



**ARTICLE**

# Research on Maneuver Decision-Making of Multi-Agent Adversarial Game in a Random Interference Environment

Shiguang Hu<sup>1,2</sup>, Le Ru<sup>1,2,\*</sup>, Bo Lu<sup>1,2</sup>, Zhenhua Wang<sup>3</sup>, Xiaolin Zhao<sup>1,2</sup>, Wenfei Wang<sup>1,2</sup> and Hailong Xi<sup>1,2</sup>

<sup>1</sup>Equipment Management and UAV Engineering College, Air Force Engineering University, Xi'an, 710051, China

<sup>2</sup>National Key Laboratory of Unmanned Aerial Vehicle Technology, Xi'an, 710051, China

<sup>3</sup>College of Information Technology, Nanjing Police University, Nanjing, 210023, China

\*Corresponding Author: Le Ru. Email: ru-le@163.com

Received: 15 July 2024 Accepted: 29 August 2024 Published: 15 October 2024

## ABSTRACT

The strategy evolution process of game players is highly uncertain due to random emergent situations and other external disturbances. This paper investigates the issue of strategy interaction and behavioral decision-making among game players in simulated confrontation scenarios within a random interference environment. It considers the possible risks that random disturbances may pose to the autonomous decision-making of game players, as well as the impact of participants' manipulative behaviors on the state changes of the players. A nonlinear mathematical model is established to describe the strategy decision-making process of the participants in this scenario. Subsequently, the strategy selection interaction relationship, strategy evolution stability, and dynamic decision-making process of the game players are investigated and verified by simulation experiments. The results show that maneuver-related parameters and random environmental interference factors have different effects on the selection and evolutionary speed of the agent's strategies. Especially in a highly uncertain environment, even small information asymmetry or miscalculation may have a significant impact on decision-making. This also confirms the feasibility and effectiveness of the method proposed in the paper, which can better explain the behavioral decision-making process of the agent in the interaction process. This study provides feasibility analysis ideas and theoretical references for improving multi-agent interactive decision-making and the interpretability of the game system model.

## KEYWORDS

Behavior decision-making; stochastic evolutionary game; nonlinear mathematical modeling; multi-agent; maneuver

## 1 Introduction

Multi-agent systems play a crucial role in modern complex systems, such as autonomous driving of unmanned vehicles, energy management in smart grids, financial market trading strategies, and coordinated confrontation of drone swarms. In these domains, multiple agents need to coordinate and make decisions to achieve optimal overall benefits. The coordinated confrontation of drone swarms in the aerial domain, as a typical application, involves not only various maneuvers and tactical strategies of the agents but also the interactive behaviors of both sides in a constantly changing dynamic



environment. Each agent must make real-time decisions in a complex and dynamic environment to ensure maximum overall effectiveness. Ensuring that agents make optimal decisions in dynamic environments has gradually become a hot topic in current academic research.

In the aerial confrontation scenario, the strategy interaction and behavior decision-making of agents are highly complex and rapidly dynamic macro-process [1]. Influenced by the dynamic changes of the situational environment, the asymmetry of the game player's perceived information, the random and sudden situation, other external interference, and many other factors, the game evolution process of agents is highly uncertain [2-4]. These uncertain emergencies are difficult to predict effectively, which will directly affect the strategy formulation and action execution of game players and have a significant impact on the final results of the game. The level of decision-making and the formulation of operational strategies can largely determine the outcome of the game and even affect the development of the macro confrontation scenario [5]. Therefore, conducting in-depth research on the behavioral strategy evolution mechanisms of aerial game agents in environments with random disturbances, exploring the behavioral learning processes of intelligent agents, and identifying and predicting potential adversarial mechanisms and behavior patterns can lead to a better understanding and prediction of various random emergent situations in real-world aerial confrontation scenarios. This research can reveal how tactics and strategies of intelligent agents undergo natural selection and evolution under different environments and conditions. Consequently, it can aid in the development of more efficient and intelligent decision support systems, enabling strategies to adapt to continuously changing situational environments. This, in turn, enhances the adaptability and flexibility of autonomous decision-making systems, providing more accurate decision recommendations and theoretical foundations for command-control personnel to formulate more effective counter-strategies. Such advancements hold significant theoretical value and practical significance for the improvement and development of multi-agent decision-making systems in the real world.

Until now, numerous scholars have conducted research and achieved corresponding results in solving decision-making problems in aerial simulation confrontation scenarios. Specific solutions include differential games, matrix games, deep reinforcement learning, heuristic algorithms [6,7], and others. In terms of differential games, relevant literature describes the strategies and behavioral state changes of both adversarial parties in continuous time by establishing differential models of dynamic systems. In the 1960s, Isaacs [8] firstly studied the pursuit-evasion maneuvering decision problem based on differential games from a mathematical perspective and proposed an analytical method for optimal strategies. Although the research results were not entirely satisfactory due to model simplifications and the limitations of mathematical methods, it provided inspiration for future studies. Garcia et al. [9] investigated the active defense cooperative differential game problem involving three game participants, focusing on the target differential game problem of active control. Park et al. [10] proposed a method based on differential game theory to study within-visual-range air combat for Unmanned Combat Aerial Vehicles (UCAVs). This algorithm employs a hierarchical decision-making structure and uses differential game theory to compute the scoring function matrix, thereby solving for the optimal maneuvering strategies to address dynamic and challenging combat situations. These studies primarily analyze the aerial combat process in continuous time and space, enabling precise simulation of dynamic changes during air battles. However, the modeling process is complex, the models are highly dependent, the solution difficulty is high, real-time performance is low, and it is challenging to address multi-dimensional and multi-variable problems.

In terms of matrix games, related research employs a discretization approach to describe the strategy space and payoffs of the opposing sides. Austin et al. [11] was the first to utilize matrix game methods to study one-on-one low-altitude air combat maneuvering decisions in a hilly battlefield

environment. Li et al. [4] proposed a constrained-strategy matrix game approach for generating intelligent decisions for multiple UCAVs in complex air combat environments using air combat situational information and time-sensitive information. Li et al. [12] proposed a dimensionality reduction-based matrix game solving algorithm for solving large-scale matrix games in a timely manner. The modeling process of matrix games is relatively straightforward, allowing for an intuitive demonstration of the strategy choices and payoff situations of both players. However, it heavily relies on prior knowledge, lacks flexibility in application, and struggles to effectively handle continuous state and action spaces. Additionally, it is unable to respond in real-time to complex dynamic adversarial environments, making it challenging to solve large-scale scenario problems.

In terms of deep reinforcement learning, relevant research primarily utilizes deep neural networks and reinforcement learning algorithms, which are capable of addressing adversarial problems in high-dimensional and complex environments. Cao et al. [13] addressed the problem that UCAVs are difficult to quickly and accurately perceive situational information and autonomously make maneuvering decisions in modern air wars that are susceptible to the influence of complex factors, and proposes an UCAVs maneuvering decision-making algorithm combining deep reinforcement learning and game theory. Liles et al. [14] developed a Markov Decision Process model of the stochastic dynamic allocation problem for the efficient and intelligent multi-program allocation decision-making problem faced by air force combat managers and uses approximate dynamic programming techniques to learn to find a high-quality solution for the Air Battle Management Problem (ABMP). Léchevin et al. [15] proposed a hierarchical decision-making and information system designed to provide coordinated aircraft path planning and deception warfare missions in real time. Zhang et al. [16] constructed beyond-visual-range air combat training environment and proposed a heuristic Q-network method that incorporates expert experience. This method aims to improve the efficiency of reinforcement learning algorithms in exploring strategy spaces, achieving self-learning of air combat maneuver strategies. Li et al. [17] proposed a UCAV autonomous maneuvering decision-making method based on game theory that considers the partially observable state of the adversary. Poropudas et al. [18] presented a new game-theoretic approach for validating discrete-event air combat simulation models and simulation-based optimization. Sun et al. [19] proposed a novel Multi-Agent Hierarchical Policy Gradient (MAHPG) Algorithm that learns various strategies and outperforms expert cognition through adversarial self-game learning. Deep reinforcement learning exhibits strong adaptability, allowing for continuous improvement of strategies during the training process. It is capable of handling problems with continuous time and continuous state spaces. However, the training process requires substantial amounts of data and computational resources. The results of model training are highly dependent on the quality and diversity of the training data, which may lead to overfitting or underfitting issues. Furthermore, the “black box” nature of the model limits its credibility and interpretability, making it challenging to explain and understand the specific formation process of the strategies.

Additionally, there are other relevant literature analyses that discuss the dynamic evolution process of aerial adversarial games. Hu et al. [20] studied the evolutionary mechanism of Unmanned Aerial Vehicle (UAV) swarm cooperative behavioral strategies in communication-constrained environments. Gao et al. [21] proposed an evolutionary game theory-based target assignment method for multi-UAV networks in 3D scenarios. Sheng et al. [22] proposed a posture evolution game model for the UAV cluster dynamic attack-defense problem. However, these studies did not take into account the interference of random environmental factors during confrontation.

To a certain extent, these methods provide effective solutions for multi-agent game decision-making, but there are also some limitations. In particular, the existing research results are mainly

based on the premise of complete rationality of the players, focusing on solving the optimal strategy in a particular moment or situation, i.e., Nash equilibrium, failing to give full consideration to the dynamic changes in the process of confrontation, and neglecting the stochastic complexity of the environment and the uncertainty of information interference, lacking the analysis of the specific influencing factors of the dynamic game process and unable to analyze and study the whole process of the player's coping strategy selection from a global perspective. In contrast, stochastic evolutionary game theory takes into account the bounded rationality and dynamic learning processes of agents, as well as the impact of random disturbances. This allows for the simulation of participants' strategy selection processes over time, reflecting the dynamic changes in the maneuvering behaviors of agents in aerial confrontation. Therefore, this paper investigates the strategy interactions and decision-making behaviors of bounded rationality game agents in simulated adversarial scenarios under the influence of random disturbances, based on stochastic evolutionary game theory. Specifically, the main research content and contributions of this paper are as follows:

(1) Considering the risks that random disturbances in complex situational environments may pose to the autonomous decision-making of game players, as well as the impact of participants' maneuvering behaviors on the state changes of game players in simulated adversarial scenarios, a nonlinear mathematical model is established to describe the strategy decision-making process of participants in this context.

(2) Through simulation experiments, the interactive relationships of strategy selection among game players, the stability of strategy evolution, and the dynamic decision-making processes were studied and validated. The study delves into the mechanisms by which certain factors influence agent behavior choices in specific environments. The results indicate that environmental disturbance factors and parameters related to agents' maneuvering actions have varying impacts on the selection and evolution speed of agents' strategies.

(3) This research explores the decision-making behaviors of multiple agents in aerial combat games, providing a new perspective for the study of agent strategy selection mechanisms. It offers a reliable quantitative analysis foundation for the observed results and fills a research gap in this field. The findings not only help decision-makers understand the dynamic changes in strategies under different factors and environmental conditions but also provide feasible analytical approaches and theoretical references to enhance the interpretability of multi-agent interactive decision-making and game system models.

## 2 Maneuvers Discussion

In the multi-agent simulation confrontation scenario, the behavioral decision-making of the agent is a highly complex game process involving the interaction of multiple variables and factors, and the state and behavior of the agent evolve with the escalation of the conflict between both game players and the change of the situation. Among them, the maneuvers executed by the agent are crucial to the implementation of their maneuvering strategies, and the quality of the execution of these maneuvers not only directly affects the effectiveness of the strategies but also has a significant impact on the game state of the entire simulation confrontation.

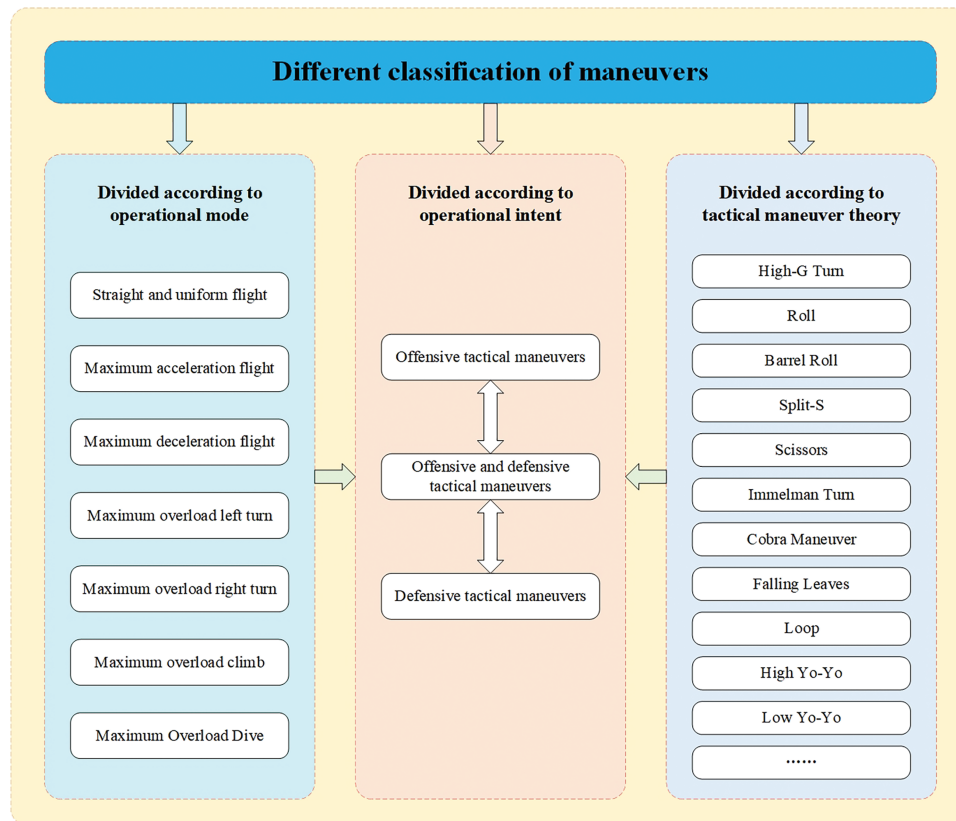
Maneuvers of the agent refer to the flight techniques and maneuvering strategies adopted by the agent in conflict to gain an advantage, avoid risk, or achieve a specific purpose. Currently, there are two common ways of dividing maneuvers: basic maneuvers based on the operational mode [23], and typical maneuvers based on tactical maneuver theory. The National Advisory Committee for Aeronautics scholars has classified the most commonly used maneuvers into seven basic maneuvers



based on the operational mode [24,25] including straight line uniform flight, maximum acceleration flight, maximum deceleration flight, maximum overload left turn, maximum overload right turn, maximum overload climb, and maximum overload dive. Complex maneuvers can be generated by combining and arranging the basic maneuvers in this group. From the point of view of tactical theory and demonstration effect, the typical maneuvers include [26,27]: High-G Turn, Roll, Barrel Roll, Split-S, Scissors, Immelman Turn, Pougatcheff Cobra maneuver, Falling Leaves, Loop, High Yo-Yo, Low Yo-Yo, and others.

However, these two types of maneuvers are still essentially two-dimensional or even multi-dimensional combinations of typical basic maneuvers such as straight flight, turn, pitch, etc., and there are some limitations in the way they are divided. In the division based on the operational mode, this mode mainly focuses on the physical operation level of the agent, oversimplifies the complexity of actual confrontation, ignores the application and effect of these maneuvers in specific situations, and does not fully consider the impact of the differences in the performance of the game players on the use of complex maneuvers. In the division based on the tactical maneuver theory, the approach is highly dependent on the specific environment, equipment performance, and the posture of both sides. It is difficult to comprehensively cover all the possible changes and the opponent's reaction, and there are discrepancies and lags between daily training and practical application, which cannot reflect the latest tactical dynamics promptly. In general, the two divisions have their focuses, but they have some common problems. In the rapidly developing and highly technological modern confrontation, neither of them can fully cover all possible maneuvers, and the maneuvers are often interrelated and affect each other, so it is difficult for a single division to reflect the interaction and synergistic effect between different maneuvers. To overcome these problems, a more integrated division is needed that takes into account both the practicality and flexibility of the operation and the theoretical and practical applicability of the maneuvers while being constantly updated to accommodate technological developments and the accumulation of practical experience.

Therefore, based on these two classifications, this paper, regarding the maneuver design of the literature [28,29], combines the strategic intent during the interactive confrontation and divides the typical strategy maneuvers commonly used by players in actual confrontation based on the characteristics and expected goals of different maneuvers, which can be classified into three specific categories: offensive maneuvers, defensive maneuvers, and offensive-defensive maneuvers. Offensive maneuvers have the primary goal of maximizing strike effectiveness, focusing on using the performance advantages and firepower systems of the game players to execute precision strikes against incoming targets, mainly including direct flight turns, target tracking, and fire attacks. Defensive maneuvers have the priority goal of ensuring the game players' safety and survival, focusing on evading the opponent's attack and protecting the game players from damage. The strategy involves a high degree of maneuverability and the ability to quickly adapt to the environment and mainly includes straight flight turns, danger warnings, and circling evasions. Offensive-defensive maneuvers can be used both as offensive strategies to enhance the effectiveness of the game players in attack, and as defensive strategies to provide effective defense flight maneuvers when necessary. The strategy has a balanced performance in terms of speed, maneuverability, concealment, and reaction speed, the maneuver is mainly loop maneuvers. The specific divisions are shown in Fig. 1.



**Figure 1: Maneuvers division**

### 3 Deterministic Evolutionary Game Model

In the multi-agent simulation confrontation scenario, both game players must be based on the ever-changing environment and a flexible choice of offensive or defensive strategies to ensure that the benefits of attacking the target and their own safety and survival are maximized. This confrontation is a dynamic decision-making process, both game players will have mutual conversion and attack-defense transposition on the situation. The game players, in the choice of strategy, must weigh the potential payoffs and possible costs in the assumption of a certain risk at the same time and strive to obtain the maximum return to ensure that every maneuver and decision-making can lay the foundation for the ultimate victory. Evolutionary game theory provides a framework for explaining strategy interactions and behavioral decisions in multi-agent systems, and it helps us understand how individuals or groups form and adjust their strategies during long-term interactions by modeling natural selection and learning processes [30].

#### 3.1 Modeling

Assuming that both game players are limited rational groups with autonomous intelligent decision-making abilities, both sides are in the range of each other's effective firepower in the confrontation process, and their strategy choices have a certain dynamic evolutionary law. Specifically, in the decision-making process, agents are capable of real-time perception of environmental changes

and opponent behaviors, independently making judgments and autonomously adjusting strategies. However, constrained by incomplete information acquisition, limited computational capabilities, and finite response times, agents cannot achieve a level of fully rational decision-making. Consequently, collective decision-making often results in locally optimal or near-optimal outcomes under specific conditions. The Multi-agent Interactive Confrontation Game Evolution Model (MAICGEM) can be expressed as a quaternion array, i.e.,  $MAICGEM = \{U, P, G, S\}$ , The details are as follows:

(1)  $U = (U_{Agent\_r}, U_{Agent\_b})$  represents both game players involved in confrontation, respectively.

(2)  $P = (P_r, P_b)$  denotes the belief space of strategies involved in the game, where the  $agent\_r$  adopts offensive and defensive maneuvering strategies corresponding to choice probabilities of  $(x, 1 - x)$ , and the  $agent\_b$  adopts offensive and defensive maneuvering strategies corresponding to choice probabilities of  $(y, 1 - y)$ ,  $0 \leq x, y \leq 1$ .

(3)  $G = (G_r, G_b)$  denotes the payoff space of the players involved in the game, where  $G_r$  denotes the payoff function of the  $agent\_r$  under different maneuvering strategies,  $G_b$  denotes the payoff function of the  $agent\_b$  under different maneuvering strategies, and both game players obtain different payoffs according to different combinations of strategies.

(4)  $S = (S_r^A, S_r^D)$  denotes the space of strategy combinations involved in the game, where  $i = r, b$ ,  $S_r = (S_r^A, S_r^D)$  represents the offensive and defensive maneuver strategies adopted by the  $agent\_r$ , respectively,  $S_b = (S_b^A, S_b^D)$  represents the offensive and defensive maneuver strategies adopted by the  $agent\_b$ , respectively.

The benefits available to both players under different tactical maneuvering strategies differ, mainly: basic flight gains for straight flight turning maneuvers  $G_i^{B-F}$ , angular posture gains for target tracking maneuvers  $G_i^{C-P}$ , fire attack gains under full-process probability conditions for fire attack maneuvers  $G_i^{Missile}$ , threat determination gains of hazard warning actions  $G_i^{T-J}$ , maneuvering defense gains from hovering evasive maneuvers  $G_i^{M-D}$ , balanced gains in offensive and defensive transitions for somersault maneuvers  $G_i^{O-D-W}$ , and flight fuel consumption costs  $V_i^{Fuel}$  and missile expenditures  $V_i^{Miss}$ . The specific form is indicated below:

$$\left\{ \begin{array}{l} G_i^{C-P} = \left(1 - \frac{\theta_i + \vartheta_i}{2\pi}\right) V_i^{tar} \\ G_i^{Missile} = \prod_{j=1}^4 P_i^j * V_i^{tar} \\ G_i^{T-J} = \frac{T_i}{10} e^{-\frac{\theta_i^2}{\pi}} V_i^{tar} \\ G_i^{M-D} = \prod_{j=1}^4 (1 - P_i^j) V_i^{tar} \\ G_i^{O-D-W} = \epsilon_1 G_i^O + \epsilon_2 G_i^D \end{array} \right. , \tag{1}$$

$$\left\{ \begin{array}{l} G_i^O = G_i^{B-F} + G_i^{C-P} + G_i^{Missile} \\ G_i^D = G_i^{B-F} + G_i^{T-J} + G_i^{M-D} \end{array} \right. , \tag{2}$$

where:

$\theta_i$  is the target azimuth,  $\vartheta_i$  is the target entry angle,  $\theta_i \in [0, \pi]$ ,  $\vartheta_i \in [0, \pi]$ , the relationship between the two is shown in Fig. 2;

$V_i^{tar}$  is the players' own inherent value;

$P_i^j$  is the probability that players detect a target;

$P_i^2$  is the surge survival probability of an air-to-air missile;

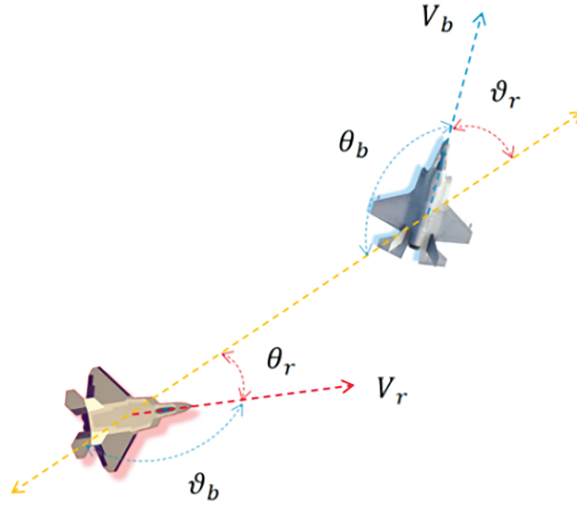
$P_i^3$  is the probability of hitting the target after an air-to-air missile breakout;

$P_i^4$  is the destructive probability of an air-to-air missile hitting a target;

$T_i$  is the time remaining to evade interceptor missiles;

$G_i^O$  represents the offensive strategy gains,  $G_i^D$  represents the defensive strategy gains.

$\epsilon_1, \epsilon_2$  are the coefficients of the proportionality of the focus on offensive-defensive action gains in the offensive-defensive switching equilibrium gains, respectively, and  $\epsilon_1 + \epsilon_2 = 1$ .



**Figure 2:** Schematic diagram of the simulation confrontation

In an evolutionary game, both game players no longer take Nash Equilibrium as the final strategy choice, but in the continuous competition, learning, and imitation, the participants with low returns will learn from the high-return players to optimize and improve their own strategies and gradually seek for Evolutionary Stable Strategy (ESS). The attack-defense game tree is shown in Fig. 3.

### 3.2 Payoff Function

From the game tree, there are four scenarios for the strategy combinations of the antagonists:  $(S_b^A, S_r^A)$ ,  $(S_b^A, S_r^D)$ ,  $(S_b^D, S_r^A)$ ,  $(S_b^D, S_r^D)$ . Due to the asymmetry of the maneuvering strategies of the two parties, the benefits that can be obtained by the two players under the conditions of different strategy combinations are different. Combined with the previous assumptions and analysis of the main strategies of both game players, the payoff matrix is established as shown in Table 1.

From Table 1, the expected gains of the *agent\_r*'s choice of offensive maneuvering strategy  $\mathcal{T}_{11}$ , and the expected gains of the defensive maneuvering strategy  $\mathcal{T}_{12}$ , and the average expected return  $\overline{\mathcal{T}}_1$  are, respectively:

$$\begin{cases} \mathcal{T}_{11} = y \left( G_r^{BF} + G_r^{CP} + G_r^{Missile} + G_r^{ODW} - V_r^{Fuel} - V_r^{Miss} \right) \\ \quad + (1 - y) \left( G_r^{BF} + G_r^{CP} + G_r^{Missile} - V_r^{Fuel} - V_r^{Miss} \right) \\ \mathcal{T}_{12} = y \left( G_r^{BF} + G_r^{TJ} + G_r^{MD} - V_r^{Fuel} \right) + (1 - y) \left( G_r^{BF} + G_r^{TJ} + G_r^{MD} - V_r^{Fuel} \right) \\ \overline{\mathcal{T}}_1 = x\mathcal{T}_{11} + (1 - x)\mathcal{T}_{12} \end{cases}, \quad (3)$$

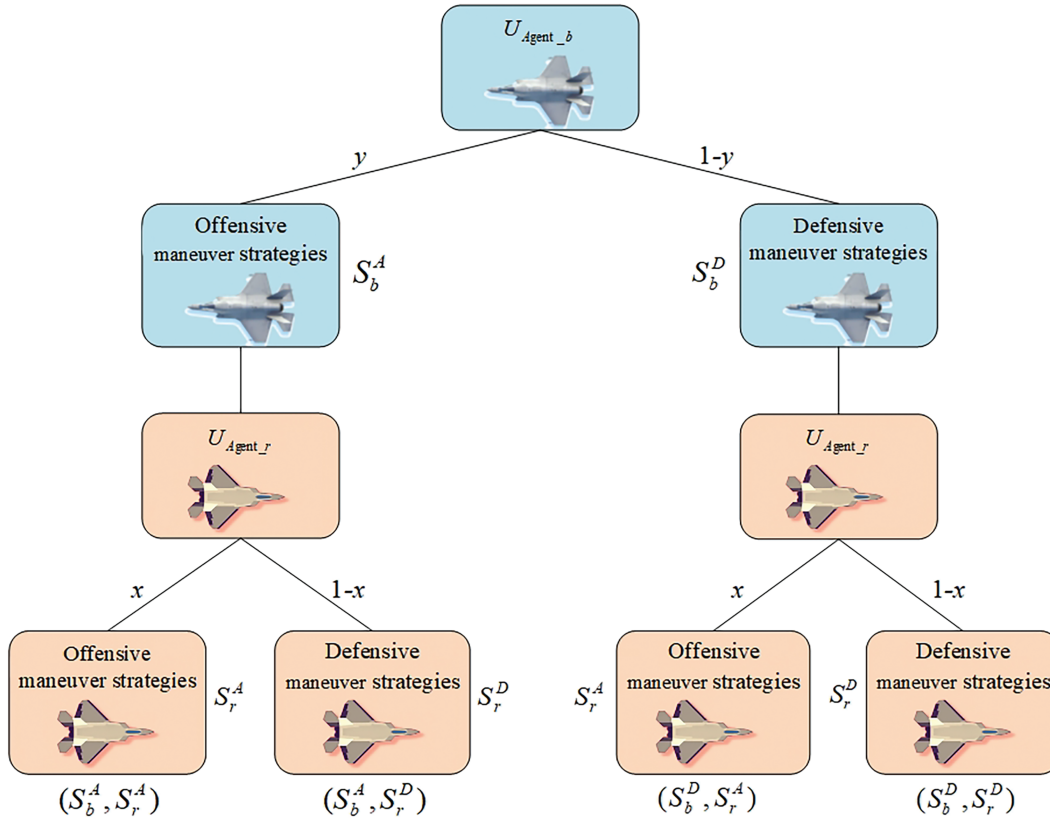


Figure 3: Confrontation game tree

Table 1: Confrontation gains matrix

Strategy combination		<i>agent_b</i>	
		$S_b^A (y)$	$S_b^D (1 - y)$
<i>agent_r</i>	$S_r^A (x)$	$G_r^{B_F} + G_r^{C_P} + G_r^{Missile} + G_r^{O_{D-W}} - V_r^{Fuel} - V_r^{Miss}$	$G_r^{B_F} + G_r^{C_P} + G_r^{Missile} - V_r^{Fuel} - V_r^{Miss}$
	$S_r^D (1 - x)$	$G_r^{B_F} + G_r^{C_P} + G_r^{Missile} + G_r^{O_{D-W}} - V_r^{Fuel} - V_b^{Miss}$	$G_r^{B_F} + G_r^{T-J} + G_r^{M-D} - V_b^{Fuel}$
		$G_r^{B_F} + G_r^{T-J} + G_r^{M-D} - V_r^{Fuel}$	$G_r^{B_F} + G_r^{T-J} + G_r^{M-D} - V_r^{Fuel}$
		$G_b^{B_F} + G_b^{C_P} + G_b^{Missile} - V_b^{Fuel} - V_b^{Miss}$	$G_b^{B_F} + G_b^{T-J} + G_b^{M-D} - V_b^{Fuel}$

Further the dynamic replication equation for the *agent\_r*'s strategy choice can be obtained as:

$$\mathcal{F}(x) = dx/dt = x(1 - x)[yG_r^{O_{D-W}} + G_r^{C_P} - G_r^{M-D} + G_r^{Missile} - G_r^{T-J} - V_r^{Miss}], \tag{4}$$

Similarly, the expected gains of the *agent\_b*'s choice of offensive maneuvering strategy  $\mathcal{T}_{21}$ , and the expected gains of the defensive maneuvering strategy  $\mathcal{T}_{22}$ , and the average expected return  $\mathcal{T}_2$  are, respectively:



$$\begin{cases} \mathcal{J}_{21} = x \left( G_b^{BF} + G_b^{CP} + G_b^{Missile} + G_b^{ODW} - V_b^{Fuel} - V_b^{Miss} \right) \\ \quad + (1-x) \left( G_b^{B-F} + G_b^{C-P} + G_b^{Missile} - V_b^{Fuel} - V_b^{Miss} \right) \\ \mathcal{J}_{22} = x \left( G_b^{B-F} + G_b^{T-J} + G_b^{M-D} - V_b^{Fuel} \right) + (1-x) \left( G_b^{B-F} + G_b^{T-J} + G_b^{M-D} - V_b^{Fuel} \right) \\ \overline{\mathcal{J}}_2 = y\mathcal{J}_{21} + (1-y)\mathcal{J}_{22} \end{cases}, \quad (5)$$

Further the dynamic replication equation for *agent\_b*'s strategy choice can be obtained as:

$$\mathcal{F}(y) = dy/dt = y(1-y)[xG_b^{O-D-W} + G_b^{C-P} - G_b^{M-D} + G_b^{Missile} - G_b^{T-J} - V_b^{Miss}], \quad (6)$$

## 4 Stochastic Evolutionary Game Model

### 4.1 Stochastic Evolutionary Game

Although the evolutionary game theory overcomes the shortcoming of the classical game theory which the decision-maker is completely rational, it discusses the strategic behavior in a deterministic system [31], and the multi-agent simulation confrontation is a macro-complex system with uncertainty and randomness in its evolving dynamical mechanism. In actual confrontation, the change of environment and the interference of other external factors have a certain degree of randomness, so the game decision-making is also affected by the dynamic change of the complex environment, and the traditional evolutionary game theory lacks the research on the strategy dynamic evolution of the game players in a random dynamic interference environment, which cannot describe the uncertainty and random dynamics of the complex environment. Gaussian white noise is a nonlinear random noise whose amplitude obeys the Gaussian distribution and power spectral density obeys the uniform distribution, which can better describe the random external environment [32,33]. Therefore, this paper introduces the Gaussian white noise, which is combined with the stochastic differential equations [34] for describing the various types of random interference factors in complex adversarial environments [35]. Considering  $x, y \in [0, 1]$ , then  $1-x$  and  $1-y$  are also non-negative,  $1-x, 1-y \in [0, 1]$ , it has no effect on the evolutionary outcome of the game system [36]. Combining the above factors, the replicated dynamic Formulas (4) and (6) are simplified as well as improved to obtain the following equations:

$$dx(t) = x(t) [yG_r^{O-D-W} + G_r^{C-P} - G_r^{M-D} + G_r^{Missile} - G_r^{T-J} - V_r^{Miss}] dt + \delta_1 x(t) d\omega(t), \quad (7)$$

$$dy(t) = y(t) [xG_b^{O-D-W} + G_b^{C-P} - G_b^{M-D} + G_b^{Missile} - G_b^{T-J} - V_b^{Miss}] dt + \delta_2 y(t) d\omega(t), \quad (8)$$

where  $\omega(t)$  is the standard one-dimensional Brownian motion, obeying the normal distribution  $N(0, t)$ , the motion is a kind of random motion without rules and order, which is able to better portray the influence of various random interference factors in complex adversarial environments;  $d\omega(t)$  denotes Gaussian white noise, which is used to represent various random interference factors in the complex adversarial environment, and its increment  $\Delta\omega(t) = \omega(t+h) - \omega(t)$  obeys a normal distribution  $N(0, \sqrt{h})$  when  $t > 0$  and the step size  $h > 0$ ;  $\delta_1$  and  $\delta_2$  are normal numbers indicating the intensity coefficient of random environmental disturbances.

Accordingly, it can be seen that Formulas (7), (8) are both one-dimensional *Itô* stochastic differential equations, which represent the stochastic differential replicated dynamic equations of the two game players after being perturbed by various random disturbances in the complex environment, respectively. *Itô* stochastic differential formula is a formulation for stochastic calculus that describes the fluctuating variation of a stochastic process [37,38]:

$$dX(t) = F(t, X(t))dt + G(t, X(t))d\omega(t), \quad (9)$$

where  $t \in [t_0, T]$ ,  $X(t_0) = X_0$ ,  $X_0 \in R$ .

Since  $It\hat{o}$  stochastic differential formula is a nonlinear stochastic equation, it cannot be solved analytically directly. So it is necessary to perform a stochastic Taylor expansion which  $t \in [t_0, T]$  is equidistributed according to the step size  $h = (T - t_0) / N$ ,  $N$  is the number of samples,  $t_n = t_0 + nh$  is the  $n$ th sampling time point, and  $n = 1, 2, \dots, N$ . Then, a stochastic Taylor expansion of [Formula \(9\)](#) is as follows [\[39\]](#):

$$X(t_{n+1}) = X(t_n) + \Gamma_0 F(X(t_n)) + \Gamma_1 G(X(t_n)) + \Gamma_{11} K^1 G(X(t_n)) + \Gamma_{00} K^0 F(X(t_n)) + R, \tag{10}$$

where,  $\Gamma_0 = h$ ,  $\Gamma_1 = \Delta\omega_n$ ,  $\Gamma_{11} = [(\Delta\omega_n)^2 - h] / 2$ ,  $\Gamma_{00} = h^2 / 2$ ,  $K^1 = \partial G(X) / \partial X$ ,  $K^0 = \partial F(X) / \partial X + (1/2) \partial^2 G(X) / \partial X^2$ ,  $R$  is the remainder term of the expansion.

Taylor expansion is the basis for the numerical solution of stochastic differential equations. In practice, we generally use discrete stochastic processes to approximate the continuous solutions of stochastic differential equations. There are two main solution methods, Euler method and Milstein method. Euler method is more direct in the solution process, and Milstein method uses  $It\hat{o}$  formula to increase the second-order terms with higher estimation accuracy. Based on this, the stochastic differential equations are solved by Milstein method, they are solved numerically after intercepting some of the terms based on the stochastic Taylor expansion. The details are as follows [\[39\]](#):

$$X(t_{n+1}) = X(t_n) + hF(X(t_n)) + \Delta\omega_n G(X(t_n)) + \frac{1}{2} [(\Delta\omega_n)^2 - h] G(X(t_n))G'(X(t_n)), \tag{11}$$

Referring to the [Formulas \(11\), \(7\) and \(8\)](#) are expanded and solved to obtain as follows:

$$\begin{aligned} x_{n+1} = & x_n + x_n [y_n G_r^{O-D-W} + G_r^{C-P} - G_r^{M-D} + G_r^{Missile} - G_r^{T-J} - V_r^{Miss}] h \\ & + \Delta\omega_n \delta_1 x_n + \frac{1}{2} [(\Delta\omega_n)^2 - h] \delta_1 x_n \delta_1 x'_n, \end{aligned} \tag{12}$$

$$\begin{aligned} y_{n+1} = & y_n + y_n [x_n G_b^{O-D-W} + G_b^{C-P} - G_b^{M-D} + G_b^{Missile} - G_b^{T-J} - V_b^{Miss}] h \\ & + \Delta\omega_n \delta_2 y_n + \frac{1}{2} [(\Delta\omega_n)^2 - h] \delta_2 y_n \delta_2 y'_n, \end{aligned} \tag{13}$$

#### 4.2 System Stability Criteria

Aiming at the possible equilibrium solutions in the confrontation game system, the stability analysis of both game players is carried out according to the stability discrimination theorem of stochastic differential equations [\[40\]](#). Let there exist a continuously differentiable smooth function  $V(t, x)$  and positive constants  $c_1, c_2$  such that [\[41\]](#):

$$\begin{cases} c_1 |x|^p \leq V(t, x) \\ c_2 |x|^p \geq V(t, x) \end{cases}, \text{ where } t \geq 0, \tag{14}$$

(1) If there exists a positive constant  $\lambda$  such that  $LV(t, x) \leq -\lambda V(t, x), t \geq 0$ , then the Zero-resolution P-order moment index of the  $It\hat{o}$  stochastic differential [Formula \(10\)](#) is stable, and holds  $E|x(t, x_0)|^p < \frac{c_2}{c_1} |x_0|^p e^{-\lambda t}, t \geq 0$ .

(2) If there exists a positive constant  $\lambda$  such that  $LV(t, x) \geq -\lambda V(t, x), t \geq 0$ , then the Zero-resolution P-order moment index of the  $It\hat{o}$  stochastic differential [Formula \(10\)](#) is unstable, and holds  $E|x(t, x_0)|^p \geq \frac{c_2}{c_1} |x_0|^p e^{-\lambda t}, t \geq 0$ .

Among them,  $LV(t, x) = V_t(t, x) + V_x(t, x)F(t, x) + \frac{1}{2}G^2(t, x) V_{xx}(t, x)$ .

For [Formulas \(7\) and \(8\)](#), let  $V_t(t, x) = x$ ,  $V_t(t, y) = y$ ,  $0 \leq x, y \leq 1$ ,  $c_1 = c_2 = 1$ ,  $p = 1$ ,  $\lambda = 1$ , we get  $LV(t, x) = F(t, x)$ , and then:

$$LV(t, x) = [yG_r^{O-D-W} + G_r^{C-P} - G_r^{M-D} + G_r^{Missile} - G_r^{T-J} - V_r^{Miss}]x, \quad (15)$$

$$LV(t, y) = [xG_b^{O-D-W} + G_b^{C-P} - G_b^{M-D} + G_b^{Missile} - G_b^{T-J} - V_b^{Miss}]y, \quad (16)$$

If the Zero-solution P-order moment index of [Formulas \(15\) and \(16\)](#) are stable, the following requirements need to be satisfied:

$$yG_r^{O-D-W} + G_r^{C-P} - G_r^{M-D} + G_r^{Missile} - G_r^{T-J} - V_r^{Miss} \leq -1, \quad (17)$$

$$xG_b^{O-D-W} + G_b^{C-P} - G_b^{M-D} + G_b^{Missile} - G_b^{T-J} - V_b^{Miss} \leq -1, \quad (18)$$

If the Zero-solution P-order moment index of [Formulas \(15\) and \(16\)](#) are unstable, the following requirements need to be satisfied:

$$yG_r^{O-D-W} + G_r^{C-P} - G_r^{M-D} + G_r^{Missile} - G_r^{T-J} - V_r^{Miss} \geq -1, \quad (19)$$

$$xG_b^{O-D-W} + G_b^{C-P} - G_b^{M-D} + G_b^{Missile} - G_b^{T-J} - V_b^{Miss} \geq -1, \quad (20)$$

Simplifying [Formulas \(17\) and \(18\)](#), respectively, then we can obtain the conditions under which the [Formulas \(15\) and \(16\)](#) satisfy the stability of the Zero-solution P-order moment index:

- When  $G_r^{O-D-W} > 0$ ,  $y \leq [-1 - (G_r^{C-P} - G_r^{M-D} + G_r^{Missile} - G_r^{T-J} - V_r^{Miss})]/G_r^{O-D-W}$ , and  $-1 - (G_r^{C-P} - G_r^{M-D} + G_r^{Missile} - G_r^{T-J} - V_r^{Miss}) - G_r^{O-D-W} \leq 0$ ;
- When  $G_r^{O-D-W} < 0$ ,  $y \geq [-1 - (G_r^{C-P} - G_r^{M-D} + G_r^{Missile} - G_r^{T-J} - V_r^{Miss})]/G_r^{O-D-W}$ , and  $-1 - (G_r^{C-P} - G_r^{M-D} + G_r^{Missile} - G_r^{T-J} - V_r^{Miss}) - G_r^{O-D-W} \geq 0$ ;
- When  $G_b^{O-D-W} > 0$ ,  $x \leq [-1 - (G_b^{C-P} - G_b^{M-D} + G_b^{Missile} - G_b^{T-J} - V_b^{Miss})]/G_b^{O-D-W}$ , and  $-1 - (G_b^{C-P} - G_b^{M-D} + G_b^{Missile} - G_b^{T-J} - V_b^{Miss}) - G_b^{O-D-W} \leq 0$ ;
- When  $G_b^{O-D-W} < 0$ ,  $x \geq [-1 - (G_b^{C-P} - G_b^{M-D} + G_b^{Missile} - G_b^{T-J} - V_b^{Miss})]/G_b^{O-D-W}$ , and  $-1 - (G_b^{C-P} - G_b^{M-D} + G_b^{Missile} - G_b^{T-J} - V_b^{Miss}) - G_b^{O-D-W} \geq 0$ .

In summary, the stochastic evolutionary game system consisting of the two parties involved in the game satisfies the condition of exponential stability of the zero-solution expectation moments as  $(a \cup b) \cap (c \cup d)$ . Similarly, the zero-solution expected moment index instability condition satisfying [Formulas \(15\) and \(16\)](#) can be obtained.

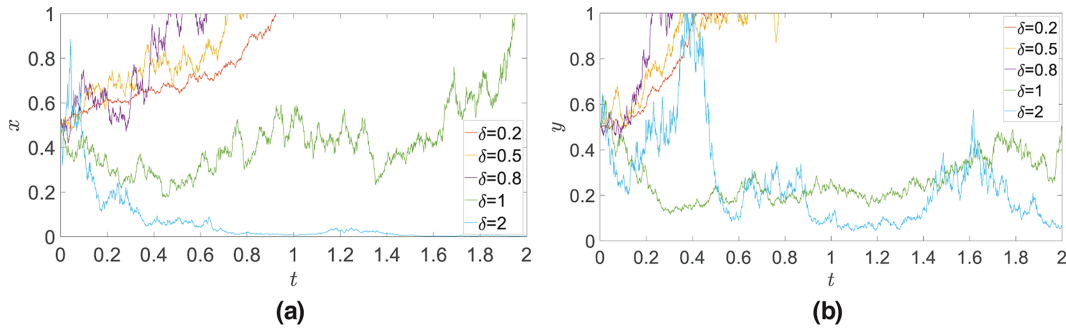
## 5 Numerical Simulation Analysis

In order to confirm the correctness of the previous theoretical derivation and verify the validity of the model, the decision-making evolution process of the two players of the game is simulated and analyzed here. Before performing the numerical simulation, it is necessary to initialize the numerical simulation by assigning values according to the system stability constraints, so the following values are taken according to the constraints in the previous section. The initial selection probability of both players of the game is set to 0.5, the intensity coefficients of random environmental disturbances are taken as  $\delta_i = 0.5, 1, 2$ , the step size of simulation is  $h = 0.001$ , and the rest of the parameters are set to values that satisfy the stability of [Formulas \(17\) and \(18\)](#) according to the constraints. Let the initial state of the confrontation game both sides are in the condition of state equivalence, at this time the specific values of the relevant parameters are given as:  $G_r^{B-F} = 10$ ,  $\theta_r = 0$ ,  $\vartheta_r = \pi$ ,  $V_r^{tar} = 100$ ,

$$P_r^1 = P_r^2 = P_r^3 = P_r^4 = 0.8, T_r = 15, V_r^{Fuel} = 1, V_r^{Miss} = 5; G_b^{B.F} = 10, \theta_b = 0, \vartheta_b = \pi, V_b^{tar} = 100, P_b^1 = P_b^2 = P_b^3 = P_b^4 = 0.8, T_b = 15, V_b^{Fuel} = 1, V_b^{Miss} = 5; \epsilon_1 = \epsilon_2 = 0.5.$$

### 5.1 Interference Intensity Factors

Fig. 4 shows the dynamic evolution process of the game players' strategy in different random environmental interference intensities. From the figure, it can be seen that the magnitude of the interference intensity has a huge impact on the evolutionary results of the system. When the interference intensity is small, the system evolution results show different degrees of volatility, but they can still maintain a certain stability and consistency in general. This stability enables the system to evolve around one or some dominantly advantageous strategies. Although there are fluctuations in the process, these fluctuations usually do not change the overall evolution direction of the system, and the evolution results show a progressive stability state. With the continuous increase of interference intensity, the internal stability of the system is destroyed, and the system may enter a more dynamic and unstable state. In this case, the strategy selection of the game players becomes more difficult to predict and control. The strategy evolution shows persistent fluctuations, decision wandering and even reversals. This phenomenon reveals the important influence of the uncertainty in random environment on strategy selection and the challenge of maintaining or adjusting strategies in complex and dynamic environments.



**Figure 4:** Dynamic evolution process of game player's strategy under uncertainty conditions (a)  $x$  evolutionary path under different interference intensities (b)  $y$  evolutionary path under different interference intensities

Moreover, the evolutionary trends and rates exhibited by the curves at different time points and under varying interference intensities differ, with some time segments even displaying reversal phenomena. These reversals are manifestations of the impact of random environmental interference on agents' decision-making process. The degree of curve fluctuation varies, with greater random interference leading to larger fluctuations and a higher probability of reversals. Furthermore, the occurrence of random environmental interference at any given moment is stochastic, with random variables at two different time points being not only uncorrelated but also statistically independent. This induces uncertainty and randomness in the system's evolutionary process both temporally and spatially, thereby increasing the overall unpredictability of the system's evolution. Consequently, strategy evolution curves exhibit diverse fluctuations or reversals at different time points. Over extended time periods, despite the influence of random environmental interference, these fluctuations typically do not alter the overall evolutionary direction of the system. The agents' strategy evolution ultimately converges to a certain strategic equilibrium state. Furthermore, the system's evolutionary process exhibits a "threshold effect" characteristic of nonlinear dynamics, with a nonlinear relationship between interference intensity and system response. As a result, the system's response characteristics differ under various interference levels. Smaller interference may induce minor fluctuations in the

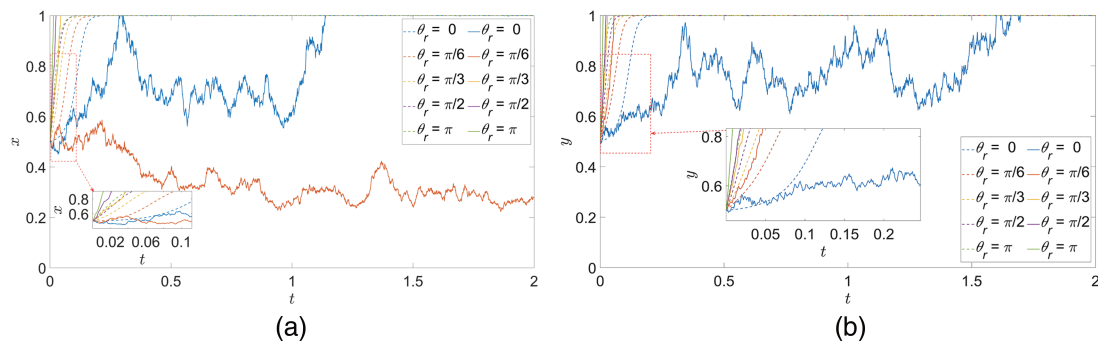
system, particularly near certain thresholds where significant changes in system behavior may occur, such as accelerated response speeds, increased amplitudes, or state transitions. Conversely, larger interference may exceed the system's threshold range, triggering more complex nonlinear behaviors and leading to system state transformations.

## 5.2 System Sensitivity Analysis

The strategy selection of the game players will be affected by the comprehensive influence of many factors. According to the previous analysis of the strategy stability of both game players, the key parameters are selected and assigned different values to further explore the impact of the value of the key factors on the strategy of the game players caused by the trajectory of the stochastic evolution of the trajectory. The selection of parameters is based on the initial assignment, and the stochastic environmental interference intensity coefficient is uniformly valued as  $\delta_i = 0.5$ , which is combined with the stochastic differential equations to analyze the reasonable values of each parameter within the constraint range. Considering that the stochastic differential equations of the two players of the game have similarity in mathematical form, the *agent\_r* player as an example only needs to analyze the evolution of the player of one side of the game, and the *agent\_b* player of the factor analysis process is similar.

### (1) Target azimuth $\theta_r$

In the case of ensuring that the initial values of other parameters remain unchanged, only the value of  $\theta_r$  is changed to analyze the impact of different values of parameter  $\theta_r$  on the evolutionary results of the game. Let  $\theta_r$  take the value  $0, \pi/6, \pi/3, \pi/2, \pi$ , respectively, the simulation results are shown in Fig. 5. Where the dashed line indicates the deterministic evolution curve and the solid line indicates the stochastic evolution curve. As can be seen from the figure, in the absence of environmental uncertainty factors interference, the initial state of the game on both sides of the players is to choose offensive tactical strategy, the change of the target azimuth did not cause subversive impact on the strategy selection of the both game players, the azimuth angle changes from small to large is only to accelerate the rate of strategy evolution. When disturbed by the environmental uncertainty factors, the strategy evolution of the *agent\_r* in the small azimuth angle was greatly affected, which might lead to the opposite strategy result, while the strategy choice of the *agent\_b* player did not undergo a big change. In the large azimuth angle, the rate of strategy evolution of the two game players was accelerated significantly.



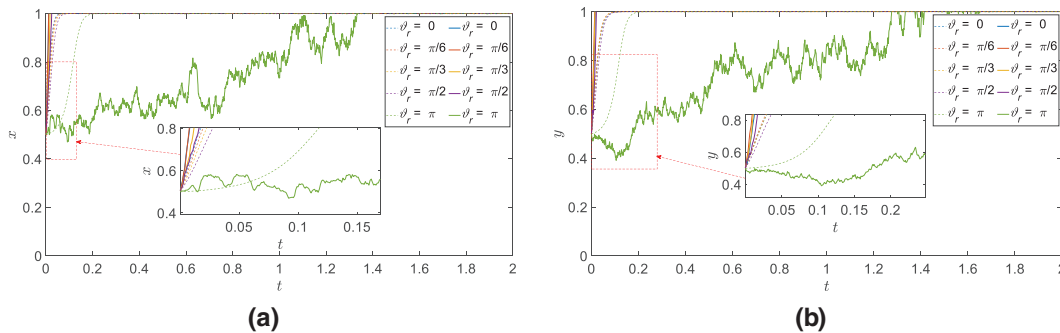
**Figure 5:**  $\theta_r$  impact on the evolution of behavioral decisions on game players (a)  $\theta_r$  impact on the evolution of behavioral decision-making in *agent\_r* (b)  $\theta_r$  impact on the evolution of behavioral decision-making in *agent\_b*



Considering the reason for this phenomenon, it is due to the fact that the target azimuth angle directly affects the relative positions of the two players, the possible engagement distance, and the acquisition and processing of information, and the change of the azimuth angle directly affects the perceptual judgment of the two sides of the game on the battlefield environment, which in turn affects the initial choice of the game player to the strategy and adjustment. In the small azimuth, it means that the two players are more directly face-to-face conflict, the two players of the game in this scenario are in the optimal working interval of the airborne electronic sensing system and air-to-air missiles, at this time there is more time and space to respond to changes in the target's strategy. In addition, due to the frequent and large amounts of information updates and rapid and dynamic strategy adjustments, decision-making requires more time to assess the effects of different strategies and possible consequences, resulting in their relatively low rate of strategy evolution. Moreover, the players under this layout is more sensitive to the perception of environmental uncertainty and random interference, which increases the complexity and unpredictability of strategy implementation, leading to greater fluctuations in its strategy selection, and its strategy adjustment becomes more conservative or hesitant, which may result in wavering and even cause strategy results opposite to those expected, further reducing the rate of strategy evolution. Under the large azimuth angle, the initial layout of the two players is more lateral or back-to-back. Although the urgency of direct conflict is reduced, the two players need to adjust their strategies more quickly to adapt to the other player's possible flanking or backward actions and to maintain their own positional advantage, which leads to a faster rate of strategy evolution. Especially the environmental uncertainty can affect the perception and judgment of the two players more significantly under the layout, which further speeds up the strategy evolution rate of the game players.

(2) Target entry angle  $\vartheta_r$

In the case of ensuring that the initial values of other parameters remain unchanged, only the value of  $\vartheta_r$  is changed to analyze the impact of different values of parameter  $\vartheta_r$  on the evolutionary results of the game. Let  $\vartheta_r$  take the value  $0, \pi/6, \pi/3, \pi/2, \pi$ , respectively, the simulation results are shown in Fig. 6. As can be seen from the figure, in the absence of environmental uncertainty factors interference, the initial state of the both game players are to choose offensive tactical strategy, the change of the target entry angle did not cause subversive impact on the strategy selection of the both game players, but only affects the rate of strategy evolution. When disturbed by environmental uncertainty factors, the environmental random interference accelerated the strategy evolution rate of the both game players in the small entry angle and reduced the strategy evolution rate in the large entry angle.

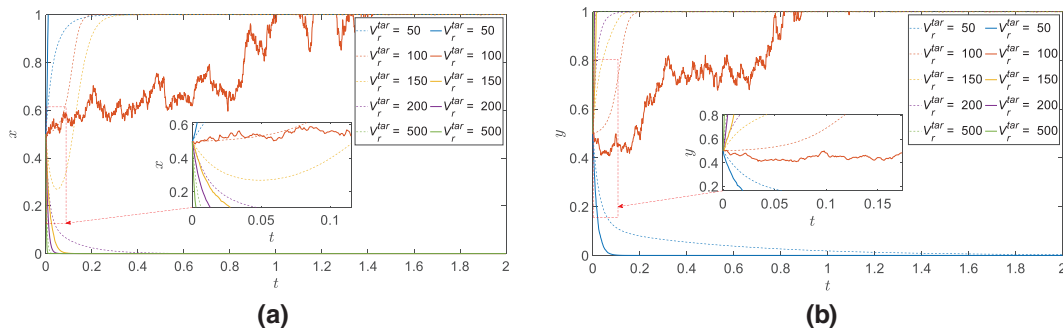


**Figure 6:**  $\vartheta_r$  impact on the evolution of behavioral decisions on game players (a)  $\vartheta_r$  impact on the evolution of behavioral decision-making in *agent\_r* (b)  $\vartheta_r$  impact on the evolution of behavioral decision-making in *agent\_b*

Consider that the phenomenon is due to the fact that different target entry angles affect the perception and judgment of both players, which in turn affects the timeliness of perception, decision-making and action. Small target entry angle means that both game players are in a tailgate situation, and the smaller the entry angle, the easier the target is to escape maneuvers. when the game players need to make quick decisions to achieve the target's strikes to intercept, especially by the random factors of the environment, the environmental disturbances may exacerbate the uncertainty of the two players of the strategy when the small entry angle, resulting in the urgency of the decision-making and the frequency of the decision-making is increased, and the need for faster information processing capabilities, which in turn cause the game players must adjust their strategies faster to adapt to the rapid response of the opponent's maneuvers and tactics as well as the constant changes of the battlefield environment. For large target entry angles, it means that both players of the game are in a head-on posture, when the missile attack area has a larger range and a better angular posture, so there is more time and space to respond to changes in the environment and target strategy.

### (3) Inherent value of the game player $V_r^{tar}$

In the case of ensuring that the initial values of other parameters remain unchanged, only the value of  $V_r^{tar}$  is changed to analyze the impact of different values of parameter  $V_r^{tar}$  on the evolutionary results of the game. Let  $V_r^{tar}$  take the value 50, 100, 150, 200, 500, respectively, the simulation results are shown in Fig. 7. As can be seen from the figure, under the premise of no environmental uncertainty factors interference, when the inherent value of the *agent\_r* is small, they tend to choose offensive tactical strategies, while the inherent value of the *agent\_b* player is higher than that of the *agent\_r*, so they tend to choose defensive tactical strategies. with the increase of the *agent\_r* player's own value, the *agent\_r* is more conservative in the selection of tactical strategies, and begins to evolve towards defensive tactical strategies, at this time the *agent\_b* side is more inclined to choose offensive tactics. When disturbed by environmental uncertainties, the *agent\_r* player's strategy evolution results in large fluctuations with the change of its own intrinsic value, which seriously affects the *agent\_r* player's strategy choice, while the *agent\_b* player's strategy evolution is relatively less affected by environmental factors and only changes the rate of strategy evolution.



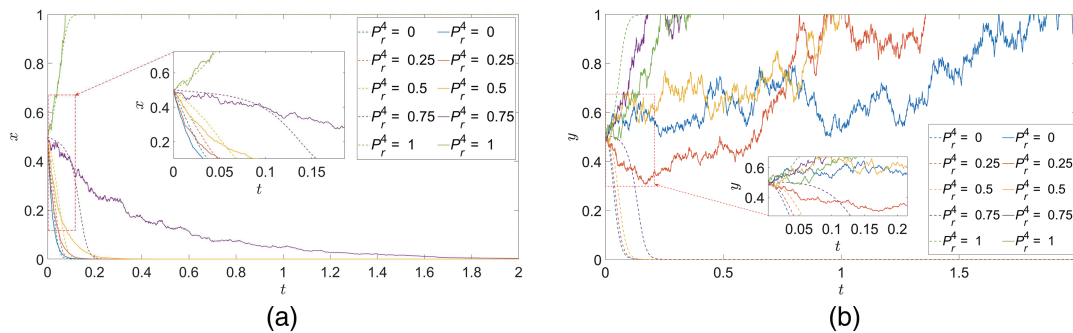
**Figure 7:**  $V_r^{tar}$  impact on the evolution of behavioral decisions on game players (a)  $V_r^{tar}$  impact on the evolution of behavioral decision-making in *agent\_r* (b)  $V_r^{tar}$  impact on the evolution of behavioral decision-making in *agent\_b*

Considering the reason for this phenomenon, it is because when the inherent value of the agent player is small, even if the agent is shot down and destroyed by the other side, its own loss is relatively small, so it is more aggressive and adventurous in the choice of strategy in order to seek to maximize the benefits or quickly change the battlefield situation. With the increase in the value of the agent player,

the cost of the loss is also increased accordingly, so it is more conservative and hesitant in the tactics and pays attention to the cost-benefit calculation trade-offs and risk management to reduce and avoid potential losses. In addition, high-value agent players are more risky in uncertain environments, and benefit analysis becomes more complex and the consideration of risk more delicate, so they are more prudent in decision-making, leading to an increase in the volatility of tactical choices. This is one of the key reasons why countries around the world are now more willing to adopt high-volume, low-cost unmanned aerial vehicle to replace manned aerial vehicle.

(4) Basic probability of fire-strike process  $P_i^A$

In the case of ensuring that the initial values of other parameters remain unchanged, only the value of  $P_i^A$  is changed to analyze the impact of different values of parameter  $P_i^A$  on the evolutionary results of the game. Taking the weapon damage probability  $P_r^A$  as a reference for illustration, so that  $P_r^A$  take the value 0, 0.25, 0.5, 0.75, 1, respectively, the simulation results are shown in Fig. 8. As can be seen from the figure, under the premise of no environmental uncertainties, the both game players tend to choose offensive tactical strategies. With the gradual increase of the weapon damage probability, the both game players tend to choose offensive tactical strategies. When interfered by environmental uncertainty factors, the random environmental interference factors have less influence on the strategy evolution of the *agent\_r* player, only causing a certain degree of interference, affecting the rate of strategy evolution, while the *agent\_b* side has a great influence, resulting in greater fluctuations in the strategy evolution results, and the reversal of the strategy results phenomenon.



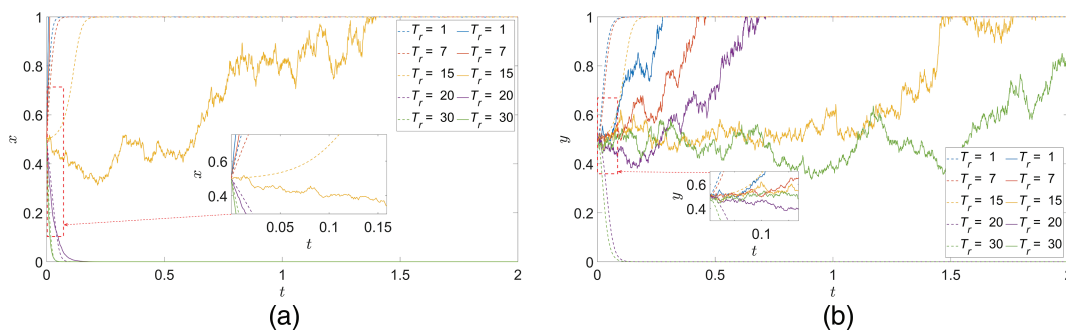
**Figure 8:**  $P_r^A$  impact on the evolution of behavioral decisions on game players (a)  $P_r^A$  impact on the evolution of behavioral decision-making in *agent\_r* (b)  $P_r^A$  impact on the evolution of behavioral decision-making in *agent\_b*

Consider that the reason for this phenomenon is that because the lower weapon damage probability means that the *agent\_r* player have difficulty effectively destroying targets when attacking, it may be safer to choose a conservative, defensive strategy, whereas the *agent\_b* player's tactical choices are dependent on predicting and reacting to the *agent\_r* player's strategy, and it may be that the *agent\_b* player does not appear to be choosing an offensive tactical strategy as expected, due to tactical flexibility, risk aversion, or other strategic considerations. The higher weapon damage probability increases the success rate of the *agent\_r* player when attacking, making attacking a more attractive option, and the *agent\_r* player begins to turn to more offensive tactics in order to gain a greater advantage in confrontation, while the *agent\_b* player responds to this threat by adjusting its tactics to be more offensive in order to try to pre-empt or at least maintain equilibrium in order to prevent being overpowered by the other side. In the face of uncertainties in environmental factors, the *agent\_r* player's strategy evolution is less affected, probably because it has already adopted a more conservative

strategy and is less sensitive to environmental changes. On the other hand, the *agent\_b* player may be more sensitive to environmental changes due to its higher weapon damage probability compared to the *agent\_r*, and the uncertainty of the environment causes the *agent\_b* player to incorrectly assess the opponent's strength at certain moments or overreact in strategy adjustment, which leads to fluctuation or even reversal of the strategy evolution results. This phenomenon also reflects that tactical decision-making on the battlefield is the result of a combination of factors, which needs to be combined with the specific battlefield environment, strategic objectives, tactical considerations, as well as the psychology and behavior of the two sides of the game, and other factors to conduct a comprehensive analysis, especially in highly uncertain environments, where even minor information asymmetries or miscalculations may have a significant impact on decision-making.

(5) Remaining time to evade missiles  $T_r$

In the case of ensuring that the initial values of other parameters remain unchanged, only the value of  $T_r$  is changed to analyze the impact of different values of parameter  $T_r$  on the evolutionary results of the game. Let  $T_r$  take the value 1, 7, 15, 20, 30, respectively, the simulation results are shown in Fig. 9. As can be seen from the figure, under the premise of no environmental uncertainties, when the remaining time for the *agent\_r* player to avoid the interception missiles is small, both game players evolve towards offensive tactical strategies, as the remaining time for the warplanes to avoid the interception increases, both game players evolve towards defensive tactical strategies. When disturbed by environmental uncertainties, the random environmental disturbances have a small impact on the strategy evolution of the *agent\_r* player, only affecting the rate of strategy evolution, while having a significant impact on the *agent\_b* player, not only affecting the rate of strategy evolution, but also resulting in greater fluctuations in the results of strategy evolution, and the reversal of the strategy results phenomenon.



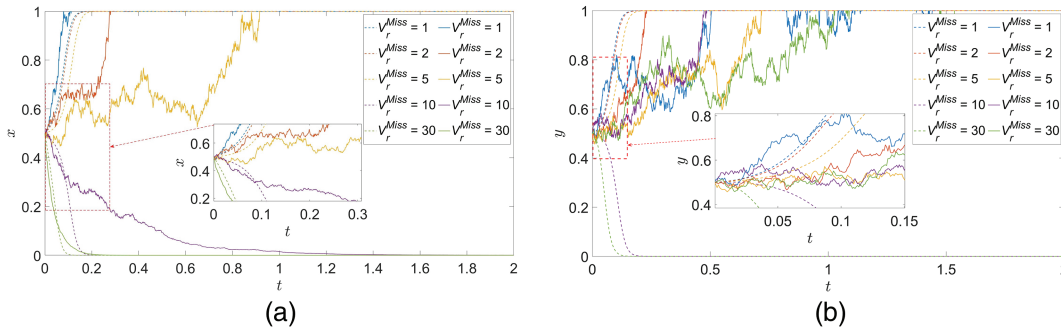
**Figure 9:**  $T_r$  impact on the evolution of behavioral decisions on game players (a)  $T_r$  impact on the evolution of behavioral decision-making in *agent\_r* (b)  $T_r$  impact on the evolution of behavioral decision-making in *agent\_b*

Consider that the reason for this phenomenon is due to the fact that when the remaining time to evade the interceptor missiles is shorter, it means that the *agent\_r* player is facing an immediate threat to its survival, and the time urgency is higher, and this urgency prompts the *agent\_r* player to adopt a more risky offensive strategy, striving to quickly break the deadlock or at least distracting the *agent\_b* player's attention in order to strive for a chance of escape, and this may lead to the fact that both the *agent\_r* and the *agent\_b* players are inclined to adopt an offensive tactical strategy in order to quickly resolve battles or gain an advantage. Moreover, the game players need to make quick decisions in a short time window, and offense is often the preferred strategy when time is of the essence. As the remaining time increases, both game players may have more time to assess the

battlefield situation and formulate their strategies, which may lead to a tendency for both players to adopt a more cautious defensive tactical strategy, with the longer period of time permitting more opportunities for both players to make tactical adjustments and reassess risks. In the face of uncertain changes in environmental factors, the *agent\_r* player's strategy evolution is less affected, which may be because the *agent\_r* has already formulated corresponding strategies based on the remaining time and these strategies are less sensitive to environmental changes. While the *agent\_b* player may show higher sensitivity to environmental uncertainty and need to continuously adjust its strategy to adapt to the changes in the *agent\_r* player's behavior, it is difficult to accurately judge the *agent\_r* player's true intentions and capabilities, which increases the complexity of strategy evolution, leading to excessive or insufficient adjustment of the strategies, resulting in fluctuations in the rate of strategy evolution and reversal of results.

(6) Missile consumption costs  $V_r^{Miss}$

In the case of ensuring that the initial values of other parameters remain unchanged, only the value of  $V_r^{Miss}$  is changed to analyze the impact of different values of parameter  $V_r^{Miss}$  on the evolutionary results of the game. Let  $V_r^{Miss}$  take the value 1, 2, 5, 10, 30, respectively, the simulation results are shown in Fig. 10. As can be seen from the figure, under the premise of no environmental uncertainties, when the missile consumption costs of the *agent\_r* player is low, both players of the game evolve towards offensive tactical strategies, as the missile consumption costs increases, both players of the game evolve towards defensive tactical strategies. When disturbed by environmental uncertainties, stochastic environmental disturbances have a small impact on the strategy evolution of *agent\_r* player, only affecting the rate of strategy evolution, while it have a great impact on the *agent\_b* player, not only affecting the rate of strategy evolution, but also causing a reversal of the results of strategy evolution.



**Figure 10:**  $V_r^{Miss}$  impact on the evolution of behavioral decisions on game players (a)  $V_r^{Miss}$  impact on the evolution of behavioral decision-making in *agent\_r* (b)  $V_r^{Miss}$  impact on the evolution of behavioral decision-making in *agent\_b*

Consider that the reason for this phenomenon is due to the fact that when the remaining time to evade the interceptor missiles is shorter, it means that the *agent\_r* player is facing an immediate threat to its survival, and the time urgency is higher, and this urgency prompts the *agent\_r* player to adopt a more risky offensive strategy, striving to quickly break the deadlock or at least distracting the *agent\_b* player's attention in order to strive for a chance of escape, and this may lead to the fact that both the *agent\_r* and the *agent\_b* players are inclined to adopt an offensive tactical strategy in order to quickly resolve battles or gain an advantage. Moreover, the game players need to make quick decisions in a short time window, and offense is often the preferred strategy when time is of the essence. As the remaining time increases, both game players may have more time to assess the



battlefield situation and formulate their strategies, which may lead to a tendency for both players to adopt a more cautious defensive tactical strategy, with the longer period of time permitting more opportunities for both players to make tactical adjustments and reassess risks. In the face of uncertain changes in environmental factors, the *agent\_r* player's strategy evolution is less affected, which may be because the *agent\_r* has already formulated corresponding strategies based on the remaining time and these strategies are less sensitive to environmental changes. While the *agent\_b* player may show higher sensitivity to environmental uncertainty and need to continuously adjust its strategy to adapt to the changes in the *agent\_r* player's behavior, it is difficult to accurately judge the *agent\_r* player's true intentions and capabilities, which increases the complexity of strategy evolution, leading to excessive or insufficient adjustment of the strategies, resulting in fluctuations in the rate of strategy evolution and reversal of results.

### 5.3 Discussion of Results

It is clear from the previous analysis that the evolution of the system will be subject to the interference of random factors in the external environment, and the greater the intensity of random environmental interference, the greater the fluctuation of the decision-making of the game players in the process of strategy evolution, and the more difficult it is for both game players to predict the changes in the environment accurately. The introduction of random disturbances increases the uncertainty of the system so that the decision-making process of the game players is no longer entirely based on rational analysis and expected returns but is affected by more uncontrollable factors. This randomness leads to the diversity and unpredictability of the system's evolutionary path.

In addition, factors such as target azimuth, target entry angle, the inherent value of the game players, the basic probability of fire-strike process, the remaining time to evade missiles, and the missile consumption costs all have different impacts on the rate of evolution and the results of the game. Specifically, factors such as target azimuth, target entry angle, and the inherent value of the game players themselves have a relatively minor impact on the strategic evolution outcomes of both parties involved in the game, but they significantly influence the rate of evolution. Conversely, factors such as the basic probability of fire-strike process, the remaining time to evade missile, and the missile consumption cost have a relatively minor impact on the strategy evolution outcomes of the parameter owner itself, but they significantly affect the other party involved in the game. Therefore, in the actual operational process, decision-makers need to combine the specific environment, strategic objectives, tactical considerations, the psychology and behavior of the two game players and other factors to conduct a comprehensive analysis. They must understand the effect of various parameters in regulating different aspects of the decision-making process of the game players, and take into account the changes caused by a multitude of factors on the evolution of decision-making on both game players. Especially in a highly uncertain environment, even a small information asymmetry or miscalculation may have a significant impact on decision-making.

The method proposed in this paper is compared with other literatures, and the results are shown in [Table 2](#). In the table, the method type refers to the theoretical approach adopted in the corresponding literature. Behavioral rationality indicates whether the literature considers the agents involved in adversarial decision-making as fully rational or boundedly rational, which affects the model's alignment with real-world problems and, consequently, the algorithm's practicality and generalizability. Information requirement denotes whether the adversarial decision-making process considers if agents can obtain complete environmental situational information. If agents can access all information in the adversarial environment, it is considered a complete information condition,

otherwise, it is an incomplete information condition. The game process refers to the phased and state-based characteristics exhibited in the air attack-defense confrontation process, indicating whether the game process is dynamic or static, with dynamic processes featuring rapid state changes and more complex solution processes. Equilibrium solution indicates whether the algorithm's equilibrium solving process is detailed and specific, as incomplete information dynamic adversarial games are more challenging to solve for equilibrium compared to complete information adversarial games and static adversarial games. Therefore, the level of detail in equilibrium solving directly affects the practicality and interpretability of the algorithm. Random environmental interference refers to whether the model considers the impact of random environmental factors on agent decision-making, which determines if the model can more accurately describe interference factors present in real-life scenarios.

**Table 2:** Comparative analysis of different models

Reference	Method type	Behavioral rationality	Information requirement	Game process	Equilibrium solution	Random interference
[6]	Heuristic algorithm	Bounded	Incomplete	Dynamic/static	Simple	No
[10]	Differential game	Perfect	Complete	Dynamic	Detailed	No
[11]	Matrix game	Perfect	Complete	Static	Simple	No
[17]	Deep reinforcement learning	Bounded	Incomplete	Dynamic	Simple	Yes
[20]	Evolutionary game	Bounded	Incomplete	Dynamic	Detailed	No
This paper	Stochastic evolutionary game	Bounded	Incomplete	Dynamic	Detailed	Yes

## 6 Conclusion

Aiming at the maneuvering decision-making problem of agents in the interaction process of multi-agents in a random interference environment, the paper considers the specific impact of maneuvering actions on the game state, analyzes the possible risks caused by random interference in the complex environment on the maneuver decision-making of the game players, establishes a stochastic evolutionary mathematical model of multi-agent game behavior strategy, and researches the strategy selection interaction relationship, evolutionary stability, and dynamic decision-making process of each game player in a random interference environment. The results show that the parameters related to the maneuvering actions of the agent have different effects on the selection and evolution speed of the agent's strategy, and the influence of environmental interference on the different parameters of the maneuver is also different. This also confirms the feasibility and effectiveness of the proposed method, which can better explain the behavior decision-making process of the agent in the interaction process.

This study quantifies the decision-making processes of agents through mathematical methods, providing a foundational model and analytical framework for subsequent research. It not only aids decision-makers in understanding the dynamics of strategies under various factors and environmental conditions but also offers feasible analytical approaches and theoretical references for improving

decision-making in multi-agent interactions and enhancing the interpretability of game system models. Compared to other methods such as differential games, matrix games, deep reinforcement learning, and genetic algorithms, the proposed method in this paper takes into account the agents' bounded rationality, dynamic learning processes, and the impact of random disturbances, making it more aligned with real-world scenarios. It can simulate the strategy selection process of participants over time, reflect the dynamic changes in agents' maneuvering behaviors during aerial combat, and describe the various tactics in air warfare. This is beneficial for predicting the evolution of tactics and identifying optimal strategy combinations, thereby assisting decision-makers in better understanding the potential behaviors and responses of both friendly and adversary forces, ultimately enhancing the scientific rigor of air combat training and strategy development.

However, the paper primarily analyzes the impact of maneuver actions in environments with stochastic disturbances on the behavioral strategies of agents. It offers a solution that is interpretable for specific problems, but its universality is somewhat limited. We are committed to expanding our research scope in future studies. In subsequent research, we will conduct more refined experimental designs and employ more complex mathematical models to further analyze the specific impacts of different factors in various contexts. Additionally, we aim to provide more detailed mechanistic explanations to uncover the specific mechanisms and principles underlying these complex relationships, thereby enhancing the specificity and operability of the research findings. We also consider combining the self-learning capability of deep reinforcement learning and the strategy analysis capability of game theory to provide a more comprehensive solution to the air combat problem.

**Acknowledgement:** We would like to thank everyone who provided comments on this work.

**Funding Statement:** The authors received no specific funding for this study.

**Author Contributions:** Study conception and design: Shiguang Hu, Le Ru, Bo Lu, Zhenhua Wang; data collection: Shiguang Hu, Xiaolin Zhao, Wenfei Wang, Hailong Xi; analysis and interpretation of results: Shiguang Hu, Le Ru, Xiaolin Zhao, Bo Lu; draft manuscript preparation: Shiguang Hu, Le Ru. All authors reviewed the results and approved the final version of the manuscript.

**Availability of Data and Materials:** The authors confirm that the data supporting the findings of this study are available within the article or its supplementary materials.

**Ethics Approval:** Not applicable.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

- [1] F. Jiang, M. Xu, Y. Li, H. Cui, and R. Wang, "Short-range air combat maneuver decision of UAV swarm based on multi-agent Transformer introducing virtual objects," *Eng. Appl. Artif. Intell.*, vol. 123, 2023, Art. no. 106358. doi: [10.1016/j.engappai.2023.106358](https://doi.org/10.1016/j.engappai.2023.106358).
- [2] Z. Ren, D. Zhang, S. Tang, W. Xiong, and S. Yang, "Cooperative maneuver decision making for multi-UAV air combat based on incomplete information dynamic game," *Def. Technol.*, vol. 27, no. 3, pp. 308–317, 2023. doi: [10.1016/j.dt.2022.10.008](https://doi.org/10.1016/j.dt.2022.10.008).
- [3] J. Xu, Z. Deng, and Q. Song, "Multi-UAV counter-game model based on uncertain information," *Appl. Math. Comput.*, vol. 366, no. 3, 2020, Art. no. 124684. doi: [10.1016/j.amc.2019.124684](https://doi.org/10.1016/j.amc.2019.124684).

- [4] S. Li, M. Chen, Y. Wang, and Q. Wu, "Air combat decision-making of multiple UCAVs based on constraint strategy games," *Def. Technol.*, vol. 18, no. 3, pp. 368–383, 2022. doi: [10.1016/j.dt.2021.01.005](https://doi.org/10.1016/j.dt.2021.01.005).
- [5] Z. Chen *et al.*, "Tripartite evolutionary game in the process of network attack and defense," *Telecommun. Syst.*, vol. 86, no. 2, pp. 351–361, May 2024. doi: [10.1007/s11235-024-01130-9](https://doi.org/10.1007/s11235-024-01130-9).
- [6] Y. Huang, A. Luo, T. Chen, M. Zhang, B. Ren and Y. Song, "When architecture meets RL plus EA: A hybrid intelligent optimization approach for selecting combat system-of-systems architecture," *Adv. Eng. Inform.*, vol. 58, Oct. 2023, Art. no. 102209. doi: [10.1016/j.aei.2023.102209](https://doi.org/10.1016/j.aei.2023.102209).
- [7] Y. Wang, X. Zhang, R. Zhou, S. Tang, H. Zhou and W. Ding, "Research on UCAV maneuvering decision method based on heuristic reinforcement learning," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–13, Mar. 2022. doi: [10.1155/2022/1477078](https://doi.org/10.1155/2022/1477078).
- [8] R. Isaacs, "A mathematical theory with applications to warfare and pursuit, control and optimization," in *Differential Games*. New York, NY, USA: Wiley, 1965.
- [9] E. Garcia, D. Casbeer, and M. Pachter, "Active target defence differential game: Fast defender case," *IET Control Theory Appl.*, vol. 11, no. 17, pp. 2985–2993, Sep. 2017. doi: [10.1049/iet-cta.2017.0302](https://doi.org/10.1049/iet-cta.2017.0302).
- [10] H. Park, B. Lee, and D. Yoo, "Differential game based air combat maneuver generation using scoring function matrix," *Int. J. Aeronaut. Space*, vol. 17, no. 2, pp. 204–213, Jun. 2016. doi: [10.5139/IJASS.2016.17.2.204](https://doi.org/10.5139/IJASS.2016.17.2.204).
- [11] F. Austin, G. Carbone, H. Hinz, M. Lewis, and M. Falco, "Game theory for automated maneuvering during air-to-air combat," *J. Guid. Control Dyn.*, vol. 13, no. 6, pp. 1143–1149, Dec. 1990. doi: [10.2514/3.20590](https://doi.org/10.2514/3.20590).
- [12] S. Li, M. Chen, and Q. Wu, "A fast algorithm to solve large-scale matrix games based on dimensionality reduction and its application in multiple unmanned combat air vehicles attack-defense decision-making," *Inform. Sci.*, vol. 594, no. 4, pp. 305–321, 2022. doi: [10.1016/j.ins.2022.02.025](https://doi.org/10.1016/j.ins.2022.02.025).
- [13] Y. Cao, Y. Kou, and A. Xu, "Autonomous maneuver decision of UCAV air combat based on double deep Q network algorithm and stochastic game theory," *Int. J. Aerosp. Eng.*, vol. 2023, pp. 1–20, Jan. 2023. doi: [10.1155/2023/3657814](https://doi.org/10.1155/2023/3657814).
- [14] J. Liles, M. Robbins, and B. Lunday, "Improving defensive air battle management by solving a stochastic dynamic assignment problem via approximate dynamic programming," *Eur J. Oper. Res.*, vol. 305, no. 3, pp. 1435–1449, 2023. doi: [10.1016/j.ejor.2022.06.031](https://doi.org/10.1016/j.ejor.2022.06.031).
- [15] N. Léchevin and C. Rabbath, "A hierarchical decision and information system for multi-aircraft combat missions," *Proc. Inst. Mech. Eng. Part G J. Aerosp.*, vol. 224, no. 2, pp. 133–148, Feb. 2010. doi: [10.1243/09544100JAERO563](https://doi.org/10.1243/09544100JAERO563).
- [16] X. Zhang, G. Liu, C. Yang, and J. Wu, "Research on air confrontation maneuver decision-making method based on reinforcement learning," *Electron.*, vol. 7, no. 11, Nov. 2018, Art. no. 279. doi: [10.3390/electronics7110279](https://doi.org/10.3390/electronics7110279).
- [17] S. Li, Q. Wu, and M. Chen, "Autonomous maneuver decision-making of UCAV with incomplete information in human-computer gaming," *Drones*, vol. 7, no. 3, Feb. 2023, Art. no. 157. doi: [10.3390/drones7030157](https://doi.org/10.3390/drones7030157).
- [18] J. Poropudas and K. Virtanen, "Game-theoretic validation and analysis of air combat simulation models," *IEEE Trans. Syst. Man Cyber.-Part A: Syst. Hum.*, vol. 40, no. 5, pp. 1057–1070, Sep. 2010. doi: [10.1109/TSMCA.2010.2044997](https://doi.org/10.1109/TSMCA.2010.2044997).
- [19] Z. Sun *et al.*, "Multi-agent hierarchical policy gradient for Air Combat Tactics emergence via self-play," *Eng. Appl. Artif. Intell.*, vol. 98, 2021, Art. no. 104112. doi: [10.1016/j.engappai.2020.104112](https://doi.org/10.1016/j.engappai.2020.104112).
- [20] S. Hu, L. Ru, and W. Wang, "Evolutionary game analysis of behaviour strategy for UAV swarm in communication-constrained environments," *IET Control Theory Appl.*, vol. 18, no. 3, pp. 350–363, Nov. 2023. doi: [10.1049/cth2.12602](https://doi.org/10.1049/cth2.12602).
- [21] Y. Gao, L. Zhang, and Q. Wang, "An evolutionary game-theoretic approach to unmanned aerial vehicle network target assignment in three-dimensional scenarios," *Mathematics*, vol. 11, no. 19, Oct. 2023, Art. no. 4196. doi: [10.3390/math11194196](https://doi.org/10.3390/math11194196).

- [22] L. Sheng, M. Shi, and Y. Qi, "Dynamic offense and defense of uav swarm based on situation evolution game," (in Chinese), *Syst. Eng. Electron.*, vol. 8, pp. 2332–2342, Jun. 2023. doi: [10.12305/j.issn.1001-506X.2023.08.06](https://doi.org/10.12305/j.issn.1001-506X.2023.08.06).
- [23] C. Liu, S. Sun, and B. Xu, "Optimizing evasive maneuvering of planes using a flight quality driven model," *Sci. China Inform. Sci.*, vol. 67, no. 3, Feb. 2024. doi: [10.1007/s11432-023-3848-6](https://doi.org/10.1007/s11432-023-3848-6).
- [24] H. Huang, H. Cui, H. Zhou, and J. Li, "Research on autonomous maneuvering decision of UAV in close air combat," in *Proc. 3rd 2023 Int. Conf. Autonom. Unmanned Syst.*, Singapore, Springer, 2023, vol. 1171. doi: [10.1007/978-981-97-1083-6\\_14](https://doi.org/10.1007/978-981-97-1083-6_14).
- [25] Z. Li, J. Zhu, M. Kuang, J. Zhang, and J. Ren, "Hierarchical decision algorithm for air combat with hybrid action based on reinforcement learning," (in Chinese), *Acta Aeronaut. et Astronaut. Sinica*, Apr. 2024. doi: [10.7527/S1000-6893.2024.30053](https://doi.org/10.7527/S1000-6893.2024.30053).
- [26] D. Lee, C. Kim, and S. Lee, "Implementation of tactical maneuvers with maneuver libraries," *Chin. J. Aeronaut.*, vol. 33, no. 1, pp. 255–270, 2020. doi: [10.1016/j.cja.2019.07.007](https://doi.org/10.1016/j.cja.2019.07.007).
- [27] H. Duan, Y. Lei, J. Xia, Y. Deng, and Y. Shi, "Autonomous maneuver decision for unmanned aerial vehicle via improved pigeon-inspired optimization," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 3, pp. 3156–3170, Jun. 2023. doi: [10.1109/TAES.2022.3221691](https://doi.org/10.1109/TAES.2022.3221691).
- [28] J. Zhu, M. Kuang, and X. Han, "Mastering air combat game with deep reinforcement learning," *Def. Technol.*, vol. 34, no. 4, pp. 295–312, 2024. doi: [10.1016/j.dt.2023.08.019](https://doi.org/10.1016/j.dt.2023.08.019).
- [29] W. Kong, D. Zhou, and Y. Zhao, "Hierarchical reinforcement learning from competitive self-play for dual-aircraft formation air combat," *J. Comput. Des. Eng.*, vol. 10, no. 2, pp. 830–859, Mar. 2023. doi: [10.1093/jcde/qwad020](https://doi.org/10.1093/jcde/qwad020).
- [30] S. Deng and Y. Yuan, "A data intrusion tolerance model based on an improved evolutionary game theory for the energy internet," *Comput. Mater. Contin.*, vol. 79, no. 3, pp. 3679–3697, Jan. 2024. doi: [10.32604/cmc.2024.052008](https://doi.org/10.32604/cmc.2024.052008).
- [31] X. Richter and J. Lehtonen, "Half a century of evolutionary games: A synthesis of theory, application and future directions," *Phil. Trans. R. Soc. B*, vol. 378, no. 1876, May 2023, Art. no. 489. doi: [10.1098/rstb.2021.0492](https://doi.org/10.1098/rstb.2021.0492).
- [32] E. L. N. Ekassi, O. Foupouapouognigni, and M. S. Siewe, "Nonlinear dynamics and Gaussian white noise excitation effects in a model of flow-induced oscillations of circular cylinder," *Chaos Solitons Fract.*, vol. 178, 2024, Art. no. 114374. doi: [10.1016/j.chaos.2023.114374](https://doi.org/10.1016/j.chaos.2023.114374).
- [33] A. V. Balakrishnan and R. R. Mazumdar, "On powers of Gaussian white noise," *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7629–7634, 2011. doi: [10.1109/TIT.2011.2158062](https://doi.org/10.1109/TIT.2011.2158062).
- [34] D. Cheng, F. He, H. Qi, and T. Xu, "Modeling, analysis and control of networked evolutionary games," *IEEE Trans. Automat. Control*, vol. 60, no. 9, pp. 2402–2415, Sep. 2015. doi: [10.1109/TAC.2015.2404471](https://doi.org/10.1109/TAC.2015.2404471).
- [35] G. Jumarie, "Approximate solution for some stochastic differential equations involving both Gaussian and Poissonian white noises," *Appl. Math. Lett.*, vol. 16, no. 8, pp. 1171–1177, 2003. doi: [10.1016/S0893-9659\(03\)90113-3](https://doi.org/10.1016/S0893-9659(03)90113-3).
- [36] H. Zhang, Y. Mi, X. Liu, Y. Zhang, J. Wang and J. Tan, "A differential game approach for real-time security defense decision in scale-free networks," *Comput. Netw.*, vol. 224, 2023, Art. no. 109635. doi: [10.1016/j.comnet.2023.109635](https://doi.org/10.1016/j.comnet.2023.109635).
- [37] J. Armstrong and A. Ionescu, "Itô stochastic differentials," *Stoch. Proc. Appl.*, vol. 171, Apr. 2024, Art. no. 104317. doi: [10.1016/j.spa.2024.104317](https://doi.org/10.1016/j.spa.2024.104317).
- [38] B. Moghaddam, A. Lopes, J. Machado, and Z. Mostaghim, "Computational scheme for solving nonlinear fractional stochastic differential equations with delay," *Stoch. Anal. Appl.*, vol. 37, no. 6, pp. 893–908, Nov. 2019. doi: [10.1080/07362994.2019.1621182](https://doi.org/10.1080/07362994.2019.1621182).
- [39] K. Kang, L. Bai, and J. Zhang, "A tripartite stochastic evolutionary game model of complex technological products in a transnational supply chain," *Comput. Ind. Eng.*, vol. 186, Dec. 2023, Art. no. 109690. doi: [10.1016/j.cie.2023.109690](https://doi.org/10.1016/j.cie.2023.109690).



- [40] J. Yang, X. Liu, and X. Liu, “Stability of stochastic functional differential systems with semi-Markovian switching and Levy noise by functional Itô’s formula and its applications,” *J. Franklin Inst.*, vol. 357, no. 7, pp. 4458–4485, May 2020. doi: [10.1016/j.jfranklin.2020.03.012](https://doi.org/10.1016/j.jfranklin.2020.03.012).
- [41] Y. Xu, B. Hu, and R. Qian, “Analysis on stability of strategic alliances based on stochastic evolutionary game including simulations,” (in Chinese), *Syst. Eng. Theory Pract.*, vol. 31, pp. 920–926, May 15, 2011. doi: [10.12011/1000-6788\(2011\)5-920](https://doi.org/10.12011/1000-6788(2011)5-920).