Defending an Asset with Partial Information and Selected Observations: A Differential Game Framework

Yunhan Huang, Juntao Chen, and Quanyan Zhu

Abstract— This paper considers a linear-quadratic-Gaussian asset defending differential game (DADG) where the attacker and the defender do not know each other's state information. However, they both know the trajectory of a moving asset. Both players can choose to observe the other player's state information by paying a cost. The defender and the attacker have to craft both control strategies and observation strategies. We obtain a closed-form feedback solution that characterizes the Nash control strategies. We show that the trajectory of the asset does not affect both players' observation choices. Moreover, we show that we can decouple the observation choices of the defender and the attacker. One can obtain the Nash observation strategies by solving two independent optimization problems. A set of necessary conditions is developed to characterize the optimal observation instances. Based on the necessary conditions, we propose an effective algorithm to compute the optimal observation instances numerically. We also present a case study to demonstrate the effectiveness of the optimal observation instances.

I. INTRODUCTION

With recent advances in Autonomous Vehicles (AV) technologies, AV application scenarios emerge in modern military operations such as surveillance, persistent area denial [1], pursuit-evasion [2], [3], and defending an asset (DA) [4], [5]. DA scenarios describe a setting where attackers attempt to intercept the asset and defenders strive to defend the asset. DA scenarios pose challenging control design problems for AVs because AVs deployed often confront intelligent rivals with mobility and strategic decision-making. Differential game theory offers the right set of theoretical underpinnings to investigate DA scenarios and develop optimal strategies for each player. Hence, several works have addressed different DA problems formulated as differential games [4]–[9]. The formulations often are referred to as Defending an Asset Differential Games (DADG). Among the various DADG models adopted, the linear-quadratic differential game formulation is favored due to its analytical friendliness [4], [7]. A common assumption taken for granted in previous studies of DADG is that state information is freely available at any time to both the attacker and the defender. However, state information, especially information regarding one's opponent, is not accessible and usually is expensive to acquire. For example, in naval warfare, the detection of aircraft carriers is a challenging task considering the vastness of the ocean and the search for adversary submarines due to their stealthiness.

The detection of such military units comes with monetary expenses, risks of exposing oneself, and loss of surveillance aircraft.

In this work, we study DADG with partial information where the attacker and the defender only have access to their state information. The attacker does not know the state information of the defender unless the attacker chooses to observe the defender's information by paying a cost. Similarly, the defender does not know the attacker's state unless the defender chooses to observe it by paying a cost. We assume that both the defender and the attacker know the trajectory of the asset. The DADG is a dynamic game with a pre-specified duration. Since the two players are constantly moving, the earlier information may deteriorate over time, and a new observation needs to be made. Thus, both players must decide when to observe and how many times to observe within the duration. At the same time, control strategies need to be developed based on the observed information.

Several papers have investigated the joint design of observation strategies and control strategies with costly observations [10]–[16]. In the 70s, Cooper studied a discrete-time optimal control problem where the controller decides at each step whether to observe a noisy observation or not [10]. Olsder later extended this study into a two-player discrete-time dynamic game setting where each player chooses when to observe, and a solution is obtained for a two-stage dynamic game [12]. More recently, [11] looked into controlled observation for continuous-time Markov decision processes, where the authors have studied its applications in queueing systems and inventory management. [16] studies a discrete-time dynamic game with controlled yet noisy observations where one player acts as a jammer that intercepts the observation of the other player. In [14], [15], and [17], Maity et al. extend the problem formulated in [12] to a linear quadratic continuous-time setting, where each player has no state information at all unless they choose to observe.

Our work is different from [14], [15], [17] in three aspects: First, in these works, players choose whether to observe the full state of the system. In our work, the defender or the attacker knows a part of the whole state (their own state) and chooses to observe the other part of the state (their opponent's state); Second, in these works, two observation scenarios are considered. 1. Each player chooses to make their own private observation. 2. Observation is available to both players only when both players choose to observe simultaneously. However, in our work, one player's observation will expose oneself to his/her opponent. Third, we fully characterize the Nash observation strategy and propose

Y. Huang and Q. Zhu are with the Department of Electrical and Computer Engineering, New York University, 370 Jay St., Brooklyn, NY 11201. {yh.huang, qz494}@nyu.edu.

J. Chen is with the Department of Computer and Information Sciences, Fordham University, New York, NY 10023. jchen504@fordham.edu

computational methods to compute these strategies.

In general, our work focuses on DADG, which has a different cost structure and system dynamics from previous works due to the nature of "defending an asset" game. In our work, both the attacker and the defender know their own information, which gives each player partial information even when they choose not to observe. They obtain complete information when they choose to observe. Both players do not share their information, which causes asymmetry of information. Third, we fully characterize the control strategies and develop a practical algorithm that calculates the observation instances numerically.

The contributions of this paper are summarized as follows. First, we abandon the common yet unrealistic assumption that information is freely available all the time in previous DADG works. A linear-quadratic-Gaussian DADG framework with controlled partial information is proposed. Second, we fully characterize the Nash control strategies and develop a set of necessary conditions that characterize the optimal observation strategy. We show the separation principle, in which the observation choices only affect the state estimate in the control strategies. Analytical results show that the observation choices are independent of the trajectory of the asset. We further show that the attacker's observation decision and the defender's observation decision are decoupled. Hence, either the attacker or the defender can make observation choices without anticipating each other's choices. As a result, we can obtain the Nash observation strategies by solving two independent optimization problems. Third, we develop an effective algorithm that can numerically compute the optimal observation instances. We demonstrate the effectiveness of the optimal observation instances by comparing them with the periodic observation instances.

The rest of the paper is organized as follows. In Sec. II, we formulate the linear-quadratic-Gaussian DADG with controlled information. Sec. III gives the theoretical results regarding the Nash control strategies and the Nash observation strategies. In Sec. IV, we conduct a case study to demonstrate the effectiveness of the optimal observation instances.

A. Notation

For a vector x or a matrix M, x' and M' represents the transpose of the vector x and the matrix M respectively. An $n \times n$ identity matrix is denoted by $\mathbb{I}_n$. $\| \cdot \|_2$ is the L-2 norm. For a vector with proper dimension, the norm $\| \cdot \|_Q$ is defined as $\|x\|_Q = x'Qx$. The set of all real numbers is denoted by $\mathbb{R}$, and $\mathbb{N}$ denotes the set of all natural numbers. The trace operator is denoted by $\mathrm{Tr}(\cdot)$. The Kronecker product is represented by $\otimes$.

II. DEFENDING AN ASSET WITH CONTROLLED INFORMATION

In this section, we formulate a DADG with controlled observation. The dynamics of each player is described by the following linear systems

$$dx_a = A_a x_a(t)dt + \tilde{B}_a u_a(t)dt + C_a dw_a(t), \quad x_a(0) = x_{a0}$$
$$dx_d = A_d x_d(t)dt + \tilde{B}_d u_d(t)dt + C_d dw_d(t), \quad x_d(0) = x_{d0} \tag{1}$$

where $x_a \in \mathbb{R}^n, x_d \in \mathbb{R}^n$ are states of the attacker and the defender; $u_a \in U_a, u_d \in U_d$ are controls inputs the corresponding players; w_a and w_d are independent standard Wiener processes. $A_a, A_d, \tilde{B}_a, \tilde{B}_d, C_a$, and C_d are real matrices with proper dimensions. The time index is denoted by t and the consider a finite-time horizon $[0, t_f]$. Let $x_s(t)$ be the location of the asset at time t and the trajectory of the asset is given and known to both the attacker and the defender. In this paper, we consider the cases of a stationary asset and an asset with an arbitrary trajectory. We assume that there is an auxiliary linear system that captures the trajectory $x_s(\cdot)$ of the asset $\dot{x}_s = A_s x_s, \; x_s(0) = x_{s0}$. This assumption is introduced for analysis purpose and is not necessary, as we will show later. From a systematic point of view, we can formulate an aggregate system as

$$dx(t) = Ax(t)dt + B_a u_a dt + B_d u_d dt + C dw(t), \text{ with } x(0) = x_0, \tag{2}$$

where $x = [x_a' \; x_d' \; x_s']'$, $B_a = [\tilde{B}_a' \; \mathbf{0} \; \mathbf{0}]'$, $B_d = [\mathbf{0} \; \tilde{B}_d' \; \mathbf{0}]'$, $w = [w_a' \; w_d' \; \mathbf{0}]'$,

$$A = \begin{bmatrix} A_a & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & A_d & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & A_s \end{bmatrix}, \quad \text{and } C = \begin{bmatrix} C_a & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & C_d & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

In this paper, we consider a situation where the defender and the attacker can select a set of time instances to observe one's opponent state. The information structure of the attacker and the defender is summarized as follows: 1. Both the defender and the attacker know the trajectory of the asset. 2. The defender and the attacker know their own state, but they don't know each other's state. 3. Each player can choose to observe the other player's state by paying a cost.

Let $\mathcal{T}_a = \{t_{1,a}, t_{2,a}, \cdots, t_{N_a,a}\}$ be the set of time instances when the attacker choose to observe. Let $\mathcal{T}_d = \{t_{1,d}, t_{2,d}, \cdots, t_{N_d,d}\}$ be that of the defender. Here, N_a and N_d are the number of observations made by the attacker and the defender respectively within time horizon $[0, t_f]$. Let $y_a(t)$ and $y_d(t)$ be the observations of the attacker and the defender respectively. We can write the above description of the information structure as

$$y_a(t) = \begin{bmatrix} \mathbb{I}_n & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbb{I}_n \end{bmatrix} x(t), \quad \text{for } t \notin \mathcal{T}_a$$

$$y_d(t) = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbb{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbb{I}_n \end{bmatrix} x(t), \quad \text{for } t \notin \mathcal{T}_d,$$

$$y_a(t) = x(t) \quad \text{for } t \in \mathcal{T}_a, \quad y_d(t) = x(t), \quad \text{for } t \in \mathcal{T}_d.$$

Let $\mathcal{I}_a(t)$ be the information the attacker has at time t and $\mathcal{I}_d(t)$ be that of the defender. Given $\mathcal{T}_a, \mathcal{T}_d$, We have

$$\mathcal{I}_a(t) = \{x_0, y_a(\tau), \tau \leq t\}, \; \mathcal{I}_d(t) = \{x_0, y_d(\tau), \tau \leq t\}. \tag{3}$$

Each player considers stationary feedback strategies γ_a and γ_d such that $u_a(t) = \gamma_a(\mathcal{I}_a(t))$ and $u_d(t) = \gamma_d(\mathcal{I}_d(t))$. We consider the objective function of the following form

$$J(\gamma_a, \gamma_d; x_0)$$
$$= \mathbb{E}\left[\int_0^{t_f} \left(u_a(t)'u_a(t) - u_d(t)'u_d(t) + \omega_a^I \|x_a(t) - x_s(t)\|_2^2 \right.\right.$$
$$\left. - \omega_d^I \|x_d(t) - x_a(t)\|_2^2 \right) dt + \omega_a \|x_a(t_f) - x_s(t_f)\|_2^2$$
$$\left. - \omega_d \|x_d(t_f) - x_a(t_f)\|_2^2 + ON_a - ON_d \,\middle|\, x(0) = x_0 \right],$$
$$(4)$$

where ω_a^I, ω_d^I, ω_a, and ω_d are the weighting coefficients that captures the trade-off between the attacker-to-defender distance and the attacker-to-asset distance. Here, the optimization variables are $u_a(u_b)$ and $\mathcal{T}_a(\mathcal{T}_b)$. The effect of $\mathcal{T}_a$ and $\mathcal{T}_b$ are not explicitly written in 4. Indeed, the choices of $\mathcal{T}_a$ and $\mathcal{T}_d$ affect $\mathcal{I}_a$ and $\mathcal{I}_d$ and hence influence the controls u_a and u_d through γ_a and γ_d. The superscript I indicates that the weight is for the intermediate cost rather than the terminal cost. In some DADG papers, w_a^I and w_d^I are set to be zero and hence attention is paid to the terminal state alone [5]. The scalar $O \geq 0$ is the cost of making observations. The attacker tries to minimize its distance to the asset while trying to avoid being intercepted by the defender. Hence, the attacker aims to minimize the objective function and the defender, however, aims to maximize it. Even though we consider the state of all players lying in the same space $\mathbb{R}^n$, the results in this paper can be easily extended to a general setting. It is tacitly assumed that the system characteristics are know to both players.

This formulation gives us a differential game with asymmetric yet controlled information. The problem formulation brings up a series of questions: When does the defender need to observe the attacker's state? For the defender, is it worth paying a cost to observe the attacker's state while the defender knows that the attacker is tracking the asset and the location of the asset is known to the defender? Are the optimal observation instances dependent on the trajectory of the asset? In the next section, we develop our main results that address these questions.

III. THEORETICAL RESULTS

In this section, we develop our main theoretical results of this paper. A close look at (4) gives the following form

$$J(\gamma_a, \gamma_d; x_0)$$
$$= \mathbb{E}\left[\int_0^{t_f} \left(u_a(t)'u_a(t) - u_d(t)'u_d(t) + \|x(t)\|_Q^2 \right) dt \right.$$
$$\left. + \|x(t_f)\|_{Q_f}^2 + ON_a - ON_d \,\middle|\, x(0) = x_0 \right],$$
$$(5)$$

where $Q = \tilde{Q}(\omega_a^I, \omega_d^I)$, $Q_f = \tilde{Q}(\omega_a, \omega_d)$ with

$$\tilde{Q}(\omega_a^I, \omega_d^I) = \begin{bmatrix} (\omega_a^I - \omega_d^I)\mathbb{I}_n & \omega_d^I \mathbb{I}_n & -\omega_a^I \mathbb{I}_n \\ \omega_d^I \mathbb{I}_n & -\omega_d^I \mathbb{I}_n & \mathbf{0} \\ -\omega_a^I \mathbb{I}_n & \mathbf{0} & \omega_a^I \mathbb{I}_n \end{bmatrix}.$$

A. The Nash Control Strategies

Applying Itô's lemma and a completion of squares on (5) yields the following lemma.

Lemma 1. *The cost functional J in (5) with state dynamics (2) has the following form*

$$J = \mathbb{E}\left[\int_0^{t_f} \|u_a(t) + B_a'K(t)x(t)\|_2^2 - \|u_d - B_d'K(t)x(t)\|_2^2 dt \right.$$
$$\left. + \|x_0\|_{K(0)}^2 + \int_0^{t_f} \mathrm{Tr}\left(K(t)CC'\right) dt + ON_a - ON_d \right],$$
$$(6)$$

where $(K(t), t \in [0, t_f])$ is symmetric and satisfies the Riccatic equation

$$\dot{K}(t) = -K(t)A - A'K(t) - Q - K(t)\left(B_d B_d' - B_a B_a'\right)K(t)$$
$$\text{with } K(t_f) = Q_f$$
$$(7)$$

The proof follows standard arguments of the "completion of squares" procedures. Readers are referred to [18, Theorem 1], [14, Theorem 3.1], or [19, Lemma 1] for specifics. The existence of bounded solutions for (7) depends on Q, Q_f, and $B_d B_d' - B_a B_a'$, which we shall discuss later. Note that the attacker aims to find an observation-dependent strategy γ_a to minimize (6) while the defender desires to maximizes (6). From (6), we know that the choice of strategies only affect the terms within the first integral. Therefore, given $\mathcal{T}_a$ and $\mathcal{T}_d$, the Nash control strategies will be of the form

$$u_a^*(t) = -B_a'K(t)\hat{x}_1(t), \quad u_d^*(t) = B_d'K(t)\hat{x}_2(t), \quad (8)$$

for some $\hat{x}_1$ and $\hat{x}_2$. The choices $\hat{x}_1$ and $\hat{x}_2$ are made by the attacker and the defender respectively such that $u_a^*(t)$ is $\mathcal{I}_a(t)$ measurable and $u_d^*(t)$ is $\mathcal{I}_d(t)$ measurable. To decompose the Nash control strategies, we split the $3n \times 3n$ matrix in (8) as

$$K(t) = \begin{bmatrix} K_{11}(t) & K_{12}(t) \\ K_{12}(t)' & K_{22}(t) \end{bmatrix},$$

where K_{11} is an $2n \times 2n$ matrix function of time t. Similarly, Q and Q_f are partitioned into Q_{ij} and $Q_{f,ij}$ for $i, j \in \{1, 2\}$. To decompose the Nash control strategy, we derive the following decomposed Riccati equations

$$\dot{K}_{11} = -K_{11}\hat{A} - \hat{A}'K_{11} - Q_{11} - K_{11}\left(\hat{B}_d\hat{B}_d' - \hat{B}_a\hat{B}_a'\right)K_{11},$$
$$\text{with } K_{11}(t_f) = Q_{f,11},$$
$$(9)$$

$$\dot{K}_{12} = -K_{12}A_s - \hat{A}'K_{12} - Q_{12} - K_{12}\left(\hat{B}_d\hat{B}_d' - \hat{B}_a\hat{B}_a'\right)K_{11},$$
$$\text{with } K_{12}(t_f) = Q_{f,12},$$

where

$$\hat{A} = \begin{bmatrix} A_a & \mathbf{0} \\ \mathbf{0} & A_d \end{bmatrix}, \quad \hat{B}_a = \begin{bmatrix} \tilde{B}_a \\ \mathbf{0} \end{bmatrix}, \quad \hat{B}_d = \begin{bmatrix} \mathbf{0} \\ \tilde{B}_d \end{bmatrix}. \quad (10)$$

Theorem 1. *Suppose that $\mathcal{T}_a$ and $\mathcal{T}_d$ are known and the trajectory $(x_s(\tau), \tau \in [0, t_f])$ is given. The DA game defined by (2) and (6) admits a Nash control strategy*

$$u_a^* = -\hat{B}_a'K_{11}\begin{bmatrix} x_a \\ \hat{x}_{1,d} \end{bmatrix} - \hat{B}_a's, \quad (11)$$

$$u_d^* = \hat{B}_d K_{11}\begin{bmatrix} \hat{x}_{2,a} \\ x_d \end{bmatrix} + \hat{B}_d's, \quad (12)$$

2368

where $(K_{11}, t \in [0, t_f])$ is the solution of the Riccati equation (9), $\hat{A}$, $\hat{B}_a$, and $\hat{B}_d$ are defined in (10), and $(s(t), t \in [0, t_f])$ is generated by

$$\dot{s} = \left[-\hat{A}' - K_{11}\left(\hat{B}_d\hat{B}'_d - \hat{B}_a\hat{B}'_a\right)\right] s - Q_{12}x_s, \quad (13)$$

with $s(t_f) = Q_{f,12}x_s(t_f)$. Moreover, the estimate $\hat{x}_{1,d}$ of defender's state evolves as

$$\dot{\hat{x}}_{1,d} = A_d\hat{x}_{1,d} + \tilde{B}_d\hat{B}'_d\left(K_{11}\begin{bmatrix} x_a \\ \hat{x}_{1,d} \end{bmatrix} + s\right) \quad (14)$$

with $\hat{x}_{1,d}(0) = x_{d0}$ and $\hat{x}_{1,d}(t) = x_d(t)$ for every $t \in \mathcal{T}_a$, and

$$\dot{\hat{x}}_{2,a} = A_a\dot{\hat{x}}_{2,a} - \tilde{B}_a\hat{B}'_a\left(K_{11}\begin{bmatrix} \hat{x}_{2,a} \\ x_d \end{bmatrix} + s\right), \quad (15)$$

with $\hat{x}_{2,a}(0) = x_{a0}$ and $\hat{x}_{2,a}(t) = x_a(t)$ for all $t \in \mathcal{T}_d$.

Proof. In the proof, we drop the time index t in some places for convenience. From [15, Proof of Theorem 2], we know that consider the game defined by (5) and (2), (8) constitutes a Nash control strategy if

$$\mathbb{E}\left[x(t) - \hat{x}_1(t)|\mathcal{I}_a(t)\right] = 0, \ \mathbb{E}\left[x(t) - \hat{x}_2(t)|\mathcal{I}_a(t)\right] = 0, \ \text{and}$$
$$\mathbb{E}\left[x(t) - \hat{x}_2(t)|\mathcal{I}_d(t)\right] = 0, \ \mathbb{E}\left[x(t) - \hat{x}_1(t)|\mathcal{I}_d(t)\right] = 0. \quad (16)$$

With this result, if we let $\hat{x}_1 = [\hat{x}'_{1,a} \ \hat{x}'_{1,d} \ \hat{x}'_{1,s}]'$ and $\hat{x}_2 = [\hat{x}'_{2,a} \ \hat{x}'_{2,d} \ \hat{x}'_{2,s}]'$, we have $\hat{x}_{1,a} = x_a$, $\hat{x}_{1,d} = \mathbb{E}[x_d|\mathcal{I}_a]$, $\hat{x}_{1,s} = x_s$ and $\hat{x}_{2,a} = \mathbb{E}[x_a|\mathcal{I}_d]$ $\hat{x}_{2,d} = x_d$, $\hat{x}_{2,s} = x_s$.

Using (9), the Nash control strategy of the attacker can be decomposed as $u^*_a = -\hat{B}'_a K_{11}[x_a' \ \hat{x}'_{1,d}]' - \hat{B}'_a K_{12}x_s$, where the attacker's control is driven by his/her state, his/her estimate of the defender's state, as well as the trajectory of the asset. To eliminate the dependence of the control on A_s (which is introduced for auxiliary purpose), we let $s(t) = K_{12}(t)x_s(t)$. Note that

$$\dot{s} = \left[-\hat{A}' - K_{11}\left(\hat{B}_d\hat{B}'_d - \hat{B}_a\hat{B}'_a\right)\right] s - Q_{12}x_s,$$

with $s(t_f) = Q_{f,12}x_s(t_f)$. Hence, the control depends on the trajectory of the asset irrespective of A_s. Under the Nash control strategy, the dynamics of the attacker is

$$\dot{x}_a = A_ax_a(t) - \tilde{B}_a\hat{B}'_a K_{11}\begin{bmatrix} x'_a \\ \hat{x}_{1,d} \end{bmatrix} - \tilde{B}_a\hat{B}'_a s + C_a dw_a(t),$$

with $x_a(0) = x_{a0}$. From (16), and using the fact $\mathbb{E}[w_a(t)|\mathcal{I}_a(t)] = 0$, we have

$$\dot{\hat{x}}_{2,a} = A_a\dot{\hat{x}}_{2,a} - \tilde{B}_a\hat{B}'_a\left(K_{11}\begin{bmatrix} \hat{x}_{2,a} \\ \mathbb{E}[\hat{x}_{1,d}|\mathcal{I}_d] \end{bmatrix} + s\right)$$
$$= A_a\dot{\hat{x}}_{2,a} - \tilde{B}_a\hat{B}'_a\left(K_{11}\begin{bmatrix} \hat{x}_{2,a} \\ x_d \end{bmatrix} + s\right)$$

with $\hat{x}_{2,a}(0) = x_{a0}$ and $\hat{x}_{2,a}(t) = x_a(t)$ for every $t \in \mathcal{T}_d$. Hence, we fully characterizes the Nash control strategy of the defender (12). Similarly, we can obtain the attacker's estimate of the defender's state, which is given by (15). $\square$

Remark 1. *The attacker's estimate of the defender's state $\hat{x}_{1,d}$ evolves according to (14). The estimate $\hat{x}_{1,d}$ does not require the attacker to know the control of the defender. Every time the attacker choose to observe, he/she receives*

the actual state of the defender, i.e., $\hat{x}_{1,d}(t) = x_d(t), \forall t \in \mathcal{T}_a$. The solution of the Riccati equation (9) may admits a finite escape time since the conditions that Q_{11} is positive semi-definite and $\hat{B}_d\hat{B}'_d - \hat{B}_a\hat{B}'_a$ is positive-definite do not hold. We can use a more lenient condition given by [20, Corollary 5.13] to check the existence of a bounded solution of (9). Due to the space constraints, we do not restate the corollary here. Instead, we provide a closed-form bounded solution of the Riccati equation for our case study in Sec. IV.

B. The Nash Observation Choices

In Theorem 1, we characterize the Nash control strategies of both players when the observation instances are given. To understand how both players would select their observation instances, we need to obtain the cost functional under the Nash control strategies for any given $\mathcal{T}_a$ and $\mathcal{T}_b$. From the decomposition of the Riccati equation in (9), we know $B'_a Kx = \hat{B}'_a K_{11}[x'_a \ x'_d]' + \hat{B}'_a s$. We further decompose K_{11} into

$$K_{11} = \begin{bmatrix} K^{ul}_{11} & K^{ur}_{11} \\ K^{ur}_{11}{}' & K^{br}_{11} \end{bmatrix},$$

where K^{ul}_{11}, K^{ur}_{11}, and K^{br}_{11} are $n \times n$ matrices. Then, the first term in (6) can be written as

$$\left\|u^*_a(t) + B'_a K(t)x(t)\right\|^2_2 = \left\|\hat{B}'_a K_{11}\begin{bmatrix} 0 \\ x_d - \hat{x}_{1,d} \end{bmatrix}\right\|^2_2$$
$$= \|\tilde{B}'_a K^{ur}_{11}(x_d - \hat{x}_{1,d})\|^2_2.$$

Similarly, we obtain

$$\left\|u^*_d(t) - B'_d K(t)x(t)\right\|^2_2 = \|\tilde{B}'_d K^{ur}_{11}{}'(x_a - \hat{x}_{2,a})\|^2_2.$$

From (1) and (14), we know that

$$\mathbb{E}\left[(x_d(t) - \hat{x}_{1,d}(t))(x_d(t) - \hat{x}_{1,d}(t))'\right]$$
$$= \int_{\tilde{t}}^t e^{A_d(s-\tilde{t})} C_d C'_d e^{A_d(s-\tilde{t})'} ds, \quad (17)$$

where $\tilde{t}$ the latest observation before t, which is dependent on t and $\mathcal{T}_a$ and is defined as $\tilde{t} = \max\{\tau \mid \tau \in \mathcal{T}_a, \tau \leq t\}$. The discussion above leads to the following corollary.

Corollary 1. *For given $\mathcal{T}_a$ and $\mathcal{T}_b$, let $\mathcal{T}_a = \{t_{1,a}.t_{2,a}, \cdots, t_{N_a,a}\}$ and $\mathcal{T}_d = \{t_{1,d}, t_{2,d}, \cdots, t_{N_d,d}\}$. Under the Nash control strategies obtained in Theorem 1, the cost functional (6) becomes*

$$J(\gamma^*_a, \gamma^*_d, x_0) = \sum_{i=0}^{N_a} \int_{t_{i,a}}^{t_{i+1,a}} \mathrm{Tr}\left[\Sigma_{1,d}(t - t_{i,a})\varphi_a(t)\right] dt$$
$$- \sum_{i=0}^{N_d} \int_{t_{i,d}}^{t_{i+1,d}} \mathrm{Tr}\left[\Sigma_{2,a}(t - t_{i,d})\varphi_d(t)\right] dt \quad (18)$$
$$+ \int_0^{t_f} \mathrm{Tr}\left(K(t)CC'\right) dt + \|x_0\|^2_{K(0)} + ON_a - ON_d,$$

where $\Sigma_{1,d}(t) = \int_0^t e^{A_d\tau} C_d C'_d e^{A_d\tau} d\tau$, $\Sigma_{2,a} = \int_0^t e^{A_a\tau} C_a C'_a e^{A_a\tau} d\tau$, $t_{0,a} = t_{0,d} = 0$, and $t_{N_a+1,a} = t_{N_d+1,d} = t_f$. Moreover, $\varphi_a(t) = K^{ur}_{11}(t)' \tilde{B}_a \tilde{B}'_a K^{ur}_{11}(t)$ and $\varphi_d(t) = K^{ur}_{11}(t) \tilde{B}_d \tilde{B}_d{}' K^{ur}_{11}(t)$.

Corollary 1 presents the cost functional under the Nash control strategies. Among the six terms in (18), only the first and the last two terms are associated with $\mathcal{T}_a$ and $\mathcal{T}_d$. Note that the objective of the attacker is to find a set of observation instances $\mathcal{T}_a$ that minimizes J, while the defender aims to maximize J. Hence, to decide their observation instances, the attacker and the defender only have to consider the first two terms and the last two terms. Moreover, the effect of the two players' observation can be decoupled, by which we mean $\tilde{J}(\mathcal{T}_a, \mathcal{T}_b) = \tilde{J}_a(\mathcal{T}_a) - \tilde{J}_b(\mathcal{T}_b)$, where

$$
\begin{aligned}
\tilde{J}_a(\mathcal{T}_a) &= \sum_{i=0}^{N_a} \int_{t_{i,a}}^{t_{i+1,a}} \mathrm{Tr}\left[\Sigma_{1,d}(t - t_{i,a})\varphi_a(t)\right] dt + O N_a, \\
\tilde{J}_d(\mathcal{T}_d) &= \sum_{i=0}^{N_d} \int_{t_{i,d}}^{t_{i+1,d}} \mathrm{Tr}\left[\Sigma_{2,a}(t - t_{i,d})\varphi_d(t)\right] dt + O N_d.
\end{aligned}
\tag{19}
$$

Remark 2. *Corollary 1 shows that the optimal choices of observation instances do not depend on the trajectory of the asset, by which we mean no matter how the asset moves, the defender and the attacker's choices of observation instances will not be affected. This is due to the fact that both players know the trajectory of the asset. The relative position between the asset and the attacker can be estimated unbiasedly by the defender. The cost is captured by the variance of the estimate error which is independent of the asset's trajectory.*

Remark 3. *Since $\tilde{J}(\mathcal{T}_a, \mathcal{T}_d)$ can be decomposed into $\tilde{J}_a(\mathcal{T}_a) - \tilde{J}_d(\mathcal{T}_d)$, the Nash observation strategies that solve $\min_{\mathcal{T}_a} \max_{\mathcal{T}_d} \tilde{J}(\mathcal{T}_a, \mathcal{T}_d)$ can be obtained by solving $\min_{\mathcal{T}_a} \tilde{J}_a(\mathcal{T}_a)$ and $\min_{\mathcal{T}_b} \tilde{J}_d(\mathcal{T}_d)$. That means the defender and the attacker can make independent observation choices by solving two independent optimization problems. The independence comes from the fact that the defender and the attacker have independent dynamics in (1).*

Remark 4. *To provide insights on the cost functional, we take the attacker as an example. In the first term of $\tilde{J}_a(\mathcal{T}_a)$, we know*

$$
\begin{aligned}
&\mathrm{Tr}\left[\Sigma_{1,d}(t - \tilde{t})\varphi_a(t)\right] \\
&= \mathbb{E}\left[(x_d(t) - \hat{x}_{1,d}(t))' K_{11}^{ur\,'} \tilde{B}_a \tilde{B}_a' K_{11}^{ur} (x(t) - \hat{x}_{1,d}(t))\right],
\end{aligned}
\tag{20}
$$

where $\Sigma_{1,d}(t - \tilde{t})$ is the variance of the estimation error $x_d - \hat{x}_{1,d}$ at time and $\tilde{t}$ is the latest observation instance before time t. The term in (20) captures the the instantaneous cost at time t induced by the mismatch between the actual state of the defender and the attacker's estimate. The observation choices are control-aware by which we mean the estimation error is scaled by the matrix $K_{11}^{ur\,'} \tilde{B}_a \tilde{B}_a' K_{11}^{ur}$ and he matrix assign more weight to the estimation error corresponding to the states that are more informative to control needs. From (17), we know that the estimation error accumulates until the attacker makes an observation. As a result, the observation clears the estimation error. However, each observation made is subject to a cost O. Hence, the attacker has to make observation decision strategically over time. Overall, the observation decision has to consider the trade-off between the estimation error and the number of observations. Moreover, the observation instances need to be well designed by both players to minimize the corresponding integral terms in (19).

Since solving the Nash observation game is equivalent to solving two independent optimization problems, we focus on solving the attacker's optimization problem. One can obtain the result for the defender similarly. The solution of $\min_{\mathcal{T}_a} \tilde{J}_a(\mathcal{T}_a)$ involves two components: the optimal number of observations N_a^* and a set of optimal observation instances $\mathcal{T}_a^* = \{t_{1,a}^*, t_{2,a}^*, \cdots, t_{N_a^*,a}^*\}$ (with a slight abuse of notation here). Define

$$
\begin{aligned}
f_a^*(N_a) &:= \min_{t_1, \cdots, t_{N_a}} f_a(t_1, t_2, \cdots, t_{N_a}) \\
&s.t. \quad t_0 = 0, \; t_{N_a} = t_f, \\
&\quad\quad t_i \leq t_{i+1}, i = 0, 1, \cdots, N_a + 1,
\end{aligned}
\tag{21}
$$

with $f_a := \sum_{i=0}^{N_a} \int_{t_i}^{t_{i+1}} \mathrm{Tr}\left[\Sigma_{1,d}(t - t_i)\varphi_a(t)\right] dt$. From [21, Proposition 8.5.12], we know that if $\Sigma_1 \geq \Sigma_2$, $\mathrm{Tr}[\Sigma_1 M] \geq \mathrm{Tr}[\Sigma_1 M]$ for a positive semi-definite matrix M. Note that $\varphi_a(t)$ is positive semi-definite and $\Sigma_{1,d}(t) > \Sigma_{1,d}(t')$ for $t > t'$. Hence, $f_a^*(N_a)$ is a decreasing function of N_a, which aligns our intuition that the more observations received, the better the control would be. After solving (21), it remains to find the optimal number of observations N_a^* that minimizes $f_a(N_a) + O N_a$. In the following theorem, we show that there always exists a minimizer for the optimization problem in (21), which can be characterized by a set of necessary conditions.

Theorem 2. *There always exists a solution, denoted by N_a^* and $t_1^*, t_2^*, \cdots, t_{N_a^*}^*$, that solves $\min_{\mathcal{T}_a} \tilde{J}_a(\mathcal{T}_a)$. Furthermore, the optimal number of observations N_a^* is bounded, i.e.,*

$$
N_a^* < \frac{1}{O} \int_0^{t_f} \mathrm{Tr}\left[\Sigma_{1,d}(t)\varphi_a(t)\right] dt.
\tag{22}
$$

And the optimal observation instances $t_1^*, t_2^*, \cdots t_{N_a^*}^*$ need to satisfy

$$
\begin{aligned}
&\int_{t_{i-1}^*}^{t_i^*} \mathrm{Tr}\left[e^{A_d(t_i^* - t)} C_d C_d' e^{A_d(t_i^* - t)'} \varphi_a(t_i^*)\right] dt \\
&= \int_{t_i^*}^{t_{i+1}^*} \mathrm{Tr}\left[e^{A_d(t - t_i^*)} C_d C_d' e^{A_d(t - t_i^*)'} \varphi_a(t_i^*)\right] dt,
\end{aligned}
\tag{23}
$$

for $i = 1, 2, \cdots, N_a^*$.

The proof is presented in Appendix A. From Theorem 2, we know that the optimal number of observations N_a^* is bounded and inversely proportional to the observation cost O. From (23), one can say that the optimal observation instances are spread out over the horizon $[0, t_f]$. Given a limited number of observations, it is unwise to allocate two observation instances in a short period of time. The effect of control is applied via $\varphi_a(t)$. For some period when $\varphi_a(t)$ is large, e.g., $\varphi_a(t) \geq \varphi(t')$, for $t \in [\tau_1, \tau_2]$ and $t' \notin [\tau_1, \tau_2]$. Then in this period, the attacker tend to observes more frequently. For example, if the goal of the attacker is to hit the asset at time t_f, the attacker may need to observe more frequently at the end of the game.

As is shown in Appendix A, the differential of f_a can be calculated analytically. The second-order differential can also

be calculated analytically. Hence, we can leverage either first-order methods or second-order methods [22] to numerically compute $t_1^*, t_2^*, \cdots, t_{N_a}^*$. Also, (23) indicates that once t_1^* is provided, $t_2^*, \cdots, t_{N_a}^*$ can be computed easily. Based on this feature, we propose a binary search algorithm that solves problem (21). In Algorithm 1, we aim to find a $t_1^\star$ such that $|t_1^\star - t_1^*| < \epsilon/2$. Line 1 initializes all the parameters in (21). Line 2 sets the initial low bound t_{low} and upper bound t_{up} of t_1^* to be 0 and t_f respectively. The initial guess of t_1 is $(0 + t_f)/2$. Line 5 computes the left-hand side of (23), which we rewrite as

$$l_a(t_{i-1}, t_i) = \int_{t_{i-1}}^{t_i} \mathrm{Tr}\left[e^{A_d(t_i - t)} C_d C_d' e^{A_d(t_i - t)'} \varphi_a(t_i) \right] dt. \tag{24}$$

Line 6 computes the right-hand side integral in (23) from t_i to t_f, which we write as

$$r_a(t_i, t_f) = \int_{t_i}^{t_f} \mathrm{Tr}\left[e^{A_d(t - t_i)} C_d C_d' e^{A_d(t - t_i)'} \varphi_a(t) \right] dt. \tag{25}$$

Line 7-11 says for any $t_i, i = 1, 2, \cdots, N_a$ that is computed based on our guess t_1, if $r_a(t_i, t_f) < l_a(t_{i-1}, t_i)$, then our guess t_1 is larger than t_1^*. Hence, we set the upper bound t_{up} as t_1 and reset out guess t_1 as $t_1 = (t_{low} + t_1)/2$. Then we break the for loop and start with our new guess t_1. Line 12 computes the next observation instance using (23). Line 13-21 says that when the for loop gets to $i = N_a$, we compute $t_{N_a + 1}$. If $t_{N_a + 1} < t_f$, our guess t_1 must be smaller than t_1^*. Hence, we set $t_{low} = t_1$, let our new guess to be $t_1 = (t_{up} + t_1)/2$, and breaks the for loop. If $t_{N_a + 1} = t_f$ (it is impossible that $t_{N_a + 1} > t_f$ due to our operations in Line 5-11), then $t_1 = t_1^*$. Hence, we set $t_{low} = t_{up} = t_1$ to leave the while loop. Since the while ends when $|t_{up} - t_{low}| < \epsilon$, we can ensure $|t_1^\star - t_1^*| < \epsilon/2$, where t_1^* is the optimal first observation instance and $t_1^\star$ is the first observation instance found using Algorithm 1. The number of iterations needed for the while loop is less than $\min\{n \mid t_f/2^n \le \epsilon\}$. For example, only 20 iterations are needed to achieve $\epsilon = 10^{-5}$ when $t_f = 10$. Once $t_1^\star$ is obtained, the rest observation instances can be computed easily using (23). Note that with $f_a^*(N_a)$ being computed for some small N_a, a bound similar to yet tighter than (22) can be developed. For example, when $f_a^*(N_a)$ is computed for $N_a = 1, 2, 3$, if $N_a^* > 3$, we have $f_a^*(3) + 3O > ON_a^*$, i.e., $N_a^* - 3 \le f_a^*(3)/O$. Hence, we only need to compute $f_a^*(N_a)$ for a very limited number of N_a.

In the next section, we provide case studies to demonstrate the computation and the effectiveness of observation instances and offer more insights.

IV. Case Studies

We consider a "simple motion" dynamics of two players on a 2-D plane. This dynamic model has been widely adopted by existing works [3]–[5], [7]. The dynamics (1) becomes

$$\begin{aligned} dx_a &= b_a \cdot \mathbb{I}_2 u_a(t)dt + C_a dw_a(t), \quad x_a(0) = x_{a0}, \\ dx_d &= b_d \cdot \mathbb{I}_2 u_d(t)dt + C_d dw_d(t), \quad x_d(0) = x_{d0}, \end{aligned} \tag{26}$$

where $x_a \in \mathbb{R}^2$ and $x_d \in \mathbb{R}^2$, b_a and b_d are scalars that describe the maneuverability of the attacker and the defender

---

Algorithm 1 Optimal Observation Instances Algorithm Based on Binary Search

1: Initialize $A_d, C_d, N_a, \varphi_a(\cdot), t_f$, and tolerate, $\epsilon > 0$
2: Set $t_{low} = 0$, $t_0 = 0$, $t_{up} = t_f$, and $t_1 = (t_{up} + t_{low})/2$
3: **while** $|t_{up} - t_{low}| > \epsilon$ **do**
4:     **for** $i = 1, \cdots, N_a$ **do**
5:         Compute val $= l_a(t_{i-1}, t_i)$ defined in (24)
6:         Compute val$' = r_a(t_i, t_f)$ defined in (25)
7:         **if** val$'$ < val **then**
8:             $t_{up} = t_1$
9:             $t_1 = (t_{low} + t_1)/2$
10:            **break**
11:         **end if**
12:         Compute t_{i+1} using (23)
13:         **if** $i = N_a$ **then**
14:            **if** $t_{i+1} < t_f$ **then**
15:               $t_{low} = t_1$
16:               $t_1 = (t_{up} + t_1)/2$
17:               **break**
18:            **else**
19:               $t_{low} = t_{up} = t_1$
20:            **end if**
21:         **end if**
22:     **end for**
23: **end while**
24: **return** $t_1^\star = (t_{low} + t_{up})/2$

---

respectively. Suppose the attacker and the defender only care the terminal state, i.e., $\omega_d^I = \omega_a^I = 0$, and hence $Q = \mathbf{0}$.

Proposition 1. *For system (26) with $\omega_d^I = \omega_a^I = 0$, the Riccati equation admits a bounded closed-form solution*

$$K_{11}(t) = k(t) \begin{bmatrix} \kappa_1(t) & -1 \\ -1 & \kappa_2(t) \end{bmatrix} \otimes \mathbb{I}_2, \tag{27}$$

where

$$\kappa_1(t) = -b_d^2(t - t_f)\omega_a + 1 - \omega_a/\omega_b,$$
$$\kappa_2(t) = b_a^2(t - t_f)\omega_a + 1,$$
$$k(t) =$$
$$\frac{\omega_a \omega_d}{\left[\omega_a b_a^2(t - t_f) + 1\right]\left[-\omega_a \omega_d b_d^2(t - t_f) + \omega_d - \omega_a\right] - \omega_a}.$$

The optimal observation instances $t_1^*, t_2^*, \cdots, t_{N_a}^*$ need to satisfy

$$k^2(t_i^*)(t_i^* - t_{i-1}^*) = \int_{t_i^*}^{t_{i+1}^*} k^2(t)dt. \tag{28}$$

Proof. Under this setting, the riccati equation (10) becomes

$$\dot{K}_{11} = -K_{11}\left(\hat{B}_d\hat{B}_d' - \hat{B}_a\hat{B}_a'\right)K_{11}, \quad \text{with } K_{11} = Q_{f,11}.$$

Note that $Q_{f,11}$ is invertible. Indeed, we have

$$Q_{f,11}^{-1} = \begin{bmatrix} \frac{1}{\omega_a} & \frac{1}{\omega_a} \\ \frac{1}{\omega_a} & \frac{1}{\omega_a} - \frac{1}{\omega_d} \end{bmatrix} \otimes \mathbb{I}_2.$$
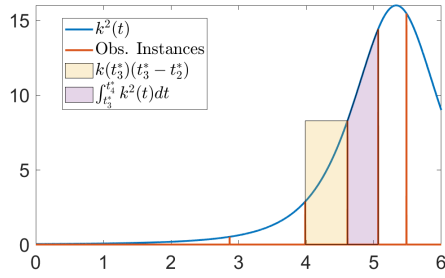
Fig. 1: $k^2(t)$ is defined in (27) and represented by blue line. The optimal observation instances $t_i^*, i = 1, \cdots, 5$ are represented by red impulses with height adjusted to $k^2(t_i^*)$. The area of the khaki box being equal to that of the lavender area illustrates the necessary conditions (28).

Fig. 2: The optimal number of observations for the attacker under different observation costs (left). The costs f_a under two choices of observation instances: the periodic instances (marked red) and the optimal instances (marked blue) computed using Algorithm 1 (right).

Let $\Gamma = K_{11}^{-1}$ for some $[t_f - \Delta t, t_f]$. Since $K_{11} K_{11}^{-1} = \mathbb{I}_4$,

$$\frac{d}{dt} K_{11}(t) K_{11}^{-1}(t) = \dot{K}_{11} K_{11}^{-1} + K_{11} \dot{K}_{11}^{-1} = 0,$$

which gives $\dot{K}_{11}^{-1} = -K_{11}^{-1} \dot{K}_{11} K_{11}^{-1}$. Hence,

$$\dot{\Gamma} = -K_{11}^{-1} \dot{K}_{11} K_{11}^{-1} = \hat{B}_a \hat{B}_a' - \hat{B}_d \hat{B}_d' = \begin{bmatrix} b_a^2 & 0 \\ 0 & -b_d^2 \end{bmatrix} \otimes \mathbb{I}_2.$$

Since $\Gamma(t_f) = Q_{f,11}^{-1}$, we obtain

$$\Gamma(t) = \begin{bmatrix} b_a^2(t - t_f) + \frac{1}{\omega_a} & \frac{1}{\omega_a} \\ \frac{1}{\omega_a} & -b_d^2(t - t_f) + \frac{1}{\omega_a} - \frac{1}{\omega_d} \end{bmatrix} \otimes \mathbb{I}_2.$$

Since $K_{11} = \Gamma^{-1}$, we have (27).

Then, we have $\varphi_a(t) = K_{11}^{ur}{}' \tilde{B}_a \tilde{B}_a K_{11}^{ur} = k^2(t) b_a^2 \mathbb{I}_2$ and similarly, $\varphi_b(t) = k^2(t) b_d^2 \mathbb{I}_2$. Note that $A_a = A_d = 0$, which gives $\Sigma_{1,d}(t) = t C_d C_d'$ and $\Sigma_{2,a} = t C_a C_a'$. Hence,

$$f_a^*(N_a) = \text{Tr}[b_a^2 C_d C_d'] \cdot \min_{t_1, \cdots, t_{N_a}} \sum_{i=0}^{N_a} \int_{t_i}^{t_{i+1}} k^2(t)(t - t_i) dt.$$

Furthermore, (23) becomes

$$\int_{t_{i-1}^*}^{t_i^*} \text{Tr}\left[C_d C_d' k^2(t_i^*) b_a^2 \right] dt = \int_{t_i^*}^{t_{i+1}^*} \text{Tr}\left[C_d C_d' k^2(t) b_a^2 \right] dt,$$

$$\Rightarrow k^2(t_i^*)(t_i^* - t_{i-1}^*) = \int_{t_i^*}^{t_{i+1}^*} k^2(t) dt.$$

$\square$

Other parameters are set to be $\omega_d = 3, C_d = C_a = 2\mathbb{I}$, $b_a = b_d = 0.5$. We use Algorithm 1 to compute the optimal observation instances for the attacker. In fig. 1, we first plot the Riccati equation component $k^2(t)$ over $[0, t_f]$. The value of $k^2(t)$ captures the importance level of an observation at time t for the attacker. For example, $k^2(t)$ attains its highest value around 5.3. In an optimal solution, observations occur more frequently around time 5.3. In fact, an observation at the beginning is not as valuable as an observation near the terminal time, as we can see from the curve of $k^2(\cdot)$. This is because $\omega_d^I = \omega_d^I = 0$ and the attacker only cares about the relative positions between him/her and his/her opponents as well as the asset. We next present the optimal observation
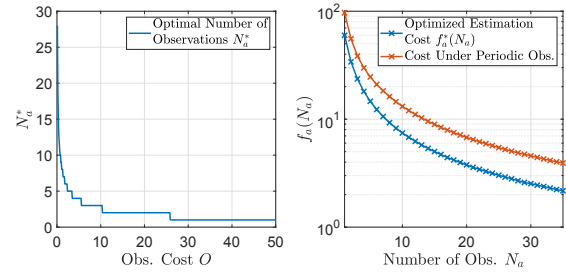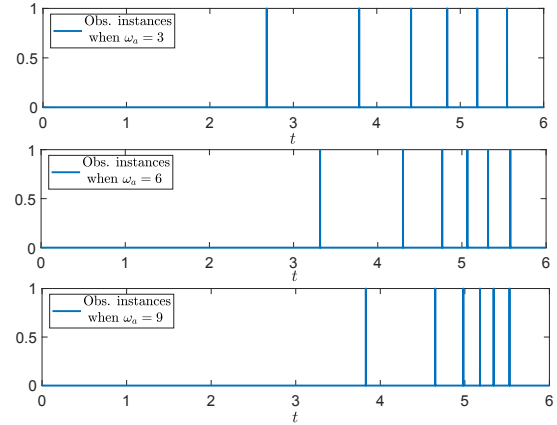

Fig. 3: When $N_a = 5$, the optimal observation instances under different ω_a.

instances when the number of observations the attacker can take is limited to $N_a = 5$. The first observation occurs late at $t = 2.8682$ and the fourth and the fifth observation instances are the closest. The distribution of the observation instances illustrates how $k^2(\cdot)$ affect the attacker's observation choices. The khaki box represents the left-hand side of (28) and the lavender area represents the right-hand side of (28). The necessary conditions (28) requires the areas of the two areas to be equal. This equality holds for each such neighboring areas.

In Fig. 2, the left figure shows that the optimal number of observations is inversely proportional to the observation cost. That means the defender can limit the performance of the attacker by increasing the attacker's observation cost. The increase of observation cost can be done, for example, by leveling up the defender's stealthiness. We compare two sets of observation instances in the right plot of Fig. 2. One is the optimal observation instances $t_1^\star, t_2^\star, \cdots, t_{N_a}^\star$ calculated using Algorithm 1 and one is the periodic observation instances chosen as $t_i = i \cdot t_f / (N_a + 1)$. The optimal observation instances induce less cost than the periodic observation instances given the same number of observations N_a. More precisely, by adopting the optimal observation instances, the attacker can reduce at least 30% of the cost generated by adopting the periodic observation instances. Similar results

can also be obtained for the defender.

In Fig. 3, we present the optimal time instances under different ω_a. Note that ω_a is the weight assigned to the terminal distance between the attacker and the defender, as one can see in (4). Even though the location of the asset is fully known to the attacker and the observations are made to observe the defender's location, the weight assigned to the distance between the attacker and the asset still affect the choice of the optimal observation instances. As ω_a increases, the attacker tends to observe late and observation instances get closer to each other. This is due to the fact that ω_a increases the terminal cost while the transient cost is zero because $\omega_a^I = \omega_d^I = 0$.

V. Conclusions

In this paper, we look into a DADG with partial information and selected observations. Due to the LQG formulation, both players' observation choices are independent of the asset's trajectory. We decouple the Nash observation game into two separate optimization problems. We exploit a set of necessary conditions to develop an effective algorithm to compute the optimal observation instances. Case studies show that the optimal observation instances outperform the periodic observation strategy by at least 30%. Future works can focus on studying noisy observation settings and investigating statistical results regarding success and capturability rates.

References

[1] Y. Liu, J. B. Cruz, and C. J. Schumacher, "Pop-up threat models for persistent area denial," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 2, pp. 509–521, 2007.

[2] I. E. Weintraub, M. Pachter, and E. Garcia, "An introduction to pursuit-evasion differential games," in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 1049–1066.

[3] S. Talebi, M. A. Simaan, and Z. Qu, "Cooperative, non-cooperative and greedy pursuers strategies in multi-player pursuit-evasion games," in *2017 IEEE Conference on Control Technology and Applications (CCTA)*. IEEE, 2017, pp. 2049–2056.

[4] 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.

[5] E. Garcia, D. W. Casbeer, M. Pachter, J. W. Curtis, and E. Doucette, "A two-team linear quadratic differential game of defending a target," in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 1665–1670.

[6] 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, 2018.

[7] ——, "Defense of a target against intelligent adversaries: A linear quadratic formulation," in *2020 IEEE Conference on Control Technology and Applications (CCTA)*. IEEE, 2020, pp. 619–624.

[8] L. Liang, F. Deng, Z. Peng, X. Li, and W. Zha, "A differential game for cooperative target defense," *Automatica*, vol. 102, pp. 58–71, 2019.

[9] M. Pachter, E. Garcia, and D. W. Casbeer, "Differential game of guarding a target," *Journal of Guidance, Control, and Dynamics*, vol. 40, no. 11, pp. 2991–2998, 2017.

[10] C. Cooper and N. Hahi, "An optimal stochastic control problem with observation cost," *IEEE Transactions on Automatic Control*, vol. 16, no. 2, pp. 185–189, 1971.

[11] Y. Huang, V. Kavitha, and Q. Zhu, "Continuous-time markov decision processes with controlled observations," in *2019 57th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2019, pp. 32–39.

[12] G. J. Olsder, "On observation costs and information structures in stochastic differential games," in *Differential Games and Applications*. Springer, 1977, pp. 172–185.

[13] Y. Huang and Q. Zhu, "Infinite-horizon linear-quadratic-gaussian control with costly measurements," *arXiv preprint arXiv:2012.14925*, 2020.

[14] D. Maity and J. S. Baras, "Strategies for two-player differential games with costly information," in *2016 13th International Workshop on Discrete Event Systems (WODES)*. IEEE, 2016, pp. 211–216.

[15] ——, "Linear quadratic stochastic differential games under asymmetric value of information," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 8957–8962, 2017.

[16] Y. Huang, Z. Xiong, and Q. Zhu, "Ccross-layer coordinated attacks on cyber-physical systems: A lqg game framework with controlled observations," in *2021 European Control Conference (ECC)*. IEEE, 2021, available at: https://arxiv.org/abs/2012.02384.

[17] D. Maity, A. Raghavan, and J. S. Baras, "Stochastic differential linear-quadratic games with intermittent asymmetric observations," in *2017 American Control Conference (ACC)*. IEEE, 2017, pp. 3670–3675.

[18] Y. Chen, T. T. Georgiou, and M. Pavon, "Covariance steering in zero-sum linear-quadratic two-player differential games," in *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 8204–8209.

[19] Y. Huang and Q. Zhu, "A pursuit-evasion differential game with strategic information acquisition," *arXiv preprint arXiv:2102.05469*, 2021.

[20] J. Engwerda, *LQ dynamic optimization and differential games*. John Wiley & Sons, 2005.

[21] D. S. Bernstein, *Matrix mathematics: theory, facts, and formulas*. Princeton university press, 2009.

[22] P. E. Gill, W. Murray, M. H. Wright *et al.*, *Numerical linear algebra and optimization*. Addison-Wesley Redwood City, CA, 1991, vol. 1.

Appendix

A. Proof of Theorem 2

Proof. First, note that $f_a(t_1, t_2, \cdots, t_{N_a})$ is differentiable over $t_1, t_2, \cdots, t_{N_a}$. Indeed, using Leibniz integral rule, we obtain

$$
\begin{aligned}
&\frac{\partial}{\partial t_i} f_a(t_1, \cdots, t_{N_a}) \\
&= \frac{\partial}{\partial t_i} \left[\int_{t_{i-1}}^{t_i} \mathrm{Tr}[\Sigma_{1,d}(t - t_{i-1})\varphi_a(t)] dt \right. \\
&\quad \left. - \int_{t_i}^{t_{i+1}} \mathrm{Tr}[\Sigma_{1,d}(t - t_i)\varphi_a(t)] dt \right] \\
&= \int_{t_{i-1}}^{t_i} \mathrm{Tr}\left[e^{A_d(t_i-t)} C_d C_d' e^{A_d(t_i-t)'} \varphi_a(t_i) \right] dt \\
&\quad - \int_{t_i}^{t_{i+1}} \mathrm{Tr}\left[e^{A_d(t-t_i)} C_d C_d' e^{A_d(t-t_i)'} \varphi_a(t) \right] dt.
\end{aligned}
$$

Hence, f_a is continuous over $t_1, t_2, \cdots, t_{N_a}$. Besides, the constraint set in (21) is closed and bounded (hence compact). Then by Weierstrass extreme value theorem, there exists at least one minimizer $t_1^*, t_2^*, \cdots, t_{N_a}^*$ for problem (21) for any given N_a. To obtain the first-order optimality condition, we let $\frac{\partial}{\partial t_i} f_a(t_1, t_2, \cdots, t_{N_a}) = 0$, which yields (23). Further, notice that $\tilde{J}_a(\mathcal{T}_a^*) \leq \tilde{J}_a(\varnothing) = f_a^*(0) = \int_0^{t_f} \mathrm{Tr}[\Sigma_{1,d}(t)\varphi_a(t)]$, and $\tilde{J}_a(\mathcal{T}_a^*) > ON_a^*$. Hence, we have $ON_a^* < \int_0^{t_f} \mathrm{Tr}[\Sigma_{1,d}(t)\varphi_a(t)]$, which gives (22). Since $N_a^* \in \mathbb{N}$ is bounded and $f^*(N_a)$ for any N_a, there always exist N_a^* and $t_1^*, \cdots, t_{N_a}^*$ that solve $\min_{\mathcal{T}_a} \tilde{J}_a(\mathcal{T}_a)$. $\square$