A Nonlinear Finite-Time Robust Differential Game Guidance Law

Axing Xi and Yuanli Cai

School of Automation Science and Engineering, Xi’an Jiaotong University, Xi’an 710049, China
* Correspondence: ylicai@mail.xjtu.edu.cn

Abstract: In this paper, a robust differential game guidance law is proposed for the nonlinear zero-sum system with unknown dynamics and external disturbances. First, the continuous-time nonlinear zero-sum differential game problem is transformed into solving the nonlinear Hamilton–Jacobi–Isaacs equation, a time-varying cost function is developed to reflect the fixed terminal time, and the robust guidance law is developed to compensate for the external disturbance. Then, a novel neural network identifier is designed to approximate the unknown nonlinear dynamics with online weight tuning. Subsequently, an online critic neural network approximator is presented to estimate the cost function, and time-varying activation functions are considered to deal with the fixed final time problem. An adaptive weight tuning law is given, where two additional terms are added to ensure the stability of the closed-loop nonlinear system and so as to meet the terminal cost at a fixed final time. Furthermore, the uniform ultimate boundedness of the closed-loop system and the critic neural network weights estimation error are proven based upon the Lyapunov approach. Finally, some simulation results are presented to demonstrate the effectiveness of the proposed robust differential game guidance law for nonlinear interception.

Keywords: guidance law; neuro-dynamic programming; differential game; nonlinear system

1. Introduction

In the modern battlefield, when the missile executes the interception mission, the fuel consumption, rapid attitude maneuver, and other uncertain factors may lead to the nonlinear dynamic characteristics of the system. Moreover, external disturbances, which may cause the system to lose control, are important factors in practice applications. Furthermore, the missile intercepts maneuvering targets with a fixed terminal time, which is a more effective attack. Therefore, it is of great significance to study the finite-time robust differential game law in nonlinear circumstances with external disturbance.

Traditional interception guidance strategies include proportional navigation guidance (PNG) [1] and augmented proportional navigation guidance (APNG) [2,3]. In these works, the interception problem was simplified into a two-dimensional plane and the target aircrafts were assumed to be non-maneuvering targets. With the development of the modern control theory, optimal guidance laws (OGLs), sliding mode control (SMC) guidance laws, and linear quadratic differential game (LQDG) guidance laws are investigated to solve the missile interception problem. In [4], first-order interceptor dynamics were taken into account, and an optimal guidance law (OGL) was proposed to control the impact time and impact angle. In [5], a combination of a line-of-sight (LOS) rate shaping technique and a second-order SMC was proposed. However, the target position was assumed to be known in advance. In [6], three-party pursuit and evasion guidance strategies were derived from a linear dynamic system, analytical solutions were derived via the Game Algebraic Riccati equation (GARE) and LQDG guidance laws were proposed. However, it is almost impossible, or even impossible, to acquire the analytic solution, due to the inherent nonlinearity. To circumvent the inherent nonlinearity of the system, in [7], an
SDRE guidance law was proposed by solving the nonlinear Hamilton–Jacobi–Isaacs (HJI) equation. However, the external disturbance was neglected and it was difficult for the analytical solution to find its Nash equilibrium. Recently, an artificial neural network [8] has been applied to solve the nonlinear HJI equation for an unknown system. However, the finite-horizon optimal guidance law still remains an unsolved problem.

To solve the GARE, adaptive dynamic programming (ADP) techniques serve as powerful tools to solve the optimal control problems. In [9], a novel online mode-free reinforcement learning algorithm was proposed to solve the multiplayer non-zero-sum games, and the algebraic Riccati equation (ARE) was solved by an iterative algorithm. However, external disturbances were not considered. Unfortunately, in practice applications, external disturbances always exist. In [10], the author proposed a data-driven value iteration (VI) algorithm to handle the adaptive CT linear optimal output regulation. The optimal feedback control gain was learned by an online value iteration algorithm. However, the iterative algorithm requires a significant number of iterations within a fixed sampling interval to guarantee the stability of the system. In the above references, no matter what the mode-free integral RL algorithm or the data-driven value iteration algorithm, all the proposed methods were based on a solvable ARE. However, in most practice applications, external disturbances and parameter uncertainties always exist, which the ARE cannot obtain, in addition to the optimal controller.

For other contemporary approaches related to ADP, one can refer to refs. [11,12]. However, these are not applicable to practice applications with unknown or inaccurate system dynamics. The neuro-dynamic programming (NDP) technique is a typical method used to solve the nonlinear Hamilton–Jacobi–Isaacs (HJI) equation. In [13], the policy or value iteration-based NDP scheme was proposed to obtain the finite-horizon ε-optimal control for the discrete-time nonlinear system by using offline neural network (NN) training. In [14], an online algorithm based on policy iteration was proposed to attain the synchronous optimal policy with infinite horizon time for nonlinear systems with known dynamics. However, inadequate iterations within a sampling interval may lead to instability. To avoid this problem, in [15], a time-based NDP method was studied, and iteration-based optimal solutions were replaced by using the previous history of system states and cost function approximations. However, this NDP scheme was not suitable for finite-time nonlinear control; furthermore, external disturbances were not considered and only infinite-time optimal control was studied. In [16], an online concurrent learning NDP algorithm was presented to solve the two-player zero-sum game of nonlinear CT systems with unknown system dynamics and three NN approximators were tuned to learn about the value function corresponding to the optimal control strategies. However, two or more NNs led to an increase in computational complexity.

Although the above research has solved the HJB problem to a certain extent, unfortunately, there are few kinds of literature involved in the missile–target interception problem. In [17], a novel sliding mode adaptive neural network guidance law was proposed to intercept highly maneuvering targets. Aimed at the external disturbance caused by the target maneuvering, the RBF neural network was adopted to eliminate estimation errors without prior information about the target. Similarly, in [18], an adaptive NN-based scheme with an estimation cost function was proposed for solving the interception problem of the spacecraft with limited communication and external disturbances, but it did consider the target maneuverability. However, in these studies, only the missile interception strategy is considered, and the target control strategy is ignored.

In this paper, the nonlinear system of missile–target engagement with external disturbances is considered, and a time-varying cost function is designed for satisfying the terminal interception time. The uncertain nonlinear two-player zero-sum game is developed via the HJI equation based on differential game theory. The main contributions of this paper include two aspects. First, for solving the external disturbances problem, unlike the work [14], an extended robust interception guidance strategy for the missile to intercept the maneuvering target within a fixed final time is proposed. Second, two novel NNs are
designed, one online NN identifier is designed to approximate the unknown nonlinear system, and the other critic NN is developed to approximate the cost function without policy or value iterations, while online learning is adopted. Finally, the nonlinear finite-time robust differential game guidance law is proposed.

The advantages of the proposed method of this paper are listed as follows:

(1) The nonlinear guidance law for missile–target engagement within a fixed interception time is studied by using the differential game theory based on neuro-dynamic programming (NDP). More importantly, a time-varying cost function is reconstructed by a critic NDP with two additional terms added to ensure the stability of the nonlinear system and to meet the fixed interception time, which implies that the missile can intercept the target at different terminal times.

(2) In practical applications, there are always external disturbances, and a robust interception guidance law is proposed to deal with this problem. Furthermore, inspired by the work [19], our proposed method extends the controller by considering the target.

(3) Unlike the discrete system, our proposed method is the CT. Moreover, compared with the existing work [15,16], a clear advantage of our method is that a simpler critic NN structure is designed; thus, the computational burden is alleviated.

The remainder of this paper is organized as follows. The statement of the normal guidance problem is presented in part II. The robust control strategy of the nonlinear system with external disturbances is developed in part III. In part IV, an online NN identifier and a novel NDP-based approximator are presented. The stability of the nonlinear system is proved in part V. The nonlinear model of the missile–target engagement is established and numerical experiments are carried out to evaluate the performance of the proposed robust differential game guidance strategy in part VI. Part VII presents some conclusions.

2. Problem Formulation

In this paper, the two players are the missile and the target, which are described in detail in part VI. For the finite-time nonlinear two-player zero-sum differential game, the object of the missile input \( \pi(t) \) is to minimize the cost function, while the target input \( \pi(t) \) is to maximize the cost function in a specified time. The continuous-time (CT) uncertain nonlinear two-player zero-sum differential game is now presented as

\[
x(t) = f(x(t)) + g(x(t))(\pi(t) + \bar{\pi}(x)) + k(x(t))(\pi(t) + \bar{\pi}(x))
\]  

(1)

where \( x(t) \in R^n \), \( \pi(t) \in R^m \), and \( \bar{\pi}(t) \in R^l \) represent the system state vector, the control input of the missile and the target, respectively. \( f(x) \in R^n \) denotes the internal system dynamics, \( g(x) \in R^{n \times m} \), and \( k(x) \in R^{n \times q} \) are the control coefficient matrices, with \( f(x), g(x) \) and \( k(x) \) are locally Lipschitz, respectively. \( \bar{\pi}(x) \) and \( \bar{\pi}(x) \) represent external disturbances, which are both bounded by known functions \( d_M(x) \) and \( p_M(x) \), i.e.,

\[
\|\bar{\pi}(x)\| \leq d_M(x) \text{ and } \|\bar{\pi}(x)\| \leq p_M(x).
\]

Furthermore, we assume that \( d(x) = R_1^2 \bar{\pi}(x) \) and \( p(x) = \gamma \bar{\pi}(x) \), with \( R_1 \) and \( \gamma \) are symmetric positive definite matrices, respectively.

The nominal system (without external disturbances) of the system (1) can be described as

\[
\dot{x}(t) = f(x(t)) + g(x(t))u(t) + k(x(t))w(t)
\]  

(2)

We assume that \( f + gu + kw \) is Lipschitz continuous on a set \( \Omega \) and the system (2) is controllable.

Considering external disturbances in the nonlinear system (1), for the nominal system (2), the finite-time two-player zero-sum differential game cost function is defined as

\[
V(x, t_0) = \phi(x(t_f), t_f) + \int_{t_0}^{t_f} r(x(t), u(t), w(t))dt
\]

\[
V(x, t_f) = \phi(x(t_f), t_f)
\]  

(3)
where \( r(x(t), u(t), w(t)) = Q(x) + u^TR_1u - \frac{1}{2}w^T\gamma w, \) with \( Q(x) = \delta^2_M(x) + R_2p^2_M(x) + x^TQ_1x; \) \( Q_1 \) is a semi-positive function. The terminal cost, external disturbances, control efforts of the missile and the target, the system state, and fixed terminal time are chosen as the performance evaluation indicators. \( \varphi(x(t), t_f) \) reflects the terminal cost between the missile with the target. The term \( Q(x) \) reflects external disturbances and the system state simultaneously, which is positively defined. Moreover, \( R_2 \) represents the influence of the target disturbance. \( R_1 \) reflects the missile control effort. \( \gamma \) reflects the target control effort. All parameters \( Q(x), R_1, R_2, \) and \( \gamma \) are positively defined. The goal of this paper is to find the saddle point of the cost function (3).

**Remark 1.** First, unlike the infinite-time scenario, the terminal cost \( V(x, t_0) \) is a time-varying function. \( \varphi(x(t), t_f) \) is needed to guarantee the finite-time scenario. Second, external disturbances are considered in the cost function via adopting a positive constant \( R_2 \) and the robust control problem can be addressed by designing the finite-time guidance strategy of the nominal system (2).

Assuming that \( V(x, t) \in C^1 \), an infinitesimial equivalent to (3) can be derived as

\[
- \frac{\partial V(x, t)}{\partial t} = r(x, u, d) + \frac{\partial V^T(x, t)}{\partial x} (f(x) + g(x)u + k(x)w)
\]

when \( t_0 = t_f \), the terminal cost function can be expressed as

\[
V(x, t_f) = \Psi(x(t_f), t_f)
\]

The Hamiltonian function of the nonlinear system (2) can be defined as

\[
H(x, u, v) = V_1 + r(x, u, v) + V^T_x(f(x) + g(x)u(t) + k(x)w(t))
\]

where \( V_1 = \frac{\partial V(x, t)}{\partial t} \) and \( V_x = \frac{\partial V(x, t)}{\partial x} \). It can be clearly observed that the Hamiltonian function includes a time-dependent term \( V_1 \).

In the Nash equilibrium theory, the saddle point with respect to the optimal control pair \((u(t), w(t))\) can be obtained by

\[
H(x, u^*, w) \leq H(x, u, w^*) \leq H(x, u, w^*)
\]

If one recalls the classical optimal game theory, both optimal controllers can be solved by using stationary conditions \( \frac{\partial H(u)}{\partial u} = 0 \) and \( \frac{\partial H(w)}{\partial w} = 0 \), which yields

\[
\begin{align*}
 u^*(x) &= -\frac{1}{2}R_1^{-1}g^T(x)V^*_x \\
 w^*(x) &= \gamma^2k(x)V^*_x
\end{align*}
\]

where \( V^*(x, t) \) is the optimal two-player zero-sum game cost function, which is the saddle point of the cost function, such that

\[
V^*(x, t) = \max_{u(t)} V(x, t)_{\text{opt}}
\]

\[
V^*(x, t_f) = \varphi(x(t_f), t_f)
\]

By substituting optimal strategy (8) into Equation (4), the Hamilton–Jacobi–Isaacs (HJI) equation reduces to

\[
V^*_f + Q(x) + V^*_x f(x) - \frac{1}{2}V^*_x g(x)R_1^{-1}g^T(x)V^*_x
+ \frac{\gamma^2}{2}V^*_x k(x)k^T(x)V^*_x = 0
\]
Remark 2. For the linear system case, the HJB equation can be easily solved by the Riccati equation [19]. However, it is difficult or even impossible to attain the mathematical solution of the HJI Equation (10), when system dynamics exist in nonlinear terms. Moreover, the fixed final time $t_f$ is provided in this nonlinear system for solving the inadequate iterations problem, a novel time-based online optimal guidance law design is proposed and the system dynamics are demonstrated.

3. The Nonlinear Finite-Time Robust Differential Game Guidance Law

In this part, the nonlinear finite-time robust differential game guidance law is presented. First, for coping with the robust stabilization problem of the system (1) with external disturbances, a robust controller is designed. Then, the finite-time NDP-based optimal guidance strategy is designed.

3.1. Robust Controller Design of Uncertain Nonlinear Differential Games

By extending the work [20], two feedback gains $\pi_1$ and $\frac{1}{\gamma_2}$ are added to the optimal feedback control (8) of the system (2) for the missile and the target, respectively. The robust optimal feedback control yields as follows:

$$
\begin{align*}
\pi^*(x) &= \pi_1 u^*(x) = -\frac{1}{2} \pi_1 R_1^{-1} g^T(x) V_x^* \\
\varpi^*(x) &= \frac{1}{2\gamma_2} \varpi^*(x) = \frac{1}{2\gamma_2} \gamma^2 k^T(x) V_x^*
\end{align*}
$$

(11)

Here, some lemmas are presented for indicating that the robust optimal control has an infinite gain margin.

Lemma 1. For the nominal system (2), the optimal control strategy given by (11) can ensure that the closed-loop system is asymptotically stable for $\pi_1 \geq \frac{1}{2}$ and $\pi_2 \geq 1$.

Proof of Lemma 1. The optimal cost function $V^*(x,t) = J^*(x,t)$ is selected as the Lyapunov function. In light of (3), it is easily found that $V^*(x,t)$ is positive definite. By combining (10) and (11), the derivative of the $V^*(x,t)$ along the trajectory of the closed-loop system yields

$$
\dot{V}^*(x,t) = V_t^* + V_x^* T(f(x) + g(x) \pi(t) + k(x) \varpi(t))
$$

$$
= V_t^* - Q(x) - V_t^* + \frac{1}{4} V_x^* T g(x) R_1^{-1} g^T(x) V_x^*
$$

$$
- \frac{2}{\gamma_2} V_x^* T k(x) k^T(x) V_x^* - \frac{1}{2} \pi_1 g(x) R_1^{-1} g^T(x) V_x^* + \frac{1}{2\gamma_2} \gamma^2 k^T(x) k(x) V_x^*
$$

(12)

Hence, whenever $\pi_1 \geq \frac{1}{2}, \pi_2 \geq 1$ and $x \neq 0$, $\dot{V}^*(x,t) < 0$. □

Theorem 1. For the system (1), there exists two positive gains, $\pi_1^* \geq 1$ and $\pi_2^* \geq \frac{2R_2}{2R_1+1}$, with $R_2 > 1$, such that for any $\pi_1 > \pi_1^*$ and $\pi_2 > \pi_2^*$, the robust control (11) ensures that the closed-loop system (1) is asymptotically stable.

Proof of Theorem 1. The optimal cost function $L(x,t) = V^*(x,t)$ is selected as the Lyapunov function, and the derivative of the $V^*(x,t)$ along the trajectory of the closed-loop system can be obtained as

$$
L(x,t) = V_t^* + V_x^* T( f(x) + g(x)(\pi(t) + \varpi(t)) + k(x)(\varpi(t) + \varpi(t)))
$$

(13)

Based on (12), (13) can be rewritten as
\[ \dot{x}(x, t) = -Q(x) - \frac{1}{2} \left( \pi_1 - \frac{1}{2} \right) \| V_x^T g(x) R_1^{-\frac{1}{2}} \|^2 \leq -\frac{1}{2} \left( \pi_1 - \frac{1}{2} \right) \left( \| V_x^T g(x) \gamma \|^2 + V_x^T g(x) d(x) + V_x^T k(x) p(x) \right) \]

\[ \leq -x^T Q_1 x - \left( \frac{1}{2} \left( \pi_1 - \frac{1}{2} \right) \| V_x^T g(x) R_1^{-\frac{1}{2}} \|^2 + \frac{1}{2} \left( \pi_1 - \frac{1}{2} \right) \| V_x^T g(x) \gamma \|^2 - \| V_x^T g(x) R_1^{-\frac{1}{2}} \| d_M(x) \right) \]

\[ -\| V_x^T k(x) \gamma \| d_M(x) + d_M(x) + R_2 d_M(x) \right) \]

If \( z = \left[ d_M(x), p_M(x), \| V_x^T g(x) R_1^{-\frac{1}{2}} \|, \| V_x^T k(x) \gamma \| \right] \); thus, (14) can be rewritten as

\[ \dot{L}(x, t) \leq -x^T Q_1 x - z^T \Theta z \]

where \( \Theta = \begin{bmatrix} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & R_2 & 0 & -\frac{1}{2} \\ -\frac{1}{2} & 0 & 1 & 0 \\ 0 & -\frac{1}{2} & 0 & 1 - \frac{1}{\pi_1} \end{bmatrix} \).

According to the Lyapunov function, the determinant \( \Theta \geq 0 \) represents the \( \dot{L}(x, t) \leq 0 \) and implies that the closed-loop system is asymptotically stable. Thus, it can be concluded that \( \pi_1^* \geq 1 \) and \( \pi_2^* \geq \frac{2R_2}{2R_2 + 1} \) can ensure the positive definiteness of \( \Theta \). When \( \pi_1 > \pi_1^* \) and \( \pi_2 > \pi_2^* \), the closed-loop system is asymptotically stable. \( \square \)

3.2. Finite-Time NDP-Based Optimal Guidance Strategy

First, a novel online NN identifier is proposed to approximate the unknown system dynamics. Next, a critical NDP-based approximator is utilized to estimate the cost function within a fixed final time and an online adaptive weight tuning law is proposed with additional terms to guarantee the stability of the nonlinear system. Finally, combining the identified system and estimated cost function, the finite-horizon optimal differential guidance strategy is derived.

3.2.1. NN Identifier

System dynamics are necessary for developing guidance laws of the nonlinear two-player zero-sum differential games. However, system dynamics may be unknown in practice applications. To overcome this problem, a novel online NN identifier is designed. Based on the NN universal function approximation property, the nonlinear system can be represented as

\[ f(x) = W_f^T \sigma_f(x) + \epsilon_f \]
\[ g(x) = W_g^T \sigma_g(x) + \epsilon_g \]
\[ k(x) = W_k^T \sigma_k(x) + \epsilon_k \]

where \( W_f \in R^{f \times N}, W_g \in R^{g \times N}, \) and \( W_k \in R^{k \times N} \) are ideal weight matrices. \( \sigma_f(x) \in R^N, \sigma_g(x) \in R^N, \) and \( \sigma_k(x) \in R^N \) denote NN activation function vectors, \( N \) is the number of hidden layer neurons, and \( \epsilon_f, \epsilon_g, \) and \( \epsilon_k \) represent NN approximation errors.

Then, the nominal system (2) can be represented by using (16) as

\[ \dot{x} = f(x) + g(x) u + k(x) w = \begin{bmatrix} W_f^T & W_g^T & W_k^T \end{bmatrix} \begin{bmatrix} \sigma_f(x) & 0 & 0 \\ 0 & \sigma_g(x) & 0 \\ 0 & 0 & \sigma_k(x) \end{bmatrix} \begin{bmatrix} 1 \\ u \\ w \end{bmatrix} + \epsilon_f + \epsilon_g u + \epsilon_k w = W^T \sigma(x) + \epsilon \]

Because the ideal NN weights are typically unknown, we define the state estimator as follows:

\[ \hat{x} = W^T \sigma(x) + K \hat{x} \]
where \( \hat{W}_I \) represents the estimate of the \( W_I \). \( \hat{x} = x - \hat{x} \in \mathbb{R}^N \) denotes the state estimation error. \( K \) is a design parameter, which can guarantee the stability of the NN identifier.

From (17) and (18), the derivative of the state estimation error yields

\[
\dot{\tilde{x}} = x - \hat{x} = \tilde{W}_I^T \sigma_{I(x)} \tilde{x} + \epsilon_I - K \tilde{x} \tag{19}
\]

Inspired by [9], in order to make the approximated NN identifier weight matrix close to its ideal value, the online tuning law is given by

\[
\hat{W}_I(t) = -\alpha_1 \hat{W}_I(t) + \sigma_{I(x)} \tilde{\epsilon}_I \tag{20}
\]

where \( \alpha_1 \) is the learning rate of the critical NN.

Next, by defining \( \tilde{W}_I = W_I - \hat{W}_I \), the identifier weight estimation error yields the following equation by using (20):

\[
\tilde{W}_I(t) = \alpha_1 \hat{W}_I(t) - \sigma_{I(x)} \tilde{\epsilon}_I \tag{21}
\]

**Theorem 2.** For the NN identifier (18), let the initial ideal critic NN weight \( W_I \) reside in a compact set by selecting the proposed NN weight tuning law provided (20). There exists a positive tuning parameter \( (\alpha_1 > 0) \). Then, the identification (19) and the weight estimation errors \( \tilde{W}_I(t) \) are uniformly ultimately bounded (UUB) with a fixed terminal time.

**Proof of Theorem 2.** Let the Lyapunov function candidate be as follows:

\[
J_I = \frac{1}{2} \tilde{x}^T \tilde{x} + \frac{1}{2} \text{tr} \left( \tilde{W}_I^T \tilde{W}_I \right) \tag{22}
\]

Then,

\[
\dot{J}_I = \tilde{x}^T \tilde{\dot{x}} + \text{tr} \left( \tilde{W}_I^T \tilde{W}_I \right) \leq \frac{1}{2} \| \tilde{x} \|^2 + \frac{1}{2} \| \tilde{\epsilon}_I \|^2 - \lambda_{\text{min}}(K) \| \tilde{x} \|^2 + \alpha_1 \text{tr} \left( \tilde{W}_I^T \tilde{W}_I \right) - \alpha_1 \text{tr} \left( \tilde{W}_I^T \tilde{W}_I \right)
\]

\[
\leq -\left( \lambda_{\text{min}}(K) - \frac{1}{2} \right) \| \tilde{x} \|^2 - \frac{\alpha_1}{2} \| \tilde{W}_I \|^2 + \epsilon_{IM} \tag{23}
\]

where \( \epsilon_{IM} = \alpha_1 \| W_I \|^2 + \frac{1}{2} \| \epsilon_I \|^2 \), the eigenvalue \( K \) and \( \alpha_1 \) are the design parameters that guarantee the stability of the system. Therefore, when \( \lambda_{\text{min}}(K) \geq \frac{1}{2} \), and \( \alpha_1 \geq 0 \), the following inequalities hold

\[
\| \tilde{x} \| > \sqrt{\frac{\epsilon_{IM}}{\lambda_{\text{min}}(K) - \frac{1}{2}}} \quad \text{or} \quad \| \tilde{W}_I \| > \sqrt{\frac{2\epsilon_{IM}}{\alpha_1}} \tag{24}
\]

It can be observed from (24) that \( \| \tilde{x} \| \) bound can be decreased by increasing the eigenvalue \( K \). Therefore, \( \| \tilde{x} \| \) is quantified by selecting the minimum eigenvalue \( \lambda_{\text{min}}(K) \). According to (19) and the relationship between \( \| \tilde{x} \| \) and \( \| \tilde{W}_I \| \), the smaller \( \| \tilde{x} \| \) will enforce the \( \| \tilde{W}_I \| \) to converge into a small bound. Therefore, we have \( J_I \leq 0 \) and it can be concluded that \( \| \tilde{x} \| \) and \( \| \tilde{W}_I \| \) are UUB.

This completes the proof. \( \square \)
3.2.2. NDP-Based Guidance Strategy

In order to confront the nonlinear HJI function, according to the universal approximation property of the neural network, the optimal cost function can be reconstructed by a critic NDP on a compact set as

\[
V(x, t) = W_V^T h(x, t_f - t) + \epsilon_V(x, t)
\]

(25)

where \( W_V \in R^{V \times N} \) is the ideal weight matrix. \( h(x, t_f - t) \in R^N \) denotes NN activation function vectors, \( N \) is the number of hidden layer neurons, and \( \epsilon_V \) represents the NN approximation error.

Thus, the terminal cost function can be expressed as

\[
V(x, t_f) = W_V^T h(x(t_f)), 0 + \epsilon_V(x, t_f)
\]

(26)

**Remark 3.** The critic NN activation function \( h(x, t) \) and its gradient \( \nabla h(x, t) \) are upper bounded, i.e., \( \|h(x, t)\| \leq h_M \) and \( \|\nabla h(x, t)\| \leq h_{DM} \), with \( h_M \) and \( h_{DM} \) positive constants. The critic NN weight \( W \) is upper bounded, i.e., \( \|W_V\| \leq W_{VM} \), with \( W_{VM} \) being a positive constant. The critic NN approximation error \( \epsilon_V(x) \) and its gradient \( \nabla \epsilon_V(x) \) are upper bounded, i.e., \( \|\epsilon_V(x)\| \leq \epsilon_{VM} \) and \( \|\nabla \epsilon_V(x)\| \leq \epsilon_{DM} \), with \( \epsilon_{VM} \) and \( \epsilon_{DM} \) positive constants. Obviously, the activation function is a time-carying function, which can maintain the fixed final time.

Next, the partial derivation of \( V(x, t) \) with respect to \( x \) and \( t \) can be obtained, respectively.

\[
\begin{align*}
V_x &= \nabla h^T(x, t_f - t) W_V + \nabla_x \epsilon_V(x, t) \\
V_t &= \nabla h^T_t(x, t_f - t) W_V + \nabla_t \epsilon_V(x, t)
\end{align*}
\]

(27)

where \( \nabla h^T_x(x, t_f - t) = \partial h(x, t_f - t) / \partial x \), \( \nabla h^T_t(x, t_f - t) = \partial h(x, t_f - t) / \partial t \), \( \nabla_x \epsilon_V(x, t) = \partial \epsilon_V(x, t) / \partial x \) and \( \nabla_t \epsilon_V(x, t) = \partial \epsilon_V(x, t) / \partial t \).

Therefore, by substituting (27) into (8), we then obtain the differential game guidance strategy as

\[
\begin{align*}
\begin{cases}
 v^*(x) = -\frac{1}{\gamma} R^{-1}_1 g^T(x) V^*_V = -\frac{1}{\gamma} R^{-1}_1 g^T(x) \left( \nabla h^T_x(x, t_f - t) W_V + \nabla_x \epsilon_V(x, t) \right) \\
w^*(x) = \gamma^2 k^T(x) V^*_V = \gamma^2 k^T(x) \left( \nabla h^T_x(x, t_f - t) W_V + \nabla_x \epsilon_V(x, t) \right)
\end{cases}
\end{align*}
\]

(28)

By substituting (28) into (10), the HJI function can be rewritten as

\[
\begin{align*}
H(x, u, w) &= W_V^T \nabla h^T_x(x, t_f - t) + Q(x) + W_V^T \nabla h^T_x(x, t_f - t) f(x) \\
&- \frac{1}{4} W_V^T \nabla h^T_x(x, t_f - t) D_1(x) \nabla h^T_x(x, t_f - t) W_V + \frac{\gamma^2}{2} W_V^T \nabla \epsilon_V(x, t) k(x) k^T(x) \nabla h^T_x(x, t_f - t) W_V + \epsilon_{HJB}(x, t)
\end{align*}
\]

(29)

where \( D_1(x) = g(x) R^{-1}_1 g^T(x) \) and

\[
\begin{align*}
\epsilon_{HJB}(x, t) &= \nabla_x \epsilon_V(x, t) + \frac{1}{2} W_V^T h(x, t_f - t) \nabla_x \epsilon_V(x, t) D_1(x) \nabla_x \epsilon_V(x, t) - \gamma^2 W_V^T h(x, t_f - t) \nabla_x \epsilon_V(x, t) D_1(x) \nabla_x \epsilon_V(x, t) \\
&- \frac{1}{2} \nabla_x \epsilon_V(x, t) D_1(x) \nabla h^T_x(x, t_f - t) W_V - \frac{1}{2} \nabla_x \epsilon_V(x, t) D_1(x) \nabla x \epsilon_V(x, t) \\
&+ \gamma^2 \nabla_x \epsilon_V(x, t) k(x) k^T(x) \nabla h^T_x(x, t_f - t) W_V + \frac{\gamma^2}{2} \nabla_x \epsilon_V(x, t) k(x) k^T(x) \nabla x \epsilon_V(x, t) \\
&+ \frac{1}{2} W_V^T \nabla h(x, t_f - t) D_1(x) \nabla x \epsilon_V(x, t) \\
&+ \nabla x \epsilon_V(x, t) f(x) - \frac{1}{4} W_V^T \nabla h(x, t_f - t) k(x) k^T(x) \nabla x \epsilon_V(x, t)
\end{align*}
\]

(30)
Because ideal NN weights are typically unknown, we define the estimated cost function $\hat{V}(x, t)$ as follows

$$\hat{V}(x, t) = \hat{W}_V^T h\left(x, t_f - t\right) \tag{31}$$

The estimated terminal cost function is

$$\hat{V}(x, t_f) = \hat{W}_V^T h\left(\hat{x}(t_f), 0\right) \tag{32}$$

where $\hat{W}_V$ denotes the estimate of the $W_V$ and $h\left(\hat{x}(t_f), 0\right)$ is the activation function with the estimated terminal state $\hat{x}(t_f)$.

Next, the partial derivation of the estimated cost function $\hat{V}(x, t)$ with respect to $x$ and $t$ can be obtained, respectively.

$$\hat{V}_x = \nabla h_t^T(x, t_f - t) \hat{W}_V$$
$$\hat{V}_t = \nabla h_t^T(x, t_f - t) \hat{W}_V \tag{33}$$

where $\hat{V}_t = \frac{\partial \hat{V}(x, t)}{\partial t}$ and $\hat{V}_x = \frac{\partial \hat{V}(x, t)}{\partial x}$.

Then, by applying (33) to (8), the estimated differential game guidance strategy can be rewritten as

$$\begin{cases}
\hat{u}(x) = -\frac{1}{2} R_1^{-1} g^T(x) \hat{V}_x = -\frac{1}{2} R_1^{-1} g^T(x) (\nabla h_t^T(x, t_f - t)) \hat{W}_V \\
\hat{w}(x) = \gamma^2 k^T(x) \hat{V}_x = \gamma^2 k^T(x) (\nabla h_t^T(x, t_f - t)) \hat{W}_V \tag{34}
\end{cases}$$

By applying (34) to (10), the estimated HJB function yields

$$\hat{H}(x, u, w) = \hat{W}_V^T \nabla h_t^T(x, t_f - t) + \hat{W}_V^T \nabla h_x(x, t_f - t) \hat{f}(x)$$
$$-\frac{1}{4} \hat{W}_V^T \nabla h_x(x, t_f - t) \hat{D}_1(x) \nabla h_t^T(x, t_f - t) \hat{W}_V + \frac{\gamma^2}{2} \hat{W}_V^T \nabla h_x(x, t_f - t) \hat{k}(x) \hat{k}^T(x) \nabla h_t^T(x, t_f - t) \hat{W}_V = e_c \tag{35}$$

In order to obtain the optimal differential game guidance strategy, we define the estimated terminal cost error as

$$e_{tf} = \Psi\left(x(t_f), 0\right) - \hat{W}_V^T h\left(\hat{x}(t_f), 0\right) = \hat{W}_V^T h\left(x(t_f), 0\right) + \hat{W}_V^T h\left(\hat{x}(t_f), 0\right) + e_{tf} \tag{36}$$

where $\hat{h}\left(x(t_f), 0\right) = h\left(x(t_f), 0\right) - h\left(\hat{x}(t_f), 0\right)$.

Moreover, in order to guarantee the estimated NN weight $\hat{W}_V$, which is approximated to the ideal NN weight $W_V$, and after combining the time-varying nature of the cost function and the estimated terminal cost error, the total NN approximation error is defined as

$$E_{total} = \frac{1}{2} \hat{e}_c^2 + \frac{1}{2} e_{tf}^2 \tag{37}$$

By using the gradient descent algorithm, a novel simplified weight tuning law is proposed with additional terms to ensure the stability of the nonlinear system (1), as follows:

$$\hat{W}_V = -a_2 \left(\frac{\omega + \frac{1}{2} sng(\omega)}{1 + \omega^2}\right) e_c + a_3 \left(\frac{\xi + \frac{1}{2} sng(\xi)}{1 + \xi^2}\right) e_{tf}^2$$
$$+ \frac{\gamma^2}{2} \nabla h_t^T(x, t_f - t) \left(\hat{D}_1(x) - 2 \gamma^2 \hat{k}^T(x) \hat{k}^T(x)\right) Q(x, t) \tag{38}$$

where $\omega = \nabla h_t^T(x, t_f - t) + \nabla h_x(x, t_f - t) \hat{f}(x) + \frac{1}{2} \nabla h_x(x, t_f - t) \hat{k}(x) \hat{k}^T(x) \hat{D}_1(x) \hat{k}^T(x)$
$$\nabla h_t^T(x, t_f - t) \hat{W}_V, \ \xi = h\left(\hat{x}(t_f), t_f\right)$$. 
Remark 4. It is important to mention that the first term in (38) is used to minimize the squared residual error. The second term is used to minimize the terminal cost estimation error. The last term is used to guarantee that the system states remain bounded.

4. Stability Analysis

In order to prove the stability of the weight tuning law and the nonlinear system by choosing the optimal strategy, and without loss of generality, the weight estimation error of the critic NN is defined as \( \tilde{W}_V = W_V - \hat{W}_V \); then, we have \( \dot{\tilde{W}}_V = -\tilde{W}_V \). Therefore, the approximated HJI is

\[
\dot{H}(x, u, w) = -\tilde{W}_V^T \nabla h^T \left( x, t_f - t \right) - \tilde{W}_V^T \nabla h \left( x, t_f - t \right) \tilde{f}(x) - W_V^T \nabla h \left( x, t_f - t \right) \tilde{f}(x) \\
+ \frac{1}{2} \tilde{W}_V^T \nabla h \left( x, t_f - t \right) \left( \tilde{D}_I(x) - 2\gamma^2 \tilde{k}^T(x) \tilde{k}^T(x) \right) \nabla h^T \left( x, t_f - t \right) \tilde{W}_V \\
+ \frac{1}{2} W_V^T \nabla h \left( x, t_f - t \right) \left( \tilde{D}(x) - 2\gamma^2 \tilde{k}^T(x) \tilde{k}(x) - 4\gamma^2 \tilde{k}^T(x) \tilde{k}(x) \right) \nabla h^T \left( x, t_f - t \right) W_V + \epsilon_{HJB}
\]

(39)

Assumption 1. For the nonlinear system (1) with the cost function (2) and the optimal guidance law (34), let the value function \( V(x, t) \) be the Lyapunov function and continuously differentiable, when

\[
\dot{V}(x, t) = V_x(x, t) \dot{x} + V_t(x, t) + V_s(x, t)(f(x) + g(x)u^* + k(x)\omega^*) < 0
\]

(40)

Exits and is it clear that the inequality holds

\[
V_t(x, t) + V_s(x, t)(f(x) + g(x)u^* + k(x)\omega^*) < -Q(x, t)
\]

(41)

Theorem 3. For the nonlinear system (1) with the ideal HJI Equation (9), let the updated law for the NN-based identifier and NDP-based cost function approximator be provided by (20) and (38), respectively and the estimated optimal guidance laws are given in (34). The existence of the positive constants \( B_{Q,x}, B_{W \Bar{V}}, B_{x} \) and \( B_{W, x} \) mean that the identification error, weight estimation error and the controller are UUB.

Proof of Theorem 3. Select the Lyapunov candidate function as

\[
J_V = J_{aV} + J_{bV} + J_{cV} + J_{dV}
\]

(42)

where

\[
J_{aV} = \frac{a_2}{4} (\tilde{x}^T \tilde{A} \tilde{x})^2 + \frac{a_3}{4} (\text{tr} (\tilde{W}^T_I \tilde{A} \tilde{W}_I))^2, \quad J_{bV} = \frac{1}{2} \tilde{W}_V^T \tilde{W}_V, \quad J_{cV} = a_4 V(x, t), \quad J_{dV} = \frac{1}{2} \tilde{x}^T \tilde{G} \tilde{x} + \frac{1}{2} \text{tr} (\tilde{W}_I^T \tilde{A} \tilde{W}_I).
\]

It can be observed that \( J_V > 0 \).

First, the derivation of \( J_{aV} \) with respect to time is given by

\[
\dot{J}_{aV} \leq \|A\|^2 a_2 (\tilde{x}^T \tilde{x}) \dot{\tilde{x}}^T \tilde{x} + \|A\|^2 a_3 \text{tr} (\tilde{W}^T_I \tilde{W}_I \dot{W}_I) \text{tr} (\tilde{W}_I^T \tilde{W}_I) \\
\leq -a_2 (\lambda_{\min}(K_f) - \frac{3}{2} - \frac{a_3}{2a_2}) \|\tilde{x}\|^4 + \frac{1}{2} \|A\epsilon_I\|^2 \|\tilde{x}\|^2
\]

(43)

where \( a_2 > 0, a_3 > 0 \) and \( a_f > 4 + \frac{a_2 + a_3}{4a_2} \|\tilde{x}\|^4, K_I \) satisfies \( \lambda_{\min}(K_f) > \frac{3}{2} + \frac{a_3}{4a_2}, \epsilon_{CM} = \frac{a_3 a_f}{4} \|A\|^2 W_{f,M}^4 + \frac{\|A\|^2 a_2}{4} \|\epsilon_I\|^2 \).

The derivation of \( J_{bV} \) is governed by...
\[ \dot{J}_{BV} = \dot{W}_V \Pi \dot{W}_V \]

\[ \leq \| \Pi \| \left( \frac{\dot{W}_V^T \alpha x_2 \frac{\tilde{\omega}}{1 + \tilde{\omega}^T \omega}}{1 + \tilde{\omega}^T \omega} \right) e_c - \alpha_3 \dot{W}_V^T \frac{\tilde{\xi}}{1 + \tilde{\omega}^T \omega} e_{t_f} - \frac{\alpha_4}{2} \| \dot{W}_V \| \| \nabla h_x(x, t_f - t) \| \left( \dot{D}_1(x, t) + \dot{k}(x)k^T(x) \right) Q(x, t) \]

\[ \leq -\alpha_3 \| \Pi \| \frac{\dot{W}_V^T \alpha x_2 \frac{\tilde{\omega}}{1 + \tilde{\omega}^T \omega}}{1 + \tilde{\omega}^T \omega} + \frac{\alpha_2 W_{2M}^2 \lambda_{max}(R^{-1}) \nabla h_x \dot{\alpha} \| \dot{W}_V \| ^4}{1 + \tilde{\omega}^T \omega} \]

\[ -\frac{\alpha_4}{2} \| \dot{W}_V \| \| \nabla h_x(x, t_f - t) \| \left( \dot{D}_1(x, t) - 2\gamma^2 \dot{k}(x)k^T(x) \right) Q(x, t) + \epsilon_{VH} \]

where

\[ 0 < \alpha_2 \leq \frac{3\alpha_2 (1 + \tilde{\omega}^T \omega)^2}{\lambda_{max} \left( \nabla h_x(x, t_f - t) \right)} \left( \dot{D}_1(x, t) - 2\gamma^2 \dot{k}(x)k^T(x) \right) \nabla h_x(x, t_f - t) \] (44)

and

\[ \epsilon_{VH} = \frac{\alpha_2 \| \Pi \| \dot{W}_V^T \alpha x_2 \frac{\tilde{\omega}}{1 + \tilde{\omega}^T \omega}}{1 + \tilde{\omega}^T \omega} + \frac{3}{2} \left( \dot{W}_V^T h(x(t_f), 0)h^T(x(t_f), 0)W_V + \epsilon_r \right) \]

\[ + \frac{15}{2} \left( \alpha_3 \frac{\dot{\omega}}{1 + \tilde{\omega}^T \omega} \right) \| \Pi \| \| \dot{W}_V \| ^2 + \frac{1}{2} \lambda_{min}(R^{-1}) \frac{\alpha_2 W_{2M}^2 \lambda_{max}(R^{-1}) \nabla h_x \dot{\alpha} \| \dot{W}_V \| ^4}{1 + \tilde{\omega}^T \omega} \]

\[ + \frac{\alpha_2}{2} \left( \frac{\tilde{\xi}}{1 + \tilde{\omega}^T \omega} \right) \| \Pi \| \| \dot{W}_V \| ^2 + \frac{\alpha_3}{2} \| \nabla h_x(x, t_f - t) \| \left( \dot{D}_1(x, t) - 2\gamma^2 \dot{k}(x)k^T(x) \right) Q(x, t) + \epsilon_{VH} \]

\[ -\frac{\alpha_2}{2} \left( \lambda_{min}(K_I - 3) - \frac{\alpha_2}{8\alpha_3} \right) \| \Pi \| \| \dot{W}_V \| ^2 + \frac{\alpha_2}{2} \left( \lambda_{min}(K_I - 3) - \frac{\alpha_2}{8\alpha_3} \right) \| \Pi \| \| \dot{W}_V \| ^2 + \epsilon_{IM} \]

\[ \leq \frac{\alpha_4}{4} \left( \frac{\tilde{\omega}}{1 + \tilde{\omega}^T \omega} \right) \| \Pi \| \| \dot{W}_V \| ^2 + \frac{\alpha_4}{4} \left( \frac{\tilde{\omega}}{1 + \tilde{\omega}^T \omega} \right) \| \Pi \| \| \dot{W}_V \| ^2 + \epsilon_{IM} \]

(46)

\[ \epsilon_{IM} = \epsilon_{FM} + \epsilon_{M} u_{M}^* + \epsilon_{M} d_{M}^* \]

where \( u^* \) and \( d^* \) are the upper bounds of the optimal guidance strategy.
Therefore, when the following conditions hold, the first derivation of $J_V$ is less than zero.

$$\|Q(x,t)\| > \sqrt{\frac{2\rho TC}{a_1}} = B_{Q_1}$$

$$\|\tilde{W}_V\| > \sqrt{\frac{2\rho TC}{a_2Vr_2}} = B_{W_V}$$

$$\|\tilde{x}\| > \sqrt{\frac{a_{TC}I_{\min}(k_{\min})\|x\|}{16\rho TC}} = B_{Z_1}$$

This completes the proof. $\square$

**Remark 5.** The eigenvalue $K_a$, $a_1$, $a_2$, $a_3$, and $a_4$ are the tuning parameters for guaranteeing the lower bound of $B_{Q_2}$, $B_{W_V}$, $B_Z$ and $B_{W_m}$, which can quantify the bound of the system. In additions, from the proof, we can observe that the estimated optimal guidance laws given in (34) can ensure the system is UUB. Thus, by combining the robust optimal feedback control (11), the complete nonlinear finite-time robust differential game guidance law is provided.

5. Application

A missile–target engagement scenario is considered in this section. The engagement geometry of the missile–target is shown in Figure 1, where the X-Y plane represents the Cartesian reference frame. The variables $V$ and $A$ denote the speed and the normal acceleration of the missile and the target, respectively. $\alpha$ and $\beta$ denote the flight angles of the missile and the target, respectively. The variables $r$ and $\theta$ represent the missile–target distance and the line of sight (LOS) angle, and LOS angular rate is $\dot{\theta}$ donated by $\sigma$. $u$ and $w$ are the control vectors perpendicular to the velocity of the missile, and the target.

![Figure 1. Missile–target engagement geometry.](image)

The engagement occurs in the terminal guidance phase and all the participators are assumed to neglect the effect of gravity and have constant velocity. The nonlinear kinematics of the missile and the target, in a polar coordinate system, is given by

$$V_r = \dot{r} = V_T\cos(\beta - \theta) - V_M\cos(\alpha - \theta)$$

(47)

$$\sigma = \dot{\theta} = (V_T\sin(\beta - \theta) - V_M\sin(\alpha - \theta))/r$$

(48)

where $V_r$ represents the closing velocity.

The first-order dynamics of the missile are considered, and the motions of the missile are as follows:

$$\dot{x}_M = V_M\cos\alpha$$

(49)

$$\dot{y}_M = V_M\sin\alpha$$

(50)

$$\dot{a} = \frac{a_M}{V_M}$$

(51)

$$\dot{a}_M = \frac{u_M - a_M}{\tau_M}$$

(52)

where $(x_M, y_M)$ is the position of the missile along with the Cartesian reference frame. $a_M$ is the lateral acceleration of the missile and $\tau_M$ is a time constant.
Similarly, the motion equations of the target are as follows:

\[
\begin{align*}
\dot{x}_T &= V_T \cos \beta \\
\dot{y}_T &= V_T \sin \beta \\
\dot{\beta} &= \frac{a_T}{V_T} \\
\dot{a}_T &= \frac{u_T - a_T}{\tau_T}
\end{align*}
\]  

(53) (54) (55) (56)

where \((x_T, y_T)\) is the position of the target along with the Cartesian reference frame. \(a_M\) is the lateral acceleration of the target and \(\tau_T\) is a time constant.

To obtain the capture zone, the guidance principle is adopted as follows.

**Definition 2.** Zero effort miss distance (ZEMD) is the closest distance between the missile and the target at an instant \(t\), while the missile and the target do not impose any control and the agents continue to perform their scheduled maneuver strategy from the current time to the endgame. The ZEMD is computed as

\[
r_{\text{miss}}(t) = \frac{r^2 \sigma}{\sqrt{V_r^2 + r^2 \sigma^2}}
\]

(57)

In this scenario, the missile applies its optimal strategy (34) to minimize the ZEMD, while the target applies its optimal strategy (34) to maximize the ZEMD. In addition, from (57), we can observe that if \(\sigma\) tends to be zero, the ZEMD also tends to be zero. Meanwhile, the closing \(V_r\) is also less than zero. Therefore, to ensure that the missile successfully intercepts the target, the following two conditions hold, which will be verified in simulations:

\[
\sigma \to 0, \ V_r < 0
\]

(58)

Based on the two conditions mentioned above, by choosing the \(x = [\theta \ \sigma]^T\) as the system state, differentiating the Equation (48) with respect to time, the nonlinear system can be given by

\[
\begin{cases}
\dot{x}_1 = x_2 \\
\dot{x}_2 = -\frac{2V_r}{r} x_2 - \frac{\cos(\beta - \theta)}{r} u_M + \frac{\cos(\alpha - \theta)}{r} w_T
\end{cases}
\]

(59)

**Remark 6.** From (59), it can be observed that when \(r\) is close to 0, the nonlinear terms \(f(x), g(x),\) and \(k(x)\) tend to be infinite. It will mean that the system is broken and the proposed differential game guidance laws do not work. Thus, a minimized ZEMD \((r_{\text{min}})\) should be designed, which means that when the displacement between the missile and the target is less or equal to the \(r_{\text{min}} = 0.5\) m, the missile completes the interception mission. Moreover, in the practice application, the missile has an actual killing radius; thus, this design is reasonable.

**Remark 7.** From (59), it also can be found that when \(|\beta - \theta| = \pi/2\) and \(|\alpha - \theta| = 0\), no matter what the proposed guidance laws change, the nonlinear system is an unstable equilibrium. Therefore, the domain where the differential game-based guidance laws are applicable is given by

\[
\varphi = \left\{ x : |(\beta - \theta)| \neq \frac{\pi}{2}, |(\alpha - \theta)| \neq 0, r \neq 0, V_r < 0 \right\}
\]

(60)

6. Simulation Results

In this section, some experiments are designed to verify the proposed finite-time robust differential game guidance strategy. To further explain the performance of the proposed differential game guidance strategy based on NDP, some simulation experiments are designed and analyzed. The initial engagement occurs in the terminal guidance phase.
The speeds of the missile and the target are \( V_M = 700 \text{ m/s} \), and \( V_T = 400 \text{ m/s} \), respectively. The initial position of the target is at (0 m, 0 m), and the position of the missile is (2500 m, 0 m). The initial flight path angles of the missile and the target are \( \alpha = 70^\circ \) and \( \beta = 150^\circ \), respectively. The time constants are \( \tau_M = 0.1 \) and \( \tau_T = 0.1 \), respectively. The weight parameters \( R_1 = 5, R_2 = 50, \gamma = 10 \) and \( t_f = 7 \text{ s} \) are set. \( Q_x = 20(x_1^2 + x_2^2 + \tau^2)^2 \), where \( \tau = t_f - t \).

Furthermore, the initial parameters of the critic NN are selected inside \( w_V \in [0 1] \) randomly and the critic NN identifier activation function is designed as \( \sigma_{I(s)} = [1 x_1^2 x_1 x_2^2 x_2^3 x_1^4 x_2^4] \), which refers to [16,21]. In addition, the polynomial basis functions can better approximate the nonlinear system. The initial parameters of the critic NN are \( w_V = [10 10 10 10 10 10 10 10] \) and the critic NN vector activation function for estimating the cost function is designed as

\[
\begin{align*}
h(x, t_f - t) &= [x_1^2 \exp(-\tau) x_2^2 \exp(-\tau) x_1 x_2 \exp(-\tau) x_1^3 x_2^3 x_1 x_2^2] T. 
\end{align*}
\]

The learning rates are \( \alpha_1 = 0.01, \alpha_2 = 0.05, \alpha_3 = 0.55 \) and \( \alpha_4 = 0.15 \), respectively. The experiment environments are performed on a PC platform with i7-9750H CPU, using Matlab 2020b.

### 6.1. Effect of the Proposed Finite-Time Differential Game Guidance Strategy without Unknown Uncertainties

In this case, the performance of the proposed finite-time differential game guidance strategy is verified without unknown uncertainties and the missile and the target choose the optimal guidance strategy (34). More importantly, the ZEMD and the lateral acceleration of the agents are the performance evaluation indicators and we have rewritten the physical meaning. The performance of the ZEMD tends to be zero, indicating that the missile intercepts the target successfully. The performance of the lateral accelerations of the agents is limited in the range \( \pm 100 \text{ g} \) and stable change indicates that the agents can be reasonably controlled. The simulation results for this engagement scenario are shown in Figures 2–4.

![Figure 2](image-url)

(a) Engagement trajectory  
(b) The relative distance

**Figure 2.** Engagement scenario for the proposed guidance strategy.
which leads the ZEMD to be zero according to the Equation (57) and Figure 2b. Furthermore, it also verifies Definition 2, which mentioned that critic NN weights can guarantee the stability of the closed-loop nonlinear system.

The simulation results for this engagement scenario are shown in Figure 3. Curves of the angular rate and the relative velocity for the proposed guidance strategy are presented in Figure 3. The angular rate is observed to be zero at about 4.5 s, indicating that the state variable (the angular rate) bound is zero according to the Equation (57) and Figure 2b. σ tends to be zero, which can ensure that the missile intercepts the target. Furthermore, the negative $V_r$ guarantees the relative distance between the missile and the target tends to be zero in Figure 3b. Moreover, sharp changes between $\sigma$ and $V_r$ are reasonable due to the system characteristics. Furthermore, it also verifies Definition 2, which mentioned that $\sigma \to 0$, and $V_r < 0$.

Figure 4 presents lateral acceleration curves of the target and the missile. It can be observed that both lateral accelerations are maintained within a reasonable range, which can ensure that the missile intercepts the target. However, due to the system characteristics, the lateral accelerations decrease sharply at the end of the engagement.

The convergence curves of critic NN weights are presented in Figure 5 and it can be observed that the critic NN weights finally converge to stable coefficients. The result reveal that critic NN weights can guarantee the stability of the closed-loop nonlinear system.
6.2. Engagement without Unknown Uncertainties for a Maneuvering Target

To further verify the effectiveness of the proposed guidance strategy dealing with the target executing other forms of target maneuvers, the following experiment is presented. In this experiment, the target is expected to perform a sin-wave maneuver with a magnitude of 10 g and the missile still selects the guidance law (34). The simulation results for this engagement scenario are shown in Figures 6–8.

**Figure 5.** Convergence curves of critic NN weights.

**Figure 6.** Engagement scenario for maneuvering target.

**Figure 7.** Curves of the angular rate and the relative velocity for a maneuvering target.
In this experiment, a uniform distribution between −0.2 and 0.2, are considered in both input vectors. The simulation results for this engagement scenario are shown in Figures 9–11.

Figure 8. Acceleration demand of the missile and the target for a maneuvering target.

Figure 6a shows the engagement trajectories of the missile confronting the target, which is expected to perform a sin-wave maneuver by selecting the proposed guidance law (34). Figure 6b presents the change in the relative distance between the missile and the target and it can be observed that the final ZEMD is less than 0.5 m. The results reveal that the missile intercepts the maneuvering target successfully.

Curves of the angular rate and the range rate are presented in Figure 7. The results reveal the same conclusion as case 1.

Figure 8 presents the lateral acceleration curves of the target and the missile. It can be observed that both lateral accelerations are maintained within a reasonable range, which can ensure that the missile intercepts the target. Furthermore, compared with Figure 4a, due to the maneuvering target, higher acceleration demands of the missile are needed.

6.3. Engagement with Unknown Uncertainties for a Maneuvering Target

To further verify how the proposed finite-time robust guidance strategy deals with unknown uncertainties, the following experiment is presented. In this experiment, the missile selects the robust differential game guidance law (11), and the target is expected to perform a sin-wave maneuver with a magnitude of 10 g. Furthermore, external disturbances, with a uniform distribution between −0.2 and 0.2, are considered in both input vectors. The simulation results for this engagement scenario are shown in Figures 9–11.

Figure 9. Engagement scenario for the target’s maneuverability.
Figure 10. Curves of the angular rate and the relative velocity with unknown uncertainties.

Figure 11. Acceleration demands of the missile and the target.

Figure 9a shows the trajectories of the missile confronting the target, which is expected to perform a sin-ware maneuver by selecting the proposed guidance law. Figure 9b presents the change in the relative distance between the missile and the target and it can be observed that the miss distance is less than 0.5 m. By combining (a) and (b), the results reveal that the missile intercepts the maneuvering target successfully.

Curves of the angular rate and the range rate are presented in Figure 10. The results reveal the same conclusion as case 1.

Figure 11 presents the acceleration demands and the control requirements of the target and the missile. It can be observed that the lateral accelerations are maintained within a
reasonable range, which can that ensure the missile intercepts the target. Furthermore, when the missile experiences external disturbances, the proposed robust differential game guidance strategy can successfully intercept maneuvering targets with external disturbances.

### 6.4. Effect of the Proposed Robust Optimal Differential Game Guidance Strategy with Other Methods

To confirm the advantage of the proposed robust optimal differential game guidance law, we offer a comparative experiment. In this experiment, the target selects the proposed robust optimal differential game guidance strategy (11), while the missile chooses the proposed differential game guidance strategy (11), i.e., the OGL in [4], and the conventional differential game guidance law (CDGGL) in [18], respectively. To further illustrate the advantages of the proposed differential game guidance strategy, the control effort \( J \) is defined as \( J = \int_0^\infty u^T(\tau)u(\tau)\,d\tau \). The comparison results for this engagement scenario are shown in Figure 12.

#### Figure 12. Engagement scenario and the control effort of the missile.

The engagement trajectories and the control effort of the missile are presented in Figure 12a. It can be observed that the OGL, the CDGGL, and the proposed robust optimal differential game guidance strategy both can successfully intercept the target, with a miss distance equal to 0.985, 1.218, and 0.0024 m, respectively. Furthermore, the curves of the missile control effort are presented in Figure 12b. It can be shown that the minimal control effort is our proposed robust optimal differential game guidance strategy, the second control effort is the CDGGL in [18], and the maximum control effort is the OGL in [4]. All control efforts are limited to the setting range (100 g). More importantly, it also can be observed that before the missile intercepts the target, the control effort we proposed is always at its minimum at different times. For example, when \( t = 4.8 \text{ s} \), the control effort of our method is minimal. Thus, the missile uses our method to intercept the target, which saves more energy. Furthermore, the OGL in [4] is larger than the proposed robust optimal differential game guidance strategy, which means that the missile may not successfully intercept the target when using the OGL in some acceleration-limited scenes. In general, our proposed guidance law is superior to the OGL in [4] and the CDGGL in [18]. To further illustrate the superiority of the proposed robust optimal differential game guidance strategy, the energy consumption and the simulation time are compared in 100 average experiments. The comparison results are shown in Table 1.
Table 1. Comparison results of OGL, CDGGL, and our method.

| Method              | Control Effort | Computation Time (t) | Number of Tests |
|---------------------|----------------|----------------------|-----------------|
| OGL in [4]          | 980.56         | 1.647                | 100             |
| CDGGL in [18]       | 570.64         | 1.083                | 100             |
| Our method          | 218.845        | 0.451                | 100             |

From Table 1, it is easily found that less control effort is needed in our proposed method compared with the OGL in [4] and the CDGGL in [18], which implies that our proposed guidance law can reduce unnecessary energy consumption. The missile can intercept the target with minimal control. Moreover, the computation time of our proposed method is shorter compared with the OGL in [4] and the CDGGL in [18], which means that our proposed method can make the confrontation strategy in the shortest time. In general, our proposed guidance law can intercept targets with minimal control effort and a fast response.

7. Conclusions

In this paper, a finite-time robust differential game law for the nonlinear two-player zero-sum game is proposed for unknown system dynamics with external disturbance. The robustness and optimality of the guidance strategy are proven under a time-varying cost function. In order to solve the fixed terminal time constraint, the HJI equation is solved by a critic NN with time-varying activation functions, and extra weight tuning terms are introduced to guarantee the stability of the closed-loop system and the minimization of the HJI approximation error. The NN identifier estimates the nonlinear system with an online tuning law, which can be subsequently utilized in the finite-time guidance strategy design. The proposed scheme yields an online guidance strategy design that enjoys great practical benefits. Finally, in the missile–target engagement experiments, the missile with the proposed differential game guidance law successfully intercepts the target in different forms of maneuvers and all the results verify the theoretical claims.

Author Contributions: Conceptualization, A.X. and Y.C.; methodology, A.X.; software, A.X.; validation, A.X.; formal analysis, A.X.; investigation, A.X. and Y.C.; resources, Y.C.; data curation, A.X.; writing—original draft preparation, A.X.; writing—review and editing, A.X.; visualization, A.X.; supervision, Y.C.; project administration, Y.C.; funding acquisition, Y.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Program of China under grant (2018YFB1700100).

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Guo, Y.; Li, X.; Zhang, H.; Cai, M.; He, F. Data-Driven Method for Impact Time Control Based on Proportional Navigation Guidance. *J. Guid. Control Dyn.* 2022, 43, 955–966. [CrossRef]
2. Franzini, G.; Tardioli, L.; Pollini, L.; Innocenti, M. Visibility Augmented Proportional Navigation Guidance. *J. Guid. Control Dyn.* 2018, 41, 987–995. [CrossRef]
3. Ghosh, S.; Ghose, D.; Raha, S. Capturability of Augmented Pure Proportional Navigation Guidance Against Time-Varying Target Maneuver. *J. Guid. Control Dyn.* 2014, 37, 780–787. [CrossRef]
4. Chen, X.; Wang, J. Optimal-control based guidance law to control both impact time and impact angle. *Aerosp. Sci. Technol.* 2019, 84, 454–463. [CrossRef]
5. Harl, N.; Balakrishnan, S.N. Impact Time and Angle Guidance with Sliding Mode Control. *IEEE Trans. Control Syst. Technol.* 2012, 20, 1436–1449. [CrossRef]
6. Re, R.; Rp, R. Three-Party Differential Game Theory Applied to Missile Guidance Problem. In *Differential Game Theory with Applications to Missiles and Autonomous Systems Guidance*; John Wiley & Sons: Hoboken, NJ, USA, 2017; p. 102.
7. Bardhan, R.; Ghose, D. Intercepting maneuvering target with specified impact angle by modified SDRE technique. In Proceedings of the 2012 American Control Conference (ACC), Montreal, QC, Canada, 27–29 June 2012; pp. 4613–4618.

8. Al-Tamimi, A.; Lewis, F.L.; Abu-Khalaf, M. Model-free Q-learning designs for linear discrete-time zero-sum games with application to H-infinite control. *Automatica* 2007, 43, 473–481. [CrossRef]

9. Xin, X.; Tu, Y.; Stojanovic, V.; Wang, H.; Shi, K.; He, S.; Pan, T. Online reinforcement learning multiplayer non-zero sum games of continuous-time Markov jump linear systems. *Appl. Math. Comput.* 2022, 1, 126537. [CrossRef]

10. Jiang, Y.; Gao, W.; Na, J.; Zhang, D.; Hämäläinen, T.T.; Stojanovic, V.; Lewis, F.L. Value iteration and adaptive optimal output regulation with assured convergence rate. *Control Eng. Pract.* 2022, 121, 105042. [CrossRef]

11. Xu, Z.; Li, X.; Stojanovic, V. Exponential stability of nonlinear state-dependent delayed impulsive systems with applications. *Nonlinear Anal. Hybrid Syst.* 2021, 42, 101088. [CrossRef]

12. Gao, W.; Jiang, Z.-P. Adaptive Optimal Output Regulation of Time-Delay Systems via Measurement Feedback. *IEEE Trans. Neural Netw. Learn. Syst.* 2019, 30, 938–945. [CrossRef]

13. Wang, F.Y.; Jin, N.; Liu, D.; Wei, Q.L. Adaptive dynamic programming for finite horizon optimal control of discrete-time nonlinear systems with $\varepsilon$-error bound. *IEEE Trans. Neural Netw.* 2010, 22, 24–36. [CrossRef]

14. Vamvoudakis, K.G.; Lewis, F.L. Online Solution of Nonlinear Two-Player Zero-Sum Games Using Synchronous Policy Iteration. *Int. J. Robust Nonlinear Control* 2012, 22, 1460–1483. [CrossRef]

15. Dierks, T.; Jagannathan, S. Online Optimal Control of Affine Nonlinear Discrete-Time Systems with Unknown Internal Dynamics by Using Time-Based Policy Update. *IEEE Trans. Neural Netw. Learn. Syst.* 2012, 23, 1118–1129. [CrossRef] [PubMed]

16. Yasini, S.; Sistani, M.; Karimpour, A. Approximate Dynamic Programming for Two-player Zero-sum Game Related to H$_\infty$ Control of Unknown Nonlinear Continuous-time Systems. *Int. J. Control Autom. Syst.* 2015, 13, 99–109. [CrossRef]

17. Cheng, P.; Wang, H.; Stojanovic, V.; He, S.; Shi, K.; Luan, X.; Liu, F.; Sun, C. Asynchronous fault detection observer for 2-D Markov jump systems. *IEEE Trans. Cybern.* 2021, 1, 32021. [CrossRef] [PubMed]

18. Sun, J.; Liu, C.; Ye, Q. Robust differential game guidance laws design for uncertain interceptor-target engagement via adaptive dynamic programming. *Int. J. Control* 2017, 5, 990–1004. [CrossRef]

19. Xie, H.; Wu, B.; Liu, W. Adaptive Neural Network Model-based Event-triggered Attitude Tracking Control for Spacecraft. *Int. J. Control Autom. Syst.* 2021, 19, 172–185. [CrossRef]

20. Wang, D.; Liu, D.; Li, H.; Ma, H. Neural-network-based robust optimal control design for a class of uncertain nonlinear systems via adaptive dynamic programming. *Inf. Sci.* 2014, 28, 167–179. [CrossRef]

21. Xu, H. Finite-horizon near optimal design of nonlinear two-player zero-sum game in presence of completely unknown dynamics. *J. Control Autom. Electr. Syst.* 2015, 36, 361–370. [CrossRef]