Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

ARTICLE TYPE

Please insert your article title here

Ye Zhang[1,2] | **Yutong Zhu**[2]

[1]Research & Development Institute of Northwestern Polytechnical University in Shenzhen, Shenzhen, China

[2]School of Astronautics, Northwestern Polytechnical University, Xi'an, China

Correspondence
Corresponding author: Ye Zhang.
Email: zhang_ye@nwpu.edu.cn

Present address
127 Youyi West Road, 710072, Xi'an, China.

Abstract

A deterministic and stochastic hybrid dynamics system is essential for dealing with the complex behaviour and stochastic optimal control among multiple agents in a reach-avoid game. Considering the heterogeneous interactions among multiple agents of a complex system, many challenges remain in implementing such a hybrid system. In this paper, a deterministic and stochastic dynamics model for the reach-avoid games is decoupled into a non-cooperative probabilistic game and a cooperative probabilistic game. Nash equilibrium and finite-time stability with collision free are guaranteed in the non-cooperative probabilistic game that is based on the strategy from zero-sum game analysis. In the cooperative probabilistic game, the DCFCI strategy is facilitated to achieve collision-free and cross-interception by using probabilistic reachability analysis to reveal the high-probability interception region presented in this paper. The feasibility of the proposed two probabilistic games is demonstrated by simulation results, where the stationed and unstationed attacker cases are successfully detected and implemented.

KEYWORDS

Reach-avoid games, Non-cooperative and cooperative probabilistic games, Probabilistic reachability, Nash equilibrium

1 | INTRODUCTION

With the development of autonomy, unmanned aerial vehicles (UAVs), especially quadrotors, have been more and more involved in applications that cannot be easily accomplished by human beings [1,2], such as search and rescue, reconnaissance, resource exploration, forest fire prevention and military missions [3,4,5]. Nevertheless, numerous reports indicate that UAVs operating in unauthorized airspace are unlikely to complete their task effectively without stable and efficient decision-making. It is therefore essential to develop effective decision-making frameworks to ensure the safe and efficient operation of UAVs in complex environments, without violating no-fly zones. Furthermore, complex systems such as air traffic and infrastructure often exhibit complex behaviours arising from heterogeneous interactions [6,7], which are essentially hybrid in nature. The uncertainty of interleaved discrete and continuous evolution also leads to the emergence of stochastic systems models, which contain both deterministic and stochastic components. Since the reach-avoid game problem deals with the determination of initial states set in the deterministic hybrid system, where at least one control strategy can be found to steer the system to a target set while guaranteeing collision avoidance [8,9]. A hybrid system of deterministic and stochastic dynamics are often more challenging than a deterministic hybrid system. Therefore, this paper mainly studies the reach-avoid game problem for deterministic and stochastic dynamics system.

The reach-avoid game problem involving a defender and an attacker has been studied in the literature as a zero-sum differential game using different approaches including optimal control [10,11], reachability analysis [12] and model predictive control [13]. Due to the curse of dimensionality, these methods extend the 1-vs-1 reach-avoid game to multi-agent systems using a "divide and conquer" approach. For a given multi-agent systems, all possible 1-vs-1 non-cooperative games between the defender and the attacker are solved first. Next, this is extended to cooperative games with multi-agent distributed control.

In Ref. [14], optimal coordinated control policies have been synthesized, demonstrating the ability of two camera-equipped UAVs to cooperatively track a moving ground target, resulting in both reduced geolocation error and distance coordination. Conversely,

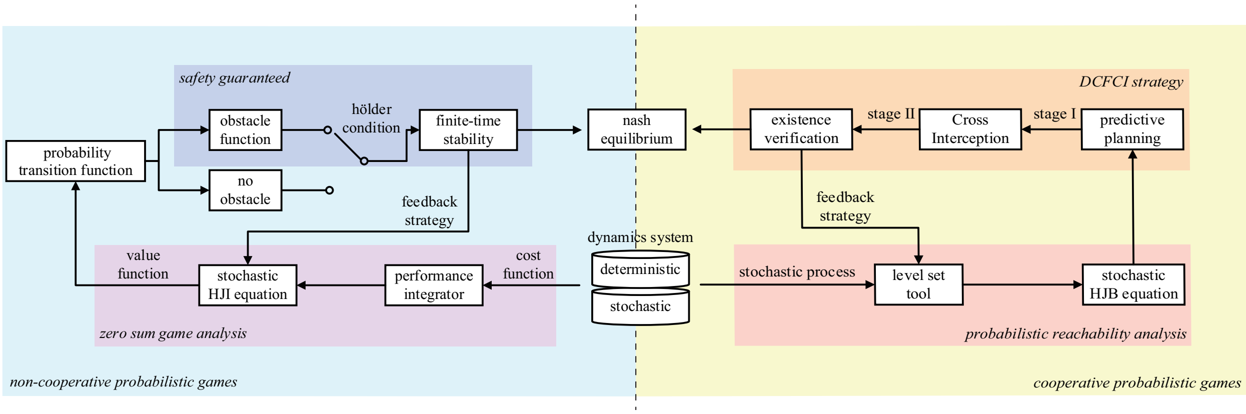


FIGURE 1 An overview of deterministic and stochastic dynamics model for reach-avoid games.

a substantial body of research has been dedicated to pursuit-evasion games. Bakolas and Tsiotras[15] propose a problem in which the pursuer aims to capture the evader. However, the aforementioned framework does not consider the presence of obstacles. Most non-cooperative games have been studied in a deterministic setting, ignoring the uncertainty of system dynamics and the influence of the environment[12]. Therefore, there is a need to extend the current solutions for non-cooperative games to a more general framework which takes into account the stochastic flowfield disturbances of the system dynamics. In[8], a link between the stochastic reach-avoid game problem and the stochastic optimal control problem involving discontinuous payoff functions was established. However, the approach in Ref.[8] does not take into account the interaction among multiple agents in a game. In Ref.[16], the authors consider a multiplayer reach-avoid game in three dimensions involving heterogeneous players. They initially decompose the original multiplayer game into a series of smaller games involving coalitions of three or fewer pursuers and only one evader, employing a Hamilton-Jacobi-Isaacs (HJI) equation. Subsequently, a sequential matching problem on a conflict graph comprising the pursuers' coalitions and evaders as nodes is solved in order to identify an approximate assignment that assigns the coalitions of pursuers to capture evaders. The aforementioned studies provide useful insights into the problem of the reach-avoid games, but their application is limited by the lack of consideration of collisions among multiple agents.

As the number of agents increases, it becomes increasingly difficult to control these agents centrally. Therefore, research on the control of multi-agent systems has gradually shifted to a distributed approach[17,18,19]. One of the research goals of distributed agents is to study cooperative game behaviour[20] when they need to deal with a task, including generalized Nash equilibrium[21]. However, for the reach-avoid problems, which consider both deterministic and stochastic dynamic models, few existing studies consider multi-agent systems with cooperative game behaviour and competitive behaviour. In addition, many current studies focus on trajectory planning in a stationed or unstationed agent environment[22], including our recent work on trajectory optimization[23]. In particular, some of the popular works[24,25] add game-theoretic methods to the prediction, planning and control of cooperative or competitive multi-agent reach-avoid games, because game theory is a suitable tool for modelling the interactions among multiple agents[26]. On the other hand, for distributed multi-agent reach-avoid games, stochastic reachability analysis[6,27] is an effective method for solving the safety constraints and performance of stochastic dynamic systems, and it provides a good mathematical framework. The theoretical framework for stochastic reachability is based on dynamic programming[28], which can effectively establish the connection between stochastic reachability problems and stochastic optimal control problems with discontinuous cost functions[8]. Therefore, using our previous work on Probabilistic Reachability Analysis (PRA) in stochastic reachability games[29], PRA is combined with competitiveness analysis with predictive planning in cooperative games, and propose a Defender Collision-Free Cross-Interception (DCFCI) strategy to deal with the defender game optimization problem in the reachability game. PRA is used to characterize the defender's high-probability interception of the attacker's region. The probabilistic reachability results are then extended to use DCFCI strategy of the defender team to intercept the attacker team to prevent it from reaching the target area.

Figure 1 shows an overview of the deterministic and stochastic dynamics model for the reach-avoid games in this paper. The contributions and innovations of our work are:

(i) The hybrid deterministic and stochastic reach-avoid game problem is decoupled into a Non-Cooperative Probabilistic Game (NCPG) between the defender and the attacker and a Cooperative Probabilistic Game (CPG) that includes collision avoidance, game analysis and reachability analysis.

(ii) A zero-sum game analysis is conducted in a NCPG. A stochastic HJI equation is constructed using the performance integrator. Finite-time stability is considered separately with and without obstacles by introducing a probability transition function. The strategy is updated in continuous iterations to obtain a Nash equilibrium solution.

(iii) The high probability interception region that is based on PRA using stochastic processes in CPG is calculated by establishing a stochastic Hamilton-Jacobi-Bellman (HJB) equation using the Level Set method. The proposed DCFCI strategy is used to transition from competitiveness analysis with predictive planning in stage I to complete collision free cross-intersection in stage II. The continuous iterative feedback strategy verifies the feasibility of the strategy to obtain a Nash equilibrium solution.

The paper has the following outline. In Sec. 2, a deterministic and stochastic hybrid system for reach-avoid games is given. In Sec. 3, the process of NCPG is introduced based on the zero-sum game analysis and safety guaranteed. The CPG method using the PRA and DCFCI strategy is proposed in Sec. 4. The method is demonstrated and verified through simulation results shown in Sec. 5. Finally, in Sec. 6, conclusions and future work are briefly introduced.

2 | PRELIMINARIES AND PROBLEM DESCRIPTION

2.1 | Problem description

In this section, a multi-agent reach-avoid game is considered among a team of N defenders, and a team of N attackers. Each agent is confined in a bounded, open domain $\Omega \subset \mathbb{R}^n$, which can be partitioned as a set of obstacles Ω_{obs} and free space Ω_{free}. Let $x_D, x_A \in \mathbb{R}^n$ be the state vector of the defender and attacker, respectively, which represent the position of each defender and attacker. Initial conditions of the agents are denoted by $x^0_{D_i}, x^0_{A_i} \in \Omega_{free}, i = 1, 2, \ldots, \mathbb{N}$. In 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 $t \geq 0$:

$$\begin{aligned} \dot{x}_{D_i}(t) &= v_D d_i(t), \quad x_{D_i}(0) = x^0_{D_i} \\ \dot{x}_{A_j}(t) &= v_A a_j(t), \quad x_{A_j}(0) = x^0_{A_j} \end{aligned} \tag{1}$$

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

$$\begin{aligned} \dot{x}_{D_i}(t) &= v_D d_i(t) + \sigma\left(x_{D_i}, d_i\right) dW, \quad x_{D_i}(0) = x^0_{D_i} \\ \dot{x}_{A_j}(t) &= v_A a_j(t) + \sigma\left(x_{A_j}, a_j\right) dW, \quad x_{A_j}(0) = x^0_{A_j} \end{aligned} \tag{2}$$

The agents' joint state and joint initial condition become $\mathbf{x} = (x_{D_i}, x_{A_j})$, $\mathbf{x^0} = (x^0_{D_i}, x^0_{A_j})$ respectively. In this reach-avoid game, the attacking team aims to reach the safe set as shown in Fig. 2, 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 agents. In this paper, the capture set is defined to be $C_{ij} = \left\{\mathbf{x} \in \Omega^{2N} \left| \|x_{D_i} - x_{A_j}\|_2 \leq R_C \right.\right\}$, the interpretation of which is that an attacker is captured by a defender if their relative distance is within R_C [12]. Next, Consider a reach-avoid stochastic process $\mathbf{x}_{t,x}$

$$\begin{aligned} \dot{\mathbf{x}} &= \mathcal{P}\left(\mathbf{x}, u, d\right) + \sigma\left(\mathbf{x}, t\right) \\ &= f(x) + g(x)u(t) + k(x)d(t) + \sigma\left(x, t\right), \quad t \geq 0 \end{aligned} \tag{3}$$

where u is the control input of the defender, d is the control input of the attacker, and f, g, k are continuous action functions on $\mathbb{R}^n$ so as to respectively maximize and minimize the probability of finishing the game at the finite time. Consider that $\Omega_{obs} = \bigcup_{i=1}^{N^O} \mathcal{O}_i$ is the region occupied by the obstacles, where N^O being the fixed number of obstacles, $\mathcal{O}_i = \{x \in \Omega : O_i(x - q_i) \leq c_i\}$ being the region occupied by each obstacle i, where $q_i \in \Omega$ is a random variable generated from a marked Binomial point process [30], representing the center of obstacle, and O_i being a Hölder condition function determining the type of the obstacle, and c_i a positive constant setting the size of the obstacle. In each game initialization, the agents have available information provided by the tuple $(q_i, e_i, O_i(x - q_i))$, where e_i is a finite set of points (cluster) associated with the obstacle i. Note that the terms $O_i(x - q_i)$ and e_i operate as marks attached to the point location q_i.

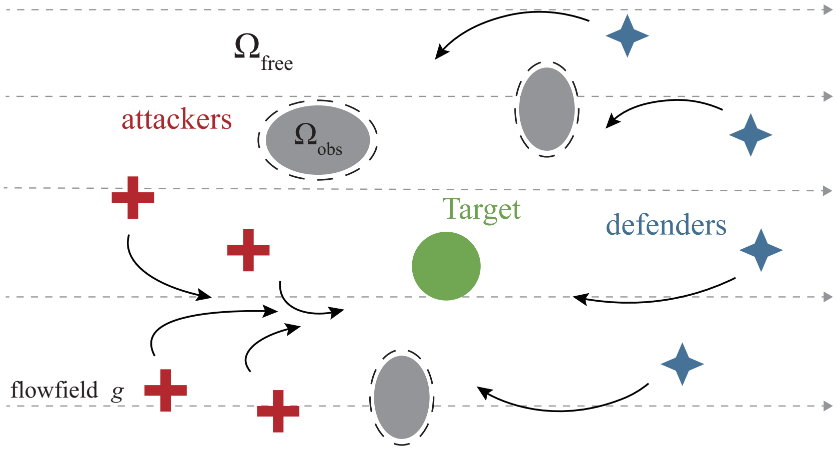


FIGURE 2 Illustration of a multi-agent reach-avoid game using game theory and reachability analysis based on dynamics flowfields.

3 | NON-COOPERATIVE PROBABILISTIC GAMES WITH STATIONED ATTACKER

3.1 | Zero-sum game analysis

Considering the system (3) with a cost function given by

$$J\left(\mathbf{x^0}, u(\cdot), d(\cdot)\right) = \int_0^{\infty} L\left(x, u(t), d(t)\right)\, dt \tag{4}$$

with the performance integrand

$$L(x, u, d) = L_1(x) + u^T R_u u - d^T R_d d \tag{5}$$

where $L_1 : \mathbb{R}^n \to \mathbb{R}$ is continuous on $\mathbb{R}^n$ with $L_1(0) = 0$ and $R_j \succ 0, j \in \{u, d\}$. We shall now provide the following definitions adopted from[31,32] to introduce the concept of finite-time stability.

Definition 1. [31] The zero solution $x(t) = 0$ to Eq. (3) with $u(t) = 0$ and $d(t) = 0$ is finite-time stable if it is Lyapunov stable and finite-time convergent, i.e., for all $x_0 \in \mathcal{N}\backslash\{0\}$, where $\mathcal{N} \subseteq \mathbb{R}^n$ is some open neighborhood of the origin, $\lim_{t\to T(x(0))} x(t) = 0$, where $T(\cdot)$ is the probability transition function such that $T(x_0) \in [0, 1]$, $x_0 \in \mathcal{N}$. The zero solution $x(t) = 0$ to Eq. (3) with $u(t) = 0$ and $d(t) = 0$ is globally finite-time stable if it is finite-time stable with $\mathcal{N} = \mathbb{R}^n$.

Definition 2. [32] Let $\mathcal{A} \subset \mathbb{R}^n$ be a compact set such that $0 \in \mathcal{A}$ and let $\mathcal{N} \subseteq \mathbb{R}^n$ be an open neighborhood of $\mathcal{A}$. The compact set $\mathcal{A}$ is finite-time attractive with respect to Eq. (3) with $u(t) = 0$ and $d(t) = 0$ if for every $x_0 \in \mathcal{N}$, the solution $x(t)$, $t \geq 0$, satisfies $dist(x(t), \mathcal{A}) = 0$, $t \geq T(x_0)$. Furthermore, the compact set $\mathcal{A}$ is globally finite-time attractive if it is finite-time attractive with $\mathcal{N} = \mathbb{R}^n$.

Now, for every initial condition $x_0 \in \mathbb{R}^n$, define $\mathcal{F}_u(x_0)$ and $\mathcal{F}_d(x_0)$ to be the set of feedback strategies of the defender and the attacker, respectively, and define $\mathcal{F}(x_0) = \{u(\cdot) \in \mathcal{F}_u(x_0), d(\cdot) \in \mathcal{F}_d(x_0)\}$: $x(\cdot)$ given by Eq. (3) satisfies $x(t) \to 0$ as $t \to T(x_0)$ to be the set of pairs $(u(\cdot), d(\cdot))$ of finite-time stabilizing control strategies. Furthermore, let $u(\cdot) \in \mathcal{U}_D(x_0) \subseteq \mathcal{F}_u(x_0)$ and $d(\cdot) \in \mathcal{U}_A(x_0) \subseteq \mathcal{F}_d(x_0)$ be such that $\mathcal{U}_D(x_0) \times \mathcal{U}_A(x_0) = \mathcal{F}(x_0)$, $x_0 \in \mathbb{R}^n$. Next, our problem is stated as a finite-time stable non-cooperative games.

Considering the system (3) end time evolution of non-cooperative games associated with the performance measure (4) wherein the defender seeks to minimize it, whereas the attacker seeks to maximize it. The objective amounts to finding a globally finite-time stabilizing saddle point solution $(u^*(\cdot), d^*(\cdot)) \in \mathcal{F}(x_0)$, rendering the equilibrium point of the closed-loop system (3) with $u(\cdot) = u^*(\cdot)$ and $d(\cdot) = d^*(\cdot)$ globally finite-time stable while $(u^*(\cdot), d^*(\cdot))$ is a saddle point of the performance index (4) for every $x_0 \in \mathbb{R}^n$. 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. The value of the game stems from the min-max problem

$$V(x_0, t) = \max_{d(\cdot)\in\mathcal{U}_A(x_0)} \min_{u(\cdot)\in\mathcal{U}_D(x_0)} \mathbb{E}\left[J\left(x_0, u(\cdot), d(\cdot)\right)\right],\ x_0 \in \mathbb{R}^n \tag{6}$$

and the Hamilton function is defined as

$$H\left(x,u,d,\dot{V}(x,t)^T\right) = L(x,u,d) + \dot{V}(x,t)\left[\mathcal{P}(x,u,d) + \sigma(x,t)\right] \tag{7}$$

Applying the stationarity conditions to the Hamilton function (7), the pair of feedback Nash strategies are obtained

$$\begin{aligned} u^*(x) &= \arg\min H\left(x,u,d,\dot{V}(x,t)^T\right) \\ &= -\frac{1}{2}R_u^{-1}g(x)^T\dot{V}(x,t)^T \end{aligned} \tag{8}$$

and

$$\begin{aligned} d^*(x) &= \arg\max H\left(x,u,d,\dot{V}(x,t)^T\right) \\ &= \frac{1}{2}R_d^{-1}k(x)^T\dot{V}(x,t)^T \end{aligned} \tag{9}$$

Substituting Eqs. (8) and (9) into Eq. (7), the HJI equation is obtained

$$\begin{aligned} L_1(x) + \dot{V}(x,t)\left[f(x)+\sigma(x)\right] - \frac{1}{4}\dot{V}(x,t)g(x)R_u^{-1}g(x)^T\dot{V}(x,t)^T + \frac{1}{4}\dot{V}(x,t)k(x)R_d^{-1}k(x)^T\dot{V}(x,t)^T = 0 \\ \Rightarrow H\left(x,u^*,d^*,\dot{V}(x,t)^T\right) = \max\min H\left(x,u,d,\dot{V}(x,t)^T\right) \\ = 0,\ x_0 \in \mathbb{R}^n \end{aligned} \tag{10}$$

The first theorem provides sufficient conditions for guaranteeing finite-time stability along with the existence of the Nash equilibrium.

Theorem 1. *Considering the controlled nonlinear dynamics system (3) with performance index (4). Assume that there exist a radially unbounded continuously differentiable function $V:\mathbb{R}^n\to\mathbb{R}$, real numbers $\alpha>0$ and $\beta\in(0,1)$, and continuous feedback strategies such that*

$$\begin{aligned} & u^*(0)=0, \quad d^*(0)=0 \\ & V(0)=0, \quad V(x)>0 \\ & \dot{V}(x)\left[\mathcal{P}\left(x,u^*(x),d^*(x)\right)+\sigma(x,t)\right] \le -\alpha V^{\beta}(x) \\ & H\left(x,u^*,d^*,\dot{V}(x)^T\right)=0 \\ & H\left(x,u,d^*,\dot{V}(x)^T\right)\ge 0 \\ & H\left(x,u^*,d,\dot{V}(x)^T\right)\le 0 \end{aligned} \tag{11}$$

with the feedback strategies $u^*(\cdot)=u^*(x(\cdot))$ and $d^*(\cdot)=d^*(x(\cdot))$, the zero solution $x(t)=0$ to

$$\dot{\mathbf{x}} = \mathcal{P}\left(\mathbf{x},u^*,d^*\right)+\sigma(\mathbf{x},t),\ t\ge 0 \tag{12}$$

is globally finite-time stable. Moreover, there exists a probability transition function $T:\mathbb{R}^n\to[0,1]$ such that

$$T(x_0) \le \frac{1}{\alpha(1-\beta)}\left|V(x_0)\right|^{1-\beta},\ x_0\in\mathbb{R}^n \tag{13}$$

and if $x_0\in\mathbb{R}^n$, then the pair of feedback strategies $(u^*(\cdot),d^*(\cdot))$ is the Nash equilibrium of the defender-attacker non-cooperative games in the sense that

$$J\left(x_0,u^*(\cdot),d^*(\cdot)\right) = \min_{u(\cdot)\in\mathcal{U}_D(x_0)}\ \max_{d(\cdot)\in\mathcal{U}_A(x_0)} \mathbb{E}\left[J\left(x_0,u(\cdot),d(\cdot)\right)\right] \tag{14}$$

and the Nash value is

$$J\left(x_0,u^*(\cdot),d^*(\cdot)\right) = V(x_0) \tag{15}$$

Proof. Global finite-time stability along with the existence of a probability transition function $T:\mathbb{R}^n\to[0,1]$ satisfying Eq. (13) follow Eq. (11) accorcding to Definition 1. Next, let $(u(x(\cdot)), d(x(\cdot)))\in\mathcal{F}(x_0)$, and let $x(t), t\ge 0$, be the solution to Eq. (3). Then, using Eq. (7), it follows that

$$H\left(x,u,d,\dot{V}(x)^T\right) = L(x,u,d)+\dot{V}(x),\ t\ge 0 \tag{16}$$

and let us extend Eq. (16) from $0 \to t$ to $t \to \infty$ to get

$$\int_0^\infty H\left(x(t), u(t), d(t), \dot{V}(x(t))^T\right) dt = \int_0^\infty L(x(t), u(t), d(t))\, dt + \int_0^\infty \dot{V}(x(t))\, dt \tag{17}$$

However, since $(u(\cdot), d(\cdot)) \in \mathcal{F}(x_0)$, it follows that $\lim_{t\to T(x_0)} x(t) = \lim_{t\to\infty} x(t) = 0$, and since $V(\cdot)$ is continuous, $\lim_{t\to\infty} V(x(t)) = V(\lim_{t\to\infty} x(t)) = 0$, which in turn implies that $\int_0^\infty \dot{V}(x(t))\, dt = -V(x_0)$. Therefore, Eq. (17) reduces to

$$\int_0^\infty H\left(x(t), u(t), d(t), \dot{V}(x(t))^T\right) dt = \int_0^\infty L(x(t), u(t), d(t))\, dt - V(x_0) \tag{18}$$

and let $u(\cdot) = u^*(x(\cdot))$ and $d(\cdot) = d^*(x(\cdot))$, and since $(u^*(\cdot), d^*(\cdot)) \in \mathcal{F}(x_0)$, Eq. (15) originates from Eq. (18) using Eq. (11). Next, let $u(\cdot) \in \mathcal{U}_D(x_0)$ and $d(\cdot) \in \mathcal{U}_A(x_0)$, which implies that $(u(\cdot), d^*(\cdot)) \in \mathcal{F}(x_0)$. Using Eqs. (11) and (18) yields

$$\begin{aligned} J\left(x_0, u(\cdot), d^*(\cdot)\right) &= V(x_0) + \int_0^\infty H\left(x(t), u(t), d^*(t), \dot{V}(x(t))^T\right) dt \\ &\geq V(x_0) \end{aligned} \tag{19}$$

similarly,

$$\begin{aligned} J\left(x_0, u^*(\cdot), d(\cdot)\right) &= V(x_0) + \int_0^\infty H\left(x(t), u^*(t), d(t), \dot{V}(x(t))^T\right) dt \\ &\leq V(x_0) \end{aligned} \tag{20}$$

consequently, in view of Eqs. (15), (19) and (20), we can get

$$J\left(x_0, u^*(\cdot), d(\cdot)\right) \leq J\left(x_0, u^*(\cdot), d^*(\cdot)\right) \leq J\left(x_0, u(\cdot), d^*(\cdot)\right) \tag{21}$$

for all $x_0 \in \mathbb{R}^n$, $u(\cdot) \in \mathcal{U}_D(x_0)$ and $d(\cdot) \in \mathcal{U}_A(x_0)$, which proves Eq. (14). □

3.2 | Collision free with safety guaranteed

In this section, the concept of a finite-time stable zero-sum game is introduced for ensuring a collision-free non-cooperative game. Firstly, the agent is assumed to detect the obstacle i.

Proposition 1. *Hölder condition: The obstacle $O_i(\cdot)$ is continuous on $x \in \Omega$ and for any boundary set $\partial\mathcal{O}_i$ there exist Hölder condition $C = C(\partial\mathcal{O}_i) \in (0, \infty)$ such that*

$$\left|O_i(x) - O_i(y)\right| \leq C\left|x - y\right|^a$$

for any $x \in \partial\mathcal{O}_i$ and $y \in \partial\mathcal{O}_i$, where C and a are non-negative constants. Thus, in the set of obstacles $\mathcal{O}_i$ we suppose $O_i(x - q_i) \Rightarrow \left|O_i(x) - O_i(y)\right|$ and let $c_i \Rightarrow C\left|x - y\right|^a$ such that $O_i(x - q_i) \leq c_i$ holds. In turn, $\mathcal{O}_i^{free}$ is setted to represent the collision-free space.

Considering that the set $\mathcal{O}_i^{free}$ related to the obstacle i represents the collision-free space and is defined as the zero-super level set of a Hölder condition function $O_i(x - q_i) : \mathbb{R}^n \to \mathbb{R}$, i.e.,

$$\begin{aligned} \mathcal{O}_i^{free} &= \{x \in \mathbb{R} : O_i(x - q_i) \geq c_i\} \\ \partial\mathcal{O}_i^{free} &= \{x \in \mathbb{R} : O_i(x - q_i) = c_i\} \\ \text{int}\left(\mathcal{O}_i^{free}\right) &= \{x \in \mathbb{R} : O_i(x - q_i) > c_i\} \end{aligned} \tag{22}$$

where $\partial\mathcal{O}_i^{free}$ and $\text{int}\left(\mathcal{O}_i^{free}\right)$ denote the boundary set and the interior set of $\mathcal{O}_i^{free}$, respectively. For ease of notation, we let $\mathcal{H} = \text{int}\left(\mathcal{O}_i^{free}\right)$.

In view of an unknown environment, the agent should make decisions in finite time aiming at establishing optimality along with guaranteeing collision avoidance, which amounts to rendering the safe set $\mathcal{H}$ positively invariant[33] with respect to the system (3), that is, if $x_0 \in \mathcal{H} \to x(t) \in \mathcal{H}$, $t \geq 0$. In the sequel, for every initial condition $x_0 \in \mathcal{H}$, the set of pairs of feedback strategies associated with the safe set is defined as $\mathcal{F}(x_0, \mathcal{H}) = \{u(\cdot) \in \mathcal{F}_u(x_0), d(\cdot) \in \mathcal{F}_d(x_0)\}$: $x(\cdot)$ given by Eq. (3) satisfies $x(t) \to 0$ as $t \to T(x_0)$ and $x(t) \in \mathcal{H}$, $t \geq 0$. Specifically, the set $\mathcal{F}(x_0, \mathcal{H})$ comprises the pairs of finite-time stabilizing control strategies $(u(\cdot), d(\cdot))$ rendering both the arising closed-loop system (3) finite-time stable and the safe set $\mathcal{H}$ positively invariant with respect to closed-loop system (3). Moreover, let $\mathcal{U}_D(x_0, \mathcal{H}) \subseteq \mathcal{F}_u(x_0)$ and $\mathcal{U}_A(x_0, \mathcal{H}) \subseteq \mathcal{F}_d(x_0)$ be such that $\mathcal{U}_D(x_0, \mathcal{H}) \times \mathcal{U}_A(x_0, \mathcal{H}) = \mathcal{F}(x_0, \mathcal{H})$, $x_0 \in \mathcal{H}$.

Now, our problem is stated as a finite-time stable zero-sum game. Considering the system (3) end time evolution of non-cooperative games associated with the performance measure (4) wherein the defender seeks to minimize it, whereas the attacker seeks to maximize it. Suppose that the agent has detected the obstacle, and considers the associated safe set $\mathcal{H}$. Then, the objective amounts to finding a safe finite-time stabilizing saddle point solution $\left(u_s^*(\cdot), d_s^*(\cdot)\right) \in \mathcal{F}(x_0, \mathcal{H})$, rendering the equilibrium point of the closed-loop system (3) with $u(\cdot) = u_s^*(\cdot)$ and $d(\cdot) = d_s^*(\cdot)$ finite-time stable and the safe set $\mathcal{H}$ positively invariant with respect to the closed-loop system (3) while $\left(u_s^*(\cdot), d_s^*(\cdot)\right)$ is a saddle point of the performance index (4) for every $x_0 \in \mathcal{H}$.

In light of the above, a finite-time stable zero-sum game is mathematically described by an optimal control problem of the form

$$V(x_0, t) = \max_{d(\cdot)\in\mathcal{U}_A(x_0,\mathcal{H})} \min_{u(\cdot)\in\mathcal{U}_D(x_0,\mathcal{H})} \mathbb{E}\left[J\left(x_0, u(\cdot), d(\cdot)\right)\right], \; x_0, x(t) \in \mathcal{H}, \; t \geq 0 \tag{23}$$

However, finding a feedback solution to a state-constrained differential game is a computationally intractable task. Thus, to circumvent that, the latter can be equivalently converted into an unconstrained optimal control problem by integrating an obstacle function $B(\cdot)$ connected with the safe set $\mathcal{H}$ into the performance integrand (5), thereby yielding

$$L_B(x, u, d) = L_1(x) + \epsilon B(x) + u^T R_u u - d^T R_d d \tag{24}$$

where $\epsilon > 0$ is a factor to enable a trade-off between safety and optimality by determining the relative dominance of the obstacle function $B(\cdot)$ to the running cost $L_B(\cdot)$. As a consequence, the safe Hamilton function is now defined as

$$H_B\left(x, u, d, \dot{V}(x,t)^T\right) = L_B(x, u, d) + \dot{V}(x,t)\left[\mathcal{P}(x, u, d) + \sigma(x,t)\right] \tag{25}$$

Definition 3. Let $B : \mathcal{H} \to \mathbb{R}$ be a continuous function on $\mathcal{H}$ and consider that $B(\cdot)$ allows the following properties.
(1) It vanishes at the equilibrium point of the system (3), that is, $B(0) = 0$.
(2) It takes positive values everywhere on $\mathcal{H}$ except the origin, that is, $B(0) > 0$, $x \in \mathcal{H}\backslash\{0\}$.
(3) It approaches infinity as the state approaches the boundary of the safe set $\mathcal{H}$, that is, $\lim_{x\to\partial\mathcal{H}} B(x) = \infty$.

Considering the properties of obstacle function in Definition 3 and $B(0) = 0$ needs to be satisfied, and $B(x) > 0$, $x \in \mathcal{H}\backslash\{0\}$, and $\sup_{x\in\mathcal{H}} B(x)$ exists. Therefore, we can let $B(x) = \log\left(1 + 1/\left(O_i(x - q_i) - c_i\right)\right)$, $x \in \mathcal{H}$. Based on the above, the safety-aware performance measure associated with the safe set $\mathcal{H}$ is provided by

$$J_B(x_0, u, d) = \int_0^\infty L_B\left(x(t), u(t), d(t)\right)\, dt \tag{26}$$

thus, Eq. (23) can give rise to the following unconstrained optimal control problem

$$V_B(x_0, t) = \max_{d(\cdot)\in\mathcal{U}_A(x_0,\mathcal{H})} \min_{u(\cdot)\in\mathcal{U}_D(x_0,\mathcal{H})} \mathbb{E}\left[J_B\left(x_0, u(\cdot), d(\cdot)\right)\right], \; x_0, x(t) \in \mathcal{H}, \; t \geq 0 \tag{27}$$

Note that the saddle point solution $\left(u_s^*(\cdot), d_s^*(\cdot)\right)$ and the value of the game $J_B\left(x_0, u_s^*(\cdot), d_s^*(\cdot)\right)$ can be termed as the safe Nash equilibrium and the safe Nash value related to the safe set $\mathcal{H}$, respectively. The following theorem provides sufficient conditions for characterizing the safe Nash equilibrium by establishing finite-time stability along with positive invariance of the safe set $\mathcal{H}$. As illustrated in Fig. 3, the structure of the NCPG includes five steps: from the hybrid system to the verification, where the finite-time stability analysis and the Nash equilibrium solution process are shown in the control structure below.

Theorem 2. *Considering the controlled nonlinear dynamics system (3) with performance index (26). Let $x_0 \in \mathcal{H}$ and assume that there exist a continuously differentiable function $V_B : \mathcal{H} \to \mathbb{R}$, real numbers $\alpha > 0$ and $\beta \in (0, 1)$, and continuous feedback* *strategies such that*

$$\begin{aligned} & u_s^*(0) = 0, \quad d_s^*(0) = 0 \\ & V_B(0) = 0, \quad V_B(x) > 0, \quad x \in \mathcal{H}\backslash\{0\} \\ & \dot{V}_B(x)\left[\mathcal{P}\left(x, u_s^*(x), d_s^*(x)\right) + \sigma(x,t)\right] \leq -\alpha V^{\beta}(x) \\ & \sup_{x\in\mathcal{H}} \left|L\left(x, u_s^*(x), d_s^*(x)\right)\right| < \infty \\ & H_B\left(x, u_s^*, d_s^*, \dot{V}_B(x)^T\right) = 0 \\ & H_B\left(x, u_s, d_s^*, \dot{V}_B(x)^T\right) \geq 0 \\ & H_B\left(x, u_s^*, d_s, \dot{V}_B(x)^T\right) \leq 0 \end{aligned} \tag{28}$$

with the feedback strategies $u_s^*(\cdot) = u_s^*(x(\cdot))$ *and* $d_s^*(\cdot) = d_s^*(x(\cdot))$, *the zero solution* $x(t) = 0$ *to*

$$\dot{\mathbf{x}} = \mathcal{P}\left(\mathbf{x}, u_s^*, d_s^*\right) + \sigma(\mathbf{x}, t), \ t \geq 0 \tag{29}$$

is finite-time stable and $\mathcal{H}$ *is a positively invariant set with respect to Eq. (29). Moreover, there exist an open neighborhood* $\mathcal{H}_0 \subset \mathcal{H}$ *of the origin and a probability transition function* $T_B : \mathcal{H}_0 \to [0,1]$ *such that*

$$T_B(x_0) \leq \frac{1}{\alpha(1-\beta)} \left|V_B(x_0)\right|^{1-\beta}, \ x_0 \in \mathcal{H}_0 \tag{30}$$

and if $x_0 \in \mathcal{H}_0$, *then the pair of feedback strategies* $\left(u_s^*(\cdot), d_s^*(\cdot)\right)$ *is the Nash equilibrium of the safe defender-attacker non-cooperative games in the sense that*

$$J_B\left(x_0, u_s^*(\cdot), d_s^*(\cdot)\right) = \min_{u(\cdot)\in\mathcal{U}_D(x_0,\mathcal{H})} \max_{d(\cdot)\in\mathcal{U}_A(x_0,\mathcal{H})} \mathbb{E}\left[J_B\left(x_0, u(\cdot), d(\cdot)\right)\right] \tag{31}$$

and the Nash value is

$$J_B\left(x_0, u_s^*(\cdot), d_s^*(\cdot)\right) = V_B(x_0) \tag{32}$$

Proof. Finite-time stability along with the existence of a probability transition function $T_B(\cdot)$, as well as the existence of the safe Nash equilibrium using identical arguments as in the proof of Theorem 1.

However, to complete the proof, that the set $\mathcal{H}$ is positively invariant with respect to Eq. (29) needs to be demonstrated. To this end, let $x_0 \in \mathcal{H}$ and let $x(t)$, $t \geq 0$, be the solution to Eq. (29). Assume that there exists $t^* > 0$ such that $\lim_{t\to t^*} dist\,(x(t), \partial\mathcal{H}) = 0$.

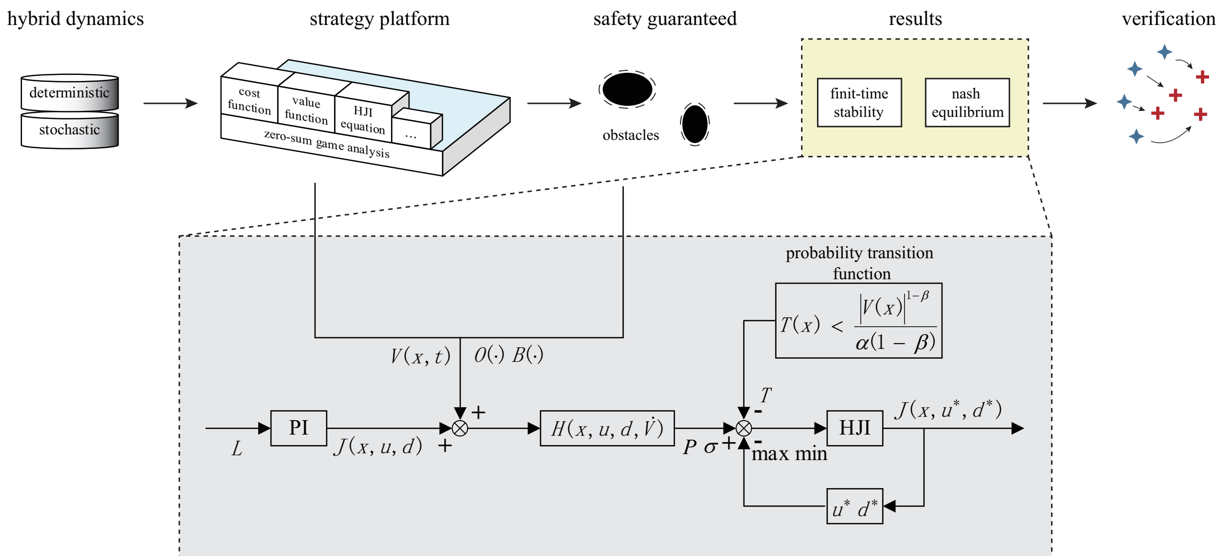


FIGURE 3 NCPG strategy control structure.

Now, integrating Eq. (28) over $[0, t^*]$ yields

$$
\begin{aligned}
H_B\left(x(t), u_s^*(x(t)), d_s^*(x(t)), V_B'^T(x)\right) &= L_B\left(x(t), u_s^*(x(t)), d_s^*(x(t))\right) + \dot{V}(x(t)) \\
\int_0^{t^*} H_B\left(x(t), u_s^*(x(t)), d_s^*(x(t)), V_B'^T(x)\right)\, dt &= \int_0^{t^*} L_B\left(x(t), u_s^*(x(t)), d_s^*(x(t))\right)\, dt + \int_0^{t^*} \dot{V}(x(t))\, dt \\
0 &= \int_0^{t^*} L\left(x(t), u_s^*(x(t)), d_s^*(x(t))\right)\, dt + \int_0^{t*} \epsilon B(x(t))\, dt + V_B\left(x(t^*)\right) - V_B(x_0) \\
V_B\left(x(t^*)\right) &= \left[V_B(x_0) - \int_0^{t^*} L\left(x(t), u_s^*(x(t)), d_s^*(x(t))\right)\, dt\right] - \epsilon \int_0^{t*} B(x(t))\, dt
\end{aligned}
\tag{33}
$$

where the term $V_B(x_0) - \int_0^{t^*} L\left(x(t), u_s^*(x(t)), d_s^*(x(t))\right)\, dt$ is finite by Eq. (28). On the contrary, since $\lim_{t\to t^*} dist\,(x(t), \partial\mathcal{H}) = 0$, it follows that $\lim_{t\to t^*} B(x(t)) = \infty$, which in turn implies that $\epsilon \int_0^{t*} B(x(t))\, dt = \infty$. Therefore, $V_B(x(t^*)) = -\infty$, leading to a contradiction. Consequently, it turns out that if $x_0 \in \mathcal{H}$, then the solution $x(t)$ to Eq. (29) cannot escape $\mathcal{H}$, and thus the safe set $\mathcal{H}$ is positively invariant with respect to Eq. (29). □

4 | COOPERATIVE PROBABILISTIC GAMES WITH UNSTATIONED ATTACKER

4.1 | Probabilistic reachability analysis

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)[34]. Combining the optimal control framework with the time-evolution of the PDF will give rise to the field of PRA. The PRA 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 $\mathcal{R}$ at time $t = 0$ and in a final state $\mathbf{x} = \left(x_p, x_t\right)$ at time $t = T$.

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

$$\dot{\mathbf{x}}(t) = \mathcal{P}(\mathbf{x}, u, d) + \sigma(\mathbf{x}, t) \tag{34}$$

Because the dynamics are stochastic, it is no longer possible to minimize the value function. Instead, the expected payoff is minimized over all possible further realizations of the Wiener process

$$V(t, x) = \min_{u\in\mathcal{U}_D} \max_{d\in\mathcal{U}_A} \mathbb{E}\left[\phi\left(\mathbf{x}_{t,x}(T)\right)\right] \tag{35}$$

with $\phi(x) := \mathbf{1}_{\mathcal{K}}(x)$, where

$$\mathbf{1}_{\mathcal{K}}(x) = \begin{cases} 1 & \text{if } x \in \mathcal{K} \\ 0 & \text{if } x \notin \mathcal{K} \end{cases} \tag{36}$$

where $\mathcal{K}$ is the non-empty target set in $\mathbb{R}^n$. Consider the reachable set $\mathcal{R}$ under probability of success ρ, or the set of initial conditions x for which the probability that there exists a trajectory $\mathbf{x}_{t,x}$ that reaches target set $\mathcal{K}$ at time T, associated with the control $u \in \mathcal{U}_D$ and $d \in \mathcal{U}_A$ is at least ρ[35]

$$\mathcal{RA}_t^\rho = \left\{x \in \mathbb{R}^n \,\middle|\, \exists u \in \mathcal{U}_D, d \in \mathcal{U}_A, \mathbf{P}\left[\mathbf{x}_{t,x}(T) \in \mathcal{K} \text{ and } \mathbf{x}_{t,x}(T) \notin \mathcal{A}\right] > \rho\right\} \tag{37}$$

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

$$\mathcal{RA}_t^\rho = \left\{x \in \mathbb{R}^n \,\middle|\, V(x, t) > \rho\right\} \tag{38}$$

An intuitive derivation can be made if it is assumed that the value function is differentiable

$$V(x, t) = \min_{u\in\mathcal{U}_D} \max_{d\in\mathcal{U}_A} \left\{\mathbb{E}\left[V\left(x + \Delta\mathcal{P}(x, u, d) + \xi, t + \Delta t\right)\right]\right\} \tag{39}$$

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.

$$V(x+\Delta x, t+\Delta t) \approx V(x,t) + V_t(x,t)\Delta t + V_x(x,t)\Delta x + \frac{1}{2}\left(V_{xx}\Delta x^2 + 2V_{xt}\Delta x\Delta t + V_{tt}\Delta t^2\right) + \mathcal{Q}(\delta^3) \tag{40}$$

where $\Delta x = \Delta\mathcal{P}(x,u,d) + \xi$. Keeping all terms of order $\mathcal{Q}(\Delta t)$, we have the Dynkin Operator[8]

$$\mathbb{E}[V] = V(x,t) + \Delta\mathcal{P}(x,u,d)V_x(x,t) + \frac{1}{2}Tr\left(\sigma^2(x,t)V_{xx}(x,t)\right) \tag{41}$$

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

$$\frac{V(x,t) - V(x,t+\Delta t)}{\Delta t} = \min_{u\in\mathcal{U}_D}\max_{d\in\mathcal{U}_A}\left\{V_x\mathcal{P}(x,t)^T + \frac{1}{2}Tr\left(\sigma(x,t), \sigma(x,t)^T V_{xx}\right)\right\} \tag{42}$$

For the limit of $\Delta t \to 0$ the value function is characterized by the viscosity solution to the stochastic HJB equation[8]

$$\frac{\partial V}{\partial t}(x,t) + \sup\left\{\frac{\partial V}{\partial x}(x,t)\mathcal{P}(x,u,d)\right\} + \frac{1}{2}Tr\left\{\sigma(x,t), \sigma(x,t)^T\frac{\partial^2 V}{\partial x^2}(x,t)\right\} = 0 \tag{43}$$

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 method. In contrary to the deterministic reachable set, the value function in the PRA is not represented by a signed distance function. Instead, an indicator function (36) is used. The regularization of the indicator function is done by

$$V^{\varepsilon}(x) = 1 - \min\left(1, \max\left(0, -\frac{1}{\varepsilon}dist(x,\mathcal{K})\right)\right) \tag{44}$$

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

In a reach-avoid game, the goal of the attacker is to avoid being intercepted by the defender while reaching the target set. This reach set $\mathcal{R}$ is represented by the attacker being inside $\mathcal{K}$. On the way to target, the attacker must avoid being intercepted by the defender. All 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, the reach set and the avoid set are defined as

$$\mathcal{R} = \left\{\mathbf{x}\in\Omega^2 \,\middle|\, x_A \in \mathcal{K} \wedge \|x_A - x_D\|^2 > R_C\right\} \cup \left\{\mathbf{x}\in\Omega^2 \,\middle|\, x_D \in \Omega_{obs}\right\} \tag{45}$$

$$\mathcal{A} = \left\{\mathbf{x}\in\Omega^2 \,\middle|\, \|x_A - x_D\|^2 \leq R_C\right\} \cup \left\{\mathbf{x}\in\Omega^2 \,\middle|\, x_A \in \Omega_{obs}\right\} \tag{46}$$

Given these sets, the corresponding level set representations Eqs. (37) and (43) can be solved. If $\mathbf{x^0} \in \mathcal{RA}_t^{\rho}$, then the attacker is guaranteed to win the game. The explicit winning strategy can be obtained from Ref.[36]

$$A^*(x_A, x_D, t) = -v_A\frac{s_d(x_A, x_D, t)}{\|s_d(x_A, x_D, t)\|^2} \tag{47}$$

where $s = (s_d, s_u) = \frac{\partial V}{\partial(x_A, x_D)}$. Similarly, if $\mathbf{x^0} \notin \mathcal{RA}_t^{\rho}$, then the defender is guaranteed to win the game. The explicit winning strategy can be obtained

$$D^*(x_A, x_D, t) = v_D\frac{s_u(x_A, x_D, t)}{\|s_u(x_A, x_D, t)\|^2} \tag{48}$$

4.2 | Stochastic dynamics flowfield

In addition, the dynamics are partly deterministic and partly stochastic as the flowfield disturbances. Before the defender reaches the attacker, it not only needs to avoid obstacles, but also needs to try to overcome the effects caused by the disturbance. The disturbance of the flowfields are nonlinear. $g(x, y)$ denotes the flowfields current at position (x, y)[37]. The current is supposed to flow with constant direction, with the magnitude of g increasing in distance from the middle of the flows:

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

similarly, to describe the uncertainty influenced by the Brownian motion W, the diffusion term is considered

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

It is assumed that defenders can change their directions α instantaneously. The complete dynamics of the defenders are given by

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

where $\alpha \in [-\pi, \pi]$ is the direction of the defenders with respect to the x axis, where α depends on the term $\frac{s_u(x_A, x_D, t)}{\|s_u(x_A, x_D, t)\|_2}$ from Eq. (48).

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. 4.1, this probability can be characterized as the level set of a function. Stochastic HJB equation can be updated from Eq. (43) as follows

$$\begin{aligned} &\frac{\partial V}{\partial t}(x, y, t) + \sup_{\alpha \in [-\pi, \pi]} \left\{ \frac{\partial V}{\partial x}(x, y, t) \left(1 + ay^2 + v_D \cos(\alpha)\right) + \frac{\partial V}{\partial y}(x, y, t) v_D \sin(\alpha) \right\} + \frac{1}{2} Tr \left\{ \sigma_x^2 \frac{\partial^2 V}{\partial x^2}(x, y, t) \right\} \\ &+ \frac{1}{2} Tr \left\{ \sigma_y^2 \frac{\partial^2 V}{\partial y^2}(x, y, t) \right\} = 0 \end{aligned} \tag{49}$$

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

$$\begin{aligned} \alpha^*(x, y, t) &= \arg\max_{\alpha \in [-\pi, \pi]} \left(\frac{\partial V}{\partial x}(x, y, t) \cos(\alpha) + \frac{\partial V}{\partial y}(x, y, t) \sin(\alpha) \right) \\ &= \arctan_\alpha \left(\frac{\partial_y V}{\partial_x V} \right) (x, y, t) \end{aligned} \tag{50}$$

Therefore, the stochastic HJB equation can be simplified to

$$\begin{aligned} &\frac{\partial V}{\partial t}(x, y, t) + \frac{\partial V}{\partial x}(x, y, t) \left(1 + ay^2\right) + \frac{1}{2} Tr \left\{ \sigma_x^2 \frac{\partial^2 V}{\partial x^2}(x, y, t) \right\} + \frac{1}{2} Tr \left\{ \sigma_y^2 \frac{\partial^2 V}{\partial y^2}(x, y, t) \right\} \\ &+ v_D \|\nabla V(x, y, t)\| = 0 \end{aligned} \tag{51}$$

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

4.3 | Competitiveness analysis with predictive planning

In this section, more agents with distributed control are added and the agents not communicating with each other are assumed. In order to compute the optimal strategy for the defender team, all defenders need to predict the strategies and actions of the other defenders and make plans based on these predictions. It is called a Generalized Nash equilibrium. However, the allocation of defenders to attackers is based solely on the cost metric associated with attacker interception. In this paper, a Defender Collision-Free Cross-Interception (DCFCI) strategy is proposed for defenders to intercept attackers in order to prevent the attackers from reaching the target area, such that the defender-to-attacker assignment considers not only the movement trajectory cost but also possible collisions among defenders on the optimal trajectory. In particular, this strategy are decoupled into two stages. Stage I will find a feasible solution set with predictive planning, and stage II will further calculate an improved solution set with cross-interception. Firstly, for each agent in the distributed multi-agent system, let $(x_i(t), u_i(t))$ denotes the relative coordinates and control input of the defender $i \in \{1, \ldots, N\}$, and let $\breve{u}_i$ denotes the strategies of all other defenders except the defender i, i.e., $\breve{u}_i = (u_1, \ldots, u_{i-1}, u_{i+1}, \ldots, u_n)$. If the strategy set of i-th defender $\Theta_i(\breve{u}_i) \subseteq \mathbb{R}^{n_i}$ depends on other defenders strategies, then the strategy set of DCFCI can be denoted as $\Theta_i(\breve{u}_i) = \{(x_i, u_i) \in \mathbb{R} | (x_i, u_i, \breve{u}_i) \in \Theta\}$, where $\Theta \subseteq \mathbb{R}^n$ is a nonempty, closed set representing the joint constraints of all defenders $i \in \{1, \ldots, N\}$. Θ is allowed to be nonconvex[38]. The set $\Theta_i(\breve{u}_i)$ is closed because of the property of Θ. In light of the above, DCFCI strategy is mathematically described by an optimal control problem of the form

$$\text{s.t. } (x_i, u_i) \in \Theta_i(\breve{u}_i) \quad V(x) = \min_{u_i} \mathbb{E}\left[J_i\left(x_i, u_i, \breve{u}_i\right)\right] \tag{52}$$

Proposition 2. *A strategy profile $(x_i^*, u_i^*) \in \Theta(x_i^*, u_i^*)$ is called as a generalized Nash equilibrium solution of DCFCI strategy, if and only if (x_i^*, u_i^*) satisfying the following inequality*

$$J_i\left(x_i^*, u_i^*, \breve{u}_i^*\right) \leq J_i\left(x_i, u_i, \breve{u}_i^*\right), \ \forall i \in \{1, \ldots, N\} \tag{53}$$

where (x_i^*, u_i^*) and (x_i, u_i) are the optimal strategy and strategy of the defender i, respectively. $\breve{u}_i^*$ summarizes the optimal strategies of all other defenders except the defender i.

The general form of constraints of DCFCI strategy can be define as

$$\Theta_i(\breve{u}_i) := \{(x_i, u_i) \in \mathbb{R}^{n_i} | h_i(x_i, u_i, \breve{u}_i) \geq 0\} \tag{54}$$

where the h_i depending on the decision of other defenders is called a joint constraint of DCFCI strategy. Since our target is to minimize the objective function J_i, this optimization problem is solved by Nikaido-Isoda function[39], which can be defined as

$$\Psi\left(\Theta^k, \Theta^{k+1}\right) := \sum_{i=1}^{N} \left[J_i\left(x_i^{k+1}, u_i^{k+1}, \breve{u}_i^k\right) - J_i\left(x_i^k, u_i^k, \breve{u}_i^k\right)\right] \tag{55}$$

where Θ^k and Θ^{k+1} are two feasible solution sets (as two stages: stage I and stage II) for the DCFCI strategy, and Θ^{k+1} is the improvement of Θ^k in the objective function of the i-th defender while all the other defenders keep their strategy unchanged. Using Nikaido-Isoda function, we define

$$\aleph(x_i, u_i) := \sup_{\Theta^{k+1}} \Psi\left(\Theta^k, \Theta^{k+1}\right), \ (x_i, u_i) \in \Theta \tag{56}$$

then, it is not difficult to see that $\aleph(x_i, u_i)$ is nonnegative for Θ^k, and that (x_i^*, u_i^*) is a solution of DCFCI strategy if and only if $(x_i^*, u_i^*) \in \Theta(x_i^*, u_i^*)$ and $\aleph(x_i^*, u_i^*)$. Therefore, finding a solution of DCFCI strategy is equivalent to computing a global minimum of the optimization problem

$$\min \aleph(x_i, u_i) \quad \text{s.t. } (x_i, u_i) \in \Theta \tag{57}$$

However, some disadvantages[40] can be found from a mathematical point of view. On the one hand, given a vector (x_i, u_i), the supremum in Eq. (56) may not exist unless additional assumptions hold, and on the other hand, this supremum, if it exists, is usually not attained at a single point which, in turn, implies that the mapping $\aleph$ and, therefore, also the corresponding optimization Eq. (57) is non-differentiable in general. In other words, the feasible solution set in stage I may lead to non-differentiability in the optimization in order to meet the collision avoidance. Therefore, the cross-interception strategy in stage II overcomes these shortcomings. Specifically, a simple regularization of Eq. (55) is exploited to allows the transition from stage I to stage II. To this end, let $\gamma > 0$ be a fixed parameter and define

$$\Psi\left(\Theta^k, \Theta^{k+1}\right) := \sum_{i=1}^{N} \left[J_i\left(x_i^{k+1}, u_i^{k+1}, \breve{u}_i^k\right) - J_i\left(x_i^k, u_i^k, \breve{u}_i^k\right) - \gamma \| \left(x_i^{k+1}, u_i^{k+1}, \breve{u}_i^k\right) - \left(x_i^k, u_i^k, \breve{u}_i^k\right) \|^2\right] \tag{58}$$

Moreover, for $(x_i, u_i) \in \Theta$ we have

$$\begin{aligned}
\aleph_\gamma(x_i, u_i) : &= \max_{\Theta^{k+1}} \Psi_\gamma\left(\Theta^k, \Theta^{k+1}\right) \\
&= \max_{\Theta^{k+1}} \sum_{i=1}^{N} \left[J_i\left(x_i^{k+1}, u_i^{k+1}, \breve{u}_i^k\right) - J_i\left(x_i^k, u_i^k, \breve{u}_i^k\right) - \gamma \| \left(x_i^{k+1}, u_i^{k+1}, \breve{u}_i^k\right) - \left(x_i^k, u_i^k, \breve{u}_i^k\right) \|^2\right] \\
&= \sum_{i=1}^{N} \left\{J_i\left(x_i^{k+1}, u_i^{k+1}, \breve{u}_i^k\right) - \min_{\Theta^{k+1}} \left[J_i\left(x_i^k, u_i^k, \breve{u}_i^k\right) + \gamma \| \left(x_i^{k+1}, u_i^{k+1}, \breve{u}_i^k\right) + \left(x_i^k, u_i^k, \breve{u}_i^k\right) \|^2\right]\right\}
\end{aligned} \tag{59}$$

be the value function and a number of properties of the mapping $\aleph_\gamma$ are demonstrated in the following

Theorem 3. *The function $\aleph_\gamma$ satisfies:*
(i) $\aleph_\gamma(x_i, u_i) \geq 0$ for all $(x_i, u_i) \in \Theta$;
(ii) (x_i^*, u_i^*) is a generalized Nash equilibrium solution of DCFCI strategy if and only if $(x_i^*, u_i^*) \in \Theta(x_i^*, u_i^*)$ and $\aleph_\gamma(x_i^*, u_i^*) = 0$.

Proof. (i) For all $(x_i, u_i) \in \Theta$, we have $\aleph_\gamma(x_i, u_i) = \max_{\Theta^{k+1}} \Psi_\gamma\left(\Theta^k, \Theta^{k+1}\right) \geq \Psi_\gamma\left(\Theta^k, \Theta^k\right) = 0$.

(ii) Assume that (x_i^*, u_i^*) is a generalized Nash equilibrium solution of DCFCI strategy. Then $(x_i^*, u_i^*) \in \Theta(x_i^*, u_i^*)$ and $J_i\left(x_i^*, u_i^*, \breve{u}_i^*\right) \leq J_i\left(x_i, u_i, \breve{u}_i^*\right)$ for all $i \in \{1, \ldots, N\}$ from Eq. (53). Thus,

$$\begin{aligned}\Psi_\gamma\left(\Theta^{k+1}, \Theta^*\right) &= \sum_{i=1}^{N}\left[J_i\left(x_i^*, u_i^*, \breve{u}_i^k\right) - J_i\left(x_i^{k+1}, u_i^{k+1}, \breve{u}_i^k\right) - \gamma\left\|\left(x_i^*, u_i^*, \breve{u}_i^k\right) - \left(x_i^{k+1}, u_i^{k+1}, \breve{u}_i^k\right)\right\|^2\right] \\ &\leq 0\end{aligned} \tag{60}$$

where $J_i\left(x_i^*, u_i^*, \breve{u}_i^k\right) - J_i\left(x_i^{k+1}, u_i^{k+1}, \breve{u}_i^k\right) \leq 0$ and this implies

$$\aleph_\gamma(x_i^*, u_i^*) = \max_{\Theta^{k+1}} \Psi_\gamma\left(\Theta^{k+1}, \Theta^*\right) \leq 0 \tag{61}$$

with part (i), we have $\aleph_\gamma(x_i^*, u_i^*) = 0$.

In turn, assume that $(x_i^*, u_i^*) \in \Theta(x_i^*, u_i^*)$ and $\aleph(x_i^*, u_i^*)$. Then, $\Psi_\gamma\left(\Theta^{k+1}, \Theta^*\right) \leq 0$ holds. We let $\lambda \in (0, 1)$ be random and define

$$(x_j^{k+1}, u_j^{k+1}) = \begin{cases}(x_j^*, u_j^*), & \text{if } j \neq i \\ \lambda(x_i^*, u_i^*) + (1-\lambda)(x_i^k, u_i^k), & \text{if } j = i\end{cases} \tag{62}$$

where the convexity of the set Θ_i implies that $(x_j^{k+1}, u_j^{k+1}) \in \Theta_j$ for all $j \in \{1, \ldots, N\}$. For this particular (x_i^{k+1}, u_i^{k+1}), it therefore obtains

$$\begin{aligned}0 &\geq \Psi_\gamma\left(\Theta^{k+1}, \Theta^*\right) \\ &= J_i\left(x_i^*, u_i^*, \breve{u}_i^*\right) - J_i\left(\lambda(x_i^*, u_i^*) + (1-\lambda)(x_i^k, u_i^k), \breve{u}_i^k\right) - \gamma(1-\lambda)^2\left\|\left(x_i^*, u_i^*, \breve{u}_i^*\right) - \left(x_i^{k+1}, u_i^{k+1}, \breve{u}_i^k\right)\right\|^2 \\ &\geq (1-\lambda)J_i\left(x_i^*, u_i^*, \breve{u}_i^*\right) - (1-\lambda)J_i\left(x_i^k, u_i^k, \breve{u}_i^k\right) - \gamma(1-\lambda)^2\left\|\left(x_i^*, u_i^*, \breve{u}_i^*\right) - \left(x_i^{k+1}, u_i^{k+1}, \breve{u}_i^k\right)\right\|^2\end{aligned} \tag{63}$$

from the convexity of J_i with respect to (x_i^k, u_i^k). Dividing both sides by $(1-\lambda)$ and then letting $\lambda \to 1^-$ shows that

$$\begin{aligned}J_i\left(x_i^*, u_i^*, \breve{u}_i^*\right) - J_i\left(x_i^k, u_i^k, \breve{u}_i^k\right) &\leq \gamma(1-\lambda)\left\|\left(x_i^*, u_i^*, \breve{u}_i^*\right) - \left(x_i^{k+1}, u_i^{k+1}, \breve{u}_i^k\right)\right\|^2 \\ J_i\left(x_i^*, u_i^*, \breve{u}_i^*\right) &\leq J_i\left(x_i^k, u_i^k, \breve{u}_i^k\right)\end{aligned} \tag{64}$$

since this holds for all $(x_i^k, u_i^k) \in \Theta_i$ for all $i \in \{1, \ldots, N\}$, it follows that (x_i^*, u_i^*) is a generalized Nash equilibrium solution of DCFCI strategy. □

5 | SIMULATION

Following the works[15,41], the dynamics of the defender and the attacker are described by $\dot{x}_D(t) = u_D(t)$ and $\dot{x}_A(t) = d_A(t)$. The tracking error whose dynamics are given by $\dot{x}(t) = f(x(t)) + u_D(t) - d_A(t)$, where $f(x) := 1 - 0.01x^2$ is a wind disturbance with nonlinear dynamics flowfield. The terms composing the running cost (24) are given by $L_1(x) = (1/2)\|x\|$, $R_u = (1/2)I$, and $R_d = I$.

The NCPG is validated when the multi-agent system contains obstacles and stationary attackers in the following simulations. In our experiments, in the region $[0, 120] \times [0, 120]$ multiple defenders are setted, their attackers to be reached and some obstacles. The relevant information about the defenders, attackers and obstacles is shown in Table 1 and 2. For the purpose of comparison, the non-cooperative game method using non-probability[41] is chosen as a baseline. In this baseline method, an approximate solution to the HJI equation is obtained using a learning-based algorithm.

5.1 | Stationed attacker

In order to demonstrate the advantages of our method, the initial position of the defender, the attacker task and the obstacles is stationed in the next simulations. Figure 4 shows the game of six defenders with their attackers in an unknown complex

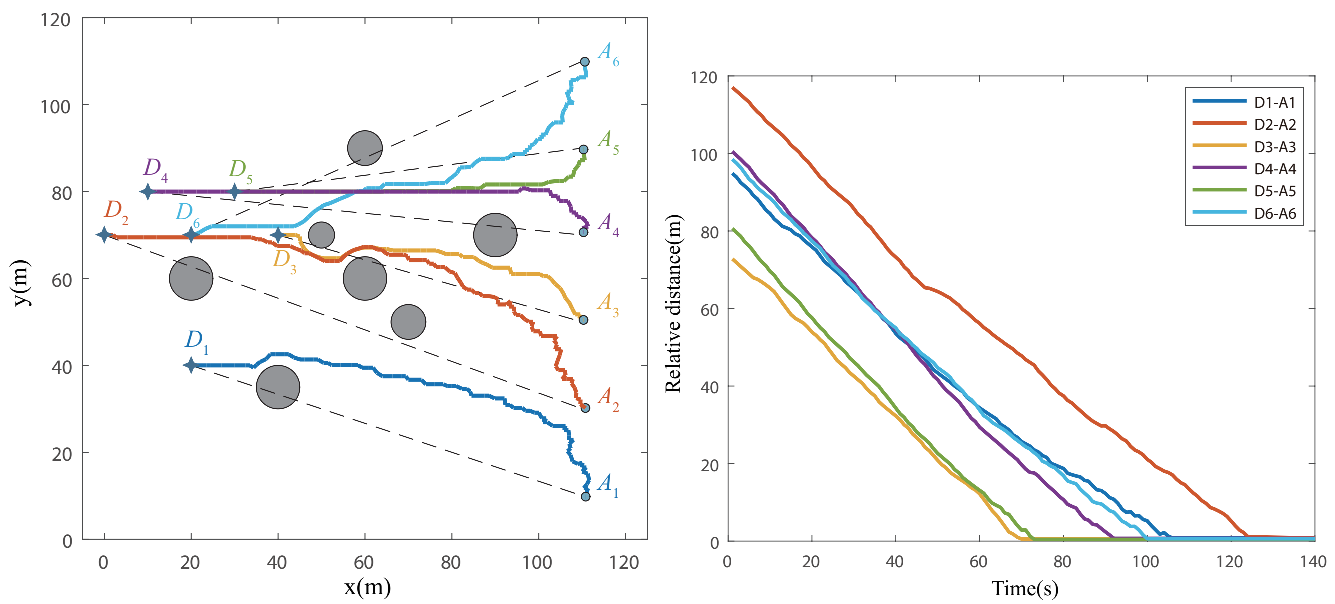
TABLE 1 Information of multiple defenders and attackers.

Object No.	Start position	Attacker position
Defender 1	[20,40]	[110,10]
Defender 2	[0,70]	[110,30]
Defender 3	[40,70]	[110,50]
Defender 4	[10,80]	[110,70]
Defender 5	[30,80]	[110,90]
Defender 6	[20,70]	[110,110]

TABLE 2 Information of obstacles.

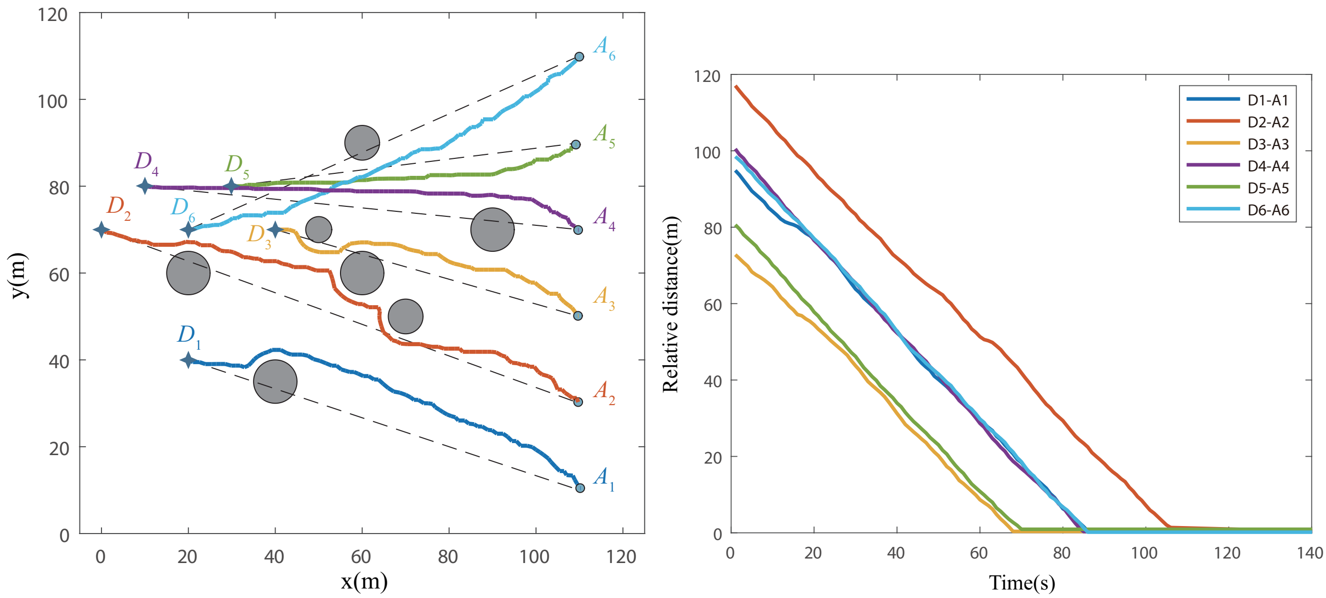
Obstacle No.	Position	Influence radius
Obstacle 1	[20,60]	5
Obstacle 2	[40,35]	5
Obstacle 3	[50,70]	3
Obstacle 4	[60,60]	5
Obstacle 5	[60,90]	4
Obstacle 6	[70,50]	4
Obstacle 7	[90,70]	5

environment using baseline method. Figure 5 shows the game situation of six defenders with their attackers in an unknown complex environment using our method. The grey dashed line represents the predefined reference trajectory r^d (without considering obstacles). Compared to the baseline method, our method is able to reach the Nash equilibrium earlier and smoothly. Figure 6 shows the real trajectory deviation from the predefined reference trajectory due to the disturbance of the flowfield. The pink region is the time period when the dynamics flowfield disturbance is applied. Obviously, our method has stronger stability compared to the baseline method.



(a) The trajectory diagram. (b) Relative distance of the defenders to their attackers.

FIGURE 4 Multiple defenders in an unknown complexed environment game with the attacker using baseline method[41]. The position of the attackers is stationed, so that the attackers are affected by the defenders counteracting on the defenders.



(a) The trajectory diagram.

(b) Relative distance of the defenders to their attackers.

FIGURE 5 Multiple defenders in an unknown complexed environment game with the attacker using our method. The position of the attackers is stationed, so that the attackers are affected by the defenders counteracting on the defenders.

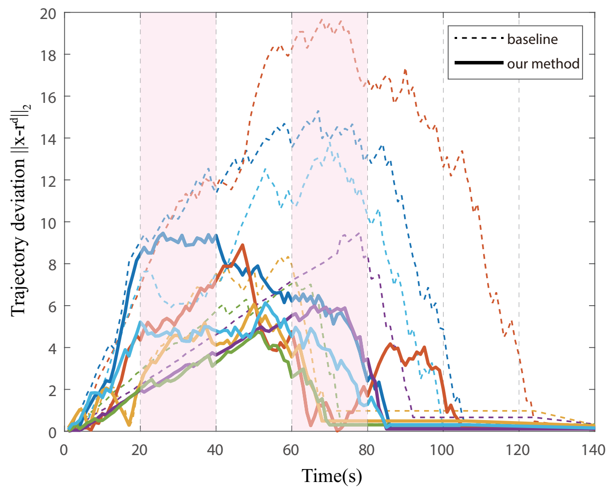


FIGURE 6 The real trajectory deviation from the predefined reference trajectory due to the disturbance of the flowfield.

5.2 | Unstationed Attacker

In this section the stationed attacker is setted to the unstationed attacker that the defender needs to capture before the attacker moves to the target. In addition, obstacles in the environment are retained to make it more difficult for the defender to reach the unstationed attacker. To present our results more clearly, the number of defenders is reduced to three. The initial positions of the defenders are set to $x_D = [40, 20, 40]$, $y_D = [100, 120, 0]$. The positions of the obstacles are set to $x_o = [30, 60, 70]$, $y_o = [110, 90, 30]$ and their radius of influence are $[5, 10, 10]$. The target is set in a circular area centred at $(80, 60)$ with a radius of 10. The initial positions of the unstationed attackers are set as $x_A = [125, 0, 125]$, $y_A = [100, 20, 20]$. The unstationed attackers are all moved in the direction closest to the target. Firstly, the baseline method is shown in Fig. 7. Obviously, although the defenders succeed in reaching the attackers, the targets have already entered the target.

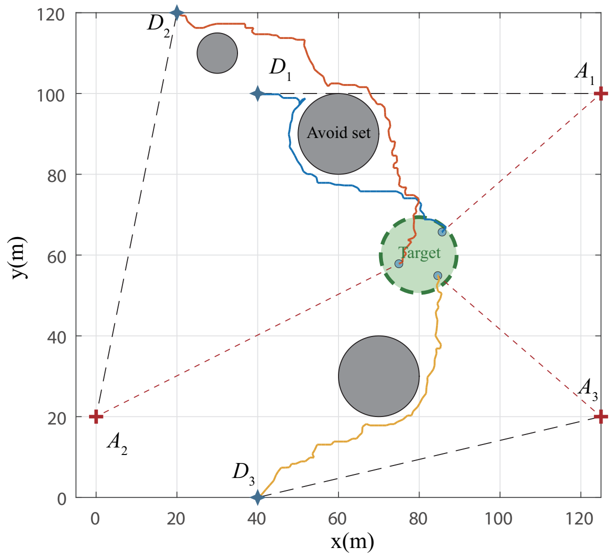
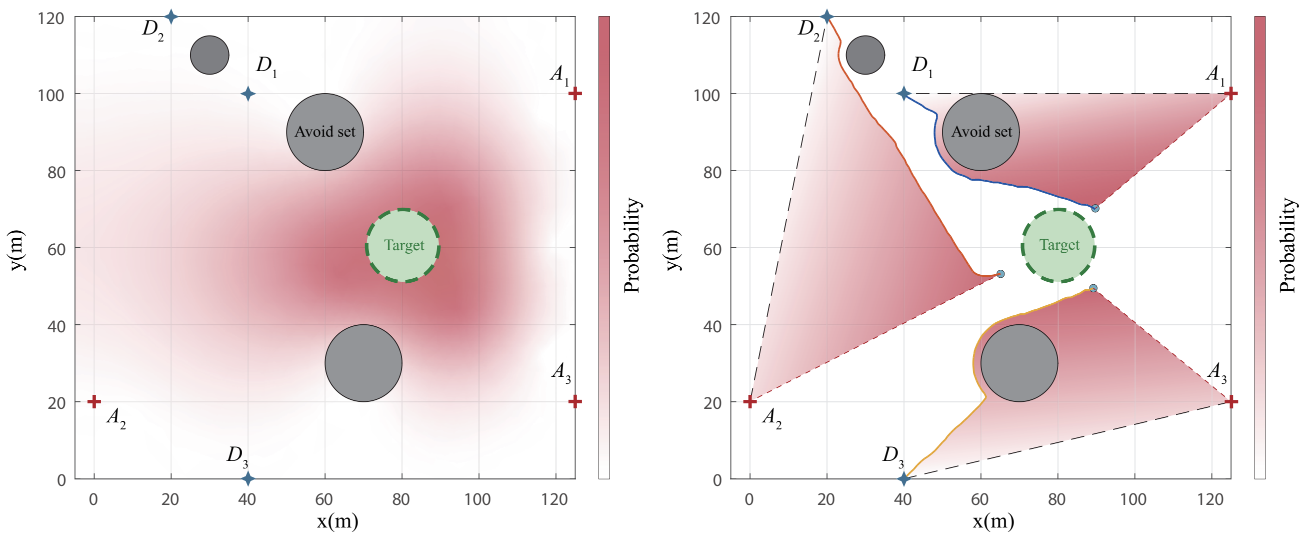


FIGURE 7 The defenders arrive at the unstationed attackers using baseline method. The green region is the target. The grey area is the avoid set (obstacles). The red dashed line is the trajectory of the unstationed attacker. The black dashed line is the initial reference trajectory.



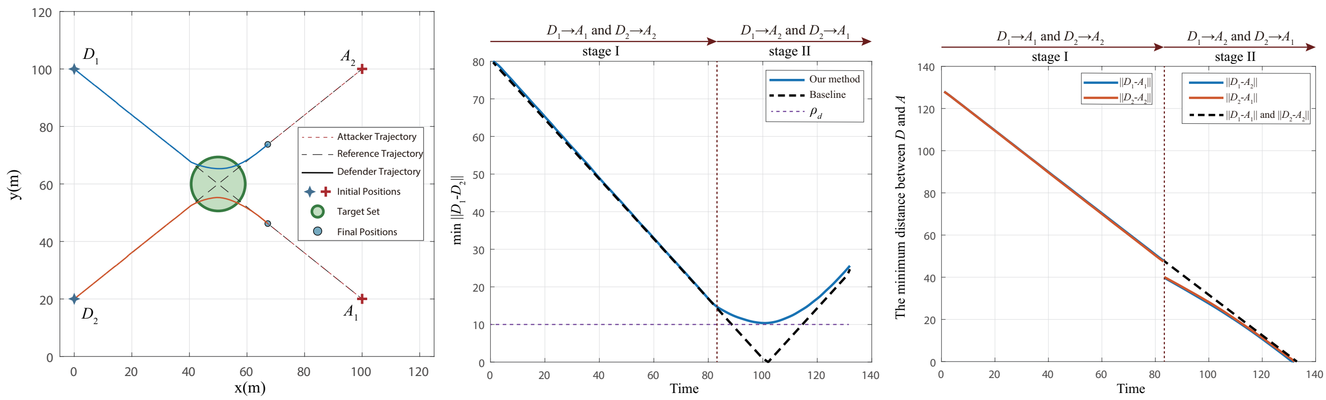
(a) High probability of reaching region.

(b) The defenders arrive at the unstationed attackers using our method.

FIGURE 8 The computation of the high probability region optimizes the game strategy for the defenders to reach the unstationed attackers. The dark pink region is the high probability of the defenders reaching the unstationed attackers region. As the distance among the defenders and the unstationed attackers decrease, the probability of the defenders reaching the attackers increases.

Figure 8 shows that the task of the defenders reaching the unstationed attackers using PRA is completed before the unstationed attackers reach the target. Our previously proposed method in probabilistic reachability in stochastic multiplayer reach-avoid games is used to compute the region of high probability of the defenders reaching the unstationed attackers as shown in Fig. 8(a). As the colour deepens, the probability of the defenders reaching the attackers increases. The defenders recalculate the game trajectory based on the calculated high probability region using our method as shown in Fig. 8(b). As the distances among the defenders and the unstationed attackers decrease, the probability of the defenders reaching the attackers increases.

In addition, we have distributed control over the defenders and assume that these defenders do not communicate with each other. Similarly, the unstationed attackers are controlled using time-optimal control. Consider a scenario with two defenders (black) and two unstationed attackers (red), as shown in Fig. 9(a), where the defenders first use the game strategy to method the initially matched game attackers, and there is no active collision-avoidance control among the defenders. When the DCFCI strategy is used, the defenders' closest unstationed attacker at this point changes from stage I to stage II according to Eqs. (58) and (59), and the closest attacker is reassigned for the DCFCI strategy. The optimal trajectory of the defenders is solid, while the dashed trajectory is the suboptimal trajectory corresponding to the allocation of the unstationed attacker by the defenders who do not know about the collision under the same initial conditions. The original method results in $D_1 \to A_1$ and $D_2 \to A_2$, while the results change to $D_1 \to A_2$ and $D_2 \to A_1$ when using the DCFCI strategy with predictive planning and cross-interception.



(a) The trajectory diagram.

(b) Minimum distance between defender 1 and defender 2.

(c) Minimum distance among the defenders and unstationed attackers.

FIGURE 9 Defenders reach the unstationed attackers before the unstationed attackers reach the target ensuring collision avoidance and time optimality between defenders. (defender: solid line, unstationed attacker: red dashed line)

Figure 9(b) shows the distance between defender 1 and defender 2 including suboptimal trajectory (baseline) and optimal trajectory with cross-interception (DCFCI strategy, our method) from stage I to stage II. The purple dashed line in Fig. 9(b) indicates the minimum safe distance ρ_d between defenders. Note that the defenders collide with each other at the baseline, which does not account for possible future collisions between the defenders. On the other hand, our method helps avoid imminent collisions between defenders and guarantees optimality in time. The whole process is divided into two stages. Compared to stage I, the defenders in stage II are reassigned game objectives. In Fig. 9(c), the minimum distance between the defenders and their game objectives is shown. The same game time as the baseline method is guaranteed in stage II while avoiding collisions.

6 | CONCLUSION

A deterministic and stochastic dynamics model for reach-avoid games is decoupled into a non-cooperative probabilistic game and a cooperative probabilistic game. With the aid of zero-sum game analysis stochastic HJI equations are constructed in the non-cooperative probabilistic game. The probability transition function facilitates finite-time stability with safety guaranteed. In the cooperative probabilistic game, probabilistic reachability analysis is used to obtain high-probability interception regions. In addition, the proposed DCFCI strategy effectively accomplishes cross-interception with collision avoidance. Finally, in order to obtain the Nash equilibrium solution, both the non-cooperative and cooperative probabilistic games are continuously iterated with feedback strategies. In this paper, two cases with stationed and unstationed attackers are considered as a verification of the method. Future work will continue to probabilistically solve the HJI equations to address the limitations of multi-agent high-dimensional games.

ACKNOWLEDGMENTS

This work was supported by the Guangdong Basic and Applied Basic Research Foundation under the Grant No. 2021A1515110569, and 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.

CONFLICT OF INTEREST

The authors declare no potential conflict of interests.

REFERENCES

1. Sujit P, Saripalli S, Sousa JB. Unmanned aerial vehicle path following: A survey and analysis of algorithms for fixed-wing unmanned aerial vehicless. *IEEE Control Systems Magazine.* 2014;34(1):42–59.
2. Motlagh NH, Taleb T, Arouk O. Low-altitude unmanned aerial vehicles-based internet of things services: Comprehensive survey and future perspectives. *IEEE Internet of Things Journal.* 2016;3(6):899–922.
3. Roberge V, Tarbouchi M, Labonté G. Comparison of parallel genetic algorithm and particle swarm optimization for real-time UAV path planning. *IEEE Transactions on industrial informatics.* 2012;9(1):132–141.
4. Dong X, Yu B, Shi Z, Zhong Y. Time-varying formation control for unmanned aerial vehicles: Theories and applications. *IEEE Transactions on Control Systems Technology.* 2014;23(1):340–348.
5. Hayat S, Yanmaz E, Muzaffar R. Survey on unmanned aerial vehicle networks for civil applications: A communications viewpoint. *IEEE Communications Surveys & Tutorials.* 2016;18(4):2624–2661.
6. Abate A, Prandini M, Lygeros J, Sastry S. Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems. *Automatica.* 2008;44(11):2724–2734.
7. Hespanha JP, Cassandras CG, Blom HA, et al. *Stochastic hybrid systems*. CRC Press, 2018.
8. Esfahani PM, Chatterjee D, Lygeros J. The stochastic reach-avoid problem and set characterization for diffusions. *Automatica.* 2016;70:43–56.
9. Leung K, Schmerling E, Zhang M, et al. On infusing reachability-based safety assurance within planning frameworks for human–robot vehicle interactions. *The International Journal of Robotics Research.* 2020;39(10-11):1326–1345.
10. Pachter M, Garcia E, Casbeer DW. Differential game of guarding a target. *Journal of Guidance, Control, and Dynamics.* 2017;40(11):2991–2998.
11. Harris MW. Abnormal and singular solutions in the target guarding problem with dynamics. *Journal of Optimization Theory and Applications.* 2020;184:627–643.
12. Chen M, Zhou Z, Tomlin CJ. Multiplayer reach-avoid games via low dimensional solutions and maximum matching. In: IEEE. 2014:1444–1449.
13. Lee S, Dullerud GE, Polak E. On the real-time receding horizon control in harbor defense. In: IEEE. 2015:3601–3606.
14. Quintero SA, Papi F, Klein DJ, Chisci L, Hespanha JP. Optimal UAV coordination for target tracking using dynamic programming. In: IEEE. 2010:4541–4546.
15. Bakolas E, Tsiotras P. Relay pursuit of a maneuvering target using dynamic Voronoi diagrams. *Automatica.* 2012;48(9):2213–2220.
16. Yan R, Duan X, Shi Z, Zhong Y, Bullo F. Matching-based capture strategies for 3D heterogeneous multiplayer reach-avoid differential games. *Automatica.* 2022;140:110207.
17. Wang C, Xie G. Limit-cycle-based decoupled design of circle formation control with collision avoidance for anonymous agents in a plane. *IEEE Transactions on Automatic Control.* 2017;62(12):6560–6567.
18. Wang C, Xia W, Xie G. Limit-cycle-based design of formation control for mobile agents. *IEEE Transactions on Automatic Control.* 2019;65(8):3530–3543.
19. Bai X, Yan W, Cao M, Xue D. Distributed multi-vehicle task assignment in a time-invariant drift field with obstacles. *IET Control Theory & Applications.* 2019;13(17):2886–2893.
20. Riehl J, Ramazi P, Cao M. A survey on the analysis and control of evolutionary matrix games. *Annual Reviews in Control.* 2018;45:87–106.
21. Benenati E, Ananduta W, Grammatico S. Optimal selection and tracking of generalized Nash equilibria in monotone games. *IEEE Transactions on Automatic Control.* 2023.
22. Sadigh D, Sastry S, Seshia SA, Dragan AD. Planning for autonomous cars that leverage effects on human actions.. In: . 2. Ann Arbor, MI, USA. 2016:1–9.
23. Zhang Y, Zhu Y, Li H, Wang J. A hybrid optimization algorithm for multi-agent dynamic planning with guaranteed convergence in probability. *Neurocomputing.* 2024;592:127764.
24. Liniger A, Lygeros J. A noncooperative game approach to autonomous racing. *IEEE Transactions on Control Systems Technology.* 2019;28(3):884–897.
25. Bhattacharyya RP, Phillips DJ, Wulfe B, Morton J, Kuefler A, Kochenderfer MJ. Multi-agent imitation learning for driving simulation. In: IEEE. 2018:1534–1539.
26. Nishimura H, Schwager M. Active motion-based communication for robots with monocular vision. In: IEEE. 2018:2948–2955.
27. Summers S, Lygeros J. Verification of discrete time stochastic hybrid systems: A stochastic reach-avoid decision problem. *Automatica.* 2010;46(12):1951–1961.
28. Lavaei A, Soudjani S, Abate A, Zamani M. Automated verification and synthesis of stochastic hybrid systems: A survey. *Automatica.* 2022;146:110617.
29. Yin M, Chu QP, Zhang Y, Niestroy MA, Visser dCC. Probabilistic flight envelope estimation with application to unstable overactuated aircraft. *Journal of Guidance, Control, and Dynamics.* 2019;42(12):2650–2663.
30. Baddeley A, Bárány I, Schneider R. Spatial point processes and their applications. *Stochastic Geometry: Lectures Given at the CIME Summer School Held in Martina Franca, Italy, September 13–18, 2004.* 2007:1–75.
31. Bhat SP, Bernstein DS. Finite-time stability of continuous autonomous systems. *SIAM Journal on Control and optimization.* 2000;38(3):751–766.
32. Polyakov A. Nonlinear feedback design for fixed-time stabilization of linear control systems. *IEEE transactions on Automatic Control.* 2011;57(8):2106–2110.
33. Haddad WM, Chellaboina V. *Nonlinear dynamical systems and control: a Lyapunov-based approach.* Princeton university press, 2008.

34. Brandt v. dR, Visser dC. Safe flight envelope uncertainty quantification using probabilistic reachability analysis. *IFAC-PapersOnLine.* 2018;51(24):628–635.
35. Assellaou M, Bokanowski O, Zidani H. Error estimates for second order Hamilton-Jacobi-Bellman equations. Approximation of probabilistic reachable sets. *Discrete and Continuous Dynamical Systems-Series A.* 2015;35(9):3933–3964.
36. Huang H, Ding J, Zhang W, Tomlin CJ. A differential game approach to planning in adversarial scenarios: A case study on capture-the-flag. In: IEEE. 2011:1451–1456.
37. Sun W, Tsiotras P, Lolla T, Subramani DN, Lermusiaux PF. Multiple-pursuer/one-evader pursuit–evasion game in dynamic flowfields. *Journal of guidance, control, and dynamics.* 2017;40(7):1627–1637.
38. Dreves A, Gerdts M. A generalized Nash equilibrium approach for optimal control problems of autonomous cars. *Optimal Control Applications and Methods.* 2018;39(1):326–342.
39. Nikaidô H, Isoda K. Note on non-cooperative convex games. *Pacific Journal of Mathematics.* 1955.
40. Yuan Q, Li S, Wang C, Xie G. Cooperative-competitive game based approach to the local path planning problem of distributed multi-agent systems. In: IEEE. 2020:680–685.
41. Kokolakis NMT, Vamvoudakis KG. Safety-aware pursuit-evasion games in unknown environments using gaussian processes and finite-time convergent reinforcement learning. *IEEE Transactions on Neural Networks and Learning Systems.* 2022.
42. Esfahani PM, Chatterjee D, Lygeros J. On a problem of stochastic reach-avoid set characterization. In: IEEE. 2011:7069–7074.