Probabilistic Reachability in Stochasitic Multiplayer Reach-Avoid Games

IEEE Publication Technology, *Staff, IEEE,*

***Abstract*—The probabilistic reachability problem, which involves the computation of probabilistic reachable sets, is studied for nondeterministic reach-avoid games. We consider a multiplayer reach-avoid game with an equal number of attackers and defenders moving on a 2D plane with obstacles. Our work provides a first attempt to address such a problem in stochastic environment and has shown some preliminary results.**

***Index Terms*—Reachability, Reach-avoid games, Stochastic differential equations, Level set method.**

NOMENCLATURE

N	number of players
x_{A_i}, x_{D_j}	states of the attackers and defenders, respectively, m
v_A, v_D	maximum speed of the attackers and defenders, m/s
a_i, d_j	control functions of attackers and defenders, respectively
W	brownian motion
Ω	bounded, open domain
Ω_{obs}	set of obstacles
Ω_{free}	free space
C_{ij}	capture set
R_C	relative distance of capture, m
$\mathbf{x}_{t,x}^{u,d}$	controlled stochastic process
u, d	joint control input of the attacking team and the defending team, respectively
$\mathcal{U}$, $\mathcal{D}$	sets of the joint admissible control inputs of the attacking team and the defending team, respectively
ρ	probability of winning
ϕ	wiener process
ξ	zero mean process
$\mathbb{E}$	expectation
$\mathbf{1}_{\mathcal{K}}$	indicator function
V	value function
u^*, d^*	optimal control input for the attacking team and the defending team, respectively
$\mathcal{K}$, $\mathcal{A}$	reach set and avoid set
A^*, D^*	winning strategy of the attacker and the defender
g	flowfield current
f, σ	deterministic and stochastic function
α, β	direction of the attacker and the defender with respect to the x-axis, rad
$E_{(e_{i,j})}$	database
$e_{i,j}$	capture relationship between the attacker and the defender
Ξ_j	set of all attackers that beat the defender D_j
Ξ_i	set of all defenders that capture an attacker A_i

*This work was supported by the Guangdong Basic and Applied Basic Research Foundation under the Grant No. 2021A1515110569, the National Natural Science Foundation of China under the Grant No. 52202502 and the Practice and Innovation Funds for Graduate Students of Northwestern Polytechnical University under the Grant No. PF2024040. (*Corresponding author: Ye Zhang.*)

Yutong Zhu is with the School of Astronautics, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: yutong.zhu@mail.nwpu.edu.cn).

Ye Zhang is with the Research & Development Institute of Northwestern Polytechnical University in Shenzhen, Shenzhen 518063, China; the School of Astronautics, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: zhang_ye@nwpu.edu.cn).

I. INTRODUCTION

The theory of differential games has been applied in a wide variety of fields such as robotics, spacecraft, aircraft control and military combat. In differential games, the multiplayer reach-avoid (RA) game (or pursuit-evasion game) is defined between two adversarial teams of cooperating players, where one team aims to reach a certain target while the other team aims to prevent the success of their components. The game is an excellent proxy for studying the tools needed to make intelligent automation system a reality.

The multiplayer RA game has been explored in several ways. Many of the previous research focus more on the multi-agent coordination aspect of the game while simplifying assumptions on the adversarial aspect [1]. In more complex RA games such as air combat, where the roles of the players may change over time. Adversary actions are predicted based upon expected opponent strategies, with feedback and re-planning used to adjust for deviations from the predicted actions at run-time [2]. In a more general situation where no prior information on each side is known, the ideal framework to use is the Hamilton-Jacobi-Isaacs (HJI) reachability approach, in which an HJI partial differential equation (PDE) is solved to obtain optimal strategies for both teams [3].

In Ref. [4], the authors address the RA game, whereby each pair of defenders and attackers operates within a compact domain with obstacles, utilizing a HJI reachability approach to provide a solution. This solution is then employed in the multiplayer setting, where defenders are assigned against attackers via a graph-theoretic maximum matching. In Refs. [5]–[7], the authors propose utilizing the maximum matching method to extend the outcome of a 1 vs. 1 RA game to a multi-agent scenario through a maximum matching approach. In Ref. [8], a mixed-integer programming model is proposed to determine the assignment logic between pairs of vehicles when there are more than two vehicles, thereby ensuring safety. These methods are based on the outcome of a 1 vs. 1 RA game and could be enhanced by integrating more tractable HJ reachability results. The above research provides useful insights into the RA problem, but its application is limited by the use of deterministic system dynamics motion models and the lack of consideration of defender collisions.

Most RA games are studied in a deterministic setting, ignoring the influence of uncertainty of system dynamics and the environment [9]. Moreover, complex systems such as those pertaining to air traffic and infrastructure often exhibit complex behaviours that arise from heterogeneous interactions [10], [11]. These behaviours are essentially hybrid in nature. Additionally, the uncertainty inherent to the interleaved discrete and continuous evolution of such systems also leads to the emergence of stochastic systems models, which contain both deterministic and stochastic components. The RA games problem deals with the determination of initial states set in the deterministic hybrid system. This system can be defined as one in which at least one control strategy can be found to steer the system to a target set while guaranteeing collision avoidance [12], [13]. A hybrid system comprising both deterministic and stochastic dynamics is often more challenging to analyse than a deterministic hybrid system. Consequently, this paper primarily examines the RA game problem for deterministic and stochastic dynamics systems. Hence, there is a necessity to extend current solutions to RA games into a more general framework where stochastic part of the system dynamics is considered. In Ref. [14], a connection between stochastic RA problems and stochastic optimal control problems involving discontinuous payoff functions are established. However, the approaches in Ref. [14] did not consider the interactions among multiple players of the games. Inspired by [9], our research attempt to address the problem of multiplayer stochastic RA game using probabilistic reachability analysis.

Figure 1 shows an overview of database building and online implementation based on deterministic and stochastic dynamics model for the RA games in this paper. The contributions and innovations of our work are:

(i) The probabilistic reachability problem in the nondeterministic RA game involving the computation of probabilistic reachable sets is considered and decoupled into a stochastic RA game problem with the same number of defenders and attackers offline and a matching problem with different number of defenders and attackers online.

(ii) An offline database of stochastic RA games is built using probabilistic reachability analysis for RA games with the same number of defenders and attackers based on nonlinear dynamics flowfield.

(iii) A rematching game strategy with different number of defenders and attackers is proposed using an online maximum matching method based on an offline database.

The paper has the following outline. In Sec. II, a general framework of a multiplayer stochastic RA game is given. In Sec. III, the process of probabilistic reachability analysis based on the RA game and the dynamics flowfield is introduced. The method of defender-attacker pair online rematching using offline database are proposed in Sec. IV. The method is

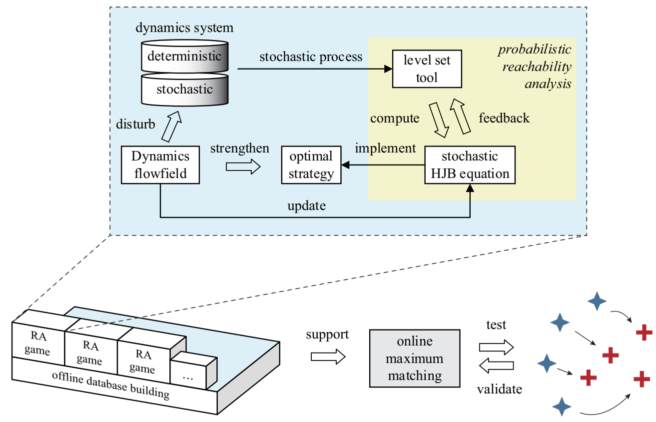


Fig. 1. An overview of database building and online implementation based on deterministic and stochastic dynamics model for the RA games.

demonstrated and validated through simulation results are shown in Sec. V. Finally, in Sec. VI, the full paper is briefly summarized, followed by future work.

II. Stochastic Reach-Avoid Games

We consider a multiplayer reach-avoid game between a team of N attackers, and a team of N defenders. Each player is confined in a bounded, open domain $\Omega \subset \mathbb{R}^2$, which can be partitioned as a set of obstacles Ω_{obs} and free space Ω_{free}. Let $x_{A_i}, x_{D_j} \in \mathbb{R}^2$ denote the states of the attackers and defenders respectively. Initial conditions of the players are denoted by $x^0_{A_i}, x^0_{D_j} \in \Omega_{free}, i = 1, 2, \ldots, \mathbb{N}$. In stochastic reach-avoid games, the dynamic of the system is composed of a deterministic part and a stochastic part. The deterministic part of the dynamics are defined by the following decoupled system for $t \geq 0$:

$$\begin{aligned} \dot{x}_{A_i}(t) = v_A a_i(t), \quad x_{A_i}(0) = x^0_{A_i} \\ \dot{x}_{D_j}(t) = v_D d_j(t), \quad x_{D_j}(0) = x^0_{D_j} \end{aligned} \tag{1}$$

where $a_i(\cdot), d_j(\cdot)$ represent the control functions of attackers and defenders respectively. The attackers have the same maximum speed v_A and the defenders have the same maximum speed v_D. Assuming the movement of both attackers and defenders are influenced by the Brownian motion W, the dynamics of the players can be denoted by the following stochastic differential equations:

$$\begin{aligned} \dot{x}_{A_i}(t) = v_A a_i(t) + \sigma\left(x_{A_i}, a_i\right) dW, \quad x_{A_i}(0) = x^0_{A_i} \\ \dot{x}_{D_j}(t) = v_D d_j(t) + \sigma\left(x_{D_j}, d_j\right) dW, \quad x_{D_j}(0) = x^0_{D_j} \end{aligned} \tag{2}$$

The players' joint state and joint initial condition become $\mathbf{x} = (x_{A_i}, x_{D_j})$, $\mathbf{x^0} = (x^0_{A_i}, x^0_{D_j})$ respectively. In this reach-avoid game, the attacking team aims to reach the target as shown in Fig. 3, a compact subset of the domain, without getting captured by the defenders. The capture conditions are formally described by the capture sets $C_{ij} \subset \Omega^{2N}$ for the pairs of the players. In this paper, we define the capture sets to be $C_{ij} = \left\{ \mathbf{x} \in \Omega^{2N} \left| \|x_{A_i} - x_{D_j}\|_2 \leq R_C \right. \right\}$, the interpretation of which is that an attacker is captured by a defender if their relative distance is within R_C [9]. Consider the special case in which each team only has one player, we can first compute the probabilistic reachable set for the attacker which predicts the region of winning with high probability. Then we extend the Probabilistic HJI framework to deal with the multiplayer reach-avoid games.

III. Reachability Analysis

A. Probabilistic Hamilton-Jacobi-Isaacs Reachability

Stochastic Differential Equations (SDEs) can be used to describe the behavior of a stochastic process. Instead of simulating each realization of the stochastic process, the forward Kolmogorov equation (or Fokker-Planck equation) can be used to describe the time rate of change of the Probability Density Function (PDF) [15]. Combining the optimal control framework with the time-evolution of the PDF will give rise to the field of probabilistic reachability analysis. The probabilistic reachability analysis will compute the transition probability to a certain state. Alternatively, the value function will be equal to the conditional probability of being in the initial set $\mathcal{K}$ at time $t = 0$ and in a final state $\mathbf{x}$ at time $t = T$.

Consider a controlled stochastic process $\mathbf{x}_{t,x}^{u,d}$

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}, u, d) + \sigma(\mathbf{x}, t), \forall t \in [0, T] \tag{3}$$

where $u \in \mathcal{U}$ is the joint control input of the attacking team, and $d \in \mathcal{D}$ is the joint control input of the defending team. The sets $\mathcal{U}$ and $\mathcal{D}$ represent the sets of the joint admissible control inputs of the attacking team and the defending team, respectively. Because the dynamics are stochastic, it is no longer possible to minimize the value function. Instead, the expected pay-off is minimized over all possible further realizations of the Wiener process

$$V(t, x) = \min_{u\in\mathcal{U}} \max_{d\in\mathcal{D}} \mathbb{E}\left[\phi\left(\mathbf{x}_{t,x}^{u,d}(T)\right)\right] \tag{4}$$

Let $\mathcal{K}$ be the non-empty target set in $\mathbb{R}^d$, $\rho \in [0, 1]$ and $t \leq T$. Consider the reachable set Ω_t under probability of success ρ, or the set of initial conditions x for which the probability that there exists a trajectory $\mathbf{x}_{t,x}^{u,d}$ that reaches set $\mathcal{K}$ at time T, associated with the admissible control $u \in \mathcal{U}$ and $d \in \mathcal{D}$ is at least ρ [16]

$$\begin{aligned}\mathcal{RA}_t^\rho = \Big\{ x \in \mathbb{R}^d \,|\exists u \in \mathcal{U}, d \in \mathcal{D},\ \mathbf{P}\Big[\mathbf{x}_{t,x}^{u,d}(T) \in \mathcal{K} \\ \text{and } \mathbf{x}_{t,x}^{u,d}(T) \notin \mathcal{A}\Big] > \rho \Big\}\end{aligned} \tag{5}$$

The sets $\mathcal{RA}_t^\rho$ can be characterized by using the Level Set approach

$$\mathcal{RA}_t^\rho = \left\{x \in \mathbb{R}^d, V(x,t) > \rho\right\} \tag{6}$$

with $\phi(x) := \mathbf{1}_{\mathcal{K}}(x)$. An intuitive derivation can be made if it is assumed that the value function is differentiable:

$$V(x, t) = \min_{u\in\mathcal{U}} \max_{d\in\mathcal{D}} \left\{\mathbb{E}\left[V\left(x + \Delta f(x, u, d) + \xi, t + \Delta t\right)\right]\right\} \tag{7}$$

with $\xi \sim \mathcal{N}(0, \sigma(x,t))$ a zero mean process which is normally distributed with the standard deviation a function of both state and time. Taking the Taylor expansion of the value function by means of Itô's calculus. Since dx^2 is of the order dt because of the Wiener process, the expansion must be performed up to the second order.

$$\begin{aligned}V(x + \Delta x, t + \Delta t) &\approx V(x,t) + V_t(x,t)\Delta t + V_x(x,t)\Delta x \\ &+ \frac{1}{2}\left(V_{xx}(x,t)\Delta x^2 + 2V_{xy}\Delta x \Delta t + V_{tt}(x,t)\Delta t^2\right) + \mathcal{O}(\delta^3)\end{aligned} \tag{8}$$

where $\Delta x = \Delta f(x, u, d) + \xi$. Keeping all terms of order $\mathcal{O}(\Delta t)$, we have the Dynkin Operator [14]:

$$\begin{aligned}\mathbb{E}[V] = V(x,t) + \Delta f(x,u,d)V_x(x,t) \\ + \frac{1}{2}Tr\left(\sigma^2(x,t)V_{xx}(x,t)\right)\end{aligned} \tag{9}$$

Substituting into $V(x,t)$ and dividing by Δt

$$\begin{aligned}&\frac{V(x,t) - V(x, t+\Delta t)}{\Delta t} \\ &= \min_{u\in\mathcal{U}} \max_{d\in\mathcal{D}} \left\{V_x f(x,t)^T + \frac{1}{2}Tr\left(\sigma(x,t), \sigma(x,t)^T V_{xx}\right)\right\}\end{aligned} \tag{10}$$

For the limit of $\Delta t \to 0$ the value function is characterized by the viscosity solution to the stochastic HJB equation [14]

$$\begin{aligned}&\frac{\partial V}{\partial t}(x,t) + \sup\left\{\frac{\partial V}{\partial x}(x,t) f(x,u,d)\right\} \\ &+ \frac{1}{2}Tr\left\{\sigma(x,t)\sigma(x,t)^T \frac{\partial^2 V}{\partial x^2}(x,t)\right\} = 0\end{aligned} \tag{11}$$

The stochastic HJB equation can be solved using a similar method as for the deterministic HJB equation. Next to the convection term, an additional diffusion term is added to the Level-Set (LS) method. In contrary to the deterministic reachable set, the value function in the probabilistic reachability analysis is not represented by a signed distance function. Instead, an indicator function is used [14]:

$$\mathbf{1}_{\mathcal{K}}(x) = \begin{cases} 1 & \text{if } x \in \mathcal{K} \\ 0 & \text{if } x \notin \mathcal{K}\end{cases} \tag{12}$$

The regularization of the indicator function is done by:

$$V^\varepsilon(x) = 1 - \min\left(1, \max\left(0, -\frac{1}{\varepsilon}dist(x, \mathcal{K})\right)\right) \tag{13}$$

where $V^\varepsilon(x)$ is $\frac{1}{\varepsilon}$-Lipschitz continuous.

The optimal control input for the attacking team is given by [17]:

$$u^*(\mathbf{x}, t) = \arg\min_{u\in\mathcal{U}} \max_{d\in\mathcal{D}} s(\mathbf{x},t)^T\left[f(\mathbf{x}, u, d) + \sigma(\mathbf{x}, t)\right] \tag{14}$$

where $s = \frac{\partial V}{\partial \mathbf{x}}$ and $t \in [0, T]$.

Similarly, the optimal control input for the defending team is given by

$$d^*(\mathbf{x}, t) = \arg\max_{d\in\mathcal{D}} s(\mathbf{x},t)^T\left[f(\mathbf{x}, u^*, d) + \sigma(\mathbf{x}, t)\right] \tag{15}$$

Taking $T \to \infty$, we obtain the set of initial conditions from which the attackers are guaranteed to win. We denote this set $\mathcal{RA}_\infty^\rho(\mathcal{K}, \mathcal{A})$. The set of initial conditions from which the defenders are guaranteed to win is given by all points not in $\mathcal{RA}_\infty^\rho$. For an N vs. N game on a two-dimensional domain $\Omega \subset \mathbb{R}^2$, the reachable set $\mathcal{RA}_\infty^\rho$ is 4ND.

B. The Reach-Avoid Game

In a two-player reach-avoid game, the goal of the attacker is to avoid being captured by the defender while reaching the target set $\mathcal{T}$. This reach set $\mathcal{K}$ is represented by the attacker being inside $\mathcal{T}$. On the way to $\mathcal{T}$, the attacker must avoid being captured by the defender. This is represented by the set C.

Both players need to avoid the obstacle Ω_{obs}, which can be viewed as the locations in Ω where the players has zero velocity. In particular, the defender wins if the attacker is in Ω_{obs} and vice versa. Thus, we define the reach set and the avoid set as

$$\begin{aligned}\mathcal{K} = &\left\{\mathbf{x} \in \Omega^2 \,|x_A \in \mathcal{T} \wedge \|x_A - x_D\|_2 > R_C\right\} \\ &\cup \left\{\mathbf{x} \in \Omega^2 \,|x_D \in \Omega_{obs}\right\}\end{aligned} \tag{16}$$

$$\begin{aligned}\mathcal{A} = &\left\{\mathbf{x} \in \Omega^2 \,|\|x_A - x_D\|_2 \leq R_C\right\} \\ &\cup \left\{\mathbf{x} \in \Omega^2 \,|x_A \in \Omega_{obs}\right\}\end{aligned} \tag{17}$$

Given these sets, we can define the corresponding level set representations $\mathbf{P}_{\mathcal{K}}$, $\mathbf{P}_{\mathcal{A}}$, and solve (11). If $\Omega \subset \mathbb{R}^2$, the result is $\mathcal{RA}^{\rho}_{\infty} \in \mathbb{R}^4$, a 4D reach-avoid set with the level set representation $V(\mathbf{x}, \infty)$. The attacker wins if and only if $\mathbf{x^0} = (x^0_{A_i}, x^0_{D_j}) \in \mathcal{RA}^{\rho}_{\infty}$.

If $\mathbf{x^0} \in \mathcal{RA}^{\rho}_{\infty}$, then the attacker is guaranteed to win the game. Applying Eq. (14) to the two-player game, we obtain the explicit winning strategy is given in Ref. [17]:

$$A^*(x_A, x_D, t) = -v_A \frac{s_u\left(x_A, x_D, t\right)}{\|s_u\left(x_A, x_D, t\right)\|_2} \tag{18}$$

where $s = (s_u, s_d) = \frac{\partial V}{\partial(x_A, x_D)}$.

Similarly, if $\mathbf{x^0} \notin \mathcal{RA}^{\rho}_{\infty}$, then the defender is guaranteed to win the game. Applying Eq. (12) to the two-player game, we obtain the explicit winning strategy given in:

$$D^*(x_A, x_D, t) = v_D \frac{s_d\left(x_A, x_D, t\right)}{\|s_d\left(x_A, x_D, t\right)\|_2} \tag{19}$$

C. Dynamics Flowfields

The dynamics of the flowfields are nonlinear. We let $g(x, y)$ denote the flowfields current at position (x, y) [18], [19]. We assume that the current flows with constant direction, with the magnitude of g increasing in distance from the middle of the flows:

$$g(x, y) := \begin{bmatrix} 1 + ay^2 \\ 0 \end{bmatrix} \tag{20}$$

similarly, to describe the uncertainty influenced by the Brownian motion, we consider the diffusion term

$$\sigma(x, y) := \begin{bmatrix} \sigma_x & 0 \\ 0 & \sigma_y \end{bmatrix} \tag{21}$$

We assume that attackers and defenders can change their directions α, β instantaneously. The complete dynamics of the attacker and the defender are given by

$$\begin{bmatrix} dx_A \\ dy_A \end{bmatrix} = \begin{bmatrix} 1 + ay^2 + v_A\cos(\alpha) \\ v_A\sin(\alpha) \end{bmatrix} ds + \begin{bmatrix} \sigma_x & 0 \\ 0 & \sigma_y \end{bmatrix} dW \tag{22}$$

$$\begin{bmatrix} dx_D \\ dy_D \end{bmatrix} = \begin{bmatrix} 1 + ay^2 + v_D\cos(\beta) \\ v_D\sin(\beta) \end{bmatrix} ds + \begin{bmatrix} \sigma_x & 0 \\ 0 & \sigma_y \end{bmatrix} dW \tag{23}$$

where $\alpha, \beta \in [-\pi, \pi]$ is the direction of the attacker and the defender with respect to the x axis. We can get α, β based on the term $\frac{s_u(x_A, x_D, t)}{\|s_u(x_A, x_D, t)\|_2}$ and $\frac{s_d(x_A, x_D, t)}{\|s_d(x_A, x_D, t)\|_2}$ from Eqs. (18) and (19).

Obviously, the probability of success starting from some initial position in the whole region depends on starting point (x, y). As shown in the previous Sec. III-A, this probability can be characterized as the level set of a function. We can update stochastic HJB equation based on Eq. (11) as follows

$$\begin{aligned}&\frac{\partial V}{\partial t}(x, y, t) + \sup_{\alpha \in [-\pi, \pi]} \left\{\frac{\partial V}{\partial x}(x, y, t)\left(1 + ay^2 + v_A\cos(\alpha)\right)\right. \\ &\left.+ \frac{\partial V}{\partial y}(x, y, t) v_A \sin(\alpha)\right\} + \frac{1}{2}Tr\left\{\sigma_x^2 \frac{\partial^2 V}{\partial x^2}(x, y, t)\right\} \\ &+ \frac{1}{2}Tr\left\{\sigma_y^2 \frac{\partial^2 V}{\partial y^2}(x, y, t)\right\} = 0\end{aligned} \tag{24}$$

$$\begin{aligned}&\frac{\partial V}{\partial t}(x, y, t) + \sup_{\beta \in [-\pi, \pi]} \left\{\frac{\partial V}{\partial x}(x, y, t)\left(1 + ay^2 + v_D\cos(\beta)\right)\right. \\ &\left.+ \frac{\partial V}{\partial y}(x, y, t) v_D \sin(\beta)\right\} + \frac{1}{2}Tr\left\{\sigma_x^2 \frac{\partial^2 V}{\partial x^2}(x, y, t)\right\} \\ &+ \frac{1}{2}Tr\left\{\sigma_y^2 \frac{\partial^2 V}{\partial y^2}(x, y, t)\right\} = 0\end{aligned} \tag{25}$$

It can be shown that the orientation controller value for the attacker and the defender maximizing the above Dynkin operator is

$$\begin{aligned}\alpha^*(x, y, t) :&= \arg\max_{\alpha \in [-\pi, \pi]} \left(\frac{\partial V}{\partial x}(x, y, t)\cos(\alpha)\right. \\ &\left.+ \frac{\partial V}{\partial y}(x, y, t)\sin(\alpha)\right) \\ &= \arctan_{\alpha}\left(\frac{\partial_y V}{\partial_x V}\right)(x, y, t)\end{aligned} \tag{26}$$

$$\begin{aligned}\beta^*(x, y, t) :&= \arg\max_{\beta \in [-\pi, \pi]} \left(\frac{\partial V}{\partial x}(x, y, t)\cos(\beta)\right. \\ &\left.+ \frac{\partial V}{\partial y}(x, y, t)\sin(\beta)\right) \\ &= \arctan_{\beta}\left(\frac{\partial_y V}{\partial_x V}\right)(x, y, t)\end{aligned} \tag{27}$$

Therefore, the stochastic HJB equation can be simplified to

$$\begin{aligned}&\frac{\partial V}{\partial t}(x, y, t) + \frac{\partial V}{\partial x}(x, y, t)\left(1 + ay^2\right) \\ &+ \frac{1}{2}Tr\left\{\sigma_x^2 \frac{\partial^2 V}{\partial x^2}(x, y, t)\right\} + \frac{1}{2}Tr\left\{\sigma_y^2 \frac{\partial^2 V}{\partial y^2}(x, y, t)\right\} \\ &+ v_A\|\nabla V(x, y, t)\| = 0\end{aligned} \tag{28}$$

$$\begin{aligned}&\frac{\partial V}{\partial t}(x, y, t) + \frac{\partial V}{\partial x}(x, y, t)\left(1 + ay^2\right) \\ &+ \frac{1}{2}Tr\left\{\sigma_x^2 \frac{\partial^2 V}{\partial x^2}(x, y, t)\right\} + \frac{1}{2}Tr\left\{\sigma_y^2 \frac{\partial^2 V}{\partial y^2}(x, y, t)\right\} \\ &+ v_D\|\nabla V(x, y, t)\| = 0\end{aligned} \tag{29}$$

where $\nabla V(x,y,t)=\left[\frac{\partial V}{\partial x}(x,y,t)\frac{\partial V}{\partial y}(x,y,t)\right]$.

IV. Database Building and Online Implementation

A. Database Building

HJ reachability analysis is computationally intractable when the total number of attackers and defenders is greater than three. Therefore, based on the 1 vs. 1 reach-avoid results, a database containing offline HJ reachability for different 1 vs. 1 scenarios can be designed and established. Based on our previous work on safe flight envelope prediction system [20], we propose to use a database to maximize the number of attackers that can be assigned to a defender. This database-based extension is defined as follows:

$$\text{maximize} \quad \sum_{i,j} e_{i,j}, \quad e_{i,j} \subseteq E_{(e_{i,j})} \tag{30a}$$

$$\text{subject to} \quad e_{i,j} \in \{0,1\}, \quad \forall i,j \tag{30b}$$

$$\sum_{i} e_{i,j} \leq 1, \quad \forall j \tag{30c}$$

$$\sum_{j} e_{i,j} \leq 1, \quad \forall i \tag{30d}$$

$$e_{i,j} = 0, \quad \forall i \in \Xi_j, \forall j \tag{30e}$$

$$e_{i,j} = 1, \quad \forall i, \forall j \in \Xi_i \tag{30f}$$

The database $E_{(e_{i,j})}$ is determined by pairs of two-player reach-avoid games in the database, where $e_{i,j}$ denotes the capture relationship between the attacker A_i and the defender D_j in the database. When $e_{i,j} = 1$, it indicates that the defender D_j is assigned to capture the attacker A_i and $e_{i,j} = 0$ otherwise.

We will then construct the constraints for this database. Firstly, the constraint (30c) limits the total number of attackers assigned to the defender D_j at any time to at most one. Secondly, the constraint (30d) ensures that an attacker can be captured by at most one defender to avoid the case where multiple defenders are assigned to only one attacker. This upper bound is limited by the computational power of the HJ reachability analysis.

Let Ξ_j be the set of all attackers that beat the defender D_j in a 1 vs. 1 game in this database. Ξ_i be the set of all defenders that capture an attacker A_i in a 1 vs. 1 game. The constraint (30e) ensures that by directly setting $e_{i,j} = 0$, a defender Dj cannot be assigned to an attacker A_i, i.e. it cannot win a 1 vs. 1 game. The constraint (30f) ensures that by directly setting $e_{i,j} = 1$, a defender D_j can be assigned to an attacker A_i, i.e. it can win a 1 vs. 1 game. The set Ξ_j and Ξ_i can be easily constructed using the pre-computed reach-avoid set $\mathcal{RA}^{\rho}_{\infty}$ in the database. Thus, if the j-th defender D_j is assigned to defend the i-th attacker A_i with maximum matching, the strategy based on Eqs. (18) and (19) that guarantees that A_i never reaches the target is

$$D_j^*(x_{A_i}, x_{D_j}, t) = v_D \frac{s_d\left(x_{A_i}, x_{D_j}, t\right)}{\|s_d\left(x_{A_i}, x_{D_j}, t\right)\|_2} \tag{31}$$

Pairs of 1 vs. 1 reach-avoid game results can be easily computed, but adding single versus multiple pairs of HJ reachability analysis computations may be difficult to handle. Therefore, it is important to compute as many 1 vs. 1 reach-avoid game results as possible so that the database can provide as much information as possible while maintaining a reasonable amount of data for a defender to capture multiple attackers.

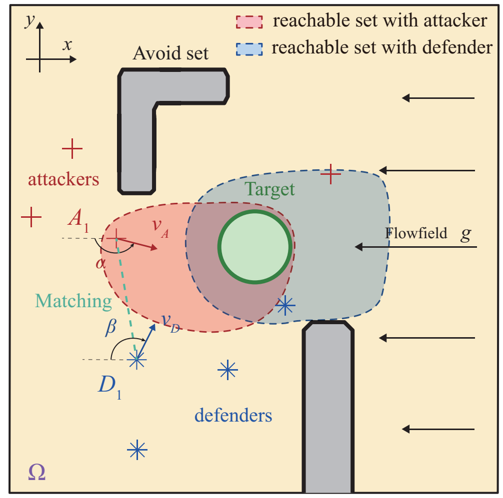


Fig. 2. A multiplayer reach-avoid game with dynamics Flowfield and maximum matching. The Flowfield $g(x,y)$ can be considered as a dynamics disturbance. The attacker and the defender must overcome the disturbance of Flowfield to complete their own tasks.

B. Online Maximum Matching

Our database building solution in Sec. IV-A determines the capture matching of an attacker to each defender that is part of a maximum pairing based on the current state of all players, and the matching does not change until the current capture task is completed. However, the bipartite graph and its maximum matching values can be updated online as player's states change over time in the game. In the online maximum matching problem, we have an uncertain bipartite graph $\mathcal{G}$. In addition to needing to complete the capture task that have already been matched, a defender must perform online rematching to complete additional capture task. Therefore, we iteratively perform matching updates based on the information in the database after each time step to account for these changes. Our online maximum matching algorithm can be summarised as shown in Algorithm 1.

When $\Delta t \rightarrow 0$, the offline database building process computes a bipartite graph and its maximum matching as a function of time. When the maximum matching is not unique in the database, the defender can choose a different maximum matching and still be guaranteed to prevent the same number of attackers from reaching the target.

On the other hand, it is possible that the size of the maximum matching increases with time. This can happen if the joint configuration of the players causes the generated bipartite graph to have a larger maximum matching value than before, since the size of the maximum matching only gives an upper bound on the number of attackers that can reach the target. Therefore, we design the online rematching process

Algorithm 1 Online Maximum Matching

Require: $\mathcal{RA}_\infty^\rho$

Initialization: states of all players ($\mathbf{x}$)
Offline database
for Maximum capacity of the database **do**
  Update new player positions
  Construct a bipartite graph with two sets of nodes A_i, D_j
  Determine whether D_j can win against A_i, for all j
  Construct the bipartite graph $\mathcal{G}$ and find its maximum matching
  For duration Δt, the optimal control inputs and trajectories for each attacker and defender are computed via Eq. (31)
end for
Online rematching
if the number of attackers increase **then**
  while not all attackers captured **do**
    Determine the constraint set Ξ_j and Ξ_i
    Apply optimal controls for defenders via database
    Update states of all attackers and defenders
  end while
else
  Use offline database directly
end if

based on the offline database. A defender, after completing the previous capture task, uses the samples in the database to recalculate the next attacker that can be captured and rematches with it until there are no capturable attackers.

Due to the continuity and optimality of the probabilistic HJ controller, the defender is able to ensure the success of each capture task with high probability. Thus, our method guarantees that the maximum number of attackers reaching the target does not increase as the defender faces more attackers. Instead, it decreases as the game plays out and new attackers are assigned. Hence, the whole schematic of our method is shown in Fig. 2.

V. Simulation

For the following numerical simulations we fix the diffusion term $\sigma_x = 0.5$ and $\sigma_y = 0.2$ in (21), and consider a non-uniform dynamics flow with $a = 0.01$ in (20). We illustrate our stochastic reachability in the example below. The probabilistic HJI reachable sets are calculated using the Level Set Toolbox [21]. We calculate the reachable set by incrementing T until convergence. The computation is done on a 4D grid with 50 nodes on each dimension. The example is shown in Fig. 3. There are four attackers and four defenders playing on a square domain with obstacles and all players have equal speeds. To visualize the 4D set in 2D, we view the RA set at the slices representing the initial positions of different players. Unlike deterministic RA games, where reachable sets decides the winning or losing conditions of the players, the probabilistic reachable sets indicate the probability of winning under the influence of stochastic process on the RA game.

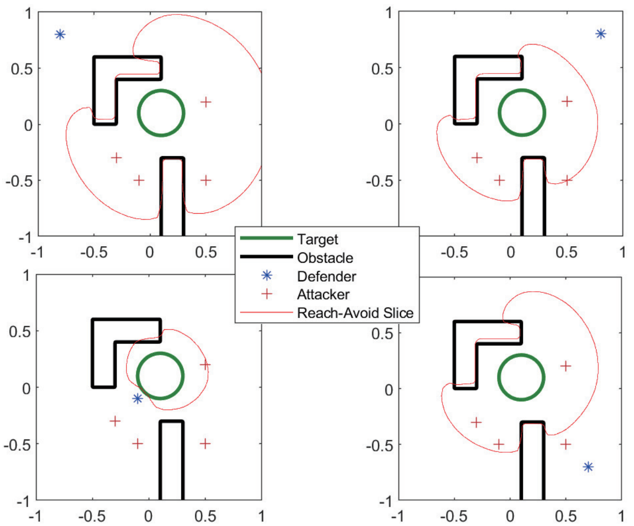
In each subplot of Fig. 3(a), the 2D slice shows the probabilistic "winning" regions ($\rho = 0.8$) of the attacker with respect to each defender. Attackers whose initial positions lie in the region have the probability $\rho = 0.8$ of winning against the defender. For example, in the upper left subplot of Fig. 3(a), since the defender is too far away from the target, the probability of losing to all four attackers is 0.8, while in the lower left subplot, the defender wins against three of the four attackers with the probability of 0.8 since it locates near the target. Hence, the shape of the probabilistic reachable sets in Fig. 3(a) is determined by the initial positions of four defenders. Similarly, Fig. 3(b) shows the probabilistic "winning" regions of the defenders, in which the defenders are guaranteed to win against the attacker with high probability of 0.8. As shown in the upper right subplot of Fig. 3(b), for example, only one of the four defenders have the high probability of winning the attacker due to its short distance to the target.

Figure 4 shows the bipartite graph (thin green solid line) and the maximum matching value (thick green dashed line) obtained after applying the algorithm from the offline database described in Sec. IV-A. The maximum matching size is 4, which means that no attacker can reach the target.

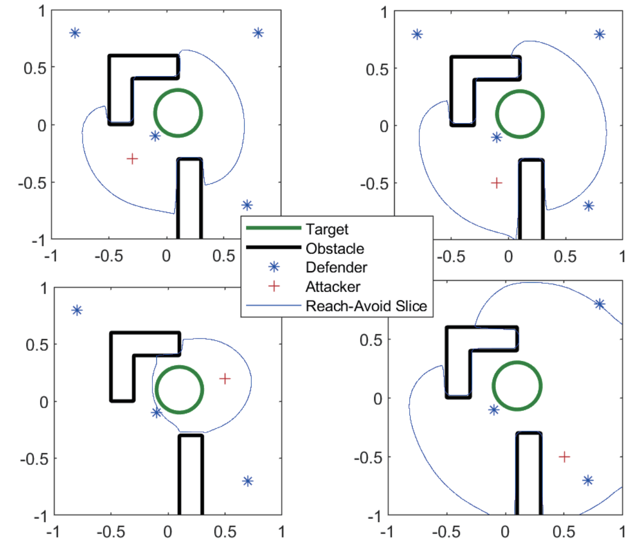
Figure 5(a) shows the results of a 4 vs. 4 reach-avoid game simulation performed over the course of 0.8 time units to illustrate the pairing of defender-attacker in an offline database. The probabilistic HJ reachability analysis makes the Δt time units vary continuously to compute a bipartite graph and its maximum matching. The initially unmatched defender D_2 plays optimally against the closest attacker according to Eq. (31). The attacker adopts a suboptimal strategy by choosing the shortest path to reach the target while avoiding obstacles and disregarding control inputs from other players in the game. The green dotted line in Fig. 5(a) indicates the matching relationship between the defender and the attacker at this moment.

At $t = 0$ (top left graph), the initial size of the maximum matching is 3. At $t = 0.2$ (top right graph), the player moves further. At $t = 0.5$ (bottom left graph), attacker A_2 is in a losing position against defender D_2 because the attacker is not playing optimally. As a result, the maximum matching now assigns attacker A_2 to defender D_2. In this way a perfect matching is created, preventing any attacker from reaching the target. Finally the whole process ends at $t = 0.8$ (bottom right graph) and the results are recorded in the offline database.

Figure 5(b) shows the 2 vs. 4 probabilistic HJ reach-avoid game scenario. The initial size of the maximum match at $t = 0$ (top left graph) is 2. Using the optimal control obtained by Eq. (31) through the offline database in our method, using the computed reach-avoid set $\mathcal{RA}_\infty^\rho$ defender D_1 is assigned to capture attacker A_1 and defender D_2 is assigned to capture attacker A_3. At $t = 0.525$ (bottom left graph), the attacker A_1 and A_3 are captured, respectively. Meanwhile, the offline database continues to be used to reassign defenders and the remaining uncaptured attackers, and the maximum matching is updated online. At $t = 0.875$ (bottom right graph), the remaining attackers are also successfully captured by the defender.



(a) Reachable sets w.r.t four different initial positions of defenders.



(b) Reachable sets w.r.t four different initial positions of attackers.

Fig. 3. 2D slices of reachable sets in a 4 vs. 4 stochastic reach-avoid game.

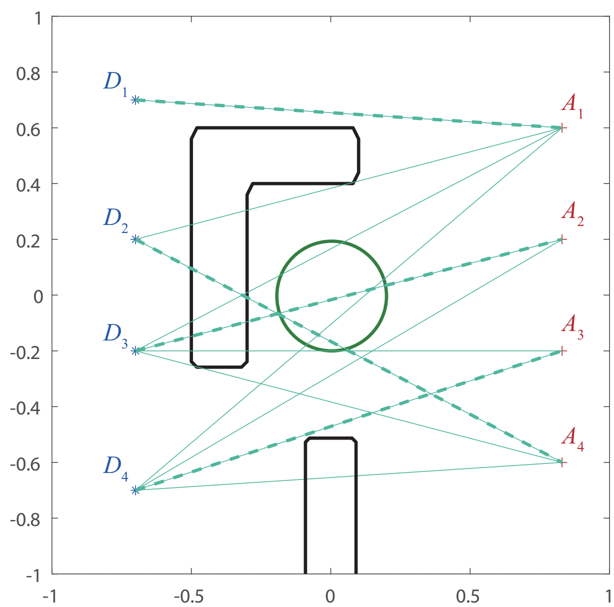


Fig. 4. Bipartite graph and maximum matching results. Each edge (thin green solid line) connects a defender to an attacker, and the other defender is guaranteed to win, creating a bipartite graph. A maximum matching of size 4 (thick green dashed line) means that no attacker can reach the target.

In contrast, with the baseline method in the offline database, a defender can only capture a attacker as shown in Fig. 5(a). Since any 1 vs. 1 task in the offline database is valid, initially the defender is assigned to capture the attacker with the highest current capture probability. Through multiple rematchings, the defenders complete all capture tasks with high probability.

VI. Conclusion

This extended abstract gives a brief introduction to the work we've been doing on probabilistic reachability analysis in stochastic reach-avoid games, and we extend this work further. Compared to the current offline maximum matching, we propose online maximum matching by constructing an offline database and online rematching. In the simulation part, we first construct the offline database. By solving the 4D HJI PDE, we can determine the probability of winning for both attackers and defenders in optimal conditions through the boundaries of probabilistic reachable sets. Further, we use the offline database to implement the time-varying multiple players game by our proposed online maximum matching. Future work will focus on the extension to the offline database to approximate high-dimensional games by neural-network in order to achieve high-dimensional game storage for our offline database.

References

[1] M. G. Earl and R. D'Andrea, "A decomposition approach to multi-vehicle cooperative control," *Robotics and Autonomous Systems*, vol. 55, no. 4, pp. 276–291, 2007.

[2] J. S. McGrew, J. P. How, B. Williams, and N. Roy, "Air-combat strategy using approximate dynamic programming," *Journal of guidance, control, and dynamics*, vol. 33, no. 5, pp. 1641–1654, 2010.

[3] R. Isaacs, "Differential games, new york, 1965," *Translated under the title Differentsial'nye igry, Moscow: Mir*, 1967.

[4] M. Chen, Z. Zhou, and C. J. Tomlin, "Multiplayer reach-avoid games via pairwise outcomes," *IEEE Transactions on Automatic Control*, vol. 62, no. 3, pp. 1451–1457, 2016.

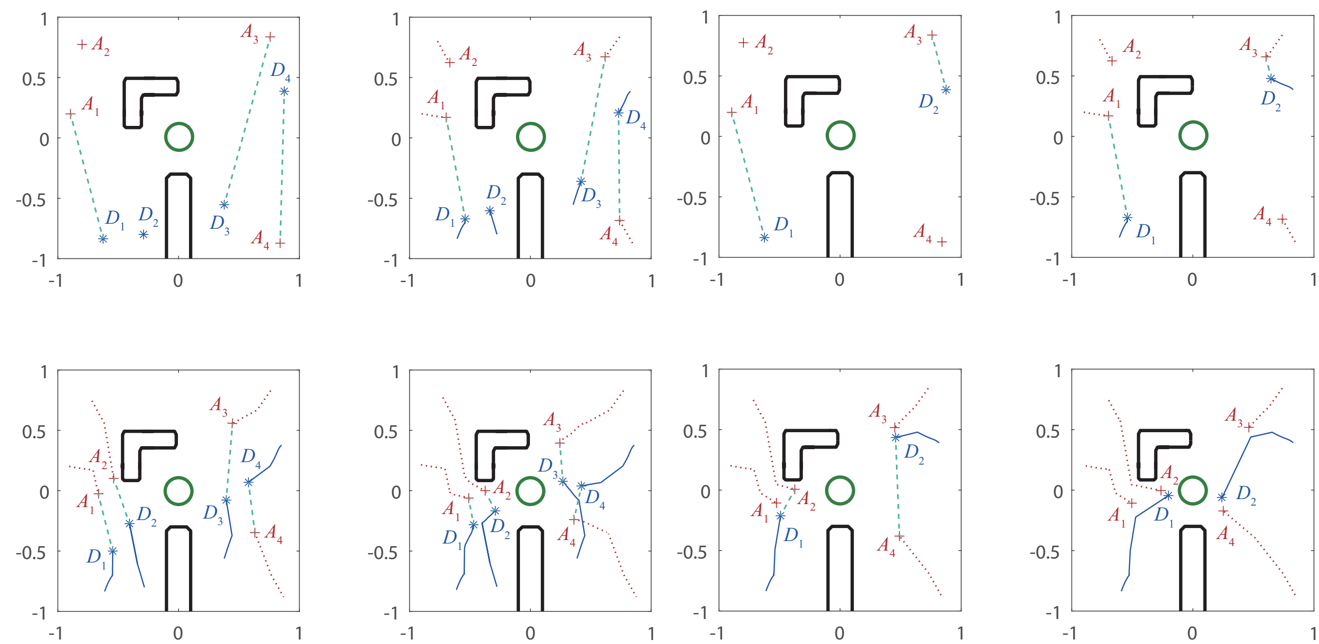
[5] R. Yan, X. Duan, Z. Shi, Y. Zhong, and F. Bullo, "Matching-based capture strategies for 3d heterogeneous multiplayer reach-avoid differential games," *Automatica*, vol. 140, p. 110207, 2022.

[6] M. Chen, Z. Zhou, and C. J. Tomlin, "A path defense approach to the multiplayer reach-avoid game," in *53rd IEEE conference on decision and control*. IEEE, 2014, pp. 2420–2426.

[7] A. Pierson, Z. Wang, and M. Schwager, "Intercepting rogue robots: An algorithm for capturing multiple evaders with multiple pursuers," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 530–537, 2016.

[8] M. Chen, J. C. Shih, and C. J. Tomlin, "Multi-vehicle collision avoidance via hamilton-jacobi reachability and mixed integer programming," in *2016 IEEE 55th Conference on Decision and Control (CDC)*. IEEE, 2016, pp. 1695–1700.

[9] M. Chen, Z. Zhou, and C. J. Tomlin, "Multiplayer reach-avoid games via low dimensional solutions and maximum matching," in *2014 American control conference*. IEEE, 2014, pp. 1444–1449.



(a) (b)

Fig. 5. (a) The case of 1 vs. 1 maximum matching in offline databases. Defenders can calculate the optimal results by probabilistic HJ reachability analysis. (b) The case of 1 vs. 2 maximum matching in online rematching via the offline database. The defender can update the bipartite graph and the maximum matching in real time using the algorithm described in Sec. IV-B. The defender finally completes the capture task.

[10] A. Abate, M. Prandini, J. Lygeros, and S. Sastry, "Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems," *Automatica*, vol. 44, no. 11, pp. 2724–2734, 2008.

[11] J. P. Hespanha, C. G. Cassandras, H. A. Blom, J. Lygeros, J.-P. Katoen, M. Prandini, A. van der Schaft, Y. Wardi, and P. J. Mosterman, *Stochastic hybrid systems*. CRC Press, 2018.

[12] A. Lavaei, S. Soudjani, A. Abate, and M. Zamani, "Automated verification and synthesis of stochastic hybrid systems: A survey," *Automatica*, vol. 146, p. 110617, 2022.

[13] K. Leung, E. Schmerling, M. Zhang, M. Chen, J. Talbot, J. C. Gerdes, and M. Pavone, "On infusing reachability-based safety assurance within planning frameworks for human–robot vehicle interactions," *The International Journal of Robotics Research*, vol. 39, no. 10-11, pp. 1326–1345, 2020.

[14] P. M. Esfahani, D. Chatterjee, and J. Lygeros, "The stochastic reach-avoid problem and set characterization for diffusions," *Automatica*, vol. 70, pp. 43–56, 2016.

[15] R. van den Brandt and C. de Visser, "Safe flight envelope uncertainty quantification using probabilistic reachability analysis," *IFAC-PapersOnLine*, vol. 51, no. 24, pp. 628–635, 2018.

[16] M. Assellaou, O. Bokanowski, and H. Zidani, "Error estimates for second order hamilton-jacobi-bellman equations. approximation of probabilistic reachable sets," *Discrete and Continuous Dynamical Systems-Series A*, vol. 35, no. 9, pp. 3933–3964, 2015.

[17] H. Huang, J. Ding, W. Zhang, and C. J. Tomlin, "A differential game approach to planning in adversarial scenarios: A case study on capture-the-flag," in *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011, pp. 1451–1456.

[18] W. Sun, P. Tsiotras, T. Lolla, D. N. Subramani, and P. F. Lermusiaux, "Multiple-pursuer/one-evader pursuit–evasion game in dynamic flowfields," *Journal of guidance, control, and dynamics*, vol. 40, no. 7, pp. 1627–1637, 2017.

[19] P. Cardaliaguet, M. Quincampoix, and P. Saint-Pierre, "Optimal times for constrained nonlinear control problems without local controllability," *Applied Mathematics and Optimization*, vol. 36, pp. 21–42, 1997.

[20] Y. Zhang, C. de Visser, and Q. Chu, "Database building and interpolation for an online safe flight envelope prediction system," *Journal of Guidance, Control, and Dynamics*, vol. 42, no. 5, pp. 1166–1174, 2019.

[21] I. M. Mitchell *et al.*, "A toolbox of level set methods," *UBC Department of Computer Science Technical Report TR-2007-11*, vol. 1, p. 6, 2007.