Differential Game Theoretic Control for Satellite Rendezvous and Docking

Abinay J. Brown
Georgia Institute of Technology, Atlanta, Georgia, 30332

This research demonstrates the application of differential game theory for satellite rendezvous, docking, and formation flying. This project employs the two-player Zero-Sum and Non-Zero-Sum game for rendezvous and docking in the presence of orbit perturbations as well as for cooperative docking. The Zero-Sum approach models the players as chaser control input and the worst-case perturbations and solved using the Game-Algebraic-Riccati-Equation (GARE) and a Synchronous Online Learning method. For the Non-Zero-Sum game case, the players are the chaser and target spacecraft that cooperate to perform docking while minimizing their individual control efforts. Simulations indicate robust performance in regulation and tracking for docking and formation flying, with the Zero-Sum GARE approach handling worst-case scenarios effectively. The Synchronous Online Learning method, while successful for docking, required significant parameter tuning and struggled with tracking. Non-Zero-Sum scenarios showcased successful cooperative docking under perturbations. The results validate the differential game-theoretic approach for realistic autonomous space maneuvers. Future work includes refining the Synchronous Online Learning approach and extending these methods to multi-satellite systems for enhanced formation flying and distributed sensing.

I. Nomenclature

x	=	relative distance in x direction
y	=	relative distance in y direction
z	=	relative distance in z direction
μ	=	Earth's gravitational parameter
ω	=	orbit angular velocity
ω_c	=	orbit angular velocity of chaser
i_c	=	Orbit inclination
J_2	=	Earth oblateness perturbation coefficient
c	=	Sedgwick-Schweighart coefficient
s	=	Sedgwick-Schweighart coefficient
k	=	Sedgwick-Schweighart coefficient
f	=	system of dynamic equations
g	=	control input or disturbance matrix
Q	=	State penalty weight matrix
R	=	Control input weight matrix
H	=	Hamiltonian Function
V	=	Value or Cost function
u	=	Control input acceleration
d	=	Disturbance input acceleration
γ	=	H-infinity gain
ϕ	=	basis functions vector
W	=	Neural-Network Weights
a	=	gradient descent rates
F	=	update law weights

II. Introduction

Rendezvous and docking are critical phases of space missions that require two spacecraft—a target and a chaser—to be placed in the same orbit and perform orbit phasing maneuvers such that their position and velocity vectors nearly match; the relative position and velocity between them are effectively zero. An extension of this concept is maintaining a specific relative distance or angular phase in the same orbit, which defines spacecraft formation flying. The purposes of rendezvous and docking include in-space spacecraft assembly and servicing, space-station crew transportation, and planetary return missions, while formation flying finds applications in Earth observation, communications, and navigation.

A key challenge in rendezvous and docking missions is performing orbit phasing maneuvers accurately in the presence of perturbing forces such as J2 perturbations, atmospheric drag, solar radiation pressure, and third-body gravitational influences. Many existing approaches simplify the dynamics into linear problems, often neglecting these perturbing forces to facilitate computational feasibility [1].

In this research project, we adopt a differential game-theoretic approach [2] to address the rendezvous and docking problem. In the first part, we assume a Zero-Sum differential game, utilizing an H-infinity-based control scheme to derive the optimal control input for the chaser satellite under worst-case perturbations by solving the GARE (Game Algebraic Riccati Equation) as well a synchronous online adaptive solution. In the second part, we assume a two-player Non-Zero-Sum cooperative differential game, where both spacecraft are collaborating directly to perform docking while minimizing their own control efforts.

III. Methodology

A. Problem Setup

For relative satellite motion, the Clohessy-Wiltshire equations [3] are popular due to its simplicity and assumptions made for first hand approximation. But in this project, the Sedgwick-Schweighart relative equations of motion are considered over the standard CW equations as it is more accurate as it models the effects of J2 perturbation about a given orbit [4].

1. Equations of Motion

The Sedgwick-Schweighart equations 1, 2, and 3 are linear once the coefficients are substituted in from 4 to 6. The coordinate frame is fixed at the Local-Vertical-Local-Horizontal (LVLH) Frame of the Target Spacecraft.

$$\ddot{x} - 2(\omega c)\dot{y} - (5c^2 - 2)\omega^2 x = u_x, \tag{1}$$

$$\ddot{y} + 2(\omega c)\dot{x} = u_y, \tag{2}$$

$$\ddot{z} + k^2 z = u_z, \tag{3}$$

$$c = \sqrt{1+s}, \tag{4}$$

$$s = \frac{2J_2R_e^2}{8r_c^2}\left[1 + 3\cos(2i_c)\right], \tag{5}$$

$$k = \omega_{orbC}\sqrt{1+s} + \frac{3J_2\omega_{orb}R_e^2}{2r_c^2}\left[\cos(i_c)\right]. \tag{6}$$

By decomposing the above equations 1, 2, and 3, into a system of first order differential equations, we should get 6 equations and 6 state variables; the relative xyz position and velocity.

2. Two-player differential game

Firstly, the Zero-Sum game scenario is considered, the two players are the chaser control thrust u and the perturbing forces d. The cost or the value function for the Zero-Sum case is given in 9 [5], where the optimal cost is finding the min-max solution; where the u player is minimizing the control input while the d player is maximizing the disturbance. The equations of motion are rewritten in the standard form below, where state $x(t) \in \mathbb{R}^n$, output $y(t) \in \mathbb{R}^p$, and smooth function $f(x(t)) \in \mathbb{R}^n$, $g(x) \in \mathbb{R}^{n\times m}$, $k(x) \in \mathbb{R}^{n\times q}$, $h(x) \in \mathbb{R}^{p\times n}$. This system has two inputs, the control input $u(x(t)) \in \mathbb{R}^m$ and the disturbance input $d(x(t)) \in \mathbb{R}^q$.

$$\dot{x} = f(x) + g(x)u(x) + k(x)d(x), \tag{7}$$

$$y = C(x). \tag{8}$$

$$V(x(0), u, d) = \int_0^\infty \left(Q(x) + u^T Ru - \gamma^2 \|d\|^2\right) dt \tag{9}$$

$$V^* = \min_u \max_d \int_0^\infty \left(Q(x) + u^T Ru - \gamma^2 \|d\|^2\right) dt. \tag{10}$$

Secondly, the Non-Zero-Sum game scenario is considered, where the two players are the target and chaser spacecraft control inputs u_1 and u_2 that are cooperating to achieve docking while minimizing their own control efforts. Here, the perturbing forces are incorporated into the equation of motion as nonlinear terms instead of a separate disturbance player like in the zero-sum case. The cost or value function for this case is given below.

$$\dot{x} = f(x) + g_1(x)u_2(x) + g_2(x)u_2(x), \tag{11}$$

$$y = C(x). \tag{12}$$

$$V_i(x(0), u_1, u_2) = \int_0^\infty \left(Q_i(x) + \sum_{j=1}^{N} u_j^T R_{ij} u_j\right) dt, i \in \{1, 2\} \tag{13}$$

$$V_i^*(x(t), u_1, u_2) = \min_{u_i} \int_t^\infty \left(Q_i(x) + \sum_{j=1}^{2} u_j^T R_{ij} u_j\right) d\tau, \quad i \in \{1, 2\}. \tag{14}$$

B. Zero-Sum Differential Game

1. Game-Algebraic-Riccati-Equation

Since the SS dynamics considered are linear and can be written in the form of 15 and 16, we can reformulate the traditional Algebraic Riccati Equation for the Differential Game case given in 18.

$$\dot{x} = Ax + B_1 u + B_2 d \tag{15}$$

$$z = \begin{bmatrix} Cx \\ Du \end{bmatrix} \tag{16}$$

$$V(x(t), u, d) = \frac{1}{2} \int_t^\infty \left(x^T C^T Cx + u^T D^T Du - \gamma^2 \|d\|^2\right) d\tau. \tag{17}$$

$$A^T P + PA + C^T C + \gamma^{-2} PB_1 B_1^T P - PB_2 (D^T D)^{-1} B_2^T P = 0. \tag{18}$$

The GARE equation can be solved using the Kleinman algorithm. First a $P^{(i)}$ is selected such that $\tilde{A}$ in 19 has eigenvalues in the left half plane, then iterated by solving the Lyapunov Equation in 20 for $P^{(i+1)}$ until P converges [6].

$$\tilde{A}_i = A - B_2 (D^T D)^{-1} B_2^T P^{(i)}. \tag{19}$$

$$\tilde{A}_i^T P^{(i+1)} + P^{(i+1)} \tilde{A}_i + C^T C + P^{(i)} B_2 (D^T D)^{-1} B_2^T P^{(i)} = 0 \tag{20}$$

Then the optimal control and disturbance input equations are found by finding the partial derivative of the Hamiltonian from the cost function in 17 and applying stationary conditions to get the equations 21 and 22.

$$u^*(x) = -(D^T D)^{-1} B_1^T P \tag{21}$$

$$d^*(x) = \frac{1}{\gamma^2} B_2^T P x \tag{22}$$

It is important to note that the SS dynamics and therefore 19 by default have two eigenvalues at the origin by nature indicating neutral stability so this would make the lyapunov equation unsolvable, so it is important to subtract a very small regularization term ϵ from 19 at each iteration to move the 0 poles slightly to the left to ensure the lyapunov equation can be solved for.

2. Synchronous Online Learning

An alternative approach is to solve the problem by approximating the Value function as a Neural Networks. This would require three networks an actor, critic, and disturbance network [7], where the weights are learned online using gradient descent update equations. First we can write the Hamiltonian of the equation as given in equation 23 and then take the partial with respect to control and disturbance variables and then apply stationary conditions to determine the contol input and disturbance equations given in equations 24 and 25 [8].

$$H = Q(x) + u^T R u - \gamma^2 ||d||^2 + \nabla V\ (f(x) + g(x)u + k(x)d) = e_1 \tag{23}$$

$$\frac{\partial H}{\partial u} = 0 \implies u = -\frac{1}{2} R^{-1} g^T(x) \nabla V_1(x) \tag{24}$$

$$\frac{\partial H}{\partial d} = 0 \implies d = \frac{1}{2\gamma^2} k^T(x) \nabla V_1(x) \tag{25}$$

We can then approximate the value functions as a NN as given in 26, now plugging into 27 and 28 we get the control laws. Here, we maintain three sets of weight W_1 as the critic, W_2 as the actor weights, and W_3 as the disturbance weights.

$$\widehat{V}_1(x) = \widehat{W}_1^T \phi_1(x) \tag{26}$$

$$u = -\frac{1}{2} R^{-1} g^T(x) \nabla \phi_1^T(x) W_2 \tag{27}$$

$$d = \frac{1}{2\gamma^2} k^T(x) \nabla \phi_1^T(x) W_3 \tag{28}$$

The three weights are updated online using equations 29, 30, 31.

$$\dot{W}_1 = -a_1 \frac{\sigma_2}{\left(\sigma_2^T \sigma_2 + 1\right)^2} \left[\sigma_2^T \hat{W}_1 + Q(x) - \gamma^2 \|\hat{d}\|^2 + \hat{u}^T R \hat{u}\right] \tag{29}$$

$$\dot{W}_2 = -a_2 \left\{ F_2 \hat{W}_2 - F_2 \bar{\sigma_2}^T \hat{W}_1 - \frac{1}{4} \bar{D}_1(x) \hat{W}_2 m^T(x) \hat{W}_1 \right\} \tag{30}$$

$$\dot{W}_3 = -a_3 \left\{ F_4 \hat{W}_3 - F_3 \bar{\sigma_2}^T \hat{W}_1 + \frac{1}{4\gamma^2} \bar{E}_1(x) \hat{W}_3 m^T(x) \hat{W}_1 \right\} \tag{31}$$

Where $\bar{D}_1(x) \equiv \nabla\phi_1(x) g(x) R^{-1} g^T(x) \nabla\phi_1^T(x)$, $\bar{E}_1(x) \equiv \nabla\phi_1(x) k k^T \nabla\phi_1^T(x)$, $m \equiv \frac{\sigma_2}{\left(\sigma_2^T \sigma_2+1\right)^2}$, $F_1 > 0,\ F_2 > 0,\ F_3 > 0,\ F_4 > 0$ are tuning parameters. $\bar{\sigma}_2 = \frac{\sigma_2}{\sigma_2^T \sigma_2 + 1}$. Where F_1, F_2, F_3, F_4 are tuning parameters, and $a_1, a_2, anda_3$ are update rates. For this simulation, the basis functions are selected as the kronecker product of the state vector x with itself: $\vec{x} \bigotimes \vec{x}$, where the non-repeating polynomial elements are stacked to form the basis vector ϕ. Since this state vector contains six elements, x, y, z position and velocity, this results in a 21 element basis vector ϕ. Additionally, persistence of excitation noise is added to the control input for sufficient exploration of weights that best approximate the value function.

C. Non-Zero-Sum Differential Game

1. Synchronous Online Learning

The approach here is similar to that of the Zero-Sum case, with some modifications to the formulation. First we write the hamiltonians in 32 and 33 and derive the control inputs in 34 and 35 by applying stationary conditions to the partial derivatives w.r.t the control inputs [7].

$$H_1 = Q_1(x) + u_1^T R_{11} u_1 + d_2^T R_{12} d_2 + \nabla V_1 \left(f(x) + g(x)u_1 + k(x)d_2 \right) = e_1 \tag{32}$$

$$H_2 = Q_2(x) + u_1^T R_{21} u_1 + d_2^T R_{22} d_2 + \nabla V_2 \left(f(x) + g(x)u_1 + k(x)d_2 \right) = e_2 \tag{33}$$

$$\frac{\partial H_1}{\partial u_1} = 0 \implies u_1(x) = -\frac{1}{2} R_{11}^{-1} g^T(x) \nabla V_1 \tag{34}$$

$$\frac{\partial H_2}{\partial u_2} = 0 \implies d_2(x) = -\frac{1}{2} R_{22}^{-1} k^T(x) \nabla V_2 \tag{35}$$

Then we approximate the Value functions as NN with basis function vector ϕ and weight $\widehat{W}$ given in 36 and 37. Here we maintain four NN weights, two critics $\widehat{W}_1$ and $\widehat{W}_2$ and two actors $\widehat{W}_3$ and $\widehat{W}_4$, a critic and actor weights for each player.

$$\widehat{V}_1(x) = \widehat{W}_1^T \phi_1(x), \tag{36}$$

$$\widehat{V}_2(x) = \widehat{W}_2^T \phi_2(x). \tag{37}$$

We apply the control inputs based on the actor weights given in 38 and 39 and continually update the critic and actor weights in real time.

$$u_3(x) = -\frac{1}{2} R_{11}^{-1} g^T(x) \nabla \phi_1^T \widehat{W}_3, \tag{38}$$

$$d_4(x) = -\frac{1}{2} R_{22}^{-1} k^T(x) \nabla \phi_2^T \widehat{W}_4. \tag{39}$$

The update laws for the actor and critic weights are derived by finding the gradient solution to $E = \frac{1}{2} e_1^T e_1 + \frac{1}{2} e_2^T e_2$ and are given below in 40, 41, 42, and 43.

$$\dot{\widehat{W}}_1 = -a_1 \frac{\sigma_3}{(\sigma_3^T \sigma_3 + 1)^2} \left[\sigma_3^T \widehat{W}_1 + Q_1(x) + u_3^T R_{11} u_3 + d_4^T R_{12} d_4 \right] \tag{40}$$

$$\dot{\widehat{W}}_2 = -a_2 \frac{\sigma_4}{(\sigma_4^T \sigma_4 + 1)^2} \left[\sigma_4^T \widehat{W}_2 + Q_2(x) + u_3^T R_{21} u_3 + d_4^T R_{22} d_4 \right] \tag{41}$$

where $\sigma_3 = \nabla \phi_1 (f + g u_3 + k d_4)$ and $\sigma_4 = \nabla \phi_2 (f + g u_3 + k d_4)$.

$$\dot{\widehat{W}}_3 = -a_3 \left\{ (F_2 \widehat{W}_3 - F_1 \sigma_3^T \widehat{W}_1) - \frac{1}{4} \left(\nabla \phi_1 g(x) R_{11}^{-T} R_{21} R_{11}^{-1} g^T(x) \nabla \phi_1^T \widehat{W}_3 m_2^T \widehat{W}_2 + \bar{D}_1(x) \widehat{W}_3 m_1^T \widehat{W}_1 \right) \right\}. \tag{42}$$

$$\dot{\widehat{W}}_4 = -a_4 \left\{ (F_4 \widehat{W}_4 - F_3 \sigma_4^T \widehat{W}_2) - \frac{1}{4} \left(\nabla \phi_2 k(x) R_{22}^{-1} R_{12} R_{22}^{-1} k^T(x) \nabla \phi_2^T \widehat{W}_4 m_4^T \widehat{W}_1 + \bar{D}_2(x) \widehat{W}_4 m_2^T \widehat{W}_2 \right) \right\}. \tag{43}$$

$$\bar{D}_1(x) \equiv \nabla \phi_1(x) g(x) R_{11}^{-1} g^T(x) \nabla \phi_1^T(x), \quad \bar{D}_2(x) \equiv \nabla \phi_2(x) k R_{22}^{-1} k^T \nabla \phi_2^T(x), \tag{44}$$

$$m_1 \equiv \sigma_3 / (\sigma_3^T \sigma_3 + 1)^2, \quad m_2 \equiv \sigma_4 / (\sigma_4^T \sigma_4 + 1)^2, \tag{45}$$

$$\bar{\sigma}_3 = \sigma_3 / (\sigma_3^T \sigma_3 + 1), \quad \bar{\sigma}_4 = \sigma_4 / (\sigma_4^T \sigma_4 + 1) \tag{46}$$

$F_1 > 0, F_2 > 0, F_3 > 0, F_4 > 0$ are tuning parameters, as well as a_1, a_2, a_3, and a_4 which are the gradient descent update rates. The proofs of the update laws are presented in. It is important that Persistence of Excitation noise is added to the control inputs u_3 and d_4 for sufficient exploration of the weights that approximate the value function. Often times, the weights diverge, if the noise is insufficient as it is unable to find the best weights for the control inputs. In this simulation case, the P.E. noise is added to the control inputs for the duration of an entire orbit, so that there's sufficient exploration of the weights that can regulate the system, and then the noise is turned off.

Here, the basis functions are selected as the list of unique elements from the Kronecker Product of the state vector with itself $\vec{x} \bigotimes \vec{x}$. The rationale being that polynomials produced are every possible permutation of the elements of the state vector. This results in 21 basis functions for the NN approximation.

D. Simulation

In this project, four simulation cases are tested. The first is the Zero-Sum case solved using GARE as a regulation problem to testing docking (relative position and velocity converge to zero), then used again for a tracking problem to test formation flying, by using an integral action setpoint tracking to maintain a separation distance between two satellites, three x, y, and z error states are appended to the original state-space to convert to a tracking problem. Next, the Zero-Sum game case is solved using the Online Synchronous approach as a regulation problem, to test docking maneuver. Finally, the Non-Zero-Sum game case is solved similarly to test cooperative docking. For all of the simulation cases, we are assuming two satellites in a Sun-Synchronous Orbit with a separation distance of 100m and velocity of 1m/s in the x, y, and z and about the LVLH frame of the Target satellite. The parameters used in the simulation is listed in the table 1 below.

Table 1 Simulation Parameters

Parameter	Value
μ	$3.986 * 10^{14} \frac{km^3}{s^2}$
J_2	$1.08263 * 10^{-3}$
R_e	6378 km
i	98^o
h	700 km
γ	100
a1	10
a2	10
a3	10
a4	10
F1	100
F2	100
F3	100
F4	100

IV. Results

A. Zero-Sum GARE Simulation

1. Regulation

The docking problem is simulated as a regulation problem and the relative position and velocity converges to zero signifying successful docking as seen in figure 1 and 2. The optimal control inputs and disturbance is given in 3 and 4 for the GARE solution found using the Kleinman algorithm.

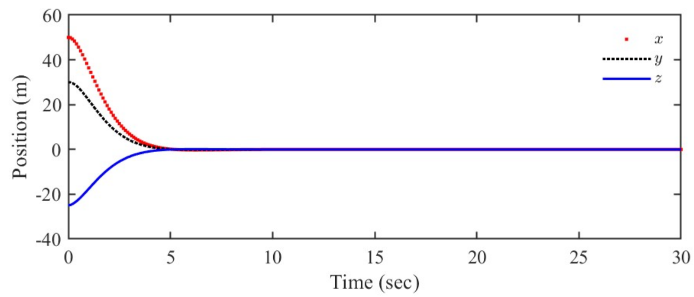

Fig. 1 Position regulation using GARE

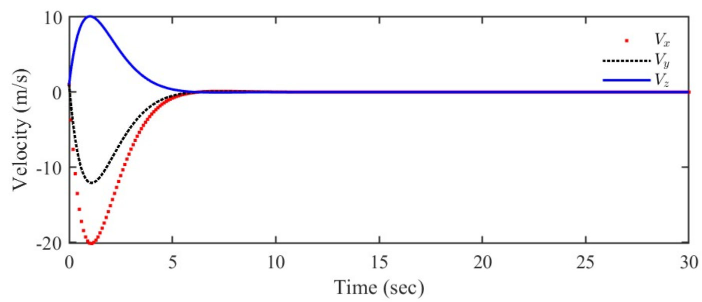

Fig. 2 Velocity regulation using GARE

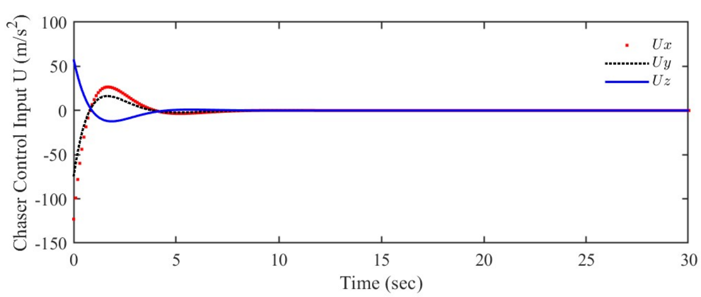

Fig. 3 Optimal control input u* for Regulation using GARE

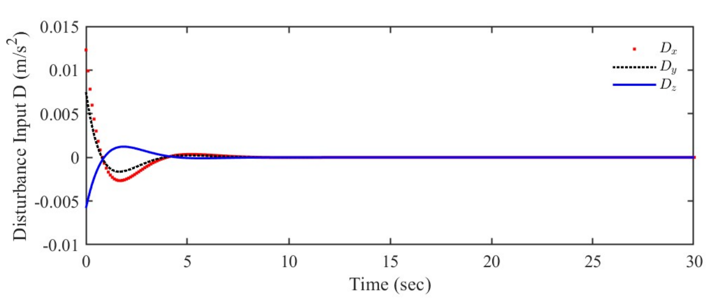

Fig. 4 Optimal disturbance d* for Regulation using GARE

2. Tracking

For the formation flying problem, where the satellites need to maintain a relative position of 10m in the x, y, and z direction, the results are given below in figure 5 and 6. The position successfully tracks 10m while the velocity drops to zero. The optimal control and disturbance inputs for the solution is depicted in figure 7 and 8.

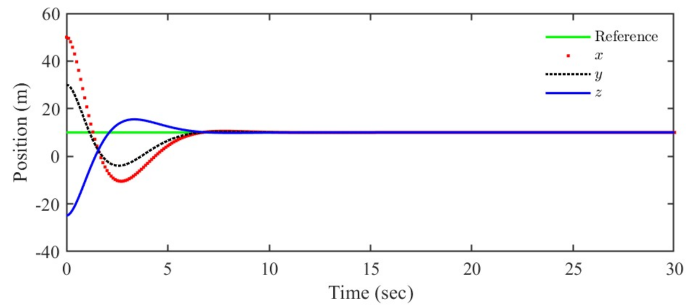

Fig. 5 Position tracking using GARE

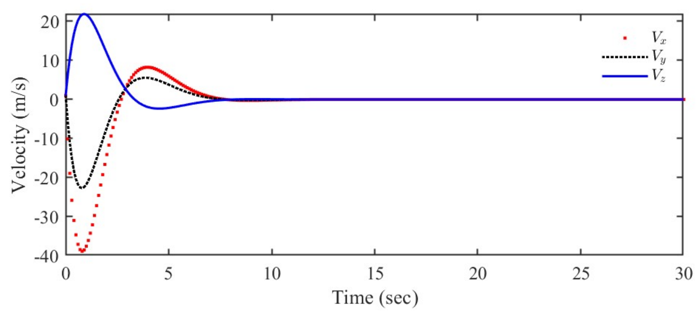

Fig. 6 Velocity result from tracking using GARE

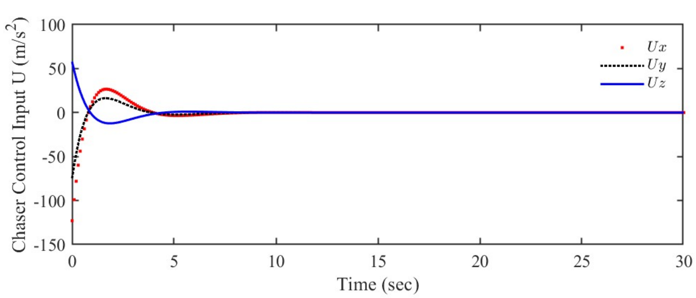

Fig. 7 Optimal control u* using GARE for tracking

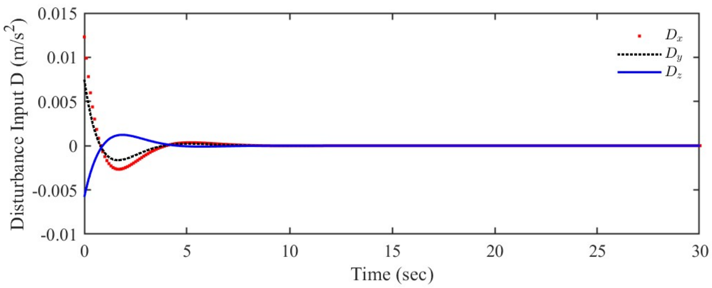

Fig. 8 Optimal disturbance d* using GARE for tracking

B. Zero-Sum Synchronous Online Learning

The Synchronous Online Learning is simulated for the regulation problem, the position and velocity converged to zero as the Persistence of Excitation Noise is cut off as seen in figure 9 and 10. This results converge to zero signifying successful docking maneuver. The outputs of x, y, z variables are mostly overlapping since they are plotted on the same plot and start with the same initial conditions of the relative positions in each direction.

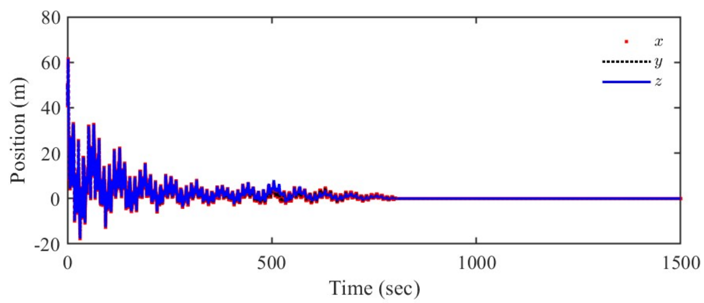

Fig. 9 Synchronous Online Learning results for position regulation

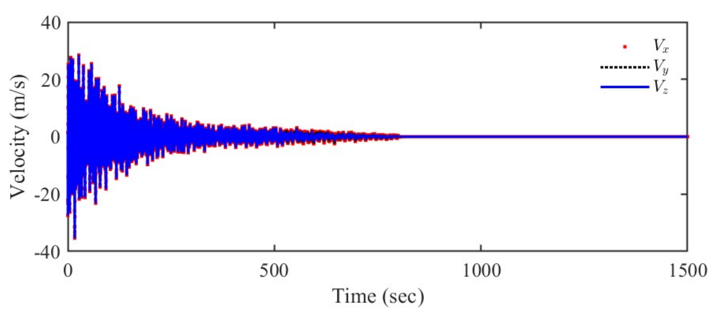

Fig. 10 Synchronous Online Learning results for velocity regulation

The chaser's control input is given below in 11 and the disturbance is given in 12. The P.E. causes the input and the disturbance to shoot to a very high value, this can simply be tuned to get desired result.

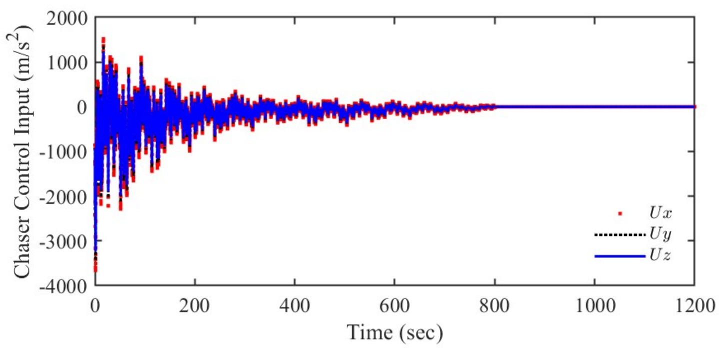

Fig. 11 Synchronous Online Learning control input u* for regulation

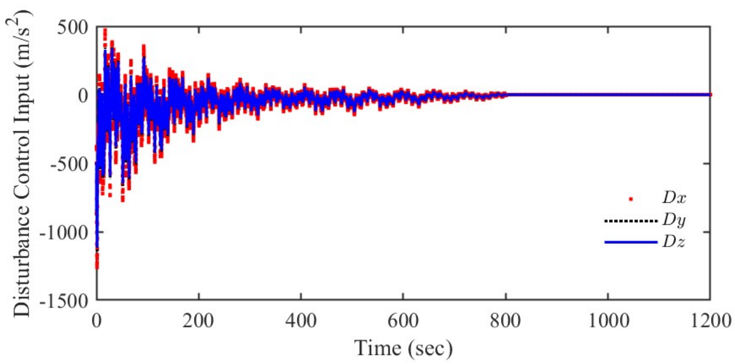

Fig. 12 Synchronous Online Learning control input d* for regulation

C. Non-Zero-Sum Cooperative Game

The Non-Zero-Sum game is simulated where both the chaser and target spacecrafts are cooperating to perform the docking maneuver. The docking is successful as the relative position and velocity drops to 0 as seen in figure 13 and 14.

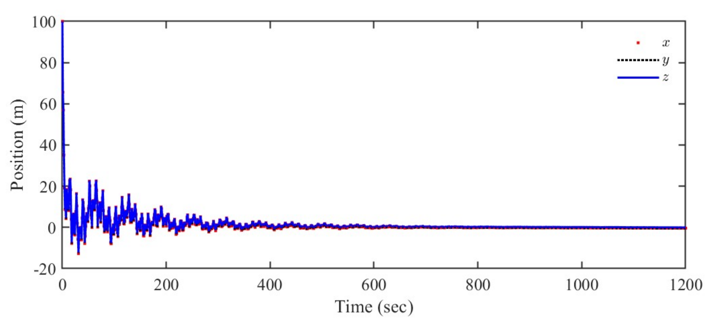

Fig. 13 Non-Zero-Sum Synchronous Online Learning position regulation

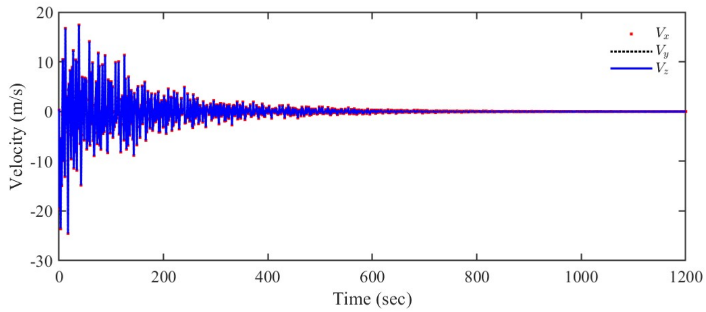

Fig. 14 Non-Zero-Sum Synchronous Online Learning velocity regulation

The optimal chaser and the target control inputs are shown in figure 15 and 16. The P.E. has a high amplitude but it can be tuned so it can converge a faster.

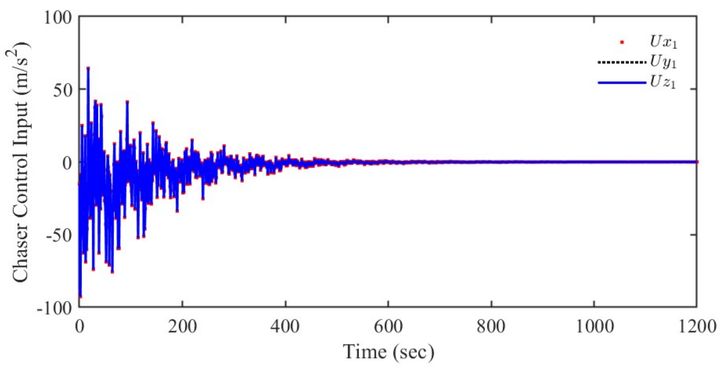

Fig. 15 Synchronous Online Learning control input for chaser

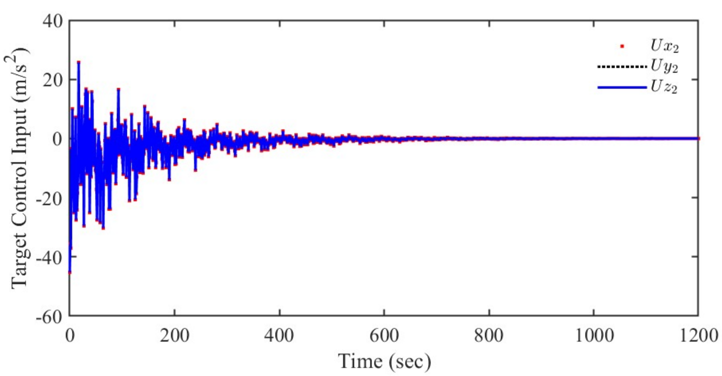

Fig. 16 Synchronous Online Learning control input for target

V. Discussion

From the results, it is evident that a differential-game theoretic approach can be effectively applied to satellite rendezvous, docking, and formation flying scenarios. The GARE (Generalized Algebraic Riccati Equation) approach provided a straightforward solution under the assumption of linear dynamics. It proved to be effective in modeling perturbations as a worst-case disturbance player, and it performed well in both regulation and tracking problems. Similarly, the Synchronous Online Learning Approach showed strong performance for both the zero-sum and non-zero-sum cases in the regulation problem. However, it required significant parameter tuning and careful testing of various persistence-of-excitation functions to ensure sufficient exploration of the weights for value function approximation. Attempts to extend the Synchronous Online Learning approach to the tracking problem using integral-action setpoint control was challenging because of difficulty in tuning, resulting in divergence of the weights and therefore results for this case are not presented in this paper.

To further improve the results for the Synchronous Online Learning approach, future work could include experimenting with various basis functions, reducing the size of the basis function vector, and refining parameter tuning. The initial rationale for selecting polynomials of the state variables as the basis functions is because the state-space dynamics are expressed as a linear combination of state variables. Perturbing forces, such as atmospheric drag and solar radiation pressure are of the polynomial form, making it a good choice of basis functions. Moreover, it is crucial to ensure that the persistence-of-excitation noise remains bounded within the operational range of a spacecraft's thruster output to maintain realistic control.

The different approaches presented in this paper are applicable based on certain assumptions that can be made. For the zero-sum game problem, the assumption is that accurately modeling environmental perturbations—such as atmospheric drag or solar radiation pressure—is challenging due to high variability and uncertainty in atmospheric density and other factors. Therefore, the orbit perturbations are modeled as worst-case disturbances and modeled as a worst-case disturbance player, and solved accordingly. In contrast, the non-zero-sum game assumes that perturbing forces can be modeled as nonlinear terms in the equations of motion, and results in a more efficient docking maneuver because of cooperation between the chaser and target spacecraft.

An extension of the two-player differential game to a multiplayer differential game has applications in multi-satellite formation flying for satellite constellations. In this context, each satellite could cooperate to maintain a specified separation distance from others while in orbit, ensuring optimal formation configurations. This approach can prove to be useful in distributed sensing, communications, and Earth observation missions.

VI. Conclusion

This research project successfully demonstrated the use of differential-game-theoretic control approach for spacecraft rendezvous, docking, and formation flying. The Game-Algebraic-Riccati-Equation (GARE) based solution provided straightforward and robust control under the assumption of linear dynamics, proving useful in handling worst-case perturbations and performing well in both regulation and tracking cases. The Synchronous Online Learning Approach also worked well for regulation in both zero-sum and non-zero-sum cases, though it required careful parameter tuning and sufficient persistence-of-excitation for convergence for the regulation problem. Extending it to the tracking problem proved challenging due to divergence issues. The Zero-Sum game assumption proved to be useful in cases where it's difficult to model the orbit perturbation such as atmospheric drag or solar radiation pressure due to high variability in atmospheric conditions and therefore modeling the perturbations as a worst-case disturbance player, while the Non-Zero-Sum approach allowed for efficient cooperation between between both spacecraft during docking. Future work could focus on improving the Synchronous Online Learning Approach by testing different basis functions, reducing dimensionality, and fine-tuning parameters. Additionally, extending this framework to multiplayer scenarios could support formation flying for satellite constellations, ensuring stable separation distances and efficient control. Overall, this research highlights the adaptability of differential game theory in addressing spacecraft maneuvers in the presence of orbit perturbations as well as cooperative space missions.

Acknowledgments

Abinay J. Brown would like to thank Dr. Kyriakos Vamvoudakis for providing the opportunity, guidance, resources, and support throughout this research project and for imparting knowledge and giving him a foundational understanding on Optimal Control Theory, differential game theory, and reinforcement learning principles.

References

[1] Wiesel, W. E., *Spaceflight Dynamics*, 3rd ed., Aphelion Press, 2010.

[2] Isaacs, R., *Differential Games: A Mathematical Theory with Applications to Warfare and Pursuit, Control and Optimization*, John Wiley & Sons, New York, 1965.

[3] Clohessy, W. H., and Wiltshire, R. S., "Terminal Guidance System for Satellite Rendezvous," *Journal of the Aerospace Sciences*, Vol. 27, No. 9, 1960, pp. 653–658. https://doi.org/10.2514/8.8704.

[4] Schweighart, S., and Sedwick, R., "High-Fidelity Linearized J2 Model for Satellite Formation Flight," *Proceedings of the AAS/AIAA Space Flight Mechanics Meeting*, Santa Barbara, CA, 2001, pp. 1157–1178.

[5] Zhang, Z., Huang, B., and Chen, R., "An Efficient Algorithm for Solving Coupled Riccati Equations in Zero-Sum Games," *Information Sciences*, Vol. 221, 2013, pp. 229–238. https://doi.org/10.1016/j.ins.2012.08.022.

[6] Kleinman, D. L., "On an Iterative Technique for Riccati Equation Computations," *IEEE Transactions on Automatic Control*, Vol. 13, No. 1, 1968, pp. 114–115. https://doi.org/10.1109/TAC.1968.1098594.

[7] Vamvoudakis, K. G., and Lewis, F. L., *Optimal Adaptive Control and Differential Games by Reinforcement Learning Principles*, Control, Robotics and Sensors, Institution of Engineering and Technology (IET), 2021.

[8] Lewis, F. L., Vrabie, D., and Syrmos, V. L., *Optimal Control*, 3rd ed., Wiley, Hoboken, NJ, 2012.