The Pursuit-Evasion-Defense Differential Game in Dynamic Constrained Environments

Jaime F. Fisac

S. Shankar Sastry

Abstract— Dynamic multi-player games are powerful abstractions of important real-world problems involving multiple interacting agents in both cooperative and adversarial settings. This paper studies a three-player differential pursuit-evasion game in which a pursuer aims to capture a fleeing evader while a third player, the defender, cooperates with the latter by attempting to intercept or delay the pursuer to avoid capture. Our analysis considers time-varying dynamics and allows the presence of possibly moving obstacles in the domain. We apply a recent theoretical result to express the outcome of the game through the solution of a double-obstacle Hamilton-Jacobi-Isaacs variational inequality, and propose a novel approach to break down the problem into two simpler two-player games with dynamic targets and constraints, which can be solved at a much lower cost. Although conservative, this method guarantees correctness of the computed winning region and strategy for the evader-defender team when a feasible escape solution is found. We demonstrate both the full solution and the approximation method through a numerical example.

I. INTRODUCTION

The study of dynamic reach-avoid games has developed considerably in the last few decades and has played a crucial role in many important engineering problems, including collision avoidance [1],[2],[3], target surveillance [4], energy management [5], and safe reinforcement learning [6],[7]. In the two-player reach-avoid formulation, one seeks to determine the set of states from which one of the players can successfully drive the system to a target set while remaining inside some state constraints at all times, in spite of the opposing actions of the other player: this set is referred to as the *reach-avoid set* (or *capture basin*) of the target under the constraints. Targets typically describe desired waypoints or operating conditions and constraints can model obstacles in the environment or forbidden configurations. A particularly interesting class of reach-avoid games, originally studied in detail by Isaacs [8], are the so called games of pursuit, in which a pursuer aims to capture a fleeing evader by reducing the distance between them. Both players are allowed to move on the domain with given dynamics.

Although analytic solutions can be found in simple cases, the complexity of these problems, which may present non-linear dynamics, bounded domains and obstacles, generally requires computing solutions numerically. In many important cases, the reach-avoid set can be obtained through a dynamic programming approach, by finding the viscosity solution to

the corresponding Hamilton-Jacobi-Isaacs (HJI) equation in the form of a variational inequality. While computationally intensive, numerical solution methods for Hamilton-Jacobi equations, such as [1], [9], [10], are able to accurately solve problems with low dimensionality.

The classic two-player reach-avoid game has been rigorously studied and solved using the Hamilton-Jacobi framework, for the time-invariant case with static targets and constraints. In [11], previous regularity difficulties at the constraint boundaries were overcome through the introduction of an auxiliary value function (proposed in [12] for the single-player optimal control setting). A first approach to extend these results to time-varying systems was proposed in [13], which required augmenting the state space with time, reducing time-dependence to state-dependence; this, however, entailed significant computational expense due to the curse of dimensionality. In recent work with our collaborators, we extended the Hamilton-Jacobi formulation to incorporate time-varying dynamics, targets and constraints without the need for such state augmentation [14]. This has opened new possibilities including the study of games in dynamic environments and the obtention of solutions for certain large-scale problems like safe trajectory planning for multi-vehicle systems [3]. These problems constitute highly current and relevant challenges, such as the definition of an unmanned traffic management (UTM) paradigm to enable safe operation of autonomous aircraft in low-altitude civil airspace [15].

Multi-player reach-avoid games have also been subject to growing attention in recent years. In [16] a decentralized control scheme for multiple pursuers based on a Voronoi cell decomposition was shown to guarantee capture in finite time of a single evader in a bounded, convex polytope in the plane; these guarantees were extended in [17] to domains with an escape window for the evader, through a modified scheme in which one of the pursuers guards the exit to prevent the evader from escaping while the other pursuers ensure capture. Games with multiple evaders have considered different kinds of evader behavior: previous work in [18] assumed random evader motion, while some recent work [19] considered multiple selfish evaders in a heterogeneous herd, each trying to avoid its own capture. Cooperation between evaders to avoid a hidden pursuer has been studied in [20], where the evaders aim to maximize the overall capture time for the entire group. A cooperative setting inspired by predator-prey behaviors in nature is investigated for one-pursuer, two-evader games in [21], where the evader that is not being targeted harasses the pursuer to increase its cost in the hope of dissuading it from continuing the pursuit.

Most recently, there has been particular interest in augmenting pursuit-evasion games with a third kind of player, called the *defender*, whose aim is to prevent the pursuer from capturing the evader by either intercepting it or delaying it enough for the evader to escape. If the evader and the defender are assumed to have perfect coordination, it is possible to treat them as a single player with two separate control inputs. A first approach in [22],[23] assumes a particular type of feedback control scheme (pure pursuit and proportional navigation) for the pursuer, while in [24] the authors switch to a differential game setting that they refer to as "active target defense". Motivated by aircraft defense from missiles, their analysis assumes unconstrained space with isotropic, time-invariant dynamics. However, we find that for many important real-world applications, including aerospace scenarios like maneuvering in the proximity of other aircraft, it is necessary to provide a more generalized treatment allowing for the presence of possibly moving obstacles and time-dependent dynamics. Given the latest theoretical advances and the current relevance of these challenges, we feel that the time is propitious for pursuing this line of research.

The present work makes two main contributions. First, we provide a novel method, based on the recent theoretical results in [14], to compute the victory domains and guaranteed winning strategies for both teams through the solution of a new type of Hamilton-Jacobi-Isaacs variational inequality. Next, we propose a principled approach to decompose the full problem into two smaller two-player games that can be solved at a much lower computational expense, allowing the obtention of timely solutions that are conservative from the point of view of the evader-defender team.

II. PROBLEM FORMULATION

A. System Dynamics

We formulate the pursuit-evasion-defense problem as a differential reach-avoid game, where one of the players is attempting to drive the system into some target set without leaving a constraint set, while the other player attempts to hinder it: in this setting, the first player is the pursuer, and the second is comprised by the evader-defender team.

Let $t = 0$ be the start time of the game and $t = T > 0$ be the end time. Let the state $x := (e, p, d)$ encode the positions of the *evader*, *pursuer* and *defender* in the n-dimensional domain (typically $n = 2$ or $n = 3$); that is, $e, p, d \in \mathbb{R}^n$ and $x \in \mathbb{R}^{3n}$. Each player i can choose an input signal $u_i(\cdot)$ from the set $\mathbb{U}_i$ of measurable functions from $[0, T]$ to a nonempty compact set $\mathcal{U}_i \subset \mathbb{R}^{n_i}$. Consider the system dynamics:

$$\dot{x}(t) = f\big(x(t), u_e(t), u_p(t), u_d(t), t\big), \quad \text{a.e. } t \in [0, T], \quad (1)$$

where the flow field $f : \mathbb{R}^{3n} \times \mathcal{U}_e \times \mathcal{U}_p \times \mathcal{U}_d \times [0, T] \to \mathbb{R}^{3n}$, is assumed to be uniformly continuous, with

$$|f(x, u_e, u_p, u_d, t)| < C,$$
$$|f(x, u_e, u_p, u_d, t) - f(\tilde{x}, u_e, u_p, u_d, t)| \leqslant L|x - \tilde{x}|, \quad (2)$$

for some $C > 0, L > 0$ and all $t \in [0, T], x, \tilde{x} \in \mathbb{R}^{3n}$, $u_e \in \mathcal{U}_e, u_p \in \mathcal{U}_p, u_d \in \mathcal{U}_d$. Under these conditions, system trajectories are well defined and continuous.

Further, we assume that the dynamics of the different players are in fact decoupled from each other. The dynamics of the system can then be decomposed as:

$$\dot{e} = f_e(e, u_e, t), \quad (3a)$$
$$\dot{p} = f_p(p, u_p, t), \quad (3b)$$
$$\dot{d} = f_d(d, u_d, t), \quad (3c)$$

where f_e, f_p, f_d inherit the boundedness and continuity of f.

B. Winning Conditions: Targets and Constraints

We will now introduce the appropriate framework to describe the target and constraint sets for the pursuit-evasion-defense game. Given a closed set $\mathcal{M} \subseteq \mathbb{R}^m$, its associated signed distance function $d_\mathcal{M} : \mathbb{R}^m \to \mathbb{R}$ is given by:

$$d_\mathcal{M}(z) := \begin{cases} \inf_{y \in \mathcal{M}} |z - y|, & z \in \mathbb{R}^m \backslash \mathcal{M}, \\ -\inf_{y \in \mathbb{R}^m \backslash \mathcal{M}} |z - y|, & z \in \mathcal{M}, \end{cases}$$

where $|\cdot|$ denotes a norm on the vector space; we will let this be the Euclidean norm throughout the paper.

We define the upper hemicontinuous[1] set-valued maps $\mathcal{T}, \mathcal{K} : [0, T] \to 2^{\mathbb{R}^{3n}}$ which respectively assign a target set $\mathcal{T}_t \subset \mathbb{R}^{3n}$ and a constraint set $\mathcal{K}_t \subset \mathbb{R}^{3n}$ in the joint state space to each time $t \in [0, T]$. Requiring that $\mathcal{T}_t, \mathcal{K}_t$ are closed for all t, we can construct the space-time sets

$$\mathbb{T} := \bigcup_{t \in [0,T]} \mathcal{T}_t \times \{t\}, \qquad \mathbb{K} := \bigcup_{t \in [0,T]} \mathcal{K}_t \times \{t\},$$

which are then closed subsets of $\mathbb{R}^{3n} \times [0, T]$ (see [14] for a proof).

The closed sets $\mathbb{T}$ and $\mathbb{K}$ can then be implicitly characterized as the subzero regions of two Lipschitz functions $l : \mathbb{R}^{3n} \times [0, T] \to \mathbb{R}$ and $g : \mathbb{R}^{3n} \times [0, T] \to \mathbb{R}$ respectively, that is, $\exists L_l, L_g > 0 : \forall (x, t), (\tilde{x}, \tilde{t}) \in \mathbb{R}^{3n} \times [0, T]$,

$$|l(x, t) - l(\tilde{x}, \tilde{t})| \leqslant L_l |(x, t) - (\tilde{x}, \tilde{t})|,$$
$$|g(x, t) - g(\tilde{x}, \tilde{t})| \leqslant L_g |(x, t) - (\tilde{x}, \tilde{t})|, \quad (4)$$

so that

$$\begin{aligned} (x, t) \in \mathbb{T} &\iff l(x, t) \leqslant 0, \\ (x, t) \in \mathbb{K} &\iff g(x, t) \leqslant 0. \end{aligned} \quad (5)$$

These functions always exist, since we can simply choose the signed distance functions $l(x, t) = d_\mathbb{T}(x, t)$ and $g(x, t) = d_\mathbb{K}(x, t)$, which are Lipschitz continuous by construction, i.e. they are the infimum of point-to-point distances.

In an analogous fashion, we can define the set of possibly moving obstacles in the domain by a closed upper hemicontinuous map $\mathcal{O} : [0, T] \to 2^{\mathbb{R}^n}$ and its associated space-time set $\mathbb{O} \subseteq \mathbb{R}^n \times [0, T]$. Note the different dimensionality in this case, since this object is defined in terms of the domain $\mathbb{R}^n$ and not the joint state space $\mathbb{R}^{3n}$. It is important not to confuse the obstacle set $\mathbb{O}$ with the constraint set $\mathbb{K}$, since, as we will see, the obstacle $\mathbb{O}$ will contribute to both the constraint $\mathbb{K}$ and the target $\mathbb{T}$.

---

[1] A set-valued map $\mathcal{M} : [0, T] \to 2^{\mathbb{R}^m}$ is upper hemicontinuous (also called upper semicontinuous) if for any open neighborhood V of $\mathcal{M}(t)$ there is an open neighborhood U of t such that $\mathcal{M}(\tau) \subseteq V \ \forall \tau \in U$.

The above definition of obstacles, targets and constraints is flexible and allows for unconnected and non-convex sets with a variety of behaviors, including changing topologies over time (e.g. a target splitting into multiple separate sets or disappearing entirely). We can use these sets to describe the objective and restrictions of the game from the point of view of the pursuer, namely the pursuer will win the game if it succeeds in driving the system state to the target set at some time $t \in [0, T]$ while avoiding leaving the constraint set at any prior instant $\tau \in [0, t]$.

The set $\mathcal{T}_t$ must therefore capture all winning configurations for the pursuer at time t. We define a *capture radius* $r_C > 0$ such that the pursuit is considered successful if the pursuer and the evader ever come as close as this distance; in addition, if the evader or the defender breach the domain constraints (run into an obstacle) we will also consider the game to be won by the pursuer; finally, the evader-defender team loses if the evader and the defender come within a distance of r_A or less of each other (we refer to this quantity as the *accident radius*). The target set $\mathcal{T}_t$ is thus given by

$$\mathcal{T}_t := \big\{(e, p, d) \in \mathbb{R}^{3n} : |e - p| \leqslant r_C\big\}$$
$$\cup \big\{(e, p, d) \in \mathbb{R}^{3n} : |e - d| \leqslant r_A\big\}$$
$$\cup \big\{(e, p, d) \in \mathbb{R}^{3n} : e \in \mathcal{O}_t\big\} \quad (6)$$
$$\cup \big\{(e, p, d) \in \mathbb{R}^{3n} : d \in \mathcal{O}_t\big\},$$

which leads us to define the target function $l(x, t)$ as

$$l(x, t) := \min\big\{|e - p| - r_C, |e - d| - r_A, d_\mathbb{O}(e, t), d_\mathbb{O}(d, t)\big\}. \quad (7)$$

The constraint set $\mathcal{K}_t$, on the other hand, will be the complement in $\mathbb{R}^{3n}$ of all conditions that make the pursuer lose the game instantaneously by failing to keep the necessary separation with obstacles or with the defender. We define the *interception radius* r_I as the minimum separation between pursuer and defender admissible to the pursuer. Let

$$\mathcal{K}_t := \big\{(e, p, d) \in \mathbb{R}^{3n} : |p - d| \geqslant r_I\big\}$$
$$\cap \big\{(e, p, d) \in \mathbb{R}^{3n} : p \notin \overset{\circ}{\mathcal{O}}_t\big\}, \quad (8)$$

and hence[2] we can define the constraint function $g(x, t)$ as

$$g(x, t) := \max\big\{|p - d| - r_I, -d_\mathbb{O}(p, t)\big\}. \quad (9)$$

Note that under these definitions l, g satisfy (4) and (5).

C. Value of the Game and Player Strategies

Given a particular system trajectory $x(\cdot)$, the outcome of the game can be characterized by the functional:

$$\mathcal{V}\big(x(\cdot)\big) := \min_{t \in [0, T]} \max\left\{l(x(t), t), \max_{\tau \in [0, t]} g(x(\tau), \tau)\right\}. \quad (10)$$

Note that this value function is constructed to be negative or zero exactly when the trajectory starting at $(x, 0)$ reaches

the target without *previously* breaching the constraints, as desired. Therefore the pursuer wins the game for those realizations $x(\cdot)$ for which $\mathcal{V}\big(x(\cdot)\big) \leqslant 0$.

We work in a perfect information setting, assuming that the system dynamics and state at each time are fully known to all players. For the analysis in this paper we will allow the pursuer to choose its instantaneous control input u_p after observing the action (u_e, u_d) played by the evader-defender team. Formally, we define the set of *nonanticipative strategies* for the pursuer as

$$\Gamma := \big\{\gamma : \mathbb{U}_e \times \mathbb{U}_d \to \mathbb{U}_p \mid$$
$$\forall t \in [0, T], \ \forall u_e(\cdot), \hat{u}_e(\cdot) \in \mathbb{U}_e, \ \forall u_d(\cdot), \hat{u}_d(\cdot) \in \mathbb{U}_d,$$
$$\big(u_e(\tau) = \hat{u}_e(\tau) \wedge u_d(\tau) = \hat{u}_d(\tau) \text{ a.e. } \tau \in [0, t]\big)$$
$$\Rightarrow \big(\gamma[u_e, u_d](\tau) = \gamma[\hat{u}_e, \hat{u}_d](\tau) \text{ a.e. } \tau \in [0, t]\big)\big\}. \quad (11)$$

The above generally results in an advantage to the pursuer, since it can adapt its instantaneous input to the one played by the evader-defender team. However, in our particular setting, Isaac's condition [8] holds due to the decoupling of the dynamics, and the order of the optimizations is inconsequential; that is, we could have instead let the evader and the defender use nonanticipative strategies, with identical results. The value of the game is therefore well defined as:

$$V(x, t) := \inf_{\gamma(\cdot) \in \Gamma} \sup_{\substack{u_e(\cdot) \in \mathbb{U}_e \\ u_d(\cdot) \in \mathbb{U}_d}} \mathcal{V}\big(x(\cdot)\big), \quad (12)$$

where, for notational simplicity, we let $x(\cdot)$ denote the particular trajectory resulting from evader, defender and pursuer control signals $u_e(\cdot), u_d(\cdot)$ and $\gamma[u_e, u_d](\cdot)$, respectively.

In the remainder of this section, we will give a characterization of the winning region for the pursuer, that is, the set of starting configurations at $t = 0$ (or, more generally, intermediate configurations at $t \in [0, T]$) from which the pursuer has a guaranteed winning strategy to capture the evader by the end of the game.

Definition 1: We say that a point in space-time (x, t) is in the *reach-avoid tube* $\mathbb{RA}$ of the target $\mathbb{T}$ under constraints $\mathbb{K}$ when the system trajectory $x(\cdot)$, with all players acting optimally as above, reaches $\mathbb{T}$ at some time $\tau \in [t, T]$ while remaining in $\mathbb{K}$ for all time $s \in [t, \tau]$.

$$\mathbb{RA} := \big\{(x, t) \in \mathbb{R}^{3n} \times [0, T] : \exists \gamma(\cdot) \in \Gamma_t,$$
$$\forall u_e(\cdot) \in \mathbb{U}_e, \forall u_d(\cdot) \in \mathbb{U}_d,$$
$$\exists \tau \in [t, T], x(\tau) \in \mathcal{T}_\tau \ \wedge \ \forall s \in [t, \tau], x(s) \in \mathcal{K}_s\big\}.$$

Definition 2: At any fixed time $t \in [0, T]$, we say that a point x is in the *reach-avoid slice* $\mathcal{RA}_t$ if (x, t) is in the reach-avoid tube $\mathbb{RA}$.

$$\mathcal{RA}_t = \big\{x \in \mathbb{R}^{3n} : (x, t) \in \mathbb{RA}\big\}, \qquad t \in [0, T]$$

Clearly, from the above definition, $\mathcal{RA}_t$ is the winning region for the pursuer starting at time t. On the other hand, as the game is played in a deterministic setting, if the system state is outside of this set, then there exists some control signal from the evader-defender team for which no possible

strategy on the pursuer's part will lead to capture. This means that the complement of $\mathcal{RA}_t$ is the winning region for the evader-defender team starting at time t.

We now characterize the victory domains in terms of the value of the game. The following fact is easy to establish from the above definitions.

Proposition 1: The reach-avoid tube of the space-time target set $\mathbb{T}$ under constraints $\mathbb{K}$ is equal to the zero sublevel set of the value function V.

$$\mathbb{RA} = \left\{ (x,t) \in \mathbb{R}^{3n} \times [0,T] : V(x,t) \leqslant 0 \right\}.$$

Remark 1: As we will see in the next section, V is a continuous function, and therefore $\mathcal{RA}(\cdot)$ is a closed set-valued map with $\mathcal{RA}(t) = \mathcal{RA}_t$ for all $t \in [0,T]$.

III. SOLUTION VIA DOUBLE-OBSTACLE HJI EQUATION

To find the value function $V : \mathbb{R}^{3n} \times [0,T] \to \mathbb{R}$ that will characterize the victory domains for each team, we make use of a theoretical result introduced in recent work [14], which lends itself to a straightforward computational implementation using available tools [25]. In this section we consider the solution to the full problem, in which the evader and the defender are assumed to communicate and cooperate optimally. In the next section, we will propose a decomposition method by which a guaranteed winning strategy may be found for the evader-defender team playing a suboptimal strategy based on partial cooperation.

The value function to the pursuit-evasion-defense game posed in Section II has recently been shown to be the (continuous) viscosity solution to a double-obstacle Hamilton-Jacobi-Isaacs variational inequality.

Proposition 2: The value function $V(x,t)$ for the game with outcome given by (10) is the unique viscosity solution of the variational inequality

$$\max \left\{ \min \left\{ \partial_t V + H\left(x, D_x V, t\right), l(x,t) - V(x,t) \right\}, \right.$$
$$\left. g(x,t) - V(x,t) \right\} = 0, \quad t \in [0,T], x \in \mathbb{R}^{3n}, \quad (13a)$$

with terminal condition

$$V(x,T) = \max \left\{ l(x,T), g(x,T) \right\}, \quad x \in \mathbb{R}^{3n}, \quad (13b)$$

where H is the Hamiltonian given at each (x,t) by the result of the static game

$$H(x,p,t) = \max_{\substack{u_e \in \mathcal{U}_e \\ u_d \in \mathcal{U}_d}} \min_{u_p \in \mathcal{U}_p} f(x, u_e, u_p, u_d, t) \cdot p. \quad (14)$$

The proof is provided in [14] and draws on viscosity solution theory [26], [27].

The above result allows us to readily compute the value function $V(x,t)$ by a numerical grid-based method. Let $\mathbf{i} \in I$ denote the index of the grid point in a discretized computational domain of a compact subset $\mathcal{X} \subset \mathbb{R}^{3n}$ and let $k \in \{1, ..., n\}$ denote the index of each discrete time step in a finite interval $[0,T]$. Since our computation will proceed in backward time, we will let $T = t_0 > t_1 > ... > t_n = 0$.

Algorithm 1: Numerical Double-Obstacle HJI Solution

Data: $\hat{l}(x_\mathbf{i}, t_k), \hat{g}(x_\mathbf{i}, t_k)$
Result: $\hat{V}(x_\mathbf{i}, t_k)$

Initialization
for i $\in I$ **do**

Init $\quad \hat{V}(x_\mathbf{i}, t_0) \leftarrow \max\{\hat{l}(x_\mathbf{i}, t_0), \hat{g}(x_\mathbf{i}, t_0)\};$

Value propagation
for $k \leftarrow 1$ **to** n **do**
$\quad$ **for i** $\in I$ **do**

U1 $\quad\quad \hat{V}(x_\mathbf{i}, t_k) \leftarrow \hat{V}(x_\mathbf{i}, t_{k-1})$
$\quad\quad\quad + \int_{t_k}^{t_{k-1}} \hat{H}\left(x_\mathbf{i}, D_x^+ \hat{V}(x_\mathbf{i}, \tau), D_x^- \hat{V}(x_\mathbf{i}, \tau), t_{k-1}\right) d\tau;$

U2 $\quad\quad \hat{V}(x_\mathbf{i}, t_k) \leftarrow \min \left\{ \hat{V}(x_\mathbf{i}, t_k), l(x_\mathbf{i}, t_k) \right\};$

U3 $\quad\quad \hat{V}(x_\mathbf{i}, t_k) \leftarrow \max \left\{ \hat{V}(x_\mathbf{i}, t_k), g(x_\mathbf{i}, t_k) \right\};$

To numerically solve the variational inequality (13), we use the procedure in Algorithm 1.

The method uses discretized values of the target function $\hat{l}(x_\mathbf{i}, t_k)$ and the constraint function $\hat{g}(x_\mathbf{i}, t_k)$, which can be readily pre-computed by evaluating (7) and (9) for each $\mathbf{i}, k$, given the information of the problem $r_C, r_A, r_I, \mathbb{O}$. $\hat{V}$ denotes the numerical approximation to V. $D_x^+ \hat{V}, D_x^- \hat{V}$ represent the "right" and "left" approximations of spatial derivatives. For the numerical Hamiltonian $\hat{H}$, we use the Lax-Friedrich approximation [28], [29]. For the results in this paper, we use a fifth-order accurate weighted essentially non-oscillatory scheme [10], [28] for the spatial derivatives $D_x^\pm \hat{V}$, and a third-order accurate total variation diminishing Runge-Kutta scheme [10], [30] for the time derivative $D_t \hat{V}$. The integral in the first update step (U1) is then computed numerically. These methods are implemented by means of the computational tools provided in [25].

Once the numerical value function $\hat{V}$ is computed, the winning set for the pursuer $\mathcal{RA}_t$ is approximated by its zero sublevel set $\{x \in \mathbb{R}^{3n} : \hat{V}(x,t) \leqslant 0\}$, to an arbitrary degree of precision determined by the discretization used. In addition, the guaranteed winning strategies are implicitly obtained in solving the minimax to compute the Hamiltonian $\hat{H}$ in step (U1). It follows from Proposition 2 and Algorithm 1 that, from either team's computed winning region, applying the optimizing action at each state as a time-varying feedback policy yields a guaranteed winning strategy for that team, again to an arbitrary degree of precision.

We stress that this methodology can find guaranteed winning starting configurations and strategies in a wide variety of problem settings, making no assumptions about linearity of the dynamics, convexity or connectedness of the domain or time-invariance of dynamics, targets or constraints. However, numerically solving the HJI variational inequality on a high-dimensional grid (typically 6-D for a two-dimensional domain, 9-D for a three-dimensional one) can be computa-

tionally impractical due to the curse of dimensionality. This motivates the search for useful approximations of the full solution through decomposition of the game into smaller sub-problems: the second main contribution of this paper is a principled approach to find one such approximation exploiting the same theoretical framework.

IV. SEQUENTIAL TWO-PLAYER DECOMPOSITION METHOD

We now propose a method to decompose the three-player pursuit-evasion-defense problem discussed so far into two smaller two-player problems than can be solved sequentially to obtain a conservative approximation of the solution from the point of view of the evader-defender team, assuming suboptimal cooperation between the two players. The proposed approach lets the evader execute its best open-loop evasion strategy assuming that the pursuer is playing optimally against it, and neglecting the presence of the defender, until the pursuer is intercepted or a viable escape route is otherwise opened by the defender. The strength of this approach, letting aside its relevance as a model of real-world situations (an evader with limited sensing capabilities), is that it provides a lower bound on the performance of the evader-defender team: indeed, if an escaping strategy is found under these pessimistic assumptions, the team can always execute it and win the game.

1) Pursuit-Evasion with no Defender: We first consider the two-player pursuit-evasion game with no defender; since we are allowing moving obstacles, we use the time-varying reach-avoid formulation in [14], as we did in the full problem. The state of the system is $x_1 := (e, p) \in \mathbb{R}^{2n}$ with dynamics given by (3a),(3b). Given the dynamic obstacle $\mathbb{O}$, we define

$$l_1(x_1, t) := \min \{|e - p| - r_C, d_{\mathbb{O}}(e, t)\}, \quad (15a)$$

$$g_1(x_1, t) := -d_{\mathbb{O}}(p, t). \quad (15b)$$

We can readily solve (13) for this system (numerically, by Algorithm 1) and obtain the value function $V_1(x_1, t)$, whose zero sublevel set $\mathbb{RA}_1$ is the victory domain for the pursuer in the absence of the defender. Importantly, the complement $\mathbb{RA}_1^C$ is the set of configurations (e, p, t) from which the evader can avoid capture without the defender's aid. In particular, if $(e_0, p_0) \notin \mathcal{RA}_1(0)$, the evader can execute its best strategy as given implicitly by (14) in Proposition 2 and win the game, in which case there is no need to solve the second game, as the defender is not required to intervene.

2) Pursuit-Defense with Hybrid Open-Loop Evader: The basic approach now is to fix the trajectory of the evader, as determined by the first game, and transform it into a moving target for the pursuer, as well as a moving obstacle for the defender. The second reach-avoid game is played with state $x_2 := (p, d)$ and dynamics (3b),(3c). Let the initial position of the evader be e_0 and let $u_e^*[\tilde{p}_0](\cdot)$ be its optimal control signal for the first problem with the pursuer starting at $\tilde{p}_0$.

Definition 3: We say that the evader is playing a *hybrid open-loop strategy* in the following sense: the evader initially executes the fixed trajectory $\bar{e}_0[\tilde{p}_0](t)$ corresponding to its optimal open-loop control signal $u_e^*[\tilde{p}_0](\cdot)$, which

assumes that the pursuer starts at position $\tilde{p}_0$ and pursues optimally with no defender; if at any time $\tau \in [0, T]$, $V_1(\bar{e}_0[\tilde{p}_0](\tau), p, \tau) > 0$ for the true position p of the pursuer, the evader will become aware of p and switch to its optimal evasion trajectory $\bar{e}_\tau[p](t)$ from its current position.

Proposition 3: Using the above hybrid open-loop strategy, the evader is guaranteed to win the game if a switch happens.

Proof: Let the switch take place at time $\tau \in [0, T]$, with $V_1(e, p, \tau) > 0$: the result then follows from Proposition 1. ∎

A caveat under our formulation, of course, is that in switching to this winning strategy the evader may blindly run into the defender. To ensure that this does not happen, we can, somewhat conservatively, forbid switching if the defender lies in conflict with the potential path of the evader after the switch. To do this, we need to define additional sets.

Given the initial evader trajectory $\bar{e}_0(\tau)$, let $\bar{e}_\tau[p](t)$ for $t \in [\tau, T]$ be the trajectory followed by the evader if it were to switch at exactly time τ with the pursuer at p. This induces a danger region for the defender

$$\mathcal{E}[p](\tau) = \mathcal{E}_\tau^p := \{\bar{e}_\tau[p](t) + r_A u : t \in [\tau, T], |u| = 1\}. \quad (16)$$

If the defender lies in this set, then the switch *may* lead a collision, if the defender cannot get clear of the evader in time. The evader will only attempt to switch if the pursuer leaves its winning region for the first game. Hence define:

$$\Delta := \{(p, \tau) \in \mathbb{R}^n \times [0, T] : V_1(\bar{e}_0(\tau), p, \tau) = 0\}, \quad (17a)$$

$$\mathbb{E} := \bigcup_{(p, \tau) \in \Delta} \{p\} \times \mathcal{E}_\tau^p \times \{\tau\} \subset \mathbb{R}^{2n} \times [0, T]. \quad (17b)$$

From continuity of $V_1(x_1, t)$ we have upper hemicontinuity of $\mathcal{E}[\cdot](\cdot)$ and closedness of Δ, hence $\mathbb{E}$ is closed. Note that $(p, d, t) \in \mathbb{E} \iff d \in \mathcal{E}_\tau^p \land (p, \tau) \in \Delta$. With this we can finally define

$$l_2(x_2, t) := \min \{|\bar{e}_0(t) - p| - r_C, |\bar{e}_0(t) - d| - r_A, d_{\mathbb{O}}(d, t)\}, \quad (18a)$$

$$g_2(x_2, t) := \max \{|p - d| - r_I, -d_{\mathbb{O}}(p, t), \\ \min\{V_1(\bar{e}_0(t), p, t), d_{\mathbb{E}}(p, d, t)\}\}. \quad (18b)$$

The minimum in (18b) captures the fact that the pursuer loses if a switch takes place, but this only happens if the pursuer leaves its winning region for the first game *and* the defender is clear of the evader's optimal evasion path.

Again, we can solve (13) (using Algorithm 1) to find the solution of this second reach-avoid game through the value function $V_2[e_0, \tilde{p}_0](x_2, t)$, which will depend on the open-loop trajectory fixed for the evader, given its initial position e_0 and the initial position it has *assumed* for the pursuer $\tilde{p}_0$. If the true initial position p_0 of the pursuer is known, we obtain the set of initial positions for the defender from which escape can be achieved through this suboptimal cooperation scheme as the set $\{d \in \mathbb{R}^n : V_2[e_0, p_0](p_0, d, 0) > 0\}$. More generally, since we are not forced to assume that the pursuer starts at any fixed position, the zero sublevel set of $V_2[e_0, \tilde{p}_0](x_2, t)$ characterizes the set $\mathbb{RA}_2$ of configurations (p, d, t) from which the pursuer will win the game if it plays optimally against the defender given the evader's hybrid

strategy. Importantly, a guaranteed winning strategy is always available to the evader-defender team (and indeed computable from (14)) if $(p, d, t) \notin \mathbb{RA}_2$, regardless of the actions of the pursuer. This fact makes the decomposition scheme presented here attractive, since it can enable computation of a viable escape strategy much faster than solving the full problem. Note that, in intrinsically boolean "games of kind" like ours, one does not care about finding a strategy that is optimal as much as one that guarantees winning the game.

V. NUMERICAL EXAMPLE

In this section we show the results of implementing our proposed method for a simulated pursuit-evasion-defense game played on a two-dimensional domain.

We consider a game of duration $T = 1$ played on a square domain $D = [-1, 1] \times [-1, 1]$ in the plane, with a central obstacle $\mathbb{O}$ that moves vertically as $\mathcal{O}(t) = [-0.2, 0.2] \times [-0.6 + b(t), 0.6 + b(t)]$, where

$$b(t) = \begin{cases} -2t, & 0 \leqslant t < 0.2, \\ -0.4, & 0.2 \leqslant t < 0.6, \\ -2.8 + 4t, & 0.6 \leqslant t < 0.8, \\ +0.4, & 0.8 \leqslant t \leqslant 1. \end{cases}$$

The pursuer can move freely on this domain at a maximum velocity that decreases with time as $v_p(t) = 3(1 - t/2)$, while the evader and defender have their motion limited to fixed paths (this allows us to study the problem on a four-dimensional grid rather than a six-dimensional one): the evader can move at a maximum speed $v_e = 2$ on a 1.2×1.6 rectangle centered at the origin, and the defender can move at a maximum speed $v_d = 2.5$ on a horizontal segment at $y_d = 0.3$. We set the capture, interception and accident radii to $r_C = 0.1$, $r_I = 0.15$, $r_A = 0.1$.

The solution to the full game is computed over 127 time steps on a $31^3 \times 86$ grid, the larger dimension corresponding to the path of the evader (longer than the side of the square). Computation using the MATLAB Level Set Toolbox [25] took 2 hours and 4 minutes on a 2013 MacBook Pro with a 3 GHz Intel Core i7 processor and 8 GB 1600 MHz DDR3 memory. Fig. 1 shows four different cross-sections of the computed four-dimensional reach-avoid set for the start of the game, at four different defender abscissae $x_{d,0} = \pm 0.8, \pm 0.4$, and a fixed evader position $e_0 = (0.4, 0.8)$. The region enclosed by the curve is therefore the set of initial positions from which the pursuer can successfully capture the evader by the end of the game while avoiding interception.

We now compare these results to the ones obtained by the decomposition method described in Section IV. We will consider the initial evader position $e_0 = (0.4, 0.8)$. We first solve the pursuit-evasion game with no defender, considering the moving obstacle $\mathbb{O}$. Solving this problem on a three-dimensional $31^2 \times 86$ grid took 3 minutes and 14 seconds on the same machine. We obtain the winning set $\mathbb{RA}_1$ for the pursuer, which can be seen for the initial time $t = 0$ in the top left plot of Fig. 2. The evader has a guaranteed escape strategy if the pursuer starts outside of this region,
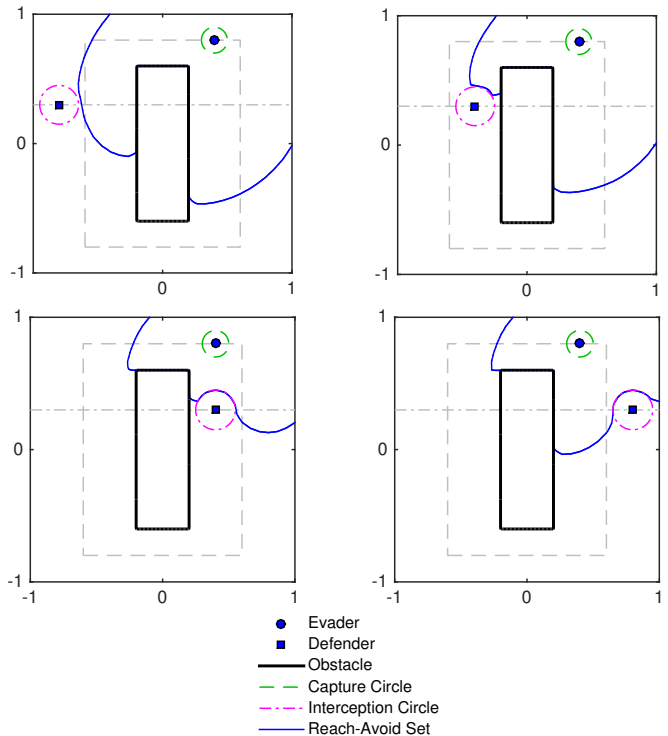

Fig. 1: Reach-avoid set at the start of the game for different defender positions and fixed evader position. The paths along which the evader and defender are allowed to travel are marked in light gray.

in which case the presence of the defender is not required; instead, we will assume that the pursuer starts at position $\tilde{p}_0 = (-0.5, 0)$, well inside the computed capture basin. Next, we simulate the game in forward time for these initial conditions to determine the evader's open-loop trajectory. It should be stressed that in this first game, the evader is fleeing from a *virtual* pursuer (whose initial position $\tilde{p}_o$ may be the known initial position p_0 of the true pursuer or a best guess made by the evader). Snapshots of the simulated trajectories of the evader and the pursuer from these starting conditions are shown in Fig. 2 (this step took roughly 10 seconds of computation). Although capture is achieved at $t = 0.56$, we let the game continue, with the evader trying to gain separation from the pursuer, who is gradually becoming slower; a second capture takes place near the end of the game resulting from a corner in the evader's constrained path.

We can then study the second game between the pursuer and the defender, incorporating the evader's trajectory as a moving target for the pursuer (with r_C) and an obstacle for the defender (with r_A), and additionally requiring the pursuer to stay inside the reach-avoid set for the first game in order to be able to achieve capture. Computation of this second game on a 31^3 grid took just under 2 minutes. The resulting winning region for the pursuer is plotted in Fig. 3 at the initial time for the fixed evader position e_0 and the same four defender positions d_0 as in the full game analysis.

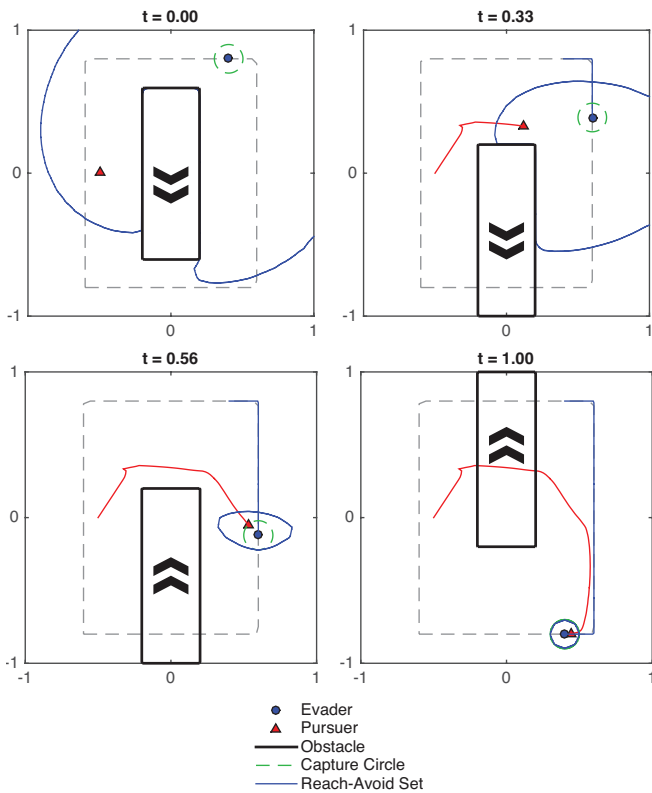Note that the result given by this method, although con-

Fig. 2: Evolution of the reach-avoid set and simulated evader and pursuer trajectories. The reach-avoid cross-section shown at each time represents the set of points from which the pursuer will succeed in capturing the evader if both play optimally for the remainder of the game.
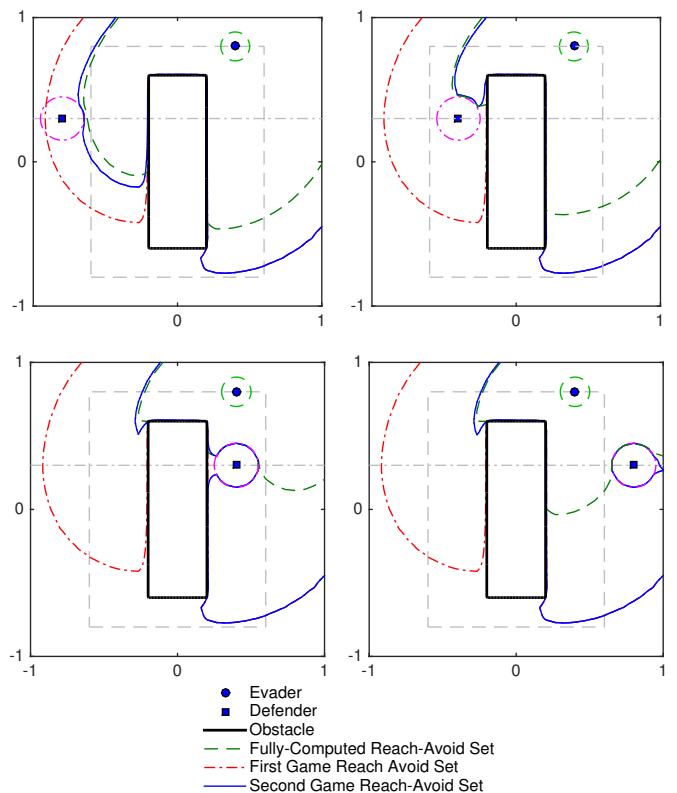

Fig. 3: Comparison of the reach-avoid sets at the start of the game computed through full reachability analysis and the sequential decomposition method. The 2-D cross-sections are shown for different defender positions. The starting position for the evader is fixed in the decomposition approach.

servative for the evader-defender team, produces a close approximation to the true reach-avoid set, especially in the neighborhood of the initial pursuer position $\tilde{p}_0$ assumed by the evader: in the perfect information setting, the starting conditions are known, and the use of this method seems appropriate for a reliable assessment of the game. It should be stressed that computation of this conservative approximation took less than 6 minutes, in contrast to the more than 2 hours taken by the full computation, which is a speed-up of over 20 times. We believe that this approach can be implemented efficiently on more powerful processors to enable close-to-real-time decision making with guarantee certificates in strategic and safety-critical scenarios involving collaboration in adversarial or uncertain environments.

VI. CONCLUSION

We have presented here an analysis of the three-player pursuit-evasion-defense differential game, allowing for dynamic obstacles and time-varying system dynamics. Formulating the problem as a reach-avoid game, we have shown how to compute the winning regions and strategies for each team through the solution to a double-obstacle Hamilton-Jacobi-Isaacs variational inequality, and have provided a

numerical implementation scheme for our algorithm.

We have further proposed a decomposition-based approach that enables relatively fast computation of an inner approximation of the winning region for the evader-defender team and the corresponding guaranteed winning strategies. To our knowledge, this is the first result enabling the obtention of winning strategies for a three-player game of pursuit of this complexity, involving a bounded domain that is effectively nonconvex and time-varying due to the presence of dynamic obstacles. Our methods have been demonstrated on a numerical example.

In future work we will study the possibility of approaching the full optimal solution to the three-player game by establishing a principled iteration procedure based on the two sub-games analyzed here. We also hope to exploit this framework to study collaboration in human-robot teams, in which the evader is human-controlled and the automation-governed defender adapts its strategy to the intent declared by or inferred from the human agent.

VII. ACKNOWLEDGEMENTS

REFERENCES

[1] I. M. Mitchell, A. M. Bayen, and C. J. Tomlin. "A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games". *IEEE Transactions on Automatic Control* 50.7 (2005), pp. 947–957.

[2] K. Margellos and J. Lygeros. "Toward 4-D Trajectory Management in Air Traffic Control: A Study Based on Monte Carlo Simulation and Reachability Analysis". *IEEE Trans. on Control Sys. Technology* 21.5 (2012).

[3] M. Chen, J. F. Fisac, S. Sastry, and C. J. Tomlin. "Safe Sequential Path Planning of Multi-Vehicle Systems via Double-Obstacle Hamilton-Jacobi-Isaacs Variational Inequality". *IEEE European Control Conference* (2015).

[4] J. H. Gillula and C. J. Tomlin. "Guaranteed Safe Online Learning via Reachability: tracking a ground target using a quadrotor". *IEEE International Conference on Robotics and Automation* (2012), pp. 2723–2730.

[5] A. Aswani, H. Gonzalez, S. S. Sastry, and C. Tomlin. "Provably safe and robust learning-based model predictive control". *Automatica* 49.5 (2013), pp. 1216–1226.

[6] J. H. Gillula and C. J. Tomlin. "Reducing Conservativeness in Safety Guarantees by Learning Disturbances Online: Iterated Guaranteed Safe Online Learning." *Robotics: Science and Systems* (2012).

[7] A. K. Akametalu, J. F. Fisac, et al. "Reachability-Based Safe Learning with Gaussian Processes". *IEEE Conference on Decision and Control* (2014).

[8] R. Isaacs. *Differential Games*. New York: John Wiley and Sons, 1965.

[9] J. Sethian. "A fast marching level set method for monotonically advancing fronts." *Proceedings of the National Academy of Sciences of the United States of America* 93.4 (1996), pp. 1591–1595.

[10] S. Osher and R. Fedkiw. *Level set methods and dynamic implicit surfaces*. Vol. 153. Springer Science & Business Media, 2003.

[11] K. Margellos and J. Lygeros. "Hamilton-Jacobi formulation for reach-avoid differential games". *IEEE Transactions on Automatic Control* 56.8 (2011), pp. 1849–1861.

[12] O. Bokanowski, N. Forcadel, and H. Zidani. "Reachability and Minimal Times for State Constrained Nonlinear Problems without Any Controllability Assumption". *SIAM Journal on Control and Optimization* 48.7 (2010), pp. 4292–4316.

[13] O. Bokanowski and H. Zidani. "Minimal time problems with moving targets and obstacles". *18th IFAC World Congress* (2011).

[14] J. F. Fisac, M. Chen, C. J. Tomlin, and S. S. Sastry. "Reach-Avoid Problems with Time-Varying Dynamics, Targets and Constraints". *18th International Conference on Hybrid Systems: Computation and Control* (2015).

[15] P. Kopardekar. *Unmanned Aerial System (UAS) Traffic Management (UTM): Enabling Low-Altitude Airspace and UAS Operations*. Tech. rep. NASA, 2014.

[16] H. Huang, W. Zhang, et al. "Guaranteed decentralized pursuit-evasion in the plane with multiple pursuers". *IEEE Conference on Decision and Control* (2011), pp. 4835–4840.

[17] S. Pan, H. Huang, et al. "Pursuit, evasion and defense in the plane". *IEEE American Control Conference*. 2012, pp. 4167–4173.

[18] J. P. Hespanha and S. Sastry. "Multiple-agent probabilistic pursuit-evasion games". *IEEE Conference on Decision and Control* 3 (1999), pp. 2432–2437.

[19] W. Scott and N. E. Leonard. "Dynamics of Pursuit and Evasion in a Heterogeneous Herd". *IEEE Conference on Decision and Control* (2014), pp. 2920–2925.

[20] S.-y. Liu, C. Tomlin, and J. K. Hedrick. "Evasion of a Team of Dubins Vehicles from a Hidden Pursuer". *IEEE International Conference on Robotics & Automation (ICRA)* (2014), pp. 6771–6776.

[21] Z. E. Fuchs, P. P. Khargonekar, and J. Evers. "Cooperative defense within a single-pursuer, two-evader pursuit evasion differential game". *IEEE Conference on Decision and Control* (2010), pp. 3091–3097.

[22] E. Garcia, D. W. Casbeer, K. Pham, and M. Pachter. "Cooperative Aircraft Defense from an Attacking Missile". *IEEE Conference on Decision and Control* (2014), pp. 2926–2931.

[23] E. Garcia, D. W. Casbeer, K. Pham, and M. Pachter. "Cooperative Aircraft Defense from an Attacking Missile using Proportional Navigation". *AIAA Guidance, Navigation, and Control Conference* (2015), pp. 1–10.

[24] M. Pachter, E. Garcia, and D. W. Casbeer. "Active Target Defense Differential Game". *52nd Annual Allerton Conference on Communication, Control, and Computing (Allerton)* (2014), pp. 46–53.

[25] I. M. Mitchell. *A toolbox of level set methods*. 2004. URL: `http://www.cs.ubc.ca/~mitchell/ToolboxLS`.

[26] L. C. Evans and P. E. Souganidis. "Differential games and representation formulas for solutions of Hamilton-Jacobi-Isaacs equations". *Indiana University mathematics journal* 33.5 (1984), pp. 773–797.

[27] E. Barron. "Differential Games with Maximum Cost". *Nonlinear analysis: Theory, methods & applications* (1990), pp. 971–989.

[28] S. Osher and C.-W. Shu. "High-Order Essentially Nonoscillatory Schemes for HamiltonJacobi Equations". *SIAM Journal on Numerical Analysis* 28.4 (1991), pp. 907–922.

[29] I. M. Mitchell. "Application of Level Set Methods to Control and Reachability Problems in Continuous and Hybrid Systems". PhD thesis. Stanford Univ., 2002.

[30] C.-W. Shu and S. Osher. "Efficient implementation of essentially non-oscillatory shock-capturing schemes". *Journal of Computational Physics* 77.2 (1988), pp. 439–471.