High-Performance Control of Mean-Field Teams in Leader-Follower Networks

Mohammad Mahdi Baharloo

A Thesis

in

The Department

of

Electrical and Computer Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Applied Science at

Concordia University

Montréal, Québec, Canada

March 2019

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Mohammad Mahdi Baharloo

Entitled: High-Performance Control of Mean-Field Teams in Leader-Follower Networks

and submitted in partial fulfillment of the requirements for the degree of

Master of Applied Science

Complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

______________________________ Chair
Dr. Rastko Selmic

______________________________ Examiner, External To the Program
Dr. Walter Lucia (CIISE)

______________________________ Examiner
Dr. Rastko Selmic

______________________________ Supervisor
Dr. Amir G. Aghdam

______________________________ Co-Supervisor
Dr.

Approved by: ______________________________
Dr. W. E. Lynch, Chair
Department of Electrical and Computer Engineering

____________20_____ ______________________________
Dr. Amir Asif, Dean
Gina Cody School of Engineering and Computer Science

Abstract

High-Performance Control of Mean-Field Teams in Leader-Follower Networks

Mohammad Mahdi Baharloo

In this thesis, a mean-field approach is used to find high-performance control strategies for multi-agent systems. The system consists of one leader and possibly many dynamically coupled followers, and all agents are affected by noise. The global objective of the multi-agent control system here is to achieve an agreement between the agents while minimizing coupled linear-quadratic cost functions for two cases: a disturbance-free system, and a system with disturbances. In the former case, the proposed solution under non-classical information structure is near-optimal, which converges to the optimal solution for a large number of followers. For the latter case, the problem is solved for three non-classical information structures, namely, mean-field sharing, partial mean-field sharing, and intermittent mean-field sharing. Using the minimax control technique, it is shown that the solution obtained for the first structure is a unique saddle-point strategy. On the other hand, it is proved that for the other two structures, the proposed solutions tend to the unique saddle-point strategy when the number of followers goes to infinity. The proposed strategies in both cases are linear, scalable and computationally efficient. The theoretical findings are verified by simulation results.

Acknowledgements

First and foremost, I would like to express my deep appreciation to my exceptional supervisor, Dr. Amir G. Aghdam for the opportunity he gave me to be a part of his research team. It was definitely a privilege to work with him and I am extremely grateful for his invaluable advice, support, and helpful guidance during my research. I have learned invaluable things from his weekly meetings.

I would like to thank Dr. Jalal Arabneydi for sharing his knowledge with me, his inputs were very crucial for the completion of this thesis. He assisted me a lot in reviewing my work and provided me with very helpful constructive criticism.

I would like to thank my beloved parents, brothers and sisters for their love and warm support.

I like to thank Mona, my love, for her presence and support in every moment of my life.

Lastly, I would like to thank Elina, my daughter, for the sacrifices she has made during my work.

List of Publications

[1] **M. Baharloo**, J. Arabneydi, and A. G. Aghdam, "Minimax Control in Linear Quadratic Mean-Field Teams for a Leader-Follower Network," submitted to *IEEE Control Systems Letters*.

[2] **M. Baharloo**, J. Arabneydi, and A. G. Aghdam, "Minimax Control in Linear Quadratic Mean-Field Teams for a Leader-Follower Network," submitted to *the 58th IEEE Conference on Decision and Control*.

[3] **M. Baharloo**, J. Arabneydi, and A. G. Aghdam, "Near-optimal control strategy in leader-follower networks: A case study for linear quadratic mean-field teams," in Proceedings of *the 57th IEEE Conference on Decision and Control, pp. 3288–3293, 2018.*

[4] J. Arabneydi, **M. M. Baharloo** and A. G. Aghdam, "Optimal distributed control for leader-follower networks: A scalable design," in Proceedings of *the 31st IEEE Canadian Conference on Electrical and Computer Engineering, pp. 1 - 4, 2018.*

Contribution of Authors

This thesis is written in a manuscript-based format. The presented papers are written by Mohammad Mahdi Baharloo and co-authored by Dr. Jalal Arabneydi and Dr. Amir G. Aghdam, all from the Department of Electrical and Computer Engineering, Concordia University. As the research supervisor, Dr. Aghdam provided guidance and direction on all aspects of the the work and gave constructive comments. Also, Dr. Arabneydi reviewed the papers and provided insightful comments and checked all simulations. All the research presented in this dissertation is performed by the author.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

There has been a growing interest recently in the control of networked systems, due to the application of this type of systems in emerging areas such as coordination of unmanned vehicles [1], environmental monitoring using sensor networks [2], and smart grids [3], to name only a few. In the control of multi-agent systems, every agent exchanges information with a subset of agents in order to properly coordinate its position and movement such that a global objective such as consensus is achieved. As a result, each agent requires some computational effort in order to determine its control action based on the information available to it. Different structures are proposed in the literature for information exchange between agents. An information pattern is said to be *classical* if all agents receive the same information and have perfect recall [49]; otherwise, it is called *non-classical*.

In developing coordination protocols for multi-agent systems, different performance measures such as minimum time or energy are sometimes considered in the literature along with some constraints on communication and computation resources. This type of limitations are often encountered in practical networks.

1.2 Leader-Follower Structure

The leader-follower structure is commonly used in the coordination control of multi-agent networks. In this type of system, each agent is either a leader or a follower, where the movement of the followers is dependent on the trajectory of the leader(s). A global objective that is of special interest is consensus, where the states of the agents are desired to converge to a common value [4–6]. If the communication graph representing the network is connected, then consensus can be reached, e.g., using a linear strategy. However, such a strategy suffers from the curse of dimensionality, in general. In other words, the strategy may not be computationally feasible when the number of agents increases. In addition, the required amount of communication between agents under this type of strategy is typically high. This can lead to some practical problems, specially given that the battery consumption of each node is closely related to the amount of its information exchange. Therefore, an optimal control strategy can be very important for the efficient resource management in the network. It is well-known that the optimal control strategy for a system with linear dynamics and a quadratic performance index under centralized information structure is a linear feedback law. According to the counterexample of [7], however, when the information structure is non-classical (e.g. decentralized), the optimal strategy may not be linear. Distributed linear quadratic control problem was investigated in [8] and [9], where it was shown that the solution is limited by the computational cost, and hence not suitable for large-scale networks. Therefore, the problem of finding a decentralized linear optimal strategy for a leader-follower network with a large number of agents is computationally challenging.

On the other hand, real-world multi-agent systems are subject to external disturbances. Thus, a practical control strategy should be robust to disturbances. One of the numerous techniques to design a robust controller in the literature, is minimax control, which has two main challenges. Computational complexity of solutions and communication constraints. To address these shortcomings, this work is focused on utilizing mean-field teams approach in minimax problems in order to propose decentralized and scalable solution.

1.3 Mean-Field Control

The concept of mean field can be traced back to the statistical mechanics literature [45, 46], where it was used as a basis for understanding phase transition phenomenon, which suggest that an atomic spin moves in the average field (or mean field) produced by all other spins. It is known that the propagation of a large number of microscopic particles (moving independently) can be modeled by a macroscopic mass that spreads over time according to the diffusion equation [48]. This property has been widely used in control theory to alleviate the computational complexity of large-scale control systems. For instance, mean-field game theory was developed to study the non-cooperative behavior of a large number of players [10, 11, 43]. The main idea in mean-field games is to exploit two key features in the presence of a large number of players: the negligible influence of individual players and the law of large numbers. The solution proposed in [10, 43, 44] is an approximate Nash equilibrium that is given in terms of two coupled forward-backward equations. The existence of such a solution is established by imposing proper fixed-point conditions and the monotonicity hypothesis.

A new line of research was recently initiated by introducing mean-field teams (MFT), which investigate the cooperative behavior of an arbitrary number of agents (which is not necessarily large). The main difference between the mean-field team and mean-field game approaches is that the MFT approach exploits the symmetry of the problem rather than the negligible influence of individual players [12–19, 50] and identifies the globally optimal solution (rather than Nash equilibrium or person-by-person optimal solution) in terms of two decoupled equations. In additions, neither fixed-point conditions nor the monotonicity hypothesis are required to find the solution of MFT. As an application of MFT, the authors in [17] investigate a network of users infected by an infectious process such as disease. The optimal and sub-optimal solutions for the network are obtained by using MFT techniques under two different information structures.

For linear quadratic control systems, the mean-field solution is optimal for any arbitrary number of agents, under mean-field sharing, and it is sub-optimal under partial mean-field sharing. This thesis is focused on linear quadratic mean-field teams, wherein mean-field type control may be viewed as a special case [42, 47].

1.4 Thesis Contributions

Two main problems are investigated in this dissertation. The first problem is concerned with a multi-agent network consisting of one leader and many identical followers, where the agents are coupled in both dynamics and cost function. Some important characteristics of the network such as convergence rate of followers to the leader or the collective behavior of followers are discussed. Then, a near-optimal strategy for a non-classical information structure is proposed by solving two decoupled Riccati equations. The computational complexity of the method is independent of the number of followers because the corresponding equations are decoupled. Then, it is shown that the proposed solution converges to the optimal strategy obtained in [15] at a rate inversely proportional to the number of followers.

The second problem studied in this dissertation is concerned with finding a robust control strategy for the multi-agent system introduced above by solving a minimax control problem under three decentralized information structures, namely, *mean-field sharing*, *partial mean-field sharing* and *intermittent mean-field sharing*. Under mean-field sharing information structure, it is proved that a unique saddle-point strategy exists. It is also shown that the proposed strategies under the other two information structures, partial mean-field sharing and intermittent mean-field sharing, converge to the saddle-point strategy as the number of followers tends to infinity. Moreover, it is shown that the robust solutions are scalable, similar to the first problem. The effectiveness of disturbance rejection and the consensus-reaching behavior of the solution are discussed and their dependency on the parameters of cost function is analyzed.

1.5 Thesis Layout

The structure of the thesis is as follows:

- **Chapter** 1 includes the motivation and background for the study, and outlines the contributions of the work. Parts of this chapter are adopted from the author's publication below:

  J. Arabneydi, **M. M. Baharloo** and A. G. Aghdam, "Optimal distributed control for leader-follower networks: A scalable design," in Proceedings of *the 31st IEEE Canadian Conference on Electrical and Computer Engineering, pp. 1 - 4, 2018.*

- **Chapter 2** studies a decentralized stochastic control system consisting of one leader and many homogeneous followers. The leader and followers are coupled in both dynamics and cost; the agent dynamics are linear and the cost function is quadratic in the states and actions of the leader and followers. By using the concept of mean-field teams, a near-optimal control strategy is proposed, which is shown to converge to the optimal solution as the number of followers increases. Three numerical examples are provided to demonstrate the efficacy of the results.

- **Chapter 3** investigates a soft-constrained minimax control problem of a leader-follower network. The network consists of one leader and an arbitrary number of followers that wish to reach consensus with minimum energy in the presence of external disturbances. The problem is solved for three non-classical information structures.

- **Chapter 4** presents the conclusion as well as the possible directions for future research in this area.

Chapter 2

Near-Optimal Control Strategy in Leader-Follower Networks: A Case Study for Linear Quadratic Mean-Field Teams

In this chapter, a decentralized stochastic control system consisting of one leader and many homogeneous followers is studied. The objective of the agents is to reach consensus while minimizing their communication and energy costs. The leader knows its local state and each follower knows its local state as well as the state of the leader. The number of required links to implement this decentralized information structure is equal to the number of followers and in the special case of a leaderless network (which can be modeled as a leader-follower network with a constant reference state trajectory as a virtual leader), no links need to exist between agents, i.e., the communication graph is not required to be connected. We propose a near-optimal control strategy that converges to the optimal solution as the number of followers increases. One of the salient features of the proposed solution is that it provides a framework to determine the convergence rate and collective behavior of the followers by choosing appropriate cost functions. In addition, the computational complexity of the proposed solution does not depend on the number of followers. Furthermore, the proposed strategy can be computed in a distributed manner, where the leader solves one Riccati equation and each follower solves two Riccati equations to determine their strategies. Numerical examples are provided to demonstrate the effectiveness of the results in the control of multi-agent systems.

This chapter is based on the following publication:

M. Baharloo, J. Arabneydi, and A. G. Aghdam, "Near-optimal control strategy in leader-follower networks: A case study for linear quadratic mean-field teams," in Proceedings of *the 57th IEEE Conference on Decision and Control, pp. 3288–3293, 2018.*

The most important additions to the above-mentioned paper are as follows. Lemmas 2.2 and 2.3 along with their proofs as well as the detailed proof of Lemma 2.4 have been added here. In addition, to compare the effect of different cost function parameters, a numerical example is presented in this chapter which does not appear in the paper.

2.1 Introduction

The control of multi-agent systems with leader-follower structure has attracted much interest in the past two decades due to its wide range of applications in various fields of science and engineering. Such applications include vehicle formation [20], sensor networks [21], surveillance using a team of unmanned aerial vehicles (UAVs) [22] and flocking [23, 24], to name only a few. In this type of problem, a group of agents (called followers) are to track another group (called leaders), while certain performance specifications are met. Different performance measures such as minimum energy, fuel or time are considered in the literature. For this purpose, limited communication and computation resources are two main challenges that need to be overcome.

To address the above challenges, the following two problems have been investigated in the context of consensus control protocols in the literature recently: (i) how the states of the followers can reach the state of the leader under communication constraints (distributed control problem), where consensus is reached under appropriate linear strategies for properly connected communication graphs [25–27], and (ii) how the states of the followers can reach the state of the leader with minimum energy consumption (optimal control problem). Note that the optimal control strategy for a quadratic performance index with linear dynamics under centralized information structure is a linear feedback rule obtained by the solution of the celebrated Riccati equation [28].

Combining the above two objectives, however, is quite challenging as it leads to a decentralized optimal control problem wherein the optimal control law is not necessarily linear [7]. Furthermore, since the dimension of the matrices in the network model increases with the number of followers, the optimal control law may be intractable for a network of large size. In this chapter, we consider a decentralized optimal control problem with a large number of followers.

For dynamically decoupled followers and also the case of a leaderless multi-agent system, [8, 9] use the control inverse optimality approach to compute the optimal distributed control strategies for special classes of communication graphs. The authors in [29] consider a large number of homogeneous followers and determine the optimal strategy by solving two coupled Riccati equations. In contrast, this work studies a leader-follower multi-agent network with coupled dynamics under a directed communication graph in which there is a direct link from the leader to each follower. In the special case of a leaderless network, the communication graph is not required to be connected. When the initial states of followers are identically and independently distributed, a near-optimal strategy is proposed for a large number of homogeneous followers by solving two decoupled Riccati equations using mean-field team approach introduced in [12] and showcased in [13–15, 30].

The remainder of this chapter is organized as follows. The problem is formulated in Section 2.2, where the main contributions of the work are also outlined. Then in Section 2.3, some important assumptions are presented and the control strategy is derived. Numerical examples are provided in Section 2.4 and finally the paper is concluded in Section 2.5.

2.2 Problem Formulation

2.2.1 Notation

In this chapter, $\mathbb{N}$ and $\mathbb{R}$ denote natural and real numbers, respectively. The short-hand notation $x_{a:b}$ is used to denote vector $(x_a, \ldots, x_b)$, $a \leq b \in \mathbb{N}$. For any $k \in \mathbb{N}$, $\mathbb{N}_k$ is the finite set of integers $\{1, 2, \ldots, k\}$. $\mathrm{Tr}(\cdot)$ is the trace of a matrix and $\mathrm{var}(\cdot)$ is the covariance of a random vector.

2.2.2 Dynamics

Consider a multi-agent system consisting of one leader and $n \in \mathbb{N}$ followers. Denote by $x_t^0 \in \mathbb{R}^{d_x}$, $u_t^0 \in \mathbb{R}^{d_u}$, and $w_t^0 \in \mathbb{R}^{d_x}$, $d_x, d_u \in \mathbb{N}$, the state, action, and noise of the leader at time $t \in \mathbb{N}$, respectively. In addition, let $x_t^i \in \mathbb{R}^{d_x}$, $u_t^i \in \mathbb{R}^{d_u}$, and $w_t^i \in \mathbb{R}^{d_x}$ denote the state, action, and noise of follower $i \in \mathbb{N}_n$ at time $t \in \mathbb{N}$, analogously. The state of the leader evolves as follows:

$$x_{t+1}^0 = A_t^0 x_t^0 + B_t^0 u_t^0 + D_t^0 \bar{x}_t + w_t^0, \tag{2.1}$$

where $\bar{x}_t := \frac{1}{n}\sum_{i=1}^n x_t^i$ is the average of the states of the followers at time t, and will hereafter be called mean-field [12]. Similarly, the state of each follower $i \in \mathbb{N}_n$ evolves as follows:

$$x_{t+1}^i = A_t x_t^i + B_t u_t^i + D_t \bar{x}_t + E_t x_t^0 + w_t^i. \tag{2.2}$$

Let $\mathbb{N}_T$, $T \in \mathbb{N}$, be the control horizon. It is assumed that the primitive random variables $\{x_1^0, x_1^1, \ldots, x_1^n, w_1^0, w_1^1, \ldots, w_1^n, \ldots, w_T^0, w_T^1, \ldots, w_T^n\}$ are defined on a common probability space, and are mutually independent.

2.2.3 Information structure

At time $t \in \mathbb{N}$, the leader observes its state x_t^0 and chooses its action u_t^0 according to a control law $g_t^0 : \mathbb{R}^{d_x} \to \mathbb{R}^{d_u}$, i.e.,

$$u_t^0 = g_t^0(x_t^0). \tag{2.3}$$

In addition, for any $i \in \mathbb{N}_n$, follower i observes its state x_t^i as well as the state of the leader x_t^0 at time t, and decides its action u_t^i as follows:

$$u_t^i = g_t^i(x_t^i, x_t^0), \tag{2.4}$$

where $g_t^i : \mathbb{R}^{d_x} \times \mathbb{R}^{d_x} \to \mathbb{R}^{d_u}$ is the control law. Note that the control actions (2.3) and (2.4) have a non-classical decentralized information structure.

Remark 2.1. The number of links required to implement the above information structure is n, which is the number of followers. In addition, no information about the states of the followers is communicated; hence, the privacy of the followers is preserved.

Remark 2.2. It is worth highlighting the difference between the coupling in dynamics given by (2.1) and (2.2), that refers to the physical interactions among agents, and information coupling in (2.3) and (2.4), that attributes to the information exchange.

The set of all control laws $\mathbf{g} := \{g^0_{1:T}, g^1_{1:T}, \dots, g^n_{1:T}\}$ is called the strategy of the network. The objective of the followers is to track the leader in an energy-efficient manner. To this end, the following cost function is defined:

$$
\begin{aligned}
J_T(\mathbf{g}) = \mathbb{E}\Big[\sum_{t=1}^{T} &(x_t^0)^\intercal Q_t^0 x_t^0 + (\bar{x}_t - x_t^0)^\intercal F_t (\bar{x}_t - x_t^0) + (u_t^0)^\intercal R_t^0 u_t^0 \\
&+ \frac{1}{n}\sum_{i=1}^{n} (x_t^i)^\intercal Q_t x_t^i + (x_t^i - x_t^0)^\intercal P_t (x_t^i - x_t^0) + (u_t^i)^\intercal R_t u_t^i \\
&+ \frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} (x_t^i - x_t^j)^\intercal H_t (x_t^i - x_t^j)\Big],
\end{aligned} \tag{2.5}
$$

where the expectation is taken with respect to the probability measures induced by the choice of strategy $\mathbf{g}$, and $\{Q_t^0, F_t, R_t^0, Q_t, P_t, R_t, H_t\}$ are symmetric matrices of appropriate dimensions. It is to be noted that the rate of convergence of the followers to the leader depends on the matrices P_t and F_t. Moreover, the collective behavior of the followers changes by matrix H_t.

Problem 2.1. *Consider the above leader-follower system with dynamics* (2.1) *and* (2.2) *and information structure* (2.3) *and* (2.4). *We are interested to find an* $\varepsilon(n)$*-optimal strategy* $\mathbf{g}^*_\varepsilon$ *such that for every strategy* $\mathbf{g}$,

$$J_T(\mathbf{g}^*_\varepsilon) \leq J_T(\mathbf{g}) + \varepsilon(n), \tag{2.6}$$

where $\varepsilon(n) \in [0, \infty)$ *and* $\lim_{n\to\infty} \varepsilon(n) = 0$.

Remark 2.3. It is to be noted that if matrices B^0_t, D^0_t, Q^0_t and R^0_t are zero, Problem 2.1 reduces to the optimal control of a leaderless multi-agent network. In that case, x^0_t represents the desired reference trajectory, and as noted before, the followers do not share anything with each other once they receive the reference trajectory information x^0_t, according to (2.4).

2.2.4 Main challenges and contributions

There are two main challenges in finding a solution to Problem 2.1. The first one is concerned with non-classical information structure, as the optimal strategy under this type of information structure is not necessarily linear [7]. The second challenge is the curse of dimensionality as the matrices in Problem 2.1 are fully dense, yet their dimension increases with the number of followers.

In their previous work [15], the authors show that if the mean-field $\bar{x}_t$ is available to the leader and followers, then the optimal solution is unique and linear. However, collecting and sharing the mean-field among all agents is not cost-efficient, in general, specially when the number of followers n is large. It is shown in the next section that the effect of such information sharing on the performance of the network is negligible when the number of followers is large enough.

2.3 Main Results

In this section, we propose a strategy and compute its performance with respect to the optimal performance, and show that the difference between them converges to zero at rate $\frac{1}{n}$.

For the sake of clarity in the notation, we use letters s and v to denote the states and actions, respectively, under the optimal strategy. Therefore, from (2.1) and (2.2), the dynamics of the leader and followers at time $t \in \mathbb{N}_T$ under the optimal strategy are given by

$$s^0_{t+1} = A^0_t s^0_t + B^0_t v^0_t + D^0_t \bar{s}_t + w^0_t, \tag{2.7}$$

$$s^i_{t+1} = A_t s^i_t + B_t v^i_t + D_t \bar{s}_t + E_t s^0_t + w^i_t, \quad i \in \mathbb{N}_n, \tag{2.8}$$

where $\bar{s}_t := \frac{1}{n}\sum_{i=1}^{n} s_t^i$. Similarly to [15], define the following matrices:

$$\bar{A}_t := \begin{bmatrix} A_t^0 & D_t^0 \\ E_t & A_t + D_t \end{bmatrix}, \quad \bar{B}_t := \begin{bmatrix} B_t^0 & \mathbf{0}_{d_x \times d_u} \\ \mathbf{0}_{d_x \times d_u} & B_t \end{bmatrix}, \tag{2.9}$$

$$\bar{Q}_t := \begin{bmatrix} Q_t^0 + P_t + F_t & -P_t - F_t \\ -P_t - F_t & Q_t + P_t + F_t \end{bmatrix}, \bar{R}_t := \begin{bmatrix} R_t^0 & \mathbf{0}_{d_u \times d_u} \\ \mathbf{0}_{d_u \times d_u} & R_t \end{bmatrix}. \tag{2.10}$$

Assumption 2.1. *Matrices $Q_t + P_t + H_t$ and $\bar{Q}_t$ are positive semi-definite and matrices R_t and R_t^0 are positive definite.*

Define two decoupled Riccati equations such that for any $t \in \mathbb{N}_T$,

$$\breve{M}_t = -A_t^\intercal \breve{M}_{t+1} B_t \left(B_t^\intercal \breve{M}_{t+1} B_t + R_t\right)^{-1} B_t^\intercal \breve{M}_{t+1} A_t + A_t^\intercal \breve{M}_{t+1} A_t + Q_t + P_t + H_t, \tag{2.11}$$

$$\bar{M}_t = -\bar{A}_t^\intercal \bar{M}_{t+1} \bar{B}_t \left(\bar{B}_t^\intercal \bar{M}_{t+1} \bar{B}_t + \bar{R}_t\right)^{-1} \bar{B}_t^\intercal \bar{M}_{t+1} \bar{A}_t + \bar{A}_t^\intercal \bar{M}_{t+1} \bar{A}_t + \bar{Q}_t, \tag{2.12}$$

where $\breve{M}_{T+1} = \mathbf{0}_{d_x \times d_x}$ and $\bar{M}_{T+1} = \mathbf{0}_{2d_x \times 2d_x}$. According to [12, 15], the optimal performance J_T^* is obtained under the following linear strategies:

$$v_t^0 = \bar{L}_t^{1,1} s_t^0 + \bar{L}_t^{1,2} \bar{s}_t, \tag{2.13}$$

$$v_t^i = \breve{L}_t s_t^i + \bar{L}_t^{2,1} s_t^0 + (\bar{L}_t^{2,2} - \breve{L}_t)\bar{s}_t, \tag{2.14}$$

where $\breve{L}_t$ and $\bar{L}_t =: \begin{bmatrix} \bar{L}_t^{1,1} & \bar{L}_t^{1,2} \\ \bar{L}_t^{2,1} & \bar{L}_t^{2,2} \end{bmatrix}$ can be found by using these formulas:

$$\breve{L}_t = -\left(B_t^\intercal \breve{M}_{t+1} B_t + R_t\right)^{-1} B_t^\intercal \breve{M}_{t+1} A_t,$$
$$\bar{L}_t = -\left(\bar{B}_t^\intercal \bar{M}_{t+1} \bar{B}_t + \bar{R}_t\right)^{-1} \bar{B}_t^\intercal \bar{M}_{t+1} \bar{A}_t. \tag{2.15}$$

2.3.1 Solution of Problem 2.1

The following standard assumptions are imposed on the model.

Assumption 2.2. *The initial states of the followers are i.i.d. with mean $\mu_x \in \mathbb{R}^{d_x}$ and finite covariance matrix $\sum_x \in \mathbb{R}^{d_x} \times \mathbb{R}^{d_x}$. In addition, the local noises of the followers are i.i.d. with zero mean and finite covariance matrix $\sum_w \in \mathbb{R}^{d_x} \times \mathbb{R}^{d_x}$.*

Assumption 2.3. *Matrices $A_t, A_t^0, B_t, B_t^0, D_t, D_t^0, E_t, F_t, Q_t,\ Q_t^0,\ R_t,\ R_t^0,\ \sum_x$ and $\sum_w$ do not depend on the number of followers n.*

Define a stochastic process $z_{1:T}$ such that $z_1 := \mu_x$ and for any $t \in \mathbb{N}_T$:

$$z_{t+1} = (A_t + B_t \bar{L}_t^{2,2} + D_t) z_t + (B_t \bar{L}_t^{2,1} + E_t) x_t^0. \tag{2.16}$$

Note that the leader and followers can compute z_t under the information structures (2.3) and (2.4). Given the matrix gains defined in (2.15), the following strategies are proposed:

$$u_t^0 = \bar{L}_t^{1,1} x_t^0 + \bar{L}_t^{1,2} z_t, \tag{2.17}$$

$$u_t^i = \breve{L}_t x_t^i + \bar{L}_t^{2,1} x_t^0 + (\bar{L}_t^{2,2} - \breve{L}_t) z_t, \quad i \in \mathbb{N}_n. \tag{2.18}$$

At any time $t \in \mathbb{N}_T$, define the following relative errors e_t^0, e_t and ζ_t:

$$e_t^0 := s_t^0 - x_t^0, \quad e_t := \bar{s}_t - z_t, \quad \zeta_t := \bar{x}_t - z_t. \tag{2.19}$$

Lemma 2.1. *The relative errors defined in (2.19) evolve linearly in time as follows:*

$$\begin{bmatrix} e_{t+1}^0 \\ e_{t+1} \\ \zeta_{t+1} \end{bmatrix} = \tilde{A}_t \begin{bmatrix} e_t^0 \\ e_t \\ \zeta_t \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{d_x \times 1} \\ \bar{w}_t \\ \bar{w}_t \end{bmatrix}, \tag{2.20}$$

where $\bar{w}_t := \frac{1}{n} \sum_{i=1}^n w_t^i$ *and*

$$\tilde{A}_t := \begin{bmatrix} A_t^0 + B_t^0 \bar{L}_t^{1,1} & B_t^0 \bar{L}_t^{1,2} + D_t^0 & -D_t^0 \\ B_t \bar{L}_t^{2,1} + E_t & A_t + B_t \bar{L}_t^{2,2} + D_t & \mathbf{0}_{d_x \times d_x} \\ \mathbf{0}_{d_x \times d_x} & \mathbf{0}_{d_x \times d_x} & A_t + B_t \breve{L}_t + D_t \end{bmatrix}. \tag{2.21}$$

Proof. From (2.7) and (2.13):

$$s_{t+1}^0 = (A_t^0 + B_t^0 \bar{L}_t^{1,1}) s_t^0 + (B_t^0 \bar{L}_t^{1,2} + D_t^0) \bar{s}_t + w_t^0. \tag{2.22}$$

Also, it results from (2.1) and (2.17) that

$$x_{t+1}^0 = (A_t^0 + B_t^0 \bar{L}_t^{1,1}) x_t^0 + B_t^0 \bar{L}_t^{1,2} z_t + D_t^0 \bar{x}_t + w_t^0. \tag{2.23}$$

Similarly, from (2.8) and (2.14):

$$\bar{s}_{t+1} = A_t \bar{s}_t + B_t \bar{v}_t + D_t \bar{s}_t + E_t s_t^0 + \bar{w}_t, \tag{2.24}$$

where $\bar{v}_t := \frac{1}{n} \sum_{i=1}^n v_t^i$ is given by:

$$\bar{v}_t = \breve{L}_t \bar{s}_t + \bar{L}_t^{2,1} s_t^0 + (\bar{L}_t^{2,2} - \breve{L}_t) \bar{s}_t = \bar{L}_t^{2,1} s_t^0 + \bar{L}_t^{2,2} \bar{s}_t. \tag{2.25}$$

Substituting (2.25) in (2.24) yields:

$$\bar{s}_{t+1} = (A_t + B_t \bar{L}_t^{2,2} + D_t) \bar{s}_t + (B_t \bar{L}_t^{2,1} + E_t) s_t^0 + \bar{w}_t. \tag{2.26}$$

In addition, from (2.2) and (2.18), one arrives at:

$$\bar{x}_{t+1} = A_t \bar{x}_t + B_t \bar{u}_t + D_t \bar{x}_t + E_t x_t^0 + \bar{w}_t, \tag{2.27}$$

where $\bar{u}_t := \frac{1}{n} \sum_{i=1}^n u_t^i$ is as follows:

$$\bar{u}_{t+1} = \breve{L}_t \bar{x}_t + \bar{L}_t^{2,1} x_t^0 + (\bar{L}_t^{2,2} - \breve{L}_t) z_t. \tag{2.28}$$

From (2.27) and (2.28), it results that:

$$\bar{x}_{t+1} \;=\; A_t\bar{x}_t \;+\; B_t\breve{L}_t\bar{x}_t \;+\; B_t\bar{L}_t^{2,1}x_t^0 \;+\; B_t(\bar{L}_t^{2,2} \;-\; \breve{L}_t)z_t \;+\; D_t\bar{x}_t \;+\; E_tx_t^0 \;+\; \bar{w}_t. \quad (2.29)$$

Equations (2.19), (2.22) and (2.23) lead to:

$$\begin{aligned}
e_{t+1}^0 &= (A_t^0 + B_t^0\bar{L}_t^{1,1})s_t^0 + B_t^0\bar{L}_t^{1,2}\bar{s}_t + D_t^0\bar{s}_t + w_t^0 \\
&\quad - (A_t^0x_t^0 + B_t^0\bar{L}_t^{1,1}x_t^0 + B_t^0\bar{L}_t^{1,2}z_t + D_t^0\bar{x}_t + w_t^0) \\
&= (A_t^0 + B_t^0\bar{L}_t^{1,1})e_t^0 + (B_t^0\bar{L}_t^{1,2} + D_t^0)e_t - D_t^0\zeta_t. \qquad (2.30)
\end{aligned}$$

Moreover, it results from (2.16), (2.19) and (2.26) that:

$$\begin{aligned}
e_{t+1} &= A_t\bar{s}_t + B_t\bar{L}_t^{2,2}\bar{s}_t + D_t\bar{s}_t + B_t\bar{L}_t^{2,1}s_t^0 + E_ts_t^0 + \bar{w}_t \\
&\quad - (A_tz_t + B_t\bar{L}_t^{2,2}z_t + D_tz_t + B_t\bar{L}_t^{2,1}x_t^0 + E_tx_t^0), \qquad (2.31) \\
&= (A_t + B_t\bar{L}_t^{2,2} + D_t)e_t + (B_t\bar{L}_t^{2,1} + E_t)e_t^0 + \bar{w}_t. \qquad (2.32)
\end{aligned}$$

As a consequence of (2.16), (2.19) and (2.29), the following equation is obtained:

$$\begin{aligned}
\zeta_{t+1} &= (A_t + B_t\breve{L}_t + D_t)\bar{x}_t + B_t(\bar{L}_t^{2,1} - \breve{L}_t)x_t^0 + B_t\bar{L}_t^{2,2}z_t \\
&\quad + E_tx_t^0 + \bar{w}_t - (A_t + B_t\bar{L}_t^{2,2} + D_t)z_t - (B_t\bar{L}_t^{2,1} + E_t)x_t^0 \\
&= (A_t + B_t\breve{L}_t + D_t)\zeta_t + \bar{w}_t, \qquad (2.33)
\end{aligned}$$

and this completes the proof. ■

Lemma 2.2. *Let Assumption 2.2 hold. At any time $t \in \mathbb{N}_T$, $\mathbb{E}\big[e_t^0\big] = \mathbb{E}\big[e_t\big] = \mathbb{E}\big[\zeta_t\big] = \mathbf{0}_{d_x \times 1}$.*

Proof. Initially at $t = 1$, $\mathbb{E}\big[e_1^0\big] = \mathbb{E}\big[s_1^0 - x_1^0\big] = \mathbf{0}_{d_x \times 1}$, $\mathbb{E}\big[e_1\big] = \mathbb{E}\big[\bar{s}_1 - z_1\big] = \mu_x - \mu_x = \mathbf{0}_{d_x \times 1}$, and $\mathbb{E}\big[\zeta_1\big] = \mathbb{E}\big[\bar{x}_1 - z_1\big] = \mu_x - \mu_x = \mathbf{0}_{d_x \times 1}$. Since the relative errors evolve linearly according to Lemma 2.1 and also $\mathbb{E}\big[\bar{w}_t\big] = 0$ according to Assumption 2.2, we have $\mathbb{E}\big[e_t^0\big] = \mathbb{E}\big[e_t\big] = \mathbb{E}\big[\zeta_t\big] = \mathbf{0}_{d_x \times 1}$ at any $t \in \mathbb{N}_T$. ■

Now, for any follower $i \in \mathbb{N}_n$, define the following variables at time $t \in \mathbb{N}_T$:

$$\breve{x}_t^i := x_t^i - \bar{x}_t, \breve{u}_t^i := u_t^i - \bar{u}_t, \breve{s}_t^i := s_t^i - \bar{s}_t, \breve{v}_t^i := v_t^i - \bar{v}_t. \tag{2.34}$$

Lemma 2.3. *At any time $t \in \mathbb{N}_T$, $\breve{x}_t^i = \breve{s}_t^i$ and $\breve{u}_t^i = \breve{v}_t^i$.*

Proof. The lemma is proved by induction on noting that initially $\breve{x}_1^i = \breve{s}_1^i = x_1^i - \bar{x}_1$ because $x_1^i = s_1^i$. It follows from (2.14) and (2.18) that $\breve{u}_1^i = \breve{v}_1^i = \breve{L}_1(x_1^i - \bar{x}_1)$. Suppose $\breve{x}_t^i = \breve{s}_t^i$ and $\breve{u}_t^i = \breve{v}_t^i$. It is now desired to show that $\breve{x}_{t+1}^i = \breve{s}_{t+1}^i$ and $\breve{u}_{t+1}^i = \breve{v}_{t+1}^i$. From (2.2) and (2.8) and the induction assumption at $t = 1$, one arrives at:

$$\breve{s}_{t+1}^i = A_t \breve{s}_t^i + B_t \breve{v}_t^i + \breve{w}_t^i = A_t \breve{x}_t^i + B_t \breve{u}_t^i + \breve{w}_t^i = \breve{x}_{t+1}^i, \tag{2.35}$$

where $\breve{w}_t^i := w_t^i - \bar{w}_t$. Also, it is implied from (2.14), (2.18) and (2.35) that:

$$\breve{v}_{t+1}^i = \breve{L}_{t+1} \breve{s}_{t+1}^i = \breve{L}_{t+1} \breve{x}_{t+1}^i = \breve{u}_{t+1}^i. \tag{2.36}$$

$\blacksquare$

Lemma 2.4. *Let ΔJ denote the discrepancy between the performance of the optimal strategies (2.13) and (2.14), and that of the proposed strategies (2.17) and (2.18). If Assumption 2.2 holds, then ΔJ is a quadratic function of the relative errors in (2.19), i.e.,*

$$\Delta J = \mathbb{E}\Big[\sum_{t=1}^{T}\begin{bmatrix} e_t^0 & e_t & \zeta_t \end{bmatrix}^\intercal \tilde{Q}_t \begin{bmatrix} e_t^0 & e_t & \zeta_t \end{bmatrix}\Big], \tag{2.37}$$

where

$$\tilde{Q}_t := \begin{bmatrix} -\bar{Q}_t - \bar{L}_t^\intercal \bar{R}_t \bar{L}_t & \mathbf{0}_{2d_x \times d_x} \\ \mathbf{0}_{d_x \times 2d_x} & Q_t + P_t + F_t + \breve{L}_t^\intercal R_t \breve{L}_t \end{bmatrix}. \tag{2.38}$$

Proof. From (2.5), we have

$$\begin{aligned}\Delta J = \mathbb{E}\Big[\sum_{t=1}^{T}(x_t^0)^\intercal Q_t^0 x_t^0 + (u_t^0)^\intercal R_t^0 u_t^0 + (\bar{x}_t - x_t^0)^\intercal F_t(\bar{x}_t - x_t^0)\\ + \frac{1}{n}\sum_{i=1}^{n}(x_t^i)^\intercal Q_t x_t^i + (x_t^i - x_t^0)^\intercal P_t (x_t^i - x_t^0) + (u_t^i)^\intercal R_t u_t^i\\ + \frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(x_t^i - x_t^j)^\intercal H_t (x_t^i - x_t^j)\Big]\\ - \mathbb{E}\Big[\sum_{t=1}^{T}(s_t^0)^\intercal Q_t^0 s_t^0 + (v_t^0)^\intercal R_t^0 v_t^0 + (\bar{s}_t - s_t^0)^\intercal F_t(\bar{s}_t - s_t^0)\\ + \frac{1}{n}\sum_{i=1}^{n}(s_t^i)^\intercal Q_t s_t^i + (s_t^i - s_t^0)^\intercal P_t (s_t^i - s_t^0) + (v_t^i)^\intercal R_t v_t^i\\ + \frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(s_t^i - s_t^j)^\intercal H_t (s_t^i - s_t^j)\Big]. \end{aligned} \tag{2.39}$$

The above equation can be re-written in terms of the variables defined in (2.34) as follows:

$$
\begin{aligned}
\Delta J = \mathbb{E}\Big[&\sum_{t=1}^{T} (x_t^0)^\intercal Q_t^0 x_t^0 + (u_t^0)^\intercal R_t^0 u_t^0 + (\bar{x}_t - x_t^0)^\intercal F_t (\bar{x}_t - x_t^0) \\
&+ \frac{1}{n}\sum_{i=1}^{n} (\breve{x}_t^i + \bar{x}_t)^\intercal Q_t (\breve{x}_t^i + \bar{x}_t) + (\breve{x}_t^i + \bar{x}_t - x_t^0)^\intercal P_t (\breve{x}_t^i + \bar{x}_t - x_t^0) \\
&+ (\breve{u}_t^i + \bar{u}_t)^\intercal R_t (\breve{u}_t^i + \bar{u}_t) + \frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} (\breve{x}_t^i - \breve{x}_t^j)^\intercal H_t (\breve{x}_t^i - \breve{x}_t^j) \Big] \\
&- \mathbb{E}\Big[\sum_{t=1}^{T} (s_t^0)^\intercal Q_t^0 s_t^0 + (v_t^0)^\intercal R_t^0 v_t^0 + (\bar{s}_t - s_t^0)^\intercal F_t (\bar{s}_t - s_t^0) \\
&+ \frac{1}{n}\sum_{i=1}^{n} (\breve{s}_t^i + \bar{s}_t)^\intercal Q_t (\breve{s}_t^i + \bar{s}_t) + (\breve{s}_t^i + \bar{s}_t - s_t^0)^\intercal P_t (\breve{s}_t^i + \bar{s}_t - s_t^0) \\
&+ (\breve{v}_t^i + \bar{v}_t)^\intercal R_t (\breve{v}_t^i + \bar{v}_t) + \frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} (\breve{s}_t^i - \breve{s}_t^j)^\intercal H_t (\breve{s}_t^i - \breve{s}_t^j) \Big]. \qquad (2.40)
\end{aligned}
$$

On the other hand, by definition the following relations hold:

$$
\sum_{i=1}^{n} \breve{x}_t^i = \sum_{i=1}^{n} \breve{s}_t^i = \mathbf{0}_{d_x \times 1}, \sum_{i=1}^{n} \breve{u}_t^i = \sum_{i=1}^{n} \breve{v}_t^i = \mathbf{0}_{d_u \times 1}. \qquad (2.41)
$$

By incorporating (2.41) in (2.40), it results that

$$
\begin{aligned}
\Delta J = \mathbb{E}\Big[\sum_{t=1}^{T} \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix}^\intercal \bar{Q}_t \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix} + \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix}^\intercal \bar{R}_t \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix} \\
+ \frac{1}{n}\sum_{i=1}^{n} (\breve{x}_t^i)^\intercal (Q_t + P_t + H_t)(\breve{x}_t^i) + (\breve{u}_t^i)^\intercal R_t (\breve{u}_t^i)\Big] \\
- \mathbb{E}\Big[\sum_{t=1}^{T} \begin{bmatrix} s_t^0 \\ \bar{s}_t \end{bmatrix}^\intercal \bar{Q}_t \begin{bmatrix} s_t^0 \\ \bar{s}_t \end{bmatrix} + \begin{bmatrix} v_t^0 \\ \bar{v}_t \end{bmatrix}^\intercal \bar{R}_t \begin{bmatrix} v_t^0 \\ \bar{v}_t \end{bmatrix} \\
+ \frac{1}{n}\sum_{i=1}^{n} (\breve{s}_t^i)^\intercal (Q_t + P_t + H_t)(\breve{s}_t^i) + (\breve{v}_t^i)^\intercal R_t (\breve{v}_t^i)\Big].
\end{aligned}
$$

According to Lemma 2.3, the above equation simplifies to

$$
\begin{aligned}
\Delta J = \mathbb{E}\Big[\sum_{t=1}^{T} \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix}^\intercal \bar{Q}_t \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix} + \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix}^\intercal \bar{R}_t \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix}\Big] \\
- \mathbb{E}\Big[\sum_{t=1}^{T} \begin{bmatrix} s_t^0 \\ \bar{s}_t \end{bmatrix}^\intercal \bar{Q}_t \begin{bmatrix} s_t^0 \\ \bar{s}_t \end{bmatrix} + \begin{bmatrix} v_t^0 \\ \bar{v}_t \end{bmatrix}^\intercal \bar{R}_t \begin{bmatrix} v_t^0 \\ \bar{v}_t \end{bmatrix}\Big].
\end{aligned} \tag{2.42}
$$

Using the definition of relative errors in (2.19), one concludes that:

$$\begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix} = \begin{bmatrix} x_t^0 \\ z_t \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{d_x \times d_x} \\ \zeta_t \end{bmatrix}, \begin{bmatrix} s_t^0 \\ \bar{s}_t \end{bmatrix} = \begin{bmatrix} e_t^0 + x_t^0 \\ e_t + z_t \end{bmatrix},$$
$$\begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix} = \bar{L}_t \begin{bmatrix} x_t^0 \\ z_t \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{d_x \times d_x} \\ \breve{L}_t \zeta_t \end{bmatrix}, \begin{bmatrix} v_t^0 \\ \bar{v}_t \end{bmatrix} = \bar{L}_t \begin{bmatrix} s_t^0 \\ \bar{s}_t \end{bmatrix}. \quad (2.43)$$

It is implied from (2.42) and (2.43) that:

$$\begin{aligned} \Delta J = & \mathbb{E}\Big[\sum_{t=1}^{T} (\begin{bmatrix} x_t^0 \\ z_t \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{d_x \times d_x} \\ \zeta_t \end{bmatrix})^\intercal \bar{Q}_t (\begin{bmatrix} x_t^0 \\ z_t \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{d_x \times d_x} \\ \zeta_t \end{bmatrix}) \Big] \\ & + \mathbb{E}\Big[\sum_{t=1}^{T} (\bar{L}_t \begin{bmatrix} x_t^0 \\ z_t \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{d_x \times d_x} \\ \breve{L}_t \zeta_t \end{bmatrix})^\intercal \bar{R}_t (\bar{L}_t \begin{bmatrix} x_t^0 \\ z_t \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{d_x \times d_x} \\ \breve{L}_t \zeta_t \end{bmatrix}) \Big] \\ & - \mathbb{E}\Big[\sum_{t=1}^{T} \begin{bmatrix} e_t^0 + x_t^0 \\ e_t + z_t \end{bmatrix}^\intercal (\bar{Q}_t + \bar{L}_t^\intercal \bar{R}_t \bar{L}_t) \begin{bmatrix} e_t^0 + x_t^0 \\ e_t + z_t \end{bmatrix} \Big]. \quad (2.44) \end{aligned}$$

Expand (2.44) as follows:

$$
\begin{aligned}
\Delta J =&\mathbb{E}\Big[\sum_{t=1}^{T}\begin{bmatrix} x_t^0 \\ z_t \end{bmatrix}^{\intercal}\bar{Q}_t\begin{bmatrix} x_t^0 \\ z_t \end{bmatrix}\Big]+2\mathbb{E}\Big[\sum_{t=1}^{T}\begin{bmatrix} x_t^0 \\ z_t \end{bmatrix}^{\intercal}\bar{Q}_t\begin{bmatrix} \mathbf{0}_{d_x\times d_x} \\ \zeta_t \end{bmatrix}\Big]\\
&+\mathbb{E}\Big[\sum_{t=1}^{T}\begin{bmatrix} \mathbf{0}_{d_x\times d_x} \\ \zeta_t \end{bmatrix}^{\intercal}\bar{Q}_t\begin{bmatrix} \mathbf{0}_{d_x\times d_x} \\ \zeta_t \end{bmatrix}\Big]+\mathbb{E}\Big[\sum_{t=1}^{T}\begin{bmatrix} x_t^0 \\ z_t \end{bmatrix}^{\intercal}\bar{L}_t^{\intercal}\bar{R}_t\bar{L}_t\begin{bmatrix} x_t^0 \\ z_t \end{bmatrix}\Big]\\
&+2\mathbb{E}\Big[\sum_{t=1}^{T}\begin{bmatrix} \mathbf{0}_{d_x\times d_x} \\ \breve{L}_t\zeta_t \end{bmatrix}^{\intercal}\bar{R}_t\bar{L}_t\begin{bmatrix} x_t^0 \\ z_t \end{bmatrix}\Big]+\mathbb{E}\Big[\sum_{t=1}^{T}\begin{bmatrix} \mathbf{0}_{d_x\times d_x} \\ \breve{L}_t\zeta_t \end{bmatrix}^{\intercal}\bar{R}_t\begin{bmatrix} \mathbf{0}_{d_x\times d_x} \\ \breve{L}_t\zeta_t \end{bmatrix}\Big]\\
&-\mathbb{E}\Big[\sum_{t=1}^{T}\begin{bmatrix} e_t^0 \\ e_t \end{bmatrix}^{\intercal}(\bar{Q}_t+\bar{L}_t^{\intercal}\bar{R}_t\bar{L}_t)\begin{bmatrix} e_t^0 \\ e_t \end{bmatrix}\Big]-2\mathbb{E}\Big[\sum_{t=1}^{T}\begin{bmatrix} e_t^0 \\ e_t \end{bmatrix}^{\intercal}(\bar{Q}_t+\bar{L}_t^{\intercal}\bar{R}_t\bar{L}_t)\begin{bmatrix} x_t^0 \\ z_t \end{bmatrix}\Big]\\
&-\mathbb{E}\Big[\sum_{t=1}^{T}\begin{bmatrix} x_t^0 \\ z_t \end{bmatrix}^{\intercal}(\bar{Q}_t+\bar{L}_t^{\intercal}\bar{R}_t\bar{L}_t)\begin{bmatrix} x_t^0 \\ z_t \end{bmatrix}\Big].
\end{aligned}\tag{2.45}
$$

The second, fifth and eighth terms in the right side of (2.45) are zero from Lemma 2.2, on noting that x_t^0 and z_t are completely known under the information structures (2.3) and (2.4). This completes the proof. ■

Theorem 2.1. *Let Assumptions 2.1 and 2.2 hold. Then,*

$$\Delta J = \operatorname{Tr}\left(\begin{bmatrix} \mathbf{0}_{d_x\times d_x} & \mathbf{0}_{d_x\times d_x} & \mathbf{0}_{d_x\times d_x} \\ \mathbf{0}_{d_x\times d_x} & \operatorname{var}(\bar{x}_1) & \operatorname{var}(\bar{x}_1) \\ \mathbf{0}_{d_x\times d_x} & \operatorname{var}(\bar{x}_1) & \operatorname{var}(\bar{x}_1) \end{bmatrix} \tilde{M}_1\right) + \sum_{t=1}^{T-1} \operatorname{Tr}\left(\begin{bmatrix} \mathbf{0}_{d_x\times d_x} & \mathbf{0}_{d_x\times d_x} & \mathbf{0}_{d_x\times d_x} \\ \mathbf{0}_{d_x\times d_x} & \operatorname{var}(\bar{w}_t) & \operatorname{var}(\bar{w}_t) \\ \mathbf{0}_{d_x\times d_x} & \operatorname{var}(\bar{w}_t) & \operatorname{var}(\bar{w}_t) \end{bmatrix} \tilde{M}_{t+1}\right), \tag{2.46}$$

where $\tilde{M}_T = \tilde{Q}_T$, and $\tilde{M}_t$ is the solution of the following Lyapunov equation for any $t \in \mathbb{N}_{T-1}$:

$$\tilde{M}_t = \tilde{A}_t^T \tilde{M}_{t+1} \tilde{A}_t + \tilde{Q}_t. \tag{2.47}$$

Proof. According to Lemma 2.4, the performance discrepancy ΔJ is a quadratic function of the relative errors, and from Lemma 2.1, the relative errors have linear dynamics. Therefore, ΔJ can be regarded as the quadratic cost of an uncontrolled linear system (where there is no control action). Thus, from the standard results in linear systems [28], ΔJ can be expressed by the Lyaponuv equation (2.47) and the covariance matrices of the initial relative errors and noises $\bar{w}_t$, $t \in \mathbb{N}_{T-1}$ as:

$$\mathbb{E}\big[[e_1^0 \quad e_1 - \mu_x \quad \zeta_1 - \mu_x]^\intercal [e_1^0 \quad e_1 - \mu_x \quad \zeta_1 - \mu_x]\big] \\ = \begin{bmatrix} \mathbf{0}_{d_x \times d_x} & \mathbf{0}_{d_x \times d_x} & \mathbf{0}_{d_x \times d_x} \\ \mathbf{0}_{d_x \times d_x} & \mathrm{var}(\bar{x}_1) & \mathrm{var}(\bar{x}_1) \\ \mathbf{0}_{d_x \times d_x} & \mathrm{var}(\bar{x}_1) & \mathrm{var}(\bar{x}_1) \end{bmatrix}, \tag{2.48}$$

and

$$\mathbb{E}\big[[\mathbf{0}_{d_x \times d_x} \quad \bar{w}_t \quad \bar{w}_t]^\intercal [\mathbf{0}_{d_x \times d_x} \quad \bar{w}_t \quad \bar{w}_t]\big] \\ = \begin{bmatrix} \mathbf{0}_{d_x \times d_x} & \mathbf{0}_{d_x \times d_x} & \mathbf{0}_{d_x \times d_x} \\ \mathbf{0}_{d_x \times d_x} & \mathrm{var}(\bar{w}_t) & \mathrm{var}(\bar{w}_t) \\ \mathbf{0}_{d_x \times d_x} & \mathrm{var}(\bar{w}_t) & \mathrm{var}(\bar{w}_t) \end{bmatrix}. \tag{2.49}$$

■

Theorem 2.2. *Let Assumptions 2.1, 2.2 and 2.3 hold. Then, the strategies proposed in (2.17) and (2.18) are $\varepsilon(n)$-optimal solutions for Problem 2.1 such that*

$$|J_T(g_\epsilon^*) - J_T^*| \leq \varepsilon(n) \in \mathcal{O}(\frac{1}{n}). \tag{2.50}$$

Proof. According to Assumption 2.2,

$$\begin{aligned}
\mathrm{var}(\bar{x}_1) &= \mathrm{var}(\frac{1}{n}\sum_{i=1}^{n} x_1^i) = \frac{n\sum_x}{n^2} = \frac{\sum_x}{n}, \\
\mathrm{var}(\bar{w}_t) &= \mathrm{var}(\frac{1}{n}\sum_{i=1}^{n} w_t^i) = \frac{n\sum_w}{n^2} = \frac{\sum_w}{n}.
\end{aligned} \tag{2.51}$$

In addition, from Assumption 2.3, matrices $\tilde{A}_t$ and $\tilde{Q}_t$ given by Lemmas 2.1 and 2.4 are independent of the number of followers n, and so is $\tilde{M}_t$. Therefore, the performance discrepancy in (2.46) converges to zero at rate $\mathcal{O}(\frac{1}{n})$ according to (2.51). ■ ■

Corollary 2.1. *For the special case of a leaderless multi-agent network, let $x_{t+1}^0 = x_t^0 = \bar{x}_1$, $t \in \mathbb{N}_T$. Then, according to Theorem 2.2, strategy (2.18) steers all the followers to the initial mean $\bar{x}_1$ as n grows to infinity. In addition, if the initial mean $\bar{x}_1$ is not known, it can be replaced by its expectation, i.e., $x_{t+1}^0 = x_t^0 = \mu_x$, $t \in \mathbb{N}_T$, and the resultant strategy (2.18) steers all the followers to the initial mean consensus as n grows to infinity, due to the strong law of large numbers.*

2.4 Numerical Examples

Example 2.1. Consider a multi-agent network with one leader and 1000 followers, where the initial state of the leader is $x_1^0 = 6$ and the initial states of the followers are chosen as uniformly distributed random variables in the interval $[0, 4]$. Let the dynamics of the leader and followers be described by (2.1) and (2.2), respectively, where

$$A_t^0 = 1, \quad B_t^0 = 0.8, \quad A_t = 1, \quad B_t = 0.9, \tag{2.52}$$

$$D_t^0 = 0.1, \quad D_t = 0.05, \quad E_t = 0.15, \quad T = 40, \tag{2.53}$$

$$w_t^0 \sim \mathcal{N}(0, 0.02), \quad w_t^i \sim \mathcal{N}(0, 0.05) \quad \forall i \in \mathbb{N}_{1000}. \tag{2.54}$$

The network objective is to minimize the cost function (2.5), where

$$Q_t^0 = 1, \quad R_t^0 = 200, \quad F_t = 20, \tag{2.55}$$

$$Q_t = 2, \quad P_t = 5, \quad R_t = 100, \quad H_t = 1. \tag{2.56}$$

The leader solves the Riccati equation (2.12) to obtain gains $\bar{L}_t^{1,1}$ and $\bar{L}_t^{1,2}$, $t \in \mathbb{N}_T$, and determines its control action according to strategy (2.17) using its local state x_t^0 as well as z_t. It is to be noted that z_t is obtained at any time t in terms of x_{t-1}^0 and z_{t-1} using (2.16). In addition, for any $i \in \mathbb{N}_n$, follower i solves two Riccati equations (2.11) and (2.12) to find $\breve{L}_t$, $\bar{L}_t^{2,1}$ and $\bar{L}_t^{2,2}$, and then computes its control action based on (2.18) using its local state x_t^i, the state of the leader x_t^0, and variable z_t. The results are depicted in Figure 2.1, where the thick curve represents the state of the leader, and thin curves are the states of the followers (to avoid a cluttered figure, only 100 followers are chosen, randomly, to display their states). It can be observed from this figure that the states of all agents (which are, in fact, the position of the agents) are convergent under the

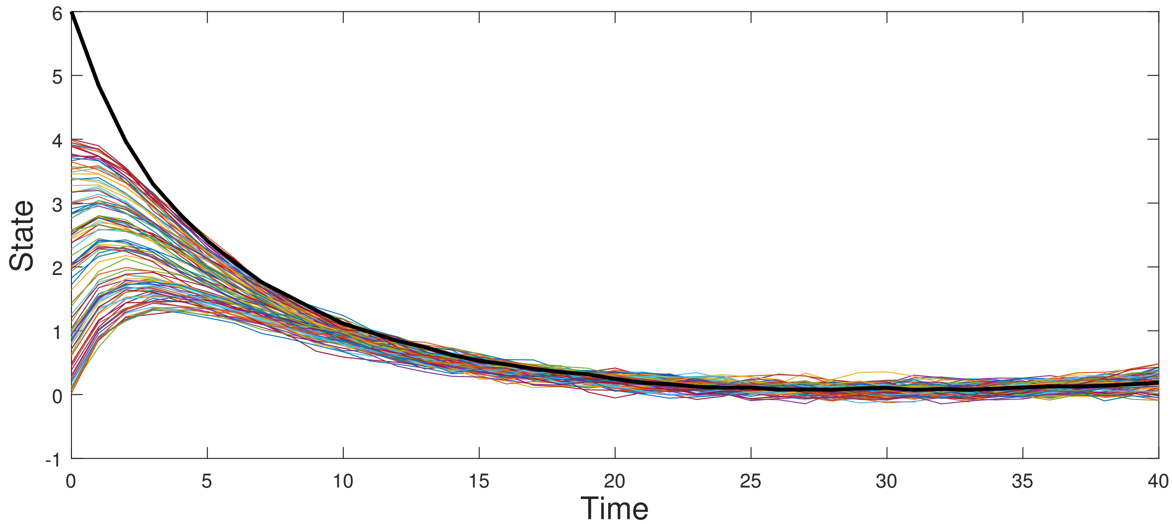


Figure 2.1: Trajectories of the leader and 100 randomly selected followers in Example 2.1

proposed control strategy, as expected, and hence consensus is achieved asymptotically.

The next example demonstrates the efficacy of the results obtained in this work, for the special case of leaderless multi-agent networks.

Example 2.2. Consider a multi-agent system consisting of 100 agents that are to track a constant reference trajectory $x_1^0 = 3$. The following parameters are used in the simulation:

$$A_t^0 = 1, \quad B_t^0 = 0, \quad D_t^0 = 0, \quad A_t = 1, \quad B_t = 0.5, \tag{2.57}$$

$$D_t = 0.05, \quad E_t = 0, \quad Q_t^0 = 0, \quad R_t^0 = 0, \tag{2.58}$$

$$Q_t = 0.1, \quad P_t = 20, \quad R_t = 100, \quad H_t = 0.5, \quad F_t = 60, \tag{2.59}$$

$$T = 40, \quad w_t^i \sim \mathcal{N}(0, 0.02) \quad \forall i \in \mathbb{N}_{100}. \tag{2.60}$$

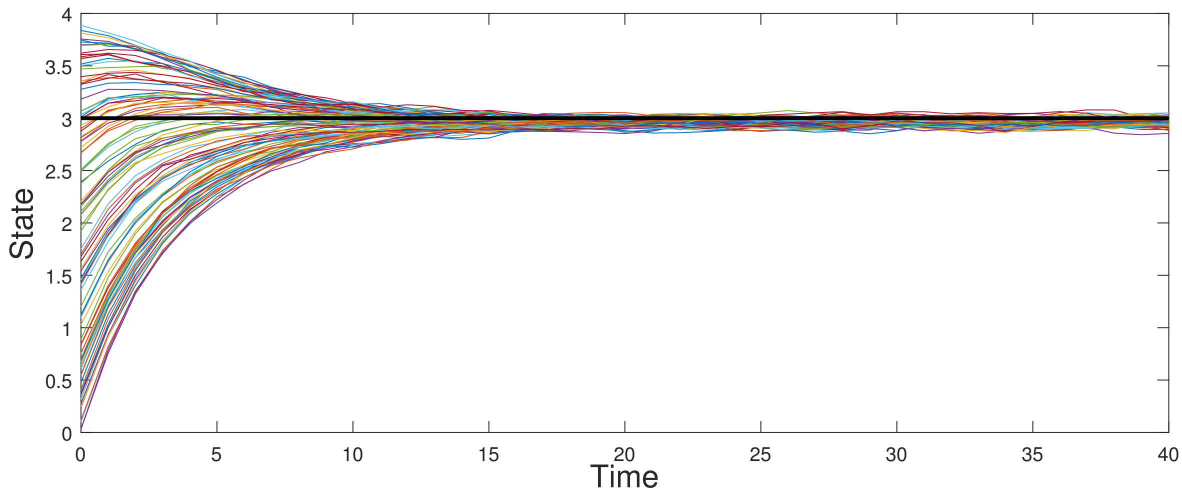


Figure 2.2: Trajectories of the followers in Example 2.2.

Similar to Example 2.1, each follower computes its control action according to (2.18). It is to be noted that in the leaderless case, the agents do not communicate as discussed in Remark 2.3. The results are given in Figure 2.2, analogously to Figure 2.1, and show that consensus is achieved as the states of all followers converge to the same value.

To compare the effect of different cost function parameters, a numerical example is presented

Example 2.3. Consider a multi-agent system with one leader and 100 followers. Assume that the state of the leader is a vector in $\mathbb{R}^2$ and that its initial is $[0 \quad 1]^\intercal$. The state of every follower, on the other hand, is a scalar, with the initial state of each one chosen as a uniformly distributed random variable in the interval $[-1, 1]$. The objective of the followers is to track the first element of the leader's state while minimizing the performance index (2.5) with the following parameters:

$$Q_t^0 = 0.1 \times \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad R_t^0 = 10^{-10} \times \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad F_t = 500, \tag{2.61}$$

$$Q_t = 0.1, \quad P_t = 50, \quad R_t = 100, \quad H_t = 10. \tag{2.62}$$

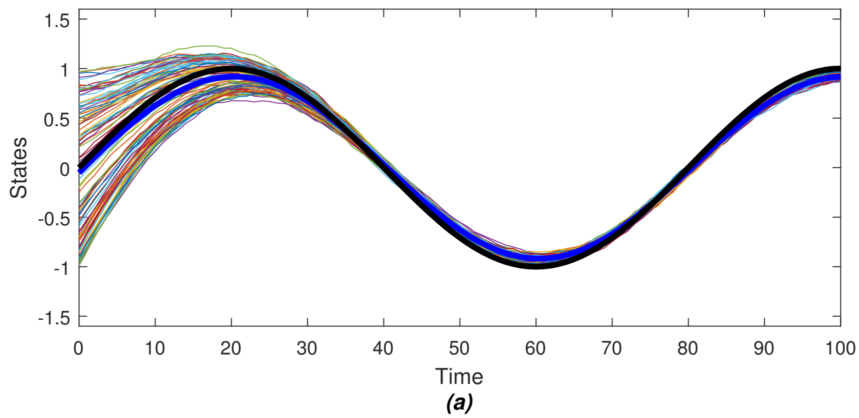
The dynamics of the leader and followers are described by (2.1) and (2.2), respectively, with the following matrices:

$$A_t^0 = \begin{bmatrix} 0.9969 & 0.0785 \\ -0.0785 & 0.9969 \end{bmatrix}, \quad B_t^0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad A_t = 1, \quad B_t = 0.1, \tag{2.63}$$

$$D_t^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad D_t = 0.01, \quad E_t = 0.02, \quad T = 100, \tag{2.64}$$

$$w_t^0 \sim \mathcal{N}(0, 0.00), \quad w_t^i \sim \mathcal{N}(0, 0.01) \quad \forall i \in \mathbb{N}_{100}. \tag{2.65}$$

The results are presented in three sets of figures. The simulations are performed by considering constant parameters in the cost function with two different values for one of the parameters H, P and F in each figure for a better comparison of the collective behavior of followers, their tendency to move towards their collective average, and the speed of convergence of the followers' states to the leader's state.



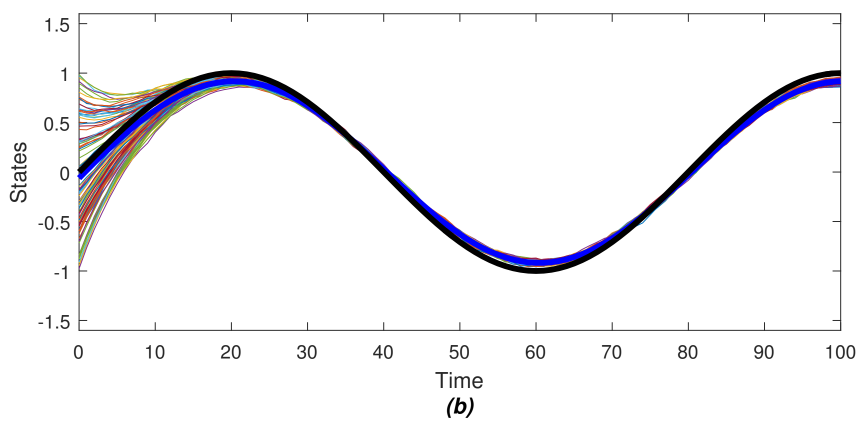


Figure 2.3: Trajectories of the leader and followers in Example 2.3 with two different values of H : $(a)H = 10$, and $(b)H = 200$.

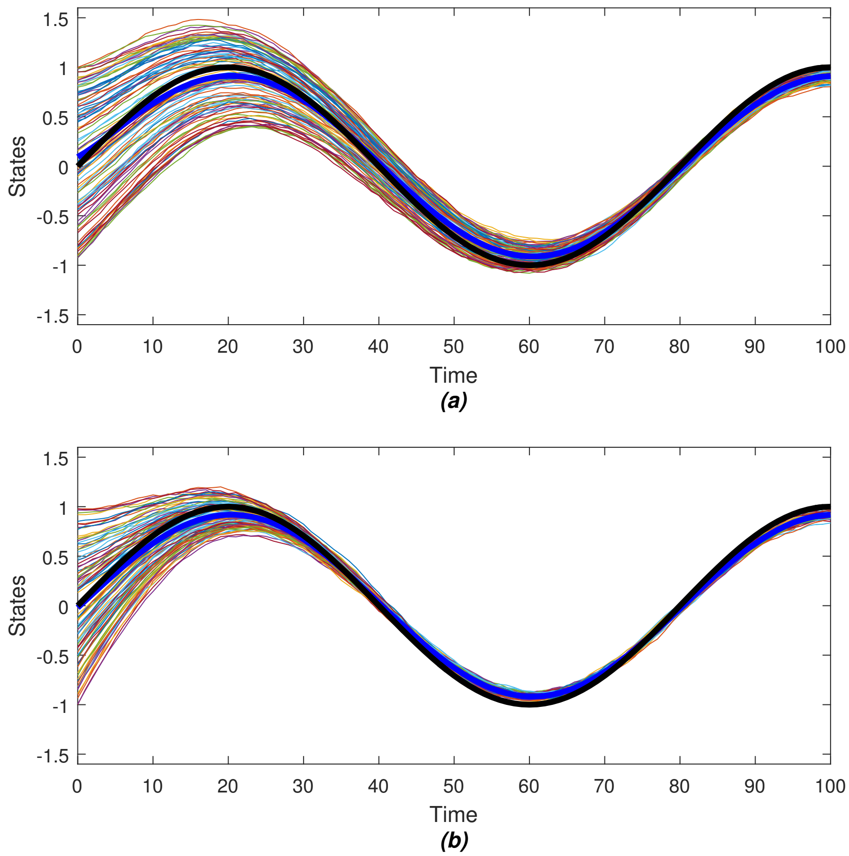


Figure 2.4: Trajectories of the leader and followers in Example 2.3 with two different values of $P : (a)P = 1$, and $(b)P = 50$.

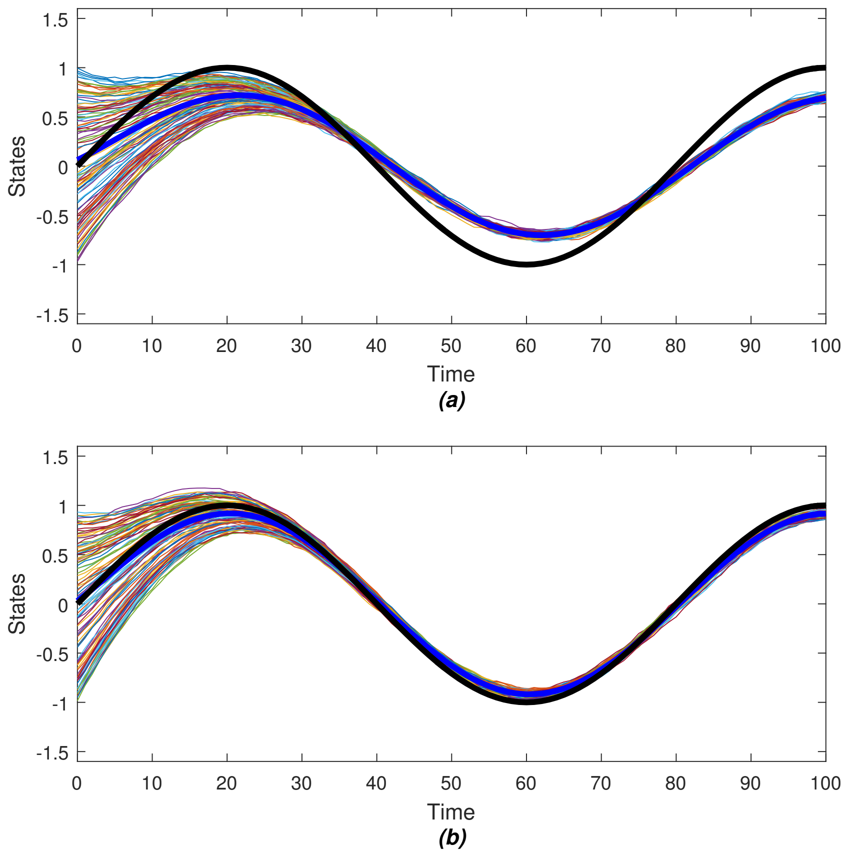


Figure 2.5: Trajectories of the leader and followers in Example 2.3 with two different values of F : $(a) F = 50$, and $(b) F = 500$.

2.5 Conclusions

A mean-field approach to the decentralized control of a leader-follower multi-agent network with a single leader is presented in this chapter, where the states of the leader and followers are coupled in the dynamics and cost. A near-optimal strategy for a non-classical information structure is proposed such that the strategy is obtained by solving two decoupled Riccati equations, where the dimension of the matrices in these equations is independent of the number of followers. This means that the proposed method is not only distributed, it is also scalable. It is shown that the proposed solution converges to the optimal strategy at a rate inversely proportional to the number of followers. The effectiveness of the results is verified by simulation, for different multi-agent settings with 1000 and 100 followers. As suggestions for future research directions, one can extend the results to the case of infinite horizon, multiple leaders, heterogeneous followers, and weighted cost functions, under standard assumptions in mean-field teams [12]. The approach is robust in the sense that the failure of a small number of followers has negligible impact on the mean-field for a network of sufficiently large population.

Chapter 3

Minimax Control in Linear Quadratic Mean-Field Teams for a Leader-Follower Network

This chapter investigates a soft-constrained minimax control problem of a leader-follower network. The network consists of one leader and an arbitrary number of followers that wish to reach consensus with minimum energy consumption in the presence of external disturbances. The leader and followers are coupled in the dynamics and cost function. Three non-classical information structures are considered: mean-field sharing, partial mean-field sharing and intermittent mean-field sharing, where the mean field refers to the aggregate state of the followers. In the mean-field sharing, every follower observes its local state, the state of the leader and the mean field; in the partial mean-field sharing, every follower observes its local state and the state of the leader, and in the intermittent mean-field sharing, the information structure switches between the above two structures. A social welfare cost function is defined, and it is shown that a unique saddle-point strategy exists which minimizes the worst-case value of the cost function under the mean-field sharing information structure. The solution is obtained by two scalable Riccati equations, which depend on a prescribed attenuation parameter, serving as a robustness factor. For the partial mean-field sharing and intermittent mean-field sharing information structures, an approximate saddle-point strategy is proposed, and its converges to the optimal saddle-point is established. Two numerical examples are provided to demonstrate the efficacy of the obtained results.

This chapter is based on following publication:

M. Baharloo, J. Arabneydi, and A. G. Aghdam, "Minimax Control in Linear Quadratic Mean-Field Teams for a Leader-Follower Network," submitted to *IEEE Control Systems Letters*.

3.1 Introduction

Recently, there has been an increasing interest in the applications of networked control systems in various engineering problems such as sensor networks [31], swarm robotics [32], unmanned aerial vehicles (UAVs) [22] and flocking [24], to name only a few. In this type of system, it is desired to achieve a global objective (such as consensus or flocking) using local control laws with limited information exchange. Many problems arise in real-world applications that are often neglected in the literature for simplicity. For instance, multi-agent networks are often subject to external disturbances, which means that a practical control strategy needs to be robust in the presence of unwanted disturbances.

Different robust control design techniques are introduced in the literature such as H_∞-control [33], risk-sensitive control [34] and minimax control [35], each of which has its own strengths and weaknesses. For example, risk-sensitive control utilizes a risk factor in order to capture the randomness of a market, that suits applications with stochastic disturbances and noises. Minimax control approach, on the other hand, models the external disturbances as an adversarial player attempting to maximize the cost of the system. In general, there are two types of formulations for the minimax control problem: (a) hard-constrained formulation, where an upper bound is set on the disturbance, and (b) soft-constrained one that penalizes the disturbance by a negative quadratic cost function [36].

There are two main challenges concerning a minimax control setting in the leader-follower problem. The first one is the computational complexity of the problem, that increases as the number of followers increases (i.e., curse of dimensionality). The second challenge is the fact that it is not always feasible to assume that the states of all followers are available, specially when the number of followers is large. In this case, a decentralized information structure is more desirable, however, it leads to a discrepancy in the followers' information. To address the above challenges, mean-field models are introduced in the literature to provide a tractable approximate solution.

The authors in [37], study a minimax mean-field type control problem in the context of social networks with a large number of homogeneous players. The proposed solution is an approximate robust mean-field equilibrium that is formulated as two coupled forward-backward partial differential equations (i.e., the Hamilton-Jacobi-Isaacs and Fokker-Planck-Kolmogorov equations). In [38], a minimax mean-field game problem is considered, and an approximate robust Nash equilibrium is obtained using two coupled forward-backward stochastic differential equations. In the minimax control problem with social cost function [39], the variational derivation and person-by-person optimality principle are employed to derive an approximate saddle-point in terms of two coupled forward-backward stochastic differential equations. The authors in [40] consider a leader-follower setting and propose an approximate robust Nash equilibrium. The method is to first solve a minimax control problem for the leader, irrespective of followers, and then solve a minimax control problem for the followers that wish to track a convex combination of the mean-field and the state of the leader in the presence of external disturbances.

In this chapter, a minimax control problem is investigated for a leader-follower network with an arbitrary number of followers. Unlike the above articles that provide an approximate solution under partial mean-field sharing information structure, we explicitly obtain the unique saddle-point strategy under mean-field sharing information structure and propose an approximate solution under partial mean-field sharing. The salient feature of the proposed solutions is that they are identified by two scalable Riccati equations that are not in the form of forward-backward coupled equations.

The remainder of the chapter is organized as follows. The problem is defined and formulated in Section 3.2. The main results are subsequently presented in Section 3.3 along with the required assumptions. In Section 3.4, two numerical examples are demonstrated, and finally in Section 3.5 some concluding remarks are provided.

3.2 Problem Formulation

3.2.1 Notation

Throughout this chapter, $\mathbb{R}$ and $\mathbb{N}$ denote, respectively, the sets of real and natural numbers. For any $k \in \mathbb{N}$, $\mathbb{N}_k$ denotes the finite set $\{1, \ldots, k\}$, and $x_{1:k}$ is short-hand notation for $\{x_1, \ldots, x_k\}$. $\mathbb{E}(\cdot)$ is the expectation of an event, $\operatorname{Cov}(\cdot)$ is the covariance matrix of a random vector, $\operatorname{Tr}(\cdot)$ is the trace of a matrix, and $\mathbf{I}$ and $\mathbf{0}$ are, respectively, identity and zero matrices.

3.2.2 Model

Consider a multi-agent system consisting of one leader and $n \in \mathbb{N}$ homogeneous followers. Let $x_t^0 \in \mathbb{R}^{\ell_x}$, $u_t^0 \in \mathbb{R}^{\ell_u}$, $d_t^0 \in \mathbb{R}^{\ell_x}$ and $w_t^0 \in \mathbb{R}^{\ell_x}$ denote, respectively, the state, action, disturbance and noise of the leader at time $t \in \mathbb{N}$, where $\ell_x, \ell_u \in \mathbb{N}$. Analogously, denote by $x_t^i \in \mathbb{R}^{\ell_x}$, $u_t^i \in \mathbb{R}^{\ell_u}$, $d_t^i \in \mathbb{R}^{\ell_x}$ and $w_t^i \in \mathbb{R}^{\ell_x}$, the state, action, disturbance and noise of follower $i \in \mathbb{N}_n$ at time $t \in \mathbb{N}$. In addition, define the aggregate state and aggregate action of followers as follows:

$$\bar{x}_t := \frac{1}{n}\sum_{i=1}^{n} x_t^i, \quad \bar{u}_t := \frac{1}{n}\sum_{i=1}^{n} u_t^i. \tag{3.1}$$

The dynamics of the leader at time $t \in \mathbb{N}$ is influenced by the aggregate state $\bar{x}_t$, the disturbance signal d_t^0 and noise w_t^0, i.e.,

$$x_{t+1}^0 = A_t^0 x_t^0 + B_t^0 u_t^0 + S_t^0 \bar{x}_t + d_t^0 + w_t^0, \tag{3.2}$$

where A_t^0, B_t^0 and S_t^0 are matrices of appropriate dimensions. Furthermore, the dynamics of follower i at time t is affected by the state of the leader x_t^0, aggregate state $\bar{x}_t$, local disturbance d_t^i and local noise w_t^i as shown below:

$$x_{t+1}^i = A_t x_t^i + B_t u_t^i + S_t \bar{x}_t + E_t x_t^0 + d_t^i + w_t^i, \quad i \in \mathbb{N}_n,\ t \in \mathbb{N}, \tag{3.3}$$

where A_t, B_t, S_t and E_t are matrices of proper dimensions. Let $T \in \mathbb{N}$ denote the control horizon, and assume that the primitive random variables

$$\{x_1^0, \{x_1^i\}_{i\in\mathbb{N}_n}, w_1^0, \{w_1^i\}_{i\in\mathbb{N}_n}, \dots, w_T^0, \{w_T^i\}_{i\in\mathbb{N}_n}\}$$

are mutually independent. In addition, it is assumed that the local noises of followers and the noise of the leader have zero mean and finite covariance matrices.

3.2.3 Admissible strategies

To be consistent with the terminology of mean-field teams [12], the aggregate state of followers is called *mean field*. It is to be noted that the term mean field has a slightly different meaning in mean-field games, where it refers to the aggregate state of the infinite population (as opposed to a finite population) of followers.

In this chapter, we consider three non-classical information structures: mean-field sharing (MFS), partial mean-field sharing (PMFS) and intermittent mean-field sharing (IMFS). In the MFS information structure, the leader has access to its local state as well as the mean-field at any time t, i.e.,

$$u_t^0 = g_t^0(x_t^0, \bar{x}_t), \tag{3.4}$$

where $g_t^0 : \mathbb{R}^{2\ell_x} \to \mathbb{R}^{\ell_u}$. Furthermore, each follower $i \in \mathbb{N}_n$ has access to its local state as well as the state of the leader and the mean-field at time t, i.e.,

$$u_t^i = g_t^i(x_t^i, \bar{x}_t, x_t^0), \tag{3.5}$$

where $g_t^i : \mathbb{R}^{3\ell_x} \to \mathbb{R}^{\ell_u}$. In the PMFS information structure, the mean-field is not observed, i.e.,

$$u_t^0 = g_t^0(x_t^0), \quad u_t^i = g_t^i(x_t^i, x_t^0), \quad \forall i \in \mathbb{N}_n. \tag{3.6}$$

The third information structure is IMFS, defined as an intermittent version of MFS and PMFS, i.e.,

$$u_t^0 = g_t^0(x_t^0, z_t), \quad u_t^i = g_t^i(x_t^i, z_t, x_t^0), \quad \forall i \in \mathbb{N}_n, \tag{3.7}$$

where $z_t := \bar{x}_t$ during the time when the information structure is MFS and $z_t := \mathbf{0}_{\ell_x \times 1}$ during the time when the information structure is PMFS. In practice, IMFS information structure is useful when the number of followers is neither that small (so that the mean-field can be shared at each time instant) nor is very large (such that the strong law of large numbers can be applied to the mean-field). In such a case, it is feasible to obtain the mean-field intermittently such that at some time instants the information structure is MFS while at some others it is PMFS.

The set of all control laws $\mathbf{g} := \{g_{1:T}^0, g_{1:T}^1, \ldots, g_{1:T}^n\}$ is called the strategy of the network.

3.2.4 Problem statement

Let the set $\mathbf{d}$ be defined as $\{d^0_{1:T}, \{d^i_{1:T}\}_{i\in\mathbb{N}_n}\}$, and $\gamma > 0$ be a given attenuation parameter. Then the cost function of the system is defined as follows:

$$\begin{aligned} J^\gamma_n(\mathbf{g},\mathbf{d}) = \mathbb{E}(\sum_{t=1}^{T} \Big[\frac{1}{n}\sum_{i=1}^{n} \big[(x^i_t)^\intercal Q_t x^i_t + (u^i_t)^\intercal R_t u^i_t - \gamma^2 (d^i_t)^\intercal d^i_t\big] \\ + (x^0_t)^\intercal Q^0_t x^0_t + (u^0_t)^\intercal R^0_t u^0_t - \gamma^2 (d^0_t)^\intercal d^0_t \\ + (\bar{x}_t - x^0_t)^\intercal F_t (\bar{x}_t - x^0_t) + \bar{x}_t^\intercal P_t \bar{x}_t + \bar{u}_t^\intercal H_t \bar{u}_t\big]\Big]), \end{aligned} \tag{3.8}$$

where $Q_t, Q^0_t, R_t, R^0_t, F_t, P_t$ and H_t are symmetric matrices of appropriate dimensions. Note that the value of γ determines the relative importance of reaching consensus and rejecting disturbance.

Problem 3.1. *Find the saddle-point strategy* $\mathbf{g}$ *under mean-field sharing information structure such that*

$$J^{\gamma,*}_n = \inf_{\mathbf{g}} \sup_{\mathbf{d}} J^\gamma_n(\mathbf{g},\mathbf{d}). \tag{3.9}$$

Problem 3.2. *Find an approximate saddle-point strategy* $\mathbf{g}_\varepsilon$ *under intermittent mean-field sharing information structure such that*

$$|\sup_{\mathbf{d}} J^\gamma_n(\mathbf{g}_\varepsilon,\mathbf{d}) - J^{\gamma,*}_n| \le \varepsilon(n), \tag{3.10}$$

where $\varepsilon(n) \ge 0$ *and* $\lim_{n\to\infty} \varepsilon(n) = 0$.

Remark 3.1. Note that IMFS information structure can be considered as a generalization of the PMFS information structure; thus, Problem 3.2 encompasses the problem of finding a near-optimal solution with PMFS structure.

Remark 3.2. It is to be noted that a leaderless network can be considered as a special case of the leader-follower system.

3.3 Main Results

Define the following matrices at any time $t \in \mathbb{N}_T$:

$$\bar{A}_t := \begin{bmatrix} A_t^0 & S_t^0 \\ E_t & A_t + S_t \end{bmatrix}, \quad \bar{B}_t := \begin{bmatrix} B_t^0 & \mathbf{0}_{\ell_x \times \ell_u} \\ \mathbf{0}_{\ell_x \times \ell_u} & B_t \end{bmatrix}, \tag{3.11}$$

$$\bar{Q}_t := \begin{bmatrix} Q_t^0 + F_t & -F_t \\ -F_t & Q_t + P_t + F_t \end{bmatrix}, \bar{R}_t := \begin{bmatrix} R_t^0 & \mathbf{0}_{\ell_u \times \ell_u} \\ \mathbf{0}_{\ell_u \times \ell_u} & H_t + R_t \end{bmatrix}. \tag{3.12}$$

Assumption 3.1. *For any time $t \in \mathbb{N}_T$, matrices Q_t and $\bar{Q}_t$ are positive semi-definite and R_t and $\bar{R}_t$ are positive definite.*

It will be shown later that Problem 3.1 under Assumption 3.1 can be cast as a strictly convex optimization problem with respect to the control actions of the leader and followers, and a strictly concave optimization problem with respect to the disturbances. Using the transformation technique introduced in [12], define the following variables:

$\breve{x}_t^i := x_t^i - \bar{x}_t$, $\breve{u}_t^i := u_t^i - \bar{u}_t$, $\breve{d}_t^i := d_t^i - \bar{d}_t$ and $\breve{w}_t^i := w_t^i - \bar{w}_t$, where $\bar{d}_t := \frac{1}{n}\sum_{i=1}^n d_t^i$ and $\bar{w}_t := \frac{1}{n}\sum_{i=1}^n w_t^i$. It follows from (3.3) that:

$$\begin{aligned}
\bar{x}_{t+1} &= (A_t + S_t)\bar{x}_t + B_t\bar{u}_t + E_t x_t^0 + \bar{d}_t + \bar{w}_t,\\
\breve{x}_{t+1}^i &= A_t\breve{x}_t^i + B_t\breve{u}_t^i + \breve{d}_t^i + \breve{w}_t^i.
\end{aligned} \tag{3.13}$$

Note that $\frac{1}{n}\sum_{i=1}^n \breve{x}_t^i = \mathbf{0}_{\ell_x\times 1}$, $\frac{1}{n}\sum_{i=1}^n \breve{u}_t^i = \mathbf{0}_{\ell_u\times 1}$, $\frac{1}{n}\sum_{i=1}^n \breve{d}_t^i = \mathbf{0}_{\ell_x\times 1}$ and $\frac{1}{n}\sum_{i=1}^n \breve{w}_t^i = \mathbf{0}_{\ell_x\times 1}$.

Rewrite the cost function defined in (3.8) in terms of the new variables as:

$$\begin{aligned}
J_n^\gamma(\mathbf{g},\mathbf{d}) = \mathbb{E}\Big(\sum_{t=1}^T \Big[\frac{1}{n}\sum_{i=1}^n (\breve{x}_t^i)^\intercal Q_t \breve{x}_t^i + (\breve{u}_t^i)^\intercal R_t \breve{u}_t^i - \gamma^2 (\breve{d}_t^i)^\intercal \breve{d}_t^i\Big]\\
+ \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix}^\intercal \bar{Q}_t \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix} + \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix}^\intercal \bar{R}_t \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix} - \gamma^2 \begin{bmatrix} d_t^0 \\ \bar{d}_t \end{bmatrix}^\intercal \begin{bmatrix} d_t^0 \\ \bar{d}_t \end{bmatrix}\Big).
\end{aligned} \tag{3.14}$$

At any time t, define the following augmented vectors:

$$\mathbf{x}_t := [\breve{x}_t^{1\intercal}, \ldots, \breve{x}_t^{n\intercal}, x_t^{0\intercal}, \bar{x}_t^\intercal]^\intercal, \tag{3.15.a}$$

$$\mathbf{u}_t := [\breve{u}_t^{1\intercal}, \ldots, \breve{u}_t^{n\intercal}, u_t^{0\intercal}, \bar{u}_t^\intercal]^\intercal, \tag{3.15.b}$$

$$\mathbf{d}_t := [\breve{d}_t^{1\intercal}, \ldots, \breve{d}_t^{n\intercal}, d_t^{0\intercal}, \bar{d}_t^\intercal]^\intercal. \tag{3.15.c}$$

Suppose for now that $\mathbf{x}_t$ is known, and solve the corresponding Isaacs' equation according to [33, Theorem 3.2]. In particular, the cost-to-go function at terminal time

T is:

$$V_T(\mathbf{x}_T) \quad = \quad \frac{1}{n}\sum_{i=1}^{n}(\breve{x}_T^i)^\intercal \breve{M}_T \breve{x}_T^i \ + \begin{bmatrix} x_T^0 \\ \\ \bar{x}_T \end{bmatrix}^\intercal \bar{M}_T \begin{bmatrix} x_T^0 \\ \\ \bar{x}_T \end{bmatrix} \ + \ \frac{1}{n}\sum_{i=1}^{n}\breve{c}_T^i \ + \ \bar{c}_T, \quad (3.16)$$

where $\breve{M}_T := Q_T$, $\bar{M}_T := \bar{Q}_T$, $\bar{c}_T := 0$ and $\breve{c}_T^i := 0, \forall i \in \mathbb{N}_n$. Suppose that the cost-to-go function takes the following form at time $t+1$:

$$\begin{aligned} V_{t+1}(\mathbf{x}_{t+1}) = \frac{1}{n}\sum_{i=1}^{n}(\breve{x}_{t+1}^i)^\intercal \breve{M}_{t+1} \breve{x}_{t+1}^i \\ + \begin{bmatrix} x_{t+1}^0 \\ \\ \bar{x}_{t+1} \end{bmatrix}^\intercal \bar{M}_{t+1} \begin{bmatrix} x_{t+1}^0 \\ \\ \bar{x}_{t+1} \end{bmatrix} + \frac{1}{n}\sum_{i=1}^{n}\breve{c}_{t+1}^i + \bar{c}_{t+1}, \end{aligned} \quad (3.17)$$

where

$$\begin{cases} \breve{M}_t := Q_t + A_t \breve{M}_{t+1} \breve{\Delta}_t^{-1} A_t^\intercal, \\ \bar{M}_t := \bar{Q}_t + \bar{A}_t \bar{M}_{t+1} \bar{\Delta}_t^{-1} \bar{A}_t^\intercal, \\ \breve{\Delta}_t := \mathbf{I}_{\ell_x \times \ell_x} + B_t R_t^{-1} B_t^\intercal \breve{M}_{t+1} - \gamma^{-2} \breve{M}_{t+1}, \\ \bar{\Delta}_t := \mathbf{I}_{2\ell_x \times 2\ell_x} + \bar{B}_t \bar{R}_t^{-1} \bar{B}_t^\intercal \bar{M}_{t+1} - \gamma^{-2} \bar{M}_{t+1}, \\ \breve{c}_t^i := \breve{c}_{t+1}^i + \mathrm{Tr}(\breve{M}_{t+1} \operatorname{Cov} \breve{w}_t^i), \\ \bar{c}_t := \bar{c}_{t+1} + \mathrm{Tr}(\bar{M}_{t+1} \operatorname{Cov}([w_t^0, \bar{w}_t])). \end{cases} \quad (3.18)$$

It is now desired to show that (3.17) holds for time t as well. It follows from Isaacs' equation that:

$$
\begin{aligned}
V_t(\mathbf{x}_t) = \sup_{\mathbf{d}_t} \inf_{\mathbf{u}_t} (\frac{1}{n} \sum_{i=1}^{n} \big[(\breve{x}_t^i)^\intercal Q_t \breve{x}_t^i + (\breve{u}_t^i)^\intercal R_t \breve{u}_t^i - \gamma^2 (\breve{d}_t^i)^\intercal \breve{d}_t^i \big] \\
+ \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix}^\intercal \bar{Q}_t \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix} + \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix}^\intercal \bar{R}_t \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix} - \gamma^2 \begin{bmatrix} d_t^0 \\ \bar{d}_t \end{bmatrix}^\intercal \begin{bmatrix} d_t^0 \\ \bar{d}_t \end{bmatrix} \\
+ \mathbb{E}\big[V_{t+1}(\mathbf{x}_{t+1}) \mid \mathbf{x}_t, \mathbf{u}_t, \mathbf{d}_t \big]).
\end{aligned} \tag{3.19}
$$

From (3.13), (3.17) and (3.19), one arrives at:

$$
\begin{aligned}
V_t(\mathbf{x}_t) = \sup_{\mathbf{d}_t} \inf_{\mathbf{u}_t} (\frac{1}{n} \sum_{i=1}^{n} \big[(\breve{x}_t^i)^\intercal Q_t \breve{x}_t^i + (\breve{u}_t^i)^\intercal R_t \breve{u}_t^i - \gamma^2 (\breve{d}_t^i)^\intercal \breve{d}_t^i \big] \\
+ \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix}^\intercal \bar{Q}_t \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix} + \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix}^\intercal \bar{R}_t \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix} - \gamma^2 \begin{bmatrix} d_t^0 \\ \bar{d}_t \end{bmatrix}^\intercal \begin{bmatrix} d_t^0 \\ \bar{d}_t \end{bmatrix} \\
+ \mathbb{E}[\frac{1}{n} \sum_{i=1}^{n} \big[(A_t \breve{x}_t^i + B_t \breve{u}_t^i + \breve{d}_t^i + \breve{w}_t^i)^\intercal \breve{M}_{t+1} \\
\times (A_t \breve{x}_t^i + B_t \breve{u}_t^i + \breve{d}_t^i + \breve{w}_t^i) \big] + \frac{1}{n} \sum_{i=1}^{n} \breve{c}_{t+1}^i + \bar{c}_{t+1} \\
+ (\bar{A}_t \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix} + \bar{B}_t \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix} + \begin{bmatrix} d_t^0 \\ \bar{d}_t \end{bmatrix} + \begin{bmatrix} w_t^0 \\ \bar{w}_t \end{bmatrix})^\intercal \bar{M}_{t+1} \\
\times (\bar{A}_t \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix} + \bar{B}_t \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix} + \begin{bmatrix} d_t^0 \\ \bar{d}_t \end{bmatrix} + \begin{bmatrix} w_t^0 \\ \bar{w}_t \end{bmatrix})]).
\end{aligned} \tag{3.20}
$$

This yields:

$$
\begin{aligned}
V_t(\mathbf{x}_t) = &\sup_{\mathbf{d}_t}\inf_{\mathbf{u}_t}\Big(\frac{1}{n}\sum_{i=1}^{n}\left[(\breve{x}_t^i)^\intercal Q_t \breve{x}_t^i + (\breve{u}_t^i)^\intercal R_t \breve{u}_t^i - \gamma^2(\breve{d}_t^i)^\intercal \breve{d}_t^i\right] \\
&+ \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix}^\intercal \bar{Q}_t \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix} + \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix}^\intercal \bar{R}_t \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix} - \gamma^2 \begin{bmatrix} d_t^0 \\ \bar{d}_t \end{bmatrix}^\intercal \begin{bmatrix} d_t^0 \\ \bar{d}_t \end{bmatrix} \\
&+ \frac{1}{n}\sum_{i=1}^{n}[(A_t\breve{x}_t^i + B_t\breve{u}_t^i)^\intercal \breve{M}_{t+1}(A_t\breve{x}_t^i + B_t\breve{u}_t^i) + (\breve{d}_t^i)^\intercal \breve{M}_{t+1}\breve{d}_t^i \\
&+ 2(\breve{d}_t^i)^\intercal \breve{M}_{t+1}(A_t\breve{x}_t^i + B_t\breve{u}_t^i) + 2\breve{w}_t^i \breve{M}_{t+1}(A_t\breve{x}_t^i + B_t\breve{u}_t^i + \breve{d}_t^i) \\
&+ 2(\breve{d}_t^i)^\intercal \breve{M}_{t+1}\breve{w}_t^i + \mathrm{Tr}(\breve{M}_{t+1}\,\mathrm{Cov}(\breve{w}_t^i))] + \begin{bmatrix} d_t^0 \\ \bar{d}_t \end{bmatrix}^\intercal \bar{M}_{t+1} \begin{bmatrix} d_t^0 \\ \bar{d}_t \end{bmatrix} \\
&+ (\bar{A}_t \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix} + \bar{B}_t \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix})^\intercal \bar{M}_{t+1}(\bar{A}_t \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix} + \bar{B}_t \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix}) \\
&+ 2 \begin{bmatrix} d_t^0 \\ \bar{d}_t \end{bmatrix}^\intercal \bar{M}_{t+1}(\bar{A}_t \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix} + \bar{B}_t \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix}) + \bar{c}_{t+1} \\
&+ 2 \begin{bmatrix} \breve{w}_t^0 \\ \bar{w}_t \end{bmatrix}^\intercal \bar{M}_{t+1}(\bar{A}_t \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix} + \bar{B}_t \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix}) + \frac{1}{n}\sum_{i=1}^{n}\breve{c}_{t+1}^i \\
&+ 2 \begin{bmatrix} d_t^0 \\ \bar{d}_t \end{bmatrix}^\intercal \bar{M}_{t+1} \begin{bmatrix} \breve{w}_t^0 \\ \bar{w}_t \end{bmatrix} + \mathrm{Tr}(\bar{M}_{t+1}\,\mathrm{Cov}([w_t^0, \bar{w}_t]))\Big).
\end{aligned} \tag{3.21}
$$

Given any disturbance vector $\mathbf{d}_t$, we now compute the gradient vector with respect to

$\mathbf{u}_t$ and set each component to zero in order to obtain the following $n+1$ equations:

$$2(\breve{u}_t^i)^\intercal R_t + 2(\breve{u}_t^i)^\intercal B_t^\intercal \breve{M}_{t+1} B_t + 2(\breve{x}_t^i)^\intercal A_t^\intercal \breve{M}_{t+1} B_t + 2(\breve{d}_t^i)^\intercal \breve{M}_{t+1} B_t = 0, \quad \forall i \in \mathbb{N}_n, \quad (3.22)$$

and

$$2 \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix}^\intercal \bar{R}_t + 2 \begin{bmatrix} u_t^0 \\ \bar{u}_t \end{bmatrix}^\intercal \bar{B}_t^\intercal \bar{M}_{t+1} \bar{B}_t + 2 \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix}^\intercal \bar{A}_t^\intercal \bar{M}_{t+1} \bar{B}_t + 2 \begin{bmatrix} d_t^0 \\ \bar{d}_t \end{bmatrix}^\intercal \bar{M}_{t+1} \bar{B}_t = 0. \quad (3.23)$$

Consequently, the optimal actions of the leader and followers can be obtained as follows:

$$\breve{u}_t^{i*} = -(R_t + B_t^\intercal \breve{M}_{t+1} B_t)^{-1} (B_t^\intercal \breve{M}_{t+1} A_t) \breve{x}_t^i - (R_t + B_t^\intercal \breve{M}_{t+1} B_t)^{-1} (B_t^\intercal \breve{M}_{t+1}) \breve{d}_t^i, \quad (3.24)$$

and

$$\begin{bmatrix} u_t^{0,*} \\ \bar{u}_t^* \end{bmatrix} = -(\bar{R}_t + \bar{B}_t^\intercal \bar{M}_{t+1} \bar{B}_t)^{-1} (\bar{B}_t^\intercal \bar{M}_{t+1} \bar{A}_t) \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix} - (\bar{R}_t + \bar{B}_t^\intercal \bar{M}_{t+1} \bar{B}_t)^{-1} (\bar{B}_t^\intercal \bar{M}_{t+1}) \begin{bmatrix} d_t^0 \\ \bar{d}_t \end{bmatrix}. \quad (3.25)$$

It is observed that the Hessian matrix is diagonal, with matrices: $R_t + B_t^\intercal \breve{M}_{t+1} B_t$ and $\bar{R}_t +$

$\bar{B}_t^\intercal \bar{M}_{t+1} \bar{B}_t$ as its diagonal terms that are positive definite according to Assumption 3.1; hence, the cost function is strictly convex in the newly defined control actions (3.15.b). By incorporating the optimal strategies (3.24) and (3.25) into (3.21), computing the gradient vector with respect to $\mathbf{d}_t$ and setting each component to zero, one arrives at the following $n+1$ equations:

$$
\begin{aligned}
&- \breve{M}_{t+1} B_t (R_t + B_t^\intercal \breve{M}_{t+1} B_t)^{-1} B_t^\intercal \breve{M}_{t+1} \breve{d}_t^i \\
&\qquad - \breve{M}_{t+1} B_t (R_t + B_t^\intercal \breve{M}_{t+1} B_t)^{-1} B_t^\intercal \breve{M}_{t+1} A_t \breve{x}_t^i \\
&\qquad\qquad - \gamma^2 \breve{d}_t^i + \breve{M}_{t+1} \breve{d}_t^i + \breve{M}_{t+1} A_t \breve{x}_t^i = 0, \quad \forall i \in \mathbb{N}_n, \quad (3.26)
\end{aligned}
$$

and

$$
\begin{aligned}
&- \bar{M}_{t+1} \bar{B}_t (\bar{R}_t + \bar{B}_t^\intercal \bar{M}_{t+1} \bar{B}_t)^{-1} \bar{B}_t^\intercal \bar{M}_{t+1} \begin{bmatrix} d_t^0 \\ \bar{d}_t \end{bmatrix} \\
&\qquad - \bar{M}_{t+1} \bar{B}_t (\bar{R}_t + \bar{B}_t^\intercal \bar{M}_{t+1} \bar{B}_t)^{-1} \bar{B}_t^\intercal \bar{M}_{t+1} \bar{A}_t \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix} \\
&\qquad\qquad - \gamma^2 \begin{bmatrix} d_t^0 \\ \bar{d}_t \end{bmatrix} + \bar{M}_{t+1} \begin{bmatrix} d_t^0 \\ \bar{d}_t \end{bmatrix} + \bar{M}_{t+1} \bar{A}_t \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix} = 0. \quad (3.27)
\end{aligned}
$$

After some manipulations described in [41, page 248], the worst-case disturbances can be obtained as:

$$
\breve{d}_t^{i*} = \gamma^{-2} \breve{M}_{t+1} \breve{\Delta}_t^{-1} A_t \breve{x}_t^i, \tag{3.28}
$$

and

$$\begin{bmatrix} d_t^{0,*} \\ \bar{d}_t^{*} \end{bmatrix} = \gamma^{-2} \bar{M}_{t+1} \bar{\Delta}_t^{-1} \bar{A}_t \begin{bmatrix} x_t^0 \\ \bar{x}_t \end{bmatrix}, \tag{3.29}$$

where $\breve{\Delta}_t$ and $\bar{\Delta}_t$ are given by (3.18). In addition, we compute the Hessian matrix which is diagonal with matrices: $\gamma^2 \mathbf{I}_{\ell_x \times \ell_x} - \breve{M}_{t+1}$ and $\gamma^2 \mathbf{I}_{2\ell_x \times 2\ell_x} - \bar{M}_{t+1}$ on its diagonal. Therefore, if these matrices are positive definite, it is concluded that the cost function is strictly concave with respect to disturbances. The recursion (3.18) is finally obtained by incorporating the worst-case disturbances (3.28) and (3.29) into the optimal strategies (3.24) and (3.25) and comparing the expressions (3.17) and (3.19) at times $t+1$ and t, respectively, which leads to the following theorem.

Theorem 3.1. *Let Assumption 3.1 hold. Then:*

1. *Given any $\gamma > 0$, Problem 3.1 admits a unique feedback saddle-point solution if matrices $\gamma^2 \mathbf{I}_{\ell_x \times \ell_x} - \breve{M}_{t+1}$ and $\gamma^2 \mathbf{I}_{2\ell_x \times 2\ell_x} - \bar{M}_{t+1}$ are positive definite for every $t \in \mathbb{N}_{T-1}$, where $\breve{M}_{t+1}$ and $\bar{M}_{t+1}$ are given by (3.18).*

2. *The saddle-point is described by:*

$$u_t^0 = \bar{L}_t^{1,1} x_t^0 + \bar{L}_t^{1,2} \bar{x}_t, \tag{3.30}$$

and for every follower:

$$u_t^i = \breve{L}_t x_t^i + \bar{L}_t^{2,1} x_t^0 + (\bar{L}_t^{2,2} - \breve{L}_t) \bar{x}_t, \quad i \in \mathbb{N}_n, \tag{3.31}$$

where $\begin{bmatrix} \bar{L}_t^{1,1} & \bar{L}_t^{1,2} \\ \bar{L}_t^{2,1} & \bar{L}_t^{2,2} \end{bmatrix} := -\bar{B}_t \bar{M}_{t+1} \bar{\Delta}_t^{-1} \bar{A}_t$ *and* $\breve{L}_t := -B_t \breve{M}_{t+1} \breve{\Delta}_t^{-1} A_t$.

3. The optimal cost function is given by:

$$J_n^{\gamma,*} \quad = \quad \frac{1}{n} \sum_{i=1}^{n} [\mathrm{Tr}(\breve{M}_1 \operatorname{Cov}(\breve{x}_1^i)) \; + \; \breve{c}_1^i] \; + \; \mathrm{Tr}(\bar{M}_1 \operatorname{Cov}([x_1^0, \bar{x}_1])) \; + \; \bar{c}_1. \quad (3.32)$$

3.3.1 Solution of Problem 3.2

We impose the following two assumptions on the model.

Assumption 3.2. *The initial states and local noises of the followers are i.i.d. random variables and independent of those of the leader.*

Assumption 3.3. *All matrices in the dynamics* (3.2) *and* (3.3) *and cost function* (3.8) *as well as the covariance matrices are independent of the number of followers.*

Let $\hat{m}_1$ be the expected value of the initial states of the followers, and $\hat{m}_t$ denote an estimate of the mean-field $\bar{x}_t$ at time t such that if the information structure is PMFS at time t:

$$\hat{m}_{t+1} \quad := \quad (A_t \; + \; S_t \; + \; B_t \bar{L}_t^{2,2}) \hat{m}_t \; + \; (B_t \bar{L}_t^{2,1} \; + \; E_t) x_t^0 \; + \; \bar{d}_t, \quad (3.33)$$

and if the information structure is MFS:

$$\hat{m}_{t+1} := \bar{x}_{t+1}. \tag{3.34}$$

Under Assumptions 3.2 and 3.3, it can be shown that $\hat{m}_{t+1}$ almost surely converges to $\bar{x}_{t+1}$ at every time instant almost surely due the strong law of large numbers, on noting that the dynamics of the mean-field under the saddle-point strategy is:

$$\bar{x}_{t+1} = (A_t + S_t + B_t \bar{L}_t^{2,2})\bar{x}_t + (B_t \bar{L}_t^{2,1} + E_t)x_t^0 + \bar{d}_t + \bar{w}_t. \tag{3.35}$$

We now replace the mean-field $\bar{x}_t$ in the saddle-point strategy of Theorem 3.1 with the estimate $\hat{m}_t$ (that is measurable with respect to IMFS information structure) to construct the following approximate saddle-point strategy:

$$v_t^0 = \bar{L}_t^{1,1} x_t^0 + \bar{L}_t^{1,2} \hat{m}_t, \tag{3.36}$$

and for every follower:

$$v_t^i = \breve{L}_t x_t^i + \bar{L}_t^{2,1} x_t^0 + (\bar{L}_t^{2,2} - \breve{L}_t)\hat{m}_t, \quad i \in \mathbb{N}_n. \tag{3.37}$$

Since the dynamics (3.2) and (3.3), cost function (3.8) and the saddle-point strategies (3.30), (3.31), (3.36) and (3.37) are bounded and continuous in $\bar{x}_t$, it results that strategies (3.36) and (3.37) are approximate saddle-point strategies. The interested reader is referred to [16] for a detailed proof in the context of optimal control, which is similar to a great extent to the convergence proof of minimax control problem considered in this subsection, but note that the Riccati equations here are different and the relative errors defined in [16] will have intermittent natures. However, these differences do not add much complexity to the convergence proof because the Riccati equations (3.18) do not depend on the number of followers according to Assumption 3.3. Hence, the rate of convergence with respect to the number of followers is $1/n$, similar to [16, Theorem 2]. Therefore, the following result is obtained.

Theorem 3.2. *Let Assumptions 3.1–3.3 hold. The strategy described by* (3.36) *and* (3.37) *is an approximate saddle-point strategy for Problem 3.2.*

3.4 Numerical Examples

In this section, two numerical examples are provided to illustrate the efficacy of the obtained results.

Example 3.1. Consider a multi-agent network consisting of one leader and 100 identical followers whose dynamics are described by equations (3.2) and (3.3), respectively, with the following numerical parameters:

$$A_t^0 = 0.85, \quad B_t^0 = 0.15, \quad A_t = 0.85, \quad B_t = 0.85, \tag{3.38}$$

$$S_t^0 = 0.03, \quad S_t = 0.1, \quad E_t = 0.01, \tag{3.39}$$

$$w_t^i \sim \mathcal{N}(0, 0.3) \quad \forall i \in \mathbb{N}_n, \quad T = 20. \tag{3.40}$$

Let the initial state of the leader be $x_1^0 = 30$ and the initial states of the followers be chosen randomly (with uniform distribution) in the interval $[0, 20]$. The followers are exposed to an external disturbance given by:

$$d_t^i = 0.6sin(t), \quad t \in \mathbb{N}_{20}, i \in \mathbb{N}_{100}. \tag{3.41}$$

The objective of the leader and followers is to minimize the cost function (3.8) under the worst-case disturbance, where at any time $t \in \mathbb{N}_{20}$:

$$R_t = 70, \quad Q_t = 8, \quad F_t = 11, \quad P_t = 0.4, \tag{3.42}$$

$$R_t^0 = 50, \quad Q_t^0 = 0.5, \quad H_t = 0.1. \tag{3.43}$$

Sample trajectories of the leader and followers are depicted in Figure 3.1. It is shown that as the attenuation parameter γ increases, the fluctuations of the mean-field decrease which means better disturbance rejection.

Example 3.2. Consider 100 followers with identical dynamics that wish to track a

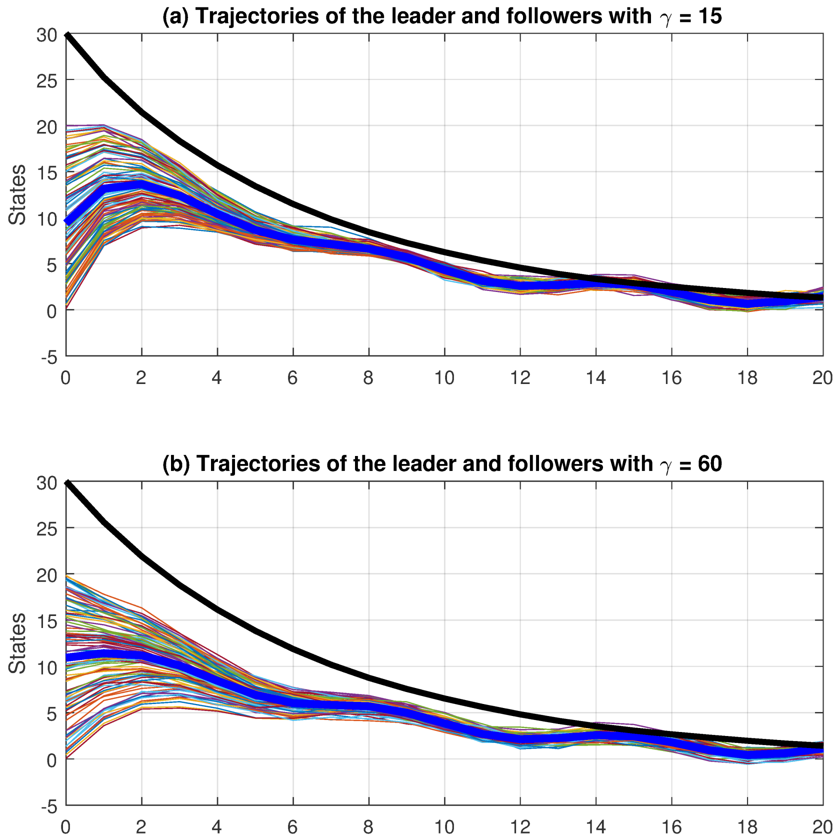


Figure 3.1: Sample trajectories of the leader and followers in Example 3.1, where thin colored curves are the states of the followers, thick blue curve is the mean-field, and thick black curve is the state of the leader.

reference signal, which may be viewed as a virtual leader with constant state $x_t^0 = 10, \forall t \in \mathbb{N}_{30}$. The initial states of the followers are chosen randomly in the interval $[0, 8]$ with a uniform distribution. The dynamics of the leader and followers are expressed by the following parameters:

$$A_t^0 = 1, \quad B_t^0 = 0, \quad A_t = 1, \quad B_t = 1, \tag{3.44}$$

$$S_t^0 = 0, \quad S_t = 0.04, \quad E_t = 0.001, \quad T = 30, \tag{3.45}$$

$$w_t^i \sim \mathcal{N}(0, 0.3), \quad \forall i \in \mathbb{N}_{100}. \tag{3.46}$$

The followers are exposed to local external disturbances given below:

$$d_t^i = 0.4sin(t), \quad i \in \mathbb{N}_{100}, \quad t \in \mathbb{N}_{30}. \tag{3.47}$$

The weight matrices in the cost function are given by:

$$R_t = 0.11, \quad Q_t = 0.01, \quad F_t = 0.07, \quad P_t = 0.001, \tag{3.48}$$

$$R_t^0 = 10^{-4}, \quad Q_t^0 = 10^{-4}, \quad H_t = 1. \tag{3.49}$$

In Figure 3.2, three sample trajectories of the states of followers are displayed for three different values of the attenuation parameter. It is observed that as the attenuation parameter increases, the disturbance is rejected more strongly at the cost of prolonging the consensus process.

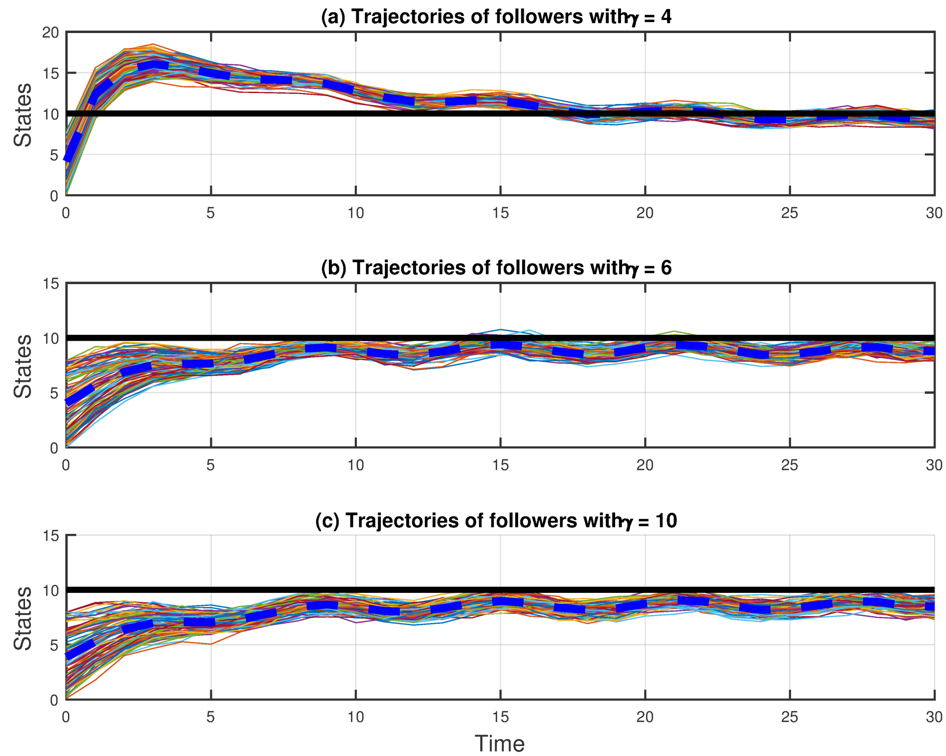


Figure 3.2: The trajectories of followers given different values of the attenuation parameter in Example 3.2

3.5 Conclusions

In this chapter, a robust control strategy was proposed for a class of leader-follower networks by solving a minimax control problem. Three decentralized information structures were studied. For mean-field sharing structure, it was shown that a unique saddle-point strategy exists under some mild assumptions. For partial mean-field sharing and intermittent mean-field sharing, the proposed strategy was shown to converge to the saddle-point strategy as the number of followers tends to infinity. The main feature of the obtained results is the fact that the solutions are identified by two Riccati equations that are not in the form of forward-backward equations, and their dimensions do not increase with the number of followers. In addition, it was numerically verified that the disturbance rejection property of the solution outweighs the consensus-reaching behavior when the attenuation parameter is large.

Chapter 4

Conclusions and Future Directions

In this thesis, the control of a leader-follower multi-agent network under non-classical information structure in the presence of noise is studied, and high-performance decentralized strategies are introduced to reach an agreement between the leader and followers. The proposed strategies are, in fact, promising alternative solutions compared to the existing techniques in the literature. In particular, using the law of large numbers in a network of many identical agents, some of the existing problems concerning communication constraints and computational complexity in this type of networks are effectively addressed.

In Chapter 2, it is desired that the leader and followers, which are assumed to be homogeneous, reach a common value while minimizing a given performance index. The dynamics of agents is coupled and dependent on the mean-field, and so is the cost function. In non-classical information structure, the agents have no access to the mean-field. Thus, using mean-field approximation, near-optimal strategies are obtained for the leader and every follower. It is shown that these solutions approach the optimal ones when there is a large number of followers. The salient feathers of the proposed strategies can be summarized as follows:

- A linear solution can be obtained under some realistic assumptions;
- not only is the solution suitable for a network with a large number of agents (as it is scalable), but also it outperforms existing control strategies for multi-agent systems, and
- convergence of the agents to consensus can be adjusted arbitrarily by proper choice of parameters.

In Chapter 3, the network described in Chapter 2 with non-classical information structure is assumed to be subject to external disturbances. The external disturbances are considered as adversarial players, penalized in the cost function by negative quadratic terms. It is shown that under some conditions, the problem has a unique saddle-point solution, in which, the value of the cost function in the worst-case scenario (in terms of disturbances) is minimized. When the mean-field is not available, a point in the neighborhood of the above solution is obtained, which is shown to be sufficiently close to the above solution when the number of followers is large enough. The main advantages of the results obtained in this chapter are given below:

- In contrast to the mean-field game solutions, the solution proposed here can be obtained by solving two decoupled equations for each follower, and as a result, it is scalable;
- the resultant strategies are linear functions of the agents' states and mean-field approximation;
- the trade off between the consensus-reaching and disturbance rejection performances is discussed, and its dependency on the attenuation parameter in the cost function is highlighted.

4.1 Future Research Directions

Some suggestions for future research in this area are outlined below:

- Extending the results to the infinite-horizon case, where some additional conditions such as stabilizability and detectability are expected to be imposed;
- exploring more efficient topologies for information exchange between the agents for the case where there are a large number of followers;
- considering a multi-agent system with more than one leader;
- investigating the problem for a network of heterogeneous followers, and

- considering the case where the communication links are subject to change.

Bibliography

[1] R. M. Murray, "Recent research in cooperative control of multi-vehicle systems," *Journal of Dynamic Systems, Measurement, and Control*, vol. 129, no. 5, pp. 571-583, 2007.

[2] P. Ogren, E. Fiorelli and N. E. Leonard, "Cooperative control of mobile sensor networks: adaptive gradient climbing in a distributed environment," *IEEE Transactions on Automatic Control*, vol. 49, no. 8, pp. 1292-1302, 2004.

[3] R. Couillet, S. M. Perlaza, H. Tembine and M. Debbah, "Electrical vehicles in the smart grid: a mean-field game analysis," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 6, pp. 1086-1096, 2012.

[4] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520-1533, 2004.

[5] L. Moreau, "Stability of multi-agent systems with time-dependent communication links," *IEEE Transactions on Automatic Control*, vol. 50, no. 2, pp. 169–182, 2005.

[6] W. Ren and R. W. Beard, "Consensus seeking in multi-agent systems under dynamically changing interaction topologies," *IEEE Transactions on Automatic Control*, vol. 50, no. 5, pp. 655-661, 2005.

[7] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal on Control*, vol. 6, no. 1, pp. 131-147, 1968.

[8] K. H. Movric and F. L. Lewis, "Cooperative optimal control for multi-agent systems on directed graph topologies," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 769-774, 2014.

[9] Y. Cao and W. Ren, "Optimal linear-consensus algorithms: an LQR perspective," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 40, no. 3, pp. 819-830, 2010.

[10] 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.

[11] M. Huang, P. E. Caines and R. P. Malhamé, "Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized ε-Nash equilibria," *IEEE Transactions on Automatic Control*, vol. 52, no. 9, pp. 1560–1571, 2007.

[12] J. Arabneydi, "New concepts in team theory: mean-field teams and reinforcement learning," *PhD thesis*, McGill University, 2016.

[13] J. Arabneydi and A. G. Aghdam, "A certainty equivalence result in team-optimal control of mean-field coupled Markov chains," in Proceedings of *the 56th IEEE Conference on Decision and Control*, Melbourne, Australia, pp. 3125-3130, December 2017.

[14] J. Arabneydi and A. G. Aghdam, "Optimal dynamic pricing for binary demands in smart grids: a fair and privacy-preserving strategy," in Proceedings of *the 2018 American Control Conference*, Milwaukee, USA, pp. 5368-5373, June 2018.

[15] J. Arabneydi, M. M. Baharloo and A. G. Aghdam, "Optimal distributed control for leader-follower networks: a scalable design," in Proceedings of *the 31st IEEE Canadian Conference on Electrical & Computer Engineering*, Quebec City, Canada, pp. 1-4, May 2018.

[16] M. Baharloo, J. Arabneydi and A. G. Aghdam, "Near-optimal control strategy in leader-follower networks: a case study for linear quadratic mean-field teams," in

Proceedings of *the 57th IEEE Conference on Decision and Control*, Miami Beach, USA, pp 3288–3293, December 2018.

[17] J. Arabneydi and A. G. Aghdam, "A mean-field team approach to minimize the spread of infection in a network," in Proceedings of *the 2019 American Control Conference*, Philadelphia, USA, July 2019, to appear.

[18] M. M. Baharloo, J. Arabneydi and A. G. Aghdam, "Minimax control in linear quadratic mean-field teams for a leader-follower network," submitted to *IEEE Control Systems Letters*.

[19] J. Arabneydi and A. Mahajan, "Linear quadratic mean-field teams: optimal and approximately optimal decentralized solutions," available at *https://arxiv.org/abs/1609.00056*, 2016.

[20] J. A. Fax and R. M. Murray, "Information flow and cooperative control of vehicle formations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1465-1476, 2004.

[21] P. Rawat, K. D. Singh, H. Chaouchi and J. M. Bonnin, "Wireless sensor networks: a survey on recent developments and potential synergies," *The Journal of Supercomputing*, vol. 68, no. 1, pp. 1-48, 2014.

[22] L. Gupta, R. Jain and G. Vaszkun, "Survey of important issues in UAV communication networks," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1123-1152, 2016.

[23] C. W. Reynolds, "Flocks, herds and schools: a distributed behavioral model," *ACM SIGGRAPH Computer Graphics*, vol. 21, no. 4, pp. 25-34, 1987.

[24] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: algorithms and theory," *IEEE Transactions on Automatic Control*, vol. 51, no. 3, pp. 401–420, 2006.

[25] A. Jadbabaie, J. Lin and A. 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.

[26] W. Ren, R. W. Beard and E. M. Atkins, "Information consensus in multi-vehicle cooperative control," *IEEE Control Systems Magazine*, vol. 27, no. 2, pp. 71-82, 2007.

[27] R. Olfati-Saber, J. Fax and R. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215-233, 2007.

[28] P. E. Caines, "Linear Stochastic Systems," *John Wiley & Sons Inc.*, 1987.

[29] F. Borrelli and T. Keviczky, "Distributed LQR design for identical dynamically decoupled systems," *IEEE Transactions on Automatic Control*, vol. 53, no. 8, pp. 1901-1912, 2008.

[30] J. Arabneydi and A. Mahajan, "Team-optimal solution of finite number of mean-field coupled LQG subsystems," in Proceedings of *the 54th IEEE Conference on Decision and Control*, Osaka, Japan, pp. 5308-5313, December 2015.

[31] H. Mahboubi, K. Moezzi, A. G. Aghdam, K. Sayrafian-Pour and V. Marbukh, "Distributed deployment algorithms for improved coverage in a network of wireless mobile sensors," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 1, pp. 163-174, 2014.

[32] Z. Shi, J. Tu, Q. Zhang, L. Liu and J. Wei, "A survey of swarm robotics system," in Proceedings of the *International Conference in Swarm Intelligence*, Berlin, Heidelberg, pp. 564–572, June 2012.

[33] T. Başar and P. Bernhard, "H-Infinity Optimal Control and Related Minimax Design Problems: a Dynamic Game Approach," *Springer Science and Business Media*, 2008.

[34] D. Jacobson, "Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games," *IEEE Transactions on Automatic Control*, vol. 18, no. 2, pp. 124-131, 1973.

[35] J. V. Neumann, “Zur theorie der gesellschaftsspiele,” *Mathematische Annalen*, vol. 100, no. 1, pp. 295–320, 1928.

[36] W. A. Van Den Broek, J. C. Engwerda and J. M. Schumacher, ”Robust equilibria in indefinite linear-quadratic differential games,” *Journal of Optimization Theory and Applications*, vol. 119, no. 3, pp. 565–595, 2003.

[37] D. Bauso, H. Tembine and T. Başar, ”Opinion dynamics in social networks through mean-field games,” *SIAM Journal on Control and Optimization*, vol. 54, no. 6, pp. 3225–3257, 2016.

[38] J. Huang and M. Huang, ”Robust mean-field linear-quadratic-Gaussian games with unknown L^2-disturbance,” *SIAM Journal on Control and Optimization*, vol. 55, no. 5, pp. 2811–2840, 2017.

[39] B. C. Wang and J. Huang, ”Social optima in robust mean-field LQG control,” in Proceedings of *the 11th IEEE Asian Control Conference*, Gold Coast, Australia, pp. 2089–2094, December 2017.

[40] M. A. Tajeddini, H. Kebriaei and L. Glielmo, ”Robust decentralised mean-field control in leader following multi-agent systems,” *IET Control Theory and Applications*, vol. 11, no. 16, pp. 2707–2715, 2017.

[41] T. Başar and G. Olsder, ”Dynamic Noncooperative Game Theory,” *Society for Industrial and Applied Mathematics*, 1998.

[42] R. Elliott, X. Li and Y. H. Ni, "Discrete time mean-field stochastic linear-quadratic optimal control problems," *Automatica*, vol. 49, no. 11, pp. 3222-3233, 2013.

[43] J. M. Lasry and P. L. Lions, "Mean-field games," *Springer Japanese Journal of Mathematics*, vol. 2, no. 1, pp. 229-260, 2007.

[44] 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*, Maui, USA, pp. 98-103, December 2003.

[45] L. P. Kadanoff, "More is the same; phase transitions and mean-field theories," *Journal of Statistical Physics*, vol. 137, no. 5, pp. 777-790, 2009.

[46] P. Weiss, "Molecular field hypothesis and ferromagnetic property," *Journal of Physics Theory and Application*, vol. 1, no. 5, pp. 77-89, 1907.

[47] J. Yong, "Linear-quadratic optimal control problems for mean-field stochastic differential equations," *SIAM on Control and Optimization*, vol. 51, no. 4, pp. 2809-2838, 2013.

[48] D. Morale, V. Capasso and K. Oelschläger, "An interacting particle system modelling aggregation behavior: from individual to populations," *Journal of Mathematical Biology*, vol. 50, no.1, pp. 49-66, 2005.

[49] H. S. Witsenhausen, "Separation of estimation and control for discrete time systems," *Proceedings of the IEEE*, vol. 59, no. 11, pp. 1557-1566, 1971.

[50] M. M. Baharloo, J. Arabneydi and A. G. Aghdam, "Minimax control in linear quadratic mean-field teams for a leader-follower network," submitted to *the 58th IEEE Conference on Decision and Control*.