Automatica 94 (2018) 26–34

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

# Optimal evasive strategies for multiple interacting agents with motion constraints[☆]

William Lewis Scott [a], Naomi Ehrich Leonard [b,*]

[a] Aerospace Engineering, University of Maryland, College Park, MD 20742, USA
[b] Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544, USA

ARTICLE INFO

*Article history:*
Received 14 November 2016
Received in revised form 17 February 2018
Accepted 23 March 2018

*Keywords:*
Autonomous mobile robots
Decentralized control
Differential games
Minimum-time control
Multi-agent systems
Optimal trajectory
Pursuit-evasion

ABSTRACT

We derive and analyze optimal control strategies for a system of pursuit and evasion with a single speed-limited pursuer, and multiple heterogeneous evaders with limits on speed, angular turning rate, and lateral acceleration. The goal of the pursuer is to capture a single evader in the minimum time possible, and the goal of each evader is to avoid capture if possible, or else delay capture for as long as possible. Optimal strategies are derived for the one-on-one differential game, and these form the basis of strategies for the multiple-evader system. We propose a pursuer strategy of *optimal target selection* which leads to capture in bounded time. For evaders, we prove how any evader not initially targeted can avoid capture. We also consider optimal strategies for agents with radius-limited sensing capabilities, proving conditions for evader capture avoidance through a local strategy of *risk reduction*. We show how evaders aggregate in response to a pursuer, much like animals behave in the wild.



## 1. Introduction

We consider a system with a single pursuer and multiple heterogeneous evader agents moving on the plane. The goal of the pursuer is to capture any *single* evader in the minimum time possible. The goal of each evader is to first of all avoid capture, and if that is not achievable to delay capture for as long as possible. The pursuer has limited speed, and the evaders have limits on speed, angular turning rate, and lateral acceleration. To analyze this system we apply the framework of *differential games* introduced by Isaacs (1965) and used to study pursuit and evasion, e.g., Başar and Oldser (1999), Elliott and Kalton (1972) and Pachter (1987). We examine optimal strategies for the one-on-one pursuit-evasion differential game under these motion constraints, and use those as building blocks for strategies in the system of multiple evaders.

In the multiple-evader system, we propose a strategy for the pursuer of *optimal target selection*, where the target is the evader that could be captured in minimum time in a one-on-one setting. For evaders, in the case of all-to-all sensing, we prove that any evader not currently the target can always choose from a set of *reactive evasion* control inputs in order to avoid capture. The currently targeted evader must use the optimal evasive strategy from the one-on-one game to delay its capture for as long as possible. We also consider the case in which the pursuer and evaders have radius-limited sensing and propose a local strategy of *risk reduction*. We prove that any evader that is not the target can avoid capture using the risk reduction strategy. The case with no constraint on turning rate is addressed in Scott and Leonard (2014).

Predator avoidance has long been considered a key factor in animal aggregation. The "selfish herd" of Hamilton (1971) is a seminal model of identical evaders on the plane. Hamilton showed that a group benefit is not necessary to explain aggregation; rather, a self-interested individual in a population stays close to others to reduce its own chance of being caught. We also consider a group of self-interested evaders, but we use continuous-time dynamics and heterogeneous evaders. We are also motivated by the problem of designing dynamics for group formation in engineered multi-agent systems. Our decentralized control law for a collective response to a moving threat accounts for practical motion constraints and provides a control mechanism for spontaneous aggregation.

Hamilton's model has been extended to include evolutionary dynamics and formation of large groups (Wood & Ackland, 2007). Numerical studies have examined properties of group motion in multiple-evader systems where biologically inspired strategies are chosen a priori: on the plane (Lee, Pak, & Chon, 2006), in discrete

☆ This research was supported in part by National Science Foundation grant ECCS-1135724 and Office of Naval Research grant N00014-14-1-0635. The material in this paper was partially presented at the 53rd IEEE Conference on Decision and Control, December 15–17, 2014, Los Angeles, CA, USA. This paper was recommended for publication in revised form by Associate Editor Michael M. Zavlanos under the direction of Editor Christos G. Cassandras.

* Corresponding author.
*E-mail addresses:* wlscott@umd.edu (W.L. Scott), naomi@princeton.edu (N.E. Leonard).

https://doi.org/10.1016/j.automatica.2018.04.008
0005-1098/

space (Vabø & Nøttestad, 1997), in three dimensions (Vabø & Skaret, 2008), with multiple pursuers (Angelani, 2012), and based on observations of crabs and shorebirds (Viscido, Miller, & Wethey, 2001). Non-spatially explicit game theoretic models of multiple-evader systems have been posed for both homogeneous evaders (Cressman & Garay, 2011), and heterogeneous evaders (Eshel, Sansone, & Shaked, 2006).

Cooperative evader strategies have been studied as differential games in systems where all evaders are captured in succession (Liu, Zhou, Tomlin, & Hedrick, 2013a, b), and in systems where evaders have defensive capabilities (Fuchs & Khargonekar, 2011). The problem of choosing the order in which to capture multiple evaders requires numerical optimization or approximate solutions for efficient computation. Because every evader will be captured and strategies are cooperative, evaders are driven apart rather than into aggregations, fundamentally different from the problem posed by Hamilton and studied in the present paper. Multiple pursuers against a single evader have been studied in many contexts (Bakolas, 2013; Bakolas & Tsiotras, 2010; Chen, Zha, Peng, & Gu, 2016; Selvakumar & Bakolas, 2016; Zhou et al., 2016), using tools such as generalized Voronoi diagrams. Oyler, Kabamba, and Girard (2016) analyzed a pursuit-evasion game on the plane in the presence of obstacles, using a time-to-reach partition to determine if two evaders can rendezvous with each other before capture by a single pursuer. Bakolas and Tsiotras (2012) considered a multi-pursuer system where the active pursuer is whoever can capture the single evader in minimum time. This is dual to the optimal target selection problem for the multiple-evader system.

We use a time-to-capture metric based on the solution to a one-on-one differential game to partition the plane into evader domains of danger in a multiple-evader system. The partition is useful both in the analysis of the pursuer strategy of optimal target selection, where the pursuer chooses its target based on which domain of danger it is in, and in each evader's strategy of reactive evasion to keep the pursuer from entering its own domain of danger.

Our analysis considers an "omnidirectional" pursuer with limited speed seeking to capture any single evader from a group of heterogeneous and non-cooperating evaders with limits on speed, angular turning rate, and lateral acceleration, motivated from legged locomotion. A study of the kinematics of horses during polo games (Tan & Wilson, 2011) indicates that grip strength and limb force limits constrain the maximum lateral acceleration during a turn. In the evader motion model, the limit on lateral acceleration serves to create a tradeoff between speed and maneuverability, as the agent cannot make a sharp turn while maintaining maximum speed.

Several recent papers examine differential games featuring steered agents with turning constraints, such as a differential-drive agent vs. an omnidirectional agent, each acting as pursuer and evader (Ruiz & Murrieta-Cid, 2016), and an omnidirectional pursuer vs. a car-like evader (Exarchos, Tsiotras, & Pachter, 2015). These types of dynamics have also been studied in minimum-time problems for a single agent, for the fixed-speed *Reeds-Shepp* vehicle (Sussmann & Tang, 1991), for a differential-drive vehicle with limited wheel speed (Balkcom & Mason, 2002), and in our own work on an agent with limited speed, turning rate, and lateral acceleration (Scott & Leonard, 2018). A biologically inspired analysis of pursuit and evasion with acceleration constraints by Howland (1974) suggests that a more agile but slower evader can escape from a fast pursuer with limited lateral acceleration by veering to the side at the last moment. Studies of evasive behavior in different animal species are reviewed in Domenici and Ruxton (2015).

Our major contributions are threefold. First, we prove an optimal strategy for a pursuer that seeks to capture, in minimum time, any single evader among multiple heterogeneous evaders moving in the plane with limits on speed, angular turning rate, and lateral acceleration. The strategy relies on the optimal solution to the corresponding one-on-one differential game, which is new relative to the literature due to the constraints imposed on the evader's motion. For the multiple-evader system, the pursuer will target one evader at a time but will switch to target another evader if and when the pursuer estimates that the other evader can be caught in the shortest time remaining. Second, we prove a reactive evasive strategy for each non-targeted evader that keeps it from becoming the target. The evasion strategies do not require cooperation and each non-targeted evader can stay close to the group and conserve energy while still avoiding capture. Third, we generalize our results to the system in which the pursuer and each evader has a limited sensing region. In this case, before using reactive evasion, each non-targeted evader responds with a risk reduction phase to decrease its chances of becoming the target. We show how each non-targeted evader will move closer to another with a lower speed limit, thus providing a distributed control mechanism for aggregation.

We define the problem and system equations in Section 2. In Section 3 we derive optimal trajectories and an evader feedback-control law for the one-on-one differential game with motion constraints. In Section 4 we prove the optimal strategies for the multiple-evader system. We introduce limits on sensing radius in Section 5 and examine evader risk reduction. We conclude in Section 6.

## 2. Problem statement and equations of motion

We consider a system on the plane with a single pursuer agent $P$ and a heterogeneous group of $n$ evader agents $E_i$. The pursuer $P$ is modeled as an agent that can freely move in any direction with maximum speed $\bar{v}_p$, position $\mathbf{r}_p(t) \in \mathbb{R}^2$ at time $t$, and velocity control input $\mathbf{u}_p(t) = (v_{x_p}(t), v_{y_p}(t))^T \in \mathbb{R}^2$ with $\|\mathbf{u}_p(t)\|_2 \le \bar{v}_p$ for all $t$. Evaders are modeled as steered agents with inputs of speed $v_i(t) \in \mathbb{R}$ and turning rate $\omega_i(t) \in \mathbb{R}$, written as $\mathbf{u}_i(t) = (v_i(t), \omega_i(t))^T$. An evader's state at time $t$ is its position $\mathbf{r}_i(t) \in \mathbb{R}^2$ and its heading angle $\theta_i(t) \in \mathbb{S}^1$.

For each evader agent $E_i$, we impose the following motion constraints:

- *Forward motion:* Speed must satisfy $v_i(t) \ge 0$ for all time $t$, such that the agent never moves in reverse.
- *Limited speed:* Let $\bar{v}_i > 0$ be the maximum speed. The speed control must satisfy $v_i(t) \le \bar{v}_i$ for all time $t$.
- *Limited turning rate:* Let $\bar{\omega}_i > 0$ be the maximum turning rate. The turning control must satisfy $|\omega_i(t)| \le \bar{\omega}_i$ for all time $t$.
- *Limited lateral acceleration:* Let $\mu_i$ represent the maximum lateral acceleration (turning traction limit). The inputs $v_i(t)$ and $\omega_i(t)$ must satisfy $|v_i(t)\omega_i(t)| \le \mu_i$ for all time $t$. We further impose the condition that $\mu_i < \bar{v}_i\bar{\omega}_i$ so that the lateral acceleration constraint is active on part of the boundary of the control domain.

We define the evader admissible control region $\Omega_{e_i} = \{\mathbf{u} = (v, \omega) \in \mathbb{R}^2 \,|\, 0 \le v \le \bar{v}_i, |\omega| \le \bar{\omega}_i, |v\omega| \le \mu_i < \bar{v}_i\bar{\omega}_i\}$ and the pursuer admissible control region $\Omega_p = \{\mathbf{u}_p \in \mathbb{R}^2 \,|\, \|\mathbf{u}_p\|_2 \le \bar{v}_p\}$. Admissible controls $\mathcal{U}_i$ for evader $E_i$ are bounded Lebesgue measurable functions from $\mathbb{R}_+$ to $\Omega_{e_i}$ and $\mathcal{U}_p$ for the pursuer from $\mathbb{R}_+$ to $\Omega_p$.

The system equations of motion are

$$\begin{aligned} \dot{\mathbf{r}}_p &= \mathbf{u}_p, && \mathbf{u}_p \in \mathcal{U}_p \\ \dot{\mathbf{r}}_i &= \begin{pmatrix} v_i \cos\theta_i \\ v_i \sin\theta_i \end{pmatrix}, && (v_i, \omega_i) \in \mathcal{U}_{e_i} \\ \dot{\theta}_i &= \omega_i, \quad \text{for } i = 1, 2, \ldots, n. && \end{aligned} \tag{1}$$

We define the pursuer's goal to be the capture of a single evader in minimum time. We define the goal of each evader to avoid capture altogether, or if that is not achievable then to delay capture for as long as possible.

## 3. Pursuit and evasion with two agents

Consider the system above with a single evader denoted by the subscript $e$, where the pursuer has a higher maximum speed $\bar{v}_p > \bar{v}_e$. In this case the pursuer can always guarantee eventual capture. To determine optimal strategies for each agent, we formulate the problem as a *differential game* with the time to capture, $T_{cap}$, as the payoff. The two agents' goals are directly opposed: the pursuer aims to minimize the time-to-capture while the evader aims to maximize time-to-capture. We define capture as the condition that the distance between the agents is equal to a capture radius, $l \geq 0$. The standard form of the pursuit-evasion differential game is described by the payoff functional with unity integral cost $L = 1$: $T_{cap}[\mathbf{q}(0), \mathbf{u}_p(\cdot), \mathbf{u}_e(\cdot)] = \int_0^T 1dt$, subject to the dynamics $\dot{\mathbf{q}}(t) = \mathbf{f}(\mathbf{q}, \mathbf{u}_p, \mathbf{u}_e)$, where $\mathbf{q} = (x_p, y_p, x_e, y_e, \theta_e)^T \in \mathbb{R}^4 \times \mathbb{S}$ and $\mathbf{f}(\mathbf{q}, \mathbf{u}_p, \mathbf{u}_e) = (v_{x_p}, v_{y_p}, v_e \cos\theta_e, v_e \sin\theta_e, \omega_e)^T$, and terminal condition $\psi(T) = 0$ for $\psi(t) = (x_p(t)-x_e(t))^2+(y_p(t)-y_e(t))^2-l^2$. We seek a pair of optimal controls $\mathbf{u}_p^*, \mathbf{u}_e^*$ such that any deviation from either player will result in a worse payoff: $T_{cap}[\mathbf{q}(0), \mathbf{u}_p^*, \mathbf{u}_e] \leq T_{cap}[\mathbf{q}(0), \mathbf{u}_p^*, \mathbf{u}_e^*] \leq T_{cap}[\mathbf{q}(0), \mathbf{u}_p, \mathbf{u}_e^*]$, $\forall \mathbf{u}_p \in \mathcal{U}_p, \mathbf{u}_e \in \mathcal{U}_e$.

The solution procedure is as follows. Candidate open-loop optimal trajectories are constructed in reverse time starting at the terminal surface, and singular surfaces where the trajectories are not uniquely defined are analyzed. From these trajectories we derive optimal control in state-feedback form. In the present system we find that for appropriate values of capture radius $l$, discussed in Section 3.5, unique optimal strategies exist everywhere except for a single dispersal surface corresponding to states where the evader faces directly towards the pursuer, and can decide whether to rotate left or right.

Define the adjoint vector as a row vector,

$$\boldsymbol{\lambda} = \frac{\partial}{\partial \mathbf{q}} T_{cap} = (\lambda_{x_p}, \lambda_{y_p}, \lambda_{x_e}, \lambda_{y_e}, \lambda_{\theta_e}). \tag{2}$$

The control Hamiltonian for the game has the form

$$\begin{aligned} H(\boldsymbol{\lambda}, \mathbf{q}, \mathbf{u}_p, \mathbf{u}_e) &= \boldsymbol{\lambda} \cdot \mathbf{f}(\mathbf{q}, \mathbf{u}_p, \mathbf{u}_e) + 1 \\ &= \lambda_{x_p} v_{x_p} + \lambda_{y_p} v_{y_p} + \lambda_{x_e} v_e \cos\theta_e \\ &\quad + \lambda_{y_e} v_e \sin\theta_e + \lambda_{\theta_e}\omega_e + 1. \end{aligned} \tag{3}$$

Then optimal control inputs $\mathbf{u}_p^*, \mathbf{u}_e^*$ are specified by the "Main Equation" of Isaacs (1965):

$$H(\boldsymbol{\lambda}, \mathbf{q}, \mathbf{u}_p^*, \mathbf{u}_e^*) = \min_{\mathbf{u}_p \in \Omega_p} \max_{\mathbf{u}_e \in \Omega_e} H(\boldsymbol{\lambda}, \mathbf{q}, \mathbf{u}_p, \mathbf{u}_e) = 0. \tag{4}$$

Note that the min and max operators commute, since the terms involving $\mathbf{u}_p$ and $\mathbf{u}_e$ are separate. The adjoint equations of motion are $\dot{\boldsymbol{\lambda}} = -\frac{\partial H}{\partial \mathbf{q}}$, and so $\dot{\lambda}_{x_p} = \dot{\lambda}_{y_p} = \dot{\lambda}_{x_e} = \dot{\lambda}_{y_e} = 0$ and $\dot{\lambda}_{\theta_e} = \lambda_{x_e} v_e \sin\theta_e - \lambda_{x_e} v_e \cos\theta_e$. As in Balkcom and Mason (2002), since $\dot{\lambda}_{\theta_e} = \lambda_{x_e}\dot{y}_e - \lambda_{y_e}\dot{x}_e$, $\lambda_{\theta_e}$ can be directly integrated (with constant of integration $\rho$):

$$\lambda_{\theta_e} = \lambda_{x_e} y_e - \lambda_{y_e} x_e - \rho. \tag{5}$$

### 3.1. Terminal conditions

We start by defining a parameterization $\mathbf{h}(\mathbf{s})$ of the capture surface in terms of parameter vector $\mathbf{s} \in \mathbb{R}^4$: $\mathbf{q}(T) = \mathbf{h}(\mathbf{s}) = (s_1 + l\cos s_4, s_2 + l \sin s_4, s_1, s_2, s_3)^T$. The value of the game, $T_{cap}$, does not depend directly on the terminal state, so all its partial derivatives with respect to the terminal surface are zero: $0 = \frac{\partial T_{cap}}{\partial s_j} = \boldsymbol{\lambda}(T) \cdot \frac{\partial \mathbf{h}}{\partial s_j}$, for $j = 1, 2, 3, 4$, providing four terminal conditions,

$$\begin{aligned} 0 &= \lambda_{x_p}(T) + \lambda_{x_e}(T) \\ 0 &= \lambda_{y_p}(T) + \lambda_{y_e}(T) \\ 0 &= \lambda_{\theta_e}(T) \\ 0 &= l(\lambda_{y_p}(T)\cos s_4 - \lambda_{x_p}(T)\sin s_4). \end{aligned} \tag{6}$$

We define the normalized adjoint values at $t = T$ as

$$\begin{aligned} \hat{\lambda}_{x_p} &= \lambda_{x_p}/\lambda_0 = \cos(s_4) \\ \hat{\lambda}_{y_p} &= \lambda_{y_p}/\lambda_0 = \sin(s_4) \\ \hat{\lambda}_{x_e} &= \lambda_{x_e}/\lambda_0 = -\cos(s_4) \\ \hat{\lambda}_{y_e} &= \lambda_{y_e}/\lambda_0 = -\sin(s_4), \end{aligned} \tag{7}$$

where $\lambda_0 = \sqrt{\lambda_{x_p}^2(T) + \lambda_{y_p}^2(T) + \lambda_{x_e}^2(T) + \lambda_{y_e}^2(T)} = \|\lambda(T)\|_2$. Normalizing by $-\lambda_0$ gives trajectories that reach the terminal surface from within the capture region.

To determine $\lambda_0$, we use Eq. (3) for $H$ at the terminal time and the optimal control inputs for each agent:

$$\begin{aligned} &H(\boldsymbol{\lambda}(T), \mathbf{h}(\mathbf{s}), \mathbf{u}_p^*, \mathbf{u}_e^*) = 0 \\ &= \lambda_0(\hat{\lambda}_{x_p} v_x^* + \hat{\lambda}_{y_p} v_y^* + \hat{\lambda}_{x_e} v_e^* \cos s_3 + \hat{\lambda}_{y_e} v_e^* \sin s_3) + 1. \end{aligned}$$

The pursuer's optimal control to minimize $H$ is given by $v_x^*(T) = -\bar{v}_p \hat{\lambda}_{x_p}$ and $v_y^*(T) = -\bar{v}_p \hat{\lambda}_{y_p}$. The evader's optimal control depends on the location of the terminal state on the terminal surface:

$$v_e^*(T) = \begin{cases} \bar{v}_e, & \cos(s_4 - s_3) < 0, \\ 0, & \cos(s_4 - s_3) \geq 0. \end{cases}$$

Thus we can solve for $\lambda_0$ through substitution:

$$\lambda_0 = \begin{cases} (\bar{v}_e \cos(s_4 - s_3) + \bar{v}_p)^{-1}, & \cos(s_4 - s_3) < 0, \\ \bar{v}_p^{-1}, & \cos(s_4 - s_3) \geq 0. \end{cases} \tag{8}$$

The "usable part" of the capture surface is the set of points $\mathbf{s}$ where the pursuer can force the state to penetrate the surface: $\min_{\mathbf{u}_p \in \Omega_p} \max_{\mathbf{u}_e \in \Omega_e} \boldsymbol{\lambda}(\mathbf{s}) \cdot \mathbf{f}(\mathbf{h}(\mathbf{s}), \mathbf{u}_p, \mathbf{u}_e) < 0$, with $\boldsymbol{\lambda}(\mathbf{s})$ defined at $t = T$ by (7), (8), and $\lambda_{\theta_e} = 0$ from (6). Since we assume a faster pursuer with $\bar{v}_p > \bar{v}_e$, the entire capture surface comprises the usable part.

### 3.2. Optimal trajectories for pursuit and evasion

Given the state at the time of capture, the trajectories for each agent can be integrated backwards in time based on the optimal controls corresponding to the adjoint vectors as computed above. From (3) the terms in $H$ corresponding to evader and pursuer control inputs are independent. Thus we can apply Pontryagin's minimum principle for each agent independently and derive optimal trajectories given the proper boundary conditions. Since $\dot{\lambda}_{x_p} = \dot{\lambda}_{y_p} = 0$, the adjoint entries for the pursuer remain constant throughout, and the pursuer will use a constant control input. To minimize $H$, the optimal path of the pursuer is to follow a straight line at full speed $\bar{v}_p$ in the direction opposite its associated adjoint vector $(\lambda_{x_p}, \lambda_{y_p})^T$.

Suppose, without loss of generality, that the evader is captured while at the origin, with its heading along the positive $x$-axis. The state at capture is given by $\mathbf{q}(T) = (l\cos s, l \sin s, 0, 0, 0)^T$, for some capture angle $s \in [-\pi, \pi]$. The value of $\boldsymbol{\lambda}$ at capture can be computed from (6) and (8) through substitution with $s = s_4 - s_3$. For trajectories ending at this capture point, the pursuer's optimal control is a constant vector $\mathbf{u}_p^* = (-\bar{v}_p \cos s, -\bar{v}_p \sin s)^T$. Integrating backwards in time with $\tau = T - t$, the pursuer's trajectory is a straight line going away from the capture point at the origin: $\mathbf{r}_p = ((l + \bar{v}_p\tau)\cos s, (l + \bar{v}_p\tau)\sin s)^T$. The evader control input $\mathbf{u}_e^* = (v_e^*, \omega_e^*)$ that maximizes $H$ depends not only on the terminal conditions but also on the current state of the system. To determine the extremal evader control, we define three state-dependent switching functions:

$$\begin{aligned}\phi_1(\mathbf{q}) &= -\cos(\theta_e - s)\\ \phi_2(\mathbf{q}) &= x_e \sin s - y_e \cos s\\ \phi_3(\mathbf{q}) &= \bar{\omega}_e|\phi_2| - \bar{v}_e\phi_1.\end{aligned} \tag{9}$$

Let $\mathrm{sgn}(z)$ be the standard sign function for $z \in \mathbb{R}$. On time intervals for which the switching functions are nonzero, the corresponding extremal controls are called generic. These fall into three categories:

- *Rotation:* When $\phi_1 < 0$, the agent rotates in place: $v_e^* = 0$ and $\omega_e^* = \bar{\omega}_e\,\mathrm{sgn}(\phi_2)$.
- *Slow turn:* When $\phi_1 > 0$ and $\phi_3 > 0$, the agent moves forward with low speed while turning at the maximum rate: $v_e^* = \mu_e/\bar{\omega}_e$ and $\omega_e^* = \bar{\omega}_e\,\mathrm{sgn}(\phi_2)$. The agent moves on a circular arc with radius $R_s = \mu_e/\bar{\omega}_e^2$.
- *Fast turn:* When $\phi_1 > 0$ and $\phi_3 < 0$, the agent moves forward at maximum speed while turning at a lower rate: $v_e^* = \bar{v}_e$ and $\omega_e^* = \mathrm{sgn}(\phi_2)\mu_e/\bar{v}_e$. The agent moves on a circular arc with radius $R_f = \bar{v}_e^2/\mu_e$.

In the case that $s = \pm\pi$, we have that $\phi_1 = 1$ and $\phi_2 = 0$ at capture. The evader control input that maximizes $H$ is not unique: any input with $v_e = \bar{v}_e$ and $\omega_e \in [-\mu_e/\bar{v}_e, \mu_e/\bar{v}_e]$ is maximizing. Integrating backwards in time, any control with $\omega_e \neq 0$ will immediately cause the evader to leave the $\phi_2 = 0$ switching surface, bringing it into a generic fast turn segment. However, should the evader use a control of $v_e = \bar{v}_e$ and $\omega_e = 0$ for an extended interval, it will remain on the switching surface. This *forward motion* evader control is optimal only when the pursuer is directly behind the evader, such that both agents are moving in the direction of the baseline vector from the pursuer to the evader. Once started, forward motion continues until capture.

### 3.3. Evader control switching times

To calculate switching times for the evader optimal control, we integrate the equations of motion backwards in time from capture at time $t = T$. Let $T_{cap}(t) = T - t$ be the time remaining until capture along a specific retro-time trajectory with pursuer and evader agent each using its optimal control. So $T_{cap} = T_{cap}(0)$. The evader's optimal trajectory will have some combination of rotation, slow turn, fast turn, and forward segments based upon the value of $s$ at capture. For $\sin s > 0$ at capture, the evader will use right turning controls ($\omega_e \leq 0$), and for $\sin s < 0$ at capture, left turning controls. For $\cos s \geq 0$ at capture, the evader's trajectory consists only of rotation. The pursuer moves in a straight line directly towards the evader. For $-1 < \cos s < 0$ at capture, the evader's trajectory ends in a fast turn. Proceeding backwards in time from capture by integrating the equations of motion (1) using fast turn input for the evader, the state crosses the $\phi_3 = 0$ switching surface at the time given by $T_{cap} = \tau_f(s) = \theta_f(s)\mu_e/\bar{v}_e$, where $\theta_f(s) = |s| - \cos^{-1}\left(\frac{\bar{v}_e\bar{\omega}_e}{\bar{v}_e\bar{\omega}_e+\mu_e}\cos s\right)$. At that time, the evader control switches to a slow turn in the same direction, for duration of $\tau_s(s) = \theta_s(s)/\bar{\omega}_e$, where $\theta_s(s) = |s| - \pi/2 - \theta_f(s)$. It is at that point that the evader crosses the $\phi_1 = 0$ switching surface, and switches to rotation control in the same direction.

When $\cos s = -1$ at capture, the evader's trajectory can end in a forward segment or a fast turn in either direction. This corresponds to a family of optimal trajectories with varying time spent in the forward segment, $\tau_d$, and either right or left turns. Going backwards in time from capture, at time $T_{cap} = \tau_d$ the evader switches to a fast turn, up to the maximum duration $\bar{\tau}_f = \tau_f(\pi)$, at which point it switches to a slow turn of up to the maximum duration $\bar{\tau}_s = \tau_s(\pi)$, then it switches to rotation.

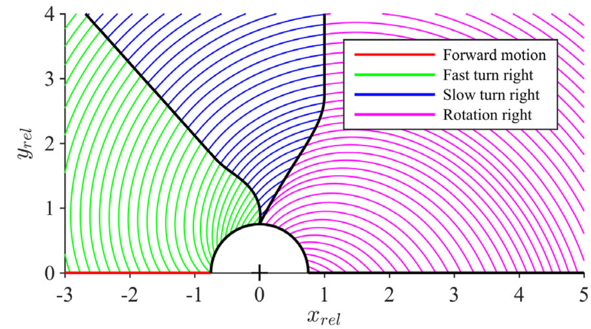


**Fig. 1.** Optimal trajectories in reduced coordinates (10). Color denotes optimal evader control input. Trajectories in the lower half plane ($y_{rel} < 0$, not shown) are mirrored about the $x_{rel}$ axis, with evader using left turns. Here $\bar{v} = 1$, $\bar{\omega} = 1$, $\mu = 0.5$, and $v_p = 1.5$. The capture radius is set to the minimum value of $l = l_c$ from (12). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

### 3.4. Evader state-feedback control law

We define a set of *reduced coordinates* $(x_{rel}, y_{rel})^T \in \mathbb{R}^2$ describing the position of the pursuer relative to a frame fixed on the evader, with the positive $x$-axis in the direction of the evader's heading:

$$\begin{pmatrix} x_{rel} \\ y_{rel} \end{pmatrix} = \begin{pmatrix} \cos\theta_e & \sin\theta_e \\ -\sin\theta_e & \cos\theta_e \end{pmatrix} \begin{pmatrix} x_p - x_e \\ y_p - y_e \end{pmatrix}. \tag{10}$$

Consider the set of all optimal trajectories $\mathbf{q}(t)$ described above in Sections 3.2 and 3.3, along with their associated optimal controls $\mathbf{u}_p^*(t)$ and $\mathbf{u}_e^*(t)$ for all $s \in [-\pi, \pi]$ and $\tau_d \geq 0$. Any ordered pair of capture angle and time-to-capture $(s, T_{cap})$ (along with forward duration $\tau_d$ for $s = \pm\pi$) corresponds to a point in the reduced space, $(x_{rel}, y_{rel})$, along with an associated optimal evader input and optimal pursuer input (transformed to the reduced coordinate frame). From these optimal trajectories together in reduced space, we can derive the optimal controls in state-feedback form: this is the inverse mapping from a point in the reduced space, pursuer position relative to the evader, to the associated optimal control.

Since the evader has a discrete set of possible optimal controls, its state-feedback control law maps a region of the reduced space to each input. Fig. 1 shows optimal trajectories in the reduced coordinates, with color denoting the optimal evader control input at each point. The locus of points in trajectories where evader control switching occurs form the switching surfaces for the evader, shown as black lines in the figure. The negative $x$-axis in the reduced coordinates corresponds to evader forward motion with the pursuer directly behind it. This line constitutes what is known as a "universal surface" in differential game theory—it is an optimal trajectory that behaves such that other optimal trajectories run into it and flow along it. Conversely the positive $x$-axis is a "dispersal surface" in that games starting with the evader facing directly towards the pursuer can proceed either with left or right turning evader trajectories, leading to capture in an equal amount of time under optimal play by both agents. The set of optimal trajectories in reduced coordinates also provides mappings from relative pursuer position to the optimal pursuer heading, and to the optimal time-to-capture $T_{cap}$, shown in Fig. 2. This mapping of the state to the value of the game plays a large role in the development of optimal strategies in the multiple-evader system presented in Section 4.

**Remark 1.** The evader optimal trajectories are equivalent to the optimal trajectories for the problem of reaching a desired point on the plane in minimum time with the same motion constraints. This also holds for the pursuer, since the fastest way to reach a point is to

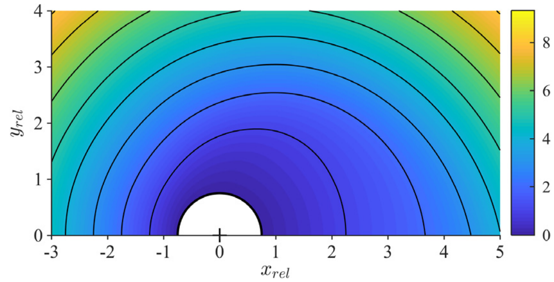


**Fig. 2.** Time-to-capture surface in reduced coordinates (10) with same parameters as Fig. 1. The $T_{cap}$ surface is mirrored about the $x_{rel}$-axis: $T_{cap}(x_{rel}, -y_{rel}) = T_{cap}(x_{rel}, y_{rel})$.

go directly towards it at full speed. The one-on-one game of pursuit and evasion is in a sense equivalent to the problem of choosing the point on the plane resulting in capture which maximizes the evader's time-to-reach and using the optimal trajectory to reach that point. In Scott and Leonard (2018) we derived minimum-time trajectories to reach a point on the plane for a steered agent with the speed, turning rate, and lateral acceleration constraints of the evader considered here.

### 3.5. Condition on capture radius

The inverse mapping described in Section 3.4 is well defined if every state in the reduced coordinates maps to a unique time-to-capture and optimal control input for each agent. This will hold if the capture radius $l$ is positive and sufficiently large. For small $l$, we find that reverse-time trajectories in the reduced space intersect not only at the dispersal surface on the positive $x$-axis, but also near the point where the evader switching surfaces meet, suggesting the presence of additional singular surfaces and possible discontinuity in the $T_{cap}$ surface. This is avoided by choosing $l$ large enough such that the $s = \pi/2$ evader rotate-only trajectory does not intersect the slow turn–rotation switching surface, except where they meet at the capture surface.

We can solve for the condition on $l$ by requiring that the slope of the slow turn–rotation switching surface is greater than that of the $s = \pi/2$ rotate-only trajectory in the upper half plane of reduced coordinates where they extend from the capture surface at $(x_{rel}, y_{rel})^T = (0, l)^T$:

$$\left.\frac{dy}{dx}\right|^{\text{switching surface}}_{x_{rel}=0, y_{rel}=l} \geq \left.\frac{dy}{dx}\right|^{\text{rotate-only traj.}}_{x_{rel}=0, y_{rel}=l}$$
$$\Leftrightarrow \quad \frac{2\bar{v}_p\bar{\omega}_e}{\mu + \bar{v}_e\bar{\omega}_e} \geq \frac{\bar{v}_p}{l\bar{\omega}_e}. \tag{11}$$

The minimum capture radius $l_c$ satisfying (11) is

$$l_c = \frac{\mu + \bar{v}_e\bar{\omega}_e}{2\bar{\omega}_e^2} = \frac{1}{2}\left(R_s + \frac{\bar{v}_e}{\bar{\omega}_e}\right). \tag{12}$$

## 4. Pursuit and evasion with multiple evaders

We now consider a system with a single fast pursuer and multiple evaders with heterogeneous control constraints. The pursuer's goal is to capture any single evader in the minimum time possible. The goal of each evader is to avoid capture by the pursuer, or else delay capture for as long as possible. Define $T_{cap,i}(t)$ as the time-to-capture evader $E_i$ for the pursuer with evader $E_i$ using their optimal control strategies from the one-on-one differential game: $T_{cap,i}(t) = T_{cap}[\mathbf{q}(t), \mathbf{u}_p^*(\cdot), \mathbf{u}_i^*(\cdot)] = \int_t^T 1dt$, under dynamics $\dot{\mathbf{q}}(t) = \mathbf{f}(\mathbf{q}, \mathbf{u}_p, \mathbf{u}_i)$ and terminal condition $\psi(T) = 0$ for pursuer and evader $E_i$. Let $T_{cap,i} = T_{cap,i}(0)$. We show that to minimize time the pursuer should choose a target evader with the lowest bounded time-to-capture ($\min_i T_{cap,i}$) and use the optimal control strategy from the one-on-one differential game. The targeted evader must use its optimal strategy from the one-on-one game, but the others can use what we call a "reactive evasion" strategy that will guarantee that they do not become the target of the pursuer. If one of these other evaders fails to use a reactive evasion strategy such that its time to capture drops below the time to capture for the current targeted evader, the pursuer should target this other evader and use the optimal control strategy from the one-on-one differential game against it. We define the optimal pursuer strategy to avoid high frequency switching of targets, which would be impractical.

### 4.1. Multiple-evader optimal pursuit

Let $\epsilon > 0$ be a small time period. We define the *multiple-evader optimal pursuit strategy* for dynamics (1) as $\mathbf{u}_p(t)$, $t \in [0, T]$, that minimizes $T$, the time to capture a *single evader*, under the constraint that switches between targeted evaders are not more frequent than $1/\epsilon$.

We define $E^*(t) = \{E_i \mid T_{cap,i}(t) = \min_j T_{cap,j}(t)\}$ to be the set of evaders with the minimum time-to-capture based on the agents' states at time $t$.

**Theorem 1.** *Consider the multiple-evader system with dynamics (1) and $\bar{v}_p > \bar{v}_i$, $\forall i$. Assume $E^*(t)$ is nonempty for all $0 \leq t \leq T$. Let $\epsilon > 0$ be a small time period. Define* recursive pursuer strategy $S_p(t_s)$ *for $0 \leq t_s < T$:*

(1) *Choose at random an evader $E_t \in E^*(t_s)$ and utilize the optimal strategy for the one-on-one game against evader $E_t$ until whichever of the following comes first: (a) $E_t$ is captured or (b) the first time $t_f > t_s + \epsilon$ such that $E_t \notin E^*(t_f)$;*
(2) *If outcome (a), the time is $T$ and the game is over; if outcome (b), let $t_s = t_f$ and return to (1), thereby switching to a new target.*

*The multiple-evader optimal pursuit strategy is $S_p(t_s)$ at $t_s = 0$. This strategy guarantees capture in time $T \leq \min_i T_{cap,i}$ and a finite number of target switches. In case of at least one target switch, $T < \min_i T_{cap,i}$.*

**Proof.** Under optimal pursuer play in a one-on-one game against an evader $E_i$, the pursuer is guaranteed to capture the evader in time $T \leq T_{cap,i}$. Conversely any evader using optimal evasive control is guaranteed to avoid capture up to time $T_{cap,i}$. Thus, the best possible strategy for the pursuer to catch an evader in minimum time is to use the optimal strategy for the one-on-one system against the evader with the minimum value of $T_{cap,i}$. In the case that multiple evaders share the same minimum time-to-capture, the target can be chosen at random from $E^*$ to achieve capture in the minimum time. If some evader does not use its optimal strategy and allows its time-to-capture to become the minimum, the pursuer will switch to targeting that evader at that time and further reduce the time it takes to capture a single evader. By pursing a target for at least $\epsilon$ units of time before switching targets, a bound on the number of switches of target is guaranteed. $\square$

### 4.2. Evader domain of danger and reactive evasion

For any given pursuer location, the optimal target is the evader $E_i$ that has the lowest value of $T_{cap,i}$. We can thus partition the plane into "domains of danger," denoted $\mathcal{D}_i$, corresponding to the points on the plane where a given evader $E_i$ has the minimum value of $T_{cap,i}$: $\mathcal{D}_i = \left\{\mathbf{r}_p \in \mathbb{R}^2 \,\middle|\, i = \operatorname{argmin}_j T_{cap,j}\right\}$. Fig. 3 shows an example of the domains of danger before and after a pursuit

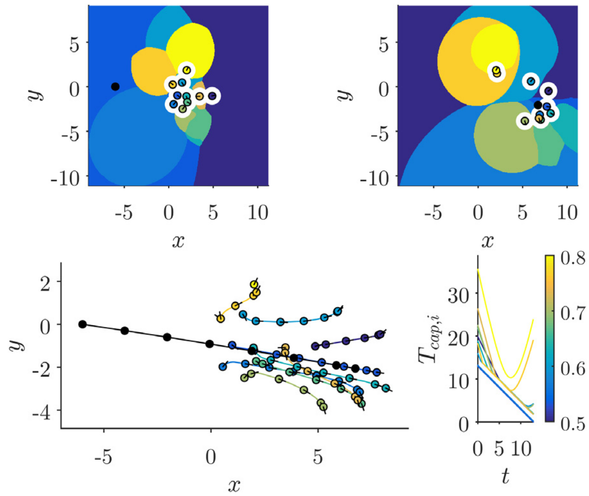


**Fig. 3.** Simulation of reactive evasion with 10 evaders. Evaders and their domains of danger are colored by evader speed; see colorbar at bottom right. $v_p = 1$. For all evaders, $l = 1$, $\omega_i = 1$, and $\mu_i = 0.5$, with $v_i$ evenly spaced in the range [0.5, 0.8]. Top left: evader domains of danger at initial time. Top right: evader domains of danger at capture. Pursuer is denoted by filled black circle. Bottom left: agent trajectories. Snapshots show evader headings every 2 s. Bottom right: evader $T_{cap}$ over the course of the chase. Note that the evader with the lowest $T_{cap}$ has its value decreasing at a constant rate of −1 second per second until capture.

for a group of evaders of different maximum speeds and common turning rate and lateral acceleration constraints. From Theorem 1, at any given time during pursuit, the optimal target of the pursuer is the evader in whose domain of danger the pursuer currently resides. If an evader can keep the pursuer from entering its domain of danger, then it will not become a target and can avoid capture. We show this is possible for any evader not initially targeted, under the assumption that each agent senses the relative position and orientation and knows the motion constraints of the others.

**Remark 2.** For evader agents with no constraints on turning rate, the domain of danger partition takes on the form of a "multiplicatively-weighted Voronoi diagram," which is described in Section 5. In the present more general system, the $T_{cap}$ surface is not radially symmetric, so the domain of danger depends not only on distance to other evaders, but on their relative headings as well.

**Theorem 2.** *Consider the multiple-evader system with a pursuer using the multiple-evader optimal pursuit strategy of Theorem 1. For each evader $E_i$ not targeted by the pursuer at time $t_0 \in [0, T]$, there exists a set of controls $(v_i, \omega_i) \in \mathcal{U}_i$ that guarantee the evader will not become the target at any future time $t > t_0$. If all non-targeted evaders employ such a strategy for the duration of a chase, the pursuer and targeted evader strategies are equivalent to those from the one-on-one game, with capture occurring at $T = \min_i T_{cap,i}$ for the initially targeted evader.*

**Proof.** Let $E_g$ be the initially targeted evader so that $T_{cap,g} = \min_i T_{cap,i}$. For every other evader $E_i$, $T_{cap,i} \geq T_{cap,g}$. $E_i$ will remain untargeted until $t$ if $T_{cap,i}(\tau) \geq T_{cap,g}(\tau)$, $\tau \in [0, t]$. If there exists $(v_i, \omega_i) \in \mathcal{U}_i$ for $E_i$ such that $\frac{d}{dt}T_{cap,i} \geq \frac{d}{dt}T_{cap,g}$ for all time $t \geq 0$, then by continuity $T_{cap,i}(t) \geq T_{cap,g}(t)$ for all $t \geq 0$.

Recall for the one-on-one game that $\boldsymbol{\lambda} = \frac{\partial}{\partial \mathbf{q}} T_{cap}$. So

$$\dot{T}_{cap} = \frac{d}{dt} T_{cap} = \frac{\partial}{\partial \mathbf{q}} T_{cap} \cdot \dot{\mathbf{q}} = \boldsymbol{\lambda} \cdot \mathbf{f}(\mathbf{q}, \mathbf{u}_p, \mathbf{u}_e). \tag{13}$$

In the one-on-one game under optimal control by both agents, $\dot{T}_{cap} = -1$ at all times, by (4). Thus $\dot{T}_{cap,g} = -1$. Any deviation from the optimal input by the pursuer (e.g., targeting some other evader) will necessarily lead to a slower decrease of $T_{cap,g}$, so a non-targeted evader $E_i$ can simply use its one-on-one optimal control to guarantee $\dot{T}_{cap,i} \geq -1$, thus avoiding becoming a target. However, this is usually not strictly necessary: a range of different evader inputs can be used to keep $\dot{T}_{cap,i} \geq -1$. We define *reactive evasion* as any evader input that satisfies $\dot{T}_{cap,i} \geq -1$. In particular, non-targeted evaders can apply a reactive evasion strategy that is input energy efficient. We show next how evader $E_i$ can compute a state-feedback control for reactive evasion.

Under optimal pursuit (of targeted evader $E_g$) the pursuer's trajectory is a straight line at maximum speed $\bar{v}_p$. So we suppose that evader $E_i$ has some estimate of the pursuer's current direction of travel $\theta_p$ based on its recent behavior. Let $v_p$ be the pursuer's speed, such that $\mathbf{u}_p = (v_p \cos\theta_p, v_p \sin\theta_p)^T = (v_{x_p}, v_{y_p})^T$.

Given the pursuer's relative position, the evader $E_i$ can compute its trajectory under optimal play, including the values of the adjoint variables that parameterize the switching functions. Let $\mathbf{u}^*_{p,i}$ be the pursuer's optimal control with respect to pursuit of $E_i$, with associated heading $\theta^*_{p,i}$. The evader must choose its input $\mathbf{u}_i$ to satisfy $\dot{T}_{cap,i}(\mathbf{q}, \mathbf{u}_p, \mathbf{u}_i) \geq \dot{T}_{cap,i}(\mathbf{q}, \mathbf{u}^*_{p,i}, \mathbf{u}^*_i)$, i.e., to satisfy

$$\boldsymbol{\lambda} \cdot \mathbf{f}(\mathbf{q}, \mathbf{u}_p, \mathbf{u}_i) \geq \boldsymbol{\lambda} \cdot \mathbf{f}(\mathbf{q}, \mathbf{u}^*_{p,i}, \mathbf{u}^*_i). \tag{14}$$

To remove the explicit dependence on $\boldsymbol{\lambda}$ from (14), we now solve for $\boldsymbol{\lambda}$ in terms of the state and optimal controls. From the derivation of the pursuer optimal control in Section 3.2, we have $\lambda_{x_p} = -\lambda_{x_i} = -\lambda_0 \cos\theta^*_{p,i}$ and $\lambda_{y_p} = -\lambda_{y_i} = -\lambda_0 \sin\theta^*_{p,i}$. Let $(x_i(T), y_i(T))^T$ be the location of the evader at capture at time $t = T$ under optimal one-on-one controls for each agent. The pursuer's position at time $t \leq T$ under optimal control $\mathbf{u}^*_{p,i}$ is

$$\begin{aligned} x_p(t) &= x_i(T) - (l + \bar{v}_p(T - t)) \cos\theta^*_{p,i}, \\ y_p(t) &= y_i(T) - (l + \bar{v}_p(T - t)) \sin\theta^*_{p,i}. \end{aligned} \tag{15}$$

From (5) and (6),

$$\lambda_{\theta_i}(t) = \lambda_0[(y_i(t) - y_i(T)) \cos\theta^*_{p,i} - (x_i(t) - x_i(T)) \sin\theta^*_{p,i}]. \tag{16}$$

By adding and subtracting $(l + \bar{v}_p(T - t)) \cos\theta^*_{p,i} \sin\theta^*_{p,i}$ from (16) and substituting with (15), we have

$$\lambda_{\theta_i}(t) = \lambda_0[(y_i(t) - y_p(t)) \cos\theta^*_{p,i} - (x_i(t) - x_p(t)) \sin\theta^*_{p,i}].$$

Finally substituting these expressions for the adjoint back into (14) and rearranging gives us

$$\begin{aligned} &\bar{v}_p - v_p \cos(\theta_p - \theta^*_{p,i}) + (v_i - v^*_i) \cos(\theta_i - \theta^*_{p,i}) + \\ &\quad (\omega_i - \omega^*_i)\left[(y_i - y_p) \cos\theta^*_{p,i} - (x_i - x_p) \sin\theta^*_{p,i}\right] \geq 0. \end{aligned} \tag{17}$$

Thus, to keep $\dot{T}_{cap,i} \geq -1$, the evader must choose its input $(v_i, \omega_i)$ to satisfy the linear inequality (17). □

**Remark 3.** An evader $E_i$ is not in danger of becoming targeted until its $T_{cap,i}$ is close to that of the target evader, at which point it must begin to use a reactive evasion strategy. For instance evader $E_i$ can wait to initiate reactive evasion until $T_{cap,i} \leq \min_j T_{cap,j} + \epsilon$ for some chosen buffer value $\epsilon > 0$. Until that point, the evader is free to use any control input, for instance a "herding" strategy of aligning heading and matching speed with neighbors, such as one adapted from the attraction–orientation–repulsion zonal model of Couzin, Krause, James, Ruxton, and Franks (2002).

## 5. Risk minimization under limited sensing

We now suppose that the pursuer and evaders have a limited sensing range. Each agent can only make use of measurements of agents located inside its sensing range. The limited sensing range may be important for modeling large groups that are widely dispersed or that are tightly packed and suffer from occlusion. We define limits on sensing and adapt the pursuit and evasion strategies introduced in Section 4 to local strategies that address the uncertainty imposed by limited sensing.

In the local (sensing-limited) system, we define $d_{sense}$ as the sensing radius for all agents. An agent's local sensing neighborhood consists of the set of agents within the sensing radius. Let $d_{ij}(t)$ be the distance between agents $i$ and $j$ at time $t$. The local neighborhood of the pursuer at time $t$ is defined as $\mathcal{N}_t(P) = \left\{E_i \mid d_{ip}(t) \le d_{sense}\right\}$. The neighborhood of evader $E_i$ at time $t$ is defined as $\mathcal{N}_t(E_i) = \left\{E_j \,\middle|\, d_{ij}(t) \le d_{sense}\right\} \cup \left\{P \mid d_{ip}(t) \le d_{sense}\right\}$.

### 5.1. Local target selection

Under local sensing, the pursuer must choose a control law based only on measurements of evaders in $\mathcal{N}_t(P)$. We assume that $\mathcal{N}_t(P)$ contains at least one evader at the start of pursuit, $t = 0$. Let $E^*_{local}(t) = \{E_i \in \mathcal{N}_t(P) | T_{cap,i}(t) = \min_{E_j \in \mathcal{N}_t(P)} T_{cap,j}(t)\}$ be the set of evaders within the pursuer's sensing neighborhood at $t$ with minimum time-to-capture. Define *local recursive pursuer strategy* $S_{p,local}(t)$ as the local equivalent to the recursive pursuer strategy $S_p(t)$ defined in Theorem 1, with $E^*(t)$ replaced with $E^*_{local}(t)$. The results of Theorem 1 hold in the local sensing system, so that capture is guaranteed by time $T \le \min_{E_i \in \mathcal{N}_t(P)} T_{cap,i}$ with a finite number of target switches, and in the case of at least one switch $T < \min_{E_i \in \mathcal{N}_t(P)} T_{cap,i}$.

Under local recursive pursuer strategy $S_{p,local}(t)$, switching can occur not only when an evader uses a suboptimal strategy and allows its $T_{cap}$ to become the minimum, but also when a new evader enters the pursuer's sensing neighborhood $\mathcal{N}_t(P)$. Any such switch to a new target with lower $T_{cap}(t)$ will decrease the remaining bound on time-to-capture. Thus the total time spent in pursuit will necessarily be less than or equal to the minimum local time-to-capture calculated at the start of pursuit.

### 5.2. Local evasion strategy

In general, not every evader will be in sensing range of the pursuer at time $t = 0$. The evader's strategy therefore must consist of two distinct phases: First, a phase of "risk reduction" prior to sensing the pursuer where the evader attempts to maneuver itself in such a way as to decrease its chance of becoming the pursuer's target, and second, a phase of reactive evasion based on the evader's strategy in the global sensing system.

Under local sensing, an evader must estimate its own domain of danger based only on the other evaders within its sensing neighborhood. Let $\mathcal{D}_{i,local}(t) = \left\{\mathbf{r}_p \in \mathbb{R}^2 \,\middle|\, i = \operatorname{argmin}_{E_j \in \mathcal{N}_t(E_i)} T_{cap,j}(t)\right\}$ be evader $E_i$'s local estimate of its domain of danger.

**Theorem 3.** *An evader's local estimate of its domain of danger includes all points in its global domain of danger $\mathcal{D}_i(t) \subset \mathcal{D}_{i,local}(t)$ at any given time.*

**Proof.** Consider an evader $E_j \notin \mathcal{N}_t(E_i)$. If there is a point $\mathbf{r}_p \in \mathcal{D}_{i,local}(t)$ such that $T_{cap,j}(t) < T_{cap,i}(t)$, then $\mathbf{r}_p \notin \mathcal{D}_i(t)$. Thus adding additional neighbors can only decrease the local estimate of the domain of danger. □

In the case that $P \notin \mathcal{N}_t(E_i)$ initially, the evader has a chance to guarantee that it will not become a target if it is able to satisfy the following two conditions:

- C1: Evader $E_i$'s local domain of danger lies entirely within its own sensing domain: $\forall \mathbf{r}_p \in \mathcal{D}_{i,local}(t)$, $\|\mathbf{r}_p - \mathbf{r}_i(t)\|_2 < d_{sense}$.
- C2: At all points on the boundary of $E_i$'s local domain of danger, the evader $E_j \in \mathcal{N}_t(E_i)$ with whom the boundary is formed is within sensing range from that point: $\forall \mathbf{r}_p \in \mathcal{D}_{i,local}(t)$, such that $T_{cap,i}(t) = T_{cap,j}(t)$, $\|\mathbf{r}_p - \mathbf{r}_j(t)\|_2 < d_{sense}$.

We denote an evader as *risk-minimized* with respect to its neighborhood at time $t$ if conditions C1 and C2 are both satisfied at time $t$.

**Theorem 4.** *For system (1) under local sensing with pursuer $P$ using the local recursive pursuer strategy $S_{p,local}(t)$, if an evader $E_i$ satisfies conditions C1 and C2 at the moment that $P$ enters the sensing neighborhood of $E_i$, then there exists a control input that guarantees that $E_i$ will avoid capture.*

**Proof.** If the evader's domain of danger extends outside its sensing neighborhood, the pursuer may come from that direction, and possibly target the evader at the moment that they enter each other's sensing neighborhoods. C1 guarantees the evader will be able to sense the pursuer and start a reactive evasion strategy before the pursuer enters the evader's domain of danger. C2 guarantees the pursuer will choose a better target before it enters $E_i$'s domain of danger. If $E_j \notin \mathcal{N}_t(P)$, the pursuer will not choose $E_j$ as its target even if $T_{cap,j}(t) < T_{cap,i}(t)$. It may be the case that $P$ chooses $E_i$ as its target when it enters its sensing neighborhood, but if C2 is satisfied, a better target, i.e. $E_j$, will come into the pursuer's view before the pursuer can enter $\mathcal{D}_{i,local}(t)$. □

Through simulations, we observe that an evader turns towards and moves closer to a neighboring evader with a lower maximum speed to reduce the size of its own domain of danger. Analytical expressions for C1 and C2 as a function of the evader's constraint parameters have not been found. The problem is tractable in the case without turning constraints. In the following, we relax the evader turning constraints in order to derive conditions for risk minimization and the resulting aggregation.

### 5.3. Risk minimization for omni-directional evaders

Consider the case in which the evader turning rates are not constrained, i.e., $|\omega_i(t)|$ and $|v_i\omega_i(t)|$ are unbounded for all $t$. We will consider the terminal condition to be "point capture," so that the game ends when the distance between pursuer and evader reaches zero. Assume $\bar{v}_p > \bar{v}_i$ for all $i = 1, 2, \ldots, n$. The system equations of motion (1) can be written compactly as

$$\dot{\mathbf{r}}_p = \mathbf{u}_p, \;\; \dot{\mathbf{r}}_i = \mathbf{u}_i, \;\; \text{for } i = 1, 2, \ldots, n. \tag{18}$$

Let $\mathbf{r}_{ip} = \mathbf{r}_i - \mathbf{r}_p$ be the *baseline vector* from the pursuer to evader $E_i$, with associated distance $d_{ip} = \|\mathbf{r}_{ip}\|_2$, and normalized unit vector $\hat{\mathbf{r}}_{ip} = \mathbf{r}_{ip}/d_{ip}$ defined for $d_{ip} > 0$. For the one-on-one pursuit-evasion game with these dynamics, as shown in Isaacs (1965), the adjoint vector and optimal controls for both pursuer and evader are constant, with $\boldsymbol{\lambda} = (\bar{v}_p - \bar{v}_i)^{-1}(-\hat{\mathbf{r}}^T_{ip}, \hat{\mathbf{r}}^T_{ip})$, pursuer control $\mathbf{u}^*_p = \bar{v}_p\hat{\mathbf{r}}_{ip}$, and evader control $\mathbf{u}^*_i = \bar{v}_i\hat{\mathbf{r}}_{ip}$. The time to capture is $T_{cap,i}(t) = d_{ip}(t)/(\bar{v}_p - \bar{v}_i)$. These strategies are known as *classical pursuit* and *classical evasion*, respectively.

Under global sensing, Theorems 1 and 2 apply to this system. The equivalent to the reactive evasion constraint (17) of Theorem 2 for this system is derived as follows. Let $\tilde{\theta}_{p,i}$ be the angle of the pursuer's motion measured counterclockwise relative to the baseline vector $\mathbf{r}_{ip}$ for evader $E_i$, and let $\tilde{\theta}_i$ be the angle of the evader's motion relative to the same vector. The reactive evasion condition for this system is the following, from substitution of the adjoint $\boldsymbol{\lambda}$ and optimal controls $\mathbf{u}^*_p$ and $\mathbf{u}^*_i$ into (14):

$$\begin{aligned} & \lambda_0\hat{\mathbf{r}}_{ip} \cdot (\mathbf{u}_i - \mathbf{u}_p) \ge \lambda_0(\bar{v}_i - \bar{v}_p) \\ \Leftrightarrow \quad & v_i \cos\tilde{\theta}_i - v_p \cos\tilde{\theta}_{p,i} \ge \bar{v}_i - \bar{v}_p. \end{aligned} \tag{19}$$

In the case of limited sensing, with sensing radius $d_{sense}$ for all agents, the local version of Theorem 1 from Section 5.1 and Theorems 3 and 4 apply to the non-turning-limited evader system (18).

Without evader turning constraints, the evader domain of danger partition $\{\mathcal{D}_i\}_{i=1,\ldots,n}$ is equivalent to a multiplicatively-weighted Voronoi diagram (Aurenhammer & Edelsbrunner, 1984), where the weight on each evader's distance is given by $(\bar{v}_p - \bar{v}_i)$. Boundaries between domains of danger take the form of circular arcs for neighbor evaders with differing maximum speeds $\bar{v}_i \neq \bar{v}_j$, and straight lines for neighbors with equal maximum speed $\bar{v}_i = \bar{v}_j$. In the special case that all evaders have the same maximum speed, the domain of danger partition is a standard Voronoi diagram, as in Hamilton's original selfish herd model (Hamilton, 1971).

We now use the properties of the multiplicatively weighted Voronoi diagram to derive a mathematical expression for the risk minimization condition of Theorem 4. This we use to define a strategy for evader $E_i$ to decrease the size of its domain of danger and achieve risk minimization in the time period before the pursuer enters its sensing range.

**Theorem 5.** *For system (18) under local sensing with pursuer $P$ using local pursuit strategy $S_{p,local}$, let $E_f$ and $E_s$ be evaders with maximum speeds $\bar{v}_f > \bar{v}_s$. If*

$$\|\mathbf{r}_f(t) - \mathbf{r}_s(t)\|_2 < \left(\frac{\bar{v}_f - \bar{v}_s}{\bar{v}_p - \bar{v}_s}\right) d_{sense} \tag{20}$$

*at the time $t$ when $P$ first enters $\mathcal{N}_t(E_f)$, then there exists a control input that guarantees that $E_f$ will avoid capture.*

**Proof.** Let $\mathbf{r}_{fs} = \mathbf{r}_f - \mathbf{r}_s$ be the baseline vector from $E_s$ to $E_f$ with associated distance $d_{fs} = \|\mathbf{r}_{fs}\|_2$. $E_f$'s domain of danger is the interior of the Apollonius circle (Aurenhammer & Edelsbrunner, 1984) formed by the locus of points where $T_{cap,f} = T_{cap,s}$. The circle has its center at $\mathbf{r}_{Apol,fs} = \mathbf{r}_f + \frac{(\bar{v}_p - \bar{v}_f)^2}{(\bar{v}_p-\bar{v}_s)^2-(\bar{v}_p-\bar{v}_f)^2}\mathbf{r}_{fs}$ and radius $R_{Apol,fs} = \frac{(\bar{v}_p-\bar{v}_s)(\bar{v}_p-\bar{v}_f)}{(\bar{v}_p-\bar{v}_s)^2-(\bar{v}_p-\bar{v}_f)^2} d_{fs}$. The maximum distance from $E_f$ to the circle is $d_{App,fs} = \left(\frac{\bar{v}_p-\bar{v}_f}{\bar{v}_f-\bar{v}_s}\right) d_{fs}$, in the direction of $\mathbf{r}_{fs}$. Since $d_{App,fs}$ is proportional to $d_{fs}$, $E_f$ may reduce this bound on the size of its domain of danger by approaching $E_s$. Since $\bar{v}_f > \bar{v}_s$, $E_f$ can always choose its velocity such that $\dot{d}_{fs} < 0$.

We now consider the conditions C1 and C2 necessary for $E_f$ to achieve risk minimization with respect to $E_s$ under Theorem 4. For C2, the distance from $\mathbf{r}_s$ to the edge of $E_f$'s domain of danger must be less than $d_{sense}$. This is satisfied when $d_{fs} + d_{App,fs} < d_{sense}$, which is equivalent to (20). For C1, $E_f$'s domain of danger must lie within its own sensing range, so $d_{App,fs} < d_{sense}$ is the necessary condition. Since $d_{fs} \geq 0$, this is satisfied whenever C2 is satisfied, by (20). We have shown that any evader $E_f$ that satisfies (20) with a slower neighbor $E_s$ is risk minimized and will be able to avoid capture by Theorem 4. □

For an evader $E_i$ with $P \notin \mathcal{N}_t(E_i)$ at $t = 0$, we call the time interval before $P$ enters its sensing range the *risk reduction phase*. During this phase, the best strategy for the evader is to choose a slower neighbor and move towards it at maximum speed until the risk minimization condition (20) is satisfied. Fig. 4 illustrates how the domains of danger decrease in size during the risk reduction phase when all evaders with a slower neighbor use this aggregating strategy.

When pursuer $P$ enters evader $E_i$'s sensing range, $E_i$ only knows $T_{cap,j}(t)$ of its neighbors $E_j \in \mathcal{N}_t(E_i)$, and $P$ chooses its target based only on $T_{cap,j}(t)$ of its neighbors $E_j \in \mathcal{N}_t(P)$. In this context, an evader must use its best estimate of the pursuer's estimate of the minimum $T_{cap}$ in order to decide when to begin its reactive evasion strategy. $E_i$ should begin reactive evasion when $T_{cap,i}(t) - T^*_{cap}(t) \leq \delta$, for some $\delta > 0$ where $T^*_{cap}(t)$ is the minimum $T_{cap,j}(t)$ for the evaders $E_j \in \mathcal{N}_t(E_i)$. When $T_{cap,i}(t) - T^*_{cap}(t) > \delta$, $E_i$ can remain in place.

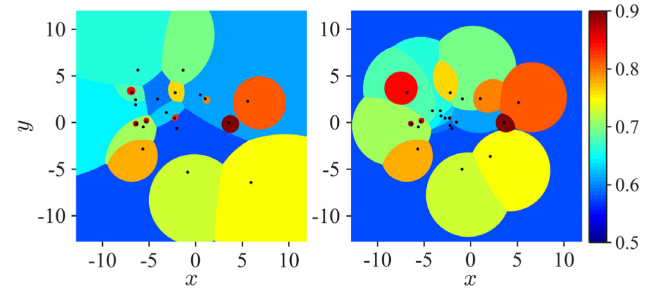


**Fig. 4.** Weighted-Voronoi domain of danger partition, calculated for a pursuer with maximum speed $\bar{v}_p = 1$ and position not sensed by evaders. Each black dot denotes the position of an evader and the color of the surrounding cell (domain of danger) indicates the evader's maximum speed. Left: initially with random initial positions. Right: after running the risk reduction strategy for locally sensing evaders with sensing radius $d_{sense} = 10$.

Consider a graph $\mathcal{G}$ where evaders act as nodes, and an edge $e_{ij}$ from evader $E_i$ to evader $E_j$ is present only if $E_i$ is risk minimized with respect to $E_j$. This forms a directed graph with edges only going from a faster evader to a slower evader. Due to that hierarchy, any connected component must contain a spanning tree with the slowest evader in the component as the root.

**Theorem 6.** *Under the local evasion strategy, an evader can only be captured if it is the slowest evader in a connected component of $\mathcal{G}$.*

**Proof.** Let $E_1$ be the target of pursuer $P$ under local sensing when $P$ enters $\mathcal{N}_{t_0}(E_1)$ at time $t_0$. If $E_1$ is risk minimized with respect to another evader $E_2$ at $t_0$, then by the definition of $\mathcal{G}$ it is not the slowest evader in its connected component, and by Theorem 5 the target of $P$ will eventually switch to another evader. If $E_1$ is not risk minimized it must be the slowest evader within its connected component, and the other evaders will be able to use reactive evasion to avoid becoming the target, leading to the capture of $E_1$. □

## 6. Discussion

We have presented solutions for a pursuit-evasion differential game with practical motion constraints for the evader not considered before in this setting. The optimal strategies for the one-on-one game were used to analyze strategies in a system with a pursuer that seeks to capture in minimum time any single evader among multiple, heterogeneous, non-cooperating evaders. We showed that the optimal strategy for the pursuer is to focus on a single evader that can be captured in the minimum time, and that non-targeted evaders are always able to avoid capture by using a strategy of reactive evasion. We showed how to compute decentralized state feedback reactive control laws. These reactive strategies allow an agent to remain still until necessary. The strategies also include herding behaviors. Because of the constraint on turning rate, the partition of the plane into domains of danger for the evaders depends not only on relative positions but also on relative headings.

In the case that agents have limited sensing range we have shown that a strategy for risk reduction provides a mechanism for group aggregation. This behavior could be leveraged in engineered multi-agent systems with limited sensing. For instance, if every agent (except for the slowest) stays close to at least one slower neighbor at all times, the sensing network will remain connected at all times based on the proof of Theorem 6. Here we have assumed a fast pursuer with sensing range equal to that of the evaders; inclusion of faster evaders with larger sensing range may open the door to additional evader strategies, such as avoidance of detection by the predator, requiring further analysis.

A weakness of the current approach is the assumption that all agents have accurate knowledge of the motion constraints and states of other agents. To address this, the strategies will need to be adapted to uncertainties in the agent estimates of these system parameters. For example, the approach of Oyler, Kabamba, and Girard (2015), provides strategies for each agent in a three-player pursuit-evasion game, which are derived based on "worst case" values for uncertain parameters. Extensions to three-dimensional spaces (Ardema & Rajan, 1987) are also of interest for air and undersea applications.

## Acknowledgments

We thank Daniel Rubenstein, Simon Levin, and Philip Holmes for motivating discussions.

## References

Angelani, L. (2012). Collective predation and escape strategies. *Physical Review Letters*, *109*(11), 118104.
Ardema, M. D., & Rajan, N. (1987). An approach to three-dimensional aircraft pursuit-evasion. *Computers & Mathematics with Applications*, *13*(1–3), 97–110.
Aurenhammer, F., & Edelsbrunner, H. (1984). An optimal algorithm for constructing the weighted Voronoi diagram in the plane. *Pattern Recognition*, *17*(2), 251–257.
Bakolas, E. (2013). Evasion from a group of pursuers with double integrator kinematics. In *Proc. IEEE Conf. Decision and Control* (pp. 1472–1477).
Bakolas, E., & Tsiotras, P. (2010). Optimal pursuit of moving targets using dynamic Voronoi diagrams. In *Proc. IEEE Conf. Decision and Control* (pp. 7431–7436).
Bakolas, E., & Tsiotras, P. (2012). Relay pursuit of a maneuvering target using dynamic Voronoi diagrams. *Automatica*, *48*(9), 2213–2220.
Balkcom, D. J., & Mason, M. T. (2002). Time optimal trajectories for bounded velocity differential drive vehicles. *International Journal of Robotics Research*, *21*(3), 199–217.
Başar, T., & Oldser, G. J. (1999). *Dynamic noncooperative game theory* (2nd ed.). New York: SIAM, (Chapter 8).
Chen, J., Zha, W., Peng, Z., & Gu, D. (2016). Multi-player pursuit–evasion games with one superior evader. *Automatica*, *71*, 24–32.
Couzin, I. D., Krause, J., James, R., Ruxton, G. D., & Franks, N. R. (2002). Collective memory and spatial sorting in animal groups. *Journal of Theoretical Biology*, *218*(1), 1–11.
Cressman, R., & Garay, J. (2011). The effects of opportunistic and intentional predators on the herding behavior of prey. *Ecology*, *92*(2), 432–440.
Domenici, P., & Ruxton, G. D. (2015). Prey behaviors during fleeing: escape trajectories, signaling and sensory defenses. In W. Cooper, & D. Blumstein (Eds.), *Escaping from predators: An integrative view of escape decisions by prey* (pp. 199–224). Cambridge: Cambridge University Press.
Elliott, R. J., & Kalton, N. J. (1972). The existence of value in differential games of pursuit and evasion. *Journal of Differential Equations*, *12*, 504–523.
Eshel, I., Sansone, E., & Shaked, A. (2006). Gregarious behaviour of evasive prey. *Journal of Mathematical Biology*, *52*(5), 595–612.
Exarchos, I., Tsiotras, P., & Pachter, M. (2015). On the suicidal pedestrian differential game. *Dynamic Games and Applications*, *5*(3), 297–317.
Fuchs, Z. E., & Khargonekar, P. P. (2011). Encouraging attacker retreat through defender cooperation. In *Proc. IEEE conf. decision and control* (pp. 235–242).
Hamilton, W. (1971). Geometry for the selfish herd. *Journal of Theoretical Biology*, *31*(2), 295–311.
Howland, H. C. (1974). Optimal strategies for predator avoidance: the relative importance of speed and manoeuvrability. *Journal of Theoretical Biology*, *47*(2), 333–350.
Isaacs, R. (1965). *Differential games: A mathematical theory with applications to warfare and pursuit, control and optimization*. New York: Wiley.
Lee, S., Pak, H., & Chon, T. (2006). Dynamics of prey-flock escaping behavior in response to predator's attack. *Journal of Theoretical Biology*, *240*(2), 250–259.
Liu, S.-Y., Zhou, Z., Tomlin, C., & Hedrick, J. K. (2013a). A gradient-based method for team evasion. In *Proc. ASME Dynamic Systems and Control Conf.* (pp. V003T36A004–V003T36A004).
Liu, S.-Y., Zhou, Z., Tomlin, C., & Hedrick, J. K. (2013b). Evasion as a team against a faster pursuer. In *Proc. IEEE American Control Conference* (pp. 5368–5373).
Oyler, D. W., Kabamba, P. T., & Girard, A. R. (2015). Dominance in pursuit-evasion games with uncertainty. In *Proc. IEEE Conf. Decision and Control* (pp. 5859–5864).
Oyler, D. W., Kabamba, P. T., & Girard, A. R. (2016). Pursuit–evasion games in the presence of obstacles. *Automatica*, *65*, 1–11.
Pachter, M. (1987). Simple-motion pursuit-evasion in the half plane. *Computers & Mathematics with Applications*, *13*(1–3), 69–82.
Ruiz, U., & Murrieta-Cid, R. (2016). A differential pursuit/evasion game of capture between an omnidirectional agent and a differential drive robot, and their winning roles. *International Journal of Control*, 1–16.
Scott, W. L., & Leonard, N. E. (2014). Dynamics of pursuit and evasion in a heterogeneous herd. In *Proc. IEEE Conf. Decision and Control* (pp. 2920–2925).
Scott, W. L., & Leonard, N. E. (2018). Minimum-time trajectories for steered agent with constraints on speed, lateral acceleration, and turning rate. *ASME Journal of Dynamic Systems, Measurement, and Control*.
Selvakumar, J., & Bakolas, E. (2016). Evasion from a group of pursuers with a prescribed target set for the evader. In *Proc. IEEE American Control Conference* (pp. 155–160).
Sussmann, H. J., & Tang, G. (1991). Shortest paths for the Reeds-Shepp car: a worked out example of the use of geometric techniques in nonlinear optimal control. In *Rutgers Center for Systems and Control Technical Report, Vol. 10* (pp. 1–71).
Tan, H., & Wilson, A. M. (2011). Grip and limb force limits to turning performance in competition horses. *Proceedings of the Royal Society of London Series B (Biological Sciences)*, *278*(1715), 2105–2111.
Vabø, R., & Nøttestad, L. (1997). An individual based model of fish school reactions: predicting antipredator behaviour as observed in nature. *Fisheries Oceanography*, *6*(3), 155–171.
Vabø, R., & Skaret, G. (2008). Emerging school structures and collective dynamics in spawning herring: A simulation study. *Ecological Modelling*, *214*(2), 125–140.
Viscido, S., Miller, M., & Wethey, D. (2001). The response of a selfish herd to an attack from outside the group perimeter. *Journal of Theoretical Biology*, *208*(3), 315–328.
Wood, A., & Ackland, G. (2007). Evolving the selfish herd: emergence of distinct aggregating strategies in an individual-based model. *Proceedings of the Royal Society of London Series B (Biological Sciences)*, *274*(1618), 1637–1642.
Zhou, Z., Zhang, W., Ding, J., Huang, H., Stipanović, D. M., & Tomlin, C. J. (2016). Cooperative pursuit with Voronoi partitions. *Automatica*, *72*, 64–72.

**William Lewis Scott** received his Bachelors in Mechanical Engineering from the University of Delaware in 2011. He received his M.A. (2013) and Ph.D. (2017) both from the Princeton University Dept. of Mechanical and Aerospace Engineering. He is currently a postdoctoral associate in the Collective Dynamics and Control Lab at the University of Maryland Dept. of Aerospace Engineering and Institute for Systems Research. His research interests include control of multi-agent systems, differential games, legged locomotion, and soft robotics.

**Naomi Ehrich Leonard** received the B.S.E. degree in mechanical engineering from Princeton University in 1985 and the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 1991 and 1994. From 1985 to 1989, she worked as an Engineer in the electric power industry. Leonard is the Edwin S. Wilsey Professor of Mechanical and Aerospace Engineering, Associated Faculty of Applied and Computational Mathematics, and Director of the Council on Science and Technology at Princeton University. Leonard's research and teaching are in control and dynamical systems with current interests in coordinated control for multi-agent systems, mobile sensor networks, collective animal behavior, and human decision-making dynamics.