Herding an Adversarial Attacker to a Safe Area for Defending Safety-Critical Infrastructure

Vishnu S. Chipade and Dimitra Panagou

***Abstract*— This paper investigates a problem of defending safety-critical infrastructure from an adversarial aerial attacker in an urban environment. A circular arc formation of defenders is formed around the attacker, and vector-field based guidance laws herd the attacker to a predefined safe area in the presence of rectangular obstacles. The defenders' formation is defined based on a novel vector field that imposes super-elliptic contours around the obstacles, to closely resemble their rectangular shape. A novel finite-time stabilizing controller is proposed to guide the defenders to their desired formation, while avoiding obstacles and inter-agent collisions. The efficacy of the approach is demonstrated via simulation results.**

I. Introduction

Advancements in and popularity of swarm technology pose increased risk to safety-critical infrastructure such as government facilities, airports, military bases etc. Even unintentional intruders such as bird flocks entering airports cause significant monetary loss [1]. Therefore, defending safety-critical infrastructure from attacks is critically important, particularly in crowded urban areas.

Approaches employing physical interception strategies [2], [3] to counteract adversarial agents (attackers) at low altitudes may not be desired due to human presence. Herding can be an indirect way of guiding the attackers to some safe area assuming risk-averse attackers who move away from the defending agents (defenders).

Various approaches have been proposed for the herding problem, including biologically-inspired robotic herding [4]–[6], and game-theoretic intruder herding [7]. In [8], [9] the authors discuss herding using a switched systems approach; the herder chases targets sequentially by switching among the targets so that certain dwell-time conditions are satisfied to guarantee stability of the resulting trajectories. However, no input constraints are considered, while the number of chased targets is limited. The authors in [10] use approximate dynamic programming to obtain control policies for the herder to chase a target agent to a goal location. However, none of these papers considered obstacles in the environment.

In [11] a cage of robots is used to herd sheep to a goal location. An RRT approach is used to find a motion plan for the robots while maintaining the cage. The herding approach discussed in [6], [12] uses a circular arc formation to influence the nonlinear dynamics of the herd based on a potential-field approach. The authors design a point-offset control to guide the herd close to a specified location. However, both papers treat robots as point masses and no inter-agent collision is addressed. In our work, we assume that the agents are of known circular footprints, and we consider the inter-agent collision avoidance. The robotic herding problem discussed in [13] uses an n-wavefront algorithm to herd a flock of birds away from an airport, where the intruders on the boundary of the flock are influenced based on the locations of the airport and a safe area. The authors also provide stability and performance guarantees in [14], as well as experimental results in [15].

In this paper, we address the problem of protecting safety-critical infrastructure (called the protected area) from an attacker, in the presence of rectangular obstacles. We utilize the idea of forming a circular arc of defenders around the attacker, in order to influence the motion trajectories of the attacker [6], [12]. In our formulation, the number of defenders depends on the obstacle geometry and the parameters of the attacker's repulsion law. In addition, we guarantee that the defenders should converge to their desired positions in finite time to ensure the defenders' timely convergence before the attacker reaches the protected area. Finite-time specifications have been considered recently in multi-agent control problems, e.g., finite-time tracking under leader-follower setting [16], finite-time containment [17]. For an overview on finite-time stability, the interested reader is referred to [18]. Here we employ the idea of finite-time stabilizing controller inspired from [19] to guide the defenders to their desired positions. We also propose a novel vector field around the rectangular obstacles to be used as reference for the formation of the defenders. The vector field has no other singular points except at the center of the safe area. As thus, it provides a safe and globally attractive motion plan for the formation of the defenders to herd the attacker to the safe area.

Compared to the authors' earlier work [20], [21] the main contributions of this paper are: (1) We consider rectangular static obstacles, which is a more realistic model for urban environments compared to circular obstacles, and design repulsive vector fields around them using super-elliptic contours. Under proper blending with attractive vector fields, we show that the resulting vector

The authors are with the Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI, USA; (vishnuc, dpanagou)@umich.edu

This work has been funded by the Center for Unmanned Aircraft Systems (C-UAS), a National Science Foundation Industry/University Cooperative Research Center (I/UCRC) under NSF Award No. 1738714 along with significant contributions from C-UAS industry members.

field is a globally safe motion plan for the formation of the defenders around the attacker. (2) We design a finite-time stabilizing, state-feedback controller to force the defenders to form and maintain their desired formation while avoiding inter-agent collisions.

The rest of the paper is structured as follows: Section II describes the mathematical modeling and problem statement. The assumptions on the attacker's control strategy are discussed in Section III. The technical developments on the herding strategy and the vector-field based motion planning for the defenders are provided in Section IV and V, respectively. Section VI presents the finite-time tracking controller that guides the defenders to their desired formation. Convergence and safety are formally proved in Section VII, while simulations are provided in Section VIII. The conclusions and our thoughts on future work are discussed in Section IX.

II. Modeling and Problem Statement

Notation: Vectors are denoted by bold letters ($\mathbf{r}$). Script letters denote sets ($\mathcal{P}$). $R^{b_1}_{b_2}(t)$ and $E^{b_1}_{ok}(t)$ (Eq. 17) are the Euclidean distance between object b_2 and b_1, and the Super-elliptic distance between b_1 and $\mathcal{O}_k$, respectively, at time t. The argument t would be omitted when clear from the context. $\sigma^{b_1}_{b_2}(\delta)$ is a blending function [20], characterized by a triplet $\delta_\sigma = (\delta^m, \bar{\delta}, \delta^u)$ and defined in Eq. 1, corresponding to the field around the object b_2 for the object b_1 which is at a distance δ from b_2.

$$\sigma^{b_1}_{b_2}(\delta) = \begin{cases} 1, & \delta^m \leq \delta \leq \bar{\delta}; \\ A^{b_1}_{b_2}\delta^3 + B^{b_1}_{b_2}\delta^2 + C^{b_1}_{b_2}\delta + D^{b_1}_{b_2}, & \bar{\delta} \leq \delta \leq \delta^u; \\ 0, & \delta^u \leq \delta; \end{cases} \tag{1}$$

The coefficients $A^{b_1}_{b_2}, B^{b_1}_{b_2}, C^{b_1}_{b_2}, D^{b_1}_{b_2}$ are chosen as: $A^{b_1}_{b_2} = \frac{2}{(\delta^u-\bar{\delta})^3}$, $B^{b_1}_{b_2} = \frac{-3(\delta^u+\bar{\delta})}{(\delta^u-\bar{\delta})^3}$, $C^{b_1}_{b_2} = \frac{6\delta^u\bar{\delta}}{(\delta^u-\bar{\delta})^3}$, $D^{b_1}_{b_2} = \frac{(\delta^u)^2(\delta^u-3\bar{\delta})}{(\delta^u-\bar{\delta})^3}$, so that (1) is a $\mathcal{C}^1$ function. The argument δ can be $R^{b_1}_{b_2}$ or $E^{b_1}_{ok}$, depending on the objects under consideration, and is omitted when clear from the context.

We consider an environment $\mathcal{W} \subseteq \mathbb{R}^2$ with N_o rectangular obstacles, a protected area $\mathcal{P} \subseteq \mathcal{W}$ defined as $\mathcal{P} = \{\mathbf{r} \in \mathbb{R}^2 \mid \|\mathbf{r} - \mathbf{r}_p\| \leq \rho_p\}$, and a safe area $\mathcal{S} \subseteq \mathcal{W}$ defined as $\mathcal{S} = \{\mathbf{r} \in \mathbb{R}^2 \mid \|\mathbf{r} - \mathbf{r}_s\| \leq \rho_s\}$, where $(\mathbf{r}_p, \rho_p)$ and $(\mathbf{r}_s, \rho_s)$ are the centers and radii of the corresponding areas, respectively. An attacker $\mathcal{A}$ and N_d defenders $\mathcal{D}_j$, $j \in I_d = \{1, 2, ..., N_d\}$, are operating in $\mathcal{W}$. $\mathcal{A}$ and $\mathcal{D}_j$ are modeled as discs of radii ρ_a and $\rho_d \leq \rho_a$, respectively, under single integrator dynamics:

$$\dot{\mathbf{r}}_a = \mathbf{v}_a, \tag{2}$$

$$\dot{\mathbf{r}}_{dj} = \mathbf{v}_{dj}, \tag{3}$$

$$\|\mathbf{v}_a\| \leq v_{max_a}, \ \|\mathbf{v}_{dj}\| \leq v_{max_{dj}}, \tag{4}$$

where $v_{max_a} < v_{max_{dj}}$, $\mathbf{r}_a = [x_a \ y_a]^T$, $\mathbf{r}_{dj} = [x_{dj} \ y_{dj}]^T$ are the position vectors of $\mathcal{A}$ and $\mathcal{D}_j$, respectively, with respect to (w.r.t.) a global inertial frame $\mathcal{F}_g(\hat{\mathbf{i}}, \hat{\mathbf{j}})$; $\mathbf{v}_a = [v_{x_a} \ v_{y_a}]^T$, $\mathbf{v}_{dj} = [v_{x_{dj}} \ v_{y_{dj}}]^T$ are their control velocity vectors, respectively, whose norms are bounded by v_{max_a} and $v_{max_{dj}}$. We also assume the following:

***Assumption** 1:* Every defender $\mathcal{D}_j$ can sense the position $\mathbf{r}_a$ of the attacker $\mathcal{A}$ once $\mathcal{A}$ lies inside a circular sensing-zone $SZ = \{\mathbf{r} \in \mathbb{R}^2 | \ \|\mathbf{r} - \mathbf{r}_p\| \leq \rho^s_d\}$ around $\mathcal{P}$, where ρ^s_d is radius of the sensing-zone.

We consider static obstacles $\mathcal{O}_k$ of rectangular shape, with their edges aligned with the axes of $\mathcal{F}_g$, defined as:

$$\mathcal{O}_k = \{\mathbf{r} \in \mathbb{R}^2 | \, |x - x_{ok}| \leq \frac{w_{ok}}{2}, |y - y_{ok}| \leq \frac{h_{ok}}{2}\}, \tag{5}$$

where $\mathbf{r}_{ok} = [x_{ok} \ y_{ok}]^T$ is the center, w_{ok} and h_{ok} are the lengths of $\mathcal{O}_k$ along $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$, $\forall k \in I_o = \{1, 2, ..., N_o\}$.

In order to defend $\mathcal{P}$, $\mathcal{D}_j$ should safely herd $\mathcal{A}$ to $\mathcal{S}$. To this end, we consider the following problem:

***Problem** 1:* Find desired positions $\mathbf{r}_{gj}$ and velocities $\dot{\mathbf{r}}_{gj}$ for the formation of $\mathcal{D}_j, \forall j \in I_d$, to guide $\mathcal{A}$ into $\mathcal{S}$ and keep it there for all future times, and control actions $\mathbf{v}_{dj}$ such that $\mathcal{D}_j$ safely converge to $\mathbf{r}_{gj}$ in finite time.

In other words, $\forall j \in I_d$ find $\mathbf{v}_{dj}$ such that 1) $\exists T^c_d, T_s > 0 : \ R^{dj}_{gj}(t) = 0, \ \forall t \geq T^c_d$ and $\|\mathbf{r}_a(t) - \mathbf{r}_s\| \leq \rho_s$, $\forall t \geq T_s$; 2) $\forall t \geq 0$ and $\forall k \in I_o$: $E^{dj}_{ok}(t) > \xi^{dj,m}_{ok}$, and $E^a_{ok}(t) > \xi^m_{ok}$, where $\xi^{dj,m}_{ok}$, ξ^m_{ok} are the minimum allowed super-elliptic distances, respectively; 3) $\forall t \geq 0$: $R^{dj}_{dl}(t) > R^{d,m}_d$, $\forall l \neq j$ and $R^a_{dj} \geq R^{a,m}_d$, where $R^{d,m}_d$, $R^{a,m}_d$ are the corresponding minimum allowed distances.

III. Attacker's Control Strategy

The attacker $\mathcal{A}$ aims to reach $\mathcal{P}$ while avoiding $\mathcal{D}_j$. The control action of $\mathcal{A}$ is modeled using a vector-field approach [21], as follows: $\mathcal{A}$ assumes that $\mathcal{D}_j$ and any static obstacles that it senses within its sensing radius ρ^s_a are of circular shape. Therefore, it moves along the direction prescribed by the vector field defined as:

$$\mathbf{F}^a = \prod_{k \in \mathcal{N}_{ao}} (1 - \sigma^a_{ok}(R^a_{ok})) \prod_{j \in \mathcal{N}_{ad}} (1 - \sigma^a_{dj}(R^a_{dj}))\mathbf{F}^a_p + \sum_{k \in \mathcal{N}_{ao}} \sigma^a_{ok}(R^a_{ok})\mathbf{F}^a_{ok} + \sum_{j \in \mathcal{N}_{ad}} \sigma^a_{dj}(R^a_{dj})\mathbf{F}^a_{dj}, \tag{6}$$

where $\mathcal{N}_{ad} = \{j \in I_d \mid \|\mathbf{r}_a - \mathbf{r}_{dj}\| \leq \rho^s_a\}$ and $\mathcal{N}_{ao} = \{k \in I_o \mid \|\mathbf{r}_a - \mathbf{r}_{ok}\| \leq \rho^s_a\}$. The vector field (6) comprises attractive term $\mathbf{F}^a_p = \frac{\mathbf{r}_p - \mathbf{r}_a}{\|\mathbf{r}_p - \mathbf{r}_a\|}$ and repulsive terms $\mathbf{F}^a_{ok} = -\frac{\mathbf{r}_{ok} - \mathbf{r}_a}{\|\mathbf{r}_{ok} - \mathbf{r}_a\|}$, $\mathbf{F}^a_{dj} = -\frac{\mathbf{r}_{dj} - \mathbf{r}_a}{\|\mathbf{r}_{dj} - \mathbf{r}_a\|}$,. $\forall k \in \mathcal{N}_{ao}$ and $\forall j \in \mathcal{N}_{ad}$, respectively. The blending functions $\sigma^a_{ok}(R^a_{ok})$, $\sigma^a_{dj}(R^a_{dj})$ are defined as in Eq. 1 with ($R^{a,m}_{ok}$, $\bar{R}^a_{ok}$, $R^{a,u}_{ok}$) and ($R^{a,m}_d$, $\bar{R}^a_d$, $R^{a,u}_d$) as the corresponding δ_σ triplets, where $R^{a,m}_{ok}$ is the minimum allowed distance between $\mathcal{A}$ and $\mathcal{O}_k$.

***Assumption** 2:* The attacker $\mathcal{A}$ has a non-zero speed when it is close to an obstacle $\mathcal{O}_k$ or a defender $\mathcal{D}_j$ i.e., $\|\mathbf{v}_a\| > 0$ when $\{\exists k \in I_o : \sigma^a_{ok} \neq 0\} \vee \{\exists j \in I_d : \sigma^a_{dj} \neq 0\}$.

***Lemma** 1:* In the absence of defenders $\mathcal{D}_j$ in $\mathcal{W}$ and if the obstacles $\mathcal{O}_k$ satisfy: $\|\mathbf{r}_{ok} - \mathbf{r}_{ol}\| \geq R^{a,u}_{ok} + R^{a,u}_{ol}$ for all $k, l \in I_o$ and $k \neq l$, then the vector field $\mathbf{F}^a : \mathbb{R}^2 \to \mathbb{R}^2$ as defined in Eq. (6) is a safe and almost globally attractive field in $\mathcal{W}^a_s = \mathcal{W} \setminus \bigcup_{k \in I_o} \{\mathbf{r} \in \mathbb{R}^2 | \, \|\mathbf{r} - \mathbf{r}_{ok}\| < R^{a,m}_{ok}\}$ to

the center $\mathbf{r}_p$ of $\mathcal{P}$, except for a set initial conditions of measure zero.

Proof: Refer to Theorem 4 in [21]. ■

The control action of the attacker $\mathcal{A}$ is given as:

$$\dot{\mathbf{r}}_a = \begin{cases} \frac{\mathbf{F}^a}{\|\mathbf{F}^a\|}\|\mathbf{v}_a\|, & \mathbf{F}^a \neq \mathbf{0}; \\ \begin{bmatrix}\cos(\theta^- + \epsilon_a)\\ \sin(\theta^- + \epsilon_a)\end{bmatrix}\|\mathbf{v}_a\|, & \mathbf{F}^a = \mathbf{0}. \end{cases} \tag{7}$$

with $\theta^- = \theta(t^-)$, where t is the time when $\mathbf{F}^a = \mathbf{0}$ and $\epsilon_a << 1$. The attacker $\mathcal{A}$ moves slightly away from the direction imposed by $\mathbf{F}^a$ when in the neighborhood of $\mathbf{F}^a = \mathbf{0}$, to get out of the deadlock situation.

IV. Herding Strategy

We now develop the strategy to herd $\mathcal{A}$ to the safe area $\mathcal{S}$. We build on the work in [6], [12], in which $\mathcal{D}_j$ are positioned along a circular arc formation around $\mathcal{A}$ to direct it towards $\mathcal{S}$ as shown in Fig. 1. By doing so the defenders generate a sufficient repulsive force on $\mathcal{A}$ such that $\mathcal{A}$ does not aim to escape even through the gaps between the defenders.

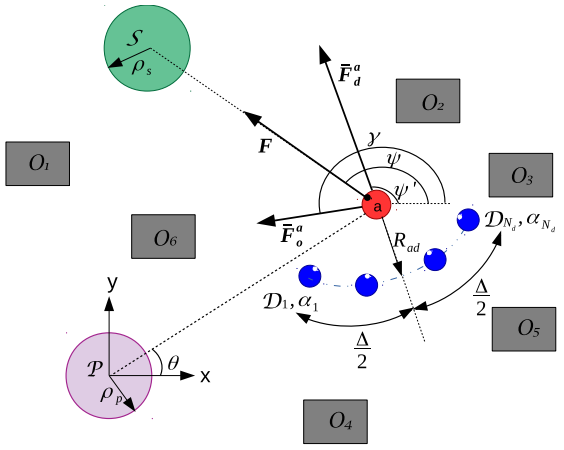

Fig. 1: Circular formation around the attacker

More specifically, we place the N_d defenders $\mathcal{D}_j$ along the circular arc of radius R_{ad} in symmetric formation, so that the angle α_j is chosen as:

$$\alpha_j = \psi' + \pi + \Delta_j, \text{ where } \Delta_j = \Delta\frac{(2j - N_d - 1)}{(2N_d - 2)}, \tag{8}$$

for some angle ψ' along which the resultant repulsive vector field from the defenders is pointed, where Δ is the total spread angle of $\mathcal{D}_j$'s around $\mathcal{A}$. For a given number N_d, a minimum separation $R_d^{a,m}$ allowed between $\mathcal{A}$ and $\mathcal{D}_j$, and a minimum separation $R_d^{d,m} > \rho_d$ allowed among $\mathcal{D}_j$, the minimum spread required is:

$$\Delta_{min} = (N_d - 1)\cos^{-1}\left(1 - \frac{(R_d^{d,m})^2}{2(R_d^{a,m})^2}\right).$$

We choose $R_{ad} = \frac{R_d^{a,m} + \bar{R}_d^a}{2} \leq \bar{R}_d^a$. With this choice of R_{ad}, we have $\sigma_{dj}^a = 1$ for all $j \in I_d$ and hence the vector field $\mathbf{F}^a$ becomes:

$$\mathbf{F}^a = \sum_{k\in\mathcal{N}_{ao}} \sigma_{ok}^a \mathbf{F}_{ok}^a + \sum_{j=1}^{N_d}\begin{bmatrix}C(\alpha_j)\\ S(\alpha_j)\end{bmatrix}, \tag{9}$$

where $C(\cdot) = \cos(\cdot)$ and $S(\cdot) = \sin(\cdot)$.

Let the repulsive fields $\mathbf{F}_{ok}^a$ due to all $\mathcal{O}_k$ in $\mathcal{A}$'s sensing radius add up to give: $\bar{\mathbf{F}}_o^a = \bar{F}_o^a\begin{bmatrix}C(\gamma)\\ S(\gamma)\end{bmatrix}$. We then obtain

$$\mathbf{F}^a = \bar{F}_o^a\begin{bmatrix}C(\gamma)\\ S(\gamma)\end{bmatrix} + \frac{S(\frac{N_d\Delta}{2N_d-2})}{S(\frac{\Delta}{2N_d-2})}\begin{bmatrix}C(\psi')\\ S(\psi')\end{bmatrix}. \tag{10}$$

Let $\bar{\Delta}_S = \frac{S(\frac{\Delta}{2N_d-2})}{S(\frac{N_d\Delta}{2N_d-2})}$, then $\mathbf{F}^a$ can be rewritten as:

$$\mathbf{F}^a = \bar{F}_o^a\begin{bmatrix}C(\gamma)\\ S(\gamma)\end{bmatrix} + \bar{\Delta}_S\begin{bmatrix}C(\psi')\\ S(\psi')\end{bmatrix}. \tag{11}$$

To guide $\mathcal{A}$ along a direction ψ of our choice, we need:

$$\frac{\mathbf{F}^a}{\|\mathbf{F}^a\|} = \begin{bmatrix}C(\psi)\\ S(\psi)\end{bmatrix}. \tag{12}$$

This is achieved by choosing ψ' such that:

$$\bar{\Delta}_S\left(S(\psi') - T(\psi)C(\psi')\right) = \bar{F}_o^a\left(T(\psi)C(\gamma) - S(\gamma)\right), \tag{13}$$

$$\dot{\psi}' = \frac{\dot{\bar{F}}_o^a(T(\psi)C(\gamma) - S(\gamma)) - \bar{F}_o^a\dot{\gamma}(T(\psi)S(\gamma) + C(\gamma))}{\bar{\Delta}_S(C(\psi') + T(\psi)S(\psi'))} + \frac{\sec^2(\psi)\dot{\psi}\left(\bar{F}_o^a C(\gamma) + \bar{\Delta}_S(C(\psi')\right)}{\bar{\Delta}_S(C(\psi') + T(\psi)S(\psi'))}, \tag{14}$$

where $T(\cdot) = \tan(\cdot)$. We choose Δ and N_d such that $\bar{\Delta}_S > \bar{F}_o^a$ ensuring $\dot{\psi}'$ is finite and bounded by $\dot{\psi}'_{max}$.

The desired (goal) position and velocity of $\mathcal{D}_j$ in the circular arc formation are then obtained as:

$$\mathbf{r}_{gj} = \mathbf{r}_a + R_{ad}\begin{bmatrix}C(\psi' + \pi + \Delta_j)\\ S(\psi' + \pi + \Delta_j)\end{bmatrix}, \tag{15}$$

$$\dot{\mathbf{r}}_{gj} = \dot{\mathbf{r}}_a + R_{ad}\dot{\psi}'\begin{bmatrix}-S(\psi' + \pi + \Delta_j)\\ C(\psi' + \pi + \Delta_j)\end{bmatrix}. \tag{16}$$

V. Guiding the Formation in Obstacle Environment

In this section, we design the reference heading ψ for the formation of the attacker $\mathcal{A}$ and the defenders $\mathcal{D}_j$ such that the obstacles $\mathcal{O}_k$ in the environment are avoided. The formation is centered at the location of the attacker $\mathbf{r}_a$ and has radius $\rho_f = R_{ad} + \rho_d$. To avoid $\mathcal{O}_k$ we define a vector field that closely follows the rectangular shape. We use the idea of superquadric isopotential contours [22], which are defined in Eq. (17) for different values of super-elliptic distance E_{ok}:

$$E_{ok} = \left|\frac{x - x_{ok}}{a_{ok}}\right|^{2n_{ok}} + \left|\frac{y - y_{ok}}{b_{ok}}\right|^{2n_{ok}} - 1, \tag{17}$$

where $a_{ok} = \frac{w_{ok}}{2}(2)^{\frac{1}{2n_{ok}}}$ and $b_{ok} = \frac{h_{ok}}{2}(2)^{\frac{1}{2n_{ok}}}$. The shape of the contour corresponding to a fixed value of E_{ok} becomes more rectangular as n_{ok} tends to ∞, i.e. by choosing large n_{ok} we can ensure, the contours closely match the shape of the rectangular building, see Fig.2.

To account for the safe distance (R_{safe}) from the obstacle boundary and the maximum size of the formation (ρ_f), the rectangular obstacles are inflated as:

$$\mathcal{O}_k^i = \left\{ \mathbf{r} \in \mathbb{R}^2 | \, |x - x_{ok}| \leq \frac{\bar{w}_{ok}}{2}, \; |y - y_{ok}| \leq \frac{\bar{h}_{ok}}{2} \right\}, \tag{18}$$

where $\bar{w}_{ok} = w_{ok} + 2(\rho_f + R_{safe})$ and $\bar{h}_{ok} = h_{ok} + 2(\rho_f + R_{safe})$ for all $k \in I_o$. We choose constant values for n_{ok}:

$$n_{ok} \quad = \frac{1}{1 - e^{-\xi_{ok}^m}}, \; \xi_{ok}^m = \frac{1}{2} \left(\left| \frac{\bar{w}_{ok}}{w_{ok}} \right|^{2n_{ok}} + \left| \frac{\bar{h}_{ok}}{h_{ok}} \right|^{2n_{ok}} \right) - 1. \tag{19}$$

The inflated rectangular obstacles can be approximated by a bounding super-ellipse, given as:

$$\mathcal{O}_k^m = \left\{ \mathbf{r} \in \mathbb{R}^2 | E_{ok} \leq \xi_{ok}^m \right\}. \tag{20}$$

The slope of the tangent to these contours at a given point f with position vector $\mathbf{r}_f \in \mathbb{R}^2$ is obtained as:

$$\tan(\bar{\beta}_{ok}^f) = -\frac{b_{ok}^{2n_{ok}} C(\beta_{ok}^f)(C(\beta_{ok}^f))^{2n_{ok}-2}}{a_{ok}^{2n_{ok}} S(\beta_{ok}^f)(S(\beta_{ok}^f))^{2n_{ok}-2}} \tag{21}$$

where β_{ok}^f is the angle made by the position vector of f with respect to the center of $\mathcal{O}_k$ and $\hat{\mathbf{i}}$, given as: $\beta_{ok}^f = \tan^{-1}\left(\frac{y_f - y_{ok}}{x_f - x_{ok}}\right)$.

A. Repulsive vector field

The repulsive vector field around $\mathcal{O}_k^m$ is defined as:

$$\mathbf{F}_{ok}^f = \begin{bmatrix} \cos(\phi_{ok}^{f,s}) \\ \sin(\phi_{ok}^{f,s}) \end{bmatrix}, \tag{22}$$

where

$$\phi_{ok}^{f,s} = \begin{cases} \bar{\beta}_{ok}^f - \Delta\bar{\beta}_{ok}^s + \frac{\Delta\beta_{ok}^{f,s}}{\pi}(\Delta\bar{\beta}_{ok}^s - \pi), & \Delta\beta_{ok}^{f,s} < \pi; \\ \bar{\beta}_{ok}^f - \Delta\bar{\beta}_{ok}^s(\frac{\Delta\beta_{ok}^{f,s} - \pi}{\pi}), & \text{otherwise}; \end{cases} \tag{23}$$

where $\Delta\beta_{ok}^{f,s} \in [0, 2\pi]$ is the angle between the vectors $\mathbf{r} - \mathbf{r}_{ok}$ and $\mathbf{r}_s - \mathbf{r}_{ok}$, defined as: $\Delta\beta_{ok}^{f,s} = \beta_{ok}^f - \beta_{ok}^s$, and, similarly, $\Delta\bar{\beta}_{ok}^s \in [0, 2\pi]$ is the angle between the tangent to the contour corresponding to the obstacle $\mathcal{O}_k$ at the center of $\mathcal{S}$ (i.e., $\bar{\beta}_{ok}^s$) and the vector $\mathbf{r}_s - \mathbf{r}_{ok}$, shown with pink text in Fig. 2, defined as: $\Delta\bar{\beta}_{ok}^s = \bar{\beta}_{ok}^s - \beta_{ok}^s$.

B. Attractive vector field

A radially converging vector field is considered around the centre of $\mathcal{S}$, defined as:

$$\mathbf{F}_s^f = \frac{\mathbf{r}_s - \mathbf{r}_f}{\|\mathbf{r}_s - \mathbf{r}_f\|}. \tag{24}$$

We set $\mathbf{F}_s^f = \mathbf{0}$ at $\mathbf{r}_f = \mathbf{0}$, so $\mathbf{F}_s^f$ is defined everywhere.

C. Combining attractive and repulsive vector fields

In order to combine the attractive and repulsive vector fields smoothly near $\mathcal{O}_k$, a blending function $\sigma_{ok}^f(E_{ok}^f)$ is defined as in Eq. 1 with $(\xi_{ok}^m, \bar{\xi}_{ok}, \xi_{ok}^u)$ as the corresponding δ_σ triplet. This essentially means that the repulsive effect of $\mathcal{O}_k$ is confined within $(\mathcal{O}_k^m)^c \cap (\mathcal{O}_k^u)$, where $(\mathcal{O}_k^m)^c$ is the compliment of $\mathcal{O}_k^m$ and $\mathcal{O}_k^u$ is given as:

$$\mathcal{O}_k^u = \left\{ \mathbf{r} \in \mathbb{R}^2 \mid E_{ok} \leq \xi_{ok}^u \right\}. \tag{25}$$

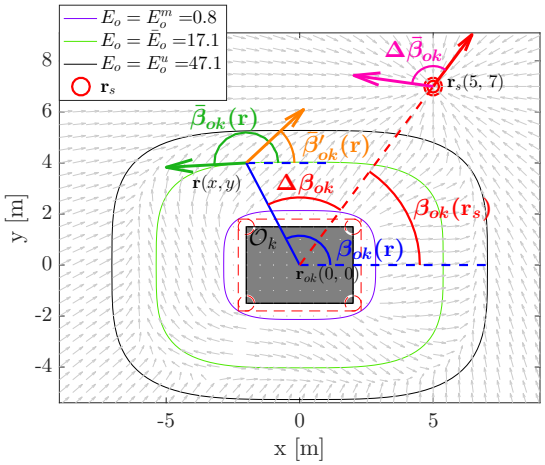

Fig. 2: Vector field definition around a rectangle

The attractive and repulsive terms are combined as:

$$\mathbf{F}^f = \prod_{k=1}^{N_o} (1 - \sigma_{ok}^f(E_{ok}^f)) \mathbf{F}_s^f + \sum_{k=1}^{N_o} \sigma_{ok}^f(E_{ok}^f) \mathbf{F}_{ok}^f. \tag{26}$$

The vector field $\mathbf{F}^f$ around $\mathcal{O}_k$ with $(w_{ok}, h_{ok}) = (4, 3)$, $\mathbf{r}_{ok} = [0, 0]^T$; and $\mathcal{S}$ with $\mathbf{r}_s = [3, 7]^T$, is shown in Fig. 2.

***Assumption** 3:* The N_o rectangular obstacles $\mathcal{O}_k$, $k \in I_o$, are spaced such that $\bigcap_{k \in I_o} \mathcal{O}_k^u = \emptyset$ and $\mathcal{S} \notin \bigcup_{k \in I_o} \mathcal{O}_k^u$, i.e., at any location, the effect of at most one obstacle is active, and there is no effect of any obstacle in the safe area $\mathcal{S}$.

Theorem 2: Let Assumption 3 holds, then the vector field $\mathbf{F}^f : \mathbb{R}^2 \to \mathbb{R}^2$ as defined in Eq. (26) is a safe, and globally attractive motion plan to the safe area $\mathcal{S}$ centered at $\mathbf{r}_s$ in $\mathcal{W}_s = \mathcal{W} \setminus \bigcup_{k \in I_o} \mathcal{O}_k^m$.

Proof: First, we show that $\mathbf{F}^f$ in $\mathcal{W}_s$ has no singular point other than $\mathbf{r}_f = \mathbf{r}_s$, to ensure that no initial conditions in $\mathcal{W}_s$ yields trajectories that end at the singular point. To verify this, we check $\left\|\mathbf{F}^f\right\|$ in 3 disjoint subsets $\mathcal{W}_{s1}$, $\mathcal{W}_{s2}$ and $\mathcal{W}_{s3}$ of $\mathcal{W}_s$ under assumption 3. Let $\sigma_{ok}^f = \sigma_{ok}^f(E_{ok}(\mathbf{r}_f))$.

- Set $\mathcal{W}_{s1} = \{\mathbf{r}_f \in \mathcal{W}_s \mid \sigma_{ok}^f = 0, \; \forall k \in I_o\}$. So, we have, $\mathbf{F}^f = \mathbf{F}_s^f$ and $\left\|\mathbf{F}^f\right\| = 1$ for all $\mathbf{r}_f \neq \mathbf{r}_s \in \mathcal{W}_{s1}$.
- Set $\mathcal{W}_{s2} = \bigcup_{k \in I_o} \{\mathbf{r}_f \in \mathcal{W}_s \mid \sigma_{ok}^f = 1\}$. So, $\exists!\; k \in I_o$ such that $\mathbf{F}^f = \mathbf{F}_{ok}^f$ and $\left\|\mathbf{F}^f\right\| = 1$ for all $\mathbf{r}_f \in \mathcal{W}_{s2}$.
- Set $\mathcal{W}_{s3} = \bigcup_{k \in I_o} \{\mathbf{r}_f \in \mathcal{W}_s \mid 0 < \sigma_{ok}^f < 1\}$. For any $\mathbf{r}_f \in \mathcal{W}_{s3}$, $\exists!\; k \in I_o$ such that $\mathbf{F}^f$ is given as: $\mathbf{F}^f = (1 - \sigma_{ok}^f)\mathbf{F}_s^f + \sigma_{ok}^f \mathbf{F}_{ok}^f$, with the norm defined as:

$$\left\|\mathbf{F}^f\right\| = \sqrt{1 + 2\sigma_{ok}^f \left(1 - \sigma_{ok}^f\right) \left(\cos(\partial\bar{\beta}_{ok}) - 1\right)}, \tag{27}$$

where $\partial\bar{\beta}_{ok}$ is the angle between the two vectors $\mathbf{F}_s^f$ and $\mathbf{F}_{ok}^f$. For $\sigma_{ok}^f \in (0, 1)$, we have $0 < \sigma_{ok}^f(1 - \sigma_{ok}^f) < \frac{1}{4}$. Now, if $\cos(\partial\bar{\beta}_{ok}) \in (-1, 1]$, then $-1 < 2\sigma_{ok}^f \left(1 - \sigma_{ok}^f\right) \left(\cos(\partial\bar{\beta}_{ok}) - 1\right) \leq 0$. This implies

$\left\|\mathbf{F}^f\right\| > 0$. Then, for $\left\|\mathbf{F}^f\right\| = 0$ it is needed that $\cos(\partial\bar{\beta}_{ok}) = -1$, i.e. $\partial\bar{\beta}_{ok} = \pm\,\pi$ and $\sigma^f_{ok} = 0.5$. Next, we show that $\partial\bar{\beta}_{ok}$ can never become $\pm\,\pi$. Let us consider $\Delta\beta^{f,s}_{ok} \geq \pi$, then $\partial\bar{\beta}_{ok}$ becomes:

$$\partial\bar{\beta}_{ok} = \tan^{-1}\left(\frac{y_s - y_f}{x_s - x_f}\right) - \bar{\beta}^f_{ok} + \Delta\bar{\beta}^s_{ok}\left(\frac{\Delta\beta^{f,s}_{ok} - \pi}{\pi}\right). \tag{28}$$

For a given $\mathbf{r}_s$, the angle $\phi^{f,s}_{ok}$ of the vector field $\mathbf{F}^f_{ok}$ only depends upon the angle β^f_{ok}. Also for a given β^f_{ok}, the angle $\tan^{-1}\left(\frac{y_s-y_f}{x_s-x_f}\right)$ becomes larger as $\|\mathbf{r}_f - \mathbf{r}_{ok}\|$ increases. So, the worst case scenario is when the contour corresponding to $E_{ok} = \xi^u_{ok}$ just touches the safe location $\mathbf{r}_s$. After parameterizing this contour in terms of β^f_{ok}, we get $x_f = x_{ok} + \bar{a}_{ok}(\cos(p_{ok}))^{\frac{1}{n_{ok}}}$ and $y_f = y_{ok} + \bar{b}_{ok}(\sin(p_{ok}))^{\frac{1}{n_{ok}}}$ where $p_{ok} = \tan^{-1}\left(\frac{(\bar{a}_{ok}\sin(\beta^f_{ok}))^{n_{ok}}}{(\bar{b}_{ok}\cos(\beta^f_{ok}))^{n_{ok}}}\right)$, $\bar{a}_{ok} = a_{ok}(1+\xi^u_{ok})^{\frac{1}{2n_{ok}}}$, and $\bar{b}_{ok} = b_{ok}(1+\xi^u_{ok})^{\frac{1}{2n_{ok}}}$. Because of the symmetry, it is sufficient to consider $\mathcal{S}$ with $\mathbf{r}_s$ in the first quadrant w.r.t $\mathbf{r}_{ok}$. The function $\partial\bar{\beta}_{ok}$ in Eq. (28) obtained in terms of β^f_{ok}, after replacing x_f, y_f and other variables, is continuous in its arguments, but it is not trivial to find the maximum value of $\partial\bar{\beta}_{ok}$. So we find the maximum and minimum $\partial\bar{\beta}_{ok}$ for $\Delta\beta^{f,s}_{ok} \in [\pi, 2\pi]$ for a given choice of β^s_{ok}. We plot the max and min values of $\partial\bar{\beta}_{ok}$ for different choices of $\beta^s_{ok} \in [0, \frac{\pi}{2}]$ in Fig. 3. It can be observed that the value never becomes π or $-\pi$. Similar analysis can be performed when $\Delta\beta^{f,s}_{ok} < \pi$ to show that $\partial\bar{\beta}_{ok}$ never becomes π or $-\pi$, which implies that $\left\|\mathbf{F}^f\right\| \neq 0$ in $\mathcal{W}_{s3}$ and hence there is no singular point in $\mathcal{W}_s$ other than $\mathbf{r}_f = \mathbf{r}_s$. Now, similar to the analysis of Lemma 1 in [21], it can be shown that the integral curves which enter the set $\mathcal{O}^u_k$ escape the boundary of $\mathcal{O}^u_k$. Also, outside $\left\{\bigcup_{k\in I_o}\mathcal{O}^u_k\right\}$, by definition the vector field $\mathbf{F}^f$ is convergent to $\mathbf{r}_s$. This proves that the vector field $\mathbf{F}^f$ is a safe and globally attractive motion plan to the safe area $\mathcal{S}$. ■

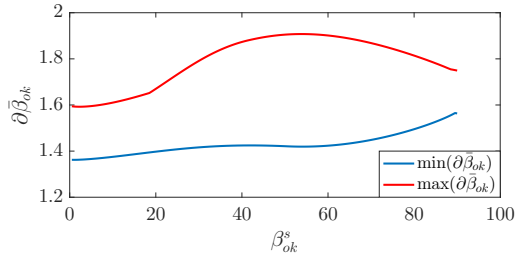

Fig. 3: Minimum and Maximum values of $\partial\bar{\beta}_{ok}$

D. Forward Invariance of Safe Area ($\mathcal{S}$)

Once $\mathcal{A}$ is in the safe area $\mathcal{S}$, which is chosen far away from the protected area $\mathcal{P}$, we gradually change the formation heading ψ to a direction which is nearly perpendicular to the direction given by the vector field $\mathbf{F}^f$, and keep applying the nearly perpendicular direction thereafter. This traps $\mathcal{A}$ inside $\mathcal{S}$, while yielding smooth controls for $\mathcal{D}_j$. Therefore ψ is designed as:

$$\psi = \begin{cases} \tan^{-1}\left(\frac{F_y}{F_x}\right), & \|\mathbf{r}_a - \mathbf{r}_s\| > \rho_s; \\ \tan^{-1}\left(\frac{F_y}{F_x}\right) + \int_{T_s}^{t} \frac{(\frac{\pi}{2}-\epsilon_s)}{\Delta T_t} d\tau, & T_s \leq t \leq T_s + \Delta T_t; \\ \tan^{-1}\left(\frac{F_y}{F_x}\right) + (\frac{\pi}{2} - \epsilon_s), & T_s + T_t \leq t; \end{cases} \tag{29}$$

where T_s is the time when $\mathcal{A}$ enters $\mathcal{S}$, ΔT_t is a transition time, and ϵ_s is a small number. Let $v_{max_d} = \min_{j\in I_d} v_{max_{dj}}$. ΔT_t and ρ_s are chosen as: $\Delta T_t \geq \frac{\pi}{2}\frac{(v_{max_d} - v_{max_a})}{R_{ad}}$ and $\rho_s \geq v_{max_a}\Delta T_t + \frac{v_{max_a}R_{ad}}{v_{max_d} - v_{max_a}}$.

VI. Reference Trajectory Tracking

To ensure that the attacker $\mathcal{A}$ is herded along the ψ direction, the defenders $\mathcal{D}_j$'s have to track $\mathbf{r}_{gj}$ (Eq. (15) and $\dot{\mathbf{r}}_{gj}$ (Eq. (16)), while avoiding collisions with $\mathcal{O}_k$'s and $\mathcal{D}_l$'s. To ensure that $\mathbf{r}_{gj}$ does not intersect any $\mathcal{O}_k$ for all $j \in I_d$, the following assumptions are made.

***Assumption** 4:* $R^{a,m}_{ok} \geq \sqrt{\bar{w}^2_{ok} + \bar{h}^2_{ok}}$, $\forall k \in I_o$.

***Assumption** 5:* $R_{safe} \geq 2\rho_d$.

A. Repulsive vector field from the obstacles for $\mathcal{D}_j$

The repulsive vector field at $\mathbf{r}_{dj} = [x_{dj}\ y_{dj}]^T$ around the obstacle $\mathcal{O}_k$ is defined as:

$$\mathbf{F}^{dj}_{ok} = \begin{bmatrix}\cos(\phi^{dj,gj}_{ok}) \\ \sin(\phi^{dj,gj}_{ok})\end{bmatrix}, \tag{30}$$

where $\phi^{dj,gj}_{ok}$ is defined similar to $\phi^{f,s}_{ok}$ as in Eq. 23.

B. Repulsive vector fields from $\mathcal{D}_l$ for $\mathcal{D}_j$

A radially repulsive vector field is considered around each one of the defenders $\mathcal{D}_l$, defined as:

$$\mathbf{F}^{dj}_{dl} = -\frac{\mathbf{r}_{dl} - \mathbf{r}_{dj}}{\|\mathbf{r}_{dl} - \mathbf{r}_{dj}\|}, \quad \forall l \neq j \in I_d. \tag{31}$$

C. Attractive vector field to the desired position of $\mathcal{D}_j$

A radially attractive vector field is considered around the desired position of the defender $\mathbf{r}_{gj}$, defined as:

$$\mathbf{F}^{dj}_{gj} = \frac{\mathbf{r}_{gj} - \mathbf{r}_{dj}}{\|\mathbf{r}_{gj} - \mathbf{r}_{dj}\|}. \tag{32}$$

D. Resultant vector field for $\mathcal{D}_j$

Blending functions $\sigma^{dj}_{ok}(E^{dj}_{ok})$ and $\sigma^{dj}_{dl}(R^{dj}_{dl})$ are defined respectively using Eq. 1 with $(\xi^{dj,m}_{ok}, \bar{\xi}^{dj}_{ok}, \xi^{dj,u}_{ok})$ and $(R^{d,m}_d, \bar{R}^d_d, R^{d,u}_d)$ as the corresponding δ_σ triplets, so that the resulting vector field is continuously differentiable. The combined vector field for $\mathcal{D}_j$ is then:

$$\begin{aligned}\mathbf{F}^{dj} = &\prod_{k=1}^{N_o}(1 - \sigma^{dj}_{ok}(E^{dj}_{ok})) \prod_{l\neq j\in I_d}(1 - \sigma^{dj}_{dl}(E^{dj}_{ok}))\mathbf{F}^{dj}_{gj} \\ &+ \sum_{k=1}^{N_o}\sigma^{dj}_{ok}(E^{dj}_{ok})\mathbf{F}^{dj}_{ok} + \sum_{l\neq j\in I_d}\sigma^{dj}_{dl}(R^{dj}_{dl})\mathbf{F}^{dj}_{dl}.\end{aligned} \tag{33}$$

Define $\mathbf{e}_{dj} = \mathbf{r}_{dj} - \mathbf{r}_{gj}$, and $\mathbf{v}^0_{dj}$ as:

$$\mathbf{v}^0_{dj} = \begin{cases} k_{d0} \tanh\left(\|\mathbf{e}_{dj}\|\right) \frac{\mathbf{F}^{dj}}{\|\mathbf{F}^{dj}\|}, & \|\mathbf{e}_{dj}\| > e^t_{dj}; \\ k_{d1} \|\mathbf{e}_{dj}\|^{k_{d2}} \frac{\mathbf{F}^{dj}}{\|\mathbf{F}^{dj}\|}, & \|\mathbf{e}_{dj}\| \leq e^t_{dj}; \end{cases} \tag{34}$$

where $0 < k_{d2} < 1$ and $k_{d0} = v_{max_{dj}} - v_{max_a} - R_{ad}\dot{\psi}'_{max}$ so that $\mathbf{v}^0_{dj}$ is bounded. To ensure continuity of $\mathbf{v}^0_{dj}$ at $\|\mathbf{e}_{dj}\| = e^t_{dj}$, the parameters k_{d1} and e^t_{dj}, are chosen as: $k_{d1} = k_{d0} \frac{\tanh(e^t_{dj})}{(e^t_{dj})^{k_{d2}}}$, $(1 - \tanh^2(e^t_{dj})) = k_{d2} \frac{\tanh(e^t_{dj})}{e^t_{dj}}$.

The control action for $\mathcal{D}_j$ is then designed as:

$$\mathbf{v}_{dj} = \begin{cases} \dot{\mathbf{r}}_{gj} + \mathbf{v}^0_{dj}, & \sigma^{dj}_{ok} = \sigma^{dj}_{dl} = 0, \forall k, l \neq j; \\ \mathbf{v}^0_{dj}, & \text{otherwise.} \end{cases} \tag{35}$$

Under $\mathbf{v}_{dj}$, $\mathcal{D}_j$ tracks the reference trajectory $\mathbf{r}_{gj}$ when it is not in conflict (i.e. corresponding σ is 0) with any $\mathcal{O}_k$ or $\mathcal{D}_l, l \neq j$, otherwise it only gets closer to $\mathbf{r}_{gj}$ by moving along $\mathbf{F}^{dj}$ until it resolves the conflict.

VII. Convergence and Safety Analysis

Theorem 3: Under the control action in Eq. (35), the system trajectories of Eq. (3) converge to piecewise continuous trajectories $\mathbf{r}_{gj}(t)$ in finite time, for all initial conditions except a set of initial conditions of measure zero, $\bar{\mathcal{D}}_0$.

Proof: With Assumptions 4 and 5, we have that $\mathbf{r}_{gj}$ never intersect the obstacles. Thus one can choose some $\xi^{dj,m}_{ok} > 0$ to generate $\mathbf{F}^{dj}_{ok}$ for safe motion of $\mathcal{D}_j$. Each defender $\mathcal{D}_j$ moves under the vector field $\mathbf{F}^{dj}$ given by Eq. (33). By following the analysis in Theorem 5 of [21], $\mathcal{D}_j$'s resolve the conflicts with other $\mathcal{D}_l$'s for all $\mathbf{r}_{dj}(0) \notin \bar{\mathcal{D}}_0$. With finite number of $\mathcal{O}$'s and $\mathcal{D}$'s, $\mathcal{D}_j$ resolves conflicts in finite time T^0_{dj}.

Now, when $\mathcal{D}_j$ is not in conflict with $\mathcal{D}_l/\mathcal{O}_k$, we have:

$$\dot{\mathbf{e}}_{dj} = \begin{cases} -k_{d0} \tanh\left(\|\mathbf{e}_{dj}\|\right) \frac{\mathbf{e}^{dj}}{\|\mathbf{e}^{dj}\|}, & \|\mathbf{e}_{dj}\| > e^t_{dj}; \\ -k_{d1} \|\mathbf{e}_{dj}\|^{k_{d2}-1} \mathbf{e}^{dj}, & \|\mathbf{e}_{dj}\| \leq e^t_{dj}. \end{cases} \tag{36}$$

Case 1: When $\left\|\mathbf{e}_{dj}(T^0_{dj})\right\| = e_{dj0} \leq e^t_{dj}$, the dynamics read: $\dot{\mathbf{e}}_{dj} = -k_{d1} \|\mathbf{e}_{dj}\|^{k_{d2}-1} \mathbf{e}^{dj}$. Then, from Lemma 1 in [19], the origin $\mathbf{e}_{dj} = \mathbf{0}$ is a finite-time stable equilibrium of the system (36). Denote the time required for convergence when $e_{dj0} = e^t_{dj}$ as $T^f_{dj} < \infty$.

Case 2: When $e_{dj0} > e^t_{dj}$, the dynamics of the system read: $\dot{\mathbf{e}}_{dj} = -k_{d0} \tanh\left(\|\mathbf{e}_{dj}\|\right) \frac{\mathbf{e}^{dj}}{\|\mathbf{e}^{dj}\|}$. From Lemma 4 in [23] there exist a finite time $T^t_{dj} = \frac{-e_{dj0}}{\tanh(e_{dj0})} \log\left(\frac{(e^t_{dj})^2}{2V(e_{dj0})}\right)$ such that $\|\mathbf{e}_{dj}(t)\| \leq e^t_{dj}$ for all $t \geq T^t_{dj}$. This combined with case 1 implies that for $e_{dj0} > e^t_{dj}$, $\exists T^c_{dj} = T^t_{dj} + T^f_{dj} + T^0_{dj} < \infty$ s.t $\|\mathbf{e}_{dj}(t)\| = 0$ for all $t \geq T^c_{dj}$. Hence $\mathbf{r}_{dj}$ converges to $\mathbf{r}_{gj}$ in finite time. ∎

To check whether $\mathcal{D}_j$ converges to r_{gj} before $\mathcal{A}$ reaches $\mathcal{P}$, we find a conservative lower bound on the time required by $\mathcal{A}$ to reach $\mathcal{P}$ in the absence of obstacles and defenders as: $T^c_a = \frac{\|\mathbf{r}_a(0) - \mathbf{r}_p\|}{v_{max_a}}$. If $\forall j \in I_d$, $T^c_{dj} < T^c_a$, then $\mathcal{D}_j$ can herd $\mathcal{A}$ to $\mathcal{S}$. Finding T^0_{dj} and a tight bound on T^c_a is an involved task and is a part of our future work.

Theorem 4: Let $\forall j \in I_d$, $\mathbf{r}_{dj}(0) \notin \bar{\mathcal{D}}_0$, $\mathbf{r}_{gj}$ is given in Eq. (15), with $\dot{\mathbf{r}}_{gj}$ as in Eq. (16), where ψ is given by Eq. (29), $\mathcal{D}_j$ uses the control action $\mathbf{v}_{dj}$ (Eq. (35)), and $\|\mathbf{r}_a(t) - \mathbf{r}_p\| > \rho_p$ at $t = T^c_d = \max_{j \in I_d} T^c_{dj}$, then $\exists T_s > 0$ s.t. $\|\mathbf{r}_a(t) - \mathbf{r}_s\| < \rho_s, \forall t \geq T_s$, i.e., $\mathcal{A}$ is taken into $\mathcal{S}$ and is trapped there for all future times.

Proof: From Theorem 3, when $\mathbf{r}_{dj}(0) \notin \bar{\mathcal{D}}_0$, $\mathcal{D}_j$ converge to $\mathbf{r}_{gj}(t)$ in finite time T^c_d for all $j \in I_d$. If $\mathcal{A}$ has not reached inside $\mathcal{P}$ at T^c_d, using Eqs. (7), (12) and (29), the dynamics of $\mathcal{A}$ reads:

$$\dot{\mathbf{r}}_a = \frac{\mathbf{F}^a}{\|\mathbf{F}^a\|} \|\mathbf{v}_a\| = \frac{\mathbf{F}}{\|\mathbf{F}\|} \|\mathbf{v}_a\|. \tag{37}$$

From Theorem 2, we have that the vector field $\mathbf{F}$ is non-vanishing at all points except $\mathbf{r} = \mathbf{r}_s$, it is safe w.r.t the obstacles and is globally convergent to $\mathbf{r}_s$. In other words, there exist an integral curve from $\mathbf{r}_a(T^c_d)$ which reaches $\mathbf{r} = \mathbf{r}_s$. Also from the Assumption 2, we have that the speed $\|\mathbf{v}_a\| \neq 0$ for all $t \geq T^c_d$. This implies that for any $\mathbf{r}_a(T^c_d) \in \mathcal{W}^a_s \subseteq \mathcal{W}_s$, $\exists T_s \geq T^c_d$ s.t. $\|r_a - r_s\| = \rho_s$, i.e, $\mathbf{r}_a$ enters $\mathcal{S}$. Once inside $\mathcal{S}$, the reference direction always points in the interior of the circle $\|\mathbf{r} - \mathbf{r}_s\| = \|\mathbf{r}_a(t) - \mathbf{r}_s\|$ and the choice of ρ_s for a given v_{max_a}, as discussed in Section V-D, ensures that $\mathcal{A}$ never escapes $\mathcal{S}$. ∎

VIII. Simulation Results

The proposed herding strategy is evaluated via simulation results. We consider $N_o = 6$ rectangular obstacles in $\mathcal{W} = \mathbb{R}^2$ with properties as: $(x_{ok}, y_{ok}, w_{ok}, h_{ok}) = \{(10, 23, 2, 3), (-6, 18, 3, 4), (11, 5, 2, 2), (15, 43, 3, 3), (-2, 45, 3, 4), (12, 60, 4, 3)\}$. $\mathcal{P}$ is centered at $[0, 0]^T$ with $\rho_p = 2$ (pink dotted circle in Fig. 4) and $\mathcal{S}$ is centered at $\mathbf{r}_s = [-5, \ 60]^T$ with $\rho_s = 5$ (green dotted circle in Fig. 4). $\mathcal{A}$, $\mathcal{D}_1$, $\mathcal{D}_2$ and $\mathcal{D}_3$ are initially located at $[20, 48]^T$, $[-10, 16]^T$, $[6, 2]^T$, $[-5, -1]^T$, respectively. The other parameters are chosen as: $\rho_a = \rho_d = 0.1$, $\rho^c_a = 10$, $R_{safe} = 2\rho_d$, $R_{ad} = 0.55$, $R^{a,m}_d = 0.3$, $R^{a,u}_d = 0.65$.

Fig. 4 shows the trajectories of $\mathcal{A}$ (in red) and all $\mathcal{D}_j$ (in blue). It can be observed that all $\mathcal{D}_j$ safely guide $\mathcal{A}$ into the safe area $\mathcal{S}$ and keep it there forever. $\mathcal{A}$ is assumed to move at its maximum speed v_{max_a} in the simulation. To assess the safety of the agents from $\mathcal{O}_k$'s, we define critical relative distances:

$$E^{ao}_{rel} = \max_{k \in I_o} \frac{\xi^m_{ok}}{E^a_{ok}}, \quad E^{do}_{rel} = \max_{k \in I_o} \max_{j \in I_d} \frac{\xi^{dj,m}_{ok}}{E^{dj}_{ok}},$$

which have to be less than 1 for safety. Similarly, we define critical relative distances as:

$$R^{dd}_{rel} = \max_{j \neq l \in I_d} \frac{R^{d,m}_d}{R^{dj}_{dl}}, \quad R^{ad}_{rel} = \max_{j \in I_d} \frac{R^{a,m}_d}{R^a_{dj}},$$

which have to be less than 1 for no inter-agent collisions. As observed in Fig. 5, all the critical relative distances are less than 1 and hence there are no collisions. Figure 5 also shows that the defenders' speeds are bounded and less than their maximum speeds.

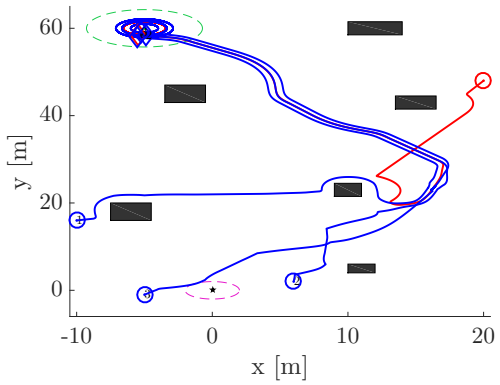

Fig. 4: Trajectories of the agents during the herding

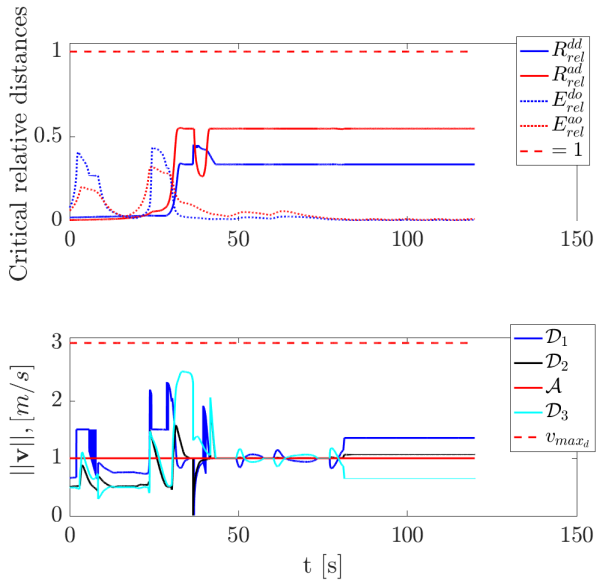

Fig. 5: Critical relative distances and control speeds

IX. Conclusions and Future Work

In this paper, we proposed novel vector-field based guidance laws for defenders herding an attacker to a safe area. Finite-time controllers are designed for the defenders to converge to a formation around the attacker, while safety and convergence are formally proved. In future, we plan to investigate the case of multiple attackers under double integrator dynamics.

ACKNOWLEDGMENT

We acknowledge the reviewers for their valuable comments.

References

[1] J. R. Allan, "The costs of bird strikes and bird strike prevention," *Human conflicts with wildlife: economic considerations*, p. 18, 2000.

[2] M. Chen, Z. Zhou, and C. J. Tomlin, "Multiplayer reach-avoid games via pairwise outcomes," *IEEE Transactions on Automatic Control*, vol. 62, no. 3, pp. 1451–1457, 2017.

[3] M. Coon and D. Panagou, "Control strategies for multiplayer target-attacker-defender differential games with double integrator dynamics," in *Decision and Control (CDC), 2017 IEEE 56th Annual Conference on*. IEEE, 2017, pp. 1496–1502.

[4] M. A. Haque, A. R. Rahmani, and M. B. Egerstedt, "Biologically inspired confinement of multi-robot systems," *International Journal of Bio-Inspired Computation*, vol. 3, no. 4, pp. 213–224, 2011.

[5] M. Evered, P. Burling, and M. Trotter, "An investigation of predator response in robotic herding of sheep," *International Proceedings of Chemical, Biological and Environmental Engineering*, vol. 63, pp. 49–54, 2014.

[6] A. Pierson and M. Schwager, "Bio-inspired non-cooperative multi-robot herding," in *IEEE International Conference on Robotics and Automation*, 2015, pp. 1843–1849.

[7] S. Nardi, F. Mazzitelli, and L. Pallottino, "A game theoretic robotic team coordination protocol for intruder herding," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 4124–4131, 2018.

[8] R. A. Licitra, Z. D. Hutcheson, E. A. Doucette, and W. E. Dixon, "Single agent herding of n-agents: A switched systems approach," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 14 374–14 379, 2017.

[9] R. A. Licitra, Z. I. Bell, E. A. Doucette, and W. E. Dixon, "Single agent indirect herding of multiple targets: A switched adaptive control approach," *IEEE Control Systems Letters*, vol. 2, no. 1, pp. 127–132, 2018.

[10] P. Deptula, Z. I. Bell, F. M. Zegers, R. A. Licitra, and W. E. Dixon, "Single agent indirect herding via approximate dynamic programming," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 7136–7141.

[11] A. Varava, K. Hang, D. Kragic, and F. T. Pokorny, "Herding by caging: a topological approach towards guiding moving agents via mobile robots," in *Proceedings of Robotics: Science and Systems*, 2017.

[12] A. Pierson and M. Schwager, "Controlling noncooperative herds with robotic herders," *IEEE Transactions on Robotics*, vol. 34, no. 2, pp. 517–525, 2018.

[13] S. Gade, A. A. Paranjape, and S.-J. Chung, "Herding a flock of birds approaching an airport using an unmanned aerial vehicle," in *AIAA Guidance, Navigation, and Control Conference*, 2015, p. 1540.

[14] ——, "Robotic herding using wavefront algorithm: Performance and stability," in *AIAA Guidance, Navigation, and Control Conference*, 2016, p. 1378.

[15] A. A. Paranjape, S.-J. Chung, K. Kim, and D. H. Shim, "Robotic herding of a flock of birds using an unmanned aerial vehicle," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 901–915, 2018.

[16] G. Wen, Y. Yu, Z. Peng, and A. Rahmani, "Distributed finite-time consensus tracking for nonlinear multi-agent systems with a time-varying reference state," *International Journal of Systems Science*, vol. 47, no. 8, pp. 1856–1867, 2016.

[17] X. Wang, S. Li, and P. Shi, "Distributed finite-time containment control for double-integrator multiagent systems," *IEEE Trans. on Cybernetics*, vol. 44, no. 9, pp. 1518–1528, 2014.

[18] S. P. Bhat and D. S. Bernstein, "Finite-time stability of continuous autonomous systems," *SIAM Journal on Control and Optimization*, vol. 38, no. 3, pp. 751–766, 2000.

[19] K. Garg and D. Panagou, "New results on finite-time stability: Geometric conditions and finite-time controllers," in *2018 American Control Conference*, June 2018, pp. 442–447.

[20] D. Panagou, "Motion planning and collision avoidance using navigation vector fields," in *Proc. of the Int. Conf. on Robotics and Automation*. IEEE, 2014, pp. 2513–2518.

[21] ——, "A distributed feedback motion planning protocol for multiple unicycle agents of different classes," *IEEE Tran. on Automatic Control*, vol. 62, no. 3, pp. 1178–1193, 2017.

[22] R. Volpe and P. Khosla, "Manipulator control with superquadric artificial potential functions: Theory and experiments," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 20, no. 6, pp. 1423–1436, 1990.

[23] K. Garg and D. Panagou, "Finite-time estimation and control for multi-aircraft systems under wind and dynamic obstacles," *Journal of Guidance, Control, and Dynamics*, pp. 1–17, 2019.