Nash Pursuit Strategy for Nonzero-Sum MPC Game via Inverse Optimal Control

Tianyu Qiu[a,b,c,d]
Shanghai Jiao Tong University
Shanghai, China
qty2898734013@sjtu.edu.cn

Han Zhang[a,c,d]
Shanghai Jiao Tong University
Shanghai, China
zhanghan_tc@sjtu.edu.cn

Jingchuan Wang[a,b,c,d]
Shanghai Jiao Tong University
Shanghai, China
jchwang@sjtu.edu.cn

***Abstract*—In this paper, a traditional one v.s. one pursuit and evasion scenario is considered. The evader aims to reach her target and avoid being captured by the pursuer, while the pursuer, without the knowledge of the evader's cost function, intends to capture the evader before she successfully escapes. We formulate this problem as a Model Predictive Control Decoupled Pursuit and Evasion Game (MPC-DPEG) model with incomplete information. We apply Inverse Optimal Control (IOC) technique to estimate the evader's cost function and give a Nash pursuit strategy for MPC-DPEG as the optimal strategy for the pursuer. Numerical examples reveal the improvement of pursuit performance of the Nash pursuit strategy based on the estimated cost function over naive pursuit strategies.**

***Index Terms*—Pursuit-evasion game, Nash equilibrium, Inverse optimal control, Model predictive control**

I. Introduction

Pursuit and evasion scenarios have been under research for decades [1], since such scenarios are widely applied in real practice, especially in the fields of aerospace, robotics and military. It is also a prototypical problem where the game theory is involved.

After the first systematical introduction of differential game in [2], Ho *et.al* [3] propose a finite horizon zero-sum linear quadratic differential game framework to depict the traditional one evader v.s. one pursuer scenario and provide a saddle solution by the variational method. Enlightened by this work, most pursuit and evasion games from then on are tackled in a zero-sum game way, and the scenarios have become more complicated to suit real practice. The concept of target (either static or moving) is studied in [4] – [7], which turns the game from *avoiding being captured* v.s. *capturing* into *reaching before being captured* v.s. *capturing before reaching* as Fig.1 shows.

However, the well-studied finite horizon zero-sum linear quadratic game model with fixed terminal time has two drawbacks:

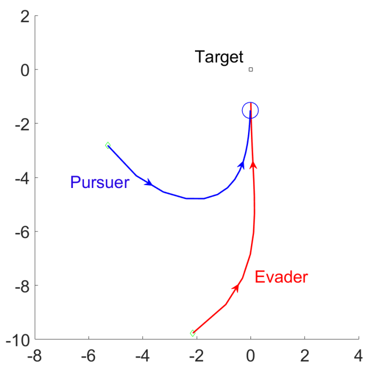


Fig. 1. Demonstration of a scenario where the evader aims to reach her target and the pursuer aims to capture the evader.

1) *the requirement of a fixed terminal time*. In real pursuit and evasion scenarios, the game never ends at a fixed terminal time. Instead, the game ends as soon as the terminal condition is met (e.g. the evader is captured or the evader has escaped), which means the terminal time of the game is dependent of the states of the evader and the pursuer. With the terminal time being indeterminate, methods such as dynamic programming or solving Ricatti equations are not applicable[1]. 2) *the requirement of zero-sum*. In real practice, the property of zero-sum is met only in limited scenarios. It may happen that the evader and the pursuer value differently on the same term of the cost function. For instance, with the introduction of the concept of target, the evader values *reaching the target* over *avoiding being captured*, while the pursuer may only care *capturing*. In these cases, a common cost function that depicts both players' behavior is not reasonable. Instead, the cost functions of the two players have to be distinct, which means the game stands in a nonzero-sum way. Worse still, the players may only know her own model well, but have no knowledge of her opponent's cost function. Without complete information of the game, general methods of deriving the Nash strategy is not feasible.

To the best of the authors' knowledge, only a few research

This work was supported by the National Natural Science Foundation of China (Grant No. 62103276 & Grant No. U2013203).

[a] Department of Automation, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China.
[b] Institute of Medical Robotics, Shanghai Jiao Tong University, Shanghai, China.
[c] Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai 200240, China.
[d] Shanghai Engineering Research Center of Intelligent Control and Management, Shanghai 200240, China.

[1]This is because that real terminal time of the game is not the nominal terminal time where the Riccati Iteration starts. The optimal strategy obtained by solving Riccati equation directly is thus flawed. So we will not demonstrate it in Section V.

consider the case of linear quadratic pursuit and evasion game with indeterminate terminal time. Although Li *et.al* [5] introduce the terminal condition of the game, it is only used to determine the winner of the game, and the terminal time of the game is still fixed. Li *et.al* [11] solve the pursuit and evasion game with indeterminate terminal time by solving high-dimensional two-point boundary value problem. However, the game is formulated in a zero-sum model and the only term in the cost function is the terminal time. The evader intends to maximize the terminal time while the pursuer wants to minimize it.

To tackle nonzero-sum scenarios, Starr *et.al* [12] first generalize the "one evader v.s. one pursuer scenario" in a nonzero-sum way and give a method of deriving Nash equilibrium of the game. Mylvaganam *et.al* [13] introduces ways of obtaining ϵ_α Nash equilibrium strategies for a class of infinite-horizon, nonzero-sum differential games. They [14] further apply nonzero-sum game theory into multi-agent collision avoidance scenarios and derives the Nash strategy. However, the nonzero-sum game with incomplete information, where the opponent's cost function is unknown to the player, is still beyond the capacity of their methods. The Nash strategy can only be obtained after techniques are applied to recover the cost function(s) of the game.

In this paper, a nonzero-sum discrete linear quadratic pursuit and evasion game with incomplete information and indeterminate terminal time is investigated. To tackle the difficulty of indeterminate terminal time, we introduces the terminal indicator functions in the finite horizon cost functions of both players. The value of terminal indicator functions change when the winning condition of either player is met. The introduction of terminal indicator functions resolves the problem of indeterminate terminal time of the game, but brings non-smoothness to the model.

To overcome non-smoothness, we apply Model Predictive Control (MPC) scheme as the solution and reformulate the game as a Model Predictive Control based Decoupled Pursuit and Evasion Game (MPC-DPEG). Several studies [8] - [9] apply MPC in the pursuit and evasion game as tracking algorithm since MPC guarantees closed-loop stability for LQR problems with positive semi-definite weighting matrix Q and any positive definite R [10]. Althougth MPC turns out to be a good controller, it requires the knowledge of the future states of both players. Consequently, the players have to know their opponent's cost function well to apply optimal strategy. For the game with incomplete information, Inverse Optimal Control (IOC) technique is thus applied to recover the unknown game model, and thus the difficulty of incomplete information is overcome.

As a summary, the game investigated in this paper is solved by three steps: 1) *Formulation of MPC-DPEG*. 2) *Recovery of the model by IOC*. 3) *Deriving the Nash strategy*. The condition for the existence and uniqueness of the Nash strategy, which serves as the constraint during the process of IOC, is also given. The performance of the Nash pursuit strategy and the comparison with naive pursuit strategy are shown by numerical examples.

II. Problem formulation

A. Decoupled pursuit-evasion game

We first formulate a general decoupled pursuit evasion game (DPEG) model. Consider that a pursuer, an evader and a static target are in an $\mathbb{R}^2$ space. The evader intends to reach the target, while keeping himself as far as possible from the pursuer. The mission of the pursuer, on the other hand, is to intercept the evader before she could get to her target. Let the state variables $x_t^e, x_t^p, x^g \in \mathbb{R}^2$ be the position of the evader, the pursuer and the target. We denote d_t^{ep} and d_t^{eg} as the distance between the evader and the pursuer and the distance between the evader and the target at time t:

$$d_t^{ep} = \|x_t^e - x_t^p\|, \quad d_t^{eg} \quad = \|x_t^e - x^g\|,$$

where $\|\cdot\|$ denotes the Euclidean norm. We denote $r^e > 0$ as the reaching radius of the evader, and $r^p > 0$ as the interception radius of the pursuer. We define the terminal indicator function $\sigma^e(\cdot), \sigma^p(\cdot) : \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2 \mapsto \{0, 1, \infty\}$ for the evader and the pursuer as follows:

$$\sigma^e(x_t^e, x_t^p, x^g) = \begin{cases} 1, & \text{if } d_t^{eg} > r^e \wedge d_t^{ep} > r^p \\ 0, & \text{if } d_t^{eg} \le r^e \wedge d_t^{ep} > r^p \\ +\infty, & \text{if } d_t^{eg} > r^e \wedge d_t^{ep} \le r^p \end{cases}. \quad (1)$$

$$\sigma^p(x_t^p, x_t^e, x^g) = \begin{cases} 1, & \text{if } d_t^{ep} > r^p \wedge d_t^{eg} > r^e \\ 0, & \text{if } d_t^{ep} \le r^p \wedge d_t^{eg} > r^e \\ +\infty, & \text{if } d_t^{ep} > r^p \wedge d_t^{eg} \le r^e \end{cases}. \quad (2)$$

Note that at each time instant, $\sigma^e(x_t^e, x_t^p, x^g)$ and $\sigma^p(x_t^p, x_t^e, x^g)$ indicate whether the game is terminated. If the game comes to an end, the value of the cost function of the player who loses the game will be set to $+\infty$. Then the pursuit and evasion game is formulated as follows:
For the evader:

$$\begin{aligned} \min_{\{x_t^e\},\{u_t^e\}} J_e = \sum_{t=t_0}^{T_f-1} & \sigma^e(x_{t+1}^e, x_{t+1}^p, x^g) \cdot \left[w_1^e \|x_{t+1}^e - x^g\|^2 \right. \\ & \left. - w_2^e \|x_{t+1}^e - x_{t+1}^p\|^2 + \|u_t^e\|^2 \right], \\ \text{s.t. } x_{t+1}^e =& x_t^e + u_t^e, \\ x_{t_0}^e =& x_0^e, \end{aligned}$$

and for the pursuer:

$$\begin{aligned} \min_{\{x_t^p\},\{u_t^p\}} J_p = \sum_{t=t_0}^{T_f-1} & \sigma^p(x_{t+1}^p, x_{t+1}^e, x^g) \\ & \cdot \left[w^p \|x_{t+1}^p - x_{t+1}^e\|^2 + \|u_t^p\|^2 \right], \\ \text{s.t. } x_{t+1}^e =& x_t^p + u_t^p, \\ x_{t_0}^p =& x_0^p, \end{aligned}$$

where $w_1^e, w_2^e, w^p > 0$ are normalized weighting parameters for corresponding terms. $x_{t_0}^e, x_{t_0}^p \in \mathbb{R}^2$ denotes the initial position of the evader and the pursuer. T_f denotes the upper limit of the game time horizon. In real practice, T_f is set to be sufficiently large to make sure that game will ends before T_f.

DPEG model is non-smooth due to the existence of the terminal indicator function $\sigma^e(\cdot)$ and $\sigma^p(\cdot)$ in both players' cost functions. It brings two problems: 1) the existence and uniqueness condition of the Nash equilibrium is difficult to analyze. 2) It is difficult to solve the closed-loop Nash equilibrium.

B. Formulation of MPC Decoupled Pursuit and Evasion Game

To tackle these problems, we apply MPC scheme to approximate the original model and formulate an MPC based Decoupled Pursuit and Evasion Game (MPC-DPEG) model: at any time instant k when the game is not ended, i.e. $s(x_k^e, x_k^p, x^g) = 1$, the cost function and dynamics for the evader is given as

$$\begin{aligned} \min_{\{x^e_{t|k}\},\{u^e_{t|k}\}} \hat{J}_k^e = & \sum_{t=k}^{k+N-1} \Big[w_1^e \| x^e_{t+1|k} - x^g \|^2 \\ & - w_2^e \| x^e_{t+1|k} - x^p_{t+1|k} \|^2 + \| u^e_{t|k} \|^2 \Big], \qquad (3) \\ \text{s.t. } x^e_{t+1|k} = & x^e_{t|k} + u^e_{t|k}, \\ x^e_{k|k} = & x^e_k. \end{aligned}$$

and the cost function and dynamics for the pursuer is given as

$$\begin{aligned} \min_{\{x^p_{t|k}\},\{u^p_{t|k}\}} \hat{J}_k^p = & \sum_{t=k}^{k+N-1} \Big[w_p \| x^p_{t+1|k} - x^e_{t+1|k} \|^2 + \| u^p_{t|k} \|^2 \Big], \\ s.t. \; x^p_{t+1|k} = & x^p_{t|k} + u^p_{t|k}, \\ x^p_{k|k} = & x^p_k. \end{aligned} \qquad (4)$$

where $x_k^e, x_k^p \in \mathbb{R}^2$ denotes the state of the evader and the pursuer at time k. In the model defined above, since $w^p > 0$, the cost function of the pursuer is a strictly convex function of u_t^p [15]. This guarantees both the existence and the uniqueness of the solution of u_t^p, given any state of the evader. In order that the solution of u_t^e also exists and is unique, given any state of the pursuer, if

$$0 < w_2^e \le w_1^e, \qquad (5)$$

which means the coefficient matrix of the quadratic term of the state variable x_t^e is semi positive-definite if the model is written in the normal form. Note that although solutions may still exist when $w_1^e < w_2^e$, we do not consider these cases, since *reaching the target* has a higher priority over *avoiding being captured* for the evader.

We further define the Nash equilibrium for MPC-DPEG as follows:

Definition 1. *The pair of control input sequences* $(\{u^{e*}_{k|k}\}, \{u^{p*}_{k|k}\})$ *is a Nash equilibrium for MPC-DPEG if*

$$\forall k, \; \begin{cases} \hat{J}_k^e(\{u^{e*}_{k|k}, u^{e*}_{k+1|k}, \cdots\}, \{u^{p*}_{k|k}, u^{p*}_{k+1|k}, \cdots\}) \\ \le \hat{J}_k^e(\{u^{e}_{k|k}, u^{e}_{k+1|k}, \cdots\}, \{u^{p*}_{k|k}, u^{p*}_{k+1|k}, \cdots\}) \\ \hat{J}_k^p(\{u^{p*}_{k|k}, u^{p*}_{k+1|k}, \cdots\}, \{u^{e*}_{k|k}, u^{e*}_{k+1|k}, \cdots\}) \\ \le \hat{J}_k^p(\{u^{p}_{k|k}, u^{p}_{k+1|k}, \cdots\}, \{u^{e*}_{k|k}, u^{e*}_{k+1|k}, \cdots\}) \end{cases}. \qquad (6)$$

Note that $(u^{e*}_{k|k}, u^{p*}_{k|k})$ is the first element of the Nash equilibrium of $(\hat{J}_k^e, \hat{J}_k^p)$ for all k satisfying $s(x_k^e, x_k^p, x^g) = 1$, and $(u^{e*}_{k|k}, u^{p*}_{k|k})$ is the control input pair that is actually applied. The Nash equilibrium for MPC-DPEG inherits the property of classic Nash equilibrium that no player can improve her performance by unilaterally deviating from the Nash equilibrium. If both players apply MPC strategies, it is optimal for the player as long as her opponent sticks to Nash strategy.

With such definition, we are able to solve MPC-DPEG with indeterminate terminal time by iteratively solve for open-loop Nash solution until the value of the terminal indicator functions (1) and (2) change. Note that The first elements of the open-loop Nash solution will be the new state of both players and they will plan based on "one step update" until the game ends. Moreover, the solution is intrinsically a closed-loop control strategy.

C. IOC Technique

In MPC-DPEG defined in (6), both players' plan based on the prediction of each other's future states, which requires the knowledge of the opponent's cost function. In order to solve for the Nash strategy for the game with incomplete information, where everything else except the evader's cost function is known to the pursuer, she must find methods to recover the evader's cost function. The Inverse Optimal Control technique is applied for the recovery.

IOC technique claims that the player behaves optimally, following the instruction of the unknown cost function. It does "inverse engineering" to estimate the cost function based on the knowledge of the optimal trajectories. For pursuit and evasion game models, the evader always follow an optimal trajectory according to the unknown cost function. IOC technique enables the pursuer to estimate the evader's cost function, and then the Nash pursuit strategy can thus be derived.

III. Nash solution of MPC-DPEG

For MPC-DPEG defined in (3), (4) we are able to derive Nash strategy analytically as follows:

Proposition 1. *For MPC-DPEG, the open-loop Nash strategy satisfies the linear matrix equation (LME)*

$$\mathcal{F}(w_1^e, w_2^e, w^p) \cdot S_k = C_k(w_1^e, x_k^e, x_k^p, x^g), \forall k = t_0, \cdots, T_f - 1 \qquad (7)$$

where

$$\mathcal{F}(w_1^e, w_2^e, w^p) = \left[\begin{array}{c:c} \mathcal{F}_e(w_1^e, w_2^e) & \begin{matrix} w_2^e I & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{matrix} \\ \hdashline \begin{matrix} -w^p I & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{matrix} & \mathcal{F}_p(w^p) \end{array} \right],$$

$$\mathcal{F}_e(w_1^e, w_2^e) = \left[\begin{array}{c:c}(w_1^e - w_2^e)I & \Gamma_1 \\ \hdashline I & \Gamma_2\end{array}\right],\ \mathcal{F}_p(w^p) = \left[\begin{array}{c:c}w^p I & \Gamma_1 \\ \hdashline I & \Gamma_2\end{array}\right],$$

$$\Gamma_1 = \begin{bmatrix} -I & I & & \\ & -I & \ddots & \\ & & \ddots & I \\ & & & -I \end{bmatrix},$$

$$\Gamma_2 = \begin{bmatrix} I_2 & & & \\ I_2 & I_2 & & \\ \vdots & \vdots & \ddots & \\ I_2 & I_2 & \cdots & I_2 \end{bmatrix},$$

$$S_k = \begin{bmatrix} x_{k+1|k}^{e*} \\ \vdots \\ x_{k+N|k}^{e*} \\ \hdashline \lambda_{k|k}^{e*} \\ \vdots \\ \lambda_{k+N-1|k}^{e*} \\ \hdashline x_{k+1|k}^{p*} \\ \vdots \\ x_{k+N|k}^{p*} \\ \hdashline \lambda_{k|k}^{p*} \\ \vdots \\ \lambda_{k+N-1|k}^{p*} \end{bmatrix},\ C_k(w_1^e, x_k^e, x_k^p, x^g) = \begin{bmatrix} w_1^e \cdot x^g \\ \vdots \\ w_1^e \cdot x^g \\ \hdashline x_k^e \\ \vdots \\ x_k^e \\ \hdashline \mathbf{0} \\ \hdashline x_k^p \\ \vdots \\ x_k^p \end{bmatrix},$$

where $\lambda_{t|k}^e, \lambda_{t|k}^p \in \mathbb{R}^2, t = k, \cdots, k+N-1$ *are Lagrange multipliers.* C_k *is the vector concerning states of both players and the evader's goal* x^g*. Moreover, the Nash strategy exists and is unique if*

$$\mathcal{F}(w_1^e, w_2^e, w^p) \text{ is invertible.} \tag{8}$$

Proof. Since $w^p > 0$, the cost function of the pursuer (4) is convex, which means that the optimal trajectory of the pursuer always exists and is unique, no matter what trajectory the evader chooses. Given that $w_1^p > w_2^p > 0$, the cost function of the evader (3) is also convex. To solve for the optimal trajectories, we construct Lagrangian for both players:

$$\mathcal{L}_{e,k} = J_{e,k} - \sum_{t=k}^{k+N-1} {\lambda_{t|k}^e}^T (x_{t+1|k}^e - x_t^e - u_{t|k}^e),$$
$$\mathcal{L}_{p,k} = J_{p,k} - \sum_{t=k}^{k+N-1} {\lambda_{t|k}^p}^T (x_{t+1|k}^p - x_t^p - u_{t|k}^p).$$

By letting $\frac{\partial \mathcal{L}_{i,k}}{\partial u_{t|k}^i} = 0$, we obtain $\forall t = k, \cdots, k+N-1$

$$u_{t|k}^{i*} = -\lambda_{t|k}^{i*},\ i = e, p. \tag{9}$$

substituting into the dynamics, we have

$$x_{t|k}^i = x_k^i - \sum_{n=0}^{t-k-1} \lambda_{k+n|k}^i,\ i = e, p. \tag{10}$$

By letting $\frac{\partial \mathcal{L}_{e,k}}{\partial x_{t|k}^e} = \frac{\partial \mathcal{L}_{p,k}}{\partial x_{t|k}^p} = 0$, we obtain $\forall t = k+1, \cdots, k+N$

$$\begin{aligned} w_1^e(x_{t|k}^e - x^g) - w_2^e(x_{t|k}^e - x_{t|k}^p) + \lambda_{t|k}^e - \lambda_{t-1|k}^e = 0. \\ w_p(x_{t|k}^p - x_{t|k}^e) + \lambda_{t|k}^p - \lambda_{t-1|k}^p = 0 \end{aligned}, \tag{11}$$

where $\lambda_{k+N|k}^e = \lambda_{k+N|k}^p = \mathbf{0}$. Combining equations (9)-(11), we are able to construct the linear matrix equation (7).

Given that (5) is satisfied, the solution of the two players exists and is unique as long as the opponent's states are given. However, it does not trivially mean that the combined LME (7) is also solvable to get the optimal control inputs of the evader and the pursuer simultaneously. S_k in (7) exists if

$$C_k(w_1^e, x_k^e, x_k^p, x^g) \in range(\mathcal{F}(w_1^e, w_2^e, w^p)),$$

which, combining (5), contribute to the sufficient condition of the existence and the uniqueness of the solution. □

IV. Estimation of the evader's cost function

In order to recover the model of MPC-DPEG. IOC technique is applied to estimate the unknown weighting parameters w_1^e and w_2^e. IOC utilize previous trajectory record of the evader as the reference trajectories. Then w_1^e and w_2^e is then estimated by solving the following optimization problem:

$$\begin{aligned} \min_{\{\hat{x}_t^{e,i}\}, w_1^e, w_2^e} \quad & \mathcal{L}oss = \sum_{i=1}^{M} \sum_{t=t_0+1}^{T_f} \|\hat{x}_t^{e,i} - x_t^{e,i}\|^2, \\ \text{s.t.} \quad & 0 \le w_2^e \le w_1^e, \\ & \mathcal{F}(w_1^e, w_2^e, w^p) \cdot S_k = C_k(w_1^e, x_k^e, x_k^p, x^g), \\ & \forall k = t_0, \cdots, T_f - 1, \end{aligned}$$

where $M \in \mathbb{N}^*$ is the number of reference trajectories (data). Note that $\{\hat{x}_t^{e,i}\}$ is obtained by solving for S_k iteratively. We loosen the restriction of w_1^e, w_2^e of positive numbers to non-negative numbers, in order to have a closed-form constraint for optimization. This does not interfere the result. With the estimated weighting parameters for the evader's cost function, the pursuer is then able to apply a Nash strategy for DPEG.

V. Numerical Examples

In this section, numerical examples of solving MPC-DPEG with the evader's cost function unknown and indeterminate terminal time are given. The process is divided into two steps: 1) recovery of the evader's cost function, 2) solving for the Nash pursuit strategy.

A. Recovery of the evader's cost function

We set $w_1^e = 2 \times 10^{-2}, w_2^e = 1 \times 10^{-2}, w^p = 0.1, r^e = r^p = 0.3$, $A = B = I_2$. The target is located at the origin (i.e. $x^g = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$) and the planning horizon N for MPC-DPEG is set to be 10, without loss of generality. We set the initial distance of the evader and the target $\|x_{e0} - x^g\| = 10$ and the initial distance of the pursuer and the target $\|x_{p0} - x^g\| = 6$ at t_0, and thus randomly generate 100 scenarios. Fig. 2 illustrates the random starting position of the evader and the pursuer. To generate reference trajectories, the pursuer's

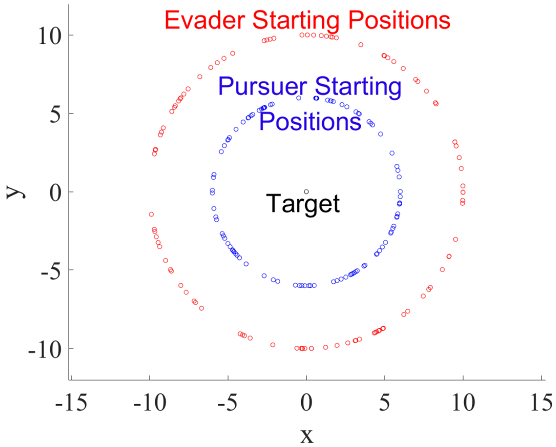


Fig. 2. 100 pair of random starting positions of evader and pursuer.

TABLE I
IOC RESULTS

	real value	estimated value	No. of Iterations
w_1^e	2×10^{-2}	2.0002×10^{-2}	36
w_2^e	1×10^{-2}	0.9996×10^{-2}	

trajectories are generated by applying a naive pursuit strategy, that is, at each time instant t, if the game is not ended, the pursuer goes straight towards the evader with the control input $u_t^p = \alpha(x_t^e - x_t^p)$, where $0 < \alpha < 1$ is the approaching rate. Without loss of generality, we set the approaching rate $\alpha = \frac{1}{15}$. The evader's reference trajectories are generated by solving the MPC optimal control problem (3).

With generated data, we apply IOC to estimate the value of w_1^e and w_2^e. We set the initial value $w_1^{e(0)} = 5 \times 10^{-2}, w_1^{p(0)} = 5 \times 10^{-3}$. IOC results are shown in Table I. The optimization finishes within 36 iterations and the estimated values are quite close to the real values, which reveals the accuracy and feasibility of IOC techniques.

B. Nash Strategy for DPEG

With the estimated weighting parameters, we solve the pursuit and evasion scenarios with both naive pursuit strategy and the Nash pursuit strategy.

Although the starting positions are randomly generated, the scenarios can be classified into three classes, according to the angle between the line connecting the target and the evader and the line connecting the target and the pursuer $\langle x_0^e - x_g, x_0^p - x_g\rangle$:
Class A: $0 \leq \langle x_0^e - x_g, x_0^p - x_g\rangle < \frac{\pi}{3}$ (Fig. 3a and Fig. 3b).
Class B: $\frac{\pi}{3} \leq \langle x_0^e - x_g, x_0^p - x_g\rangle < \frac{2\pi}{3}$ (Fig. 3c and Fig. 3d).
Class C: $\frac{2\pi}{3} \leq \langle x_0^e - x_g, x_0^p - x_g\rangle \leq \pi$ (Fig. 3e and Fig. 3f). In order to give a straight view of the performance of the Nash pursuit strategy and naive pursuit strategy, for each class, a prototypical scenario is illustrated for comparison.

For the scenario in Class A, the pursuer succeeds in capturing the evader by applying either strategy. However, the place

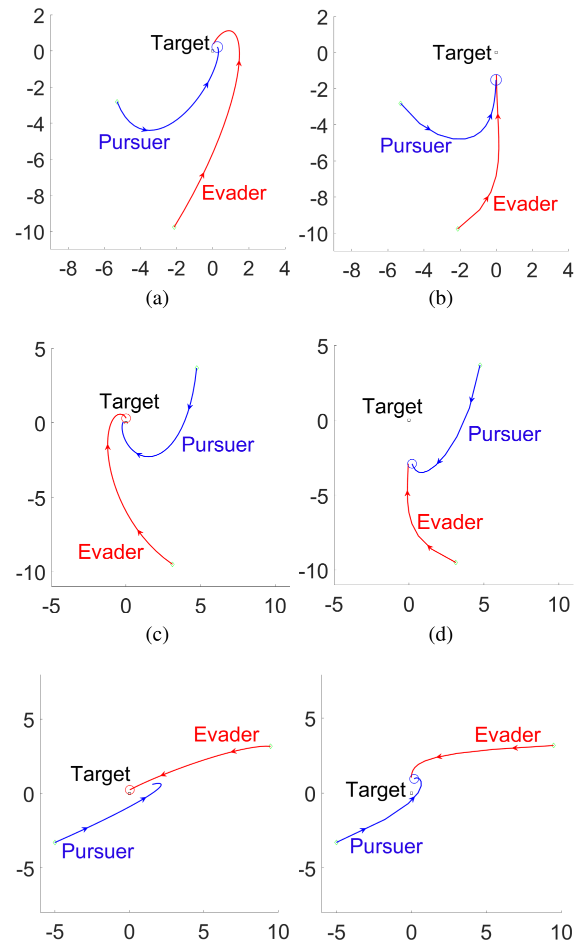


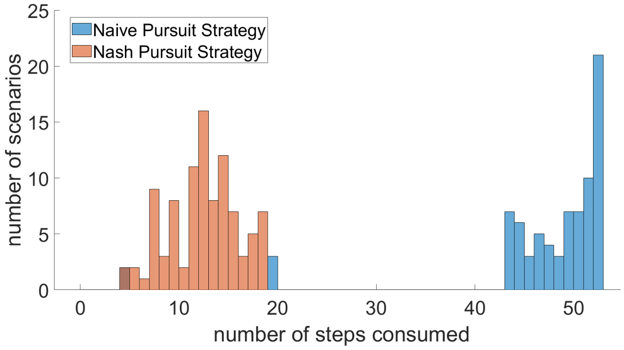
Fig. 3. Demonstration of pursuit performance of three scenarios when the pursuer applies different strategies. 3a,3c,3e: Naive pursuit strategy; 3b,3d,3f: the Nash Pursuit strategy.

where the pursuer who applies naive pursuit strategy capture the evader is quite close to the target. It is quite risky as if r^e takes a larger value, the evader would win the game.

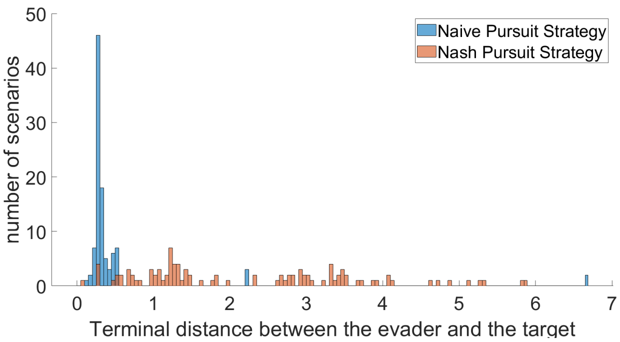
For the scenario in Class B, the pursuer fails to capture the evader in time when applying naive pursuit strategy (Fig. 3c). In contrast, the Nash pursuit strategy guarantee the success of capture.

Scenarios in Class C are the worst for the pursuer, since the initial distance between the pursuer and the evader is the longest of all classes. Pursuer who applies naive strategy suffers from low success rate of capture in similar cases (Fig. 3e). However, the Nash pursuit strategy still enables the pursuer to capture the evader in time (Fig. 3f), even under such unfavorable situation.

Stats of all the winning cases in the randomly generated 100 scenarios are demonstrated in Table II, providing direct evidence of improvement of pursuit performance of the Nash pursuit strategy over naive strategy. The pursuer who applies the Nash pursuit strategy wins 96 of 100 cases, while the pursuer applies naive pursuit strategy only wins 78 of 100



(a)



(b)

Fig. 4. 4a: Stats of time consumption to capture the evader, the smaller the better; 4b: Stats of distance between the evader and the target at terminal time, the larger the better.

TABLE II
NASH PURSUIT STRATEGY V.S. NAIVE PURSUIT STRATEGY

	No. of capture	Avg. capture time	Avg. final dist $\|x_e - x^g\|$
Naive	78	46.44	0.583
Nash	**96**	**12**	**2.288**

cases. Moreover, the Nash pursuit strategy performs better in average capture time, consuming only one fourth of the time that naive pursuit strategy spends. Histogram (Fig. 4a) demonstrate the distribution of time consumption of the two strategies, revealing that the pursuer spends less time capturing the evader. On the other hand, the Nash pursuit strategy also outperforms in the distance between the target and the place where the evader is captured, which is also reflected in the histogram (Fig. 4b).

VI. CONCLUSIONS AND FUTURE WORK

A. Conclusions

We formulate a nonzero-sum linear quadratic pursuit and evasion game with incomplete information (the opponent's cost function is unknown). The cost functions of game is also non-smooth due to the existence of terminal indicator functions, which makes the problem difficult to solve. To address the difficulty, we apply MPC to approximate the original game model. To address the problem of incomplete information, IOC technique is applied to recover the cost function of the evader, so that the pursuer has a clear access to the evader's future states. This makes it possible for the pursuer to solve for the Nash pursuit strategy. Numerical examples indicates that the recovery of the evader's cost function by IOC is sufficiently accurate, and simulation results reveal the improvement of pursuit performance when applying the Nash pursuit strategy over naive pursuit strategy.

B. Future work

Currently, the Nash solution of one pursuer v.s. one evader MPC-DPEG in $\mathbb{R}^2$ is given. In future research, more complicated scenarios will be studied. The game may be further conducted in a constrained environment, where both the evader and the pursuer are only allowed to move in restricted areas (e.g. due to the presence of obstacles). Moreover, the number of pursuers and evaders may be increased so that the game will become the team of pursuers v.s. the team of evaders. The study of multiple number of players or even varying number of players will broaden the field of application of MPC-DPEG.

REFERENCES

[1] I. E. Weintraub, M. Pachter, and E. Garcia, "An Introduction to Pursuit-evasion Differential Games," in Proceedings of the American Control Conference, 2020, vol. 2020-July, pp. 1049–1066.

[2] R. Isaacs, Differential Games. New York: Wiley, 1965.

[3] Y. C. Ho, A. E. Bryson, and S. Baron, "Differential Games and Optimal Pursuit-Evasion Strategies," IEEE Transactions on Automatic Control, vol. 10, no. 4, pp. 385–389, 1965.

[4] E. Garcia, D. W. Casbeer, and M. Pachter, "Design and analysis of state-feedback optimal strategies for the differential game of active defense," IEEE Transactions on Automatic Control, vol. 64, no. 2, pp. 553–568, 2019.

[5] D. Li and J. B. Cruz, "Defending an asset: a linear quadratic game approach," IEEE Transactions on Aerospace and Electronic Systems, vol. 47, no. 2, pp. 1026–1044, 2011.

[6] E. Garcia, D. W. Casbeer, M. Pachter, J. W. Curtis and E. Doucette, "A Two-team Linear Quadratic Differential Game of Defending a Target," 2020 American Control Conference (ACC), 2020, pp. 1665-1670.

[7] J. Z. Ben-Asher, S. Levinson, J. Shinar, and H. Weiss, "Trajectory shaping in linear-quadratic pursuit-evasion games," Journal of Guidance, Control, and Dynamics, vol. 27, no. 6, pp. 1102–1105, 2004.

[8] J. M. Eklund, J. Sprinkle, and S. S. Sastry, "Switched and symmetric pursuit/evasion games using online model predictive control with application to autonomous aircraft," IEEE Trans. Control Syst. Technol., vol. 20, no. 3, pp. 604–620, 2012.

[9] M. Sani, B. Robu, and A. Hably, "Pursuit-evasion game for nonholonomic mobile robots with obstacle avoidance using NMPC," 2020 28th Mediterr. Conf. Control Autom. MED 2020, pp. 978–983, 2020.

[10] M. Morari and J. H. Lee, "Model predictive control: Past, present and future," Comput. Chem. Eng., vol. 23, no. 4–5, pp. 667–682, 1999.

[11] Z. yu Li, H. Zhu, Z. Yang, and Y. zhong Luo, "A dimension-reduction solution of free-time differential games for spacecraft pursuit-evasion," Acta Astronautica, vol. 163, no. December 2018, pp. 201–210, 2019.

[12] A. W. Starr and Y. C. Ho, "Nonzero-sum differential games," J. Optim. Theory Appl., vol. 3, no. 3, pp. 184–206, 1969.

[13] T. Mylvaganam, M. Sassano, and A. Astolfi, "Constructive ε-Nash Equilibria for Nonzero-Sum Differential Games," IEEE Trans. Automat. Contr., vol. 60, no. 4, pp. 950–965, 2015.

[14] T. Mylvaganam, M. Sassano, and A. Astolfi, "A Differential Game Approach to Multi-agent Collision Avoidance," IEEE Trans. Automat. Contr., vol. 62, no. 8, pp. 4229–4235, 2017.

[15] T. Basar and G. J. Olsder, Dynamic Noncooperative Game Theory, 2nd Edition. SIAM, 1998.