

Optimizing Space Debris Cleanup: A Nonlinear Orbital Gaming Approach

Rugang Tang^{1,2}, Jiaqi Fan³, Chengfeng Luo^{2,4}, Xin Ning^{2,†} and Chih-Yung Wen¹

¹Department of Aeronautical and Aviation Engineering, The Hong Kong Polytechnic University. Hong Kong

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

³Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University. Hong Kong

⁴College of Computing and Data Science, Nanyang Technological University. Singapore

rg-polyu.tang@connect.polyu.hk · jiaqi.fan@connect.polyu.hk

lcf_nwpu@163.com · nx_nwpu@163.com · cywen@polyu.edu.hk

[†]Corresponding author

Abstract

The increasing accumulation of space debris presents a critical challenge to the sustainability and safety of space operations, necessitating the development of effective mitigation strategies. Orbital gaming offers a strategic framework for planning and executing debris removal missions, wherein spacecraft must operate efficiently to optimize debris clearance while minimizing resource expenditure. This study employs a nonlinear model of spacecraft motion, which provides a more accurate representation of general orbital relative motion compared to the Tschauner-Hempel (TH) or Clohessy-Wiltshire (CW) equations, as it accounts for nonlinear terms and reduces modeling errors.

To address the constraints imposed by fuel consumption in space missions, an objective function is formulated to balance fuel usage and target proximity. Furthermore, to overcome challenges associated with nonlinear dynamics, a single-network adaptive dynamic programming (ADP) method is proposed to derive the optimal control strategy for spacecraft navigation. The convergence of the algorithm is rigorously established, demonstrating its capability for online learning and rapid convergence. This enables the efficient computation in multi-body orbital game scenarios.

The effectiveness of the proposed algorithm is validated through extensive numerical simulations, which highlight its ability to adaptively optimize spacecraft trajectories in real-time, effectively balancing fuel efficiency and mission objectives. These simulations underscore the algorithm's robustness and scalability, showcasing its potential to significantly enhance the operational efficiency of space debris removal missions. This work not only advances the theoretical understanding of orbital dynamics and control but also provides a practical and scalable solution to the pressing issue of space debris management, thereby contributing to safer and more sustainable space exploration.

1. Introduction

1.1 Background and Motivation

The proliferation of space debris in Earth's orbit has emerged as one of the most critical challenges confronting the global space community.¹⁻⁴ Since the inception of the space age, thousands of satellites, launch vehicle upper stages, and mission-related objects have been deployed into orbit. Over time, collisions, explosions, and fragmentation events have generated millions of debris fragments, ranging from large defunct satellites to minuscule paint flecks. The majority of these objects reside in low Earth orbit (LEO), where the density of operational satellites and debris is highest. Even the smallest fragments, traveling at velocities exceeding 7 km/s, possess sufficient kinetic energy to cause catastrophic damage to active spacecraft, thereby threatening both crewed and uncrewed missions.

The escalating congestion of the orbital environment has led to an increased risk of collision, which can, in turn, trigger cascading fragmentation events—a phenomenon known as the Kessler Syndrome.^{5,6} In this scenario, each collision generates additional debris, further elevating the probability of subsequent collisions and potentially rendering entire orbital regions unusable for decades or even centuries. The ramifications of such a runaway process would be severe, jeopardizing not only scientific and commercial space activities but also critical infrastructure supporting navigation, communications, weather forecasting, and Earth observation.

SHORT PAPER TITLE

Recognizing the gravity of this situation, space agencies and international organizations have established guidelines for debris mitigation, such as post-mission disposal and passivation measures. However, these passive strategies alone are insufficient to curb the exponential growth of debris, particularly as the number of launches and mega-constellations continues to rise. Active debris removal (ADR) has thus been identified as an essential component of a comprehensive solution to preserve the long-term sustainability of space operations.^{7,8} The development of efficient, reliable, and cost-effective ADR technologies is now a top priority, necessitating advances in orbital mechanics, autonomous guidance and control, and mission planning.

In this context, the design of optimal strategies for debris removal missions is of paramount importance. Such strategies must account for the nonlinear and uncertain nature of orbital dynamics, the limited resources available to spacecraft (notably fuel), and the potential for adversarial or unpredictable behavior by non-cooperative targets. The integration of game-theoretic approaches and advanced control algorithms offers a promising avenue for addressing these challenges, enabling the formulation of robust and adaptive solutions capable of operating effectively in the complex and dynamic space environment.

1.2 Literature Review

Over the past decades, a variety of methods have been proposed for space debris removal, including mechanical capture (e.g., robotic arms, nets, harpoons), contactless techniques (e.g., laser ablation, ion beams), and drag augmentation devices.^{9–11} While these approaches have demonstrated varying degrees of technical feasibility, they often face significant limitations related to mission complexity, cost, reliability, and scalability. For instance, mechanical capture methods require precise rendezvous and docking maneuvers, which are challenging in the presence of tumbling or non-cooperative debris. Contactless methods, on the other hand, may be constrained by power requirements and effectiveness at different altitudes.

Concurrently, substantial research has focused on the modeling and control of spacecraft relative motion.^{12–14} Classical approaches such as the Clohessy-Wiltshire (CW) and Tschauner-Hempel (TH) equations provide linearized descriptions of relative dynamics, suitable for near-circular orbits and small separations. However, these models become increasingly inaccurate for larger distances, high-eccentricity orbits, or aggressive maneuvers, necessitating the adoption of nonlinear models that capture the full complexity of orbital motion. Recent advances in nonlinear control, adaptive dynamic programming (ADP), and reinforcement learning have enabled the development of more flexible and robust guidance strategies.

Game-theoretic frameworks have also gained traction in the context of space operations, particularly for scenarios involving multiple agents with potentially conflicting objectives. Orbital pursuit-evasion games, for example, model the interaction between a chaser spacecraft and a maneuvering target, providing a foundation for the synthesis of optimal or near-optimal control policies. Despite these advances, significant gaps remain in the integration of nonlinear dynamics, resource constraints, and multi-agent interactions within a unified and computationally tractable framework. This research aims to address these gaps by leveraging a nonlinear orbital gaming approach and advanced ADP techniques.

1.3 Contributions and Paper Outline

The principal contributions of this work are as follows:

- **Nonlinear Orbital Gaming Framework:** This work formulates the space debris removal problem as a nonlinear orbital pursuit-evasion game, leveraging a high-fidelity nonlinear relative dynamics model that captures complex orbital effects beyond traditional linearized approaches.
- **Single-Network Adaptive Dynamic Programming (ADP):** A single-network ADP algorithm is developed for real-time, near-optimal control under nonlinear dynamics and input constraints, enabling efficient online learning and policy adaptation without requiring explicit system linearization.
- **Rigorous Stability and Practical Validation:** Theoretical analysis establishes the uniform ultimate boundedness (UUB) of the closed-loop system and neural network weights, while comprehensive numerical simulations demonstrate the algorithm's effectiveness, robustness, and scalability for practical space debris removal missions.

The remainder of the paper is organized as follows. Section 2 presents the problem formulation, including the nonlinear relative dynamics model, the orbital gaming framework, and the objective function design. Section 3 details the development of the optimal control strategy using single-network ADP and the value function approximation with online learning, along with stability analysis. Section 4 reports the results of numerical simulations and discusses the performance of the proposed method. Finally, Section 5 concludes the paper and outlines directions for future research.

2. Problem Formulation and Preliminaries

2.1 Nonlinear Relative Dynamics Model

This study adopts a nonlinear model of spacecraft relative motion to accurately capture the dynamics between a deputy satellite and a chief satellite. In this context, the chief satellite represents the space debris to be removed, while the deputy satellite is the active pursuer tasked with intercepting and cleaning up the debris. Unlike the TH and CW equations, which are based on linearized assumptions, the nonlinear model incorporates higher-order terms, thereby reducing modeling errors and providing a more precise representation of the orbital environment.¹⁵ The nonlinear relative dynamics are described as follows:

$$\begin{aligned}\ddot{x} &= 2\dot{f}\dot{y} + \ddot{f}y + f^2x - \frac{\mu(r+x)}{[(r+x)^2 + y^2 + z^2]^{\frac{3}{2}}} + \frac{\mu}{r^2} \\ \ddot{y} &= -2\dot{f}\dot{x} - \ddot{f}x + f^2y - \frac{\mu y}{[(r+x)^2 + y^2 + z^2]^{\frac{3}{2}}} \\ \ddot{z} &= -\frac{\mu z}{[(r+x)^2 + y^2 + z^2]^{\frac{3}{2}}}\end{aligned}\quad (1)$$

where $[x, y, z]^T$ denotes the relative position vector of the chaser (deputy satellite) with respect to the target (chief satellite, i.e., space debris) in the local-vertical-local-horizontal (LVLH) frame; μ is the Earth's gravitational parameter; r is the reference orbital radius; f is the true anomaly, which can be computed via Newton's iterative method by solving Kepler's equation as detailed in Algorithm 1; $\dot{f}$ and $\ddot{f}$ are its first and second derivatives, respectively.

Algorithm 1 Calculation of True Anomaly from Orbital Elements

Require: Time since periapsis t , gravitational parameter μ , semi-major axis a , eccentricity e ($0 \leq e < 1$), mean anomaly at epoch M_0 (default 0), tolerance tol (default 10^{-10})

- 1: Compute mean motion: $n \leftarrow \sqrt{\mu/a^3}$
 - 2: Compute mean anomaly: $M \leftarrow M_0 + n \cdot t$
 - 3: Initialize eccentric anomaly: $E \leftarrow M$
 - 4: **repeat**
 - 5: $\Delta \leftarrow E - e \cdot \sin(E) - M$
 - 6: $E \leftarrow E - \Delta / (1 - e \cdot \cos(E))$
 - 7: **until** $|\Delta| < tol$
 - 8: Compute true anomaly: $f \leftarrow 2 \cdot \text{atan2}(\sqrt{1+e} \cdot \sin(E/2), \sqrt{1-e} \cdot \cos(E/2))$
 - 9: **Return** f
-

For comparison, the linearized CW equations, which are valid for near-circular orbits and small relative distances, are given by:

$$\begin{aligned}\ddot{x} &= 2n\dot{y} + 3n^2x \\ \ddot{y} &= -2n\dot{x} \\ \ddot{z} &= -n^2z\end{aligned}\quad (2)$$

where $n = \sqrt{\mu/r^3}$ is the mean motion of the reference orbit.

Compared to the linearized CW equations, the nonlinear model offers several advantages for space debris removal missions:

- **Higher Accuracy:** The nonlinear model captures higher-order effects and remains valid across a broader range of orbital regimes, including high-eccentricity orbits and large relative distances, where linear models become inaccurate.
- **Reduced Modeling Error:** By incorporating nonlinear gravitational and kinematic terms, the model reduces approximation errors, resulting in more precise trajectory predictions and control.
- **Broader Applicability:** The nonlinear formulation accommodates aggressive maneuvers and complex mission scenarios that cannot be addressed by linear models, thus enhancing the robustness and reliability of debris removal strategies.

SHORT PAPER TITLE

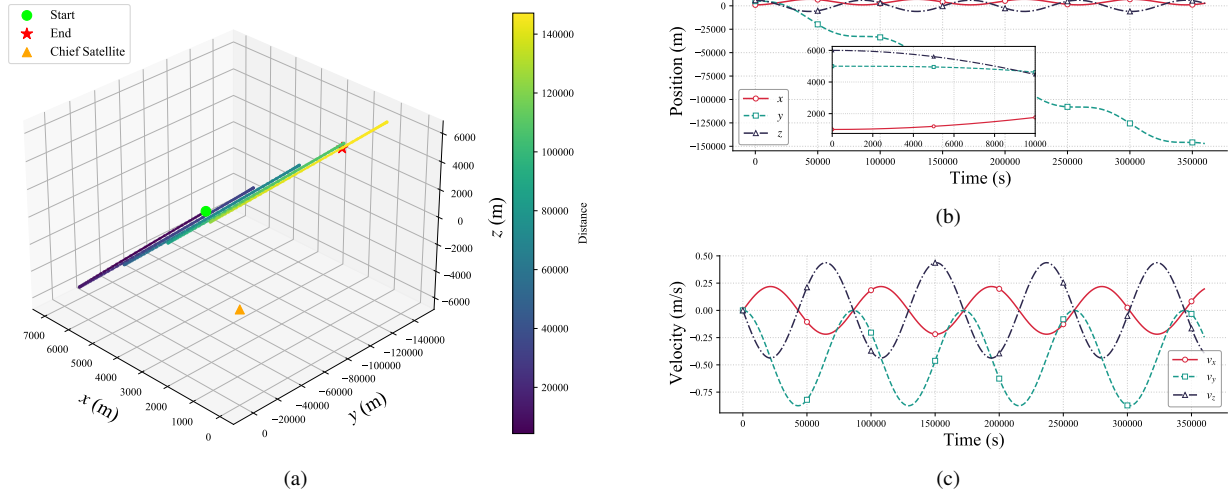


Figure 1: Comparison of trajectories using the CW equations. The initial orbital and state parameters for the chief (space debris) and deputy (chaser) satellites used in the simulations are as follows: the chief satellite is placed in an elliptical orbit with a semi-major axis of $a = 42,166.3$ km, eccentricity $e = 0.1$, and inclination $i = 45^\circ$. The deputy satellite is initialized with a relative position of $[x, y, z] = [1000, 5000, 6000]$ m and a relative velocity of $[\dot{x}, \dot{y}, \dot{z}] = [0, 0, 0]$ m/s in the LVLH frame. (a) Three-dimensional trajectory of the deputy satellite, (b) Relative position in the LVLH frame, and (c) Velocity components in the LVLH frame. Constants: $\mu = 3.986004418 \times 10^{14} \text{ m}^3/\text{s}^2$ (Earth's gravitational constant).

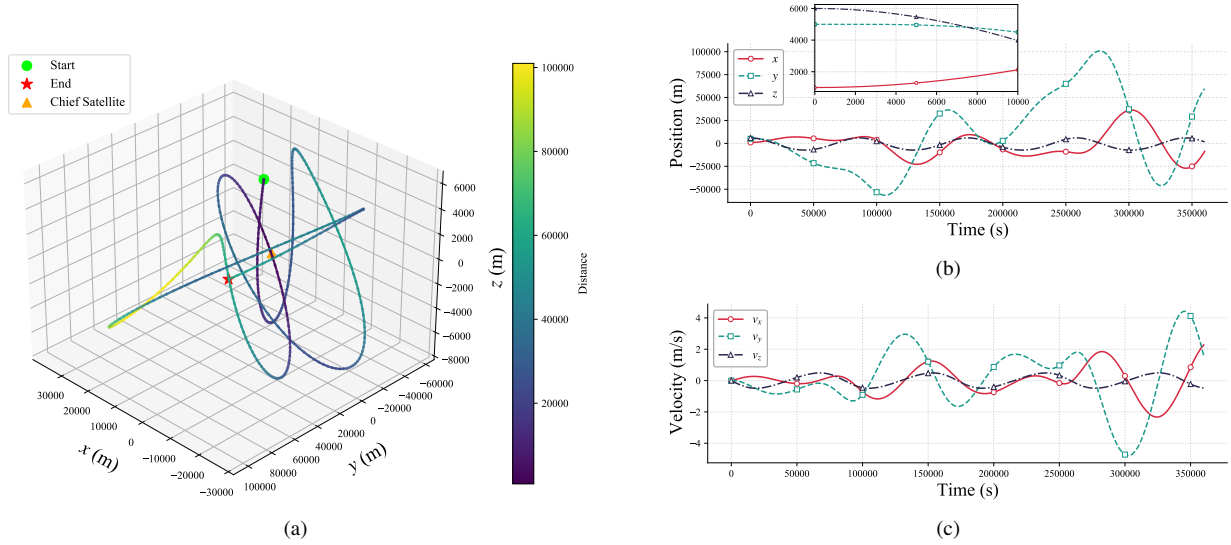


Figure 2: Comparison of trajectories using the nonlinear model under the same initial conditions as Figure 1. (a) Three-dimensional trajectory of the deputy satellite, (b) Relative position in the LVLH frame, and (c) Velocity components in the LVLH frame. The nonlinear model provides improved fidelity in capturing the true dynamics, particularly for larger separations and over extended time horizons.

The aforementioned advantages render the nonlinear model particularly well-suited for the development and optimization of advanced guidance and control algorithms in realistic space environments. Figures 1 and 2 provide a comparative illustration of the trajectories, relative positions, and velocities derived from the CW equations and the nonlinear model. As evidenced by these simulation results, while the CW equations initially exhibit trends similar to those of the nonlinear model, discrepancies accumulate over time due to linearization errors. For spacecraft operations, even minor deviations can result in substantially increased fuel consumption, given that fuel constitutes one of the most critical and limited resources in space missions. Consequently, the nonlinear model demonstrates superior fidelity in capturing the true dynamics of satellite motion, particularly in scenarios characterized by larger separations or complex maneuvers over extended durations.¹⁵ In light of these considerations, this study focuses on the nonlinear relative

dynamics model, while ensuring that computational efficiency is maintained within practical limits.

2.2 Orbital Gaming Framework for Debris Removal

Consider a satellite with continuous thrust orbiting the Earth.¹² Satellite P acts as the pursuer (deputy satellite), aiming to intercept the target satellite E (chief satellite, i.e., space debris), which may represent either space debris or a non-cooperative satellite. The satellites must account for both the interception objective and the minimization of energy consumption.

For analytical tractability and to facilitate the subsequent game-theoretic formulation, the nonlinear relative dynamics can be expressed in a compact state-space form:

$$\dot{X} = F(X) + g \cdot u \quad (3)$$

where $X = [x, y, z, \dot{x}, \dot{y}, \dot{z}]^T$ denotes the state vector of the deputy satellite in the LVLH frame, and $u = [u_x, u_y, u_z]^T$ represents the control input (thrust vector) of the deputy satellite. The function $F(X)$ encapsulates the nonlinear orbital dynamics as defined in (1):

$$F(X) = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ 2\dot{f}\dot{y} + \ddot{f}y + \dot{f}^2x - \frac{\mu(r+x)}{[(r+x)^2+y^2+z^2]^{3/2}} + \frac{\mu}{r^2} \\ -2\dot{f}\dot{x} - \ddot{f}x + \dot{f}^2y - \frac{\mu y}{[(r+x)^2+y^2+z^2]^{3/2}} \\ -\frac{\mu z}{[(r+x)^2+y^2+z^2]^{3/2}} \end{bmatrix} \quad (4)$$

and the input matrix g is given by

$$g = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (5)$$

This formulation facilitates the explicit integration of the pursuer's control actions into the system dynamics, thereby establishing a unified framework for subsequent optimal control design and game-theoretic analysis.

2.3 Objective Function Design

To address the dual objectives of fuel efficiency and target proximity, the following objective function is formulated:

$$J(X(t), u) \triangleq \int_0^\infty r(X(\tau), u(\tau)) d\tau \quad (6)$$

where τ is the time integration variable. To account for input constraints, we impose $u_k \leq u^{\text{sat}}$ for $k = x, y, z$, where u^{sat} denotes the respective saturation limits. The performance function $r(X(t), u(t))$ is defined as

$$r(X(t), u) \triangleq Q(X(t)) - 2 \int_0^u u^{\text{sat}} R_u \tanh^{-1}(\bar{u}) d\bar{u} \quad (7)$$

where $R_u \in \mathbb{R}^{3 \times 3}$ is a positive definite matrix and $\bar{u}$ is the integration variable. The function $Q(X(t)) : \mathbb{R}^6 \rightarrow \mathbb{R}$ is a positive function that satisfies the following assumption:

Assumption 1. (¹⁶) There exist constants $\underline{q}, \bar{q} \in \mathbb{R}_{>0}$ such that $\underline{q}\|X(t)\|^2 \leq Q(X(t)) \leq \bar{q}\|X(t)\|^2$ for all $X(t)$, $t \geq 0$.

This objective function explicitly incorporates fuel consumption constraints by penalizing excessive control usage, thereby ensuring that the resulting trajectory balances mission success with the resource limitations inherent in space operations.

3. Optimal Control Strategy using Single-Network Adaptive Dynamic Programming

In this section, optimal control strategies for both the pursuer satellite and the evader are established. The single-network adaptive dynamic programming (ADP) approach enables the derivation of near-optimal controllers under nonlinear orbital dynamics and input constraints. By leveraging value function approximation and online learning, the proposed method provides a computationally efficient solution for real-time trajectory optimization. The stability and convergence properties of the algorithm are rigorously analyzed, ensuring reliable performance in practical space debris removal scenarios.

3.1 Optimal Control Development

The value of the game stems from the mini-max problem. For simplicity of expression, t will be omitted below.

$$V(X) \triangleq \max_{u(\cdot)} \min_{v(\cdot)} J(X, u). \quad (8)$$

Define the Hamiltonian function as

$$H(X, u, V(X)) \triangleq \nabla^T V(X) (F(X) + u) + r(X, u). \quad (9)$$

The optimal control function for the given problem can be obtained by stationary condition $\nabla_u H(X, u, V(X)) = 0$, which leads to the following optimal control for the pursuer:

$$\begin{aligned} u^*(X) &\triangleq \arg \max H(X, u, V(X)) \\ &= -u^{\text{sat}} \tanh \left(\frac{1}{2u^{\text{sat}}} R_u^{-1} g^T \nabla V(X) \right). \end{aligned} \quad (10)$$

The optimal controllers derived in (10) necessitate explicit knowledge of the value function $V(X)$ and its gradient, which are typically intractable to obtain in closed form for nonlinear systems. Therefore, this work reformulates the problem as follows: Given the system dynamics in (3), the objective is to design approximately optimal controllers for both the pursuer and the evader by leveraging the objective function in (6) and employing advanced function approximation techniques.

3.2 Value Function Approximation via Single-Network ADP

According to the Weierstrass Approximation Theorem,¹⁷ there exists a complete set of independent basis functions that enables the uniform approximation of the value function and its gradient as follows:

$$V(X) = \sum_{j=1}^{\infty} w_j \varphi_j(X) = \sum_{j=1}^N w_j \varphi_j(X) + \sum_{j=N+1}^{\infty} w_j \varphi_j(X), \quad (11)$$

which can be equivalently expressed as

$$V(X) = W^T \phi(X) + \varepsilon(X), \quad (12)$$

where $W \in \mathbb{R}^N$ is the ideal weight vector, $\phi(X) = [\varphi_1(X), \varphi_2(X), \dots, \varphi_N(X)]^T$ is the vector of basis functions, and $\varepsilon(X)$ denotes the residual approximation error, which uniformly converges to zero as $N \rightarrow \infty$. The gradient of the value function is thus given by

$$\nabla V(X) = \nabla^T \phi(X) W + \nabla^T \varepsilon(X). \quad (13)$$

Let $\sigma = \nabla \phi[F(X) + g \cdot u]$. The residual error in the Hamiltonian function due to the approximation error ε is defined as

$$\varepsilon_H = W^T \sigma + r. \quad (14)$$

Let $\hat{W}$ denote the estimated weights of the neural network (NN), and define the estimation error as $\tilde{W} = W - \hat{W}$. The approximate optimal control strategy is then given by

$$\hat{u}(X) = u^{\text{sat}} \tanh \left(\frac{1}{2u^{\text{sat}}} R_u^{-1} g^T \nabla^T \phi(X) \hat{W} \right). \quad (15)$$

The residual error in the Hamiltonian function, arising from both the NN estimation error and the approximation error, is expressed as

$$\bar{\varepsilon}_H = \hat{W}^T \sigma + r. \quad (16)$$

The objective of the NN parameter update is to determine $\hat{W}$ that minimizes the squared residual error:

$$E = \frac{1}{2} \bar{\varepsilon}_H^\top \bar{\varepsilon}_H. \quad (17)$$

The following normalized gradient descent algorithm is adopted as the tuning law:

$$\dot{\hat{W}} = -\alpha \frac{\partial E}{\partial \hat{W}} = -\alpha \frac{\sigma}{(1 + \sigma^\top \sigma)^2} (\sigma^\top \hat{W} + r), \quad (18)$$

where $\alpha > 0$ is a positive learning rate.

Assumption 2. ⁽¹⁸⁾ The signal $\bar{\sigma} = \sigma/(\sigma^\top \sigma + 1)$ is persistently exciting over the interval $[t, t + T]$; that is, there exist constants $\beta_1 > 0$, $\beta_2 > 0$, and $T > 0$ such that, for all t ,

$$\beta_1 I \leq M_0 \equiv \int_t^{t+T} \bar{\sigma}(\tau) \bar{\sigma}^\top(\tau) d\tau \leq \beta_2 I. \quad (19)$$

Assumption 3. ^(16, 18, 19) The function $F(X)$ is bounded as $\|F(X)\| < b^F \|X\|$. Furthermore, the NN approximation error and its gradient, as well as the NN activation functions and their gradients, are bounded: $\|\varepsilon\| < b^\varepsilon$, $\|\nabla \varepsilon\| < b^{\nabla \varepsilon}$, $\|\phi(X)\| < b^\phi$, and $\|\nabla \phi(X)\| < b^{\nabla \phi}$.

It is noted that ε_H is bounded by b^{ε_H} if Assumption 3 holds.

3.3 Stability Analysis

This subsection establishes the stability properties of the proposed single-network ADP-based control scheme for the nonlinear orbital dynamics system. We begin by introducing the concept of uniform ultimate boundedness (UUB), followed by a key lemma and a main theorem that collectively guarantee the boundedness of both the system state and the neural network (NN) weight estimation errors under the proposed adaptive law.

Definition 1 (Uniform Ultimate Boundedness). The system $\dot{\mathbf{x}} = f(\mathbf{x})$ is said to be uniformly ultimately bounded (UUB) if there exists a compact set $S \subset \mathbb{R}^n$ such that for any initial condition $\mathbf{x}(0) \in S$, there exist constants $B > 0$ and $T = T(B, \mathbf{x}(0)) > 0$ satisfying $\|\mathbf{x}(t) - \mathbf{x}_e\| \leq B$ for all $t \geq T$, where $\mathbf{x}_e$ denotes the equilibrium point.

Lemma 1. ^(20–22) Consider the nonlinear system in (3). Suppose there exists a continuously differentiable, radially unbounded Lyapunov candidate function $V(X)$ such that

$$\dot{V} = \nabla^\top V(X) (F(X) + g \cdot u - g \cdot v) < 0. \quad (20)$$

Then, there exists a positive definite matrix $\Pi \in \mathbb{R}^{n \times n}$ such that

$$\nabla J^* \Pi \nabla V = Q(X) - 2 \int_0^u u^{\text{sat}} R_u \tanh^{-1}(\bar{u}) d\bar{u} + 2 \int_0^v v^{\text{sat}} R_v \tanh^{-1}(\bar{v}) d\bar{v}, \quad (21)$$

and

$$\nabla^\top V(X) (F(X) + g \cdot u - g \cdot v) = -\nabla V^\top \Pi \nabla V, \quad (22)$$

where $\Pi_{\min} \leq \|\Pi\| \leq \Pi_{\max}$ for bounded X .

Theorem 1. Consider the nonlinear system described by (1). Under Assumptions 1–3, if the control input $\hat{u}(X)$ in (15) is updated according to the NN tuning law in (18), then both the system state X and the NN weight estimation error $\tilde{W}$ are uniformly ultimately bounded (UUB).

Proof. Define the composite Lyapunov candidate function:

$$L = V(X) + \frac{1}{2} \tilde{W}^\top \alpha^{-1} \tilde{W}, \quad (23)$$

where $V(X)$ is the value function and $\tilde{W} = W - \hat{W}$ is the NN weight estimation error. The time derivative of L is

$$\dot{L} = \dot{V}(X) + \tilde{W}^\top \alpha^{-1} \dot{\tilde{W}}. \quad (24)$$

SHORT PAPER TITLE

Using the NN update law and the relationship $\bar{\varepsilon}_H = \varepsilon_H - \tilde{W}^\top \sigma$, we have

$$\dot{\tilde{W}} = \alpha \frac{\sigma}{(1 + \sigma^\top \sigma)^2} (\varepsilon_H - \tilde{W}^\top \sigma) = \alpha \Xi \varepsilon_H - \alpha \Xi \Xi^\top \tilde{W}, \quad (25)$$

where $\Xi = \frac{\sigma}{1 + \sigma^\top \sigma}$. Applying Young's inequality yields

$$\frac{\tilde{W}^\top \Xi \varepsilon_H}{1 + \sigma^\top \sigma} \leq \frac{1}{2} \tilde{W}^\top \Xi \Xi^\top \tilde{W} + \frac{1}{2} \|b^{\varepsilon_H}\|^2. \quad (26)$$

Substituting into the derivative of L and using Lemma 1, we obtain

$$\dot{L} \leq -\Pi_{\min} \|\nabla V\|^2 - \frac{1}{2} \lambda_{\min}(\Xi \Xi^\top) \|\tilde{W}\|^2 + \frac{1}{2} \|b^{\varepsilon_H}\|^2. \quad (27)$$

Therefore, $\dot{L} \leq 0$ whenever either

$$\|\nabla V\| \geq \frac{1}{\sqrt{2\Pi_{\min}}} \|b^{\varepsilon_H}\| \quad (28)$$

or

$$\|\tilde{W}\| \geq \frac{1}{\sqrt{\lambda_{\min}(\Xi \Xi^\top)}} \|b^{\varepsilon_H}\|. \quad (29)$$

This implies that both ∇V and $\tilde{W}$ are UUB. Since $V(X)$ is continuously differentiable and radially unbounded, the boundedness of ∇V ensures the boundedness of X , completing the proof. $\square$

4. Numerical Simulations and Validation

This section presents comprehensive numerical simulations to validate the effectiveness and robustness of the proposed single-network ADP approach for optimal control in space debris removal missions. The simulations are designed to demonstrate the algorithm's capability to handle nonlinear orbital dynamics, adapt to varying mission scenarios, and achieve efficient debris interception while minimizing control effort.

Table 1: Simulation Parameters and Settings

Parameter	Value
Chief satellite semi-major axis a	42,166.3 km
Chief satellite eccentricity e	0.1
Chief satellite inclination i	45°
Initial deputy position $[x_0, y_0, z_0]^\top$	$[100,000, 450,000, 0]^\top$ m
Initial deputy velocity $[\dot{x}_0, \dot{y}_0, \dot{z}_0]^\top$	$[0, 0, 0]^\top$ m/s
Control input bound u^{sat}	0.1 m/s ²
State weighting Q	diag(1, 1, 1, 1, 1, 1)
Control weighting R	diag(1, 1, 1)
Simulation duration t_{end}	36,000 s
Termination distance	10,000 m
NN dimension	20
NN learning rate (gain)	1
Initial NN weights W_0	$0.001 \times \mathbf{1}_{20}$

The nonlinear relative dynamics model described in (3) is employed, with simulation parameters and neural network (NN) configurations summarized in Table 1. All simulations are conducted in a Python 3.9 environment, utilizing the single-network ADP algorithm detailed in Section 3. The neural network basis functions are constructed using the state variables $x_1 = x$, $x_2 = y$, $x_3 = \dot{x}$, and $x_4 = \dot{y}$. The z -direction is omitted from the network since it is relatively independent and its initial value is set to zero in the simulations. The basis functions employed in the neural network include quadratic, cross, and quartic terms of the state variables, specifically:

$$\phi(X) = [x_1^2, x_2^2, x_3^2, x_4^2, x_1x_2, x_1x_3, x_1x_4, x_2x_3, x_2x_4, x_3x_4, x_1^4, x_2^4, x_3^4, x_4^4, x_1^2x_2^2, x_1^2x_3^2, x_1^2x_4^2, x_2^2x_3^2, x_2^2x_4^2, x_3^2x_4^2]^\top$$

where $X = [x_1, x_2, x_3, x_4]^T$. This selection of basis functions enables the neural network to approximate the value function and its gradient with sufficient flexibility for the nonlinear orbital dynamics considered in this study.

Figures 3 and 4 illustrate the efficacy of the proposed ADP-based control law. Figure 3 depicts the trajectories of both the deputy satellite and the target debris in the Geocentric Celestial Reference System (GCRS), demonstrating that the deputy successfully intercepts the target along a dynamically feasible path in the inertial frame. Figure 4 further presents the three-dimensional trajectory and its projection onto the x - y plane in the LVLH frame, highlighting the smooth and efficient interception maneuver achieved by balancing fuel efficiency and interception accuracy.

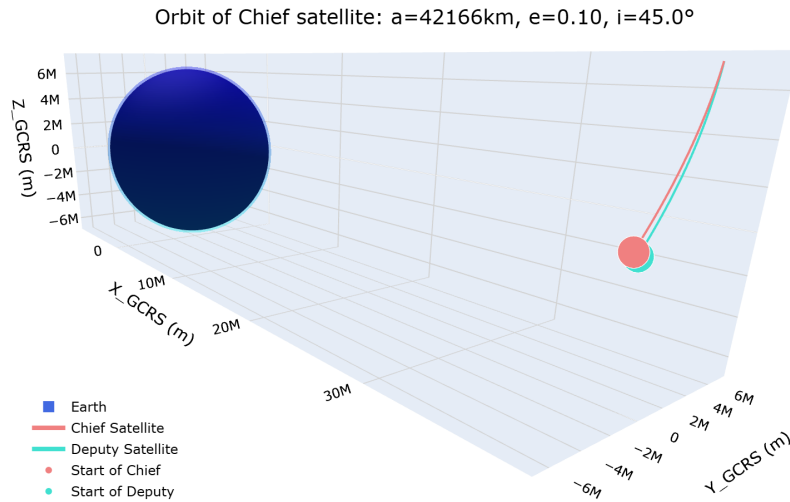


Figure 3: Trajectories of both the deputy satellite and the target debris in the Geocentric Celestial Reference System (GCRS).

Figure 5 provides the time histories of the relative position and velocity components in the LVLH frame. The position components exhibit monotonic convergence toward the target, confirming the effectiveness of the control strategy. It is noted that the velocity components do not fully converge within the simulation horizon, primarily due to the lower weighting assigned to velocity in the objective function relative to position. The selection of these weights can be tailored to specific mission requirements: increasing the velocity weighting enhances economic efficiency by minimizing fuel consumption, while prioritizing position weighting enables rapid interception. This flexibility allows the control strategy to be adapted for diverse operational objectives.

The evolution of the neural network weights is depicted in Figure 6. The weights converge rapidly and remain stable throughout the simulation, indicating that the learning process is both efficient and robust. This stability is crucial for ensuring consistent control performance during extended mission durations.

Figure 7 presents the control input profiles generated by the ADP controller. The control inputs remain within the prescribed bounds, and their smooth variation reflects the controller's ability to generate feasible and efficient thrust commands without inducing chattering or excessive actuation.

Collectively, the simulation results validate the capability of the proposed single-network ADP algorithm to achieve real-time, near-optimal control for nonlinear orbital pursuit-evasion scenarios. The algorithm demonstrates strong adaptability, rapid convergence, and robustness to nonlinearities, making it a promising approach for practical space debris removal missions. The ability to learn and update control policies online is particularly advantageous in dynamic and uncertain space environments, where mission parameters and environmental conditions may change unpredictably.

5. Conclusion and Future Work

This paper has presented a single-network adaptive dynamic programming (ADP) approach for optimal control in nonlinear orbital gaming scenarios, with a particular focus on space debris removal missions. The proposed method leverages a nonlinear relative dynamics model and advanced function approximation techniques to derive near-optimal control strategies for both pursuer and evader satellites. The stability and convergence properties of the algorithm are rigorously established, ensuring reliable performance in practical applications.

Numerical simulations demonstrate the effectiveness of the proposed ADP algorithm in achieving efficient interception of space debris while minimizing fuel consumption. The results highlight the algorithm's capability to handle

SHORT PAPER TITLE

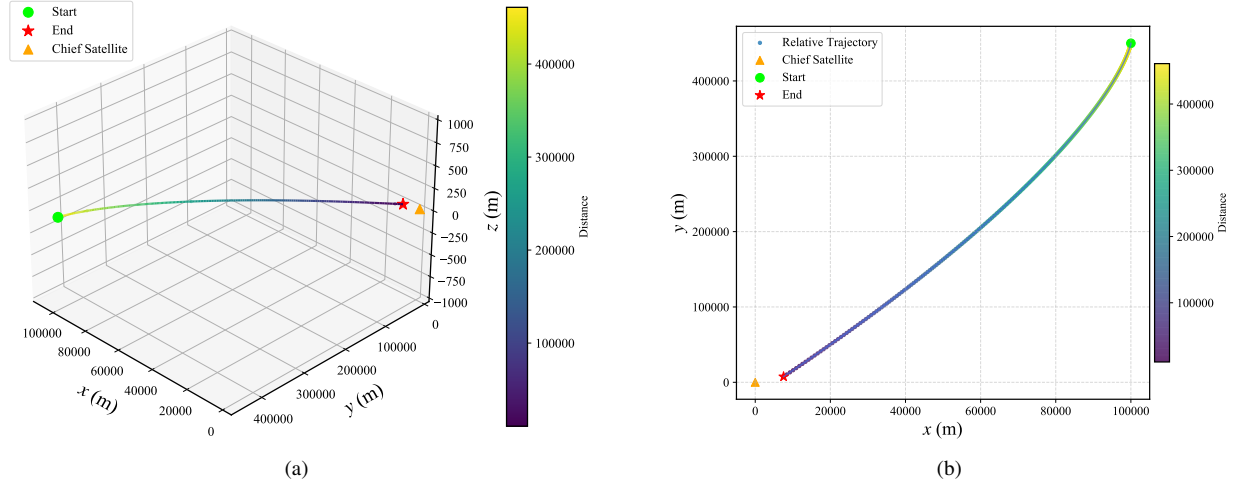


Figure 4: (a) Three-dimensional trajectory of the deputy satellite using ADP; (b) Relative position in the LVLH frame.

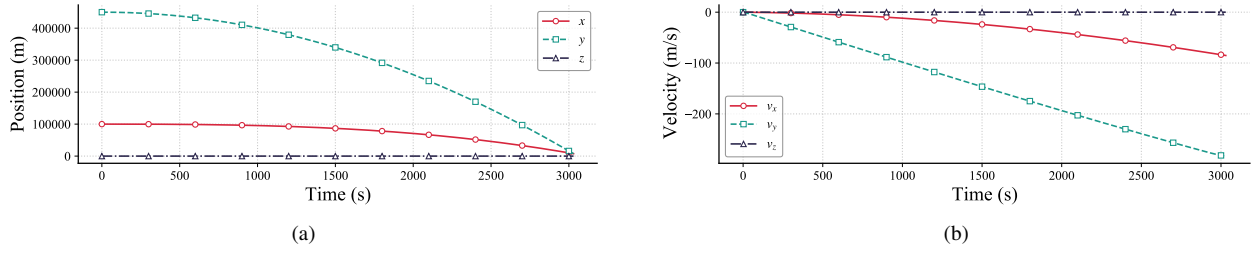


Figure 5: (a) Relative position in the LVLH frame; (b) Velocity components in the LVLH frame using ADP.

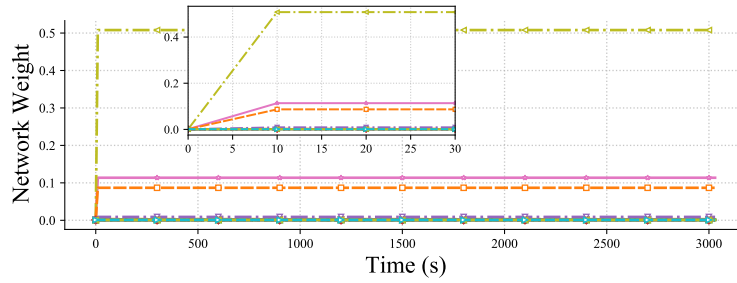


Figure 6: Neural network weights evolution using ADP.

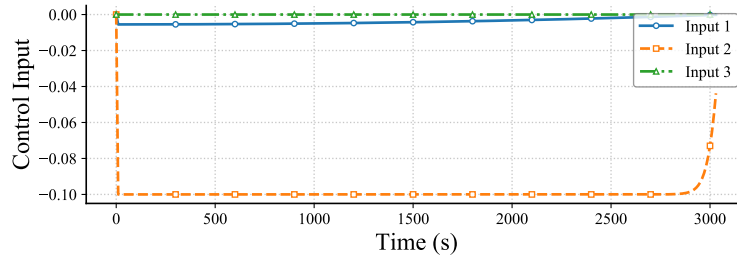


Figure 7: Control input profile using ADP.

nonlinear dynamics, thereby confirming its suitability for real-time applications in space operations. The use of a single neural network for value function approximation simplifies implementation and enhances computational efficiency, making the approach feasible for onboard deployment in space missions.

Future research will focus on extending the proposed framework to multi-satellite scenarios, where multiple

pursuers and evaders interact within complex orbital environments. Additionally, the integration of advanced machine learning techniques for adaptive learning and real-time decision-making will be explored to further enhance the robustness and scalability of the algorithm in dynamic space missions. The incorporation of uncertainty modeling and robust control strategies will also be investigated to improve the performance of the single-network ADP method in real-world applications.

6. Acknowledgments

This research was supported, in part by Research Center of Unmanned Autonomous Systems, Hong Kong Polytechnic University under Grant P0046487, in part by the Doctor Dissertation of Northwestern Polytechnical University under Grant CX2023037.

References

- [1] Robert G Ryan, Eloise A Marais, Chloe J Balhatchet, and Sebastian D Eastham. Impact of rocket launch and space debris air pollutant emissions on stratospheric ozone and global climate. *Earth's Future*, 10(6):e2021EF002612, 2022.
- [2] Omer F Keles. Telecommunications and space debris: Adaptive regulation beyond earth. *Telecommunications Policy*, 47(3):102517, 2023.
- [3] Francesca Letizia, Manuel Metz, Pascal Faucher, Holger Krag, Liu Jing, Maria Antonia Ramos, Noelia Sanchez Ortiz, Thomas Schildknecht, and Smiriti Srivastava. Improving the knowledge of the orbital population new technical means of space debris monitoring. *Acta Astronautica*, 223:734–740, 2024.
- [4] Fumihiro Hayashi, Arata Kioka, Takuma Ishii, and Takumu Nakamura. Unveiling the resource potential of space debris: A forecast of valuable metals to 2050. *Waste Management*, 193:376–385, 2025.
- [5] Donald J Kessler, Nicholas L Johnson, JC Liou, and Mark Matney. The kessler syndrome: implications to future space operations. *Advances in the Astronautical Sciences*, 137(8):2010, 2010.
- [6] Jakub Drmola and Tomas Hubik. Kessler syndrome: System dynamics model. *Space Policy*, 44:29–39, 2018.
- [7] J-C Liou, Nicholas L Johnson, and NM Hill. Controlling the growth of future leo debris populations with active debris removal. *Acta Astronautica*, 66(5-6):648–653, 2010.
- [8] Christophe Bonnal, Jean-Marc Ruault, and Marie-Christine Desjean. Active debris removal: Recent progress and current trends. *Acta Astronautica*, 85:51–60, 2013.
- [9] Alexander Ledkov and Vladimir Aslanov. Review of contact and contactless active space debris removal approaches. *Progress in Aerospace Sciences*, 134:100858, 2022.
- [10] Man Ru, Ying Zhan, Bin Cheng, and Yu Zhang. Capture dynamics and control of a flexible net for space debris removal. *Aerospace*, 9(6):299, 2022.
- [11] VV Svtina and MV Cherkasova. Space debris removal—review of technologies and techniques. flexible or virtual connection between space debris and service spacecraft. *Acta Astronautica*, 204:840–853, 2023.
- [12] Haoran Gong, Shengping Gong, and Junfeng Li. Pursuit–evasion game for satellites based on continuous thrust reachable domain. *IEEE Transactions on Aerospace and Electronic Systems*, 56(6):4626–4637, 2020.
- [13] Zhenyu Zhuo, Lan Du, Xiaofei Lu, Jian Chen, and Zhuowei Cao. Smoothed lv distribution based three-dimensional imaging for spinning space debris. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–13, 2022.
- [14] Chaoyu Mu, Shuo Liu, Ming Lu, Zhaoyang Liu, Lei Cui, and Ke Wang. Autonomous spacecraft collision avoidance with a variable number of space debris based on safe reinforcement learning. *Aerospace Science and Technology*, 149:109131, 2024.
- [15] Joshua Sullivan, Sebastian Grimberg, and Simone D’Amico. Comprehensive survey and assessment of spacecraft relative motion dynamics models. *Journal of Guidance, Control, and Dynamics*, 40(8):1837–1859, 2017.

SHORT PAPER TITLE

- [16] Patryk Deptula, Hsi-Yuan Chen, Ryan A Licitra, Joel A Rosenfeld, and Warren E Dixon. Approximate optimal motion planning to avoid unknown moving avoidance regions. *IEEE Transactions on Robotics*, 36(2):414–430, 2019.
- [17] Murad Abu-Khalaf and Frank L Lewis. Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network hjb approach. *Automatica*, 41(5):779–791, 2005.
- [18] Kyriakos G Vamvoudakis and Frank L Lewis. Online actor–critic algorithm to solve the continuous-time infinite horizon optimal control problem. *Automatica*, 46(5):878–888, 2010.
- [19] Kyriakos G Vamvoudakis, Frank L Lewis, and Greg R Hudas. Multi-agent differential graphical games: Online adaptive learning solution for synchronization with optimality. *Automatica*, 48(8):1598–1611, 2012.
- [20] Huaguang Zhang, Lili Cui, and Yanhong Luo. Near-optimal control for nonzero-sum differential games of continuous-time nonlinear systems using single-network adp. *IEEE Transactions on Cybernetics*, 43(1):206–216, 2012.
- [21] Majid Mazouchi, Mohammad Bagher Naghibi-Sistani, and Seyed Kamal Hosseini Sani. A novel distributed optimal adaptive control algorithm for nonlinear multi-agent differential graphical games. *IEEE/CAA Journal of Automatica Sinica*, 5(1):331–341, 2017.
- [22] Haoming Zou and Guoshan Zhang. Dynamic event-triggered-based single-network adp optimal tracking control for the unknown nonlinear system with constrained input. *Neurocomputing*, 518:294–307, 2023.