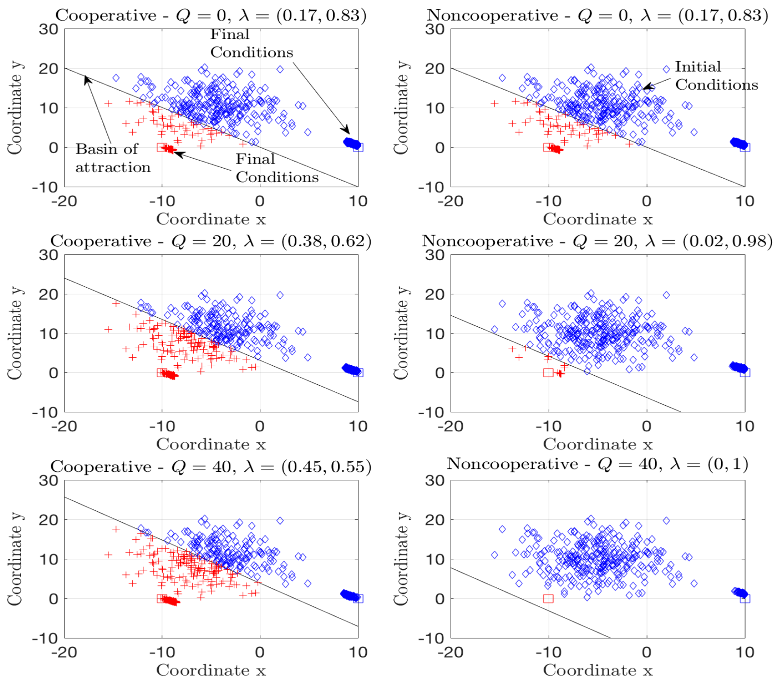

Figure 4.1 Comparison of the cooperative and non-cooperative behaviors in the collective discrete choices problems.

4.6 Conclusion

We present in this chapter the cooperative DCDC model. We show that the naïve approach to computing an exact solution becomes quickly intractable and its implementation requires a significant amount of communication as the number of agents increases sufficiently. Instead, we develop via the MFG methodology a set of decentralized simple strategies that qualify as socially optimal as the size of the population increase to infinity.

Hitherto, the agents make their choices of alternatives under the social pressure. In the following chapter, we study situations where the individual choices are shaped by both the

[1]The centralized and decentralized solutions are computed on a 2.7 GHz Intel Core i5 processor.

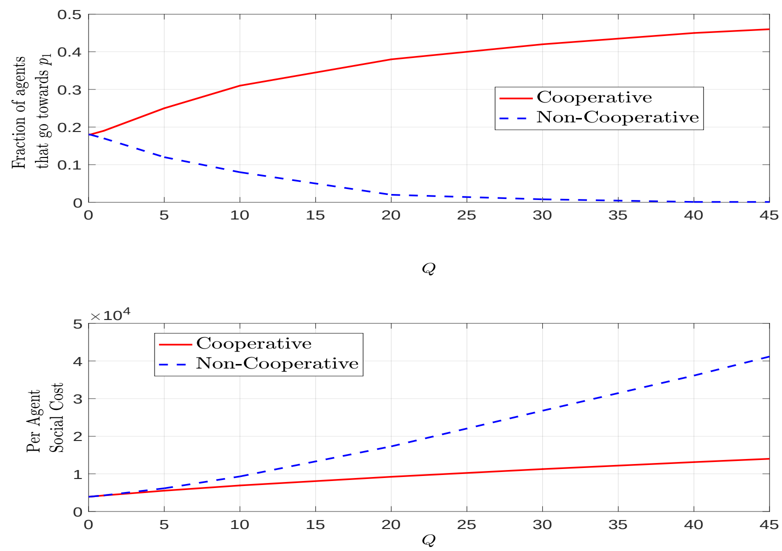

Figure 4.2 Comparison of the choices' distributions and per agent costs in the cooperative and non-cooperative cases.

social and advertising effects.

CHAPTER 5 DYNAMIC COLLECTIVE DISCRETE CHOICE MODEL WITH AN ADVERTISER

We discuss in Chapter 1 the relation between the DCDC problems and elections. These models, however, do not take into consideration the candidates' electoral campaigns that influence the voters' opinions. The aim of this chapter is to extend the non-cooperative degenerate DCDC model to account for such situations. More precisely, we consider a dynamic choice problem, where the consumers choose between two alternatives under the social and advertising effects. The latter is the effort exerted by an advertiser to encourage the consumers to choose a specific alternative. We formulate in Section 5.1 the DCDCA problem as a Stackelberg competition involving an advertiser and a large number of consumers. The advertiser makes some investments to generate an advertising effect and influence the paths of the consumers to ultimately choose a specific alternative. Subsequently, the consumers choose between two alternatives under the social and advertising effects. In Section 5.2, we study the continuum of consumers case via the MFG methodology. We show that the advertiser can always make optimal investments when the consumers' a priori opinions towards the alternatives are sufficiently *diverse*. In other words, we show that there exists a Stackelberg solution (See definition 4) if the consumers' initial states are sufficiently spread. The corresponding strategies, when applied by a finite number of consumers, constitute an approximate Stackelberg solution. In Section 5.3, we study the special case where a priori opinions have a certain uniform distribution. We give an explicit form of the advertiser's optimal investment strategy, and the consumers' optimal choices. Moreover, we anticipate the choices' distribution under the social and advertising effects. Section 5.4 provides some numerical simulation results, while Section 5.5 concludes this chapter.

5.1 Mathematical model

We formulate the DCDCA problem as a Stackelberg competition involving N consumers and an advertiser of respective dynamics,

$$\frac{d}{dt}x_i = Ax_i + Bu_i, \qquad 1 \leq i \leq N, \tag{5.1}$$

$$\frac{d}{dt}y = A_0 y + B_0 v, \tag{5.2}$$

where $x_i \in \mathbb{R}^n$ and $u_i \in \mathcal{U}$ are the state and control input of the consumer i. The set of admissible strategies $\mathcal{U}$ is defined in (2.3), Chapter 2. $y \in \mathbb{R}^{n_1}$ and $v \in L_2([0,T],\mathbb{R}^{m_1})$ are the state and control input of the advertiser. We assume that the initial conditions $x_i(0)$, $1 \le i \le N$, are i.i.d. random vectors on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with distribution P_0. The consumers and advertiser are associated with the following individual costs,

$$J_i(u_i, \bar{x}, v) = \mathbb{E}\left(\int_0^T \left\{\|x_i - \alpha\bar{x} - K(p_2)y\|_Q^2 + \|u_i\|_R^2\right\} \mathrm{dt} + \min_{j=1,2} \|x_i(T) - p_j\|_M^2\right), \tag{5.3}$$

$$J_0(v, \bar{x}) = \mathbb{E}\left(\int_0^T \|v\|_{R_0}^2 \mathrm{dt} + \|\bar{x}(T) - p_2\|_{M_0}^2\right), \tag{5.4}$$

for $1 \le i \le N$, where $\alpha \ge 0$, $Q \succeq 0$, $R \succ 0$, $R_0 \succ 0$, $M \succ 0$, $M_0 \succ 0$, $\bar{x} = 1/N \sum_{i=1}^N x_i$, and p_1 and p_2 are the alternatives. We explain the meaning of these costs via the following example.

Example 1 (Ministry of health campaign against smoking)**.** *Consider a group of N teenagers choosing before a time T between smoking ($p_1 = 1$) or not smoking ($p_2 = -1$). At time t, teenager i's smoking inclination is modeled by a variable $x_i(t) \in [-1, 1]$, where the value -1 corresponds to a nonsmoker, while 1 represents a full smoker. The effort exerted by i at time t to change its smoker status is modeled by $u_i(t) \in \mathbb{R}$. For example, $|u_i|$ would represent the amount of money spent per unit of time paid to increase (buying extra cigarettes) or decrease (medical treatment) its smoker status. On the other hand, the government rate of investments against smoking is modeled by a variable $v \in \mathbb{R}$. The variable y represents the effectiveness of the advertising investment. The influence exerted by the advertisement on the teenagers' smoking status is modeled by $K(p_2)y$, where $K(p_2) := p_2 =$ "Do Not Smoke" $= -1$. A teenager, in the process of choosing between not smoking or smoking, minimizes the cost* (5.3)*, which penalizes along the path the deviation from the peers smoking status $\bar{x}$ and the government nonsmoking advertisement $K(p_2)y$, as well as the effort to change the smoking status. Moreover, the teenager should be by time T a smoker (p_1) or nonsmoker (p_2) lest he/she be considered indecisive by its peers. Thus, lack of a decision by time T is strongly penalized in the final cost. On the other hand, the government tries to minimize its advertisement investments (the running cost of* (5.4)*), and should convince by time T the teenagers to be nonsmokers. Failure to sway a majority of teenagers away from smoking results in a strong penalty in the final cost.*

5.2 Mean field Stackelberg competition

In our Stackelberg competition, the advertiser plays first, and then the consumers make their choices. Thus, the game is solved as follows. Given an advertiser's investment strategy v, the consumers solve a degenerate non-cooperative min-LQG game parametrized by v. If for each admissible v the consumers' game has a unique Nash equilibrium $(u_i^*(v), u_{-i}^*(v))$, then the advertiser knows how the consumers will respond to his/her investment strategies. Technically speaking, he/she constructs a map $Nash$ that maps its strategy v to the corresponding consumers' Nash equilibrium $Nash(v) = (u_i^*(v), u_{-i}^*(v))$. Finally, an advertiser's optimal strategy is an optimal control law of $J_0(v, \bar{x}(v))$, where $\bar{x}(v)$ is the average of the consumers under their Nash equilibrium $Nash(v)$. The main challenge here is to derive conditions for the existence of a unique Nash equilibrium to hold. Indeed, we saw in Section 2.2 that the Nash equilibria of the degenerate non-cooperative min-LQG games may not exist.

Instead, we solve the game via the MFG methodology. We start by assuming a continuum of consumers. Following the results of Chapter 2, the consumers' Nash equilibria are totally determined by the sustainable mean trajectories. If for each v there exists a unique Nash equilibrium, then the advertiser constructs a map that maps v to the consumers' sustainable trajectory $\bar{x}_v$. The advertiser minimizes then $J_0(v, \bar{x}_v)$. Once the infinite population game is solved, we show that the mean-field strategies, when applied by the advertiser and a finite number of consumers, constitute an approximate Stackelberg solution.

5.2.1 Consumers' game

We start by assuming a continuum of consumers. Given the advertiser's strategy v and corresponding state y, the consumers solve a degenerate min-LQG game parametrized by v. The results of Chapter 2 show that this game is fully characterized by the sustainable mean trajectories. To compute these trajectories, and the consumers' optimal strategies, we define Γ, β^j and δ^j to be the unique solutions of (2.20)-(2.22), where we replace $\bar{x}$ in by $\alpha\bar{x} + K(p_2)y$. The functions Γ, β^j and δ^j, which are related to the LQR problem that solves a generic consumer when only p_j is available, depend on $\alpha\bar{x} + K(p_2)y$ rather than $\bar{x}$. Indeed, when the advertiser's state and consumers' mean are fixed, a generic consumer optimally tracks $\alpha\bar{x} + K(p_2)y$. We recall that the basin of attraction $D(\bar{x}, y)$ of p_1 is the set of points in $\mathbb{R}^n$, such that, if a consumer is initially in this region, then the optimal LQR cost associated with p_1 is less than that associated with p_2. By solving the equations of β^j and δ^j, one can show that $D(\bar{x}, y)$ has the following form,

$$D(\bar{x}, y) = \{x_0 \in \mathbb{R}^n | \beta' x_0 \leq \delta + \Delta(\alpha\bar{x} + K(p_2)y)\}, \tag{5.5}$$

where $\beta = \beta^1(0) - \beta^2(0)$ does not depend on $\bar{x}$ and y. $\delta^2(0) - \delta^1(0) := \delta + \Delta(\alpha\bar{x} + K(p_2)y)$, where δ is the term that does not depend on $\bar{x}$ and y, and Δ is a linear form on $L_2([0,T], \mathbb{R}^n)$, defined as follows,

$$\Delta(x) = (p_1 - p_2)' M \int_T^0 \int_T^\eta \phi(\eta, T)' B R^{-1} B' \phi(\eta, \sigma) Q x(\sigma)\, \mathrm{d}\sigma \mathrm{d}\eta, \tag{5.6}$$

with ϕ is the unique solution of $\frac{d}{dt}\phi(t,s) = \left(\Gamma(t) B R^{-1} B' - A'\right) \phi(t,s)$, $\phi(s,s) = I_n$.

Using techniques similar to those used in Theorems 4 and 11, one can construct a one-to-one map between the sustainable mean trajectories $\bar{x}$ and the fixed points of a finite dimensional map, and show the existence of a sustainable trajectory under the following Assumptions.

Assumption 12. *The following Riccati equation has a unique solution:*

$$\frac{d}{dt}\gamma = -\gamma A - A'\gamma + \gamma B R^{-1} B' \gamma - (1-\alpha) Q, \qquad \gamma(T) = M. \tag{5.7}$$

Assumption 13. *We assume that P_0 is such that the P_0-measure of hyperplanes is zero.*

Note that if $\alpha \leq 1$, then (5.7) has a unique solution (Anderson and Moore, 2007, page 23). The finite dimensional map is defined as follows. $\forall \lambda \in [0,1]$,

$$F(\lambda, y) = P_0\left(D\left(\bar{x}^\lambda, y\right)\right), \tag{5.8}$$

where

$$\begin{aligned} \bar{x}^\lambda(t) &= R_1(t,0)\mathbb{E}x(0) + R_2(t)\left(\lambda p_1 + (1-\lambda) p_2\right) + \Xi(y)(t) \\ \frac{d}{dt} R_1(t,s) &= \left(A - B R^{-1} B' \gamma\right) R_1(t,s), \qquad R_1(s,s) = I_n \\ \frac{d}{dt} R_2(t) &= \left(A - B R^{-1} B' \gamma(t)\right) R_2(t) + B R^{-1} B' R_1(T,t)' M, \qquad R_2(0) = 0 \\ \Xi(y)(t) &= -\int_0^t \int_T^\sigma R_1(t,\sigma) B R^{-1} B' R_1(\tau,\sigma)' Q K(p_2) y(\tau) \mathrm{d}\tau \mathrm{d}\sigma, \end{aligned} \tag{5.9}$$

and $x(0)$ is a generic consumer's initial state, which has a distribution P_0. We summarize the solution of the consumers' game in the following theorem.

Theorem 16 (Consumers' game solution)**.** *Under Assumptions 12 and 13, the following statements hold:*

(i) $\bar{x}(t)$ is a sustainable mean trajectory if and only if

$$\bar{x}(t) = \bar{x}^\lambda, \tag{5.10}$$

where $\lambda = F(\lambda, y)$, i.e. λ is fixed point of $\lambda \mapsto F(\lambda, y)$ defined in (5.8)*.*

(ii) *The function* $\lambda \mapsto F(\lambda, y)$ *has at least one fixed point* λ*, equivalently there exists at least one sustainable mean trajectory.*

(iii) *A generic consumer's best response to a sustainable mean trajectory* $\bar{x}$ *is given by,*

$$u_*(t, x, \bar{x}, x(0), y) = \begin{cases} -R^{-1}B'\left(\Gamma(t)x + \beta^1(t)\right), & \text{if } x(0) \in D(\bar{x}, y) \\ -R^{-1}B'\left(\Gamma(t)x + \beta^2(t)\right), & \text{if } x(0) \notin D(\bar{x}, y), \end{cases} \tag{5.11}$$

where $x(0)$ *is the generic consumer's initial state.*

Theorem 16 states that there exists at least one Nash equilibrium for each advertiser's strategy v. But, the advertiser needs to anticipate *uniquely* the consumers' behavior in response to his/her investment strategy v. The following assumption provides a condition under which the uniqueness of the consumers' Nash equilibria holds.

Assumption 14. *We assume that* $\bar{F}(s) := P_0\left(\beta' x_0 \leq \delta + s\right)$ *is differentiable and* $\left|\frac{d}{ds}\bar{F}(s)\right| < \frac{1}{|\alpha\Delta(R_2(p_1 - p_2))|}$ *(here the linear form* Δ *defined in* (5.6) *acts on the function* $R_2(t)(p_1 - p_2)$*).*

Noting that the function $\frac{d}{ds}\bar{F}$ is the probability density function of $\beta' x(0) - \delta$, Assumption 14 requires that the consumers' a priori opinions $x(0)$ be sufficiently spread in the direction β. For example, if the consumers' a priori opinions $x(0)$ are distributed according to the normal distribution $\mathcal{N}(\mu_0, \Sigma_0)$, then $\beta' x(0) - \delta$ is distributed according to $\mathcal{N}(\beta'\mu_0 - \delta, \beta'\Sigma_0\beta)$, and the corresponding probability density function has a maximum $1/\left(\sqrt{2\pi\beta'\Sigma_0\beta}\right)$. In this case, Assumption 14 is satisfied if $2\pi\beta'\Sigma_0\beta > \left(\alpha\Delta\left(R_2(p_1 - p_2)\right)\right)^2$. Under Assumption 14, the function $\lambda \mapsto F(\lambda, y)$ is a contraction. Indeed, $\frac{d}{d\lambda}F = \alpha\Delta\left(R_2(p_1 - p_2)\right)\frac{d}{ds}\bar{F}$, which under Assumption 14 has an absolute value strictly less than one. Therefore, we can state the following theorem.

Theorem 17 (Unique consumers' behavior)**.** *Under Assumptions 12, 13, and 14, given the advertiser's strategy* v*,* $\lambda \mapsto F(\lambda, y)$ *has a unique fixed point. Thus, the consumers' limiting game admits a unique Nash equilibrium.*

To summarize, if the consumers have diverse a priori opinions toward the alternatives, then for each advertising strategy v, the distribution of the consumers' choices (i.e. $(\lambda, 1 - \lambda)$ for a fixed point λ of $\lambda \mapsto F(\lambda, y)$), and the corresponding mean trajectory $\bar{x}^\lambda$ defined in (5.9) are unique. In other words, the consumers react macroscopically in a unique way to an investment strategy. On the individual level, a consumer chooses alternative p_1 if his/her a priori opinion $x(0)$ is in the basin of attraction $D(\bar{x}^\lambda, y)$. Otherwise, he/she chooses p_2.

5.2.2 Advertiser's game

Having determined the consumers' individual and macroscopic (mean trajectory) responses to the investment strategies, we turn now to the problem of finding an optimal investment policy v. We assume in the rest of this chapter that Assumptions 12, 13, and 14 hold to guaranty the existence of a unique Nash equilibrium for each $v \in L_2([0,T],\mathbb{R}^{m_1})$. The sustainable mean trajectory that corresponds to v is denoted by $\bar{x}_v$ in this section. It is the mean trajectory of the consumers under their best responses $u_*(t,x,\bar{x}_v,x(0),y)$ to it. Thus, the advertiser solves the following optimal control problem:

$$\begin{aligned} &\min_{v\in L_2([0,T],\mathbb{R}^{m_1})} J_0\left(v,\bar{x}_v\right) \\ \text{s.t. } &\frac{d}{dt}y = A_0 y + B_0 v \text{ and } \frac{d}{dt}\bar{x}_v = A\bar{x}_v + B\mathbb{E}u_*(t,x,\bar{x}_v,x(0),y). \end{aligned} \tag{5.12}$$

In the following theorem, we show that if the consumers' a priori opinions are sufficiently diverse, then the advertiser can find an optimal investment policy v_*. Afterwards, we characterize in Theorem 19 this strategy as the co-state of $(y_{v_*},\bar{x}_{v_*})$, where y_{v_*} is the advertiser's optimal state that corresponds to v_*. This allows us to derive explicit optimal investment policies in some situations, for example, in case the consumers' initial states are uniformly distributed in the direction β (See Section 5.3 below).

Theorem 18 (Existence of an optimal investment strategy)**.** *Under Assumptions 12, 13 and 14, the advertiser's optimal control problem* (5.12) *has an optimal control law* v_*.

Proof. The cost functional J_0 is positive and coercive with respect to $v \in L_2([0,T],\mathbb{R}^{m_1})$, i.e. $\lim_{\|v\|_{L_2}\to\infty} J_0(v,\bar{x}_v)/\|v\|_{L_2} = \infty$. Thus, it is sufficient to show that J_0 is continuous in the reflexive Banach space $L_2([0,T],\mathbb{R}^{m_1})$ w.r.t. v. Tonelli's existence theorem (Clarke, 2013, Theorem 5.51) guaranties then the existence of a minimum for J_0. The state y is continuous with respect to v. The fixed points $\lambda(y)$ of F are also continuous with respect to v. In fact, suppose v and v' two investment strategies in $L_2([0,T],\mathbb{R}^{m_1})$, and denote y and y' the corresponding advertiser's states, and λ and λ' the corresponding fixed points of $\lambda \mapsto F(\lambda,y)$ and $\lambda' \mapsto F(\lambda',y')$. We have

$$\begin{aligned} |\lambda-\lambda'| &= |F(\lambda,y)-F(\lambda',y')| \leq |F(\lambda,y)-F(\lambda',y)| + |F(\lambda',y)-F(\lambda',y')| \\ &\leq \sup_{\alpha\in[0,1]}\left|\frac{dF}{d\lambda}(\alpha,y)\right| |\lambda-\lambda'| + |F(\lambda',y)-F(\lambda',y')|. \end{aligned} \tag{5.13}$$

Therefore,

$$\left(1 - \sup_{\alpha\in[0,1]} \left|\frac{dF}{d\lambda}(\alpha, y_2)\right|\right) |\lambda - \lambda'| \leq |F(\lambda', y) - F(\lambda', y')|. \tag{5.14}$$

Under Assumption 14, $\sup_{\alpha\in[0,1]} \left|\frac{dF}{d\lambda}(\alpha, y)\right| < 1$. Moreover, under Assumption 13, $\bar{F}$ is continuous. Δ is continuous with respect to the norm $\|\cdot\|_{L_2}$. Hence, F is continuous with respect to y, and $|F(\lambda', y) - F(\lambda', y')|$ converges to zero as $\|y - y'\|_{L_2}$ converges to zero. Therefore, the fixed points λ of F are continuous. In view of (5.10) and the continuity of the fixed points λ, $\bar{x}_v(T)$ is continuous. Therefore, J_0 is continuous. □

In the following theorem, we characterize an optimal strategy v_* as the co-state of $(\bar{x}_{v_*}, y_{v_*})$. Given an optimal control law v_*, we define the co-state equations:

$$-\frac{d}{dt}P = A_0'P + \mathcal{L}_1^*(W)(t), \qquad P(T) = 0 \tag{5.15}$$

$$-\frac{d}{dt}W = \mathcal{L}_2^*(W)(t), \qquad W(T) = M_0(\bar{x}_{v_*}(T) - p_2), \tag{5.16}$$

where for all $z \in L_2([0,T], \mathbb{R}^n)$

$$\mathcal{L}_1^*(z)(t) = K(p_2)'Q\int_0^t R_1(t,\sigma)BR^{-1}B'z(\sigma)\mathrm{d}\sigma + \xi^* K(p_2)'H(t)\int_0^T R_1(T,\sigma)BR^{-1}B'z(\sigma)\mathrm{d}\sigma$$

$$\mathcal{L}_2^*(z)(t) = (A - BR^{-1}B'\gamma)'z(t) + \xi^*\alpha H(t)\int_0^T R_1(T,\sigma)BR^{-1}B'z(\sigma)\mathrm{d}\sigma, \tag{5.17}$$

$$\text{with } \xi^* = \frac{d\bar{F}}{ds}\left(\Delta\left(\alpha\bar{x}_{v_*} + K(p_2)y_{v_*}\right)\right).$$

Here, $H(t) = Q\int_0^t \phi(\eta,t)'BR^{-1}B'\phi(\eta,T)\mathrm{d}\eta M(p_1 - p_2)(p_1 - p_2)'M$, Δ defined in (5.6) and ϕ defined below (5.6).

Theorem 19 (Characterization of the optimal investment strategies)**.** *Under Assumptions 12, 13, and 14, if v_* is an optimal control law of* (5.12) *and the corresponding equations* (5.15)-(5.16) *have a unique solution (P, W), then*

$$v_* = -R_0^{-1}B_0'P. \tag{5.18}$$

Proof. We derive the condition on v_* (5.18) by studying the first variation of the cost functional in (5.12) with respect to a perturbation $v = v_* + \eta\delta v$, where $\eta \in \mathbb{R}$, and $\delta v \in L_2([0,T], \mathbb{R}^{m_1})$. To this end, we need to derive at first an explicit form of the constraint on $\bar{x}_v$. We have that $\bar{x}_v = \bar{x}^\lambda$ defined in (5.9), where λ is the unique fixed point of

$\lambda \mapsto F(\lambda, y)$. By taking the derivative of $\bar{x}^\lambda$ with respect to time, we obtain that,

$$\frac{d}{dt}\bar{x}_v = \mathcal{L}(\bar{x}_v, y)(t), \tag{5.19}$$

where

$$\begin{aligned}\mathcal{L}(\bar{x}_v, y)(t) &= (A - BR^{-1}B'\gamma)\bar{x}_v + BR^{-1}B'R_1(T,t)'\bar{F} \circ \Delta\Big(\alpha\bar{x}_v + K(p_2)y)\Big)M(p_1 - p_2) \\ &\quad + BR^{-1}B'R_1(T,t)'Mp_2 - BR^{-1}B'\int_T^t R_1(\sigma,t)'QK(p_2)y(\sigma)\mathrm{d}\sigma.\end{aligned} \tag{5.20}$$

We compute now the Gâteaux derivatives (Clarke, 2013) of y and $\bar{x}$ at v_* in the direction δv:

$$\begin{aligned}\frac{d}{d\eta}y_{v^*+\eta\delta v}\Big|_{\eta=0} &:= \delta y \\ \frac{d}{d\eta}\bar{x}_{v^*+\eta\delta v}\Big|_{\eta=0} &:= \delta\bar{x},\end{aligned} \tag{5.21}$$

where,

$$\frac{d}{dt}\delta y = A_0\delta y + B_0\delta v, \quad \delta y(0) = 0 \tag{5.22}$$

$$\frac{d}{dt}\delta\bar{x} = \mathcal{L}_1(\delta y)(t) + \mathcal{L}_2(\delta\bar{x})(t), \quad \delta\bar{x}(0) = 0, \tag{5.23}$$

and $\mathcal{L}_1$ (resp. $\mathcal{L}_2$) is a continuous linear operator from the Hilbert space $L_2([0,T],\mathbb{R}^{n_1})$ (resp. $L_2([0,T],\mathbb{R}^n)$) to $L_2([0,T],\mathbb{R}^n)$, such that for all $z_1 \in L_2([0,T],\mathbb{R}^{n_1})$ and $z_2 \in L_2([0,T],\mathbb{R}^n)$,

$$\begin{aligned}\mathcal{L}_1(z_1)(t) &= -BR^{-1}B'\int_T^t R_1(\sigma,t)'QK(p_2)z_1(\sigma)\mathrm{d}\sigma \\ &\quad + \xi^*\Delta\left(K(p_2)z_1\right)BR^{-1}B'R_1(T,t)'M(p_1 - p_2)\end{aligned} \tag{5.24}$$

$$\mathcal{L}_2(z_2)(t) = \Big(A - BR^{-1}B'\gamma(t)\Big)z_2(t) + \alpha\xi^*\Delta(z_2)BR^{-1}B'R_1(T,t)'M(p_1 - p_2) \tag{5.25}$$

Using Fubini-Tonelli's theorem (Rudin, 1987), one can show that the adjoint operators of $\mathcal{L}_1$ and $\mathcal{L}_2$ are respectively $\mathcal{L}_1^*$ and $\mathcal{L}_2^*$ defined in (5.17). We recall from (Rudin, 1991) that the adjoint operator of a linear continuous operator $\mathcal{G}$ defined from the Hilbert space $(H_1, \langle,\rangle_1)$ into the Hilbert space $(H_2, \langle,\rangle_2)$ is the linear continuous operator $\mathcal{G}^*$ defined from the Hilbert space $(H_2, \langle,\rangle_2)$ into the Hilbert space $(H_1, \langle,\rangle_1)$ and satisfying for all $x \in H_1$ and $y \in H_2$ $\langle\mathcal{G}(x), y\rangle_2 = \langle x, \mathcal{G}^*(y)\rangle_1$.

The Gâteaux derivative of J_0 is

$$\delta J_0 = \frac{d}{d\eta} J_0\Big(v_* + \eta\delta v, \bar{x}_{v_*+\eta\delta v}\Big)\Big|_{\eta=0} = \Big\langle R_0 v_*, \delta v\Big\rangle_{L_2} + (\bar{x}_{v_*}(T) - p_2)' M_0 \delta\bar{x}(T). \tag{5.26}$$

We have

$$\frac{d}{dt}(\delta y' P) = \delta v' B_0' P - \delta y' \mathcal{L}_1^*(W)(t) \tag{5.27}$$

$$\frac{d}{dt}(\delta\bar{x}' W) = \mathcal{L}_1(\delta y)(t)' W + \mathcal{L}_2(\delta\bar{x})(t)' W - \delta\bar{x}' \mathcal{L}_2^*(W)(t). \tag{5.28}$$

By integrating (5.27) from 0 to T, we get $0 = \Big\langle B_0' P, \delta v\Big\rangle_{L_2} - \Big\langle \mathcal{L}_1^*(W), \delta y\Big\rangle_{L_2}$. Similarly, by integrating (5.28), we obtain that

$$\begin{aligned}\delta\bar{x}(T)' M(\bar{x}_{v_*}(T) - p_2) &= \Big\langle \mathcal{L}_1(\delta y), W\Big\rangle_{L_2} + \Big\langle \mathcal{L}_2(\delta\bar{x}), W\Big\rangle_{L_2} - \Big\langle \delta\bar{x}, \mathcal{L}_2^*(W)\Big\rangle_{L_2} \\ &= \Big\langle \mathcal{L}_1^*(W), \delta y\Big\rangle_{L_2}.\end{aligned} \tag{5.29}$$

Therefore, $\delta J_0 = \Big\langle B_0' P, \delta v\Big\rangle_{L_2} + \Big\langle R_0 v_*, \delta v\Big\rangle_{L_2}$. By optimality, $\delta J_0 = 0$, $\forall \delta v \in L_2([0,T], \mathbb{R}^{m_1})$. Hence, $v_* = -R_0^{-1} B_0' P$. □

Theorem 18 states that the advertiser can act optimally. But, it doesn't give any indication on how to compute an optimal investment strategy v_*. Theorem 19, however, provides a formula (5.18) for v_*. As a result, the computation of v_* requires solving the state equations (5.2) and (5.19), which are coupled with the co-state equations (5.15)-(5.16) through the optimal control law (5.18). In Section 5.3, we study a special case where the optimal strategies v_* can be computed explicitly.

Before moving to the next section, we give a sufficient condition for the existence and uniqueness of the solutions to (5.15)-(5.16) to hold. This condition is required to apply the results of Theorem 19 later. Given the function W, equation (5.15) is a linear ODE which has a unique solution. Thus, it is sufficient to study the second equation (5.16). We define the matrix

$$\Sigma = \alpha \int_0^T \int_\sigma^T R_1(T,\sigma) B R^{-1} B' R_1(T,\sigma)' R_1(\tau, T)' H(\tau) \mathrm{d}\tau \mathrm{d}\sigma. \tag{5.30}$$

Assumption 15. *Either ξ^* is equal to zero or $1/\xi^*$ is not an eigenvalue of Σ, where ξ^* is defined in* (5.17).

Assumption 15 can be satisfied, for example, in the following two cases:

1. If the initial spread of the consumers is sufficient ($d\bar{F}/ds$ is low enough).
2. If $d\bar{F}/ds$ is bounded, and T is small enough.

In fact, $\xi^*\Sigma$ is in both cases negligible with respect to I_n. Hence, $1/\xi^*$ is not an eigenvalue of Σ.

Lemma 8. *Under Assumption 15,* (5.16) *has a unique solution.*

Proof. The idea of the proof is to replace the term

$$\int_0^T R_1(T,\sigma)BR^{-1}B'W(\sigma)\mathrm{d}\sigma \tag{5.31}$$

in the expression of $\mathcal{L}_2^*(W)$ by an assumed known constant K_1. Equation (5.16) is then a linear ODE parameterized by K_1, whose solution is a linear operator of K_1. By replacing this solution in the term (5.31) and by requiring that K_1 is equal to (5.31), one can show that the unique solution of (5.16) is $W(t) = R_1(T,t)'\Big(\alpha\xi^* \int_t^T R_1(\sigma,T)'H(\sigma)\mathrm{d}\sigma Y + M_0(\bar{x}_{v_*}(T) - p_2)\Big)$, where Y is the unique solution (under Assumption 15) of the following linear algebraic equation:

$$(I_n - \xi^*\Sigma)Y = \int_0^T R_1(T,\sigma)BR^{-1}B'R_1(T,\sigma)'\mathrm{d}\sigma M_0(\bar{x}_{v_*}(T) - p_2). \tag{5.32}$$

This proves the result. □

5.2.3 Approximate Stackelberg solution

We develop in the previous sections mean-field based Stackelberg strategies for a continuum of consumers. In this section, we show that these strategies, when applied to a finite population, constitute an approximate Stackelberg solution. We adopt the following notations. Given an advertiser's strategy v, we denote by y_v the advertiser's corresponding state, $\bar{x}_v$ the consumers' sustainable mean trajectory (5.19), u_i^v the mean-field based strategy (5.11) when applied by consumer i, i.e. $u_i^v = u_*(t, x, \bar{x}_v, x_i(0), y_v)$, x_i^v the corresponding consumer i's state, and $\bar{x}_v^N = 1/N \sum_{i=1}^N x_i^v$.

Theorem 20 (Approximate Stackelberg solution)**.** *Under Assumptions 12, 13 and 14, we* *have for all* $1 \leq i \leq N$*,*

$$J_0\left(v_*, \bar{x}^N_{v_*}\right) - \epsilon_N \leq \inf_{v \in L_2([0,T],\mathbb{R}^{m_1})} J_0\left(v, \bar{x}^N_v\right) \leq J_0\left(v_*, \bar{x}^N_{v_*}\right), \tag{5.33}$$

$$J_i\left(u_i^{v_*}, \bar{x}^N_{v_*}, v_*\right) - \epsilon_N \leq \inf_{u_i \in \mathcal{U}} J_i\left(u_i, \bar{x}^N_{-i}, v_*\right) \leq J_i\left(u_i^{v_*}, \bar{x}^N_{v_*}, v_*\right), \tag{5.34}$$

where $\epsilon_N = O(1/\sqrt{N})$*,* v_* *is an optimal control law of* (5.12)*, and* $\bar{x}^N_{-i} = \frac{1}{N}\sum_{j=1,j\neq i}^N x_j^{v_*} + \frac{1}{N}x_i$*, with* x_i *the state of consumer* i *under its control input* u_i*.*

Proof. Inequalities (5.34) state that $(u_i^{v_*}, u_{-i}^{v_*})$ constitutes an ϵ_N-Nash equilibrium with respect to the costs (5.3). We refer the reader to Theorem 6 for a proof of this result. It remains to show the first inequality. Let $v \in L_2([0,T],\mathbb{R}^{m_1})$, such that $J_0(v, \bar{x}^N_v) \leq J_0(v_*, \bar{x}^N_{v_*})$. We show in the following the there exists $\epsilon_N = O(1/\sqrt{N})$ that doesn't depend on v, and such that $J_0(v, \bar{x}^N_v) \geq J_0(v_*, \bar{x}^N_{v_*}) - \epsilon_N$, which implies (5.33). We have,

$$\begin{aligned} J_0(v, \bar{x}^N_v) &= J_0(v, \bar{x}_v) + \mathbb{E}\|\bar{x}_v(T) - \bar{x}^N_v(T)\|^2_{M_0} + \mathbb{E}(\bar{x}^N_v(T) - \bar{x}_v(T))' M_0(\bar{x}_v(T) - p_2) \\ &\geq J_0(v_*, \bar{x}_{v_*}) + \mathbb{E}\|\bar{x}_v(T) - \bar{x}^N_v(T)\|^2_{M_0} + \mathbb{E}(\bar{x}^N_v(T) - \bar{x}_v(T))' M_0(\bar{x}_v(T) - p_2) \\ &= J_0(v_*, \bar{x}^N_{v_*}) - \epsilon_N, \end{aligned} \tag{5.35}$$

where the inequality follows from the fact that v_* is an optimal control law of (5.12), and

$$\begin{aligned} \epsilon_N = &-\mathbb{E}\|\bar{x}_v(T) - \bar{x}^N_v(T)\|^2_{M_0} - \mathbb{E}(\bar{x}^N_v(T) - \bar{x}_v(T))' M_0(\bar{x}_v(T) - p_2) \\ &- \mathbb{E}\|\bar{x}_{v_*}(T) - \bar{x}^N_{v_*}(T)\|^2_{M_0} + \mathbb{E}(\bar{x}^N_{v_*}(T) - \bar{x}_{v_*}(T))' M_0(\bar{x}^N_{v_*}(T) - p_2). \end{aligned} \tag{5.36}$$

Since $\bar{x}_v$ is a sustainable mean trajectory, then $\mathbb{E}x_i^v = \bar{x}_v$, for $1 \leq i \leq N$. Therefore, $\mathbb{E}(\bar{x}^N_v(T) - \bar{x}_v(T))' M_0(\bar{x}_v(T) - p_2) = 0$. The consumers' initial conditions are i.i.d., which implies that $\mathbb{E}\|\bar{x}_v(T) - \bar{x}^N_v(T)\|^2_{M_0} = \frac{1}{N}\mathbb{E}\|\mathbb{E}x_i^v(T) - x_i^v(T)\|^2_{M_0}$. Similarly, $\mathbb{E}\|\bar{x}_{v_*}(T) - \bar{x}^N_{v_*}(T)\|^2_{M_0} = \frac{1}{N}\mathbb{E}\|\mathbb{E}x_i^{v_*}(T) - x_i^{v_*}(T)\|^2_{M_0}$. Furthermore, Cauchy-Schwarz inequality implies,

$$\begin{aligned} &\mathbb{E}(\bar{x}^N_{v_*}(T) - \bar{x}_{v_*}(T))' M_0(\bar{x}^N_{v_*}(T) - p_2) \\ &\leq \|M_0\| \left(\mathbb{E}\|\bar{x}^N_{v_*}(T) - \bar{x}_{v_*}(T)\|^2\right)^{1/2} \left(\mathbb{E}\|\bar{x}^N_{v_*}(T) - p_2\|^2\right)^{1/2} \\ &= \|M_0\| \left(\frac{1}{N}\mathbb{E}\|x_i^{v_*}(T) - \mathbb{E}x_i^{v_*}(T)\|^2\right)^{1/2} \left(\mathbb{E}\|\bar{x}^N_{v_*}(T) - p_2\|^2\right)^{1/2} = \frac{C_0}{\sqrt{N}}. \end{aligned} \tag{5.37}$$

C_0 does not depend on v. We have $J_0(v, \bar{x}^N_v) \leq J_0(v_*, \bar{x}^N_{v_*})$. Hence, $\int_0^T \|v\|^2 dt \leq C_1$, where C_1 does not depend on v. Noting that y and $\bar{x}_v$ satisfy (5.2) and (5.19), the boundedness of v implies that $\int_0^T \{\|\bar{x}_v\|^2 + \|y_v\|^2\} dt \leq C_2$, where C_2 does not depend on v. Finally, by implementing (5.11) in the dynamics of consumer i, and noting the boundedness of $\bar{x}_v$ and y_v, one can show that $\mathbb{E}\|x_i^v(T)\|^2 \leq C_3\mathbb{E}\|x_i(0)\|^2 + C_4$, where C_3 and C_4 do not depend on v. This implies that $\epsilon_N \leq C_5/\sqrt{N}$, for some C_5 that does not depend on v, and proves the result. □

Given an advertiser's strategy v, rational consumers seek an exact Nash equilibrium. But, as we saw in Section 2.2, an exact solution may not exist. Even if it does, it requires that each consumer know the exact a priori opinions of the other consumers, which is not realistic in a large population. Thus, the mean field based strategies are more suitable to describe the way the consumers make their choices. For these strategies are more realistic to implement and compute, and are sufficiently robust ($\epsilon-$Nash) in face of selfish deviant behaviors. Hence, we assume that the consumers make their choices of alternatives by implementing the mean-field based strategies. Furthermore, we assume that the advertiser knows that the consumers solve their game via the MFG methodology. Following these assumptions, a rational advertiser minimizes $J_0(v, \bar{x}_v^N)$ to find his/her optimal investment strategy, where $\bar{x}_v^N$ is the average state of the N consumers when they apply their mean field strategies. But, this requires that he/she observe the states of all the consumers, which is not realistic in a large population. Here again, the MFG provides a simpler alternative, the mean-field based strategy v_*, which requires that the advertiser know only the initial probability distribution of the consumers. Theorem 20 guaranties that the loss of optimality induced by v_* is negligible in a large population.

5.3 Case of uniform initial distribution

Because ξ^* defined in (5.17) is a nonlinear functional of $\bar{x}_{v_*}$ and y_{v_*}, solving (5.2)-(5.19)-(5.15)-(5.16) is not easy. Note however that $d\bar{F}/ds$ is the probability density function of $\beta' x(0) - \delta$ (see the definition of $\bar{F}$ in Assumption 14), so one can hope to compute an explicit solution when this random variable is uniformly distributed, for example. Indeed, in this case the probability density function is piecewise constant. Hence, equations (5.2)-(5.19)-(5.15)-(5.16) can be written as a pair of forward-backward linear ODE's (5.40). These equations are coupled in the boundary condition $K_{\lambda_{v_*}}$, through λ_{v_*} the probability that a generic consumer is initially in $D(\bar{x}_{v_*}, y_{v_*})$. Hence, we use similar techniques to those used in Lemma 5 and Theorem 11 to provide an explicit solution of (5.2)-(5.19)-(5.15)-(5.16), see Theorem 22 below. This solution encapsulates the advertiser's optimal investment strategy v_*, and state y_{v_*}, the consumers' sustainable mean trajectory $\bar{x}_{v_*}$, as well as the fraction of consumers that go towards p_1 under a social and advertising effects.

So we assume in this section that the consumers are "uniformly" distributed in the following sense: the random variable $\beta' x(0)$ has a uniform distribution on $[a-c/2, a+c/2]$, where $a \in \mathbb{R}$ and $c > 0$. It should be noted that Assumption 13 is satisfied for this initial distribution. The function $\bar{F}$ defined in Assumption 14 is piecewise differentiable. Therefore, we need an alternative to Assumption 14, under which the uniqueness of the consumers' Nash equilibria holds. Moreover, in order to apply the variational methods of Subsection 5.2.2, we require $\bar{F}$ to stay in a differentiable domain for all the advertiser's strategies, which is the case when the consumers are spread enough (see Lemma 9 below).

Assumption 16. *We assume that* $c > \alpha \left|\Delta\left(R_2(p_1 - p_2)\right)\right|$.

Under Assumption 16, the consumers' limiting game admits a unique Nash equilibrium for each advertiser's strategy v by virtue of Theorem 17.

Theorem 21 (Existence of an optimal investment strategy)**.** *Under Assumptions 12 and 16, the advertiser's optimal control problem* (5.12) *has an optimal control law* v_*.

Proof. Let v and v' two investment strategies in $L_2([0,T], \mathbb{R}^{m_1})$, and denote y and y' the corresponding advertiser's states and λ and λ' the corresponding fixed points of $\lambda \mapsto F(\lambda, y)$ and $\lambda' \mapsto F(\lambda', y')$, where F is defined in (5.8). We have

$$\begin{aligned} |\lambda - \lambda'| &= |F(\lambda, y) - F(\lambda', y')| \leq |F(\lambda, y) - F(\lambda', y)| + |F(\lambda', y) - F(\lambda', y')| \\ &\leq \frac{\alpha}{c}\left|\Delta\Big(R_2(t)(p_1 - p_2)\Big)\right||\lambda - \lambda'| + |F(\lambda', y) - F(\lambda', y')|. \end{aligned} \tag{5.38}$$

The rest of the proof is similar to that of Theorem 18. □

Lemma 9. *Under Assumptions 12 and 16, there exists* $c_0 > 0$ *independent of* v *such that for all* $c > c_0$*, there exists a unique consumers' Nash equilibrium corresponding to* $\lambda \in (0,1)$.

Proof. The uniqueness follows from Assumption 16. Let $v \in L_2([0,T])$. The path $\bar{x}_v$ defined in (5.10) is uniformly bounded with c (with respect to the L_2 norm). Therefore, the optimal cost $J_0(v_*, \bar{x}_{v_*}) \leq J_0(v, \bar{x}_v)$ of the advertiser's optimal control problem (5.12) is uniformly bounded with c. Hence, the optimal control laws v_* and corresponding optimal states y_{v_*} are uniformly bounded with c. Consequently, the term $\Delta\left(K(p_2) y_{v_*} + \alpha \bar{x}^\lambda\right)$, where $\bar{x}^\lambda$ is defined by (5.9), is uniformly bounded with c by a positive constant C_1. This means that $-C_1 \leq \Delta\left(K(p_2) y_{v_*} + \alpha \bar{x}^\lambda\right) \leq C_1$. Hence, $\bar{F}(-C_1) \leq F(\lambda, y_{v_*}) \leq \bar{F}(C_1)$. If we choose $-C_1 > a - c/2$ and $C_1 < a + c/2$, that is, $c > \max(2(a + C_1), 2(-a + C_1)) := c_0$, the map F defined by (5.8) takes its values in $(0,1)$. Therefore, $\lambda \mapsto F(\lambda, y)$ has a unique fixed point $\lambda \in (0,1)$. This proves the result. □

For the rest of the analysis, we assume that $c > c_0$. In this case, the unique fixed point λ corresponding to an advertiser's optimal policy v_* is in $(0,1)$. Since F is differentiable in $(0,1)$, one can use techniques similar to those used in Theorem 19 to show that v_* satisfies (5.18) (provided that the Assumptions 12 and 16 are satisfied, and $1/c$ is not an eigenvalue of Σ defined in (5.30)).

The Stackelberg competition for a continuum of consumers is fully described by the optimal state y_{v_*} which satisfies (5.2) for the control input (5.18), and the corresponding consumers' sustainable mean trajectory $\bar{x}_{v_*}$, which satisfies (5.19). To compute an explicit solution of the game, we rewrite the equations of y_{v_*} and $\bar{x}_{v_*}$ as two coupled forward-backward ODE's. Let us start by equation (5.19). Using similar techniques to those used in Lemma 1, one can derive the following equivalent representation of (5.19),

$$\begin{aligned} \frac{d}{dt}\bar{x}_{v_*} &= A\bar{x}_{v_*} - BR^{-1}B'\bar{q}_{v_*} \\ -\frac{d}{dt}\bar{q}_{v_*} &= A'\bar{q}_{v_*} + (1-\alpha)Q\bar{x}_{v_*} - QK(p_2)y_{v_*}, \end{aligned} \tag{5.39}$$

with $\bar{x}_{v_*}(0) = \mathbb{E}x(0)$, $\bar{q}_{v_*}(T) = M(\bar{x}_{v_*}(T) - \lambda_{v_*}p_1 + (1-\lambda_{v_*})p_2)$, and λ_{v_*} is the unique fixed point of $\lambda \mapsto F(\lambda, y_{v_*})$. Next, we define the states $h = (\bar{x}_{v_*}, y_{v_*}, W_1)$ and $d = (\bar{q}_{v_*}, P, W, W_2)$. Here $W_1(t) := \int_0^t R_1(T,\sigma)BR^{-1}B'W(\sigma)\mathrm{d}\sigma$ and $W_2(t) := \int_t^T R_1(T,\sigma)BR^{-1}B'W(\sigma)\mathrm{d}\sigma$ are the forward and backward propagating parts of the integral $\int_0^T R_1(T,\sigma)BR^{-1}B'W(\sigma)\mathrm{d}\sigma$ that appears in (5.15)-(5.16). The pair (h,d) satisfies

$$\begin{aligned} \frac{d}{dt}h &= K_1(t)h + K_2(t)d \\ \frac{d}{dt}d &= K_3(t)h + K_4(t)d \end{aligned} \tag{5.40}$$

with $h(0) = (\mathbb{E}x(0), y(0), 0)$ and $d(T) = K_5 h(T) + K_{\lambda_{v_*}}$, where $K_1(t) = \mathrm{diag}(A, A_0, 0)$,

$$K_2(t) = \begin{bmatrix} k_1 & 0 & 0 & 0 \\ 0 & k_2 & 0 & 0 \\ 0 & 0 & k_3 & 0 \end{bmatrix} \qquad K_5 = \begin{bmatrix} M & 0 & 0 \\ 0 & 0 & 0 \\ M_0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{5.41}$$

$$K_3(t) = \begin{bmatrix} -(1-\alpha)Q & QK(p_2) & 0 \\ 0 & 0 & k_4 \\ 0 & 0 & -\alpha H(t)/c \\ 0 & 0 & 0 \end{bmatrix}, \quad K_4(t) = \begin{bmatrix} -A' & 0 & 0 & 0 \\ 0 & -A_0' & 0 & -K(p_2)'H(t)/c \\ 0 & 0 & k_5 & -\alpha H(t)/c \\ 0 & 0 & -k_3 & 0 \end{bmatrix}, \tag{5.42}$$

$K_{\lambda_{v_*}} = -\Big(M(\lambda_{v_*}p_1 + (1-\lambda_{v_*})p_2), 0, M_0 p_2, 0\Big)$, $k_1 = -BR^{-1}B'$, $k_2 = -B_0R_0^{-1}B_0'$, $k_3 = R_1(T,t)BR^{-1}B'$, $k_4 = -K(p_2)'\Big(QR_1(t,T) + \frac{1}{c}H(t)\Big)$ and $k_5 = -(A - BR^{-1}B'\gamma)'$.

The equation system (5.40) consists of two coupled nonlinear forward-backward ODE's. The final condition $d(T)$ depends through λ_{v_*} non-linearly on the path $(\bar{x}_{v_*}(\sigma), y_{v_*}(\sigma))$, $\sigma \in [0,T]$. As in the previous chapters, we need the following assumption to decouple and solve these equations.

Assumption 17. *The following generalized Riccati equation has a unique solution*

$$\frac{d}{dt}\Pi = K_4\Pi - \Pi K_1 - \Pi K_2 \Pi + K_3, \qquad \Pi(T) = K_5. \tag{5.43}$$

Using similar techniques to those used in Lemma 5 and Theorem 11, one can show that under Assumption 17, (h,d) is a solution of (5.40) if and only if

$$h(t) = \Phi_1(t,0)h(0) + R_u(t)K_\lambda =: \left(\bar{x}^\lambda, y^\lambda, W_1^\lambda\right), \tag{5.44}$$

$$d(t) = \Pi(t)h(t) + \Phi_2(t,T)K_\lambda =: \left(\bar{q}^\lambda, P^\lambda, W^\lambda, W_2^\lambda\right), \tag{5.45}$$

where Φ_1 and Φ_2 are the state transition matrices of of $K_1 + K_2\Pi$ and $K_4 - \Pi K_2$, $R_u(t) = \int_0^t \Phi_1(t,\sigma)K_2(\sigma)\Phi_2(\sigma,T)\mathrm{d}\sigma$ and λ is a fixed point of the following final dimensional map,

$$F_u(\lambda) = \bar{F} \circ \Delta\left(\alpha\bar{x}^\lambda + K(p_2)y^\lambda\right). \tag{5.46}$$

Theorem 22 (Explicit Stackelberg solution)**.** *Under Assumptions 12, 16, and 17, the Stackelberg competition (for a continuum of consumers) has a unique solution $(v_*, \bar{x}_{v_*})$, where $v_* = -R_0^{-1}B_0'P^{\lambda_*}$ and $\bar{x}_{v_*} = \bar{x}^{\lambda_*}$, with P^{λ_*} and $\bar{x}^{\lambda_*}$ defined in* (5.44)-(5.45) *for the unique fixed point λ_* of F_u.*

Proof. By Theorem 21 we know that there exists an optimal investment strategy v_*. This strategy and the corresponding sustainable mean trajectory satisfy $v_* = R_0^{-1}B_0'P^\lambda$ and $\bar{x}_{v_*} = \bar{x}^\lambda$, where P^λ and $\bar{x}^\lambda$ are defined in (5.44)-(5.45) for a fixed point λ of F_u. It remains to show that F_u has a unique fixed point λ_*. Let λ and λ' be two distinct fixed points of F_u.

Then, λ and λ' are respectively the fixed points of $s \mapsto F(s, y^{\lambda})$ and $s \mapsto F(s, y^{\lambda'})$, where F is defined in (5.8). Following Lemma 9, λ and λ' belong to $(0, 1)$. But, $\beta' x(0)$ has a uniform distribution, which implies that F_u has a shape similar to that of the cumulative distribution function of a uniform distribution. Thus, all the real numbers in the interval $[0, 1]$ are fixed points of F_u, and as a result of F. This leads to a contradiction, and shows that F_u has a unique fixed point. □

Theorem 22 states that for a population of consumers with sufficiently diverse a priori opinions uniformly distributed in the direction β, the unique optimal investment strategy is $v_* = -R_0^{-1} B_0' P^{\lambda_*}$. This strategy convinces a fraction of the consumers to choose the advertised alternative. This fraction is equal to $1 - \lambda_*$, where λ_* is the unique fixed point of F_u. One can apply the bisection method to find λ_*. Once λ_* is computed, the agents can compute the vectors (h, d), given by (5.44)-(5.45). The advertiser can then implement its optimal strategy (5.18), where P^{λ_*} is the second component of d. Subsequently, the consumers can predict their limiting macroscopic behavior, the first component of d, and implement their optimal strategies (5.11).

5.4 Simulation results

To illustrate the collective choice mechanism in the presence of social and advertising effects, we consider a group of 6000 consumers that have opinion states initially uniformly distributed between -25 and 5. The consumers are choosing between $p_1 = -20$ and $p_2 = 20$. The social effect is represented by $\alpha \bar{x}$, where $\alpha = 0.5$. We consider two scenarios. In the first one, the consumers make their choices in the absence of an advertising effect ($K(p_2) = 0$), while in the second scenario, an advertiser advertises for p_2. The advertising effect is modeled in the cost by $K(p_2)y = p_2 y$, where y is the (influence) state of the advertiser. We set $T = 3$, $A = 0.5$, $B = 0.1$, $A_0 = -0.1$, $B_0 = 0.1$, $y(0) = 0$, $Q = 10$, $R = R_0 = 10$, and $M = M_0 = 2000$. In the absence of an advertising effect, $\lambda = 0.84$ is the unique fixed point of F_u defined in (5.46). Accordingly, 84% of the consumers choose p_1 (Fig. 5.1). On the other hand, the presence of advertising for alternative p_2 increases from 16% to 87% the fraction of consumers that go towards p_2 (Fig. 5.1).

5.5 Conclusion

We introduce in this chapter a dynamic collective choice discrete model with an advertiser. In this model, a large group of consumers choose between two alternatives while influenced by their average and an advertising effect. The latter is generated by an advertiser aiming at

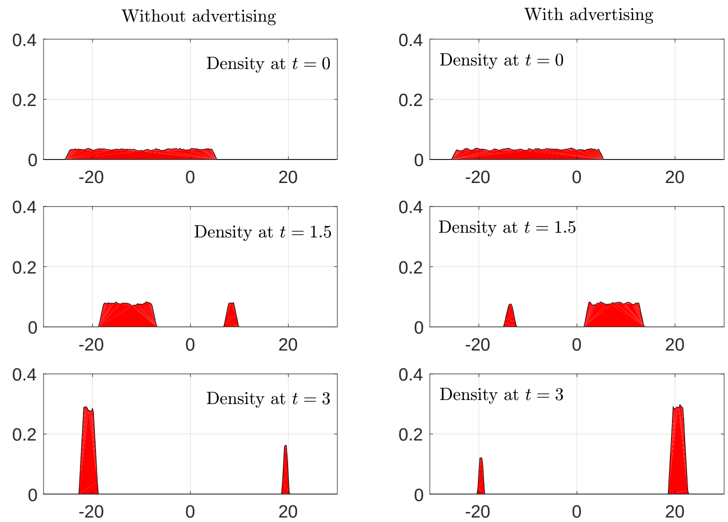

Figure 5.1 Evolution of the consumers' density in the absence and presence of an advertising effect.

convincing the population of consumers to choose p_2. We develop approximate Stackelberg solutions, which are decentralized and simple to implement. In case the consumers are initially uniformly distributed in the direction β, we provide an explicit form of the unique solution, and characterize it by the fraction of consumers that choose p_1. This fraction is the unique fixed point of a well defined finite dimensional map.

CHAPTER 6 CONCLUSION

6.1 Summary

In this thesis, we study a class of dynamic collective discrete choice problems. We formulate the non-cooperative degenerate problem as a dynamic non-cooperative game in Chapter 2. We propose a naïve approach to compute an exact Nash equilibrium, if it exists. This approach becomes computationally intractable and requires a significant amount of communication between the agents when their number is large. Alternatively, we develop via the MFG methodology a set of strategies that are simple to compute and implement. This simplification occurs at the expenses of weakening the robustness of the equilibrium. Indeed, the mean-field based strategies constitute an approximate Nash equilibrium. The robustness of these equilibria improves, however, as the number of agents increases sufficiently.

In Chapter 3, we consider the non-cooperative non-degenerate collective discrete choice problem. We solve the game via the MFG methodology. This includes solving a new class of stochastic optimal control problems, that we call min-LQG. We derive an explicit form of the optimal solution, and interpret it at each instant as a static discrete choice problem, where the cost of choosing an alternative includes an additional term to penalize premature myopic decisions. In the degenerate case, the agents make their choices of alternatives before starting to move. In the non-degenerate case, however, the agents can no longer commit to a choice from the beginning.

We formulate the cooperative collective discrete choice problem in Chapter 4. Here again, the MFG is invoked to simplify the exact solutions. In fact, Section 4.2 shows that the naïve approach to computing a social optima is computationally intractable in a large population. Section 4.3 develops a set of mean-field based strategies that are simpler to compute and implement than the exact solution. Moreover, they converge to a social optimum as the number of agents increases to infinity. We provide a numerical example showing that the cooperative strategies have the advantage of distributing much more evenly the agents between the alternatives than the non-cooperative strategies.

The non-cooperative and cooperative problems share a common feature. Their solutions (for a continuum of agents) are mapped one-to-one to the fixed points of some finite dimensional maps. The latter are the potential probability distributions of the agents' choices over the alternatives. This mapping has two important consequences. First, one can anticipate the way the population splits between the alternatives. Second, it reduces the infinite dimensional problems of finding a solution to the DCDC problems, to computing a fixed point for a finite dimensional map. Accordingly, we suggest simple numerical schemes to compute the infinite population Nash equilibria and social optima.

In Chapter 5, we extend the degenerate non-cooperative collective discrete choice problems by including an advertiser in the game. We derive conditions for an approximate Stackelberg solution to exist. When the consumers are initially uniformly distributed, we compute an explicit form of the approximate solution and characterize it by a scalar describing the way the population of consumers splits between the alternatives under the social and advertising effects.

6.2 Future directions

In the concluding section of Chapter 3, we discuss some future directions to generalize the results of the min-LQG optimal control problem. It is of interest also for future work to consider an infinite compact set of alternatives. Some interesting applications are modeling the collective tissue invasion phenomenon exhibited by some biological societies, such as cancer cells (Deisboeck and Couzin, 2009). The question that arises in such situations is: how do the cancer cells migrate from their initial positions and distribute themselves on the surface of the tissue? From a macroscopic point of view, this problem can be viewed as a dynamic optimal transport problem (Villani, 2008), where a mass with an initial distribution on some set is carried with least effort and distributed on some compact destination set according to a predefined final distribution. Indeed, we expect that, by adjusting the weights attributed to the points in the destination set, we could distribute the mass on the destination set according to a predefined final distribution. In this case, the min-LQG problem with infinite compact set of alternatives could describe the microscopic processes underlying some optimal transport problems. Even the current formulation of a finite number of alternatives could be interpreted as an optimal transport problem if we approximate the final predefined distribution by a finite number of Dirac functions.

We assume in our models that each agent interacts with all the other agents through the average state of the population. In real situations, for example elections, this interaction is local. In the future, it would be interesting to study discrete choice models with local interactions. Another future direction is to extend the current advertising model to an oligopolistic one, where multiple advertisers are competing to increase the number of their consumers.

So far, we assume in our discrete choice models that the agents know the probability distributions of their initial states and parameters. In practice, they have to learn these distributions. An interesting future direction is to incorporate a learning process (Bertsekas and Tsitsiklis, 1989) in the game, a consensus-like algorithm, for example, where the agents have to communicate their states and parameters periodically. This learning process creates in the non-cooperative case some challenging problems. First, we need to understand how this process affects the robustness of the equilibrium. Does it incite the agents to deviate from the equilibrium? Second, how does this process affect the distribution of the agents' choices over the alternatives?

In some situations, beside some personal attributes, the social effect, and the distance to the alternatives, the quality of the alternatives influences also the individual choices. For example, the quality of candidates (or their electoral programs) in elections is a major factor shaping the final choices. In a rescue mission, for instance, a collection of robots identify multiple potential sites to visit. But, they are more attracted to those that contain more victims. Usually, the quality of the alternatives are not known a priori, and are learned along the path to make a final choice. Future work could try to extend our current formulation to capture the quality of the alternatives.

Finally, throughout the thesis, we reduce the infinite dimensional problem of finding a solution to our game to computing a fixed point of a finite dimensional map. A challenging future direction is to understand the ingredients that made this reduction possible and try to generalize it to a larger class of nonlinear individual costs and dynamics. An important advantage of this approach is to simplify the numerical schemes (Achdou and Capuzzo-Dolcetta, 2010) to solve the mean field equations.

REFERENCES

D. Acemoğlu, G. Como, F. Fagnani, and A. Ozdaglar, "Opinion fluctuations and disagreement in social networks," *Mathematics of Operations Research*, vol. 38, no. 1, pp. 1–27, 2013.

Y. Achdou and I. Capuzzo-Dolcetta, "Mean field games: Numerical methods," *SIAM Journal on Numerical Analysis*, vol. 48, no. 3, pp. 1136–1162, 2010.

B. D. Anderson and J. B. Moore, *Optimal control: linear quadratic methods*. Dover, 2007.

W. E. Baker, "Floor trading and crowd dynamics," *The social dynamics of financial markets*, pp. 107–128, 1984.

T. Başar and G. J. Olsder, *Dynamic noncooperative game theory*. SIAM, 1998.

A. Bensoussan, M. Chau, and P. Yam, "Mean field stackelberg games: Aggregation of delayed instructions," *SIAM Journal on Control and Optimization*, vol. 53, no. 4, pp. 2237–2266, 2015.

——, "Mean field games with a dominating player," *arXiv preprint arXiv:1404.4148*, 2014.

D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: numerical methods*. Prentice hall Englewood Cliffs, NJ, 1989, vol. 23.

C. Bhat and J. Guo, "A mixed spatially correlated logit model: formulation and application to residential choice modeling," *Transportation Research*, vol. 38, pp. 147–168, 2004.

W. Brock and S. Durlauf, "Discrete choice with social interactions," *Review of Economic Studies*, pp. 147–168, 2001.

C. G. Broyden, "A class of methods for solving nonlinear simultaneous equations," *Mathematics of computation*, vol. 19, no. 92, pp. 577–593, 1965.

S. Camazine, P. K. Visscher, J. Finley, and R. S. Vetter, "House-hunting by honey bee swarms: collective decisions and individual behaviors," *Insectes Sociaux*, vol. 46, no. 4, pp. 348–360, November 1999.

P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions, "The master equation and the convergence problem in mean field games," *arXiv preprint arXiv:1509.02505*, 2015.

R. Carmona and F. Delarue, “Probabilistic analysis of mean-field games,” *SIAM Journal on Control and Optimization*, vol. 51, no. 4, pp. 2705–2734, 2013.

F. Clarke, *Functional analysis, calculus of variations and optimal control.* Springer Science & Business Media, 2013, vol. 264.

J. B. Conway, *A course in functional analysis.* Springer Science & Business Media, 2013, vol. 96.

M. H. DeGroot, “Reaching a consensus,” *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.

T. S. Deisboeck and I. D. Couzin, “Collective behavior in cancer cell populations,” *Bioessays*, vol. 31, no. 2, pp. 190–197, 2009.

R. Durrett, *Probability: theory and examples.* Cambridge university press, 2010.

G. M. Erickson, “Differential game models of advertising competition,” *European Journal of Operational Research*, vol. 83, no. 3, pp. 431–438, 1995.

L. Evans, *Partial Differential Equations*, ser. Graduate studies in mathematics. American Mathematical Society, 1998.

W. H. Fleming and H. M. Soner, *Controlled Markov processes and viscosity solutions.* Springer Science & Business Media, 2006, vol. 25.

G. Freiling, “A survey of nonsymmetric Riccati equations,” *Linear algebra and its applications*, vol. 351, pp. 243–270, 2002.

N. E. Friedkin and E. C. Johnsen, “Social influence and opinions,” *Journal of Mathematical Sociology*, vol. 15, no. 3-4, pp. 193–206, 1990.

D. Gomes, E. Pimentel, and V. Voskanyan, *Regularity Theory for Mean-Field Game Systems*, ser. SpringerBriefs in Mathematics. Springer, 2016.

R. Gray, A. Franci, V. Srivastava, and N. E. Leonard, “Multi-agent decision-making dynamics inspired by honeybees,” *IEEE Transactions on Control of Network Systems*, 2018.

A. Halász, M. A. Hsieh, S. Berman, and V. Kumar, “Dynamic redistribution of a swarm of robots among multiple sites,” in *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on.* IEEE, 2007, pp. 2320–2325.

R. Hegselmann and U. Krause, "Opinion dynamics and bounded confidence models, analysis, and simulation," *Journal of artificial societies and social simulation*, vol. 5, no. 3, 2002.

M. A. Hsieh, Á. Halász, S. Berman, and V. Kumar, "Biologically inspired redistribution of a swarm of robots among multiple sites," *Swarm Intelligence*, vol. 2, no. 2, pp. 121–141, 2008.

M. Huang, P. E. Caines, and R. P. Malhamé, "Individual and mass behaviour in large population stochastic wireless power control problems: centralized and Nash equilibrium solutions," in *Proceedings of the 42nd IEEE Conference on Decision and Control*, 2003, pp. 98–103.

M. Huang, R. P. Malhamé, and P. E. Caines, "Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle," *Communications in Information & Systems*, vol. 6, no. 3, pp. 221–252, 2006.

M. Huang, P. E. Caines, and R. P. Malhamé, "Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized epsilon-Nash equilibria," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1560–1571, 2007.

——, "Social optima in mean field LQG control: centralized and decentralized strategies," *IEEE Transactions on Automatic Control*, vol. 57, no. 7, pp. 1736–1751, 2012.

M. Huang, "Large-population LQG games involving a major player: the Nash certainty equivalence principle," *SIAM Journal on Control and Optimization*, vol. 48, no. 5, pp. 3318–3353, 2010.

A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on automatic control*, vol. 48, no. 6, pp. 988–1001, 2003.

O. Kallenberg, *Foundations of modern probability.* Springer Science & Business Media, 2006.

I. Karatzas and S. Shreve, *Brownian motion and stochastic calculus.* Springer Science & Business Media, 2012, vol. 113.

F. Koppelman and V. Sathi, "Incorporating variance and covariance heterogeneity in the generalized nested logit model: an application to modeling long distance travel choice behavior," *Transportation Research*, vol. 39, pp. 825–853, 2005.

J. M. Lasry and P. L. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, pp. 229–260, 2007.

——, "Jeux à champ moyen. I–le cas stationnaire," *Comptes Rendus Mathématique*, vol. 343, no. 9, pp. 619–625, 2006.

——, "Jeux à champ moyen. II–horizon fini et contrôle optimal," *Comptes Rendus Mathématique*, vol. 343, no. 10, pp. 679–684, 2006.

K. Lehrer, "Social consensus and rational agnoiology," *Synthese*, vol. 31, no. 1, pp. 141–160, 1975.

T. Liggett, *Interacting particle systems.* Springer Science & Business Media, 2012, vol. 276.

G. B. Markus and P. E. Converse, "A dynamic simultaneous equation model of electoral choice," *American Political Science Review*, vol. 73, no. 04, pp. 1055–1070, 1979.

D. McFadden, "Conditional logit analysis of qualitative choice behavior," in *Frontiers in Econometrics*, P. Zarembka, Ed. New York: New York: Academic Press, 1974, ch. 4, pp. 105–142.

S. Merrill and B. Grofman, *A Unified Theory of Voting: Directional and Proximity Spatial Models.* Cambridge University Press, 1999.

R. Nakajima, "Measuring peer effects on youth smoking behaviour," *The Review of Economic Studies*, vol. 74, no. 3, pp. 897–935, 2007.

J. Nash, "Non-cooperative games," *Annals of mathematics*, pp. 286–295, 1951.

M. Nourian and P. E. Caines, "ϵ-nash mean field game theory for nonlinear stochastic dynamical systems with major and minor agents," *SIAM Journal on Control and Optimization*, vol. 51, no. 4, pp. 3302–3331, 2013.

L. Perko, *Differential equations and dynamical systems.* Springer Science & Business Media, 2013, vol. 7.

L. Pichler, A. Masud, and L. Bergman, "Numerical solution of the Fokker–Planck equation by finite difference and finite element methods - A comparative study," in *Computational Methods in Stochastic Dynamics.* Springer, 2013, pp. 69–85.

G. Pólya, *How to solve it.* Princeton university press, 1957.

W. Rudin, *Real and complex analysis.* Tata McGraw-Hill Education, 1987.

——, *Functional analysis.* McGraw-Hill, 1991.

W. J. Rugh, *Linear system theory.* prentice hall Upper Saddle River, NJ, 1996, vol. 2.

J. Rust, "Structural estimation of markov decision processes," *Handbook of econometrics*, vol. 4, pp. 3081–3143, 1994.

T. D. Seeley, S. Camazine, and J. Sneyd, "Collective decision-making in honey bees: how colonies choose among nectar sources," *Behavioral Ecology and Sociobiology*, vol. 28, pp. 277–290, 1991.

A. V. Skorokhod, "Stochastic differential equations depending on a parameter," *Theory of Probability & Its Applications*, vol. 25, no. 4, pp. 659–666, 1981.

D. W. Stroock and S. S. Varadhan, *Multidimensional diffussion processes.* Springer Science & Business Media, 1979, vol. 233.

D. Vail and M. Veloso, "Multi-robot dynamic role assignment and coordination through shared potential fields," *Multi-robot systems*, pp. 87–98, 2003.

C. Villani, *Optimal transport: old and new.* Springer Science & Business Media, 2008, vol. 338.

J. Von Neumann and O. Morgenstern, *Theory of games and economic behavior.* Princeton university press, 2007.

H. Von Stackelberg, *Marktform und Gleichgewicht*, ser. Die Handelsblatt-Bibliothek "Klassiker der Nationalökonomie". J. Springer, 1934.

J. Yong and X. Y. Zhou, *Stochastic controls: Hamiltonian systems and HJB equations.* Springer Science & Business Media, 1999, vol. 43.

S. Yüksel and T. Başar, *Stochastic networked control systems.* Birkhäuser Basel, 2013.

APPENDIX A PUBLICATIONS

1. Rabih Salhab, Roland P. Malhamé and Jerome Le Ny, “Collective Stochastic Discrete Choice Problems: A Min-LQG Game Formulation”, Submitted to the *IEEE Transactions on Automatic Control*, August 2017.

2. Rabih Salhab, Roland P. Malhamé and Jerome Le Ny, “A Dynamic Game Model of Collective Choice in Multi-Agent Systems”, *IEEE Transactions On Automatic Control*, In Press 2017.

3. Rabih Salhab, Roland P. Malhamé and Jerome Le Ny, “A Dynamic Collective Choice Model With An Advertiser", *Dynamic Games and Applications : Special Issue on Dynamic Games in Economics and Management Science*, In Press 2018.

4. Rabih Salhab, Jerome Le Ny and Roland P. Malhamé, “Dynamic Collective Choice: Social Optima”, *IEEE Transactions On Automatic Control*, In Press 2018.

5. Rabih Salhab, Jerome Le Ny and Roland P. Malhamé, “A Dynamic Ride-Sourcing Game with Many Drivers”, Accepted in the *55th Annual Allerton Conference on Communication, Control, and Computing,* August 2017.

6. Rabih Salhab, Roland P. Malhamé and Jerome Le Ny, “A Dynamic Collective Choice Model With An Advertiser”, *55th IEEE Conference on Decision and Control (CDC)*, December 2016, Las Vegas, USA.

7. Rabih Salhab, Roland P. Malhamé and Jerome Le Ny, “A Dynamic Game Model of Collective Choice in Multi-Agent Systems”, *54th IEEE Conference on Decision and Control (CDC)*, December 2015, Osaka, Japan.

8. Rabih Salhab, Roland P. Malhamé and Jerome Le Ny, “Consensus and Disagreement in Collective Homing Problems: A Mean Field Games Formulation”, *53rd IEEE Conference on Decision and Control (CDC)*, December 2014, Los Angeles, California, USA.