

Distributed Game Strategy Design with Application to Multi-Agent Formation Control

Wei Lin, *Member, IEEE*, Zhihua Qu *Fellow, IEEE*, and Marwan A. Simaan, *Life Fellow, IEEE*

Abstract

In this paper, we consider a multi-agent formation control problem from a game theory point of view. It is well known that a major difficulty in a communication network based formation control problem is that each agent is only able to exchange information with other agents according to the communication topology. This information constraint prevents many game strategy design approaches that require individual agents to have global information from being implemented in many cases. We formulate the formation control problem in such a way that individual agents try to minimize their locally measured formation errors and to solve it as a differential game problem. We consider two cases of non-cooperative and cooperative games and propose a novel distributed design approach that utilizes the relationship between the initial and terminal state variables. This approach is applied to an illustrative formation control example among three agents and the formation errors under various scenarios are compared and analyzed.

I. INTRODUCTION

An important application of cooperative control theory [?], [?] is the multi-agent formation control problems [?], [?], [?]. This problem revolves around designing local information constrained control inputs for the agents (mobile robots, unmanned vehicles, etc) and making their positions follow a prescribed formation. There is considerable work done in this field and the

This work is supported in part by US National Science Foundation under grants ECCS-1308928 and CCF-0956501, by US Department of Energys award DE-EE0006340, by US Department of Transportations award DTRT13-G-UTC51, by L-3 Communications contract 11013I2034, and by Leidos contract P010161530.

Wei Lin is with Western Digital Corporation, Irvine, CA, 92612, USA. E-mail: weilin0929@gmail.com.

Zhihua Qu and Marwan A. Simaan are with the Department of EECS, University of Central Florida, Orlando, Florida, 32816, USA, e-mails: qu@eeecs.ucf.edu, simaan@eeecs.ucf.edu.

comprehensive review paper [?] on multi-agent distributed control systems is a good reference on recent progress in cooperative control theory including the formation control problem. Recently, multiple attempts [?], [?], [?] have been made to incorporate differential game theory [?], [?] into the formation control design problem. A fundamental connection between these two problems can be explained as follows: in a multi-agent formation control problem, all the agents are basically trying to minimize the formation errors between each other. If every agent is able to observe all the other agents, then the problem can be essentially viewed as an optimal control problem where a common performance index of global formation errors can be assigned for all the agents. However, in most of realistic applications, since the agents are usually connected through a communication network with a certain topology, they are only able to exchange information with each other in a distributed manner (i.e., to share information locally). In this case, instead of minimizing a globally defined performance index, it is more realistic for each agent to only minimize a performance index of its locally measured formation errors. This scenario is exactly equivalent to that of a differential game where each player is trying to minimize its own performance index. Therefore, most multi-agent formation control design problems can be viewed and solved as differential game problems. The game solutions including Nash equilibria for non-cooperative agents and Pareto optimality for cooperative agents can be well established. However, solving for the Nash and Pareto solutions under distributed information is not a trivial task in the multi-agent environment because of the following two issues : (1) the strategy for each agent has to conform to its local information topology constraint and (2) local strategy computing within each agent is preferred rather than centralized strategy computing where strategies are designed by a centralized designer and assigned to individual agents. In [?], [?], the first issue has been addressed and the corresponding Nash and Pareto strategies are considered for the formation control problem as a differential game. In this present paper, we will focus more on the second aspect. By investigating the structure of the game solution in more details, a strategy design approach as well as an implementation algorithm are proposed for each agent to carry out its optimal control strategy design in a fully distributed manner. The remainder of this paper is organized as follows. A brief setup of the game problem is presented in Section II. The main results of the distributed Nash and Pareto strategy designs are presented in Section ?? and ?? respectively. An illustrative example with three agents under noncooperative and cooperative cases is presented and analyzed in Section IV.

II. PROBLEM FORMULATION

We consider the finite-time formation control problem between N agents in the n -dimensional Euclidean space where the motion dynamics of each agent is described as the following double integrator model:

$$\begin{bmatrix} \dot{p}_i \\ \dot{v}_i \end{bmatrix} = \begin{bmatrix} 0 & I_n \\ 0 & 0 \end{bmatrix} \begin{bmatrix} p_i \\ v_i \end{bmatrix} + \begin{bmatrix} 0 \\ I_n \end{bmatrix} u_i \quad (1)$$

for $i = 1, \dots, N$. Vector $p_i \in \mathbb{R}^n$ and $v_i \in \mathbb{R}^n$ are the position and velocity of agent i respectively. Vector $u_i \in \mathbb{R}^n$ is the acceleration control of agent i . We assume that the agents are able to exchange information with each other through a communication network whose topology can be described by a set \mathcal{E} in graph theory. An edge $e_{ij} \in \mathcal{E}$ represents information being transmitted from agent j to agent i . The objective of the agents is to form a prescribed formation over a finite time interval, which is equivalent to individual agents minimizing the following performance indices:

Since every agent is only aware of its local communication pattern and its own performance index, this formation problem can be regarded as a differential game problem where each player has its own objectives to pursue over a time interval. In this paper, we consider two types of open-loop (where the control inputs are functions of initial states and time only) formation control approaches through differential game strategy design. Specifically, if the control inputs u_1, \dots, u_N can be looked as strategies, then for the formulated problem, strategies u_1^N, \dots, u_N^N are the Nash strategies and form a Nash equilibrium if the inequality

$$J_i(u_1^N, \dots, u_i^N, \dots, u_N^N) \leq J_i(u_1^N, \dots, u_i, \dots, u_N^N) \quad (2)$$

holds for all $u_i \in U_i$ and for all $i = 1, \dots, N$, where U_i is the open-loop control set for agent i , i.e., $U_i = \{u_i(t, x(0)) \in \mathbf{R}^r \mid t \in [0, t_f]\}$. The Nash equilibrium is an equilibrium where it is not possible for any agent to lower its performance index value by unilaterally deviating from its Nash strategy. The Nash equilibrium is a widely used solution in the situation where the players do not cooperate with each other. On the other hand, The strategies $u_0^P, u_1^P, \dots, u_N^P$ are the Pareto strategies and form a Pareto optimality if the inequality

$$J_i(u_1^P, \dots, u_N^P) < J_i(u_1, \dots, u_N) \quad (3)$$

holds for at least one $i \in \{1, \dots, N\}$. The Pareto optimality can be interpreted as a solution in which any changes made do not help lower every agent's performance index value. The Pareto optimality is a widely used solution in the situation where the players can cooperate with each other. In the following sections, two formation control approaches will be proposed based on the open-loop Nash strategy (assuming that the agents do not cooperate) and the open-loop Pareto strategy (assuming that the agents cooperate).

III. OPEN-LOOP NASH ANALYSIS FOR HOMOGENEOUS AGENTS

We consider homogeneous time varying discrete time agents' dynamics:

$$\dot{x}_i(t) = A(t)x_i(t) + B(t)u_i(t)$$

$$y_i(t) = C(t)x_i(t)$$

where $y_i \in \mathbb{R}^m$ is the output state of agent i that needs to be coordinated with other agents, such as velocity, position, etc. The initial state $x_i(0)$ is usually known to agent i only. The objective of a coordination problem is generally equivalent to minimizing by the following performance index

$$J_i = \frac{1}{2} \sum_{e_{ij} \in \mathcal{E}} \|y_i(t_f) - y_j(t_f) - \mu_{ij}\|_{F_{ij}}^2 + \frac{1}{2} \int_0^{t_f} \|u_i(t)\|_{R_i(t)}^2 dt \quad (4)$$

for $i = 1, \dots, N$, where μ_{ij} is a prescribed vector by the coordination problem, matrix F_{ij} is positive-semi definite, and matrix $R_i(t)$ is positive definite for all $t \in [0, t_f]$. Minimizing performance index J_i means that every agent will try to minimize its total coordination error measured locally according to the graph while at the same time minimizing its total control energy in the game. To facilitate Nash strategy analysis for the formulated multi-agent coordination problem, we define a matrix weighted Laplacian based on network topology and agents' performance indices.

Definition 1: For the multi-agent coordination problem on graph \mathcal{G} with performance indices in (4), a **matrix weighted Laplacian matrix** is defined as $L = [L_{ij}] \in \mathbb{R}^{Nn \times Nn}$ where

$$L_{ij} = \begin{cases} -C^T F_{ij} C & \text{if } e_{ij} \in \mathcal{E} \text{ and } j \neq i \\ 0 & \text{if } e_{ij} \notin \mathcal{E} \text{ and } j \neq i \\ \sum_{e_{ij} \in \mathcal{E}} C^T F_{ij} C & \text{if } j = i \end{cases} \quad (5)$$

The open-loop Nash equilibrium for the formulated problem is presented as follows.

Theorem 1: For the N -agent linear quadratic differential game of fixed duration $[0, t_f]$ defined by agents' dynamics (1) and performance indices (4), if the information structure of every agent is open-loop, i.e., $U_i = \{u_i(t, x(0)) \in \mathbf{R}^r \mid t \in [0, t_f]\}$, there exists a unique Nash equilibrium where corresponding open-loop control strategy vector u_i^* is given by

$$u_i^*(t, x(0)) = -R_i^{-1}B^T[e_i \otimes \Phi^T(t_f, t)]\{L(I + DL)^{-1}[\Phi(t_f, 0)x(0) + D\mu] - \mu\} \quad (6)$$

if matrix $(I + DL)$ is invertible, where

$$R = \text{blkdiag}\{R_1, \dots, R_N\}, \quad (7)$$

$$D = \text{blkdiag}\{D_1, \dots, D_N\}, \quad (8)$$

$$D_i = \int_0^{t_f} \Phi(t_f, \tau) B R_i^{-1} B^T \Phi^T(t_f, \tau) d\tau \quad (9)$$

$$x = [x_1^T, \dots, x_N^T]^T, \quad (10)$$

$$\mu = \left[\sum_{e_{1j} \in \mathcal{E}} \mu_{1j}^T, \dots, \sum_{e_{Nj} \in \mathcal{E}} \mu_{Nj}^T \right]^T, \quad (11)$$

e_i is a unit vector with the i th entry being 1 and elsewhere 0, and $\Phi(t_f, t)$ is the state-transition matrix of A .

Proof: We define the Hamiltonian for agent i as

$$H_i = \frac{1}{2} \|u_i\|_{R_i}^2 + \lambda_i^T (Ax_i + Bu_i) \quad (12)$$

where vector λ_i is the Lagrangian multiplier. Necessary conditions for the existence of open-loop Nash equilibria are

$$\dot{x}_i = \frac{\partial H_i}{\partial \lambda_i} = Ax_i + Bu_i, \quad (13a)$$

$$\dot{\lambda}_i = -\frac{\partial H_i}{\partial x_i} = -A_i^T \lambda_i, \quad (13b)$$

$$\begin{aligned} \lambda_i(t_f) &= \sum_{e_{ij} \in \mathcal{E}} C^T F_{ij} [y_i(t_f) - y_j(t_f) - \mu_{ij}] \\ &= \sum_{e_{ij} \in \mathcal{E}} L_{ij} [x_i(t_f) - x_j(t_f)] - \mu_i, \end{aligned} \quad (13c)$$

$$\frac{\partial H_i}{\partial u_i} = R_i u_i + B_i^T \lambda_i = 0, \quad (13d)$$

for all $i = 1, \dots, N$, where $\mu_i = \sum_{e_{ij} \in \mathcal{E}} C^T F_{ij} \mu_{ij}$. Therefore, the existence of the Nash equilibria is equivalent to the solvability of above equations given initial states $x_i(0)$ for all $i = 1, \dots, N$. The uniqueness of Nash equilibrium is equivalent to the uniqueness of the solution to conditions (13a) to (13d). From (13b), the trajectory of λ_i can be obtained as

$$\lambda_i(t) = \Phi^T(t_f, t) \lambda_i(t_f)$$

where . From (13d), the control vector can be obtained as

$$u_i(t) = -R_i^{-1} B^T \lambda_i(t), \quad (14)$$

Substituting the above expression into (13a) yields

$$\dot{x}_i = Ax_i - BR_i^{-1} B^T \lambda_i.$$

Based on the above equation, the state at t_f can be expressed as

$$x_i(t_f) = \Phi(t_f, 0) x_i(0) - D_i \lambda_i(t_f), \quad (15)$$

where D_i is defined in (9). Based on definition of L in (5), condition (13c) can be expressed more compactly as

$$\lambda(t_f) = Lx(t_f) - \mu,$$

where $\lambda_v = [\lambda_{v1}, \dots, \lambda_{vN}]^T$. Stacking equation (15) from $i = 1$ to $i = N$ yields

$$\begin{aligned} x(t_f) &= [I_N \otimes \Phi(t_f, 0)] x(0) - D \lambda(t_f) \\ &= [I_N \otimes \Phi(t_f, 0)] x(0) - D(Lx(t_f) - \mu) \end{aligned}$$

where $x = [x_1^T, \dots, x_N^T]^T$, and L is defined in (5). The relationship between $x(t_f)$ and $x(0)$ can be expressed as

$$(I + DL)x(t_f) = [I_N \otimes \Phi(t_f, 0)] x(0) + D\mu \quad (16)$$

If matrix $(I + DL)$ is invertible, control vector (14) can be now expressed as

$$\begin{aligned} \lambda_i(t) &= \Phi^T(t_f, t) \lambda_i(t_f) \\ &= [e_i \otimes \Phi^T(t_f, t)] \lambda(t_f) \\ &= [e_i \otimes \Phi^T(t_f, t)] [Lx(t_f) - \mu] \\ &= [e_i \otimes \Phi^T(t_f, t)] \{L(I + DL)^{-1} [\Phi(t_f, 0) x(0) + D\mu] - \mu\} \end{aligned} \quad (17)$$

Substituting the above expression into equation (14) yields (6). Therefore, the open-loop Nash equilibrium exists and is unique. ■

It is obvious that computing and implementing open-loop Nash strategy (6) requires agent i to have knowledge of full initial state vector $x(0)$ which conflicts with communication constraint of multi-agent system that each agent is only aware of information from local agents. In fact, we are able to overcome this difficulty by carrying out distributed estimation under certain spectral condition of matrix $(I + DL)$. The basic idea is as follows: with every agent having an estimate of its actual open-loop Nash strategy, we let all agents exchange their estimates over time in a distributed manner based on given network topology to achieve asymptotic convergence of the estimates to the corresponding actual open-loop Nash strategies. Toward that, we have the following result.

Theorem 2: If all eigenvalues of matrix $(I + DL)$ have positive real parts, then \hat{u}_i converges to open-loop Nash strategy u_i^* in (6), that is,

$$\lim_{\tau \rightarrow \infty} \hat{u}_i(t, x_i(0), \tau) = u_i^*(t, x(0)) \quad (18)$$

for all $i = 1, \dots, N$ given

$$\hat{u}_i(t, x_i(0), \tau) = -R_i^{-1}B^T \sum_{e_{ij} \in \mathcal{E}} L_{ij}[\hat{x}_{if}(\tau) - \hat{x}_{jf}(\tau)] - \mu_i, \quad (19)$$

where \hat{x}_{if} satisfies

$$\dot{\hat{x}}_{if}(\tau) = g_i \left\{ \Phi(t_f, 0)x_i(0) - x_i(t_f)(\tau) - D_i \sum_{e_{ij} \in \mathcal{E}} L_{ij}[x_i(t_f)(\tau) - x_{jf}(\tau)] + D_i \mu_i \right\}, \quad (20)$$

for all $i = 1, \dots, N$, where g_i is a positive scalar.

Proof: Stacking equations (20) for $i = 1, \dots, N$ yields

$$\dot{x}_f(\tau) = G \{ [I_N \otimes \Phi(t_f, 0)]x(0) - (I + DL)x_f(\tau) + D\mu \}, \quad (21)$$

where $x_f = [x_{1f}^T, \dots, x_{Nf}^T]^T$, $G = \text{blkdiag}\{g_1 I, \dots, g_N I\}$. If all eigenvalues of matrix $(I + DL)$ has positive real parts, differential equation (21) is asymptotically stable and $x_f(\tau)$ converges to $x(t_f)$ in (16) as τ goes to infinity. Hence, \hat{u}_i in (19) also converges to open-loop Nash strategy u_i^* in (6) as τ goes to infinity. Therefore, equation (18) holds for all $i = 1, \dots, N$. ■

In the above theorem, variable $\hat{u}_i(\tau)$ can be in fact regarded as an estimate of u_i^* made by agent i . Through exchanging this estimate information among themselves according to equation

(20), all agents are able to collaboratively make their estimates asymptotically converge to their open-loop Nash strategies.

Implementing equation (20) is indeed distributed because it is based on each agent's own performance index and information received from local agents only. Specifically, at time τ , agent i only needs to send out its own estimate variable $x_i(t_f)(\tau)$ to agent j for all $e_{ij} \in \mathcal{E}$ and receive $x_{jf}(\tau)$ from agent j for all $e_{ji} \in \mathcal{E}$. In equation (20), scalar coefficient g_i determines the convergence rate and a large value of g_i results in faster convergence.

As we can see, in order to make distributed open-loop Nash strategy estimation valid, all eigenvalues of matrix $(I + DL)$ need to have positive real parts. In general, this condition needs to be validated on a case by case basis based on agents' dynamics, performance indices, and graph connectivity. Nonetheless, there still exist several special cases where spectral property of $(I + DL)$ can be simply analyzed based on the structure of D and L . These scenarios are summarized as follows.

Proposition 1: All eigenvalues of $(I + DL)$ have positive real parts if either of the following holds

- (1) Graph \mathcal{G} is undirected and $F_{ij} = F_{ji}$ for all $e_{ij} \in \mathcal{E}$.
- (2) Graph \mathcal{G} is directed and weighting matrices for terminal state and control energy can be expressed as $F_{ij} = f_{ij}\bar{F}$ for all $e_{ij} \in \mathcal{E}$ and $R_i = r_i\bar{R}$ for all $i = 1, \dots, N$, respectively, where f_{ij} and d_i are positive scalars and matrix \bar{F} and \bar{R} are positive semi-definite.

Proof: For scenario (1), if graph \mathcal{G} is undirected, matrix L defined in (5) is positive semi-definite because

$$x^T L x = \frac{1}{2} \sum_{e_{ij} \in \mathcal{E}} \|x_i - x_j\|_{C^T F_{ij} C}^2 \geq 0.$$

Since matrix D is positive semi-definite, it can be written as $D = D_s^T D_s$. Moreover, eigenvalues of $DL = D_s^T D_s L$ are equivalent to those of $D_s L D_s^T$ using the fact that $E_1 E_2$ and $E_2 E_1$ have the same eigenvalues for square matrix E_1 and E_2 . Since $D_s L D_s^T$ is positive semi-definite, product DL has non-negative eigenvalues and hence eigenvalues of $(I + DL)$ are positive.

For scenario (2), if $F_{ij} = f_{ij}\bar{F}$, matrix weighted Laplacian matrix L can be expressed as $L = \bar{L} \otimes \bar{F}$ where \bar{L} is the conventional scalar weighted Laplacian matrix for graph \mathcal{G} with weights f_{ij} on edge $e_{ij} \in \mathcal{E}$. If $R_i = r_i\bar{R}$, block matrix D can be written as $D = \text{diag}\{r_1, \dots, r_N\} \otimes \bar{D}$ where \bar{D} denotes the result of (9) when $R_i = r_i\bar{R}$. Therefore, product $(I + DL)$ can be expressed

as

$$\begin{aligned} I + DL &= I + (\text{diag}\{r_1, \dots, r_N\} \otimes \bar{D})(\bar{L} \otimes \bar{F}) \\ &= I + \bar{L}_r \otimes (\bar{D}\bar{F}), \end{aligned}$$

where $\bar{L}_r = \text{diag}\{r_1, \dots, r_N\} \bar{L}$. As shown previously, product $\bar{D}\bar{F}$ has non-negative eigenvalues because \bar{D} and \bar{F} are both positive semi-definite. Moreover, it is well known that all eigenvalues of scalar weighted Laplacian matrix \bar{L} or \bar{L}_d have non-negative real parts. Therefore, according to its spectrum property, Kronecker product $\bar{L}_r \otimes (\bar{D}\bar{F})$ has non-negative real parts and hence eigenvalues of $(I + DL)$ have positive real parts. ■

Physical interpretations of the above two scenarios are given as follows: In scenario (1), communication between two connected agents on the graph is bidirectional and these pair of agents have the common weighting matrix ($F_{ij} = F_{ji}$) associated with their terminal output variable difference. In scenario (2), communication between two connected agents on the graph is directional and all agents in fact choose weighting scalars f_{ij} and r_i (instead of matrices) associated with their terminal output variable difference and control energy.

If all eigenvalues of $(I + DL)$ are verified to have positive real parts, next problem to consider is how to implement or synthesize open-loop Nash strategy in real applications. There are basically two types of implementations: One is offline and the other is online. For an offline implementation, we let all agents collaboratively communicate with each other and carry out differential equation (20) over an adequate amount of time, say τ_f , such that estimate $\hat{u}_i(\tau_f)$ is close enough to open-loop Nash strategy u_i^* , or equivalently, variable change $\|\dot{x}_{if}(\tau_f)\|$ in (20) is below a prescribed threshold. Each agent will then implement $\hat{u}_i(\tau_f)$ as its control strategy in the game. The advantage of an offline implementation is that it provides arbitrary accuracy of estimate \hat{u}_i to the actual open-loop Nash strategy. The disadvantages of such an implementation is that it requires a “warm-up” time τ_f and all agents have to know and agree on the value of τ_f . For an online implementation, all agents carry out equation (19) and (20) simultaneously once game starts. The resulting control strategies can be expressed as follows:.

$$\dot{x}_{if}(t) = g_i \{ \Phi(t_f, 0)x_i(0) - x_i(t_f)(t) - D_i \sum_{e_{ij} \in \mathcal{E}} L_{ij}[x_i(t_f)(t) - x_{jf}(t)] + D_i \mu_i \} \quad (22a)$$

$$\hat{u}_i(t) = -R_i^{-1} B^T \sum_{e_{ij} \in \mathcal{E}} L_{ij}[\hat{x}_{if}(t) - \hat{x}_{jf}(t)] - \mu_i, \quad (22b)$$

for all $i = 1, \dots, N$. The advantages of an online implementation is that it does not require any “warm-up” time and is suitable for real-time applications. The disadvantage of such an implementation is that the resulting trajectory of \hat{u}_i in (22b) over entire game interval $[0, t_f]$ is generally less accurate to actual open-loop Nash strategy than the offline one. To quantify this inaccuracy, the concept of ϵ -Nash equilibrium is introduced as follows.

Definition 2: For a given scalar ϵ , strategies $u_1^\epsilon, \dots, u_N^\epsilon$ form an open-loop ϵ -Nash equilibrium if inequality

$$J_i(u_1^\epsilon, \dots, u_i^\epsilon, \dots, u_N^\epsilon) \leq \min_{u_i \in U_i} J_i(u_1^\epsilon, \dots, u_i, \dots, u_N^\epsilon) + \epsilon \quad (23)$$

holds for all $i = 1, \dots, N$, where U_i is the open-loop control set for agent i , i.e., $U_i = \{u_i(t, x(0)) \in \mathbf{R}^r \mid t \in [0, t_f]\}$.

If $\epsilon = 0$, the above definition becomes exactly the same as the one for open-loop Nash equilibrium and hence $u_i^\epsilon = u_i^*$. Intuitively, as ϵ decreases, ϵ -Nash equilibrium will be close to the original Nash equilibrium. Therefore, the value of ϵ can be used to measure the closeness of online implementation control strategy \hat{u}_i in (22b) to u_i^* in (6). Toward that end, we show that $\hat{u}_1, \dots, \hat{u}_N$ form an ϵ -Nash equilibrium with the value of ϵ explicitly derived.

First of all, we present the following lemmas:

Lemma 1: All the eigenvalues of matrix

$$S_{jj} = (d_j^T \otimes I_r) S_j D_j^T \tilde{F}_j (d_j \otimes I_r) \quad (24)$$

has nonnegative real parts for all $j = 1, \dots, N$ where matrix S_j is defined in (??).

Proof: Substituting (??) into (24) yields

$$S_{jj} = (d_j^T \otimes I_r) S_j (d_j \otimes I_r) = (d_j^T \otimes I_r) \int_0^{t_f} \tilde{B}_j(t) R_j^{-1}(t) \tilde{B}_j^T(t) dt D_j^T \tilde{F}_j (d_j \otimes I_r).$$

Due to the definition of D_j in (9) and $\tilde{B}_j = (d_j \otimes I_r) B_j'$, the above equation becomes

$$S_{jj} = \underbrace{\int_0^{t_f} B_j'(t) R_j^{-1}(t) (B_j')^T(t) dt}_a \underbrace{(d_j^T \otimes I_r) \tilde{F}_j (d_j \otimes I_r)}_b.$$

Since both term a and term b are positive semi-definite, all the eigenvalues of matrix S_{jj} have nonnegative real parts. ■

Lemma 2: If $(-M)$ is Hurwitz for matrix M defined in (49), supposing that matrix P is the unique positive definite solution to the Lyapunov equation

$$-M^T P - P M = -I, \quad (25)$$

then for linear system

$$\dot{\theta} = -gM\theta \quad (26)$$

with the initial state $\theta(0)$, the inequalities

$$\int_0^{t_f} \|\theta\|_2^2 dt \leq \frac{2\gamma_{\max} V(0)}{g\gamma_{\min}} (1 - e^{-g\gamma_{\min} t_f}), \quad (27a)$$

$$\int_0^{t_f} \|\theta\|_2 dt \leq \frac{2\sqrt{2\gamma_{\max} V(0)}}{g\gamma_{\min}} (1 - e^{-g\gamma_{\min} t_f/2}). \quad (27b)$$

hold, where $\|\cdot\|_2$ stand for Euclidean norm, $V(0) = 1/2\theta^T(0)P\theta(0)$, $\gamma_{\min} = 1/\lambda_{\max}(P)$, $\lambda_{\max}(P)$ is the largest eigenvalue of matrix P , $\gamma_{\max} = 1/\lambda_{\min}(P)$, and $\lambda_{\min}(P)$ is the smallest eigenvalue of matrix P .

Proof: If matrix $(-M)$ is Hurwitz, we consider the quadratic Lyapunov function $V = 1/2\theta^T P\theta$ for system (26), where matrix P is the solution to the Lyapunov equation (25). The derivative of the Lyapunov function along the trajectory of system (26) is

$$\dot{V} = \frac{1}{2}\theta^T(-gM^T P - gPM)\theta = -\frac{g}{2}\|\theta\|_2^2 \leq -\frac{g}{2\lambda_{\max}(P)}\theta^T P\theta = -g\gamma_{\min} V$$

where the last inequality is due to the property $\theta^T P\theta \leq \lambda_{\max}(P)\|\theta\|_2^2$. The above differential inequality yields

$$\frac{1}{2}\lambda_{\min}(P)\|\theta\|_2^2 \leq \frac{1}{2}\theta^T P\theta = V(t) \leq e^{-g\gamma_{\min} t} V(0) \implies \|\theta\|_2^2 \leq 2\gamma_{\max} e^{-g\gamma_{\min} t} V(0).$$

Therefore, integrating the above inequality from 0 to t_f yields (27a) and the square root of the above inequality from 0 to t_f yields (27b). ■

Lemma 3: Under strategy , Δx_i can be expressed as

Proof: Since $\Delta x_i(t_f) = x_i(t_f) - x_i^*(t_f)$, we need to find expressions for both $x_i(t_f)$ and $x_i^*(t_f)$. For $x_i(t_f)$, we solve differential equation (21) as

$$\begin{aligned} \dot{x}_f(\tau) &= G \{ [I_N \otimes \Phi(t_f, 0)]x(0) - (I + DL)x_f(\tau) + D\mu \}, \\ z^f(t) &= M^{-1}z(0) + \theta(t). \end{aligned} \quad (28)$$

where $\theta(t)$ is defined in (26) with the initial state $\theta(0) = z^f(0) - M^{-1}z(0)$. The system dynamics (??) when every agent chooses strategy (22a) becomes

$$\dot{z} = - \sum_{j=1}^N \tilde{B}_j R_j^{-1} \tilde{B}_j^T D_i^T \tilde{F}_j D_j z^f \triangleq -W z^f \quad (29)$$

where W is defined in (40f). Substituting (28) into (29) yields

$$\dot{z} = -WM^{-1}z(0) - W\theta. \quad (30)$$

Since

$$\int_0^{t_f} W dt = \sum_{j=1}^N S_j F_j$$

where matrices F_j and S_j are defined in (??) and (??), integrating equation (30) from 0 to t_f yields

$$z(t_f) = \left(I - \sum_{j=1}^N S_j F_j M^{-1} \right) z(0) - \int_0^{t_f} W \theta dt. \quad (31)$$

Recalling the definition of matrix M in (49), we have

$$\left(I + \sum_{j=1}^N S_j F_j \right) M^{-1} = I \implies M^{-1} = I - \sum_{j=1}^N S_j F_j M^{-1}.$$

Hence, equation (31) becomes

$$z(t_f) = M^{-1}z(0) - \int_0^{t_f} W \theta dt. \quad (32)$$

Second, substituting u_i^* in (??) into (??) yields

$$\dot{z}_i^* = -(d_i^T \otimes I_r) \tilde{B}_i R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i [D_i z(t_f) - (d_i \otimes I_r) \Delta z_i(t_f)]$$

Integrating the above equation from 0 to t_f yields

$$\begin{aligned} z_i^*(t_f) &= z_i(0) - (d_i^T \otimes I_r) S_i D_i^T \tilde{F}_i [D_i z_i(t_f) - (d_i \otimes I_r) \Delta z_i(t_f)] \\ &= z_i(0) + S_{ii} \Delta z_i(t_f) - (d_i^T \otimes I_r) S_i F_i z(t_f) \end{aligned}$$

where S_{ii} is as defined in (24). Therefore, substituting the above equation into the expression of $\Delta z_i(t_f)$ yields

$$\Delta z_i(t_f) = z_i(t_f) - z_i^*(t_f) = z_i(t_f) - z_i(0) - S_{ii} \Delta z_i(t_f) + (d_i^T \otimes I_r) S_i F_i z(t_f)$$

After some manipulations, we arrive at

$$(I + S_{ii}) \Delta z_i(t_f) = (d_i^T \otimes I_r) [M z(t_f) - z(0)]$$

As we showed in Lemma 1, all the eigenvalues of S_{ii} in the right hand side of the above equation have nonnegative real parts and hence matrix $(I + S_{ii})$ is invertible. Therefore,

$$\Delta z_i(t_f) = (I + S_{ii})^{-1} (d_i^T \otimes I_r) [M z(t_f) - z(0)] \quad (33)$$

Therefore, substituting (32) into (33) yields the value of $\Delta z_i(t_f)$ as follows:

$$\Delta z_i(t_f) = - (I + S_{ii})^{-1} (d_i^T \otimes I_r) M \int_0^{t_f} W \theta dt. \quad (34)$$

■

Lemma 4: Δu_i

Proof: The value of Δu_i obtained as follows:

$$\begin{aligned} \Delta u_i &= u_i - u_i^* = - R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i D_i z^f + R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i D_i z^*(t_f) \\ &= - R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i D_i z^f + R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i D_i z(t_f) - R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i (d_i \otimes I_r) \Delta z_i(t_f) \\ &= - R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i D_i [z^f - z(t_f)] - R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i (d_i \otimes I_r) \Delta z_i(t_f), \end{aligned} \quad (35)$$

Recalling (??) and (32), the value of $[z^f - z(t_f)]$ in the above equation can be obtained as

$$z^f - z(t_f) = \theta + \int_0^{t_f} W \theta dt \quad (36)$$

Substituting (34) and (36) into (35) yields

$$\begin{aligned} \Delta u_i &= - R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i D_i \left(\theta + M_i \int_0^{t_f} W \theta dt \right) \\ &\triangleq - R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i D_i v_i, \end{aligned} \quad (37)$$

where $v_i = \theta + M_i \int_0^{t_f} W \theta dt$ and matrix M_i is defined in (40e). ■

We now present the ε -Nash equilibrium for the proposed strategy (22) as the following theorem:

Theorem 3: For the N -agent linear quadratic differential game of fixed duration $[0, t_f]$ defined by agents' dynamics (1) and performance indices (4), if all eigenvalues of $(I + DL)$ has positive real parts, then the control strategies $\hat{u}_1, \dots, \hat{u}_N$ in (22) form an ϵ -Nash equilibrium given

$$\epsilon = \max\{\epsilon_1, \dots, \epsilon_N\} \quad (38)$$

where

$$\epsilon_i = \frac{2 \|\tilde{F}_{Ri}\|_2 \gamma_{\max} V(0)}{g \gamma_{\min}} (1 - e^{-g \gamma_{\min} t_f}) + \frac{8 \|\tilde{M}_i\|_2 W_{\max}^2 \gamma_{\max} V(0)}{g^2 \gamma_{\min}^2} (1 - e^{-g \gamma_{\min} t_f / 2})^2, \quad (39)$$

scalars γ_{\min} , γ_{\max} , and $V(0)$ are defined in Lemma 2 with $\theta(0) = z^f(0) - M^{-1}z(0)$,

$$\tilde{F}_{Ri} = D_i^T \tilde{F}_i D_i \tilde{B}_i R_i^{-1} \tilde{B}_i^T D_i^T \tilde{F}_i D_i, \quad (40a)$$

$$\tilde{M}_i = M_{Fi} + t_f M_i^T \tilde{F}_{Ri} M_i, \quad (40b)$$

$$M_{Fi} = M^T (d_i \otimes I_r) (I + S_{ii})^{-1} \tilde{F}_{ii} (I + S_{ii})^{-1} (d_i^T \otimes I_r) M, \quad (40c)$$

$$\tilde{F}_{ii} = (d_i^T \otimes I_r) \tilde{F}_i (d_i \otimes I_r), \quad (40d)$$

$$M_i = I - (d_i \otimes I_r) (I + S_{ii})^{-1} (d_i^T \otimes I_r) M, \quad (40e)$$

$$W_{\max} = \max_{0 \leq t \leq t_f} \|W(t)\|_2, \quad W(t) = \sum_{j=1}^N \tilde{B}_j R_j^{-1} \tilde{B}_j^T D_j^T \tilde{F}_j D_j, \quad (40f)$$

and matrix S_{ii} is defined in (24).

Proof: Finding ϵ that satisfies inequalities (23) involves three steps. The first step is to obtain each agent's performance index value J_i when all agents choose control strategies (22). The second step is to obtain the best performance index value J_i^* each agent can achieve by unilaterally deviating from control strategy (22). The final step is to find ϵ_i such that $J_i \leq J_i^* + \epsilon_i$.

As the first step, agent i 's performance index value J_i under control strategy \hat{u}_i can be expressed as

$$J_i = \frac{1}{2} \sum_{e_{ij} \in \mathcal{E}} \|y_i(t_f) - y_j(t_f) - \mu_{ij}\|_{F_{ij}}^2 + \frac{1}{2} \int_0^{t_f} \|\hat{u}_i(t)\|_{R_i(t)}^2 dt \quad (41)$$

As the second step, if all agents but agent i chose strategy (22), the control strategy of agent i based on (14) to achieve the best performance index value can be expressed as

$$\begin{aligned} u_i^*(t) &= -R_i^{-1} B^T \lambda_i^*(t) \\ &= -R_i^{-1} B^T \Phi^T(t_f, t) \lambda_i^*(t_f) \\ &= -R_i^{-1} B^T \Phi^T(t_f, t) \left\{ \sum_{e_{ij} \in \mathcal{E}} L_{ij} [x_i^*(t_f) - x_j(t_f)] - \mu_i \right\} \end{aligned} \quad (42)$$

where $x_i^*(t_f)$ is agent i 's terminal state under strategy $u_i^*(t)$ in (42), $x_i(t_f)$ is agent i 's terminal state under strategy \hat{u}_i , and $x_i^*(t) = x_i(t) - \Delta x_i(t)$. Agent i 's performance index value J_i^* under $u_i^*(t)$ can be expressed as

$$J_i^* = \frac{1}{2} \sum_{e_{ij} \in \mathcal{E}} \|y_i^*(t_f) - y_j(t_f) - \mu_{ij}\|_{F_{ij}}^2 + \frac{1}{2} \int_0^{t_f} \|u_i^*(t)\|_{R_i(t)}^2 dt$$

As the third step, the value of $J_i - J_i^*$ can be calculated as follows

$$\begin{aligned}
J_i - J_i^* &= \frac{1}{2} \sum_{e_{ij} \in \mathcal{E}} [\|y_i(t_f) - y_j(t_f) - \mu_{ij}\|_{F_{ij}}^2 - \|y_i^*(t_f) - y_j(t_f) - \mu_{ij}\|_{F_{ij}}^2] \\
&\quad + \frac{1}{2} \int_0^{t_f} [\|u_i(t)\|_{R_i(t)}^2 - \|u_i^*(t)\|_{R_i(t)}^2] dt \\
&= \frac{1}{2} \sum_{e_{ij} \in \mathcal{E}} [y_i(t_f) - y_i^*(t_f)]^T F_{ij} [y_i(t_f) + y_i^*(t_f) - 2y_j(t_f) - 2\mu_{ij}] \\
&\quad + \frac{1}{2} \int_0^{t_f} [u_i(t) - u_i^*(t)] R_i [u_i(t) + u_i^*(t)] dt \\
&= \frac{1}{2} \sum_{e_{ij} \in \mathcal{E}} \Delta x_i(t_f)^T C^T F_{ij} [C \Delta x_i(t_f) + 2y_i^*(t_f) - 2y_j(t_f) - 2\mu_{ij}] \\
&\quad + \frac{1}{2} \int_0^{t_f} \Delta u_i^T R_i (\Delta u_i + 2u_i^*) dt \\
&= \frac{1}{2} \sum_{e_{ij} \in \mathcal{E}} \Delta x_i(t_f)^T L_{ij} \Delta x_i(t_f) + \sum_{e_{ij} \in \mathcal{E}} \Delta x_i(t_f)^T C^T F_{ij} [y_i^*(t_f) - y_j(t_f) - \mu_{ij}] \\
&\quad + \int_0^{t_f} \Delta u_i^T R_i u_i^* dt + \frac{1}{2} \int_0^{t_f} \Delta u_i^T R_i \Delta u_i dt
\end{aligned} \tag{43}$$

where $\Delta u_i(t) = u_i(t) - u_i^*(t)$. Since

$$\Delta \dot{x}_i = \dot{x}_i - \dot{x}_i^* = Ax_i + Bu_i - Ax_i^* - Bu_i^* = A\Delta x_i + B\Delta u_i \tag{44}$$

and $\Delta x_i(0) = x_i(0) - x_i^*(0) = 0$, terminal value $\Delta x_i(t_f)$ can be expressed as

$$\Delta x_i(t_f) = \int_0^{t_f} \Phi(t_f, t) B(t) \Delta u_i(t) dt \tag{45}$$

Substituting (42) and (45) into (43) yields

$$J_i - J_i^* = \frac{1}{2} \Delta x_i^T(t_f) \left(\sum_{e_{ij} \in \mathcal{E}} L_{ij} \right) \Delta x_i(t_f) + \frac{1}{2} \int_0^{t_f} \Delta u_i^T R_i \Delta u_i dt, \tag{46}$$

which is always nonnegative. In order to find the upper bound of $(J_i - J_i^*)$, we need to find the value of $\Delta x_i(t_f)$ and Δu_i . Given the value of $\Delta z_i(t_f)$ in (34) and Δu_i in (37), the upper bound of $(J_i - J_i^*)$ can be derived from (46) as follows:

$$\begin{aligned}
J_i - J_i^* &= \frac{1}{2} \Delta z_i^T(t_f) \tilde{F}_{ii} \Delta z_i(t_f) + \frac{1}{2} \int_0^{t_f} \Delta u_i^T R_i \Delta u_i dt \\
&= \frac{1}{2} \left(\int_0^{t_f} W \theta dt \right)^T M_{Fi} \left(\int_0^{t_f} W \theta dt \right) + \frac{1}{2} \int_0^{t_f} v_i^T \tilde{F}_{Ri} v_i dt
\end{aligned} \tag{47}$$

where matrix M_{Fi} is defined in (40c) and matrix \tilde{F}_{Ri} is defined in (40a). Since

$$\begin{aligned} \frac{1}{2} \int_0^{t_f} v_i^T \tilde{F}_{Ri} v_i dt &= \frac{1}{2} \int_0^{t_f} \left(\theta + M_i \int_0^{t_f} W \theta dt \right)^T \tilde{F}_{Ri} \left(\theta + M_i \int_0^{t_f} W \theta dt \right) dt \\ &\leq \int_0^{t_f} \theta^T \tilde{F}_{Ri} \theta dt + t_f \left(\int_0^{t_f} W \theta dt \right)^T M_i^T \tilde{F}_{Ri} M_i \left(\int_0^{t_f} W \theta dt \right) \end{aligned} \quad (48)$$

Substituting (48) into (47) yields

$$\begin{aligned} J_i - J_i^* &\leq \int_0^{t_f} \theta^T \tilde{F}_{Ri} \theta dt + \left(\int_0^{t_f} W \theta dt \right)^T \tilde{M}_i \left(\int_0^{t_f} W \theta dt \right) \\ &\leq \|\tilde{F}_{Ri}\|_2 \int_0^{t_f} \|\theta\|_2^2 dt + \|\tilde{M}_i\|_2 W_{\max}^2 \left(\int_0^{t_f} \|\theta\|_2 dt \right)^2 \end{aligned}$$

where matrix \tilde{M}_i is defined in (40b) and W_{\max} is defined in (40f). Recalling Lemma 2, substituting inequalities in (27) into the above inequality yields (39). The maximum value in $\{\varepsilon_1, \dots, \varepsilon_N\}$ will satisfy the inequalities (??). Therefore, online computing strategies (22) form an ε -Nash equilibrium. \blacksquare

Note that as shown in (39), it is clear that the value of ε_i decreases as g becomes larger. Therefore, we can claim that there exists a scalar g such that the online computing strategies in (22) forms an ε -Nash equilibrium that can be arbitrarily close to the Nash equilibrium in (??).

Since differential equation (22a) is asymptotically stable, it is clear that the larger scalar g_i is, the closer \hat{u}_i in (22b) will be to actual Nash equilibrium u_i^*

Remark 1: So far, we have not mentioned about the connectivity of the information graph which is very important to ensure that the desired formation can be successfully achieved among the agents. For the distributed Nash strategy design approach itself, it does not require any connectivity requirement on the information graph. However, for formation control purpose, the underlying graph \mathcal{G} is usually assumed to be connected (which means that there exists a directed path from every node to every other node on the graph).

Note that the terminal state estimate will approximate to the ideal Nash equilibrium strategies since the convergence of the linear system (20) is asymptotic in nature. Therefore, the strategies obtained in the above algorithm are also known to form an ϵ -Nash equilibrium. A quantitative analysis is provided in [?].

where $R = \text{diag}\{r_1, \dots, r_N\}$ and

$$M = I + W \otimes (R^{-1}L), \quad (49)$$

Since all the eigenvalues of a Laplacian matrix have nonnegative real parts, all the eigenvalues of matrix M in (49) have positive real parts and matrix M is hence invertible.

the relationship between initial state and terminal state is given by (??)

A special case to consider is when the agent does not receive any information from the other agents. In this situation, according to the expressions in (??) and (??), either the Nash or Pareto strategy becomes zero and hence the agent will keep a constant velocity, which is still a reasonable choice in practice.

Remark 2: One might be interested in the limiting case of open-loop Nash equilibrium. The resulting control trajectory will be the same as the one for closed-loop Nash equilibrium under no state disturbance such that system is asymptotically stable.

Therefore, the following implementation algorithm is proposed.

Algorithm 1:

1. All the agents initialize the state $s_i(0) = p_i(0) + t_f v_i(0)$ for all $i = 1, \dots, N$.
2. For a certain period of time (Δt) , all the agents communicate with each other through the network and update the states h_{si} and h_{vi} according to (20) and obtain the estimate of $s_i(t_f)$ and $v_i(t_f)$ for all $i = 1, \dots, N$.
3. At $t = \Delta t$, all the agents implement the open-loop Nash strategies in (??) with $s_i(t_f)$ and $v_i(t_f)$ replaced by the estimated terminal states $h_{si}(\Delta t)$ and $h_{vi}(\Delta t)$.

IV. ILLUSTRATIVE EXAMPLE

In this section, we consider an illustrative formation control example in two-dimensional space and apply the proposed distributed Nash and Pareto strategy design approaches. The dynamics of the agents are given by (1) with $n = 2$. For the ease of illustration, the horizontal axis is called the x -axis and the vertical axis is called the y -axis. Various scenarios are considered as follows.

Remark 3: If the system is simply a double integrator, we have

$$\begin{bmatrix} p_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} I - \frac{1}{6}t_f^3 L & -t_f I - \frac{1}{2}t_f^2 L \\ \frac{1}{2}t_f^2 L & I + t_f L \end{bmatrix} \begin{bmatrix} p(t_f) \\ v(t_f) \end{bmatrix} - \begin{bmatrix} -\frac{1}{6}t_f^3 I \\ \frac{1}{2}t_f^2 I \end{bmatrix} \mu,$$

Scenario 1: We start with a simple case, where there are three agents with directed communication topology as shown in Figure 1, where agent 1 as the leader does not have information injection from the other agents but only sends its information out to the other agents.



Fig. 1: Three agents with undirected graph

Hence, the Laplacian matrix defined in (5) is given by

$$\mathcal{L} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}.$$

The initial states of the three agents are given by

$$p_1 = [0 \ 0]^T, p_2 = [-3 \ 0]^T, p_3 = [3 \ 0]^T,$$

$$v_1 = [0 \ 2]^T, v_2 = [0 \ 0]^T, v_3 = [0 \ 0]^T,$$

which means their initial positions are equally spaced and agent 1's initial velocity is upward and agents 2 and 3's initial velocities are equal to zero. We assume that the desired formation among the agents is a triangle with the displacement vectors given by

$$\mu_{12} = [2 \ 2]^T, \quad \mu_{13} = [-2 \ 2]^T.$$

In performance indices (4), the parameter r_i is given by $r_1 = r_2 = r_3 = 1$ and $t_f = 3$ for the time interval of the formation control. Since the information graph is directed, we only consider the formation control using the distributed Nash strategy design in this scenario. For comparison purpose, we solve for the actual Nash strategies by directly solving the equations (??) with strategies in (??) substituted in. The Nash strategy components on x -axis and y -axis are shown as blue trajectories in Figure 2 and the agents' terminal states under the Nash strategies are

$$\begin{aligned} s_1(t_f) &= \begin{bmatrix} 0 \\ 6 \end{bmatrix}, & v_1(t_f) &= \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \\ s_2(t_f) &= \begin{bmatrix} -2.20 \\ 3.65 \end{bmatrix}, & v_2(t_f) &= \begin{bmatrix} 0.23 \\ 1.90 \end{bmatrix}, \\ s_3(t_f) &= \begin{bmatrix} 2.20 \\ 3.65 \end{bmatrix}, & v_3(t_f) &= \begin{bmatrix} -0.23 \\ 1.90 \end{bmatrix}. \end{aligned}$$

If we implement the distributed Nash strategy design algorithm described in Algorithm 1 with $\Delta t = 0.1$ and $k_1 = k_2 = k_3 = 20$ in (20), then the final terminal state estimate values are

$$\begin{aligned} h_{s1}(\Delta t) &= \begin{bmatrix} 0 \\ 5.51 \end{bmatrix}, & h_{v1}(\Delta t) &= \begin{bmatrix} 0 \\ 1.84 \end{bmatrix}, \\ h_{s2}(\Delta t) &= \begin{bmatrix} -2.19 \\ 3.16 \end{bmatrix}, & h_{v2}(\Delta t) &= \begin{bmatrix} 0.21 \\ 1.72 \end{bmatrix}, \\ h_{s3}(\Delta t) &= \begin{bmatrix} 2.19 \\ 3.16 \end{bmatrix}, & h_{v3}(\Delta t) &= \begin{bmatrix} -0.21 \\ 1.72 \end{bmatrix} \end{aligned}$$

and the corresponding control strategies (??) with $s_i(t_f)$ and $v_i(t_f)$ replaced by the estimated terminal states $h_{si}(\Delta t)$ and $h_{vi}(\Delta t)$ for $i = 1, 2, 3$ are shown as red circled trajectories in figure 2. We can see that the largest terminal state estimation error is 13.4% and the resulting control inputs slightly deviate from the original Nash strategies. This estimation error can be dramatically reduced with larger values of k_1 , k_2 , and k_3 . Therefore, if we implement the same algorithm with $\Delta t = 0.1$ and $k_1 = k_2 = k_3 = 40$, then the final estimate values becomes

$$\begin{aligned} h_{s1}(\Delta t) &= \begin{bmatrix} 0 \\ 5.89 \end{bmatrix}, & h_{v1}(\Delta t) &= \begin{bmatrix} 0 \\ 1.96 \end{bmatrix}, \\ h_{s2}(\Delta t) &= \begin{bmatrix} -2.20 \\ 3.54 \end{bmatrix}, & h_{v2}(\Delta t) &= \begin{bmatrix} 0.23 \\ 1.86 \end{bmatrix}, \\ h_{s3}(\Delta t) &= \begin{bmatrix} 2.20 \\ 3.54 \end{bmatrix}, & h_{v3}(\Delta t) &= \begin{bmatrix} -0.23 \\ 1.86 \end{bmatrix}. \end{aligned}$$

and the largest terminal state estimation error is 3%. In this case, the corresponding control strategies (??) with $s_i(t_f)$ and $v_i(t_f)$ replaced by the estimated terminal states $h_{si}(\Delta t)$ and $h_{vi}(\Delta t)$ for $i = 1, 2, 3$ are shown as green crossed trajectories in figure 2, which are almost identical to original Nash strategies. The motion trajectories under various control inputs are shown in Figure 3. Graphically, there does not exist much difference between the trajectories under the original Nash strategies and the ones under designed strategies. Quantitatively, we are able to calculate the total terminal formation error with respect to the desired displaced among

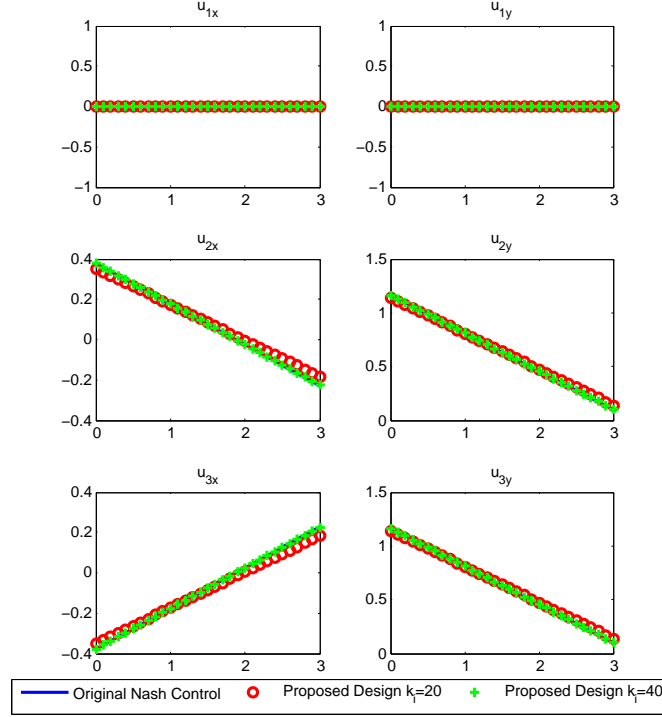


Fig. 2: Nash and proposed designed strategies

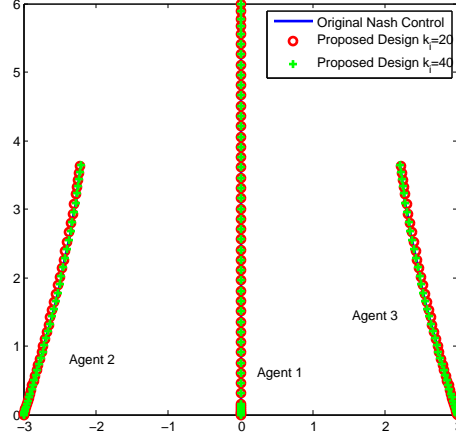


Fig. 3: Motion trajectories under different control strategies

the agents, which is

$$\eta = \sum_{i=1}^3 \|s_1(t_f) - s_i(t_f) - \mu_{1i}\|, \quad (50)$$

where $s_i(t_f)$ is agent i 's terminal position vector under certain control strategies. Therefore,

the total terminal formation errors based on different values of $r_1 = r_2 = r_3 = r$ and $k_1 = k_2 = k_3 = k$ can be summarized into Table I. From the table, it is clear to see that the total

	Nash	$k = 20$	$k = 40$
η ($r = 0.2$)	0.2162	0.2177	0.2162
η ($r = 1$)	0.8164	0.8513	0.8179
η ($r = 5$)	2.6541	2.6920	2.6560

TABLE I: Total terminal formation errors in Scenario 1

terminal formation error decreases as k_i increase because as mentioned, larger k_i will result in more accurate terminal state estimates and more accurate control inputs (with respect to the Nash strategies). It is also clear to see that the total terminal formation error increases as r increases. This is because that if agents put higher penalties on the control effort or energy (i.e., the quadratic term of the control input in (4)), they will try to reduce more on control energy instead of the terminal formation errors in order to minimize their performance indices in (4) during the time period of the formation control. Based on more simulation data, the relationship among η , k , and r can be depicted in Figure 4. As we can see in the figure, when r is small,

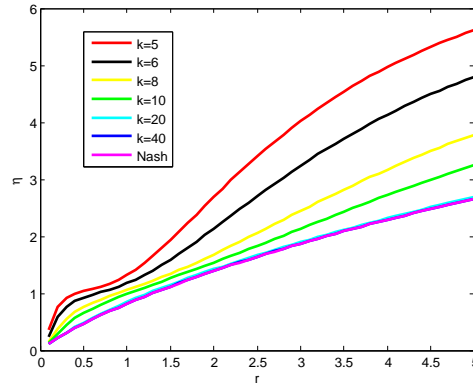


Fig. 4: Relationship among η , r , and k

the variance of the total terminal formation error η among different k is small, and when r is large, the variance of the total terminal formation error η among different k becomes large. Another application of Figure 4 is to determine the lowest update gain k needs to achieve a

desired terminal formation error. For instance, if we want to determine a update gain that can achieve a maximum terminal formation error of 2 given $r = 2$, then we can find that a value of k less than 6 is good choice from the figure. Therefore, this figure provides a intuitive way of choosing a proper update gain, especially to avoid choosing a high gain which is undesired in many real life applications.

Scenario 2: We now consider the scenario where there are three agents with undirected communication topology as shown in Figure 5. With this communication topology, both Nash



Fig. 5: Three agents with undirected graph

and Pareto strategies can be derived. The initial states of the agents and the performance index parameters are the same as the ones in Scenario 1. For both Nash strategy design and Pareto strategy design, the gain k_i in the terminal state estimation equation (20) and (??) are set equal to 40 for all $i = 1, 2, 3$. For the Pareto strategies, the convex parameters in (??) are set to be $\alpha_1 = \alpha_2 = \alpha_3 = 1/3$. The Laplacian matrix for the noncooperative in (5) and the one for the cooperative case in (??) are

$$\mathcal{L} = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathcal{L} = \begin{bmatrix} \frac{4}{3} & -\frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & 0 \\ -\frac{2}{3} & 0 & \frac{2}{3} \end{bmatrix},$$

respectively. The resulting motion trajectories under both Nash and Pareto strategies are shown in Figure 6. The total terminal formation errors under different choices of the convex parameters are summarized in Table II. As we can clearly see in the table, the agents can actually achieve lower formation error using Pareto strategies than using Nash strategies. This is because the agents are assumed to cooperate with each other in the Pareto optimality. Also note that as we place more weight on agent 1's performance index, the total formation error can be reduced. This is because it is much easier to move agent 1 alone to form the desired formation instead of moving all the three agents together.

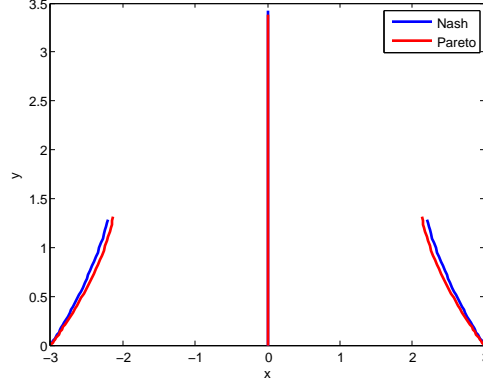


Fig. 6: Motion trajectories under Nash and Pareto strategies

	η
Nash	0.4856
Pareto ($\alpha_1 = \alpha_2 = \alpha_3 = 1/3$)	0.3033
Pareto ($\alpha_1 = 1/6, \alpha_2 = 2/3, \alpha_3 = 1/6$)	0.3128
Pareto ($\alpha_1 = 2/3, \alpha_2 = 1/6, \alpha_3 = 1/6$)	0.1791
Pareto ($\alpha_1 = 0.9, \alpha_2 = 0.05, \alpha_3 = 0.05$)	0.0602

TABLE II: Total terminal formation errors in Scenario 2

V. CONCLUSION

This paper considers the design of open-loop Nash and Pareto strategies for multi-agent formation control problems. The proposed approach utilizes the distributed estimation of the terminal state variables among the agents through network-based information exchange. The convergence rate of the terminal state estimation algorithm is determined by a scalar gain and can be made sufficiently fast to get a better approximation of the real Nash and Pareto strategies. An illustrative example is solved with two scenarios and various comparisons have been made.