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

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 IX.

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:

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

for $i = 1, \cdots, N$. 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 $u_1, \cdots, u_N$ can be looked as strategies, then for the formulated problem, strategies $u_1^N, \cdots, u_N^N$ are the Nash strategies and form a Nash equilibrium if the inequality

$$J_i(u_1^N, \cdots, u_i^N, \cdots, u_N^N) \le J_i(u_1^N, \cdots, u_i, \cdots, u_N^N) \tag{2}$$

holds for all $u_i \in U_i$ and for all $i = 1, \cdots, N$, where U_i is the open-loop control set for agent i. 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 $u_0^P, u_1^P, \cdots, u_N^P$ are the Pareto strategies and form a Pareto optimality if the inequality

$$J_i(u_1^P, \cdots, u_N^P) < J_i(u_1, \cdots, u_N) \tag{3}$$

holds for at least one $i \in \{1, \cdots, N\}$. 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. FREE TERMINAL STATE

If we let the terminal state be free, the performance indices are

$$J_i = \sum_{e_{ij} \in \mathcal{E}} \frac{1}{2} \left[\|p_i(t_f)\|^2 + \|p_i(t_f) - p_j(t_f) - \mu_{ij}\|^2 + + \|v_i(t_f)\|^2 + \|v_i(t_f) - v_j(t_f)\|^2 \right]$$
$$+ \frac{1}{2} \int_0^{t_f} \|u_i\|^2 \mathrm{d}t \tag{4}$$

for $i = 1, \cdots, N$, where $\| \cdot \|$ is the Euclidean norm, μ_{ij} is the desired displacement vector pointing from agent j to agent i, and r_i is a positive scalar. Minimizing performance index J_i means that every agent will try to minimize its locally measured total terminal formation error and terminal velocity error according to the graph while at the same time minimizing its control effort during the time interval.

Theorem 1 (Free Terminal State): 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., $U_i = \{u_i(t, p_0, v_0) \in \mathbf{R} \mid t \in [0, t_f]\}$, there exists a unique Nash equilibrium and corresponding open-loop control strategies are

$$u^*(t, p_0, v_0)$$

$$= \begin{bmatrix} (t - t_f)L & -L \end{bmatrix} \begin{bmatrix} I - \frac{1}{6}t_f^3 L & -t_f I - \frac{1}{2}t_f^2 L \\ \frac{1}{2}t_f^2 L & I + t_f L \end{bmatrix}^{-1} \left(\begin{bmatrix} p_0 \\ v_0 \end{bmatrix} + \begin{bmatrix} -\frac{1}{6}t_f^3 I \\ \frac{1}{2}t_f^2 I \end{bmatrix} \mu \right) + (t_f - t)\mu \tag{5}$$

in vector form, where

$$u_i = \begin{bmatrix} u_1 \\ \cdots \\ u_N \end{bmatrix}, \quad p = \begin{bmatrix} p_1 \\ \cdots \\ p_N \end{bmatrix}, \quad v = \begin{bmatrix} v_1 \\ \cdots \\ v_N \end{bmatrix}, \quad \mu = \begin{bmatrix} \sum_{e_{1j} \in \mathcal{E}} \mu_{1j} \\ \cdots \\ \sum_{e_{Nj} \in \mathcal{E}} \mu_{Nj} \end{bmatrix},$$

and matrix $L = [l_{ij}] \in \mathbb{R}^{N \times N}$ is the graph Laplacian matrix defined by

$$
l_{ij} = \begin{cases} -1 & \text{if } e_{ij} \in \mathcal{E} \text{ for } j \neq i \\ 0 & \text{if } e_{ij} \notin \mathcal{E} \text{ for } j \neq i \\ -\sum_{q=1,q\neq i}^{N} l_{iq} & \text{if } j = i \end{cases} .
\tag{6}
$$

Proof: We define the Hamiltonian for agent i as

$$
H_i = \frac{1}{2}\|u_i\|^2 + \lambda_{pi}^T v_i + \lambda_{vi}^T u_i
\tag{7}
$$

where vectors λ_{pi} and λ_{vi} are the Lagrangian multipliers. Necessary conditions for the existence of open-loop Nash equilibria are

$$
\dot{p}_i = \frac{\partial H_i}{\partial \lambda_{pi}} = v_i,
\tag{8a}
$$

$$
\dot{v}_i = \frac{\partial H_i}{\partial \lambda_{vi}} = u_i
\tag{8b}
$$

$$
\dot{\lambda}_{pi} = -\frac{\partial H_i}{\partial p_i} = 0,
\tag{8c}
$$

$$
\dot{\lambda}_{vi} = -\frac{\partial H_i}{\partial v_i} = -\lambda_{pi},
\tag{8d}
$$

$$
\lambda_{pi}(t_f) = \sum_{e_{ij}\in\mathcal{E}} [p_i(t_f) - p_j(t_f) - \mu_{ij}],
\tag{8e}
$$

$$
\lambda_{vi}(t_f) = \sum_{e_{ij}\in\mathcal{E}} [v_i(t_f) - v_j(t_f)],
\tag{8f}
$$

$$
\frac{\partial H_i}{\partial u_i} = u_i + \lambda_{vi} = 0,
\tag{8g}
$$

for all $i = 1, \cdots, N$. Therefore, the existence of the Nash equilibria is equivalent to the solvability of above equations given initial states $p_i(0)$ and $v_i(0)$ for all $i = 1, \cdots, N$. The uniqueness of the Nash equilibrium is equivalent to the uniqueness of the solution. From (8g), the control vector can be obtained as

$$
u(t) = -\lambda_v(t),
\tag{9}
$$

where $u = [u_1, \cdots, u_N]^T$ and $\lambda_v = [\lambda_{v1}, \cdots, \lambda_{vN}]^T$. Denoting $\lambda_p = [\lambda_{p1}, \cdots, \lambda_{pN}]^T$, state and

co-state equations (8a)-(8d) can be expressed compactly as

$$
\begin{bmatrix} \dot{p} \\ \dot{v} \\ \dot{\lambda}_p \\ \dot{\lambda}_v \end{bmatrix} = \begin{bmatrix} 0 & I & 0 & 0 \\ 0 & 0 & 0 & -I \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -I & 0 \end{bmatrix} \begin{bmatrix} p \\ v \\ \lambda_p \\ \lambda_v \end{bmatrix}
$$

The solution to above first order differential equation is

$$
\begin{bmatrix} p(t) \\ v(t) \\ \lambda_p(t) \\ \lambda_v(t) \end{bmatrix} = \begin{bmatrix} I & tI & 1/6t^3 & -1/2t^2 I \\ 0 & I & 1/2t^2 I & -tI \\ 0 & 0 & I & 0 \\ 0 & 0 & -tI & I \end{bmatrix} \begin{bmatrix} p_0 \\ v_0 \\ \lambda_p(0) \\ \lambda_v(0) \end{bmatrix} \tag{10}
$$

for $t \in [0, t_f]$. The control vector (9) can be now expressed as

$$
u = \begin{bmatrix} tI & -I \end{bmatrix} \begin{bmatrix} \lambda_p(0) \\ \lambda_v(0) \end{bmatrix}. \tag{11}
$$

From (8e)-(8f), we have

$$
\begin{bmatrix} p(t_f) \\ v(t_f) \\ \lambda_p(t_f) \\ \lambda_v(t_f) \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \\ L & 0 \\ 0 & L \end{bmatrix} \begin{bmatrix} p(t_f) \\ v(t_f) \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ I \\ 0 \end{bmatrix} \mu. \tag{12}
$$

Substituting $t = t_f$ and (15) into (10) yields

$$
\begin{bmatrix} p_0 \\ v_0 \\ \lambda_p(0) \\ \lambda_v(0) \end{bmatrix} = \begin{bmatrix} I - \frac{1}{6}t_f^3 L & -t_f I - \frac{1}{2}t_f^2 L \\ \frac{1}{2}t_f^2 L & I + t_f L \\ L & 0 \\ t_f L & L \end{bmatrix} \begin{bmatrix} p(t_f) \\ v(t_f) \end{bmatrix} - \begin{bmatrix} -\frac{1}{6}t_f^3 \\ \frac{1}{2}t_f^2 I \\ I \\ t_f I \end{bmatrix} \mu \tag{13}
$$

The control vector (11) can be further expressed as

$$
u = \begin{bmatrix} (t - t_f)L & -L \end{bmatrix} \begin{bmatrix} p(t_f) \\ v(t_f) \end{bmatrix} + (t_f - t)\mu. \tag{14}
$$

From (13), the relationship between $[p_0, v_0]^T$ and $[p(t_f), v(t_f)]^T$ can be obtained as

$$
\begin{bmatrix} p_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} I - \frac{1}{6}t_f^3 L & -t_f I - \frac{1}{2}t_f^2 L \\ \frac{1}{2}t_f^2 L & I + t_f L \end{bmatrix} \begin{bmatrix} p(t_f) \\ v(t_f) \end{bmatrix} - \begin{bmatrix} -\frac{1}{6}t_f^3 I \\ \frac{1}{2}t_f^2 I \end{bmatrix} \mu,
\tag{15}
$$

which can be rewritten as

$$
\begin{bmatrix} p(t_f) \\ v(t_f) \end{bmatrix} = \begin{bmatrix} I - \frac{1}{6}t_f^3 L & -t_f I - \frac{1}{2}t_f^2 L \\ \frac{1}{2}t_f^2 L & I + t_f L \end{bmatrix}^{-1} \left(\begin{bmatrix} p_0 \\ v_0 \end{bmatrix} + \begin{bmatrix} -\frac{1}{6}t_f^3 \\ \frac{1}{2}t_f^2 I \end{bmatrix} \mu \right).
\tag{16}
$$

The above matrix is invertible because it is product of two invertible matrices as shown below

$$
\begin{bmatrix} I - \frac{1}{6}t_f^3 & -t_f I - \frac{1}{2}t_f^2 L \\ \frac{1}{2}t_f^2 L & I + t_f L \end{bmatrix} = \begin{bmatrix} I & -t_f I \\ 0 & I \end{bmatrix} (I + W \otimes L),
\tag{17}
$$

where $\otimes$ is Kronecker product,

$$
W = \begin{bmatrix} t_f^3/3 I & t_f^2/2 I \\ t_f^2/2 I & t_f I \end{bmatrix},
\tag{18}
$$

and Laplacian matrix L is known to be positive semi-definite. Substituting (16) into (14) yields (5). Therefore, the open-loop Nash equilibrium exists and is unique. ∎

Remark 1: 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.

As we have mentioned, since each agent is only aware of its own performance index and its local communication topology, computing and implementing open-loop Nash strategy (5) conflict with such constraints. In fact, we are able to overcome this difficulty by carrying out distributed state estimation. Specifically, we let all agents exchange certain information across the communication network to estimate terminal variable values in a distributed manner. Toward that, the following theorem is propose.

Theorem 2: If agent i updates its estimation vectors $\hat{p}_i$ and $\hat{v}_i$ in a distributed manner according to

$$
\begin{bmatrix} \dot{\hat{p}}_i \\ \dot{\hat{v}}_i \end{bmatrix} = g_i \left\{ \begin{bmatrix} p_i(0) \\ v_i(0) \end{bmatrix} + \frac{1}{r_i} \begin{bmatrix} (t_f^3/3)\mu_i \\ (t_f^2/2)\mu_i \end{bmatrix} - \begin{bmatrix} \hat{p}_i \\ \hat{v}_i \end{bmatrix} - \frac{1}{r_i} W \sum_{e_{ij} \in \mathcal{E}} \left(\begin{bmatrix} \hat{p}_i \\ \hat{v}_i \end{bmatrix} - \begin{bmatrix} \hat{p}_j \\ \hat{v}_j \end{bmatrix} \right) \right\},
\tag{19}
$$

where g_i is a positive scalar, then

$$\lim_{\tau \to \infty} \begin{bmatrix} \hat{p}_i(\tau) \\ \hat{v}_i(\tau) \end{bmatrix} = \begin{bmatrix} p_i(t_f) \\ v_i(t_f) \end{bmatrix} \tag{20}$$

for $i = 1, \cdots, N$, where $p_i(t_f)$ and $v_i(t_f)$ are the terminal variables specified in (16) under open-loop Nash strategy vector (5).

Proof: Firstly, under open-loop Nash strategies (5), the relationship between initial state and terminal state is given by (15). Substituting (17) yields

$$\begin{bmatrix} I & t_f I \\ 0 & I \end{bmatrix} \begin{bmatrix} p(0) \\ v(0) \end{bmatrix} = M \begin{bmatrix} p(t_f) \\ v(t_f) \end{bmatrix} - \begin{bmatrix} (t_f^3/3)(R^{-1} \otimes I_n) \\ (t_f^2/2)(R^{-1} \otimes I_n) \end{bmatrix} \mu \tag{21}$$

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

$$M = I + W \otimes (R^{-1}L), \tag{22}$$

Since all the eigenvalues of a Laplacian matrix have nonnegative real parts, all the eigenvalues of matrix M in (22) have positive real parts and matrix M is hence invertible. Secondly, combining equations (19) for $i = 1, \cdots, N$ in a more compact vector form yields

$$\begin{bmatrix} \dot{\hat{p}} \\ \dot{\hat{v}} \end{bmatrix} = G \left(\begin{bmatrix} p(0) + t_f v(0) \\ v(0) \end{bmatrix} - M \begin{bmatrix} \hat{p} \\ \hat{v} \end{bmatrix} + \begin{bmatrix} (t_f^3/3)(R^{-1} \otimes I_n) \\ (t_f^2/2)(R^{-1} \otimes I_n) \end{bmatrix} \mu \right), \tag{23}$$

where $\hat{p} = [\hat{p}_1^T \cdots \hat{p}_N^T]^T$, $\hat{v} = [\hat{v}_1^T \cdots \hat{v}_N^T]^T$, $G = \text{blkdiag}\{g_1 I, \cdots, g_N I\}$. Since all the eigenvalues of matrix M has positive real parts, system (23) is asymptotically stable and final convergent values of $\hat{p}$ and $\hat{v}$ are identical to $p(t_f)$ and $v(t_f)$ in equation (16). Therefore, equation (20) holds for all $i = 1, \cdots, N$. ∎

IV. Fixed Terminal State

If the terminal state is fixed as

$$p_i(t_f) - p_j(t_f) = \mu_{ij}, \tag{24}$$

$$v_i(t_f) - v_j(t_f) = 0, \tag{25}$$

for all $e_{ij} \in \mathcal{E}$, then the performance indices are modified as

$$J_i = \frac{1}{2}(\|p_i(t_f)\|^2 + \|v_i(t_f)\|^2) + \frac{1}{2} \int_0^{t_f} \|u_i\|^2 \mathrm{d}t \tag{26}$$

for $i = 1, \cdots, N$, which is essentially a multi-agent minimum fuel problem. This game problem turns out to have multiple Nash equilibria.

Theorem 3 (Fixed Terminal State): For the N-agent linear quadratic differential game of fixed duration $[0, t_f]$ defined by agents' dynamics (1) and performance indices (50), if the information structure of every agent is open-loop, i.e., $U_i = \{u_i(t, p_0, v_0) \in \mathbf{R} \mid t \in [0, t_f]\}$ and the terminal states satisfy constraints (24) and (25), there exist N Nash equilibria (not necessarily distinctive) denoted by $NE_1, \cdots, NE_N$ and open-loop control strategies corresponding to NE_k for $k \in \{1, \cdots, N\}$ are given by

$$u_k^*(t, p_0, v_0) = 0 \tag{27}$$

$$u_{-k}^*(t, p_0, v_0) = \begin{bmatrix} (t - t_f)I & -I \end{bmatrix} (W \otimes L_{-k})^{-1} \cdot$$
$$\left\{ \left(\begin{bmatrix} 1 & t_f \\ 0 & 1 \end{bmatrix} \otimes L_{-k} \right) \begin{bmatrix} p_{-k}(0) \\ v_{-k}(0) \end{bmatrix} + \begin{bmatrix} l_{-k}[p_k(0) + t_f v_k(0)] - \mu_{-k} \\ l_{-k} v_k(0) \end{bmatrix} \right\} \tag{28}$$

where $u_{-k}(t, p_0, v_0)$ is subvector of $u(t, p_0, v_0)$ with $u_k(t, p_0, v_0)$ eliminated, W is defined in (18), L_{-k} is submatrix of L with the kth row and kth column eliminated, $p_{-k}(t_f)$ is subvector of $p(t_f)$ with $p_k(t_f)$ eliminated, $v_{-k}(t_f)$ is subvector of $v(t_f)$ with $v_k(t_f)$ eliminated, l_{-k} is the kth column of L with entry l_{kk} eliminated, μ_{-k} is subvector of μ with μ_k eliminated.

Proof: Given terminal state constraints (24) and (25), we define Lagrangian for agent i as

$$L_i = \sum_{e_{ij} \in \mathcal{E}} \{\gamma_{pij}^T [p_i(t_f) - p_j(t_f) - \mu_{ij}] + \gamma_{vij}^T [v_i(t_f) - v_j(t_f)]\}$$
$$+ \frac{1}{2} \int_0^{t_f} [\|u_i\|^2 + \lambda_{pi}^T(v_i - \dot{p}_i) + \lambda_{vi}^T(u_i - \dot{v}_i)] \mathrm{d}t \tag{29}$$

where vectors γ_{pi}, γ_{vi}, λ_{pi}, and λ_{vi} are Lagrangian multipliers. Total differential of L_i is

$$
\begin{aligned}
dL_i =& \sum_{e_{ij} \in \mathcal{E}} \{ \gamma_{pij}^T \delta p_i(t_f) + [p_i(t_f) - p_j(t_f) - \mu_{ij}] d\gamma_{pij} \\
& + \gamma_{vij}^T \delta v_i(t_f) + [v_i(t_f) - v_j(t_f)]^T d\gamma_{vij} \} + \int_0^{t_f} [u_i^T \delta u_i \\
& + \delta \lambda_{pi}^T (v_i - \dot{p}_i) + \lambda_{pi}^T (\delta v_i - \delta \dot{p}_i) + \delta \lambda_{vi}^T (u_i - \dot{v}_i) + \lambda_{vi}^T (\delta u_i - \delta \dot{v}_i)] \mathrm{dt} \\
=& \sum_{e_{ij} \in \mathcal{E}} \{ \gamma_{pij}^T \delta p_i(t_f) + [p_i(t_f) - p_j(t_f) - \mu_{ij}] d\gamma_{pij} \\
& + \gamma_{vij}^T \delta v_i(t_f) + [v_i(t_f) - v_j(t_f)]^T d\gamma_{vij} \} + \int_0^{t_f} [u_i^T \delta u_i \\
& + \delta \lambda_{pi}^T (v_i - \dot{p}_i) + \lambda_{pi}^T \delta v_i + \delta \lambda_{vi}^T (u_i - \dot{v}_i) + \lambda_{vi}^T \delta u_i] \mathrm{dt} - \int_0^{t_f} \lambda_{pi}^T \mathrm{d}\delta p_i - \int_0^{t_f} \lambda_{vi}^T \mathrm{d}\delta v_i \\
=& \sum_{e_{ij} \in \mathcal{E}} \{ \gamma_{pij}^T \delta p_i(t_f) + [p_i(t_f) - p_j(t_f) - \mu_{ij}] d\gamma_{pij} \\
& + \gamma_{vij}^T \delta v_i(t_f) + [v_i(t_f) - v_j(t_f)]^T d\gamma_{vij} \} + \int_0^{t_f} [u_i^T \delta u_i \\
& + \delta \lambda_{pi}^T (v_i - \dot{p}_i) + \lambda_{pi}^T \delta v_i + \delta \lambda_{vi}^T (u_i - \dot{v}_i) + \lambda_{vi}^T \delta u_i] \mathrm{dt} \\
& - [\lambda_{pi}^T \delta p_i]_0^{t_f} + \int_0^{t_f} \dot{\lambda}_{pi}^T \delta p_i \mathrm{dt} - [\lambda_{vi}^T \delta v_i]_0^{t_f} + \int_0^{t_f} \dot{\lambda}_{vi}^T \delta v_i \mathrm{dt} \\
=& \left(\sum_{e_{ij} \in \mathcal{E}} \gamma_{pij} - \lambda_{pi}(t_f) \right)^T \delta p_i(t_f) + \sum_{e_{ij} \in \mathcal{E}} [p_i(t_f) - p_j(t_f) - \mu_{ij}] d\gamma_{pij} \\
& + \left(\sum_{e_{ij} \in \mathcal{E}} \gamma_{vij} - \lambda_{vi}(t_f) \right)^T \delta v_i(t_f) + \sum_{e_{ij} \in \mathcal{E}} [v_i(t_f) - v_j(t_f)]^T d\gamma_{vij} + \int_0^{t_f} [(u_i + \lambda_{vi})^T \delta u_i \\
& + \delta \lambda_{pi}^T (v_i - \dot{p}_i) + (\lambda_{pi} + \dot{\lambda}_{vi})^T \delta v_i + \delta \lambda_{vi}^T (u_i - \dot{v}_i) + \dot{\lambda}_{pi}^T \delta p_i] \mathrm{dt}
\end{aligned}
$$

Therefore, it is straight forward to obtain necessary conditions for existence of open-loop Nash

equilibria with terminal state fixed

$$\dot{p}_i = v_i, \tag{30a}$$

$$\dot{v}_i = u_i \tag{30b}$$

$$\dot{\lambda}_{pi} = 0, \tag{30c}$$

$$\dot{\lambda}_{vi} = -\lambda_{pi}, \tag{30d}$$

$$\lambda_{pi}(t_f) = \sum_{e_{ij} \in \mathcal{E}} \gamma_{pij}, \tag{30e}$$

$$\sum_{e_{ij} \in \mathcal{E}} [p_i(t_f) - p_j(t_f) - \mu_{ij}] = 0 \tag{30f}$$

$$\lambda_{vi}(t_f) = \sum_{e_{ij} \in \mathcal{E}} \gamma_{vij}, \tag{30g}$$

$$\sum_{e_{ij} \in \mathcal{E}} [v_i(t_f) - v_j(t_f)] = 0 \tag{30h}$$

$$u_i + \lambda_{vi} = 0, \tag{30i}$$

Solution to differential equations (30a)-(30d) is the same as (10)

$$\begin{bmatrix} p(t) \\ v(t) \\ \lambda_p(t) \\ \lambda_v(t) \end{bmatrix} = \begin{bmatrix} I & tI & 1/6t^3 & -1/2t^2 I \\ 0 & I & 1/2t^2 I & -tI \\ 0 & 0 & I & 0 \\ 0 & 0 & -tI & I \end{bmatrix} \begin{bmatrix} p_0 \\ v_0 \\ \lambda_p(0) \\ \lambda_v(0) \end{bmatrix} \tag{31}$$

Based on (30e), (30g), and (31), it is straight forward to obtain

$$\lambda_p(0) = \lambda_p(t_f) = \gamma_p \tag{32}$$

$$\lambda_v(0) = \lambda_v(t_f) + t_f\lambda_p(0) = \gamma_v + t_f\gamma_p \tag{33}$$

$$p(t_f) = p_0 + t_f v_0 + \frac{t_f^3}{6}\lambda_p(0) - \frac{t_f^2}{2}\lambda_v(0)$$

$$= p_0 + t_f v_0 + \frac{t_f^3}{6}\gamma_p - \frac{t_f^2}{2}(\gamma_v + t_f\gamma_p)$$

$$= p_0 + t_f v_0 - \frac{t_f^3}{3}\gamma_p - \frac{t_f^2}{2}\gamma_v \tag{34}$$

$$v(t_f) = v_0 + \frac{t_f^2}{2}\lambda_p(0) - t_f\lambda_v(0)$$

$$= v_0 + \frac{t_f^2}{2}\gamma_p - t_f(\gamma_v + t_f\gamma_p)$$

$$= v_0 - \frac{t_f^2}{2}\gamma_p - t_f\gamma_v \tag{35}$$

where

$$\gamma_p = \begin{bmatrix} \gamma_{p1} \\ \vdots \\ \gamma_{pN} \end{bmatrix} = \begin{bmatrix} \sum_{e_{1j}\in\mathcal{E}} \gamma_{p1j} \\ \vdots \\ \sum_{e_{Nj}\in\mathcal{E}} \gamma_{pNj} \end{bmatrix}, \gamma_v = \begin{bmatrix} \gamma_{v1} \\ \vdots \\ \gamma_{vN} \end{bmatrix} = \begin{bmatrix} \sum_{e_{1j}\in\mathcal{E}} \gamma_{v1j} \\ \vdots \\ \sum_{e_{1j}\in\mathcal{E}} \gamma_{vNj} \end{bmatrix}.$$

From (30i), the control strategy is obtained as

$$u(t) = -\lambda_v(t) = t\lambda_p(0) - \lambda_v(0) = (t - t_f)\gamma_p - \gamma_v \tag{36}$$

The remaining two necessary conditions to satisfy are (30f) and (30h), which are equivalent to

$$Lp(t_f) = \mu, \tag{37a}$$

$$Lv(t_f) = 0, \tag{37b}$$

where L is the Laplacian matrix and

$$\mu = \begin{bmatrix} \sum_{e_{1j}\in\mathcal{E}} \mu_{1j} \\ \cdots \\ \sum_{e_{Nj}\in\mathcal{E}} \mu_{Nj} \end{bmatrix}$$

Since the graph is connected, rank of Laplacian matrix L is $N - 1$, indicating one variable of vector $p(t_f)$ and one of vector $v(t_f)$ in equation (59) can be arbitrarily chosen. Without loss of

generality, assume $p_k(t_f)$ and $v_k(t_f)$ are free variable entries for $k \in \{1, \cdots, N\}$. Physically, since $p_k(t_f)$ and $v_k(t_f)$ can be arbitrarily chosen, agent k is allowed to choose $u_k(t)$ arbitrarily. Therefore, a natural choice is $u_k(t) = 0$ for all $t \in [0, t_f]$ to achieve minimum fuel consumption. Based on equation (36), $u_k(t) = 0$ leads to

$$\gamma_{pk} = 0, \tag{38}$$

$$\gamma_{vk} = 0, \tag{39}$$

$$p_k(tf) = p_k(0) + t_f v_k(0), \tag{40}$$

$$v_k(tf) = v_k(0). \tag{41}$$

Arrange equation (59) into the following equivalent form

$$L_{-k} p_{-k}(t_f) = \mu_{-k} - l_{-k} p_k(t_f), \tag{42a}$$

$$L_{-k} v_{-k}(t_f) = - l_{-k} v_k(t_f). \tag{42b}$$

Substituting (34) and (35) into (64) yields

$$L_{-k}[p_{-k}(0) + t_f v_{-k}(0)] - \frac{t_f^3}{3} L_{-k} \gamma_{-pk} - \frac{t_f^2}{2} L_{-k} \gamma_{-vk} = \mu_{-k} - l_{-k} p_k(t_f) \tag{43}$$

$$L_{-k} v_{-k}(0) - \frac{t_f^2}{2} L_{-k} \gamma_{-pk} - t_f L_{-k} \gamma_{-vk} = - l_{-k} v_k(t_f) \tag{44}$$

where γ_{-pk} is the subvector of γ_p by eliminating by eliminating γ_{pk}, and γ_{-vk} is the subvector of γ_v by eliminating by eliminating γ_{vk}. A more compact expression is

$$\left(\begin{bmatrix} I & t_f I \\ 0 & I \end{bmatrix} \otimes L_{-k} \right) \begin{bmatrix} p_{-k}(0) \\ v_{-k}(0) \end{bmatrix} - (W \otimes L_{-k}) \begin{bmatrix} \gamma_{-pk} \\ \gamma_{-vk} \end{bmatrix} = \begin{bmatrix} \mu_{-k} - l_{-k} p_k(t_f) \\ -l_{-k} v_k(t_f) \end{bmatrix} \tag{45}$$

Therefore, substituting (62) and (63) into the above equation and solving for γ_{-pk} and γ_{-vk} yields

$$\begin{bmatrix} \gamma_{-pk} \\ \gamma_{-vk} \end{bmatrix} = (W \otimes L_{-k})^{-1} \left\{ \left(\begin{bmatrix} I & t_f I \\ 0 & I \end{bmatrix} \otimes L_{-k} \right) \begin{bmatrix} p_{-k}(0) \\ v_{-k}(0) \end{bmatrix} + \begin{bmatrix} l_{-k}[p_k(0) + t_f v_k(0)] - \mu_{-k} \\ l_{-k} v_k(0) \end{bmatrix} \right\} \tag{46}$$

where L_{-k} is of full rank and hence $(W \otimes L_{-k})$ is invertible. Therefore, agent k's open-loop control strategy $u_k = 0$ and the other agents' open-loop strategies (28) (obtained by and then

into (36)) form a Nash equilibrium. Since there are N choice of $k \in \{1, \cdots, N\}$, there exists N Nash equilibria. ∎

Depending on agents' initial states, the N Nash equilibria might not be distinct, that is, some of them might coincide. For the simplest case, if all agents' initial states are already in the desired formation with the same velocity to maintain it, their open-loop Nash strategies are trivial and all equal to zero. Therefore, there is only one unique Nash equilibrium. As for the case where all agents' initial states are already in desired formation expect for one, there exists two distinct Nash equilibria. One Nash equilibrium consists of a nonzero Nash strategy for the agent out of desired formation and zero ones for the rest $(N-1)$ agents in desired formation. The other Nash equilibrium consists of a zero Nash strategy for the agent out of desired formation and nonzero ones for the rest $(N-1)$ agents in desired formation. In general, one can virtually treat M agents whose initial states are already in desired formation among themselves as one large agent group as if there are only $(N-M+1)$ effective agents in the game. Therefore, the game will only render $(N-M+1)$ distinct equilibria.

Corollary 1: For the N-agent linear quadratic differential game of fixed duration $[0, t_f]$ defined by agents' dynamics (1) and performance indices (50), if the information structure of every agent is open-loop, i.e., $U_i = \{u_i(t, p_0, v_0) \in \mathbf{R} \mid t \in [0, t_f]\}$ and the terminal states satisfy constraints (24) and (25), there exist $(N-M+1)$ distinct Nash equilibria if the number of agents whose initial states are already in desired formation among themselves is M where $1 \leq M \leq N$.

Like Theorem (2), we can also propose distributed estimation for the open-loop Nash equilibrium with fixed terminal states.

Theorem 4: For any $k \in \{1, \cdots, N\}$, if agent i ($i \neq k$) updates its estimation vectors $\hat{\gamma}_{pi}$ and $\hat{\gamma}_{vi}$ in a distributed manner according to

$$\begin{bmatrix} \dot{\hat{\gamma}}_{pi} \\ \dot{\hat{\gamma}}_{vi} \end{bmatrix} = g_i \left(W_1 \sum_{e_{ij} \in \mathcal{E}, j \neq k} \begin{bmatrix} p_i(0) - p_j(0) \\ v_i(0) - v_j(0) \end{bmatrix} - W \sum_{e_{ij} \in \mathcal{E}, j \neq k} \begin{bmatrix} \hat{\gamma}_{pi} - \hat{\gamma}_{pj} \\ \hat{\gamma}_{vi} - \hat{\gamma}_{vj} \end{bmatrix} - \begin{bmatrix} l_{ik} p_k(t_f) - \mu_i \\ l_{ik} v_k(t_f) \end{bmatrix} \right),$$

(47)

where g_i is a positive scalar and

$$W_1 = \begin{bmatrix} I & t_f I \\ 0 & I \end{bmatrix},$$

then

$$\lim_{\tau \to \infty} \begin{bmatrix} \hat{\gamma}_{pi}(\tau) \\ \hat{\gamma}_{vi}(\tau) \end{bmatrix} = \begin{bmatrix} \gamma_{pi} \\ \gamma_{vi} \end{bmatrix} \tag{48}$$

for $i = 1, \cdots, N$, where γ_{pi} and γ_{vi} are the solution to (68).

Proof: Stacking (47) for all $i = 1, \cdots, N$ but $i \neq k$ yields

$$\begin{bmatrix} \dot{\hat{\gamma}}_{-pk} \\ \dot{\hat{\gamma}}_{-vk} \end{bmatrix} = G \left\{ (W_1 \otimes L_{-k}) \begin{bmatrix} p_{-k}(0) \\ v_{-k}(0) \end{bmatrix} - (W \otimes L_{-k}) \begin{bmatrix} \hat{\gamma}_{-pk} \\ \hat{\gamma}_{-vk} \end{bmatrix} - \begin{bmatrix} \mu_{-k} - l_{-k}p_k(t_f) \\ -l_{-k}v_k(t_f) \end{bmatrix} \right\} \tag{49}$$

where $G = \mathrm{blkdiag}\{g_1 I, \cdots, g_N I\}$. Since all the eigenvalues of sub-matrix Laplacian matrix L_{-k} have positive real parts, all the eigenvalues of matrix $(W \otimes L_{-k})$ have positive real parts. Therefore, system (49) is asymptotically stable and final convergent values of $\hat{\gamma}_p$ and $\hat{\gamma}_v$ are identical to $p(t_f)$ and $v(t_f)$ in equation (68). Therefore, equation (48) holds for all $i = 1, \cdots, N$. ∎

V. FREE TERMINAL TIME

In previous section, we let all agents to achieve desired formation at a fixed prescribed terminal time. We are interested in the case where the terminal time is free. Moreover, an interesting case to consider is whether agents can have different terminal times. In such case, agents might be able to achieve desired formation in a sequential manner. Toward that, we consider the following performance index

$$J_i = c_i t_f + \frac{1}{2} \int_0^{t_f} \|u_i\|^2 \mathrm{d}t \tag{50}$$

for $i = 1, \cdots, N$ with terminal state constraints (24) and (25), where $c_i \geq 0$ and terminal time t_f is a variable to be determined. The above performance index is in fact a mixture of minimum time and minimum fuel problem, indicating that all agents seek a Nash equilibrium where they can minimize the terminal formation time while at the same time minimize their energy consumption.

Theorem 5 (Free Terminal Time): For the N-agent linear quadratic differential game defined by agents' dynamics (1) and performance indices (50) with free terminal time t_f for all $i = 1, \cdots, N$, if the information structure of every agent is open-loop, i.e., $U_i = \{u_i(t, p_0, v_0) \in \mathbf{R} \mid t \in [0, t_f]\}$ and the terminal states satisfy constraints (24) and (25), there exist and open-loop control strategies corresponding to NE_k for $k \in \{1, \cdots, N\}$ are given by .

Proof: Given terminal state constraints (24) and (25), we define Lagrangian for agent i as

$$L_i = c_i t_f + \sum_{e_{ij} \in \mathcal{E}} \{\gamma_{pij}^T [p_i(t_f) - p_j(t_f) - \mu_{ij}] + \gamma_{vij}^T [v_i(t_f) - v_j(t_f)]\}$$
$$+ \frac{1}{2} \int_0^{t_f} [\|u_i\|^2 + \lambda_{pi}^T (v_i - \dot{p}_i) + \lambda_{vi}^T (u_i - \dot{v}_i)]\mathrm{d}t \qquad (51)$$

where vectors γ_{pi}, γ_{vi}, λ_{pi}, and λ_{vi} are Lagrangian multipliers. We know from calculus of variation that

$$dp_i(t_f) = \delta p_i(t_f) + \dot{p}_i(t_f)dt_f$$
$$= \delta p_i(t_f) + v_i(t_f)dt_f$$
$$dv_i(t_f) = \delta v_i(t_f) + \dot{v}_i(t_f)dt_f$$
$$= \delta v_i(t_f) + u(t_f)dt_f$$

Total differential of L_i is

$$
\begin{aligned}
dL_i =& c_i dt_f + \sum_{e_{ij}\in\mathcal{E}} \{\gamma_{pij}^T[\delta p_i(t_f) + \dot{p}_i(t_f)dt_f] + [p_i(t_f) - p_j(t_f) - \mu_{ij}]d\gamma_{pij} \\
& + \gamma_{vij}^T[\delta v_i(t_f) + \dot{v}_i(t_f)dt_f] + [v_i(t_f) - v_j(t_f)]^T d\gamma_{vij}\} \\
& + \left\{\frac{1}{2}\|u_i(t_f)\|^2 + \lambda_{pi}^T(t_f)[v_i(t_f) - \dot{p}_i(t_f)] + \lambda_{vi}^T(t_f)[u_i(t_f) - \dot{v}_i(t_f)]\right\} dt_f \\
& + \int_0^{t_f}[u_i^T\delta u_i + \delta\lambda_{pi}^T(v_i - \dot{p}_i) + \lambda_{pi}^T(\delta v_i - \delta\dot{p}_i) + \delta\lambda_{vi}^T(u_i - \dot{v}_i) + \lambda_{vi}^T(\delta u_i - \delta\dot{v}_i)]\mathrm{d}t \\
=& \left\{c_i + \sum_{e_{ij}\in\mathcal{E}}[\gamma_{pij}^T\dot{p}_i(t_f) + \gamma_{vij}^T\dot{v}_i(t_f)]\right\} dt_f + \sum_{e_{ij}\in\mathcal{E}}\{\gamma_{pij}^T\delta p_i(t_f) + [p_i(t_f) - p_j(t_f) - \mu_{ij}]d\gamma_{pij} \\
& + \gamma_{vij}^T\delta v_i(t_f) + [v_i(t_f) - v_j(t_f)]^T d\gamma_{vij}\} + \int_0^{t_f}[u_i^T\delta u_i \\
& + \delta\lambda_{pi}^T(v_i - \dot{p}_i) + \lambda_{pi}^T\delta v_i + \delta\lambda_{vi}^T(u_i - \dot{v}_i) + \lambda_{vi}^T\delta u_i]\mathrm{d}t - \int_0^{t_f}\lambda_{pi}^T\mathrm{d}\delta p_i - \int_0^{t_f}\lambda_{vi}^T\mathrm{d}\delta v_i \\
=& \left\{c_i + \sum_{e_{ij}\in\mathcal{E}}[\gamma_{pij}^T\dot{p}_i(t_f) + \gamma_{vij}^T\dot{v}_i(t_f)]\right\} dt_f + \sum_{e_{ij}\in\mathcal{E}}\{\gamma_{pij}^T\delta p_i(t_f) + [p_i(t_f) - p_j(t_f) - \mu_{ij}]d\gamma_{pij} \\
& + \gamma_{vij}^T\delta v_i(t_f) + [v_i(t_f) - v_j(t_f)]^T d\gamma_{vij}\} + \int_0^{t_f}[u_i^T\delta u_i \\
& + \delta\lambda_{pi}^T(v_i - \dot{p}_i) + \lambda_{pi}^T\delta v_i + \delta\lambda_{vi}^T(u_i - \dot{v}_i) + \lambda_{vi}^T\delta u_i]\mathrm{d}t \\
& - [\lambda_{pi}^T\delta p_i]_0^{t_f} + \int_0^{t_f}\dot{\lambda}_{pi}^T\delta p_i\mathrm{d}t - [\lambda_{vi}^T\delta v_i]_0^{t_f} + \int_0^{t_f}\dot{\lambda}_{vi}^T\delta v_i\mathrm{d}t \\
=& \left\{c_i + \sum_{e_{ij}\in\mathcal{E}}[\gamma_{pij}^T v_i(t_f) + \gamma_{vij}^T u_i(t_f)]\right\} dt_f \\
& + \left(\sum_{e_{ij}\in\mathcal{E}}\gamma_{pij} - \lambda_{pi}(t_f)\right)^T \delta p_i(t_f) + \sum_{e_{ij}\in\mathcal{E}}[p_i(t_f) - p_j(t_f) - \mu_{ij}]d\gamma_{pij} \\
& + \left(\sum_{e_{ij}\in\mathcal{E}}\gamma_{vij} - \lambda_{vi}(t_f)\right)^T \delta v_i(t_f) + \sum_{e_{ij}\in\mathcal{E}}[v_i(t_f) - v_j(t_f)]^T d\gamma_{vij} \\
& + \int_0^{t_f}[(u_i + \lambda_{vi})^T\delta u_i + \delta\lambda_{pi}^T(v_i - \dot{p}_i) + (\lambda_{pi} + \dot{\lambda}_{vi})^T\delta v_i + \delta\lambda_{vi}^T(u_i - \dot{v}_i) + \dot{\lambda}_{pi}^T\delta p_i]\mathrm{d}t
\end{aligned}
$$

Therefore, it is straight forward to obtain necessary conditions for existence of open-loop Nash

equilibria with fixed terminal state and free terminal time

$$\dot{p}_i = v_i, \tag{52a}$$

$$\dot{v}_i = u_i \tag{52b}$$

$$\dot{\lambda}_{pi} = 0, \tag{52c}$$

$$\dot{\lambda}_{vi} = -\lambda_{pi}, \tag{52d}$$

$$\lambda_{pi}(t_f) = \sum_{e_{ij} \in \mathcal{E}} \gamma_{pij}, \tag{52e}$$

$$\sum_{e_{ij} \in \mathcal{E}} [p_i(t_f) - p_j(t_f) - \mu_{ij}] = 0 \tag{52f}$$

$$\lambda_{vi}(t_f) = \sum_{e_{ij} \in \mathcal{E}} \gamma_{vij}, \tag{52g}$$

$$\sum_{e_{ij} \in \mathcal{E}} [v_i(t_f) - v_j(t_f)] = 0 \tag{52h}$$

$$c_i + \sum_{e_{ij} \in \mathcal{E}} \gamma_{pij}^T v_i(t_f) + \sum_{e_{ij} \in \mathcal{E}} \gamma_{vij}^T(t_f) u_i(t_f) + \frac{1}{2} \|u_i(t_f)\|^2 = 0 \tag{52i}$$

$$u_i + \lambda_{vi} = 0, \tag{52j}$$

Solution to differential equations (52a)-(52d) is the same as (10)

$$
\begin{bmatrix} p(t) \\ v(t) \\ \lambda_p(t) \\ \lambda_v(t) \end{bmatrix} =
\begin{bmatrix} I & tI & 1/6t^3 & -1/2t^2 I \\ 0 & I & 1/2t^2 I & -tI \\ 0 & 0 & I & 0 \\ 0 & 0 & -tI & I \end{bmatrix}
\begin{bmatrix} p_0 \\ v_0 \\ \lambda_p(0) \\ \lambda_v(0) \end{bmatrix} \tag{53}
$$

Based on (52e), (52g), and (53), it is straight forward to obtain

$$\lambda_p(0) = \lambda_p(t_f) = \gamma_p \tag{54}$$

$$\lambda_v(0) = \lambda_v(t_f) + t_f \lambda_p(0) = \gamma_v + t_f \gamma_p \tag{55}$$

$$p(t_f) = p_0 + t_f v_0 - \frac{t_f^3}{3} \gamma_p - \frac{t_f^2}{2} \gamma_v \tag{56}$$

$$v(t_f) = v_0 - \frac{t_f^2}{2} \gamma_p - t_f \gamma_v \tag{57}$$

where

$$\gamma_p = \begin{bmatrix} \gamma_{p1} \\ \vdots \\ \gamma_{pN} \end{bmatrix} = \begin{bmatrix} \sum_{e_{1j}\in\mathcal{E}} \gamma_{p1j} \\ \vdots \\ \sum_{e_{Nj}\in\mathcal{E}} \gamma_{pNj} \end{bmatrix}, \gamma_v = \begin{bmatrix} \gamma_{v1} \\ \vdots \\ \gamma_{vN} \end{bmatrix} = \begin{bmatrix} \sum_{e_{1j}\in\mathcal{E}} \gamma_{v1j} \\ \vdots \\ \sum_{e_{1j}\in\mathcal{E}} \gamma_{vNj} \end{bmatrix}.$$

From (52j), the control strategy is obtained as

$$u(t) = -\lambda_v(t) = t\lambda_p(0) - \lambda_v(0) = (t - t_f)\gamma_p - \gamma_v \tag{58}$$

The remaining two necessary conditions to satisfy are (52f) and (52h), which are equivalent to

$$Lp(t_f) = \mu, \tag{59a}$$

$$Lv(t_f) = 0. \tag{59b}$$

Since the rank of Laplacian matrix L is $N-1$, indicating one variable of vector $p(t_f)$ and one of vector $v(t_f)$ in equation (59) can be arbitrarily chosen. Without loss of generality, assume $p_k(t_f)$ and $v_k(t_f)$ are free variable entries for $k \in \{1, \cdots, N\}$. Physically, since $p_k(t_f)$ and $v_k(t_f)$ can be arbitrarily chosen, agent k is allowed to choose $u_k(t)$ arbitrarily. Therefore, a natural choice is $u_k(t) = 0$ for all $t \in [0, t_f]$ to achieve minimum fuel consumption. When $u_k(t) = 0$, condition (52i) becomes

$$0 = c_i + \sum_{e_{ij}\in\mathcal{E}} [\gamma_{pij}^T v_i(t_f) + \gamma_{vij}^T u_i(t_f)]$$

$$= c + \lambda_{pi}^T(t_f)v_i(t_f) + \lambda_{vi}^T(t_f)u_i(t_f)$$

where $c = [c_1, \cdots, c_N]^T$.

Based on equation (36), $u_k(t) = 0$ leads to

$$\gamma_{pk} = 0, \tag{60}$$

$$\gamma_{vk} = 0, \tag{61}$$

$$p_k(t_f) = p_k(0) + t_f v_k(0), \tag{62}$$

$$v_k(t_f) = v_k(0). \tag{63}$$

Arrange equation (59) into the following equivalent form

$$L_{-k}p_{-k}(t_f) = \mu_{-k} - l_{-k}p_k(t_f), \tag{64a}$$

$$L_{-k}v_{-k}(t_f) = -l_{-k}v_k(t_f). \tag{64b}$$

Substituting (34) and (35) into (64) yields

$$L_{-k}[p_{-k}(0) + t_f v_{-k}(0)] - \frac{t_f^3}{3}L_{-k}\gamma_{-pk} - \frac{t_f^2}{2}L_{-k}\gamma_{-vk} = \mu_{-k} - l_{-k}p_k(t_f) \tag{65}$$

$$L_{-k}v_{-k}(0) - \frac{t_f^2}{2}L_{-k}\gamma_{-pk} - t_f L_{-k}\gamma_{-vk} = -l_{-k}v_k(t_f) \tag{66}$$

where γ_{-pk} is the subvector of γ_p by eliminating by eliminating γ_{pk}, and γ_{-vk} is the subvector of γ_v by eliminating by eliminating γ_{vk}. A more compact expression is

$$\left(\begin{bmatrix} I & t_f I \\ 0 & I \end{bmatrix} \otimes L_{-k}\right)\begin{bmatrix} p_{-k}(0) \\ v_{-k}(0) \end{bmatrix} - (W \otimes L_{-k})\begin{bmatrix} \gamma_{-pk} \\ \gamma_{-vk} \end{bmatrix} = \begin{bmatrix} \mu_{-k} - l_{-k}p_k(t_f) \\ -l_{-k}v_k(t_f) \end{bmatrix} \tag{67}$$

Therefore, substituting (62) and (63) into the above equation and solving for γ_{-pk} and γ_{-vk} yields

$$\begin{bmatrix} \gamma_{-pk} \\ \gamma_{-vk} \end{bmatrix} = (W \otimes L_{-k})^{-1}\left\{\left(\begin{bmatrix} I & t_f I \\ 0 & I \end{bmatrix} \otimes L_{-k}\right)\begin{bmatrix} p_{-k}(0) \\ v_{-k}(0) \end{bmatrix} + \begin{bmatrix} l_{-k}[p_k(0) + t_f v_k(0)] - \mu_{-k} \\ l_{-k}v_k(0) \end{bmatrix}\right\} \tag{68}$$

where L_{-k} is of full rank and hence $(W \otimes L_{-k})$ is invertible. Therefore, agent k's open-loop control strategy $u_k = 0$ and the other agents' open-loop strategies (28) (obtained by and then into (36)) form a Nash equilibrium. Since there are N choice of $k \in \{1, \cdots, N\}$, there exists N Nash equilibria. ∎

VI. CLOSED-LOOP STRATEGIES

VII. INFINITE HORIZON FEEDBACK CASE

The performance indices are

$$J_i = \frac{r_i}{2}\int_0^{t_f}\sum_{e_{ij}\in\mathcal{E}}\frac{1}{2}\left[\|p_i(t_f) - p_j(t_f) - \mu_{ij}\|^2 + \|v_i(t_f) - v_j(t_f)\|^2\right]$$
$$+ \sum_{i=1}^{N}(w_{pi}\|p_i\|^2 + w_{vi}\|v_i\|^2) + \|u_i\|^2 \mathrm{d}t \tag{69}$$

VIII. IMPLEMENTATION ALGORITHM

Based on the above theorem, all the agents can exchange information with each other through the communication network and estimate the unknown terminal values in (**??**). 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 $i = 1, \cdots, N$.

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 (19) and obtain the estimate of $s_i(t_f)$ and $v_i(t_f)$ for all $i = 1, \cdots, N$.

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 $h_{vi}(\Delta t)$.

The above algorithm is fully distributed (except for the knowledge of time instants 0 and Δt) because agent i only need to send out its own state information h_{si} and h_{vi} to the agents that have direct information injection links from agent i and receive the state variables from the agents that have direct information links to agent i.

Remark 2: According to the algorithm, although the agents communicate with each other for a certain period of time, the final state estimate will not be exactly the same as the solution obtained by directly solving the state equations (**??**) with (**??**) since the convergence of the linear system (19) is asymptotic in nature. It is obvious that the scalar coefficient k_i in (19) determines the convergence rate of the equations. Therefore, for better approximation, large value of k_i is preferred in practice. Due to the approximation to the ideal Nash equilibrium strategies, the strategies rendered by the proposed algorithm are also known to form an ϵ-Nash equilibrium in theory. The value of ϵ measures how close the designed strategies are close to the ideal Nash strategies. A quantitative analysis on the value of ϵ is provided in [**?**].

Remark 3: 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 (19) 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 [**?**].

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.

IX. 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.

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. Hence,

Fig. 1: Three agents with undirected graph

the Laplacian matrix defined in (6) 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

$$s_1(t_f) = \begin{bmatrix} 0 \\ 6 \end{bmatrix}, \quad v_1(t_f) = \begin{bmatrix} 0 \\ 2 \end{bmatrix},$$

$$s_2(t_f) = \begin{bmatrix} -2.20 \\ 3.65 \end{bmatrix}, \quad v_2(t_f) = \begin{bmatrix} 0.23 \\ 1.90 \end{bmatrix},$$

$$s_3(t_f) = \begin{bmatrix} 2.20 \\ 3.65 \end{bmatrix}, \quad v_3(t_f) = \begin{bmatrix} -0.23 \\ 1.90 \end{bmatrix}.$$

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 (19), then the final terminal state estimate values are

$$h_{s1}(\Delta t) = \begin{bmatrix} 0 \\ 5.51 \end{bmatrix}, \quad h_{v1}(\Delta t) = \begin{bmatrix} 0 \\ 1.84 \end{bmatrix},$$

$$h_{s2}(\Delta t) = \begin{bmatrix} -2.19 \\ 3.16 \end{bmatrix}, \quad h_{v2}(\Delta t) = \begin{bmatrix} 0.21 \\ 1.72 \end{bmatrix},$$

$$h_{s3}(\Delta t) = \begin{bmatrix} 2.19 \\ 3.16 \end{bmatrix}, \quad h_{v3}(\Delta t) = \begin{bmatrix} -0.21 \\ 1.72 \end{bmatrix}$$

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

$$h_{s1}(\Delta t) = \begin{bmatrix} 0 \\ 5.89 \end{bmatrix}, \quad h_{v1}(\Delta t) = \begin{bmatrix} 0 \\ 1.96 \end{bmatrix},$$

$$h_{s2}(\Delta t) = \begin{bmatrix} -2.20 \\ 3.54 \end{bmatrix}, \quad h_{v2}(\Delta t) = \begin{bmatrix} 0.23 \\ 1.86 \end{bmatrix},$$

$$h_{s3}(\Delta t) = \begin{bmatrix} 2.20 \\ 3.54 \end{bmatrix}, \quad h_{v3}(\Delta t) = \begin{bmatrix} -0.23 \\ 1.86 \end{bmatrix}.$$

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
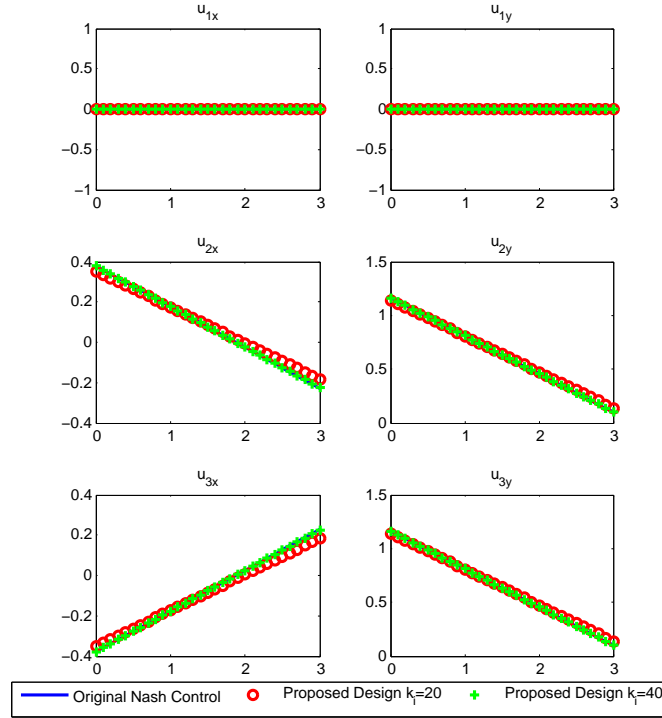

Fig. 2: Nash and proposed designed strategies

shown in Figure 3. Graphically, there does not exist much difference between the trajectories
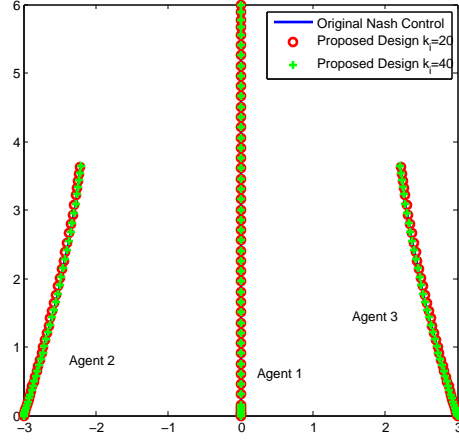

Fig. 3: Motion trajectories under different control strategies

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 the agents, which is

$$\eta = \sum_{i=1}^{3} \|s_1(t_f) - s_i(t_f) - \mu_{1i}\|, \tag{70}$$

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	$k = 20$	$k = 40$
$\eta \ (r = 0.2)$	0.2162	0.2177	0.2162
$\eta \ (r = 1)$	0.8164	0.8513	0.8179
$\eta \ (r = 5)$	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,
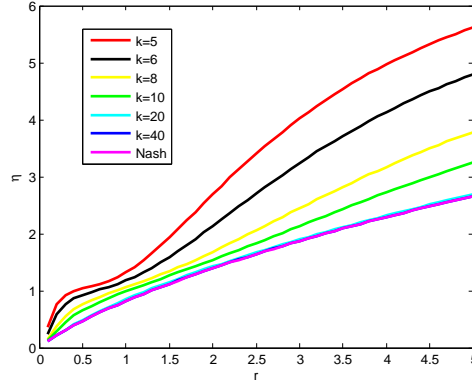

Fig. 4: Relationship among η, r, and k

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 (19) 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 (6) 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
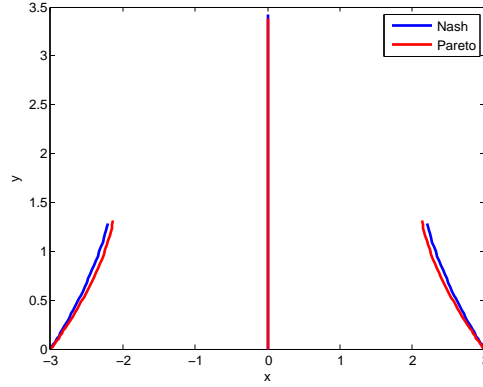

Fig. 6: Motion trajectories under Nash and Pareto strategies

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.

X. 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

	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

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.