

# Optimal Intermittent Sensing for Pursuit-Evasion Games

Dipankar Maity

**Abstract**—We consider a class of pursuit-evasion differential games in which the evader has continuous access to the pursuer’s location, but not vice-versa. There is a remote sensor (e.g., a radar station) that can sense the evader’s location upon a request from the pursuer and communicate that sensed location to the pursuer. The pursuer has a budget on the total number of sensing requests. The outcome of the game is determined by the players’ sensing and motion strategies. We obtain the equilibrium sensing and motion strategies for the players. We quantify the degradation in the pursuer’s pay-off due to its sensing limitations.

**Index Terms**—Pursuit-evasion, intermittent sensing, self-triggered control.

## I. INTRODUCTION

Pursuit-Evasion games [1] have been applied to investigate a wide class of civilian and military applications involving multi-agent interactions in adversarial scenarios [2], [3], [4]. In its simplest form the game involves a pursuing player that is tasked to capture an evading player before either the evader reaches its destination or the pursuer runs out of fuel.

While several variations ranging from complex dynamic models for the players (e.g., [5]) to complex geometry of the environment (e.g., [6]) to limited visibility of the players (e.g., [7]) have been considered, one of the prevailing assumptions have been the continuous sensing capability for the players, with the exception of [8], [9], [10] among few others. By ‘continuous sensing’ we refer to the capability that enables the players to keep their sensors turned on continuously for the entire duration of the game. Extensions of pursuit-evasion games in the context of sensing limitations have mainly considered limited sensing range [11], [12], [13] and limited field-of-view [14], but the challenges associated with the lack of continuous sensing remain unsolved.

In this letter, we revisit the classical pursuit evasion game in an obstacle-free environment where the pursuer does not have the continuous sensing capability. In particular, the pursuer relies on a remotely located sensor (e.g., a radar station) to sense the evader’s position. Upon request, the remote sensor can perfectly sense the location of the evader and share it with the pursuer.<sup>1</sup> The communication channel between the pursuer and the remote sensor is assumed to be noiseless, instantaneous (i.e., no delay), and perfectly reliable (i.e., no packet losses). The pursuer intermittently requests the evader’s location to update its pursuit strategy. Due to resource (e.g., energy) constraints, the pursuer can only make a maximum of  $n$  requests. On the other hand, the evader is able to sense the pursuer continuously and is aware of the sensing limitation of the pursuer. The objective of this work is to analyze the game under this asymmetric sensing

limitation to obtain the optimal time instances for sensing and the equilibrium motion strategies.

Our prior works [9], [15] considered a linear quadratic differential game formulation where both the agents had sensing limitations. These works were later extended to discrete time [16], infinite horizon [17], and recently to asset defense scenarios [10]. All these extensions rely on the linear-quadratic structure of the problem. Closely related to the problem considered in this letter are the works in [8] and [18] where the authors considered a self-triggered strategy to decide the sensing instances. Self-triggered strategies are conservative and often highly suboptimal. [18] extends the work of [8] by incorporating noisy sensor measurements. These works did not consider any budget on the number of sensing and therefore, the proposed sensing strategy is agnostic to the sensing budget  $n$ . With respect to these existing works, the contributions of this letter are: (i) Formulating a sensing (and fuel) limited pursuit-evasion game and analyzing the equilibrium sensing and motion strategies for the players, (ii) Demonstrating the optimality of a ‘waiting strategy’ for a pursuer equipped with only intermittent sensing, (iii) Finding the required number of sensing  $n_{\max}$  to ensure that the pursuer can perform the same as it does under continuous sensing, and (iv) Comparing the performance with existing work [8] to illustrate the reduction in the number of sensing achieved by our work.

The rest of the letter is organized as follows: In Sec. II we formulate the problem, describe our preliminary results, and derive the required number of sensing to ensure that the pursuer performs the same as it does in the continuous sensing case. Building upon this preliminary result, in Sec. III, we analyze the game and derive the associated value function. We further investigate how the value function behaves with respect to the sensing budget  $n$  in Sec. IV, and finally, we conclude the letter in Sec. V.

## II. PROBLEM FORMULATION AND PRELIMINARY RESULTS

We consider a pursuit evasion game with simple motion:

$$\begin{aligned}\dot{\mathbf{x}}_p(t) &= \mathbf{v}_p(t), \\ \dot{\mathbf{x}}_e(t) &= \mathbf{v}_e(t),\end{aligned}\tag{1}$$

where  $\mathbf{x}_p(t) \in \mathbb{R}^2$ ,  $\mathbf{x}_e(t) \in \mathbb{R}^2$  denote the locations and  $\mathbf{v}_p(t) \in \mathbb{R}^2$ ,  $\mathbf{v}_e(t) \in \mathbb{R}^2$  denote the velocities of the pursuer and the evader, respectively, at time  $t$ . We assume that the game starts at time  $t = 0$  and both the players know the initial locations  $\mathbf{x}_p(0)$  and  $\mathbf{x}_e(0)$ . The maximum speeds for the pursuer and the evader are 1 and  $\nu$ , respectively.<sup>2</sup> We consider a faster pursuer, i.e.,  $\nu < 1$ , to avoid a trivial game.

For a pursuer with capture radius  $r_{\text{cap}}$ , capture happens as soon as the distance between the pursuer and the evader

<sup>2</sup>In case the maximum speeds are  $v_{p,\max}$  and  $v_{e,\max}$ , we shall use scaling and time dilation to transform the system to obtain maximum speeds 1 and  $\nu$ , respectively. To that end, we define  $\nu = v_{e,\max}/v_{p,\max}$ ,  $c = v_{p,\max}$ ,  $\bar{\mathbf{x}}_i(t) = \mathbf{x}_i(t/c)$  and  $\bar{\mathbf{v}}_i(t) = \mathbf{v}_i(t/c)/c$  for  $i = p, e$ . With this new scaling, we obtain  $\bar{\mathbf{x}}_i = \bar{\mathbf{v}}_i$  with  $\|\bar{\mathbf{v}}_p(t)\| \leq 1$  and  $\|\bar{\mathbf{v}}_e(t)\| \leq \nu$ .

The work of D. Maity was supported, in parts, by ARL grant ARL DCIST CRA W911NF-17-2-0181.

D. Maity is with the Department of Electrical and Computer Engineering, University of North Carolina at Charlotte, NC, 28223, USA. Email: dmaity@uncc.edu

<sup>1</sup>One may alternatively also consider a scenario where the pursuer has an onboard sensor to measure the evader’s location, however, it can only use the sensor intermittently due to, for example, energy and computational constraints.

becomes  $r_{\text{cap}}$ . The pursuer's objective is to minimize its distance from the evader at the end of the game and, if possible, to capture the evader within the game duration, in which case the game ends as soon as the evader is captured. The pay-off function for this game is as follows

$$J = \begin{cases} 0, & \text{if capture happened,} \\ \|\mathbf{x}_p(t_f) - \mathbf{x}_e(t_f)\| - r_{\text{cap}}, & \text{otherwise.} \end{cases} \quad (2)$$

The evader maximizes the pay-off while the pursuer minimizes it. Although we are particularly interested in the pay-off function (2), our analysis can be extended to a more general class of pay-off functions where

$$J = \begin{cases} 0, & \text{if capture happened,} \\ \phi(\|\mathbf{x}_p(t_f) - \mathbf{x}_e(t_f)\|), & \text{otherwise.} \end{cases} \quad (3)$$

where we assume the following conditions on function  $\phi$ .

*Assumption 1:*

- (A1)  $\phi : \mathbb{R} \rightarrow \mathbb{R}_+$  is a non-decreasing function with  $\phi(x) = 0$  for all  $x \leq r_{\text{cap}}$  and  $\phi(x) > 0$  otherwise.
- (A2)  $\phi$  is convex.

Assumption (A1) incentivizes the pursuer to minimize its final distance from the evader. The convexity assumption keeps the formulation analytically tractable while encompassing a wider class of problems than the one presented in (2).

#### A. Sensing Limited Pursuit Strategy: Preliminary Results

When both players can sense each other continuously, capture happens within a duration of  $\frac{\rho_0 - r_{\text{cap}}}{1 - \nu}$ , where  $\rho_0 \triangleq \|\mathbf{x}_p(0) - \mathbf{x}_e(0)\|$  is the initial distance between the players. In this case, the equilibrium strategy for both the players is to move along the pursuer's line-of-sight unit vector  $\mathbf{r}(t) \triangleq \frac{\mathbf{x}_e(t) - \mathbf{x}_p(t)}{\|\mathbf{x}_e(t) - \mathbf{x}_p(t)\|}$  with maximum speed for all  $t$ . One may verify that  $\dot{\mathbf{r}}(t) = 0$  under this equilibrium, and both players must move along the direction  $\mathbf{r}(0)$ . Consequently, the instantaneous heading angles for both the players do not change with time. It is noteworthy that *the players do not need their opponent's location continuously to implement their equilibrium strategies*. Although it appears that the sensing capability is redundant in implementing the equilibrium strategies, however, the lack thereof is detrimental for the pursuer since the evader can obtain a better pay-off by deviating from the direction  $\mathbf{r}(0)$  when the pursuer is 'blindly' moving along  $\mathbf{r}(0)$ . On the other hand, if the evader is 'blindly' moving along  $\mathbf{r}(0)$ , the pursuer still does not have any incentive to deviate from  $\mathbf{r}(0)$  since any deviation will degrade its pay-off. Therefore, a continuous sensing capability seems crucial for the pursuer.

In this letter, we consider only an intermittent sensing capability for the pursuer. Similar sensing constrained pursuit-evasion games have been previously studied in [8] where the pursuer follows a self-triggered strategy to sense the evader intermittently and correct its heading angle. In [8], the pursuer follows the pursuit strategy

$$\mathbf{v}_p(t) = \mathbf{r}(t_k) \quad \text{for all } t \in (t_k, t_{k+1}], \quad (4)$$

where  $t_k$  denotes the  $k$ -th sensing instance. We denote the 0-th sensing instance to be the initial time, i.e.,  $t_0 = 0$ . A self-trigger function was designed in [8] to determine the sensing instances and an upper bound  $n_{\text{max}}$  on the number

of required sensing was derived:

$$n_{\text{max}} = \left\lceil \frac{\log r_{\text{cap}} - \log \rho_0}{\log(h(\nu))} \right\rceil, \quad (5)$$

$$h(\nu) = 1 - \frac{\nu(1 - \nu)\sqrt{1 - \nu^2} - (1 - \nu)(1 - \nu^2)}{2\nu^2 - 1},$$

where  $\rho_0$  is the initial distance between the players. One may verify that  $1 > h(\nu) > \nu$  for all  $\nu \in (0, 1)$  and consequently, we get  $n_{\text{max}} \geq \left\lceil \frac{\log r_{\text{cap}} - \log \rho_0}{\log \nu} \right\rceil$ . Their proposed triggering strategy [8, Eq. (6)] yields

$$t_{k+1} - t_k = f(\nu)\|\mathbf{x}_e(t_k) - \mathbf{x}_p(t_k)\|,$$

where  $f(\nu) = \frac{\sqrt{1 - \nu^2}}{\nu + \sqrt{1 - \nu^2}}$ . Notice that  $f(\nu)$  decreases with  $\nu$  and converges to 0 as  $\nu$  converges to 1. In other words, the inter sensing time  $(t_{k+1} - t_k)$  decreases with  $\nu$ . In the following we state a more efficient sensing strategy with two major properties: (i) The total required number of sensing is strictly less than the  $n_{\text{max}}$  given in (5), and (ii) The capture time does not increase due to the sensing limitation. Before stating this main result for this section, we provide the following lemma that will be used in proving our result.

*Lemma 1:* If the initial distance between the players is less than or equal to  $\frac{r_{\text{cap}}}{\nu}$ , then there is no need for sensing and the equilibrium strategy for both the players is to move along the direction of  $\mathbf{r}(0)$ .  $\triangle$

*Proof:* Under the mentioned strategies  $\mathbf{v}_p(t) = \mathbf{r}(0)$  and  $\mathbf{v}_e(t) = \nu\mathbf{r}(0)$ , the capture time is  $\frac{\rho_0 - r_{\text{cap}}}{1 - \nu}$ , where  $\rho_0$  is the initial distance between the players. To prove that these strategies form an equilibrium pair, we let the evader follow an arbitrary strategy and show that the capture time is strictly reduced. To that end, note that  $\mathbf{x}_p(t) = \mathbf{x}_p(0) + t\mathbf{r}(0)$  and  $\mathbf{x}_e(t) = \mathbf{x}_e(0) + \int_0^t \mathbf{v}_e(s)ds$ , and therefore, for all  $t$ ,

$$\begin{aligned} \|\mathbf{x}_e(t) - \mathbf{x}_p(t)\| &= \|(\rho_0 - t)\mathbf{r}(0) + \int_0^t \mathbf{v}_e(s)ds\| \\ &\leq |\rho_0 - t| + \nu t, \end{aligned}$$

where the equality holds if and only if  $\mathbf{v}_e(t) = \nu\mathbf{r}(0)$  for all  $t$ . Consequently, at  $t = \frac{\rho_0 - r_{\text{cap}}}{1 - \nu}$ , we have  $\|\mathbf{x}_e(t) - \mathbf{x}_p(t)\| \leq r_{\text{cap}}$ , which implies that capture must have happened before  $t = \frac{\rho_0 - r_{\text{cap}}}{1 - \nu}$  unless the evader followed  $\mathbf{v}_e(t) = \nu\mathbf{r}(0)$ . In a similar fashion, one may also verify that there is no incentive for the pursuer to deviate from the strategy  $\mathbf{v}_p(t) = \mathbf{r}(0)$ .  $\blacksquare$

From Lemma 1, we notice that for any initial separation of  $\rho_0 \leq \frac{r_{\text{cap}}}{\nu}$ , the capture time is  $\frac{\rho_0 - r_{\text{cap}}}{1 - \nu}$ . This is the same capture time as what we also obtain from continuous sensing. In other words, the pursuer's performance (measured by its ability to capture and the associated capture time) with or without the sensing capabilities are the same when  $\rho_0 \leq \frac{r_{\text{cap}}}{\nu}$ . We now state the following Proposition on the number of required sensing for a given arbitrary initial distance  $\rho_0$ .

*Proposition 1:* Let the pursuer simply move to the last sensed location of the evader and request for a sensing as soon as it gets there. Then, the capture time is upper bounded by  $\frac{\rho_0 - r_{\text{cap}}}{1 - \nu}$  and the total number of required sensing is  $\left\lfloor \frac{\log r_{\text{cap}} - \log \rho_0}{\log \nu} \right\rfloor$ .  $\triangle$

*Proof:* Let the distance between the players at the  $k$ -th sensing moment be denoted by  $\rho_k \triangleq \|\mathbf{x}_e(t_k) - \mathbf{x}_p(t_k)\|$ . The next sensing request happens when the pursuer reaches  $\mathbf{x}_e(t_k)$ . The time taken by the pursuer to reach the point  $\mathbf{x}_e(t_k)$  is exactly  $\rho_k$ , i.e.,  $t_{k+1} - t_k = \rho_k$ . Therefore, the

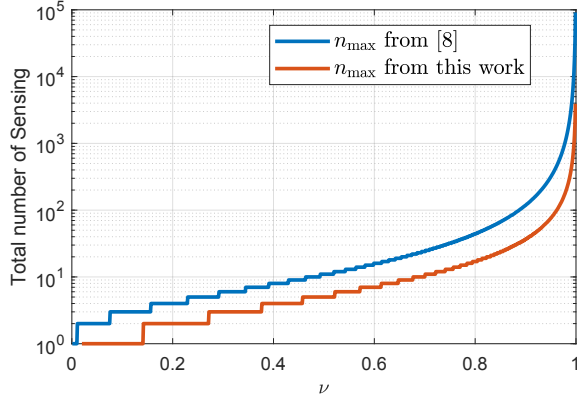


Fig. 1: The required number of sensing vs.  $\nu$ . The  $y$ -axis being in log scale shows that the required number of sensing according to [8] is several times higher than what is found in Proposition 1.

distance between the players at the  $(k+1)$ -th sensing instance is

$$\begin{aligned} \rho_{k+1} &= \|\mathbf{x}_e(t_{k+1}) - \mathbf{x}_p(t_{k+1})\| = \|\mathbf{x}_e(t_{k+1}) - \mathbf{x}_e(t_k)\| \\ &\leq \nu(t_{k+1} - t_k) = \nu\rho_k \leq \nu^{k+1}\rho_0. \end{aligned}$$

Therefore, with  $n_{\max} = \lfloor \frac{\log r_{\text{cap}} - \log \rho_0}{\log \nu} \rfloor$ , we obtain  $\rho_{n_{\max}} \leq \frac{r_{\text{cap}}}{\nu}$ . At this point, we invoke Lemma 1 to conclude that capture is inevitable without any further sensing. Furthermore from Lemma 1, starting with an initial distance of  $\rho_{n_{\max}} \leq \frac{r_{\text{cap}}}{\nu}$ , the capture time is  $\frac{\rho_{n_{\max}} - r_{\text{cap}}}{1 - \nu}$ . Therefore, the total capture time is

$$\begin{aligned} T_{\text{capture}} &= \frac{\rho_{n_{\max}} - r_{\text{cap}}}{1 - \nu} + \sum_{k=1}^{n_{\max}} (t_k - t_{k-1}) \\ &= \frac{\rho_{n_{\max}} - r_{\text{cap}}}{1 - \nu} + \sum_{k=1}^{n_{\max}} \rho_{k-1} \leq \frac{\rho_0 - r_{\text{cap}}}{1 - \nu}, \end{aligned}$$

where we have used  $t_0 = 0$  and  $\rho_k \leq \nu^k \rho_0$ . This completes the proof.  $\blacksquare$

In Fig. 1, we compare the number of required sensing proposed by [8] and by Proposition 1 to demonstrate the sensing efficiency of our proposed strategy. We notice that the difference between the  $n_{\max}$  proposed in [8] and that in Proposition 1 increases with  $\nu$  and there is an order of magnitude difference (note the log scale on the  $y$ -axis) when  $\nu$  is high. Next, we comment on the time consumed (or distance traveled) by the pursuer following the strategy in Proposition 1. This will be helpful in our subsequent analysis when we consider constraints on the fuel consumed by the pursuer (or equivalently, constraints on the game duration).

*Corollary 1:* Given  $\rho_0$  and  $r_{\text{cap}}$ , let  $m$  be the smallest integer such that  $r_{\text{cap}} > \nu^m \rho_0$ . Then,  $\max\{m-1, 0\}$  number of sensings are required. Furthermore, the pursuer will travel at most  $\frac{1-\nu^{m+1}}{1-\nu} \rho_0$  distance before capturing the evader.  $\triangle$

*Proof:* The proof follows directly from Proposition 1 where we use the inequality  $r_{\text{cap}} > \nu^m \rho_0$  to obtain the bounds on the number of sensing and the capture time.  $\blacksquare$

The results presented in this section assume that the pursuer is able to sense  $n_{\max}$  times as well as the pursuer has sufficient fuel to complete the game. However, in reality, there might be constraints on the duration of the game<sup>3</sup> as well as on the maximum number of allowed sensing. For

<sup>3</sup>Given the first order dynamics (1), a constraint on the fuel is cast as a constraint on the game duration.

the rest of the letter, we focus on a problem where the game is played for a duration of  $t_f$  and the maximum number of allowed sensing is  $n$ . The objective now is to analyze how the sensing strategy depends on  $t_f$  and  $n$ . Note that the sensing strategies proposed in [8] and in Proposition 1 are agnostic to  $t_f$  and  $n$ . In the subsequent sections, we formalize this problem and derive the optimal sensing strategy. At this point, note that when  $t_f \geq \frac{\rho_0 - r_{\text{cap}}}{1 - \nu}$  and  $n \geq \lfloor \frac{\log r_{\text{cap}} - \log \rho_0}{\log \nu} \rfloor$ , Proposition 1 is sufficient to construct a sensing strategy for the pursuer.

## B. Information Sets and Admissible Strategies

We consider the motion strategy for the pursuer to be similar to what was given in (4).

*Assumption 2:* For all  $t \in (t_k, t_{k+1}]$ , the pursuer follows

$$\mathbf{v}_p(t) = \gamma(t)\mathbf{r}(t_k) \quad (6)$$

where  $\gamma(t) \in [0, 1]$  is to be designed by the pursuer. Furthermore, the evader is aware of the pursuit strategy (6).

*Remark 1:* It is a rational choice to pick the pursuer's heading direction to be  $\mathbf{r}(t_k)$  since the pursuer is forced to operate in *open-loop* in between sensing instances as it does not have continuous access to  $\mathbf{r}(t)$ . Although it may seem that  $\gamma(t) \equiv 1$  is the best choice, we will prove that, under certain choices for the parameters,  $\gamma(t) = 0$  is optimal for some intervals of time. Note that  $\gamma(t) = 0$  represents a *waiting* behavior in the pursuer's strategy. We prove that, under certain cases, waiting has a clear advantage over continuously moving and/or sensing earlier. The waiting behavior has proven to be a crucial characteristics in some of the recent works involving sensing limited pursuit-evasion games [13].  $\triangle$

Although the dynamics (1) are deterministic, the outcome of the game is not necessarily deterministic if the players adopt randomized strategies. Therefore, the pursuer/evader minimizes/maximizes  $\mathbb{E}[J]$  where the expectation  $\mathbb{E}[\cdot]$  is with respect to the randomized strategies taken by the pursuer and the evader. Let  $\mu_e$  denote the motion strategy of the evader and  $\mu_p$  denote the sensing strategy of the pursuer. The strategies are considered to be time dependent, however, to maintain notational brevity, we will suppress such time dependencies. These strategies are measurable functions of the information sets of the players. To describe the information sets of the players, we first denote  $m(t)$  to be the total number of sensing requests up to time  $t$  and  $\mathcal{T}(t) \triangleq \{t_0, t_1, \dots, t_{m(t)}\}$  be the set of sensing instances up to time  $t$ , where  $t_0$  is the initial time, which is assumed to be 0 here, and  $t_i < t_{i+1}$  for all  $i$ , and  $t_{m(t)}$  is the latest sensing time. For all  $t$ , we have  $m(t) \leq n$  and  $t_{m(t)} \leq t$ . Furthermore, we denote  $\mathcal{I}_e(t) \triangleq \{\mathbf{x}_e(s), \mathbf{x}_p(s), \mathcal{T}(t) \mid s \leq t\}$  to be the information available to the evader at time  $t$  and  $\mathcal{I}_p(t) = \{\mathbf{x}_e(s'), \mathbf{x}_p(s), \mathcal{T}(t) \mid s' \in \mathcal{T}(t), s \leq t\}$  to be the pursuer's available information.<sup>4</sup>

For a given pair of evader and pursuer strategies  $(\mu_e, \mu_p)$ , we define  $\bar{J}(\mu_e, \mu_p) \triangleq \mathbb{E}[J]$  to be the expected pay-off from the given strategy pair. In the subsequent sections, we derive the equilibrium pair  $(\mu_e^*, \mu_p^*)$  such that for any admissible strategies  $\mu_e$  and  $\mu_p$ ,

$$\bar{J}(\mu_e^*, \mu_p) \geq \bar{J}(\mu_e^*, \mu_p^*) \geq \bar{J}(\mu_e, \mu_p^*).$$

<sup>4</sup>To be precise, the information sets are to be defined by the  $\sigma$ -algebras generated by these random variables.

### III. GAME ANALYSIS

For the subsequent analysis, we define  $V(\rho, \tau, \ell)$  to be the expected pay-off under a Nash equilibrium where  $\rho$  is the current distance between the players,  $\tau$  represents the remaining game duration, and  $\ell$  is the number of remaining sensing requests. Notice that, given the current time  $t$ , we have  $\tau = t_f - t$ . That is,

$$V(\rho, \tau, \ell) = \inf_{\mu_p} \sup_{\mu_e} \mathbb{E}[J \mid t = t_f - \tau, \|\mathbf{x}_e(t) - \mathbf{x}_p(t)\| = \rho, |\mathcal{T}(t)| = n - \ell],$$

where  $J$  is defined in (3) and  $|\mathcal{T}(t)|$  denotes the cardinality of the set  $\mathcal{T}(t)$ . Therefore,  $V(\rho_0, t_f, n) = \bar{J}(\mu_e^*, \mu_p^*)$ , where recall that  $\rho_0 = \|\mathbf{x}_p(0) - \mathbf{x}_e(0)\|$ .

To compute  $V(\rho, t_f, n)$ , we first note that, for all  $n$ ,

$$V(\rho, 0, n) = \begin{cases} 0, & \text{if } \rho \leq r_{\text{cap}}, \\ \phi(\rho), & \text{otherwise.} \end{cases} \quad (7)$$

Therefore, (7) defines a boundary condition on  $V$ . To simplify our derivation for  $V(\rho_0, t_f, n)$ , let us divide the entire game duration into intervals  $[t_0, t_1), [t_1, t_2), \dots, [t_n, t_f)$ , where  $t_i$  denotes the  $i$ -th sensing time. We construct  $V(\rho, \tau_i, n - i)$  backwards starting with  $i = n$ , where  $\tau_i = t_f - t_i$ . The following lemma computes  $V(\rho, \tau, 0)$  for any  $\rho$  and  $\tau$ .

**Lemma 2:** Let  $\rho$  be the distance between the players at the last sensing instance. Then,

$$V(\rho, \tau, 0) \leq \begin{cases} 0, & \text{if } \tau \geq \rho, \nu\rho \leq r_{\text{cap}}, \\ \phi(\nu\tau + [\rho - \tau]^+), & \text{otherwise.} \end{cases} \quad (8)$$

where we define  $[x]^+ \triangleq \max\{x, 0\}$ .  $\triangle$

**Proof: Case 1:**  $\tau \geq \rho, \nu\rho \leq r_{\text{cap}}$

In this case, the pursuer moves along  $\mathbf{r}(t_n)$  and capture is inevitable regardless of the evader's strategy.

**Case 2a:**  $\tau \geq \rho, \nu\rho > \sqrt{1 + \nu^2}r_{\text{cap}}$

In this case, the pursuer shall move to the last sensed location of the evader and stay there until the end of the game. The evader shall randomly pick either  $\mathbf{r}(t_n)^\perp$  or  $-\mathbf{r}(t_n)^\perp$  with equal probability and move along that direction with maximum speed, where  $\mathbf{r}(t_n)^\perp$  denotes a unit vector perpendicular to  $\mathbf{r}(t_n)$ . Notice that,

$$\begin{aligned} \mathbf{x}_p(t) &= \mathbf{x}_p(t_n) + \mathbf{r}(t_n) \int_{s=t_n}^t \gamma(s) ds, \\ \mathbf{x}_e(t) &= \mathbf{x}_e(t_n) + \theta\nu(t - t_n)\mathbf{r}(t_n)^\perp, \end{aligned}$$

where  $\gamma(s) = 1$  for all  $s \in [t_n, t_n + \rho]$  and zero otherwise, and  $\theta \in \{-1, 1\}$  is chosen uniform randomly by the evader at time  $t_n$ . One may verify that the minimum distance between the players during the interval  $[t_n, t_n + \rho]$  is  $\frac{\nu\rho}{\sqrt{1 + \nu^2}}$ . Since we are considering the case  $\nu\rho > \sqrt{1 + \nu^2}r_{\text{cap}}$ , capture does not happen during the interval  $[t_n, t_n + \rho]$ . To show that it is an equilibrium pair we prove that unilateral deviation in any of the player's strategy does not improve their pay-off. From the evader's perspective, since the pursuer is going to stop at  $\mathbf{x}_e(t_n)$ , the final distance between the players is independent of the evader's heading direction as long as the evader does not get captured before time  $t_n + \rho$ . Therefore, the evader cannot receive a better pay-off by deviating from this strategy. On the other hand, let us assume that the pursuer deviates from the proposed strategy and ends up at a different

point  $\mathbf{x}_p(t_n) + \alpha_1\mathbf{r}(t_n) + \alpha_2\mathbf{r}(t_n)^\perp$  where  $\alpha_1, \alpha_2 \in [-\tau, \tau]$  and  $\sqrt{\alpha_1^2 + \alpha_2^2} \leq \tau$ . One may verify that,

$$\begin{aligned} \|\mathbf{x}_e(t_f) - \mathbf{x}_p(t_f)\| &= \|(\rho - \alpha_1)\mathbf{r}(t_n) + (\theta\nu\tau - \alpha_2)\mathbf{r}(t_n)^\perp\| \\ &= \sqrt{(\rho - \alpha_1)^2 + (\theta\nu\tau - \alpha_2)^2} \triangleq g(\theta). \end{aligned}$$

Notice that  $g(\theta)$  is convex in  $\theta$ , and therefore, taking expectation of  $g(\theta)$  and using Jensen's inequality yields

$$\begin{aligned} \mathbb{E}[g(\theta)] &\geq \sqrt{(\rho - \alpha_1)^2 + \mathbb{E}[(\theta\nu\tau - \alpha_2)^2]} \\ &= \sqrt{(\rho - \alpha_1)^2 + \nu^2\tau^2 + \alpha_2^2}, \end{aligned} \quad (9)$$

where the last equality is obtained using  $\theta^2 = 1$  and  $\mathbb{E}[\theta] = 0$ . Therefore, the expected pay-off is

$$\begin{aligned} \mathbb{E}[\phi(\|\mathbf{x}_e(t_f) - \mathbf{x}_p(t_f)\|)] &\stackrel{(\dagger)}{\geq} \phi(\mathbb{E}[\|\mathbf{x}_e(t_f) - \mathbf{x}_p(t_f)\|]) \\ &\stackrel{(\ddagger)}{\geq} \phi(\sqrt{(\rho - \alpha_1)^2 + \nu^2\tau^2 + \alpha_2^2}), \end{aligned}$$

where the inequality  $(\dagger)$  is obtained by using Jensen's inequality to the convex function  $\phi$ . The second inequality  $(\ddagger)$  is obtained by combining the non-decreasing nature of  $\phi$  (Assumption 1) and (9). Since  $\phi$  is non-decreasing, we may further conclude that the optimal choices for  $\alpha_1$  and  $\alpha_2$  are  $\rho$  and 0, respectively. Consequently, the expected pay-off is  $\phi(\nu\tau)$ , which is the same as  $\phi(\nu\tau + [\rho - \tau]^+)$  since  $\tau \geq \rho$ .

**Case 2b:**  $\tau \geq \rho, r_{\text{cap}} < \nu\rho \leq \sqrt{1 + \nu^2}r_{\text{cap}}$

We prescribe the pursuer to follow the same strategy as in the last case. Notice that, since  $\nu\rho \leq (1 + \nu^2)r_{\text{cap}}$ , the evader cannot move in the perpendicular directions for the entire duration without being captured. Nonetheless, given that the evader picks a trajectory that does not lead to capture (such a trajectory exists<sup>5</sup>), the final distance between the players will be at most  $\nu\tau + (\rho - \tau)$ . Consequently, the pay-off from this case is no more than  $\phi(\nu\tau + [\rho - \tau]^+)$ .

**Case 3**  $\tau < \rho$ :

In this case, the optimal strategy for both the players is to move along  $\mathbf{r}(t_n)$  for the entire duration. One may verify that the proposed strategies constitute an equilibrium pair by considering unilateral deviations in the players' strategies. In this case,  $\|\mathbf{x}_e(t_f) - \mathbf{x}_p(t_f)\| = \rho - (1 - \nu)\tau = \nu\tau + (\rho - \tau)$  and thus, the pay-off can be written as  $\phi(\nu\tau + [\rho - \tau]^+)$ .

This completes the proof.  $\blacksquare$

Two important remarks are in order from Lemma 2, one regarding the tightness of the inequality (8) and the other regarding the optimality of a *waiting behavior* in the pursuer's strategy.

**Remark 2:** Although Lemma 2 provides an upper bound on  $V(\rho, t_n, 0)$ , the upper bound is tight everywhere except in the region  $\Omega_0 \triangleq \{t_f - t_n \geq \rho, r_{\text{cap}} < \nu\rho \leq \sqrt{1 + \nu^2}r_{\text{cap}}\}$  as discussed under **Case 2b** in Lemma 2. In practice, this region is small as can be seen in Fig. 2(b).  $\triangle$

**Remark 3:** We notice in the proof of Lemma 2 that, when  $\tau \geq \rho$ , the optimal motion strategy for the pursuer is to move to the last sensed location of the evader,  $\mathbf{x}_e(t_n)$ , and *wait* there. This is quite intuitive since moving at any arbitrary direction without the knowledge of the evader's location is harmful w.r.t. the expected pay-off. The pursuer could have also waited at the beginning (or at any other point) of the game for a duration of  $\tau - \rho$  and then move toward  $\mathbf{x}_e(t_n)$ .

<sup>5</sup>For example, let the evader pick the direction  $R(\theta)\mathbf{r}(t_n)$  with  $\theta = \cos^{-1} \nu$  where  $R(\theta) \triangleq \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$ .

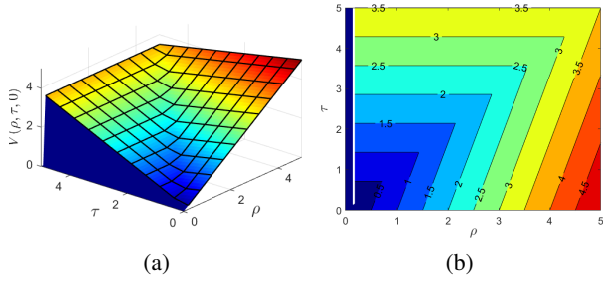


Fig. 2: We choose  $r_{\text{cap}} = 0.1$  and  $\nu = 0.7$ , and  $\phi(x) = [x - r_{\text{cap}}]^+$  for this plot. (a) The upperbound of  $V(\rho, \tau, 0)$  from Lemma 2. (b) Contour plot of  $V(\rho, \tau, 0)$  on the  $(\rho, \tau)$  plane. The numbers on each contour show the value of  $V(\rho, \tau, 0)$  along that contour. The thin vertical white slice approximately along  $\rho = 0.15$  illustrates the region  $\Omega_0$ . This is the only region in which we have an upper bound of  $V(\rho, \tau, 0)$  instead of an exact expression.

Alternatively, the pursuer could have moved for the entire duration with a slower speed. In summary, all the scenarios lead to the same conclusion that  $\gamma(t) \equiv 1$  is not an optimal strategy for the entire duration.  $\triangle$

The following theorem is the main result of this section where we provide an upper bound on  $V(\rho, \tau, \ell)$  for all  $\rho, \tau$  and  $\ell \leq n$ . The upper bound is tight everywhere except in the region  $\{\tau \geq \frac{1-\nu^{\ell+1}}{1-\nu}\rho, r_{\text{cap}} \leq \nu\rho \leq \sqrt{1+\nu^2}r_{\text{cap}}\} \subseteq \Omega_0$ .

**Theorem 1:** Let  $t_k$  be the  $k$ -th sensing time and  $\rho$  be the distance between the players at that instance. Let  $\tau = t_f - t_k$  and  $\ell = n - k$  denote the remaining time and the remaining number of sensing, respectively. Then,

$$V(\rho, \tau, \ell) \leq \begin{cases} 0, & \text{if } \tau \geq \frac{1-\nu^{\ell+1}}{1-\nu}\rho, \nu^{\ell+1}\rho \leq r_{\text{cap}}, \\ \phi(\nu\tau + \rho - \tau), & \text{if } \tau \leq \frac{1-\nu^{\ell+1}}{1-\nu}\rho, \\ \phi(\frac{1-\nu}{1-\nu^{\ell+1}}\nu^{\ell+1}\tau), & \text{otherwise.} \end{cases} \quad \triangle$$

*Proof:* (Sketch) The proof follows similar steps as Lemma 2.

**Case 1:**  $\tau \geq \frac{1-\nu^{\ell+1}}{1-\nu}\rho, \nu^{\ell+1}\rho \leq r_{\text{cap}}$

The strategy for the pursuer is to go to the last sensed location of the evader and request for sensing. The evader on the other hand shall go along one of the directions  $\pm \mathbf{r}(t_i)^\perp$  chosen uniformly randomly for the interval  $(t_i, t_{i+1}]$ , where  $i = (n - \ell), (n - \ell + 1), \dots, n$ , and we define  $t_{n+1} \triangleq t_f$ .

**Case 2:**  $\tau \leq \frac{1-\nu^{\ell+1}}{1-\nu}\rho$

The strategy for the pursuer remains the same as the last case. The strategy for the evader also remains the same for the intervals  $\{(t_i, t_{i+1}]\}_{i=n-\ell}^{n-1}$ . During the interval  $(t_n, t_f]$ , the evader shall go along  $\mathbf{r}(t_n)$  (this interval falls under **Case 3** of Lemma 2).

**Case 3:**  $\tau \geq \frac{1-\nu^{\ell+1}}{1-\nu}\rho, \nu^{\ell+1}\rho > r_{\text{cap}}$

The strategy for the pursuer is to go to  $\mathbf{x}_e(t_{n-\ell})$  and wait there for  $w = \frac{1-\nu}{1-\nu^{\ell+1}}\tau - \rho$  amount of time before requesting the first sensing. Therefore, at the moment of sensing, the remaining distance, time, sensing requests are  $\rho' = \nu \frac{1-\nu}{1-\nu^{\ell+1}}\tau$ ,  $\tau' = \frac{\nu(1-\nu^\ell)}{1-\nu^{\ell+1}}\tau$ , and  $\ell' = \ell - 1$ , respectively. Since  $\rho', \tau'$ , and  $\ell'$  satisfy  $\tau' = \frac{1-\nu^{\ell'+1}}{1-\nu}\rho'$ , we may invoke **Case 2** of this theorem to conclude  $V(\rho', \tau', \ell') = \phi(\nu\tau' + \rho' - \tau')$ . By substituting the expressions of  $\rho', \tau'$  and  $\ell'$ , we obtain  $\phi(\nu\tau' + \rho' - \tau') = \phi(\frac{1-\nu}{1-\nu^{\ell+1}}\nu^{\ell+1}\tau)$ . If  $\sqrt{1+\nu^2}r_{\text{cap}} < \nu\rho$ , the evader must move along  $\pm \mathbf{r}(t_i)^\perp$  (chosen uniformly randomly) for  $i = (n - \ell), \dots, (n - 1)$ , and consequently  $V(\rho, \tau, \ell) = \phi(\frac{1-\nu}{1-\nu^{\ell+1}}\nu^{\ell+1}\tau)$ . If  $\sqrt{1+\nu^2}r_{\text{cap}} \geq \nu\rho$ , the evader must move in a fashion to escape from being

captured in the first interval  $(t_{n-\ell}, t_{n-\ell+1}]$ . In this case,  $\phi(\frac{1-\nu}{1-\nu^{\ell+1}}\nu^{\ell+1}\tau)$  is an upper bound of  $V(\rho, \tau, \ell)$ .  $\blacksquare$

**Remark 4:**  $V(\rho_0, t_f, n)$  is a special case of Theorem 1. According to Theorem 1, the pursuer must wait before requesting the first sensing. The waiting time is constructed in such a way that the total distance traveled starting from  $t_1$  (i.e., the first sensing instance) is the same as the duration left. In this way the pursuer does not need to wait in the future. Furthermore, it can be shown that, if the pursuer requests the first sensing before waiting for  $\frac{1-\nu}{1-\nu^{\ell+1}}\tau - \rho$ , then it results in a suboptimal outcome for the pursuer. The proof can be found in the unabridged version [19].  $\triangle$

Using the result from Theorem 1, we plot  $V(\rho, \tau, \ell)$  in Fig. 3. For a given  $\rho$  and  $\tau$ ,  $V(\rho, \tau, \ell)$  is a non-increasing function of  $\ell$ , as one would have expected. In the next section, we investigate how  $V(\rho_0, t_f, n)$  behaves with respect  $n$ .

#### IV. PERFORMANCE DEGRADATION DUE TO SENSING LIMITATION

The objective of this section is to quantify the degradation in the pursuer's performance due to its sensing limitation. For an initial distance of  $\rho_0$  and a game duration of  $t_f$ , the pursuer's pay-off under continuous sensing is  $\phi([\rho_0 - (1 - \nu)t_f]^+)$ , which will serve as the baseline for quantifying the performance degradation. In particular, for a given  $\nu, \rho_0$  and  $t_f$ , we use the metric  $\delta(n) \triangleq V(\rho_0, t_f, n) - \phi([\rho_0 - (1 - \nu)t_f]^+)$  to quantify the performance degradation. Since we only have an upper bound of  $V(\rho, \cdot, \cdot)$  in  $\Omega_0$ , we exclude this region from the discussion in this section and make the following assumption.

**Assumption 3:** We assume that  $\nu\rho_0 > \sqrt{1+\nu^2}r_{\text{cap}}$ .

In the following, we first state a proposition that provides the required number of sensing ( $n_{\text{req}}$ ) to ensure  $\delta(n_{\text{req}}) = 0$ .

**Proposition 2:** For a given initial distance  $\rho_0$  and a game duration  $t_f$ , the pursuer's pay-off is  $\phi([\rho_0 - (1 - \nu)t_f]^+)$  with  $n_{\text{req}}$  sensing, where

$$n_{\text{req}} = \begin{cases} \left\lfloor \frac{\log(\rho_0 - (1-\nu)t_f) - \log(\rho_0)}{\log(\nu)} \right\rfloor, & \text{if } t_f < \frac{\rho_0 - r_{\text{cap}}}{1-\nu}, \\ \left\lfloor \frac{\log(r_{\text{cap}}) - \log(\rho_0)}{\log(\nu)} \right\rfloor, & \text{otherwise.} \end{cases} \quad (10)$$

$\triangle$   
The proof can be found in the unabridged version [19]. Proposition 2 is an extension of Proposition 1 where we now have incorporated the role of  $t_f$  into the number of required sensing. Proposition 1 was derived under the case in which the pursuer had sufficient time (i.e.,  $t_f \geq \frac{\rho_0 - r_{\text{cap}}}{1-\nu}$ ) to capture.

**Remark 5:** Since Theorem 1 provides an upper bound on  $V(\rho, \tau, n)$  in  $\Omega_0$ , the  $n_{\text{req}}$  found in Proposition 2 is an upper bound for the required number of sensing in  $\Omega_0$ .  $\triangle$

We now derive the degradation in pursuer's performance when the available number of sensing is less than  $n_{\text{req}}$ . We discuss it separately for  $t_f \geq \frac{\rho_0 - r_{\text{cap}}}{1-\nu}$  and for  $t_f < \frac{\rho_0 - r_{\text{cap}}}{1-\nu}$ . When  $t_f \geq \frac{\rho_0 - r_{\text{cap}}}{1-\nu}$ , the pursuer is able to capture the evader with  $n_{\text{req}}$  number of sensing. On the other hand, with  $n < n_{\text{req}}$  we have  $V(\rho_0, t_f, n) = \phi(\frac{1-\nu}{1-\nu^{n+1}}\nu^{n+1}t_f)$  and therefore,  $\delta(n) = V(\rho_0, t_f, n) - \phi([\rho_0 - (1 - \nu)t_f]^+) = \phi(\frac{1-\nu}{1-\nu^{n+1}}\nu^{n+1}t_f)$ .

On the other hand, when  $0 < t_f < \frac{\rho_0 - r_{\text{cap}}}{1-\nu}$ , one may verify that

$$t_f \geq \frac{\rho_0 - \nu^{n+1}\rho_0}{1-\nu}, \quad \text{and } \nu^{n+1}\rho_0 > r_{\text{cap}}$$

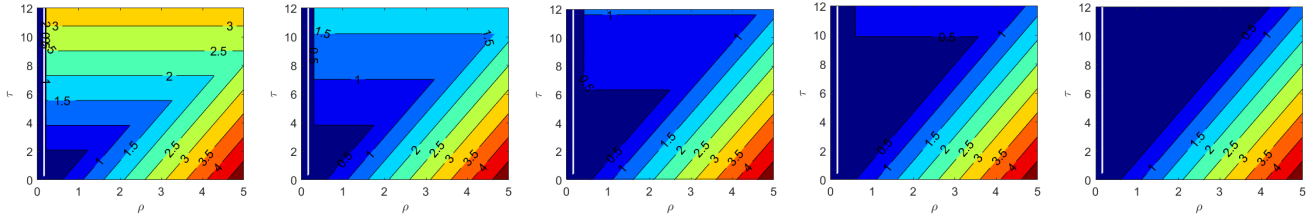


Fig. 3: Contour plots of  $V(\rho, \tau, \ell)$  from  $\ell = 1$  to  $\ell = 5$ .  $\ell$  gradually increases from the left most subfigure to the right most one.

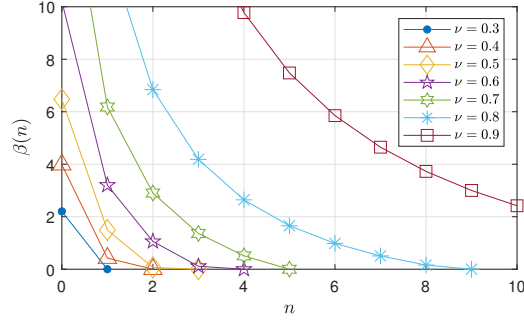


Fig. 4:  $\beta(n)$  vs.  $n$  for different values of  $\nu$ . For all the plots we pick  $r_{\text{cap}} = 0.1$ ,  $\rho_0 = 5$  and  $t_f = 0.9 \frac{\rho_0 - r_{\text{cap}}}{1 - \nu}$ .

for any integer  $n < n_{\text{req}}$ , where  $n_{\text{req}}$  is given in (10). Therefore, for all  $n < n_{\text{req}}$ , we obtain

$$\begin{aligned} \delta(n) &= \phi\left(\frac{1-\nu}{1-\nu^{n+1}}\nu^{n+1}t_f\right) - \phi(\rho_0 - (1-\nu)t_f) \\ &\geq \underbrace{\left(\frac{\nu^{n+1}}{1-\nu^{n+1}}\frac{(1-\nu)t_f}{\rho_0 - (1-\nu)t_f} - 1\right)}_{\triangleq \beta(n)} \phi(\rho_0 - (1-\nu)t_f) \end{aligned}$$

where Jensen's inequality has been used on the convex function  $\phi$ .<sup>6</sup> Notice that the coefficient  $\beta(n)$  depends on  $n$  exponentially and therefore, the degradation is exponential with  $n$ . This is demonstrated in Fig. 4.

## V. CONCLUSIONS

In this letter, we considered a sensing limited pursuit-evasion game where the pursuer is restricted to intermittent sensing. We derived the number of required sensing  $n_{\text{max}}$  in Proposition 1 to ensure capture. Next, we considered the game under explicit sensing and time/fuel budgets, i.e., the pursuer has  $t_f$  amount of time to capture the evader while using a maximum of  $n$  sensing requests. An upper bound on the value function for this game has been obtained, where the upper bound holds with an equality everywhere except in a very small region  $\Omega_0$ . Future work will focus on the case where the evader is also equipped with an intermittent sensing capability. Finding an equilibrium strategy is necessary for applications where sensing is expensive and/or undesired for both agents. In addition, tightening the upper bound of  $V$  in  $\Omega_0$  is also an open problem.

$$\dot{\mathbf{x}}_i(t) = v_i \begin{bmatrix} \cos(\theta_i(t)) \\ \sin(\theta_i(t)) \end{bmatrix}, \quad \forall i = 1, \dots, (N+1) \quad (11)$$

<sup>6</sup>For  $x \geq y > 0$ ,  $\phi(y) = \phi(\lambda x) \leq \lambda\phi(x) + (1-\lambda)\phi(0) = \lambda\phi(x)$ , where  $\lambda = \frac{y}{x}$ .

## VI. ACKNOWLEDGEMENT

The author would like to acknowledge the helpful comments and suggestions by Drs. Daigo Shishika, Alexander Von Moll, and Michael Dorothy.

## REFERENCES

- [1] R. Isaacs, *Differential games: a mathematical theory with applications to warfare and pursuit, control and optimization*. Courier Corporation, 1999.
- [2] C. Yan and T. Zhang, "Multi-robot patrol: A distributed algorithm based on expected idleness," *International Journal of Advanced Robotic Systems*, vol. 13, no. 6, 2016.
- [3] T. Alam and L. Bobadilla, "Multi-robot coverage and persistent monitoring in sensing-constrained environments," *Robotics*, vol. 9, no. 2, 2020.
- [4] E. García, A. Von Moll, D. Casbeer, and M. Pachter, "Strategies for defending a coastline against multiple attackers," in *58th Conference on Decision and Control*. IEEE, 2019, pp. 7319–7324.
- [5] W. Sun, P. Tsiotras, T. Lolla, D. N. Subramani, and P. F. Lermusiaux, "Multiple-pursuer/one-evader pursuit-evasion game in dynamic flow-fields," *Journal of guidance, control, and dynamics*, vol. 40, no. 7, pp. 1627–1637, 2017.
- [6] D. W. Oyler, P. T. Kabamba, and A. R. Girard, "Pursuit-evasion games in the presence of obstacles," *Automatica*, vol. 65, pp. 1–11, 2016.
- [7] S. Bhattacharya and S. Hutchinson, "On the existence of Nash equilibrium for a two-player pursuit-evasion game with visibility constraints," *The International Journal of Robotics Research*, vol. 29, no. 7, pp. 831–839, 2010.
- [8] S. A. Aleem, C. Nowzari, and G. J. Pappas, "Self-triggered pursuit of a single evader," in *54th Conference on Decision and Control*. IEEE, 2015, pp. 1433–1440.
- [9] D. Maity and J. S. Baras, "Optimal strategies for stochastic linear quadratic differential games with costly information," in *55th Conference on Decision and Control (CDC)*. IEEE, 2016, pp. 276–282.
- [10] Y. Huang, J. Chen, and Q. Zhu, "Defending an asset with partial information and selected observations: A differential game framework," in *60th Conference on Decision and Control*. IEEE, 2021, pp. 2366–2373.
- [11] S. D. Bopardikar, F. Bullo, and J. P. Hespanha, "Cooperative pursuit with sensing limitations," in *American Control Conference*. IEEE, 2007, pp. 5394–5399.
- [12] J. W. Durham, A. Franchi, and F. Bullo, "Distributed pursuit-evasion with limited-visibility sensors via frontier-based exploration," in *International Conference on Robotics and Automation*. IEEE, 2010, pp. 3562–3568.
- [13] D. Shishika, D. Maity, and M. Dorothy, "Partial information target defense game," in *International Conference on Robotics and Automation*. IEEE, 2021, pp. 8111–8117.
- [14] B. P. Gerkey, S. Thrun, and G. Gordon, "Visibility-based pursuit-evasion with limited field of view," *The International Journal of Robotics Research*, vol. 25, no. 4, pp. 299–315, 2006.
- [15] D. Maity and J. S. Baras, "Strategies for two-player differential games with costly information," in *13th International Workshop on Discrete Event Systems*. IEEE, 2016, pp. 211–216.
- [16] D. Maity, A. Anastasopoulos, and J. S. Baras, "Linear quadratic games with costly measurements," in *56th Conference on Decision and Control*. IEEE, 2017, pp. 6223–6228.
- [17] D. Maity and J. S. Baras, "Asymptotic policies for stochastic differential linear quadratic games with intermittent state feedback," in *25th Mediterranean Conference on Control and Automation*. IEEE, 2017, pp. 117–122.
- [18] S. A. Aleem, C. Nowzari, and G. J. Pappas, "Self-triggered pursuit of a single evader with uncertain information," *arXiv preprint arXiv:1512.06184*, 2015.
- [19] Maity, Dipankar, "Optimal Intermittent Sensing for Pursuit-Evasion Games," *arXiv preprint arXiv:2303.15598*, 2023.