Nash Strategies for Pursuit-Evasion Differential Games Involving Limited Observations

WEI LIN, Member, IEEE
Western Digital Corporation
Irvine, CA, USA

ZHIHUA QU, Fellow, IEEE
MARWAN A. SIMAAN, Life Fellow, IEEE
University of Central Florida
Orlando, FL, USA

A linear-quadratic *N*-pursuer single-evader differential game is considered. The evader can observe all the pursuers but pursuers have limited observations of themselves and the evader. The evader implements the conventional feedback Nash strategy and the pursuers implement Nash strategies based on a novel concept of best achievable performance indices. This problem has potential applications in situations where a well-equipped unmanned vehicle is evading several weakly equipped pursuing vehicles. An illustrative example is solved, and several scenarios are presented.

Manuscript received August 26, 2013; revised June 12, 2014, August 10, 2014; released for publication October 12, 2014.

DOI. No. 10.1109/TAES.2014.130569.

Refereeing of this contribution was handled by T. Robertazzi.

This work is partially supported by US National Science Foundation under grants ECCS-1308928 and CCF-0956501. A preliminary version of this paper was presented at the 2013 American Control Conference, June 2013, Washington, D.C., and is referenced in [16].

Authors' addresses: W. Lin, Western Digital Corporation, 3355 Michelson Drive, Suite 100, Irvine, CA 92612; Z. Qu, M. A. Simaan, University of Central Florida, Dept. of Electrical Engineering and Computer Science, Harris Engineering Center, Orlando, FL 32816, E-mail: (simaan@eecs.ucf.edu).

I. INTRODUCTION

Differential game theory has attracted considerable attention in aerospace systems and optimal control literature since the original work of Isaacs [11] in 1965. Since then, there has been a variety of applications of this theory, including pursuit-evasion games [1, 8, 10, 14–16, 19, 21–24] as well as military and aerospace applications [2, 6, 7, 9, 17, 18]. Basically, the pursuit-evasion problem models the process where one or more pursuers tries to chase one or more evaders while the evaders try to escape. Solving a pursuit-evasion game essentially involves developing strategies for both the pursuers and evaders such that their prescribed performance indices are optimized. Historically, pursuit-evasion games were first investigated by Isaacs [11] in the 1960s. Saddle point solutions for a type of zero-sum single-pursuer single-evader games were considered in [10]. Nonzero-sum pursuit-evasion games were introduced and investigated as an example of the Nash strategies in [23] and as an example of the leader-follower Stackelberg strategies in [21, 22]. Two pursuers and one evader games were studied in [8]. In recent years, multiplayer pursuit evasion differential games involving a number of pursuers and a single evader have received considerable attention in the literature [1, 16, 18, 24]. These types of games present numerous interesting challenges when compared to the original single-pursuer single-evader game. For the pursuers, the challenge revolves around designing coordinated strategies to accomplish their common goal of capturing the evader. For the evader, who is clearly outnumbered by the pursuers, the challenge is to maneuver around the pursuers so as to avoid being captured. Conventional multiplayer pursuit-evasion games usually assume that every player has unlimited observations of all the other players at all times and yields feedback strategies that are dependent on the initial states. While this unlimited observations assumption may hold in some cases, a realistic challenge in solving these problems is when some of the players may not have access to unlimited observations of other players. This condition might occur due to factors such as limited sensing capabilities, obstacles in the environment, or severe weather conditions. The conventional differential games approach for solving problems with limited observations yields feedback strategies that depend on the initial states of all the players [12, 13], but the requirement of all pursuers having access to the full information of initial states may not be met in certain applications.

This paper considers an *N*-pursuer single-evader game over a finite time horizon in which only the evader is assumed to have unlimited sensing capability that allows it to observe all the pursuers at all times. Each pursuer, on the other hand, has limited sensing capabilities that allows it to observe the evader and/or other pursuers only if they fall within its sensing range. A practical example of such a situation occurs when a well-equipped unmanned aerial vehicle (UAV) with a very wide range of sensing capability must evade several (possibly a large number of) weakly equipped pursuing UAVs. Because of the limited sensing capabilities of the pursuers, these types of problems need to be treated using an approach different from the probabilistic or artificial intelligence approaches that currently exist in the literature [1, 7, 9, 14, 18, 24]. This paper considers a differential game approach in formulating and solving this problem. This approach builds upon previous results obtained in [15, 16] and overcomes the requirement of full observations of the initial states by all pursuers in implementing their feedback strategies. In [15], a class of distributed game strategies for the pursuers and evader were shown to be inversely optimal with respect to certain performance indices that are typically different from the original indices. These results were then extended in [16] using an optimization approach to derive optimal strategies within this class. In this paper, we further explore the control structure of feedback Nash strategies for multipursuer single-evader differential games with unlimited observations and extend it to the case where only the pursuers have limited observation capabilities. In such a case, the novel concept of best achievable performance indices that utilize the inverse optimality is proposed to determine feedback Nash strategies that can be implemented for the given control structure and are independent of the initial states. This is a very important property of our approach in overcoming a long-standing problem of the dependence on the initial states of the feedback Nash strategies that are derived using the conventional approach.

The remainder of the paper is organized as follows. Problem formulation is presented in Section II. The feedback Nash strategies with unlimited observations are presented in Section III. Pursuers' Nash strategies with limited observations are analyzed and obtained in Section IV. An illustrative example and simulation results with different scenarios are shown in Section V. Concluding remarks are made in Section VI.

II. PROBLEM FORMULATION

In this paper, we consider a differential game problem between a single evader with unlimited sensing capability and N pursuers with limited sensing capabilities. The displacement vector z_i between pursuer i and the evader e as shown in Fig. 1 is defined as

$$z_i = x_e - x_i \qquad \forall i \in \{1, \cdots, N\}, \tag{1}$$

where $x_e \in \mathbb{R}^n$ is the evader's position vector and $x_i \in \mathbb{R}^n$ is pursuer i's position vector.

We assume that a collective objective of the pursuers is to minimize the sum of the weighted distances between the evader and themselves at a terminal time $t_f > 0$ while at the same time minimizing these distances and their control efforts over the time interval $[0, t_f]$. Hence, the group of pursuers tries to minimize the following

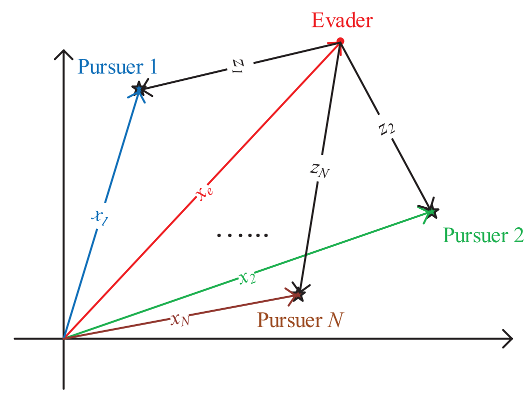


Fig. 1. Displacement vectors.

performance index:

$$J_p = \sum_{j=1}^{N} \frac{f_{p_j}}{2} \left\| z_j(t_f) \right\|^2 + \int_0^{t_f} \sum_{j=1}^{N} \left(\frac{q_{p_j}}{2} \left\| z_j \right\|^2 + \frac{r_j}{2} \left\| u_j \right\|^2 \right) \mathrm{d}t, \tag{2}$$

where u_j is pursuer j's velocity control input, $\|\cdot\|$ is the Euclidean norm, and scalars f_{p_j}, q_{p_j}, and r_j are positive weights for $j = 1, \ldots, N$. On the other hand, we assume that the evader's objective is to maximize the sum of the weighted terminal distances between the pursuers and itself while at the same time maximizing these distances and minimizing its control effort over the time interval $[0, t_f]$. Hence, the evader will try to minimize the performance index:

$$J_e = -\sum_{j=1}^{N} \frac{f_{e_j}}{2} \left\| z_j(t_f) \right\|^2 + \int_0^{t_f} \sum_{j=1}^{N} \left(-\frac{q_{e_j}}{2} \left\| z_j \right\|^2 \right) + \frac{r_e}{2} \left\| u_e \right\|^2 \mathrm{d}t, \tag{3}$$

where u_e is the evader's velocity control input and scalars f_{e_j}, q_{e_j}, and r_e are positive weights for $j = 1, \ldots, N$. To express the system dynamics more compactly, we define the vector $z = [z_1^T; \cdots\ z_N^T]^T$ which, along with (1), yields

$$\dot{z} = B_e u_e + B_p u_p. \tag{4}$$

where matrix $B_e = \mathbf{1}_N \otimes I_n$, $\mathbf{1}_N \in \mathbb{R}^{N\times 1}$ is a vector with all the entries equal to 1, $I_n \in \mathbb{R}^{n\times n}$ is the identity matrix, $\otimes$ is the Kronecker product, $u_p = [u_1^T\ \cdots\ u_N^T]^T$, and $B_p = -I_N \otimes I_n$. The performance indices (2) and (3) can be rewritten as

$$J_p = \frac{1}{2} \left\| z(t_f) \right\|_{F_p}^2 + \frac{1}{2} \int_0^{t_f} (\|z\|_{Q_p}^2 + \|u_p\|_{R_p}^2)\, \mathrm{d}t, \tag{5a}$$

$$J_e = \frac{1}{2} \left\| z(t_f) \right\|_{F_e}^2 + \frac{1}{2} \int_0^{t_f} (\|z\|_{Q_e}^2 + \|u_e\|_{R_e}^2) \mathrm{d}t, \tag{5b}$$

where $\|z\|_F^2 = z^T F z$ and

$$F_p = \text{diag}\{f_{p_1}, \cdots, f_{p_N}\} \otimes I_n,$$

$$F_e = -\text{diag}\{f_{e_1}, \cdots, f_{e_N}\} \otimes I_n$$

$$Q_p = \text{diag}\{q_{p_1}, \cdots, q_{p_N}\} \otimes I_n,$$

$$Q_e = -\text{diag}\{q_{e_1}, \cdots, q_{e_N}\} \otimes I_n$$

$$R_p = \text{diag}\{r_1, \cdots, r_N\} \otimes I_n,$$

$$R_e = r_e \otimes I_n,$$

and "diag" stands for "diagonal matrix." Hence, given the system dynamics in (4) and performance indices in (5), a differential nonzero-sum game between the group of pursuers and the evader is formed. A Nash equilibrium solution (u_p^*, u_e^*) for this game is defined by the following inequalities [23]:

$$J_p(u_p^*, u_e^*) \le J_p(u_p, u_e^*), \quad \forall u_p \in U_p \tag{6a}$$

$$J_e(u_p^*, u_e^*) \le J_e(u_p^*, u_e), \quad \forall u_e \in U_e, \tag{6b}$$

where U_p and U_e are the admissible strategy sets for the pursuers and evader, respectively.

The admissible strategy sets U_p and U_e are closely related to the information structures of the pursuers and evader. We assume that both the group of pursuers and evader have feedback information structures, which means that their admissible strategies will be functions of the displacement vector z at every instant of time t during the game. Furthermore, since the evader's observation range is assumed to be sufficiently wide, its admissible strategy can be implemented as a full feedback of $z_1, \ldots, z_N$. However, since each pursuer has only a limited observation range, its admissible strategy can only be implemented as a partial feedback of $z_1, \ldots, z_N$. We call these strategies as "structured" and refer to them as u_{pi}^s for $i = 1, \ldots, N$.

Mathematically, to accurately model the limited observation capabilities of the pursuers, we assume that pursuer i has a sensing range defined by a sensing radius $r_i > 0$. If the Euclidean distance between pursuer i and the evader is less than or equal to r_i, i.e., $\|z_i\| = \|x_i - x_e\| \le r_i$, then pursuer i is able to observe the evader, otherwise, pursuer i cannot observe the evader. Consequently, we define a binary scalar function $h_i(t)$ to represent pursuer i's ability to observe the evader at time t as follows:

$$h_i(t) = \begin{cases} 1 & \text{if } \|z_i\| \le r_i \\ 0 & \text{if } \|z_i\| > r_i \end{cases}. \tag{7}$$

Similarly, if the Euclidean distance between pursuer i and pursuer j is less than or equal to r_i, i.e., $\|z_i - z_j\| = \|x_i - x_j\| \le r_i$, then pursuer i is able to observe pursuer j, otherwise, pursuer i cannot observe pursuer j. Consequently, we can use the Laplacian matrix [5] (a widely used tool in cooperative control theory [3, 20]) to describe the observations among the pursuers at every instant of time t. The Laplacian matrix is denoted by $L(t) = [L_{ij}(t)] \in \mathbb{R}^{N \times N}$ where

$$L_{ij}(t) = \begin{cases} -1 & \text{if } \|z_i - z_j\| \le r_i \text{ for } j \neq i \\ 0 & \text{if } \|z_i - z_j\| > r_i \text{ for} j \neq i \\ -\sum_{l=1, l\neq i}^{N} L_{il} & \text{if } j = i \end{cases} \tag{8}$$

for $i, j = 1, \ldots, N$.

It is important to note at this point that although the evader has sufficiently wide observation range to observe all of the pursuers at every instant of time, we will assume that it has no information on the individual pursuers' observation radii $r_1, \ldots, r_N$ or the capabilities of the pursuers to obverse each other. Therefore, we assume that during the game process, the evader has no knowledge of the existence of limited observations among the pursuers and the overall information topology. On the other hand, for the pursuers, we assume that all are aware of their limited observation capabilities as well as the evader's unlimited observation capability.

III. FEEDBACK NASH STRATEGIES WITH UNLIMITED OBSERVATIONS

For the pursuit-evasion problem described in (4) and performance indices (5), the standard feedback Nash strategies for all players can be determined as [23]:

$$u_p^* = -R_p^{-1} B_p^T P_p z \tag{9a}$$

$$u_e^* = -R_e^{-1} B_e^T P_e z, \tag{9b}$$

where matrices P_p and P_e are solutions to the coupled differential Riccati equations

$$\dot{P}_p + Q_p - P_p B_p R_p^{-1} B_p^T P_p - P_p B_e R_e^{-1} B_e^T P_e - P_e B_e R_e^{-1} B_e^T P_p = 0 \tag{10a}$$

$$\dot{P}_e + Q_e - P_e B_p R_p^{-1} B_p^T P_p - P_p B_p R_p^{-1} B_p^T P_e - P_e B_e R_e^{-1} B_e^T P_e = 0. \tag{10b}$$

with boundary condition $P_p(t_f) = F_p$ and $P_e(t_f) = F_e$. To implement the strategies in (9a) and (9b), both the evader and pursuers require unlimited observations because solutions P_e and P_p are matrices of full nonzero entries.

In this paper, the evader is assumed to have full information about the pursuers, hence, it can employ strategy (9a). On the other hand, the pursuers are assumed to have only limited observations and hence cannot implement the strategies in (9b). To see how strategy (9b) can be modified to match the limited observations, we note that (9b) can be rewritten as:

$$
\begin{aligned}
u_{pi}^* &= -R_i^{-1}(-B_i^T)P_{pi}^T z \\
&= R_i^{-1}B_i^T \sum_{j=1}^{N} P_{ij} z_j \\
&= R_i^{-1}B_i^T \left[\sum_{j=1}^{N} P_{ij}\right] z_i + R_i^{-1}B_i^T \sum_{j=1}^{N} P_{ij}(z_j - z_i)
\end{aligned} \tag{11}
$$

for all $i = 1, \ldots, N$, where P_{pi} is the ith nN by n block column in P_p, i.e., $P_p = [P_{p1} \ldots P_{pN}]$, and P_{ij} is the jth n by n block matrix in P_{pi}. The full-information pursuer strategy (11) consists of two parts. The first part represents a feedback of z_i (which is the displacement between pursuer i and the evader), and the second part is the sum of feedbacks $(z_1 - z_i)$ up to $(z_N - z_i)$ (which are the displacement vectors between pursuer i and the rest of pursuers). Essentially, expression (11) reveals that pursuer i will chase the evader not only along its own line of sight toward the evader but also it takes into account all the lines of sight from all the pursues toward the evader. The reason for the pursuers to follow such a strategy is because the evader considers in its strategy the line-of-sight information about all the pursuers.

Inspired by expression of (11), a pursuer with limited observations can only choose a strategy based on the available information but in the form of (11). If, e.g., pursuer i can observe the evader (i.e., $\|z_i\| = \|x_i - x_e\| \le r_i$), the first part in (11) can be implemented and hence should be included. Similarly, if pursuer i can observe some of the other pursuers, the corresponding terms in the second part of (11) are implementable and hence should be included for all j satisfying $\|z_i - z_j\| = \|x_i - x_j\| \le r_i$. Such a modification represents the best possible effort by the pursuers from the information perspective, and, as mentioned earlier, the resulting strategy becomes "structured" (with respect to the available information) and will be discussed in details in the next section. Clearly, this structured strategy no longer forms a Nash equilibrium with evader's strategy (9b). Therefore, a different concept or approach is needed to analyze and design pursuers' strategy under limited observations so that they do form a Nash equilibrium with the evader's strategy (9b), which is the subject of the following section.

IV. FEEDBACK NASH STRATEGIES FOR THE PURSUERS

Motivated by the discussions in the previous section, the following structured feedback strategies are proposed for the pursuers:

$$
u_{pi}^s \triangleq h_i(t)K_{ie}(t)z_i + K_{ip}(t)\sum_{j=1}^{N} L_{ij}(t)[z_j - z_i],
$$

which, using the zero-row-sum property of the Laplacian matrix L, can be reduced to

$$
u_{pi}^s = \underbrace{h_i(t)K_{ie}(t)z_i}_{a} + \underbrace{K_{ip}(t)\sum_{j=1}^{N} L_{ij}(t)z_j}_{b} \quad \forall i = 1, \cdots, N \tag{12}
$$

where the superscript s in u_{pi}^s means that the strategy is structured, scalar $h_i(t)$ is defined in (7), scalar $L_{ij}(t)$ is defined in (8), and matrices $K_{ie} \in \mathbb{R}^{n\times n}$ and $K_{ip} \in \mathbb{R}^{n\times n}$ are feedback gains to be determined. Term a in (12) represents the component for pursuer i to chase the evader directly if it observes the evader (i.e., when $h_i(t) = 1$). Term b in (12) is also known as a cooperative protocol, and it ensures that the pursuer also uses all the available information about the rest of pursuers that it can observe in addition to the evader. When pursuer i is unable to observe the evader (i.e., when $h_i(t) = 0$), it has no choice but to merely collaborate with the nearby pursuers. Strategy (12) can be expressed in a more compact form as

$$
\begin{aligned}
u_{pi}^s &= K_{ie}[0 \ \cdots \ 0 \ (h_i I_n) \ 0 \ \cdots \ 0]z \\
&\quad + K_{ip}[(L_{i1} I_n) \ \cdots \ (L_{iN} I_n)]z \\
&\triangleq M_i C_i z,
\end{aligned} \tag{13}
$$

where $M_i = [K_{ie} \ \ K_{ip}] \in \mathbb{R}^{n\times 2n}$ and

$$
C_i = \begin{bmatrix} 0 \ \cdots \ 0 \ (h_i I_n) \ 0 \ \cdots \ 0 \\ (L_{i1} I_n) \ \cdots \ \cdots \ (L_{iN} I_n) \end{bmatrix},
$$

where h_i and L_{ij} are defined in (7) and (8), respectively. Therefore, the pursuers' control vector $u_p^s = [(u_{p1}^s)^T \ \cdots \ (u_{pN}^s)^T]^T$ can be written as $u_p^s = M_p z$, where

$$
M_p = [(M_1 C_1)^T \ \cdots \ (M_N C_N)^T]^T. \tag{14}
$$

The problem now reduces to finding a set of matrices $M_1^*, \cdots, M_N^*$ that are independent of the initial states such that feedback gain $M_p^* = [(M_1^* C_1)^T \ \cdots \ (M_N^* C_N)^T]^T$ and the resulting pursuers' strategy

$$
u_p^{s*} = M_p^* z \tag{15}
$$

can still form a Nash equilibrium with the evader's strategy u_e^* in (9b). Clearly, since the structure of u_p^s is in the output feedback type with N output feedback channels $C_1 z, \ldots, C_N z$, one possible approach to finding matrices $M_1^*, \cdots, M_N^*$ that ensure that (9b) and (15) form a Nash equilibrium is to adopt the widely used optimal output feedback or structured control design approach in [12, 13], which is also applicable to differential games. The basic idea of this approach when applied to the optimal control case is to parameterize the control inputs with given structure and optimize the feedback gain directly with respect to the given performance index. However, for a game problem, because every player's control input has influence on other players' performance indices values, all the players' structured controls and the corresponding feedback gains need to be simultaneously parameterized and optimized with respect to the given set of performance indices to obtain the Nash equilibrium. This cannot be implemented in our game setup because, as mentioned earlier, the evader will be implementing the Nash strategy (9b) and hence it is not possible to simultaneously parameterize and optimize it along with the strategy of the pursuers to form a Nash equilibrium. In what follows, we propose a novel Nash strategy design approach based on a concept of "best achievable performance indices."

Because the evader's Nash strategy in (9b) is fixed, our approach is to find a class of performance indices parameterized by the feedback gain M_p with respect to which the evader's strategy in (9b) and pursuers' strategy in (15) form a Nash equilibrium. Within this class, we then find one set of performance indices called the best achievable performance indices such that they are as close as possible to the original indices given in (5b). The pursuers' Nash strategy is then chosen to be the one with a matrix $M_p = M_p^*$ corresponding to the best achievable performance indices. To achieve this goal, we first need to find the class of performance indices corresponding to the strategies (9b) and (15). Therefore, we propose the following theorem.

THEOREM 1 For the pursuit-evasion game described by system dynamics (4) and performance indices (5), for an arbitrary set of matrices $M_1, \ldots, M_N$, the strategies u_e^* in (9b) and $u_p^s = M_p z$ form a Nash equilibrium:

$$J_p^s(u_p^s, u_e^*) \le J_p^s(u_p, u_e^*), \quad \forall u_p \in U_p \tag{16a}$$

$$J_e^s(u_p^s, u_e^*) \le J_e^s(u_p^s, u_e), \quad \forall u_e \in U_e, \tag{16b}$$

with respect to performance indices

$$J_p^s = \frac{1}{2}\left\|z(t_f)\right\|_{F_p}^2 + \frac{1}{2}\int_0^{t_f}(\|z\|_{Q_p^s}^2 - u_p^T S z - z^T S^T u_p + \|u_p\|_{R_p}^2)\mathrm{d}t, \tag{17a}$$

$$J_e^s = \frac{1}{2}\left\|z(t_f)\right\|_{F_e}^2 + \frac{1}{2}\int_0^{t_f}(\|z\|_{Q_e^s}^2 + \|u_e\|_{R_e}^2)\mathrm{d}t, \tag{17b}$$

where

$$Q_p^s = M_p^T R_p M_p - P_p B_p R_p^{-1} B_p^T P_p + Q_p \tag{18a}$$

$$S = B_p^T P_p + R_p M_p, \tag{18b}$$

$$Q_e^s = -P_e B_p R_p^{-1}(R_p M_p + B_p^T P_p) - (R_p M_p + B_p^T P_p)^T R_p^{-1} B_p^T P_e + Q_e \tag{18c}$$

matrices P_p and P_e are the solutions to (10) and matrix M_p is as given in (14).

PROOF Consider Lyapunov functions

$$V_p = \frac{1}{2} z^T P_p z \quad \text{and} \quad V_e = \frac{1}{2} z^T P_e z. \tag{19}$$

Differentiating V_p in (19) with respect to t and integrating it from 0 to t_f yield

$$\begin{aligned} V_p(t_f) - V_p(0) &= \frac{1}{2}\int_0^{t_f}[-\|z\|_{Q_p^s}^2 - \|u\|_{R_p}^2 + u_p^T S z \\ &\quad + z^T S^T u_p + 2z^T P_p B_e(u_e + R_e^{-1} B_e^T P_e z) \\ &\quad + \left\|u_p - M_p z\right\|_{R_p}^2]\mathrm{d}t. \end{aligned}$$

Hence,

$$J_p^s = V_p(0) + \int_0^{t_f} \frac{1}{2}\left\|u_p - M_p z\right\|_{R_p}^2 + z^T P_s B_e(u_e + R_e^{-1} B_e^T P_e z)\mathrm{d}t, \tag{20}$$

where J_p^s is defined in (17a). Similarly, we can show that

$$J_e^s = V_e(0) + \int_0^{t_f} \frac{1}{2}\left\|u_e + R_e^{-1} B_e^T P_e z\right\|_{R_e}^2 - z^T P_e B_p(u_p - M_p z)\mathrm{d}t, \tag{21}$$

where J_e^s is defined in (17b). Because R_p and R_e are positive definite, it is obvious from (20) and (21) that given performance indices defined in (17), the inequalities in (6) holds for $u_p = M_p z$ and $u_e^* = -R_e^{-1} B_e^T P_e$. Hence, strategies (9b) and (15) form a Nash equilibrium with respect to performance indices in (17). Clearly, if M_p can be written as$M_p = -R_p^{-1} B_p^T P_p$, then J_p^s in (17a) becomes identical to (5b) and J_e^s in (17b) becomes identical to (5b).

The evader's strategy (9b) and pursuers' strategy (15) form a Nash equilibrium with respect to (17) for any matrix M_p. The next step is to find a matrix M_p^* such that the corresponding pair of performance indices in (17) with $M_p = M_p^*$ being closest to the original performance indices J_p and J_e in (5). Toward that end, we propose the following concept of best achievable performance indices.

DEFINITION 1 Given the class of performance indices in (17), if the matrix norms of $(Q_p^s - Q_p)$, S, and $(Q_e^s - Q_e)$ are simultaneously minimized by a feedback matrix $M_p^*(t)$ for all t, then the set of performance indices J_p^{s*} and J_e^{s*} corresponding to this $M_p^*(t)$ are called the best achievable performance indices.

To find the optimal matrix $M_p^*(t)$ corresponding to the best achievable performance indices, we need to solve a multiobjective optimization problem of minimizing $\left\|Q_p^s - Q_p\right\|$, $\|S\|$, and $\left\|Q_e^s - Q_e\right\|$ simultaneously. One way to accomplish this is to minimize a convex combination of these three terms over $[0, t_f]$. Let

$$\phi(M_p(t)) = \int_0^{t_f} H(t)\mathrm{d}t, \tag{22}$$

where

$$\begin{aligned}H(t) &= \alpha_1 \left\| Q_p^s - Q_p \right\|_f^2 + \alpha_2 \|S\|_f^2 + \alpha_3 \left\| Q_e^s - Q_e \right\|_f^2 \\ &= \alpha_1 \mathrm{Tr}[(Q_p^s - Q_p)^2] + \alpha_2 \mathrm{Tr}(S^T S) \\ &\quad + \alpha_3 \mathrm{Tr}[(Q_e^s - Q_e)^2] \end{aligned} \tag{23}$$

where $\|\cdot\|_f$ is the Frobenius norm with the property $\|S\|_f^2 = \mathrm{Tr}(S^T S)$, $0 < \alpha_j < 1$ for $j = 1, 2, 3$, and $\sum_{j=1}^{3} \alpha_j = 1$. The minimization problem reduces to finding a matrix $M_p^*(t)$ such that

$$\phi(M_p^*(t)) \leq \phi(M_p(t)) \quad \forall M_p(t). \tag{24}$$

Because M_p is as defined in (14), the minimization in (24) is actually done with respect to $M_1(t), \ldots, M_N(t)$. This minimization problem is generally quite difficult to solve analytically. A possible numerical approach is to use gradient-based iterative algorithms [4] to find matrices $M_1^*(t), \cdots, M_N^*(t)$. These algorithms will require an expression for the gradient of $H(t)$ with respect to $M_i(t)$ for $i = 1, \ldots, N$. This expression can be determined from (23) as follows:

$$\begin{aligned}\nabla_{M_i} H = (d_i^T \otimes I_n)[4\alpha_1 R_p M_p (Q_p^s - Q_p) \\ + 2\alpha_2 R_p S - 4\alpha_3 B_p^T P_e (Q_e^s - Q_e)] C_i^T \end{aligned} \tag{25}$$

for all $i = 1, \ldots, N$, where M_p is defined in (14) and $d_i \in \mathbb{R}^N$ is a vector with the ith entry equal to 1 and the other entries equal to 0. Note that using this approach, the derived optimal matrices $M_1^*(t), \cdots, M_N^*(t)$ that define the Nash feedback strategies of the pursuers (15) will be independent of the initial states. As mentioned earlier in the introduction section, this is an important property of our approach because any dependence of these matrices on the initial states would render their use as feedback gains impractical. Also note that by varying the coefficients α_1, α_2, and α_3 in (23), a noninferior set of the solutions can be generated. An appropriate choice of these coefficients can be made to place a desired emphasis on the importance of minimizing each of the three terms in (23) as compared to the other two.

It is important to note at this stage that the algorithm can be implemented in a distributed manner. This is clear since it follows from (13) that gradient (25) does not require any state information but only the knowledge of $L(t)$, which, according to (8), has a finite number of values. Optimization algorithm (25) can be performed offline with respect to possible values of the Laplacian matrix $L(t)$. The proposed method can then be implemented in a setting of distributed communication and coordination with low-rate changes of $L(t)$ being relayed to all the pursuers. This can be accomplished because, as the game begins and proceeds, distributed information of relative distance $[z_i(t) - z_j(t)]$ becomes available, the row entries of $L(t)$ reflect evader/other-pursuers moving into/out of the sensor range of individual pursuers, and these mechanical-motion-driven changes in $L(t)$ are much slower than the sampling rate of locally measuring the relative distances.

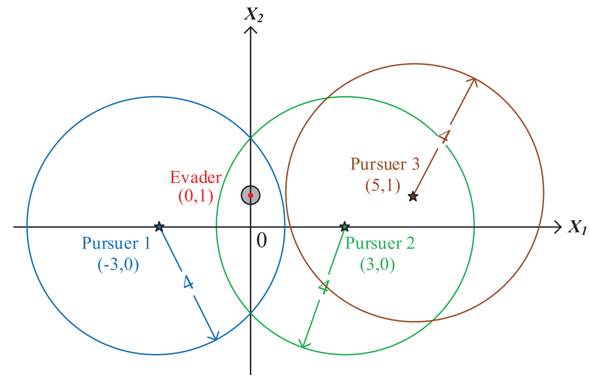


Fig. 2. Initial positions of three pursuers and single evader.

V. ILLUSTRATIVE EXAMPLE

In order to illustrate the Nash strategy obtained using the best achievable performance indices approach, let us consider a three-pursuer single-evader differential game taking place in a planar environment and defined over a time interval [0, 3]. Suppose that $x_i = [x_{i1}^T \; x_{i2}^T]^T \in \mathbb{R}^2$ represents player i's position and $u_{pi} = [u_{pi1}^T \; u_{pi2}^T]^T \in \mathbb{R}^2$ represents player i's velocity control. Hence, in equation (4), we have

$$B_e = \begin{bmatrix} I_2 \\ I_2 \\ I_2 \end{bmatrix} \quad \text{and} \quad B_p = -\begin{bmatrix} I_2 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_2 \end{bmatrix}.$$

The performance indices are given by (5) with $t_f = 3$, $F_p = Q_p = qI_6$, $R_p = I_6$, $F_e = Q_e = I_6$, and $R_e = I_2$, where q is a positive scalar that can be varied to analyze different scenarios. As shown in Fig. 2, we assume that the pursuers' initial positions are $x_1(0) = [-3 \; 0]^T$, $x_2(0) = [3 \; 0]^T$, $x_3(0) = [5 \; 1]^T$, the evader's initial position is $x_e(0) = [0 \; 1]^T$, and the pursuers' sensing radii are the same and equal to 4. Clearly, at $t = 0$, pursuer 1 can only observe the evader, pursuer 2 can observe the evader and pursers 3, and pursuer 3 can only observe pursuer 2. Further, we assume that the evader is captured if the minimum distance between the pursuers and evader is less than a capture radius $\sigma = 0.1$, which is shown as a light black circle centered at the evader in Fig. 2. In this example, we will consider two different scenarios:

1. Scenario 1: $q = 1$. Pursuers put equal emphasis on minimizing their distances to the evader and on minimizing their control effort.
2. Scenario 2: $q = 5$. Pursuers put more emphasis on minimizing their distances to the evader than on minimizing their control effort.

A. Evader's Strategy

The evader solves the coupled differential Riccati equations (10) and implements the corresponding Nash strategy. This yields solutions P_p and P_e in the following

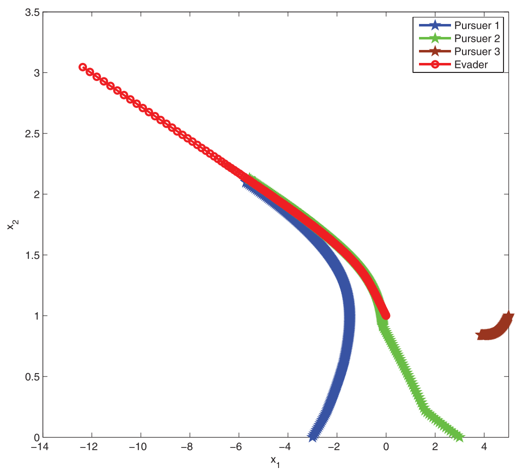


Fig. 3. Motion trajectories of pursuers and evader for Scenario 1 (see animation in Supplemental Video S1).

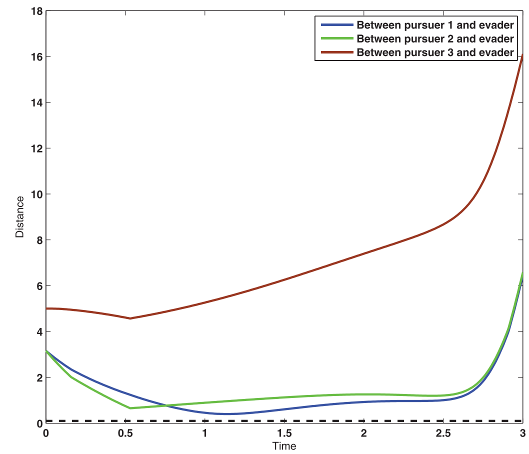


Fig. 4. Distances between pursuers and evader for Scenario 1.

form:

$$P_p = \begin{bmatrix} P_{p1} & P_{p2} & P_{p2} \\ P_{p2} & P_{p1} & P_{p2} \\ P_{p2} & P_{p2} & P_{p1} \end{bmatrix} \otimes I_2,$$

$$P_e = \begin{bmatrix} P_{e1} & P_{e2} & P_{e2} \\ P_{e2} & P_{e1} & P_{e2} \\ P_{e2} & P_{e2} & P_{e1} \end{bmatrix} \otimes I_2.$$

Hence, the evader's feedback Nash strategies (9b) in terms of z_1, z_2, and z_3 can be expressed as

$$u_e^* = (P_{e1} + 2P_{e2})(z_1 + z_2 + z_3).$$

B. Pursuers' Strategies

To derive the pursuers' strategy, we assume that for implementation purpose, the pursuers perform sensing only at discrete instants of time $t_0, t_1, \ldots, t_{299}$, where $t_0 = 0$, $t_{300} = t_f = 3$. Since $t_j - t_{j-1} = 0.01$ is quite small for all $j = 1, \ldots, 300$, we assume that the observations among the players can be regarded to be constant within such a small time interval $(t_j - t_{j-1})$. We also assume that the pursuers will carry out the proposed best achievable performance indices approach with the following (arbitrary) choice of coefficients in (23): $\alpha_1 = 1/4$, $\alpha_2 = 1/2$, and $\alpha_3 = 1/4$.

1) Scenario 1: In this scenario, the motion trajectories of the pursuers and evader over time are shown in Fig. 3. The distances between the pursuers and evader over time are shown in Fig. 4 where the capture radius $\sigma = 0.1$ is shown in terms of a dashed black horizontal line. Clearly, in this scenario, none of the pursuer is able to capture the evader when the final time $t_f = 3$ is reached. Furthermore, the change in the observations among the players is reflected in the changes of $h_i(t)$ in (7) and Laplacian matrix with $L_{ij}(t)$ defined in (8). In this scenario, the values of $h_1(t)$, $h_2(t)$, $h_3(t)$ and the value of the Laplacian matrix $L(t)$ have changed as follows:

$$h_1(t) = \begin{cases} 1 & 0 \le t \le 2.91 \\ 0 & 2.91 < t \le 3 \end{cases}$$

$$h_2(t) = \begin{cases} 1 & 0 \le t \le 2.91 \\ 0 & 2.91 < t \le 3 \end{cases}$$

$$h_3(t) = 0 \quad 0 < t \le 3$$

and

$$L(t) = \begin{cases} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} & 0 \le t \le 0.16 \\ \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} & 0.16 < t \le 0.52 \\ \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} & 0.52 < t \le 3 \end{cases}.$$

The change in $h_i(t)$ means that pursuers 1 and 2 lose observation of the evader after $t = 2.91$ while pursuer 3 was never able to observe the evader for the entire game. The change in the Laplacian matrix essentially means that only pursuers 2 and 3 can observe each other for $t \in [0, 0.16]$, pursuer 2 can observe pursuers 1 and 3 for $t \in (0.16, 0.52]$ while pursuers 1 and 3 cannot observe each other at this time interval, and only pursuers 1 and 2 can observe each other for $t \in (0.52, 3]$.

2) Scenario 2: In this scenario, the motion trajectories of the pursuers and evader are shown in Fig. 5. The distances between the pursuers and evader are shown in Fig. 6 where the capture radius $\sigma = 0.1$ is shown in terms of a dashed black horizontal line. Clearly, in this scenario, pursuer 2 is the first one to capture the evader at $t = 1.2$. During the entire game, the values of $h_1(t)$, $h_2(t)$, $h_3(t)$ and the value of the Laplacian matrix $L(t)$ have changed as

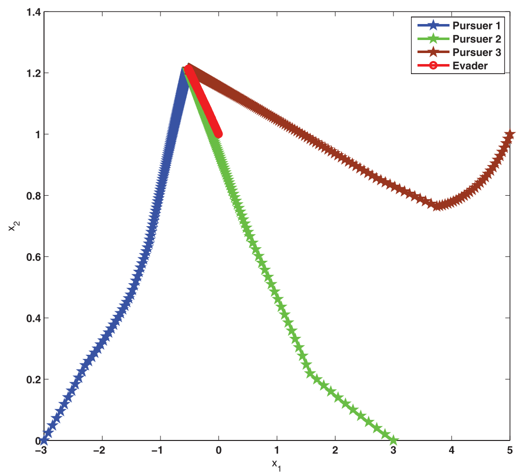


Fig. 5. Motion trajectories of pursuers and evader for Scenario 2 (See animation in Supplemental Video S2).

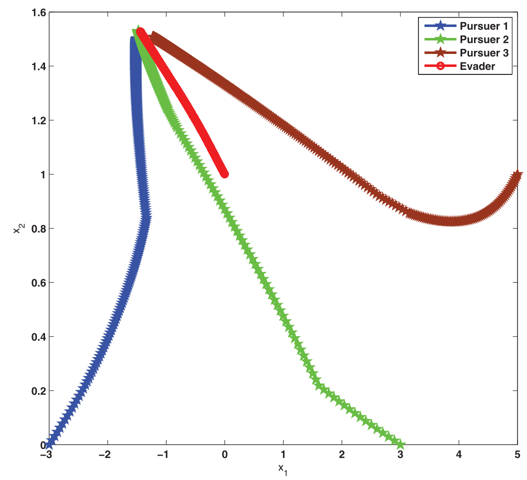


Fig. 7. Trajectories of pursuers and evader when $q = q_c = 1.38$ (See animation in Supplemental Video S3).

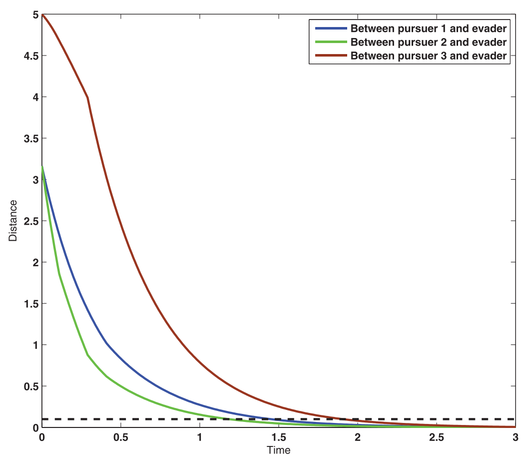


Fig. 6. Distances between pursuers and evader for Scenario 2.

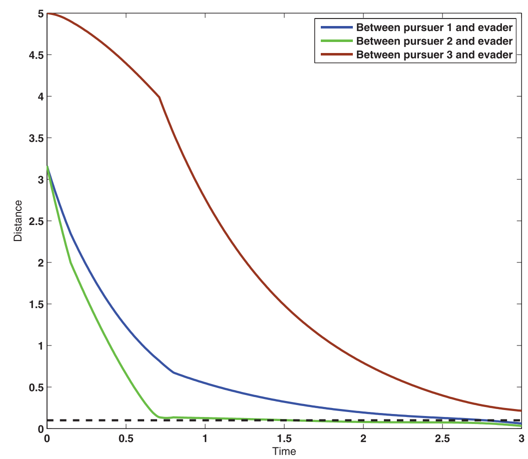


Fig. 8. Distances between pursuers and evader when $q = q_c = 1.38$.

follows:

$$h_1(t) = 1 \quad 0 < t \leq 3$$

$$h_2(t) = 1 \quad 0 < t \leq 3$$

$$h_3(t) = \begin{cases} 0 & 0 \leq t \leq 0.29 \\ 1 & 0.29 < t \leq 3 \end{cases}$$

and

$$L(t) = \begin{cases} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} & 0 \leq t \leq 0.11 \\ \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} & 0.11 < t \leq 0.41 \\ \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} & 0.41 < t \leq 3 \end{cases}$$

The change of $h_i(t)$ means that after $t = 0.29$, all the pursuers are able to observe the evader. The change of the Laplacian matrix means that only pursuers 2 and 3 can observe each other for $t \in [0, 0.11]$, pursuer 2 can observe pursuers 1 and 3 for $t \in (0.11, 0.41]$ while pursuers 1 and 3 cannot observe each other at this time interval, and all the pursuers are able to observe each other for $t \in (0.41, 3]$.

C. Remark

It would be interesting to determine a critical value q_c of q that separates the escape and capture regions of the evader. That is, if $q < q_c$, the evader escapes and if $q \geq q_c$, the evader is captured at a time instant $t \in [0, 3]$. For this game, the critical value of q has been determined to be $q_c = 1.38$. Fig. 7 shows the motion trajectories of the pursuers and evader when $q = q_c = 1.38$. Fig. 8 shows the distances between the pursuers and evader when $q = q_c = 1.38$, where the capture time occurs at $t = 1.55$.

VI. CONCLUSION

In this paper, the problem of deriving feedback Nash strategies that are independent of the initial states for an

N-pursuer single-evader differential game with quadratic performance indices over a finite time horizon is considered. The evader is assumed to have unlimited observations while each pursuer has limited observations based on its own sensing radius. Because the evader is able to observe all the pursuers at all times, it implements the standard feedback Nash strategy derived from the well-known coupled differential Riccati equations approach. However, for the pursuers whose observations are limited, we propose a novel approach for them to implement a collective strategy based on the concept of best achievable performance indices. This approach yields initial-states-independent strategies for the pursuers that satisfy a Nash equilibrium with the evader's strategy with respect to the best achievable performance indices. An illustrative example involving three pursuers and one evader is solved and simulation results corresponding to different scenarios are presented.

REFERENCES

[1] Bopardikar, S. D., Bullo, F., and Hespanha, J. P. On discrete-time pursuit-evasion games with sensing limitations. *IEEE Transactions on Robotics*, **24**, 6 (2008), 1429–1439.

[2] Bugnon, F. J., and Mohler, R. R. A controllability study for a team linear quadratic game. *IEEE Transactions on Aerospace and Electronic Systems*, **23**, 5 (1987), 686–690.

[3] Cao, Y., Yu, W., Ren, W., and Chen, G. An overview of recent progress in the study of distributed multi-agent coordination. *IEEE Transactions on Industrial Informatics*, **9**, 1 (2013), 427–438.

[4] Chong, E. K., and Zak, S. H. *An Introduction to Optimization*. New York: John Wiley & Sons, 2013.

[5] Chung, F. R. *Spectral Graph Theory*, Vol. 92. Providence, RI: AMS Bookstore, 1997.

[6] Cruz, J., Simaan, M. A., Gacic, A., and Liu, Y. Moving horizon Nash strategies for a military air operation. *IEEE Transactions on Aerospace and Electronic Systems*, **38**, 3 (2002), 989–999.

[7] Cruz, J. B., Simaan, M. A., Gacic, A., Jiang, H., Letelliier, B., Li, M., and Liu, Y. Game-theoretic modeling and control of a military air operation. *IEEE Transactions on Aerospace and Electronic Systems*, **37**, 4 (2001), 1393–1405.

[8] Foley, M., and Schmitendorf, W. A class of differential games with two pursuers versus one evader. *IEEE Transactions on Automatic Control*, **19**, 3 (1974), 239–243.

[9] Galati, D. G., and Simaan, M. A. Effectiveness of the Nash strategies in competitive multi-team target assignment problems. *IEEE Transactions on Aerospace and Electronic Systems*, **43**, 1 (2007), 126–134.

[10] Ho, Y., Bryson, A., and Baron, S. Differential games and optimal pursuit-evasion strategies. *IEEE Transactions on Automatic Control*, **AC-10**, 4 (1965), 385–389.

[11] Isaacs, R. *Differential Games*. New York: John Wiley and Sons, 1965.

[12] Levine, W., Johnson, T., and Athans, M. Optimal limited state variable feedback controllers for linear systems. *IEEE Transactions on Automatic Control*, **16**, 6 (1971), 785–793.

[13] Lewis, F., Vrabie, D., and Syrmos, V. *Optimal Control*, 3rd ed. New York: Wiley Interscience, 2012.

[14] Li, D., and Cruz, J. Defending an asset: A linear quadratic game approach. *IEEE Transactions on Aerospace and Electronic Systems*, **47**, 2 (Apr. 2011), 1026–1044.

[15] Lin, W., Qu, Z., and Simaan, M. A. A design of distributed nonzero-sum Nash strategies. In *49th IEEE Conference on Decision and Control*, Atlanta, GA, 2010, 6305–6310.

[16] Lin, W., Qu, Z., and Simaan, M. A. Multi-pursuer single-evader differential games with limited observations. In *Proceedings of the American Control Conference*, Washington, D.C., 2013, 2711–2716.

[17] Liu, Y., Cruz, J. B., and Schumacher, C. J. Pop-up threat models for persistent area denial. *IEEE Transactions on Aerospace and Electronic Systems*, **43**, 2 (2007), 509–521.

[18] Liu, Y., Qi, N., and Tang, Z. Linear quadratic differential game strategies with two-pursuit versus single-evader. *Chinese Journal of Aeronautics*, **25**, 6 (2012), 896–905.

[19] Pachter, M. Simple-motion pursuit-evasion in the half plane. *Computers & Mathematics with Applications*, **13**, 1 (1987), 69–82.

[20] Qu, Z. *Cooperative Control of Dynamical Systems: Applications to Autonomous Vehicles*. London: Springer Verlag, 2009.

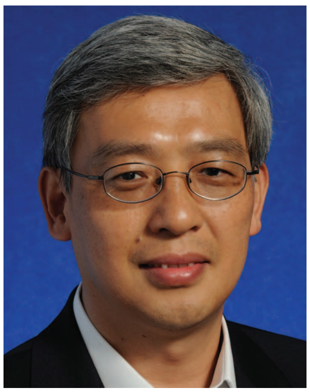
[21] Simaan, M., and Cruz, J. On the Stackelberg solution in the nonzero-sum games. *Journal of Optimization Theory and Applications*, **11**, 5 (1973), 533–555.

[22] Simaan, M., and Cruz, Jr., J. B. Additional aspects of the Stackelberg strategy in nonzero-sum games. *Journal of Optimization Theory and Applications*, **11**, 6 (1973), 613–626.

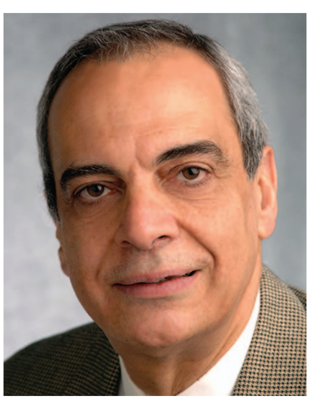
[23] Starr, A., and Ho, Y. Nonzero-sum differential games. *Journal of Optimization Theory and Applications*, **3**, 3 (1969), 184–206.

[24] Wei, M., Chen, G., Cruz, J. B., Haynes, L. S., Pham, K., and Blasch, E. Multi-pursuer multi-evader pursuit-evasion games with jamming confrontation. *AIAA Journal of Aerospace Computing, Information, and Communication*, **4**, 3 (2007), 693–706.

Wei Lin (S'10–M'14) received his B.Eng. degree in automation from Tongji University, Shanghai, China, in 2007, M.S. degree in electrical engineering from Boston University, Boston, MA, in 2009, and Ph.D. degree in electrical engineering from University of Central Florida, Orlando, FL, in 2013 in the area of control systems. He is currently working at Western Digital Corp. in Irvine, CA.



Zhihua Qu (M'90–SM'93–F'09) is the SAIC endowed professor in College of Engineering and Computer Science, professor and chair of Electrical and Computer Engineering at the University of Central Florida, and the director of the FEEDER Center (one of the Department of Energy-funded centers on distributed technologies and smart grid). His recent research focuses on networked systems and cooperative control, distributed optimization, and their applications to energy/power systems and autonomous vehicles. He is a Fellow of the IEEE and AAAS. He served on Board of Governors of IEEE Control Systems Society and as an associate editor for IEEE Transactions on Automatic Control and International Journal of Robotics and Automation (since 1995). Currently, he serves as an associate editor for Automatica and IEEE ACCESS.



Marwan A. Simaan (S'69C–M'72C–SM'79C–F'88C–LF'12) received his Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign. He joined the University of Central Florida in 2008 as the Florida 21st Century Chair and distinguished professor of Electrical Engineering and Computer Science. In 2009, he served as interim dean and then dean of the College of Engineering and Computer Science until 2012. Prior to joining the University of Central Florida, he was the Bell of PA/Bell Atlantic Professor and former chairman of the Department of Electrical Engineering at the University of Pittsburgh. His research covers a broad spectrum of topics in game theory, control, optimization, and signal processing. He has published five books, more than 125 refereed journal papers and book chapters, and more than 210 papers in conference proceedings. Dr. Simaan is a registered Professional Engineer in Pennsylvania. He is a member of US National Academy of Engineering (NAE) and a Fellow of the American Society for Engineering Education (ASEE), the American Association for the Advancement of Science (AAAS), and the American Institute for Medical and Biological Engineering (AIMBE). He currently serves or has served on numerous IEEE boards and committees, including the IEEE Fellow Committee, the IEEE Education Medal Committee, the IEEE Access Editorial Board, the IEEE Proceedings Editorial Board, and others. He was the recipient of two IEEE best paper awards from the Geoscience and Remote Sensing Society (1985) and the Industry Applications Society (1999). In 1995, he was named a Distinguished Alumnus of the Department of Electrical and Computer Engineering, University of Illinois. In 2007, he received the IEEE William E. Sayle II Award for Achievement in Education. In 2008 he received the University of Illinois, College of Engineering Award for Distinguished Service in Engineering.