A Fast Fuzzy DRL-Based Joint Beam Design and Power Allocation for Multi-Beam GEO-LEO Coexisting Satellite Networks

Quynh Tu Ngo, *Senior Member, IEEE,* Ying He, *Senior Member, IEEE,*
Beeshanga Jayawickrama, *Senior Member, IEEE,* Eryk Dutkiewicz, *Senior Member, IEEE*

***Abstract*—As demand for ubiquitous connectivity grows, integrating satellite communications into sixth-generation (6G) networks has emerged as a crucial strategy to enhance global coverage, especially in remote and underserved regions. However, achieving the stringent performance, reliability, and spectral efficiency required for 6G presents significant challenges. Coexisting geostationary (GEO) and low Earth orbit (LEO) satellite networks offer a promising solution by enabling complementary coverage and enhanced service capabilities. Nonetheless, a critical challenge is managing intersystem interference from the LEO satellite system on the GEO system when sharing spectral resources, all while maintaining the performance of both systems. To address this, this paper introduces a fast fuzzy deep reinforcement learning (DRL)-based approach for joint beam design and power allocation in multi-beam GEO-LEO coexisting satellite networks. A robust design problem of LEO beam size and power allocation is formulated to maximize the spectral efficiency of the LEO system, considering tolerable interference on the GEO system, frequency reuse schemes employed by both GEO and LEO systems, and Doppler frequency offset induced by LEO satellite movement. A fast DRL algorithm, integrating fuzzy logic, post-decision state, and deep deterministic policy gradient, is proposed to solve this problem. Numerical results demonstrate a faster learning convergence rate for the proposed DRL algorithm compared to benchmark algorithms and confirm that the proposed method enhances LEO spectral efficiency while maintaining tolerable intersystem interference on the GEO system.**

***Index Terms*—Multi-beam GEO-LEO networks, LEO beam size design, Doppler frequency offset, fuzzy logic, post-decision state, DDPG.**

I. Introduction

Satellite communications (SatCom) offer a compelling solution for the upcoming sixth-generation (6G) networks, extending global connectivity to remote and underserved regions. Among SatCom architectures, the geostationary (GEO) and low Earth orbit (LEO) coexisting satellite network (CSN) stands out as a promising configuration for 6G, leveraging GEO satellites' wide coverage and LEO satellites' low latency and high throughput capabilities [1]. However, the rapid growth of LEO satellite deployments is creating a scarcity of available spectral resources. Unlike GEO satellites, which occupy fixed positions and provide broad, stable coverage, LEO satellites operate in lower orbits, covering smaller, more dynamic areas. This setup allows for higher frequency reuse but also increases the risk of interference as more LEO satellites operate concurrently. As LEO constellations expand, they require more spectral resources to support continuous service across vast regions, prompting regulatory attention to mitigate the interference with GEO systems already using these resources.

While traditional spectrum allocation strategies have assigned different frequency bands to GEO and LEO satellites to reduce interference, several emerging and operational systems demonstrate spectrum coexistence between these orbits. For example, Starlink and OneWeb LEO constellations operate in the Ku- and Ka-bands, which are also heavily utilized by GEO satellites for fixed satellite services and broadband communications [2]. Amazon's Kuiper constellation is planned to use the Ka-band, overlapping with GEO systems in the same band [3]. Similarly, operators such as Eutelsat-OneWeb and Telesat Lightspeed manage GEO and LEO constellations within shared frequency bands [4]–[6], further emphasizing the importance of interference mitigation strategies in multi-orbit satellite networks. These real-world examples illustrate the growing trend toward spectrum sharing and the necessity for efficient coexistence mechanisms.

In GEO-LEO CSN, the International Telecommunication Union (ITU) has established regulatory guidelines to address these interference risks. These guidelines mandate that LEO systems prioritize the protection of existing GEO systems by avoiding unacceptable interference that could degrade GEO services [7]. Specifically, LEO satellites are required to limit their transmission power and manage spectral resources in a way that minimizes interference with GEO satellites, which serve as primary users of these frequencies. The ITU's regulatory framework enforces strict power limitations, frequency coordination, and spatial isolation measures between GEO and LEO satellites [7]. These constraints are crucial to ensuring that GEO systems, which often support critical applications like broadcast television, weather monitoring, and emergency communications, maintain reliable service quality despite the rapid growth of LEO deployments.

To comply with these regulations, most studies have concentrated on protecting GEO satellite systems [8], which poses significant challenges for LEO systems by often limiting their access and operational flexibility. This highlights the need for

The authors are with the School of Electrical and Data Engineering, University of Technology Sydney, Sydney, NSW 2007, Australia (Email: {QuynhTu.Ngo, Ying.He, Beeshanga.Jayawickrama, Eryk.Dutkiewicz}@uts.edu.au).

This work has been supported by the SmartSat CRC, whose activities are funded by the Australian Government's CRC Program.

innovative strategies that allow LEO systems to navigate these restrictions while maintaining service quality and coverage. Consequently, the coexistence of GEO and LEO networks in shared spectral environments requires innovative strategies that can balance the spectrum demands of LEO networks with the interference protection needs of GEO satellites.

In recent years, the application of advanced antenna technologies, particularly phased array systems, has led to significant advancements in high-throughput multi-beam satellite systems with highly adaptable beam coverage. These innovations are shaping the new space era, where multi-beam satellite systems are increasingly prevalent due to their ability to dynamically allocate bandwidth and improve spectral efficiency [9]. Phased array antennas enable precise control over beam shapes, directions, and power levels, supporting a range of capabilities, including rapid beam steering and reconfiguration [10]. Multi-beam techniques offer significant potential for implementing advanced spatial domain strategies that mitigate intersystem interference between GEO and LEO satellite networks. In addition, the 3rd Generation Partnership Project (3GPP) has recognized frequency reuse as a critical technique to enhance spectrum efficiency within non-terrestrial networks [1]. This approach, first emphasized in 3GPP Release 17, enables satellite and high-altitude platform systems to better utilize limited spectrum resources by allowing multiple transmissions to share the same frequency bands with minimal interference. By implementing frequency reuse strategies, 3GPP aims to support the increasing demand for connectivity, facilitating wider coverage and improved data capacity while maintaining reliable and efficient communication links across diverse environments. Recent studies have explored various approaches to address this issue. [11] proposed a cooperative beamforming strategy using multiple satellites and multiple reconfigurable intelligent surfaces (RIS) to manage interference between GEO and LEO networks effectively. This approach leverages coordinated beam adjustments across satellites and RIS elements to minimize interference across shared spectral bands. Authors in [12] introduced an interference-mitigation method by grouping LEO satellites and their users to prevent severe co-channel interference. They also proposed a continuous beam-pointing technique for LEO user antennas, which dynamically adjusts beam direction to avoid co-channel interference with GEO systems. This adaptive beam orientation is particularly beneficial in reducing interference in crowded frequency bands. [13] developed a framework for analyzing intersystem interference from non-geostationary orbit constellations, such as LEO systems, to GEO satellites based on spatio-temporal slices. This framework models interference patterns by accounting for the spatial and temporal variations in satellite orbits and beam coverage areas, providing insights into interference dynamics over time. [14] introduced a dynamic spectrum access scheme tailored for LEO satellites, enabling them to opportunistically reuse the frequency bands initially allocated to GEO satellites within the GEO-LEO CSN framework. This study provides the analysis of the interference experienced by GEO users as a result of the frequency reuse scheme, alongside an evaluation of LEO average throughput.

These studies, however, rely on the assumption of perfect Doppler compensation, where Doppler shifts from the high-speed movement of LEO satellites are fully mitigated. In practical GEO-LEO CSN, however, the high relative velocity of LEO satellites introduces significant Doppler shifts [15], complicating interference management between GEO and LEO systems. To address these shifts, pre-compensation techniques are often employed, adjusting transmission parameters at the LEO satellite to offset anticipated Doppler shifts. This approach is particularly effective for users located at the center of a LEO beam, where Doppler compensation can be more accurate. However, users located off-center experience residual Doppler effects, known as Doppler frequency offsets, which lead to incomplete Doppler compensation. In multi-beam satellite systems, Doppler frequency offsets pose additional challenges, especially for narrow beams. Even slight misalignments can cause the intended signal to drift into neighboring subbands, making narrow beams particularly susceptible to frequency shifts due to Doppler effects. These shifts increase the potential for inter-beam and intersystem interference, affecting both GEO users located at the center and those outside the central area of the LEO beam. Consequently, Doppler frequency offsets must be carefully considered when designing interference management strategies for GEO-LEO CSNs, to ensure reliable communication and effective mitigation of interference across these dynamic networks.

To address the above gaps, this article presents a joint design of LEO beam size and power allocation, taking into account tolerable intersystem interference, Doppler frequency offsets, and frequency reuse to maximize the spectral efficiency of the LEO system within a multi-beam GEO-LEO CSN. To our knowledge, this is the first work in the literature to incorporate Doppler frequency offsets directly into the design framework for GEO-LEO CSNs. Our design leverages deep reinforcement learning (DRL) to optimize decision-making in dynamic environments. The proposed DRL-based approach is designed for robustness, ensuring adaptability to dynamic and previously unseen network conditions. This is achieved through a fast-converging DRL algorithm that integrates fuzzy logic for adaptive action exploration, post-decision state (PDS) techniques for enhanced learning stability, and the deep deterministic policy gradient (DDPG) approach [16]. These elements collectively improve the agent's ability to learn optimal policies efficiently while maintaining high performance across diverse network conditions. The main contributions of this article are as follows.

1) We analyze the impact of Doppler frequency offsets on transmission bandwidth in a practical multi-beam GEO-LEO CSN, considering various frequency reuse schemes.

2) We develop an analytical framework to quantify intersystem interference and evaluate LEO spectral efficiency within a multi-beam GEO-LEO CSN, incorporating Doppler frequency offsets and frequency reuse schemes for both typical operational scenario and handover situation between LEO satellites.

3) We design a joint optimization strategy for LEO beam size and power allocation to maximize LEO spectral efficiency. This non-convex optimization problem is addressed using a DRL approach, where fuzzy logic is incorporated into the

action exploration process, and PDS learning is integrated into the DDPG algorithm to enhance convergence speed.

4) We evaluate the performance of our proposed approach under different frequency reuse schemes and realistic operational scenarios for LEO satellites. Additionally, we compare our method against alternative strategies, including independent power allocation and beam size optimization, as well as approaches that do not account for Doppler shift variations. The results highlight the advantages of our approach in terms of adaptability and spectral efficiency improvements.

The remainder of this article is organized as follows: Section II provides an overview of the system model, discusses the impact of Doppler frequency offsets, and outlines our analytical framework. Section III details the joint design of LEO beam size and power allocation, including the problem formulation and our proposed DRL algorithm for solving the optimization problem. Numerical results are presented in Section IV. Finally, Section V concludes the article.

II. System Model

Consider a multi-beam GEO-LEO coexisting satellite network, as illustrated in Fig. 1, comprising a GEO satellite network and a LEO satellite constellation. The GEO network features a GEO satellite S_G serving its users. Employing a multi-beam mode with frequency reuse, the GEO satellite uses a spectrum of bandwidth B_G and an ι_G-color frequency reuse scheme, where B_G is divided into ι_G sub-bands ($\iota_G > 1$). The coverage areas of the GEO spot beams on the ground are assumed to be circular, each with the same radius R_G. The LEO network comprises a LEO satellite constellation serving a specific ground area. At any given time, a LEO satellite S_L serves its users $U_{L_i}, i = 1, \ldots, N$, using the GEO spectrum. A handover occurs when the current serving LEO satellite exits the area and a new LEO satellite enters to take over[1]. Each LEO satellite is equipped with phased array antennas, which are capable of dynamically shape narrow spot beams while simultaneously maintaining a fixed gaze in a particular direction. Without loss of generality, it is assumed that all the LEO satellites have identical circular footprints with a radius F_L, $F_L > R_G$. Inside these footprints, multiple LEO spot beams are uniform hexagons, each with a radius of R_L, $R_L < R_G$, and they exhibit a non-overlapping configuration, with a uniform antenna radiation pattern within each beam's area. Let $N = 2\pi F_L^2/3\sqrt{3}R_L^2$ represent the number of LEO spot beams within a LEO satellite footprint, and assume that LEO users are distributed such that there is at least one user per spot beam coverage. All GEO and LEO users are equipped with single antennas.

Let s_G and s_{L_i} respectively represent the normalized data symbols requested by U_G and U_{L_i}, with $\mathbb{E}\left[|s_G|^2\right] = \mathbb{E}\left[|s_{L_i}|^2\right] = 1$. Let t denote the transmission time. Hence, the signal received at a GEO user U_G can be expressed as

$$y_{U_G}(t) = \sqrt{\frac{P_G G_G}{L_{GU_G}}} h_{GU_G}(t) s_G + I_{U_G}(t) + n_{U_G}(t), \quad (1)$$

[1]The operation during a handover is described separately in Section II-E1.

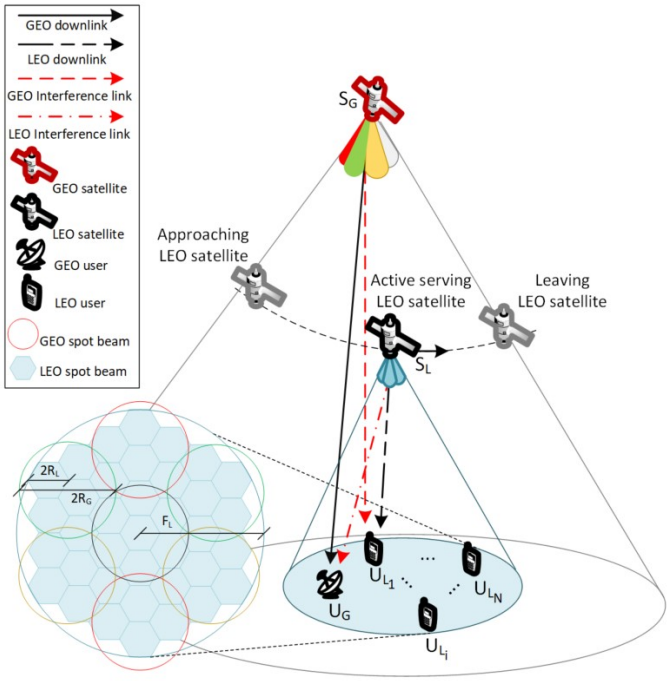


Fig. 1: A multi-beam GEO-LEO coexisting satellite network with GEO satellite employing 4-color frequency reuse scheme.

where P_G is the GEO satellite's transmit power; G_G is the GEO spot beam gain; L_{GU_G} is the free space path loss between S_G and U_G; h_{GU_G} is the channel coefficient of the link between S_G and U_G; I_{U_G} is the total interference on U_G; and n_{U_G} is additive white Gaussian noise (AWGN) at U_G with zero mean and variance $\sigma_{n_G}^2$.

The signal received at a LEO user U_{L_i} served by the i-th LEO spot beam can be expressed as

$$y_{U_{L_i}}(t) = \sqrt{\frac{P_{L_i} G_{L_i}}{L_{LU_{L_i}}}} h_{LU_{L_i}}(t) s_{L_i} e^{j2\pi\Delta f_i^D(t)t} + I_{U_{L_i}}(t) + n_{U_L}(t), \quad (2)$$

where P_{L_i} is the LEO i-th spot beam transmit power; G_{L_i} is the LEO i-th spot beam gain; $L_{LU_{L_i}}$ is the free space path loss between S_L and U_{L_i}; $h_{LU_{L_i}}$ is the channel coefficient of the link between S_L and U_{L_i}; Δf_i^D denotes the frequency offset due to Doppler shift induced by the LEO satellite movement; $I_{U_{L_i}}$ is the total interference on U_{L_i}; and n_{U_L} is AWGN at LEO user with zero mean and variance $\sigma_{n_L}^2$.

A. Spot Beam Gain

The spot beam of a GEO/LEO satellite can be modeled as

$$G_{G/L_i}(\phi) = G_{G/L_i}^{max}\left(\frac{J_1(\mu)}{2\mu} + \frac{36J_3(\mu)}{\mu^3}\right)^2, \quad (3a)$$

$$G_{G/L_i}^{max} = \eta\frac{4\pi A}{(c/f_c)^2}, \quad (3b)$$

$$\mu = 2.07123\frac{\sin(\phi)}{\sin(\phi_{3dB})}, \quad (3c)$$

where ϕ denotes the satellite's off-boresight angle in the direction of its user; ϕ_{3dB} is the off-boresight angle corresponding

to the 3 dB beamwidth; J_1 and J_3 respectively denote the first order and third order Bessel functions; G^{max}_{G/L_i} denotes the maximum antenna gain when $\phi = 0$; A and η are the antenna area and efficiency, respectively; c is the speed of light; and f_c is the carrier frequency.

B. Channel Model

Perfect channel state information (CSI) is assumed. The channel between a satellite and its user, i.e., h_{GU_G}, is modeled using the Shadowed-Rician fading, characterized by the following probability distribution function (PDF) [17],

$$f_h(x) = \varkappa \cdot \frac{e^{-x/b}}{b} \cdot {}_1F_1\left(m, 1, \varsigma x\right), \tag{4a}$$

$$\varkappa = \left(\frac{bm}{bm+w}\right)^m, \tag{4b}$$

$$\varsigma = \frac{w}{b(bm+w)}, \tag{4c}$$

where b is the average power of the scatter components; w is the average power of the line-of-sight (LoS) component; m is the Nakagami parameter; and ${}_1F_1(\cdot,\cdot,\cdot)$ is the first kind confluent hypergeometric function.

The channel power gain, i.e., $|h_{GU_G}|^2$, can be modeled using the squared Shadowed-Rician distribution characterized by the following PDF,

$$f_{h^2}(y) = \frac{\kappa}{2\sqrt{y}} \cdot \frac{e^{-\sqrt{y}/b}}{b} \cdot {}_1F_1\left(m, 1, \sigma\sqrt{y}\right). \tag{5}$$

The path loss between a satellite and its user is modeled as free space propagation loss,

$$L = \left(\frac{4\pi d}{c/f_c}\right)^2, \tag{6a}$$

$$d = 2R_E \arcsin\left(\sqrt{\sin^2\left(\frac{l^a_U - l^a_S}{2}\right) + \cos(l^a_S)\cos(l^a_U)\sin^2\left(\frac{l^o_U - l^o_S}{2}\right)}\right), \tag{6b}$$

where R_E is the Earth's radius; $\left\{l^a_S, l^o_S\right\}$ and $\left\{l^a_U, l^o_U\right\}$ denote the latitude and longitude of the satellite and its user, respectively.

C. Doppler Frequency Offset

Due to the rapid movement of LEO satellites, ground users experience a frequency shift in the received signals, known as the Doppler shift. To mitigate this effect in a LEO satellite communication system, pre-compensation techniques are utilized. These techniques are implemented by the LEO satellites before transmission, allowing them to adjust for the Doppler shift. Each LEO satellite can only compensate for the Doppler shift relative to its position and the beam center where the ground users are located. The satellite, knowing its orbital direction, can predict its future position. Additionally, it can determine the beam center's location using data from the global navigation satellite system. With perfect pre-compensation, users located at the beam center do not experience any Doppler frequency offset. However, users not situated at the beam center will experience a Doppler frequency offset.

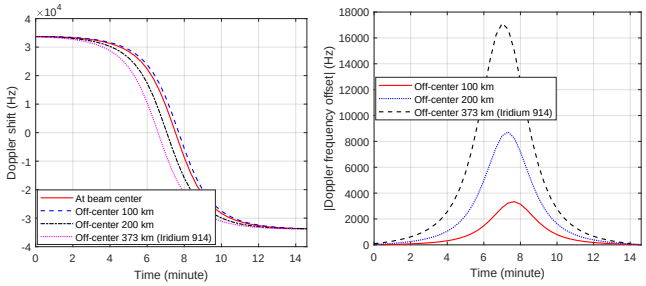
Let V_L denote the velocity of the LEO satellite S_L. The Doppler shift experienced by a user U_{L_i} located at the i-th beam center during a communication session with S_L can be expressed as

$$f^D_{i,ctr}(t) = \frac{V_L(t)}{c} f_c \cos\varphi_{ctr}(t), \tag{7}$$

where φ_{ctr} denotes the angle between the satellite's velocity vector and the line of sight from the satellite to the user. Assuming perfect pre-compensation, the Doppler frequency offset experienced by a user U_{L_i} can be expressed as

$$\Delta f^D_i(t) = \begin{cases} 0, & \text{at the center of the } i\text{-th beam} \\ \frac{V_L(t)}{c} f_c \left(\cos\varphi_{oc}(t) - \cos\varphi_{ctr}(t)\right), & \text{off-center} \end{cases} \tag{8}$$

where φ_{oc} denotes the angle between the satellite's velocity vector and the LoS to the user U_{L_i} located off-center. Fig. 2 illustrates the Doppler shift and Doppler frequency offset experienced by a user located at a beam center, off-center 100 km, off-center 200 km, and at the beam edge during a communication session with the LEO satellite Iridium-NEXT 914 at a frequency of 1.53 GHz over one complete orbital cycle.



(a) Doppler shift (b) Doppler frequency offset

Fig. 2: Doppler shift and frequency offset experienced by a user located at different position within a LEO satellite beam.

Remark 1. *The upper bound Doppler frequency offset experienced by a user U_{L_i} occurs when the user is located at the farthest edge of a beam and can be expressed as*

$$\sup\left\{\left|\Delta f^D_i\right|\right\} = \frac{V^*_{L_i}}{c} f_c \left(\frac{R_L}{\sqrt{R^2_L + A^2_L}}\right), \tag{9}$$

where $V^*_{L_i}$ denotes the velocity when the satellite is directly overhead the i-th beam center; A_L is the satellite's altitude.

Proof. When the satellite is directly overhead the beam center, the line of sight to the beam center is perpendicular to the satellite's velocity vector. In this case, $\cos\varphi_{ctr} = 0$ in (8) and the Doppler frequency offset can be written as a function of the off-center distance r,

$$\Delta f^{D*}_i = \pm \frac{V^*_{L_i}}{c} f_c \left(\frac{r}{\sqrt{r^2 + A^2_L}}\right). \tag{10}$$

With Δf_i^{D*} increasing when r increases, and $0 \leq r \leq R_L$, (9) can be derived. □

D. Effective Bandwidth

Let ι_L denote the frequency reuse scheme factor at LEO satellites. The bandwidth per beam available for LEO transmission is

$$B_{pb} = \frac{B_G}{\iota_G \iota_L}, \tag{11a}$$

$$\iota_L = \begin{cases} 1, \text{ full frequency reuse} \\ N_{GL}, \text{ no frequency reuse} \end{cases} \tag{11b}$$

where $N_{GL} = 2\pi R_G^2 / 3\sqrt{3} R_L^2$ is the number of LEO spot beams inside one GEO spot beam. In a full frequency reuse scenario, the LEO spot beams within a GEO beam utilize the same frequency band as the GEO beam. Conversely, a no frequency reuse scheme implies that the bandwidth of the GEO beam is evenly divided among all the LEO beams within that GEO beam.

The effective bandwidth per beam, B_{pb}^e, available for LEO transmission after accounting for Doppler frequency offset can be expressed as

$$B_{pb} - 2 \sup \left\{ \left| \Delta f_i^D \right| \right\} \leq B_{pb}^e \leq B_{pb}. \tag{12}$$

To illustrate the effect of Doppler shift, let's consider a cognitive satellite network with the GEO satellite being Inmarsat-4 F1 and the LEO satellite being Iridium-NEXT 914. The Inmarsat-4 F1 employs a 4-color frequency reuse scheme ($\iota_G = 4$) with each spot beam having a bandwidth of 200 kHz and a radius of 555 km [18]. The Iridium-NEXT 914 covers an area with a radius of 2350 km with 48 spot beams [19], thus there are 18 GEO spot beams within the LEO satellite's coverage and 3 LEO spot beams within a single GEO spot beam coverage. This results in a bandwidth of 66.67 kHz per beam available for the LEO satellite if it employs no frequency reuse scheme. A LEO user within that beam can experience a maximum Doppler frequency offset of 17.5 kHz, which reduces the minimum effective bandwidth per beam available for LEO transmission to 31.67 kHz. If the LEO satellite employs a full frequency reuse scheme, the effective bandwidth per beam is 165 kHz.

Remark 2. *The infimum of the LEO beam size, considering Doppler frequency offset to ensure that transmission is possible, can be expressed as*

$$\inf \{R_L\} = \sqrt{\frac{B_{pb}^2 c^2 A_L^2}{4 V_{L_i}^{*2} f_c^2 - B_{pb}^2 c^2}}. \tag{13}$$

Proof. For LEO transmission to be feasible, the effective bandwidth per beam must be positive. Therefore, the lower bound for the effective bandwidth per beam, as given in (12), must be greater than zero. This requirement leads to the result expressed in (13). □

E. Interference Model

The interference experienced by the GEO user U_G arises from LEO transmission when a LEO satellite transmits on the same sub-band used by the active GEO spot beam. The total interference on U_G in (1) can be expressed as

$$I_{U_G}(t) = \sum_{i=1}^{N_{GL}} \sqrt{\frac{P_{L_i} G_{L_i}}{L_{LU_G}}} h_{LU_G}(t) s_{L_i} e^{j2\pi \Delta f_i^D(t) t}. \tag{14}$$

Note that when a full frequency reuse scheme is employed at LEO satellite S_L, the interference signals occupy the same bandwidth as the desired GEO signal. Conversely, when a no frequency reuse scheme is employed, the interference signals occupy distinct subbands within the bandwidth of the desired signal. The corresponding signal-to-interference-plus-noise ratio (SINR) per bandwidth at U_G can be written as

$$\Gamma_{U_G} = \frac{P_G G_G L_{GU_G}^{-1} \left| h_{GU_G} \right|^2 \iota_G / B_G}{\sum_{i=1}^{N_{GL}} P_{L_i} G_{L_i} L_{LU_G}^{-1} \left| h_{LU_G} \right|^2 / B_{pb}^e + \sigma_{n_G}^2}. \tag{15}$$

For a LEO user U_{L_i}, the total interference in (2) includes interference from GEO transmission when U_{L_i} is located within the active GEO beam, as well as LEO inter-beam interference if the LEO satellite employs a full frequency reuse scheme. The total interference on U_{L_i} can be expressed as

$$\begin{aligned} I_{U_{L_i}}(t) = & \sqrt{\frac{P_G G_G}{L_{GU_{L_i}}}} h_{GU_{L_i}}(t) s_G \\ & + \alpha_f \sum_{l, l \neq i}^{N_{GL}} \sqrt{\frac{P_{L_l} G_{L_l}}{L_{LU_{L_l}}}} h_{LU_{L_l}}(t) s_{L_l} e^{j2\pi \Delta f_l^D(t) t}, \end{aligned} \tag{16}$$

where $\alpha_f \in \{0, 1\}$ denotes the LEO frequency reuse scheme, in which $\alpha_f = 0$ for no frequency reuse scheme and $\alpha_f = 1$ for full frequency reuse scheme.

The corresponding SINR per bandwidth at U_{L_i} can be written as

$$\begin{aligned} \Gamma_{U_{L_i}} = & \frac{P_{L_i} G_{L_i} L_{LU_{L_i}}^{-1} \left| h_{LU_{L_i}} \right|^2 / B_{pb}^e}{P_G G_G L_{GU_{L_i}}^{-1} \left| h_{GU_{L_i}} \right|^2 \iota_G / B_G + \sigma_{n_{U_L}}^2} \\ & + \alpha_f \sum_{l, l \neq i}^{N_{GL}} P_{L_l} G_{L_l} L_{LU_{L_l}}^{-1} \left| h_{LU_{L_l}} \right|^2 / B_{pb}^e \end{aligned}. \tag{17}$$

The spectral efficiency of LEO satellite system, η_L, is defined as the ergodic channel capacity in nat/sec/Hz [20] and can be expressed as

$$\eta_L = \sum_{i=1}^{N} \ln \left(1 + \Gamma_{U_{L_i}} \right). \tag{18}$$

1) Operation under handover scenario: During handover, to ensure continuous service, both the departing LEO satellite and the approaching one simultaneously cover the same footprint region. As a result, the interference on GEO users during this period is higher. Let $\Gamma_{U_G}^{HO}$ represent the SINR per bandwidth at U_G during handover, and let L^{ap} and L^{de} denote the approaching and departing LEO satellites, respectively.

* The SINR per bandwidth at the GEO user during handover:

$$\Gamma_{U_G}^{HO} = \frac{P_G G_G L_{GU_G}^{-1} \left|h_{GU_G}\right|^2 \iota_G / B_G}{\sum\limits_{i=1}^{N_{GL}} P_{L_i^{de}} G_{L_i^{de}} L_{L^{de}U_G}^{-1} \left|h_{L^{de}U_G}\right|^2 / B_{pb}^{e*} + \sum\limits_{i=1}^{N_{GL}} P_{L_i^{ap}} G_{L_i^{ap}} L_{L^{ap}U_G}^{-1} \left|h_{L^{ap}U_G}\right|^2 / B_{pb}^{e*} + \sigma_{n_G}^2} \tag{19}$$

The SINR per bandwidth at the GEO user during handover can be expressed as in (19) where B_{pb}^{e*} represents the effective bandwidth per beam available for LEO transmission during handover. B_{pb}^{e*} is expressed as

$$\max\left\{B_{pb} - 2\sup\left\{\left|\Delta f_i^{D,de}\right|\right\}, B_{pb} - 2\sup\left\{\left|\Delta f_i^{D,ap}\right|\right\}\right\} \leq B_{pb}^{e} \leq B_{pb} \tag{20}$$

where $\Delta f_i^{D,de}$ and $\Delta f_i^{D,ap}$ are the Doppler frequency offset due to the departing and approaching LEO satellites, respectively.

III. Joint LEO Beam Design and Power Allocation

A. Problem Formulation

Let's consider the joint design of LEO beam size and LEO transmit power allocation to maximize the spectral efficiency of the LEO system. The objective is to maximize the spectral efficiency while adhering to the constraints of maximum LEO satellite power budget, tolerable interference on the GEO user, and acceptable Doppler frequency offset. The optimization problem can be formulated as follows:

$$\max_{R_L, P_{L_i}} \ \eta_L \tag{21a}$$

$$\text{s.t.} \sum_{i=1}^{N} P_{L_i} \leq P_{L,\max}, \tag{21b}$$

$$\Gamma_{U_G} \geq \bar{\Gamma}_{U_G}, \tag{21c}$$

$$B_{pb} - 2\sup\left\{\left|\Delta f_i^{D}\right|\right\} \leq B_{pb}^{e} \leq B_{pb}, \tag{21d}$$

$$\inf\{R_L\} \leq R_L \leq R_G, \tag{21e}$$

where $P_{L,\max}$ denotes the LEO satellite's maximum transmit power, and $\bar{\Gamma}_{U_G}$ is the SINR threshold of GEO user U_G.

The constraint in (21b) is imposed to satisfy the maximum power budget at the LEO satellite. The constraint in (21c) ensures sufficient protection on GEO user, which is strictly mandatory in GEO-LEO coexisting satellite networks. Lastly, the constraints in (21d) and (21e) address the Doppler frequency offset for feasible LEO transmission.

The optimization problem in (21) is inherently non-convex due to the logarithmic relationship of the objective function with SINR and the non-linear interactions within the constraints, particularly those related to SINR and Doppler frequency offsets. Traditional convex optimization methods struggle with such complexities, often requiring problem-specific relaxations or approximations that may not generalize well to dynamic satellite environments. Heuristic optimization techniques, while feasible, lack adaptability and scalability for real-time decision-making.

To address these challenges, we employ a reinforcement learning (RL) approach with a continuous action space, which offers several advantages. First, RL allows for adaptive learning from CSI and Doppler variations during training, enabling the model to generalize to dynamic and time-varying GEO-LEO spectrum environments. Second, a continuous-action RL model enables fine-grained control over beam size and power allocation, which would be difficult to achieve with a discrete-action RL approach or conventional optimization techniques. Lastly, to mitigate sample inefficiency and stability issues often associated with RL in continuous spaces, we incorporate fuzzy logic for exploration control, post-decision state techniques for stability, and a deep deterministic policy gradient-based framework to enhance convergence.

B. Reinforcement Learning-Based Problem Reformulation

Reinforcement learning is a machine learning approach in which an agent learns to make decisions by interacting with an environment to maximize cumulative rewards. The agent observes the environment's current state, takes an action guided by a policy, and receives feedback as a reward. The environment then transitions to a new state influenced by the agent's action. This process repeats, with the agent's objective being to learn an optimal policy that maximizes cumulative rewards. Using various algorithms, the agent iteratively refines its policy based on the rewards received, thereby enhancing its decision-making capabilities over time.

In this section, problem (21) is reformulated as a RL problem following Markov decision process (MDP), within the multi-beam GEO-LEO coexisting satellite system environment. The key elements of the MDP are described below.

Action space $\mathcal{A}$: Let $a_t \in \mathcal{A}$ denote the action at time step t. Since problem (21) aims to joint design the LEO beam size and power allocation, the action a_t can be defined in terms of the beam radius $R_{L,t}$ and transmit power $\mathbf{P}_{L,t} = \left[P_{L_1,t}, \ldots, P_{L_N,t}\right]$, i.e.,

$$a_t \triangleq \left\{R_{L,t}, \mathbf{P}_{L,t}\right\}. \tag{22}$$

Reward function: The reward function evaluates how effective the LEO beam design and power allocation are after the agent takes an action. Since the goal of problem (21) is to maximize the LEO spectral efficiency, the reward at time step t can be defined as a function of spectral efficiency:

$$r_t = \begin{cases} \eta_{L,t}, & \text{if (21b)-(21e) are satisfied,} \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

The cumulative reward is given by $R_t = \sum_{i=t}^{\infty} \kappa^{i-t} r_i$, where $\kappa \in \{0, 1\}$ is the reward discount factor.

State space $\mathcal{S}$: Let $s_t \in \mathcal{S}$ denote the state at time step t. s_t includes the current CSI of GEO user and all LEO users, represented by $\mathbf{h}_t = \left[h_{G,t}, h_{L_1,t}, \ldots, h_{L_N,t}\right]$, i.e., $h_{G,t} = G_{G,t}^{1/2} L_{GU_G,t}^{-1/2} h_{GU_G,t}$, the upper bound Doppler frequency offset

$\mathbf{\Delta f}_{D,t} = \left[\sup\left\{\left|\Delta f_1^D\right|\right\}, \ldots, \sup\left\{\left|\Delta f_N^D\right|\right\}\right]$, the action from previous time step (for supporting PDS learning), and the reward received from previous time step (for guiding fuzzy action exploration). Formally, the state s_t can be written as

$$s_t \triangleq \left\{\mathbf{h}_t, \mathbf{\Delta f}_{D,t}, R_{L,t-1}, \mathbf{P}_{L,t-1}, r_{t-1}\right\}. \tag{24}$$

Since the LEO satellite is moving, its mobility is reflected in the distance variation between the LEO satellite and ground users, which in turn affects the variation of the CSIs. Thus, the mobility of the LEO satellite influences the dynamics of the RL states but does not impact the design frameworks.

C. Proposed Fuzzy Post Decision State-Deep Deterministic Policy Gradient Learning Algorithm

The proposed solution, namely fuzzy post decision state-deep deterministic policy gradient (FPDS-DDPG), leverages fuzzy logic to accelerate the learning speed of our previously developed deep post decision state-deterministic policy gradient (DPDS-DPG) [21], which is specially designed to capture the dynamic changes of the environment. DPDS-DPG builds upon PDS learning, recognizing for its rapid convergence [22], and DDPG [16], which integrates the advantages of deep Q-learning and policy gradient methods to effectively handle continuous action spaces.

In DPDS-DPG, random noise exploration introduces fixed or stochastic noise into the actions, which does not adapt to the current state or the agent's learning progress. This unstructured approach can lead to inefficient learning and suboptimal policies, as it may cause over-exploration or under-exploration, and result in slower convergence. To address these challenges, in the proposed FPDS-DDPG, random noise exploration is replaced with fuzzy logic. Fuzzy action exploration adapts the exploration process based on the current state and past performance, using predefined rules from expert knowledge or empirical data. This structured approach allows for more adaptive, context-sensitive exploration, targeting more promising areas of the state-action space and reducing the risks of over- and under-exploration. By intelligently modulating exploration, fuzzy logic facilitates faster convergence by minimizing unnecessary randomness and focusing on meaningful action variations.

FPDS-DDPG algorithm framework, illustrated in Fig. 3, employs an actor-critic architecture. The actor network determines the best action to take given the current state, while the critic network evaluates the action by estimating the Q-values, which indicate the quality of the action taken in a given state. There are two target networks, which are copies of the actor and critic networks, with delayed updates to slowly track the learned actor and critic networks, thereby stabilizing training. The agent learns from a replay buffer containing experiences $< s_t, a_t, r_t, s_{t+1} >$, which are divided into mini batches of size $\mathfrak{B}$, i.e., $< s_i, a_i, r_i, s_{i+1} >$, $i \in \mathfrak{B}$. This approach ensures stable and effective learning by breaking correlations between experiences and enabling efficient use of computational resources.

Let's define the PDS as the intermediate state immediately after the agent takes action a_t at the current state s_t and before

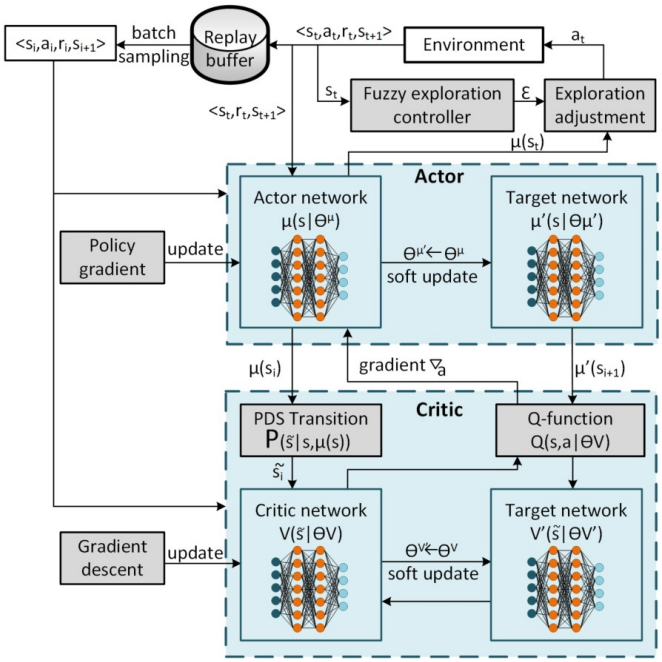


Fig. 3: Proposed FPDS-DDPG algorithm framework.

the transition to the next state s_{t+1}. The PDS at time step t, denoted as $\tilde{s}_t$, encapsulates all known information about the transition from the current state to the next state. The PDS can be defined as

$$\tilde{s}_t \triangleq \left\{\mathbf{h}_t, \mathbf{\Delta f}_{D,t}, R_{L,t}, \mathbf{P}_{L,t}, r_{t-1}\right\}. \tag{25}$$

In the current state, when the agent takes action, it receives a known reward r_t^k. The agent then transitions to the PDS $\tilde{s}_t$ and subsequently moves to the next state, receiving an unknown reward r_t^u. The state transition probability is

$$\mathbb{P}\left(s_{t+1}|s_t, \mu(s_t)\right) = \sum_{\tilde{s}_t \in \mathcal{S}} \mathbb{P}^k\left(\tilde{s}_t|s_t, \mu(s_t)\right) \mathbb{P}^u\left(s_{t+1}|\tilde{s}_t, \mu(s_t)\right), \tag{26}$$

where μ is the function that deterministically maps states to actions representing the actor network, $\mathbb{P}^k\left(\tilde{s}_t|s_t, \mu(s_t)\right)$ denotes the known transition probability from the current state to PSD with action $\mu(s_t)$ and $\mathbb{P}^u\left(s_{t+1}|\tilde{s}_t, \mu(s_t)\right)$ denotes the unknown transition probability from PDS to the next state with action $\mu(s_t)$. The corresponding reward received after the agent takes action at state s_t can be expressed as

$$r_t\left(s_t, a_t\right) = r_t^k\left(s_t, a_t\right) + \sum_{\tilde{s}_t \in \mathcal{S}} \mathbb{P}^k\left(\tilde{s}_t|s_t, \mu(s_t)\right) r_t^u\left(\tilde{s}_t, a_t\right). \tag{27}$$

The action exploration is controlled by a fuzzy exploration controller that determines the exploration factor, denoted as ε, indicating the degree of action perturbation for exploration. The fuzzy exploration controller receives the current state as input, with fuzzy variables including current CSI representing channel quality, Doppler frequency offset representing the uncertainty due to LEO satellite mobility, and the previous reward representing the effectiveness of the prior action. Fuzzy sets for all input variables and the exploration factor are

categorized as {Low, Medium, High}. Algorithm 1 defines the fuzzy rules governing exploration based on these inputs.

Algorithm 1: Fuzzy rules to govern action exploration.

Rule 1: if $\mathbf{h}_t$ is Low **OR** $\mathbf{\Delta f}_{D,t}$ is High **then**
    ε is High;
Rule 2: if $\mathbf{h}_t$ is Medium **AND** $\mathbf{\Delta f}_{D,t}$ is Medium **then**
    ε is Medium;
Rule 3: if $\mathbf{h}_t$ is High **AND** $\mathbf{\Delta f}_{D,t}$ is Low **then**
    ε is Low;
Rule 4: if r_{t-1} is Low **then**
    ε is High;
Rule 5: if r_{t-1} is Medium **then**
    ε is Medium;
Rule 6: if r_{t-1} is High **then**
    ε is Low;

Algorithm 2: Proposed FPDS-DDPG beam design and power allocation algorithm.

Input : $\mathbf{h}_t, \mathbf{\Delta f}_D$
Output: R_L^*, $\mathbf{P}_L^*$, $\hat{Q}$
Initialize actor and critic networks with random parameters Θ^μ and Θ^V, ϑ_μ and ϑ_V, target networks with parameters $\Theta^{\mu'} \leftarrow \Theta^\mu$ and $\Theta^{V'} \leftarrow \Theta^V$, replay buffer, and fuzzy exploration controller.
for *episode* $ep = 1 : N_{ep}$ **do**
    Receive initial state s_1;
    for *time-step* $t = 1 : T$ **do**
        Compute exploration factor based on s_t using fuzzy logic in **Algorithm 1**;
        Select a_t using the actor network and the exploration factor as in (28);
        Perform a_t and get reward r_t^k;
        Observe the transition from s_t to $\tilde{s}_t$ to s_{t+1};
        Update reward $r_t(s_t, a_t)$ using (27);
        Update PDS Q-value function $V(\tilde{s}_t)$ using (29);
        Update state-action Q-value function $Q(s_t, a_t|\Theta^V)$ using (30);
        Store experiences $< s_t, a_t, r_t, s_{t+1} > \rightarrow$ replay buffer;
        if *replay buffer is full* **then**
            Delete the oldest experiences;
        end if
        Sample a mini batch of experiences $< s_i, a_i, r_i, s_{i+1} >$, $i \in \mathfrak{B}$;
        Compute target values as in (31b);
        Update actor and critic networks using (31);
        Update target networks using (32);
    end for
end for

The fuzzy logic rules dynamically adjust the exploration factor based on channel state quality, Doppler-induced uncertainty, and prior action effectiveness. High Doppler offset or poor channel quality increases uncertainty, necessitating greater exploration to identify better actions. Conversely, stable conditions with low offset and strong channel quality reduce the need for exploration. Additionally, if a prior action was ineffective, higher exploration is encouraged, whereas effective actions favor lower exploration to refine the learned policy. Medium settings balance exploration in moderately uncertain conditions, ensuring efficient learning and improved policy performance.

The action taken by the agent after exploration is expressed as follows:

$$a_t = \mu\left(s_t|\Theta^\mu\right) + \varepsilon\Delta a_t, \tag{28}$$

where $\Delta a_t = 0.01 a_t$ is a predefined perturbation.

In actor-critic based algorithm, the critic network estimates Q-value using a state-action Q-value function $Q(s, a)$ that requires both state and action information. In FPDS-DDPG, on the other hand, the critic network estimates Q-value using a PDS Q-value function $V(\tilde{s}_t)$ that requires only PDS information, making the learning process more efficient. The PDS Q-value function can be written as

$$V(\tilde{s}_t) = r_t^u(\tilde{s}_t, a_t) + \kappa \sum_{s_{t+1}} \mathbb{P}^u(s_{t+1}|\tilde{s}_t, \mu(s_t)) Q(s_{t+1}, \mu(s_{t+1})). \tag{29}$$

The relationship between PDS Q-value function and state-action Q-value function can be expressed as

$$Q(s_t, a_t) = r_t^k(s_t, a_t) + \kappa \sum_{\tilde{s}_t} \mathbb{P}^k(\tilde{s}_t|s_t, a_t) V(\tilde{s}_t). \tag{30}$$

Let Θ^μ and Θ^V represent the parameters of the actor and critic networks, respectively. $\Theta^{\mu'}$ and $\Theta^{V'}$ represent the corresponding target network parameters. The actor and critic network parameters are updated using mini-batch policy gradient and gradient descent, respectively. The update functions are written as

$$\Theta^\mu \leftarrow \Theta^\mu + \frac{\vartheta_\mu}{\mathfrak{B}} \nabla_{a_i} Q\left(s_i, \mu(s_i|\Theta^\mu)|\Theta^V\right) \nabla_{\Theta^\mu} \mu(s_i|\Theta^\mu), \tag{31a}$$

$$\Theta^V \leftarrow \Theta^V - \frac{\vartheta_V}{\mathfrak{B}} \nabla_{\Theta^V} \sum_{i \in \mathfrak{B}} \Bigg[V(\tilde{s}_i|\Theta^V) - \underbrace{\left\{ r_i^k - r_i - \kappa Q\left(s_{i+1}, \mu'(s_{i+1}|\Theta^{\mu'})|\Theta^{V'}\right) \right\}}_{\text{Target values}} \Bigg]^2, \tag{31b}$$

where ϑ_μ and ϑ_V are the learning rates of the actor and critic networks, respectively. ∇ denotes the gradient. The target networks are updated as follows:

$$\Theta^{\mu'} \leftarrow \varpi\Theta^\mu + (1 - \varpi)\Theta^{\mu'}, \tag{32a}$$

$$\Theta^{V'} \leftarrow \varpi\Theta^V + (1 - \varpi)\Theta^{V'}, \tag{32b}$$

where $\varpi \ll 1$ is the target value factor, ensuring slow changes to the target values.

The operation of the proposed FPDS-DDPG-based beam design and power allocation algorithm is described in Algorithm 2.

TABLE I: Computational complexity

Training stage	**DDPG** [16]	$\mathcal{O}\left(N_{ep}T\mathfrak{B}\left(\|\mathcal{S}\|\times\|\mathcal{A}\|+2\left(\|\mathcal{S}\|N_1+\sum_{l=1}^{L_A}N_lN_{l+1}\right)+2\left(\|\mathcal{A}\|N_1+\sum_{l=1}^{L_C}N_lN_{l+1}\right)\right)\right)$
	DPDS-DPG [21]	$\mathcal{O}\left(N_{ep}T\mathfrak{B}\left(2\|\mathcal{S}\|\times\|\mathcal{A}\|+2\left(\|\mathcal{S}\|N_1+\sum_{l=1}^{L_A}N_lN_{l+1}\right)+2\left(\|\mathcal{A}\|N_1+\sum_{l=1}^{L_C}N_lN_{l+1}\right)\right)\right)$
	FPDS-DDPG	$\mathcal{O}\left(N_{ep}T\mathfrak{B}\left(2\|\mathcal{S}\|\times\|\mathcal{A}\|+N_{Fz}+2\left(\|\mathcal{S}\|N_1+\sum_{l=1}^{L_A}N_lN_{l+1}\right)+2\left(\|\mathcal{A}\|N_1+\sum_{l=1}^{L_C}N_lN_{l+1}\right)\right)\right)$
Online operation		$\mathcal{O}\left(2\left(\|\mathcal{S}\|N_1+\sum_{l=1}^{L_A}N_lN_{l+1}\right)+2\left(\|\mathcal{A}\|N_1+\sum_{l=1}^{L_C}N_lN_{l+1}\right)\right)$

D. Computational Complexity Analysis

In this section, we analyze the computational complexity during both the training and online operational stages of the proposed FPDS-DDPG and compare it with the benchmark algorithms, including DPDS-DPG [21] and the conventional DDPG [16]. The training is performed offline on the ground to leverage higher computational resources, while the online operational stage is executed onboard LEO satellites to balance computational efficiency with the resource constraints of LEO satellites.

1) Information required for algorithm execution: The execution of the FPDS-DDPG algorithm during the online operational stage requires specific information that must be acquired in advance. The key data required include the CSI of users and the upper bound Doppler frequency offset.

The CSI acquisition time depends on whether the CSI is estimated onboard or transmitted by users on the ground. For CSI estimation, a typical estimation method may take up to 1 ms, depending on the complexity of the estimation method and the onboard processing capabilities. For CSI feedback from users, the propagation delay for uplink transmission is approximately 1 ms for low-altitude (300 km) LEO satellites and 6.67 ms for high-altitude (2,000 km) LEO satellites.

The upper bound Doppler frequency offset acquisition time is negligible, as it is calculated based on satellite orbital data, which is preloaded onboard the satellites.

2) Computational complexity during training and online stages: Let L_A, L_C, and N_l denote the number of hidden layers in the actor network, the number of hidden layers in the critic network, and the number of neurons in the l-th hidden layer, respectively. During the training stage, the complexity of a single neural network per time step, i.e., the actor network is $\mathcal{O}\left(|\mathcal{S}|N_1+\sum_{l=1}^{l=L_A}N_lN_{l+1}\right)$, where the number of neurons in the input layer of the actor network corresponds to the dimension of the state space. Updating the actor and critic in conventional DDPG per time step results in a complexity of $\mathcal{O}\left(|\mathcal{S}|\times|\mathcal{A}|\right)$, while in DPDS-DPG it results in $\mathcal{O}\left(2|\mathcal{S}|\times|\mathcal{A}|\right)$ because the PDS space has the same dimension as the state space. In the proposed FPDS-DDPG, which incorporates fuzzy action exploration, the complexity of updating the actor and critic is $\mathcal{O}\left(2|\mathcal{S}|\times|\mathcal{A}|+N_{Fz}\right)$, where N_{Fz} represents the number of fuzzy rules evaluated per time step. The computational complexities of the three algorithms are summarized in Table I. Note that the complexity significantly decreases during the online operational stage, as the learning model has already been established after the training stage.

E. Robustness Analysis

To evaluate the robustness of the proposed framework, we define the robustness score $\mathfrak{R}$ as follows:

$$\mathfrak{R}=\frac{1}{N_{test}}\sum_{i=1}^{N_{test}}\left(\frac{\bar{r}_i}{\sigma_{r,i}+\rho}-\epsilon\mathcal{V}_{G,i}\right), \tag{33}$$

where N_{test} is the number of test scenarios with different environmental conditions; $\bar{r}_i$ is the average reward obtained in the i-th test scenario that measures long-term performance; $\sigma_{r,i}$ is the standard deviation of the reward in the i-th test scenario that captures stability with lower values indicate more stable convergence; ρ is a small positive constant to prevent division by zero; $\mathcal{V}_{G,i}=\frac{1}{T'}\sum_{t=1}^{T'}1(\Gamma_{U_G}<\bar{\Gamma}_{U_G})$ with $1(\cdot)$ is an indicator function that equals 1 if the GEO user protection constraint (21c) is violated at time step t, and 0 otherwise; and ϵ is a penalty coefficient that ensures strong enforcement of constraint (21c).

The robustness score evaluates the reliability, adaptability, and safety [23] of the proposed RL framework by considering three key aspects. First, performance stability (σ_r) ensures that the RL policy exhibits low reward variance across multiple runs, leading to consistent and reliable performance. Second, generalization to unseen conditions ($\bar{r}$) measures the algorithm's ability to adapt by assessing the average reward across diverse test environments, ensuring robust operation in dynamic scenarios. Finally, GEO user protection constraint satisfaction ($\mathcal{V}_G$) enforces safety by penalizing frequent violations of the GEO user's SINR constraint, guaranteeing feasibility in a coexisting GEO-LEO network. By integrating these factors, the robustness score provides a comprehensive assessment of the framework's effectiveness in real-world deployment.

IV. Numerical Results

This section evaluates the performance of the proposed FPDS-DDPG and compares it with benchmark algorithms, including DPDS-DPG [21] and the conventional DDPG [16]. The multi-beam GEO-LEO CSN simulator is configured with the key parameters listed in Table II. The region of interest is a circular area on the ground corresponding to a single LEO footprint with a radius of $2,350\ km$. The simulator's origin is set at the center of this region. The GEO satellite is configured with seven beams and utilizes a 4-color frequency

reuse scheme. Within the region of interest, both GEO and LEO users are randomly placed such that each LEO beam contains one GEO user and one LEO user. The channel coefficients between satellites and ground users are randomly generated using Shadowed-Rician fading with parameters $\{m, b, w\} = \{4, 0.126, 0.835\}$ [21]. The Doppler frequency offset is generated based on real-time orbital data from the LEO satellite Iridium-NEXT 914 [24].

TABLE II: Key parameters of the multi-beam GEO-LEO coexisting satellite network environment simulator

Description	Parameter	Value
Carrier frequency (L band)	f_c	1.53 GHz
Speed of light	c	3×10^5 km/s
GEO satellite altitude		$35,786$ km
LEO satellite altitude	A_L	785 km
LEO footprint radius	F_L	$2,350$ km
GEO spot beam radius	R_G	555 km
GEO frequency reuse factor	ι_G	4
GEO bandwidth	B_G	800 kHz
GEO transmit power per beam	P_G	100 W
GEO/LEO antenna diameter		2.4/0.2 m
Antenna efficiency	η	0.55
Noise power	$\sigma^2_{n_G} = \sigma^2_{n_{U_L}}$	-120 dBm/Hz

A. Deep Neural Network Structure

The actor and critic networks are fully connected, each consisting of an input layer, two hidden layers, and an output layer. For the actor network, the input dimension corresponds to the state space and the output dimension to the action space, while for the critic network, the input dimension corresponds to the action space, and the output dimension represents the Q-value function space.

Since the state comprises channel coefficients represented as complex-valued matrices, each complex value is split into its real and imaginary parts, which are fed separately into the network. Notably, the state and action spaces are dynamic, with dimensions that vary over time based on the LEO beam radius. This variability poses a challenge for DRL algorithms, as neural networks typically expect fixed input and output layer sizes. To address this, zero padding is applied to the input and output layers of the networks, setting the state and action space dimensions based on the corresponding $\inf\{R_L\}$ in (13). When the actual state and action space dimensions are smaller, zero padding ensures that the input and output layer dimensions remain fixed. The state space dimension is $|\mathcal{S}| = 6N+2$, while the action state space dimension is $|\mathcal{A}| = N+1$, where N varies accordingly to R_L. The Q-value space dimension is fixed at 1. All hidden layers of the actor and critic networks are identically structured, each with $2^{(6\dot{N}+2-1)-16}$ neurons [21], where $\dot{N}$ corresponds to $\inf\{R_L\}$. Each DNN uses the *tanh* activation function to handle negative inputs from the complex CSIs. The Adam optimizer, with an adaptive learning rate, is applied to both networks. The constraints in (21) are enforced at the output of the actor network, adjusting the agent's decisions on LEO beam radius and transmit power to ensure compliance. Additional hyper-parameters used are listed in Table III.

TABLE III: DRL hyper-parameters

Description	Parameter	Value
Reward discount factor	κ	0.99
Replay buffer size		100000
Mini batch size	$\mathfrak{B}$	64
Number of episode	N_{ep}	2000
Number of time step per episode	T	10000
Learning rate	$\vartheta_\mu = \vartheta_V$	0.001
Learning decaying rate		0.00001

B. DRL Implementation Description

The multi-beam GEO-LEO CSN environment simulator is implemented using MATLAB's Satellite Toolbox, providing CSIs and Doppler frequency offsets as system states for the DRL framework. The DRL framework itself is implemented in PyTorch, chosen for its flexibility and efficiency. Network parameters, architecture, and hyper-parameters are detailed in the previous subsection.

To govern exploration in the proposed FPDS-DDPG framework, we employ a fuzzy logic-based approach, where CSI, Doppler shift, and reward are classified into Low, Medium, and High categories using Gaussian membership functions. Since reward distribution is not predefined, we normalize the reward value based on the agent's observed reward history before applying the Gaussian membership function for classification. The classification follows predefined statistical distributions with means $\mu_L = 0.2$, $\mu_M = 0.5$, and $\mu_H = 0.8$, and a standard deviation of 0.1 to control the spread. These membership functions ensure a smooth transition between categories, allowing adaptive exploration based on dynamic network conditions.

PyTorch facilitates the implementation of DRL algorithms and neural network training, while MATLAB's user-friendly interface complements PyTorch's computational power. This combination streamlines the development and training of the DRL model.

C. Simulation Results

1) DRL learning performance: To evaluate the learning performance of the proposed FPDS-DDPG algorithm, we examine its learning curves alongside those of benchmark algorithms, using the environment simulator configured with the parameters listed in Table II. Figs. 4 and 5 illustrate the average rewards obtained by the learning agent per hundred episodes. It is observed that the proposed FPDS-DDPG algorithm converges more quickly than the benchmark algorithms across all training configurations. Unlike the benchmarks, which use random noise for action exploration, FPDS-DDPG leverages fuzzy action exploration, resulting in smoother, more structured action choices that focus exploration on productive areas of the action space. This approach not only reduces unproductive variability but also enables adaptive exploration intensity based on the agent's familiarity with different regions of the state space. By stabilizing reward signals and minimizing unnecessary noise, the fuzzy exploration in FPDS-DDPG accelerates convergence by making better use of environmental feedback and optimizing sample efficiency.

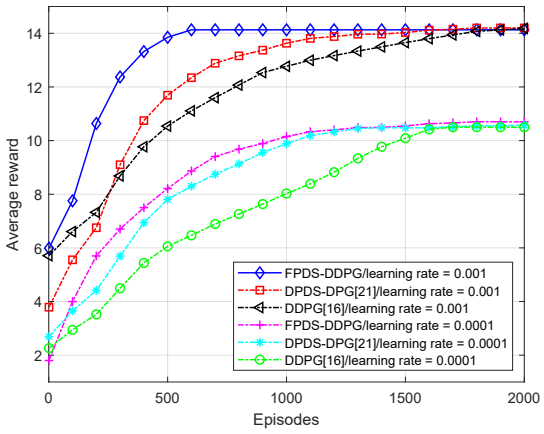


Fig. 4: Average rewards per hundred episodes under different learning rates with decaying rate fixed at 0.00001.

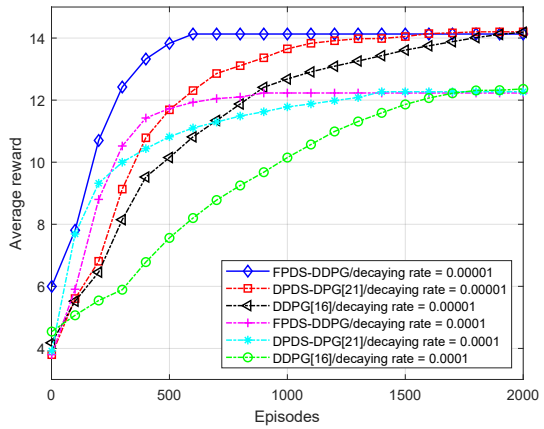


Fig. 5: Average rewards per hundred episodes under different decaying rates with learning rate fixed at 0.001.

Fig. 4 illustrates the impact of learning rate on DRL performance with a fixed decay rate of 0.00001. A smaller learning rate significantly reduces the performance of all DRL algorithms, leading to an average reward decrease of up to 33.3%. In contrast, as shown in Fig. 5, the decay rate has a less pronounced effect on performance. With a fixed learning rate of 0.001, a lower decay rate only reduces the average rewards by a maximum of 15%. From both figures, it is evident that FPDS-DDPG converges after 600 episodes, while DPDS-DPG and DDPG require 1100 and 1700 episodes, respectively, to reach convergence.

2) DRL robustness performance: To assess the robustness of the proposed FPDS-DDPG algorithm, we evaluate its performance across three distinct satellite downlink channel conditions, as characterized in [25]:

- Scenario 1: Average shadowing $(\{m, b, w\} = \{10.1, 0.126, 0.835\})$
- Scenario 2: Infrequent light shadowing $(\{m, b, w\} = \{19.4, 0.158, 1.29\})$
- Scenario 3: Frequent heavy shadowing $(\{m, b, w\} = \{0.739, 0.063, 0.00089\})$

The robustness score, as defined in (33), is computed for each scenario using parameters $\rho = 0.001$ and $\epsilon = 2$. The results are summarized in Table IV.

TABLE IV: Robustness test results

Algorithm		Sce. 1	Sce. 2	Sce. 3	$\Re$
FPDS -DDPG	$\bar{r}$	12.245	12.872	11.903	**139.030**
	σ_r	0.086	0.082	0.097	
	$\mathcal{V}_G$	3/100	3/100	4/100	
DPDS -DPG [21]	$\bar{r}$	12.192	12.767	11.890	**119.162**
	σ_r	0.101	0.103	0.102	
	$\mathcal{V}_G$	4/100	3/100	5/100	
DDPG [16]	$\bar{r}$	12.221	12.795	11.926	**78.127**
	σ_r	0.153	0.169	0.148	
	$\mathcal{V}_G$	4/100	4/100	6/100	

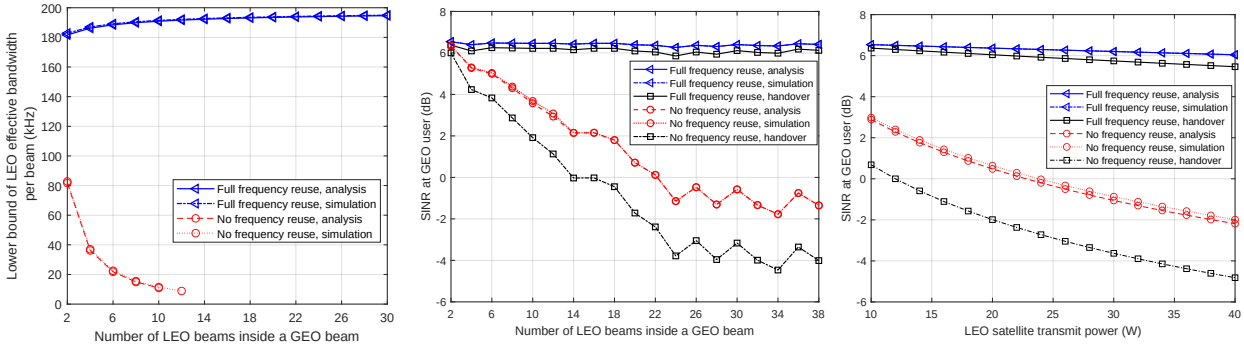
As shown in Table IV, the FPDS-DDPG algorithm achieves the highest robustness score across all test scenarios, demonstrating superior adaptability and stability compared to benchmark algorithms. This indicates its ability to maintain performance under varying channel conditions, reinforcing its robustness in dynamic satellite communication environments.

3) SINR at GEO user: The effect of LEO beam size and transmit power on a GEO user are illustrated in Fig. 6. Specially, the impact of LEO beam size on LEO effective bandwidth and the SINR at a GEO user with a total LEO transmit power budget of $20W$ are showed in Figs. 6a and 6b, respectively. Additionally, the effect of LEO transmit power on the SINR at a GEO user under the configuration of $N_{GL} = 22$ and $R_L = 130.116\,km$ is depicted in Fig. 6c.

In Fig. 6a, the lower bound of LEO effective bandwidth per beam is significantly higher when the LEO satellite employs a full frequency reuse scheme compared to when no frequency reuse scheme is used. With full frequency reuse, the lower bound of effective bandwidth slightly increases as the number of LEO beams within a GEO beam grows. In contrast, without frequency reuse, the lower bound of effective bandwidth decreases significantly as the number of LEO beams inside a GEO beam increases, and a maximum of only twelve LEO beams can be accommodated within a GEO beam. This outcome is expected: without frequency reuse, the bandwidth of a GEO beam is evenly divided among all LEO beams within it, whereas with full frequency reuse, all LEO beams share the entire bandwidth of the GEO beam.

In Fig. 6b, the SINR at the GEO user is significantly higher under full frequency reuse compared to no frequency reuse when there are four or more LEO beams inside a GEO beam. During handover, the SINR is 5% llower than in a typical operational scenario under full frequency reuse and more than 20% lower under no frequency reuse. This reduction is attributed to intersystem interference from two LEO satellites during handover, while in a typical operational scenario, there is interference from only one LEO satellite. Additionally, under the no frequency reuse scheme, when $N_{GL} \geq 24$, the SINR fluctuates, with the radius difference between various N_{GL} cases being only $3\,km$. With $N_{GL} = 2$, the SINR at the GEO user remains nearly the same across all cases.

In Fig. 6c, as the total transmit power budget at the LEO satellite varies, the SINR at the GEO user exhibits similar behavior to that observed in Fig. 6b. Specifically, the SINR under full frequency reuse remains significantly higher than that under no frequency reuse. Additionally, during handover,



(a) LEO effective bandwidth under different LEO beam sizes

(b) SINR per bandwidth under different LEO beam sizes

(c) SINR per bandwidth under different LEO transmit power

Fig. 6: SINR at GEO user under different LEO beam sizes and transmit power considering Doppler frequency offset and frequency reuse schemes during both typical operational scenario and handover situation.

the SINR is lower compared to the typical operational scenario as the total transmit power budget increases.

4) Effect of residual Doppler frequency offset: Figs. 7 and 8 illustrate the interference power experienced by the GEO user and the SINR at the LEO user, respectively, under two Doppler compensation assumptions:

- Practical Doppler compensation assumption: considering residual Doppler frequency offset as in Section II.C.
- Perfect Doppler compensation assumption: not considering residual Doppler frequency offset.

Fig. 7 shows that assuming perfect Doppler compensation results in lower interference power estimation at the GEO user – by 151.19% (in linear scale) for no frequency reuse and 58.49% (in linear scale) for full frequency reuse – compared to the practical Doppler compensation assumption. Similarly, Fig. 8 shows that the perfect Doppler compensation assumption leads to a higher SINR at the LEO user – by 30.32% (in linear scale) for no frequency reuse and 10.87% (in linear scale) for full frequency reuse.

These results demonstrate that assuming perfect Doppler compensation significantly overestimates SINR while underestimating interference power. Moreover, this assumption distorts interference estimation, as Doppler-induced frequency shifts affect different subbands non-uniformly, particularly in frequency reuse scenarios. While a simplified model with perfect Doppler compensation may be computationally appealing for first-order approximations, our findings reveal that it introduces significant inaccuracies. Therefore, incorporating a detailed Doppler model is essential for precise and realistic interference estimation, especially in GEO-LEO coexistence scenarios where accurate interference management is critical.

5) LEO spectral efficiency: The impact of LEO beam size and transmit power on the spectral efficiency of LEO system are illustrated in Figs. 9 and 10. The simulation setup is similar to that in Figs. 6, 7, and 8. In both cases, the LEO spectral efficiency increases as the number of LEO beams within a GEO beam rises and as the total LEO transmit power budget increases. Furthermore, the spectral efficiency under the no frequency reuse scheme is consistently higher than that under the full frequency reuse scheme.

To assess the effectiveness of the proposed learning-based solution, we compare its performance against three non-learning benchmarks:

- Benchmark 1: LEO beam size design based on elevation angle, without accounting for residual Doppler frequency offset, as presented in [26].
- Benchmark 2: LEO beam size design considering residual Doppler frequency offset, as proposed in [27].
- Benchmark 3: Beam power control approach from [28].

Fig. 11 illustrates the spectral efficiency of the LEO system under different beam size optimization strategies as the transmit power varies. The optimal beam radius for the full frequency reuse and no frequency reuse schemes differs slightly. However, both result in the same optimal number of LEO beams within a single GEO beam, denoted as N_{GL}^{*}. The proposed FPDS-DDPG achieves the highest spectral efficiency by jointly optimizing beam size and transmit power. In contrast, Benchmark 1 [26] and Benchmark 2 [27] focus solely on optimizing beam size. Benchmark 2 outperforms Benchmark 1 in spectral efficiency, as it determines the beam size while considering the residual Doppler frequency offset constraint, leading to a more accurate estimation on interference.

Fig. 12 illustrates the spectral efficiency of the LEO system under different transmit power optimization strategies as N_{GL} varies. The proposed approach identifies the optimal $N_{GL}^{*} = 10$. When using this value, the transmit power optimized by Benchmark 3 [28] results in a 24.25% reduction in spectral efficiency compared to the proposed approach.

6) Computational feasibility and latency analysis: To ensure the practicality of deploying the proposed FPDS-DDPG approach on LEO satellites, we analyze its computational efficiency, power consumption, and execution latency. As discussed in Section III.D, full DRL training is conducted offline on the ground, while only inference and lightweight fine-tuning occur onboard the LEO satellite. The onboard inference runs on an AMD Xilinx Zynq UltraScale+ MPSoC with a Quad ARM Cortex-A53 CPU and a deep learning processing unit (DPU) accelerator, consuming 7-10 W of power. This system exemplifies the highly integrated onboard computing platforms used in practice on LEO satellites [29].

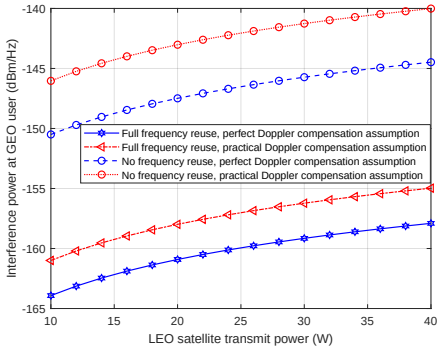


Fig. 7: Effect of residual Doppler on interference power at GEO user.

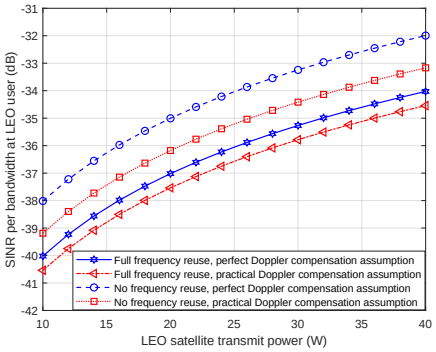


Fig. 8: Effect of residual Doppler on SINR at LEO user.

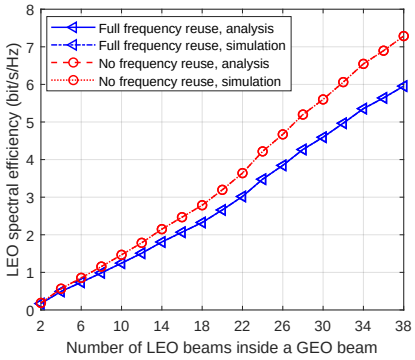


Fig. 9: LEO spectral efficiency under different beam sizes.

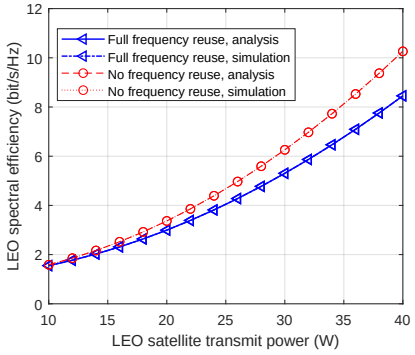


Fig. 10: LEO spectral efficiency under different transmit power.

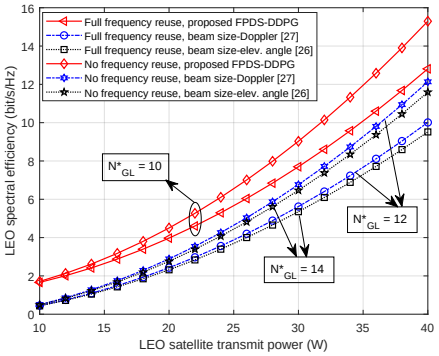


Fig. 11: LEO spectral efficiency under different beam size optimizations.

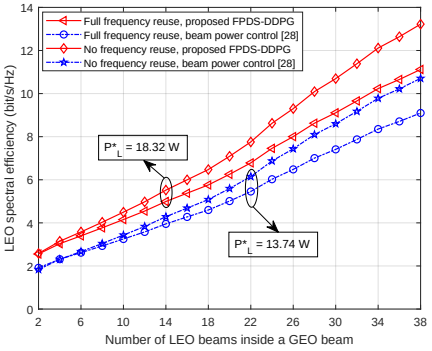


Fig. 12: LEO spectral efficiency under different transmit power optimizations.

The estimated floating-point operations (FLOP) required for a forward pass of the actor and critic networks is

$$C_{FLOP} = \mathcal{O}\Bigg(2(6N+2)N_1 + 2\sum_{l=1}^{L_A} N_l N_{l+1} + 2(N+1)N_1 + 2\sum_{l=1}^{L_C} N_l N_{l+1}\Bigg). \tag{34}$$

With $N = 32$, $L_A = L_C = 2$, $N_l = 128$, FLOP count is $C_{FLOP} = 123904$ FLOPs. If we use CPU at 50 giga FLOP per second (GFLOPS), processing time is $2.48\ \mu s$, while using DPU at 500 GFLOPS results in processing time of $0.248\ \mu s$ per inference iteration. The rerun frequency of the algorithm is dynamically adjusted based on environmental variations, ensuring efficient adaptation to changing spectrum conditions. Furthermore, computational tasks are distributed between ground stations (for offline model training) and LEO satellites (for real-time inference and fine-tuning) to balance power and resource constraints. These results demonstrate that the proposed approach is computationally feasible and meets the real-time operational requirements of LEO satellite networks.

V. Conclusions

This article introduced a novel joint design of LEO beam size and power allocation within a multi-beam GEO-LEO CSN, incorporating crucial factors such as tolerable intersystem interference, Doppler frequency offsets, and frequency reuse to maximize spectral efficiency. As the first work to directly address Doppler frequency offsets in this context, we leveraged a fast-converging DRL algorithm that combines fuzzy logic and PDS techniques with the DDPG approach. Our comprehensive analysis revealed the impact of Doppler offsets on transmission bandwidth and provided a robust framework for quantifying intersystem interference and evaluating LEO spectral efficiency across various scenarios. Simulation results validated the effectiveness and adaptability of our design, paving the way for enhanced performance in future GEO-LEO satellite communications.

References

[1] A. Guidotti, A. Vanelli-Coralli, and et.al., "Role and evolution of non-terrestrial networks toward 6G systems," *IEEE Access*, vol. 12, pp. 55 945–55 963, 2024.

[2] International Telecommunication Union, "Operation of earth stations in motion communicating with geostationary space stations in the fixed-satellite service allocations," 2019, ITU-R.

[3] Amazon, "Amazon marks breakthrough in Project Kuiper development," 2020. [Online]. Available: https://www.aboutamazon.com/news/innovation-at-amazon/amazon-marks-breakthrough-in-project-kuiper-development

[4] Eutelsat, "The world's first satellite operator with an integrated GEO-LEO infrastructure," 2023. [Online]. Available: https://www.eutelsat.com/en/group/about-us.html

[5] Telesat, "Telesat Lightspeed LEO networks," 2018. [Online]. Available: https://www.telesat.com/leo-satellites/

[6] SES, "O3b mPOWER insight paper," 2024.

[7] International Telecommunications Union, "Radio regulations. Chapter VI, Provision of servies and stations. Article 22, Space Services," (2016). ITU-R.

[8] N. Heydarishahreza, T. Han, and N. Ansari, "Spectrum sharing and interference management for 6G LEO satellite-terrestrial network integration," *IEEE Commun. Surveys Tuts.*, pp. 1–1, 2024.

[9] Z. Xiao, Z. Han, A. Nallanathan, O. A. Dobre, B. Clerckx, J. Choi, C. He, and W. Tong, "Guest editorial special issue on antenna array enabled space/air/ground communications and networking," *IEEE J. Select. Areas Commun.*, vol. 40, no. 10, pp. 2767–2772, 2022.

[10] A. Hills, J. M. Peha, and J. Munk, "Feasibility of using beam steering to mitigate Ku-band LEO-to-GEO interference," *IEEE Access*, vol. 10, pp. 74 023–74 032, 2022.

[11] Z. Zheng, W. Jing, Z. Lu, Q. Wu, H. Zhang, and D. Gesbert, "Cooperative multi-satellite and multi-RIS beamforming: Enhancing LEO satcom and mitigating LEO-GEO intersystem interference," *IEEE J. Select. Areas Commun.*, pp. 1–1, 2024.

[12] M. He, G. Cui, M. Wu, and W. Wang, "Collaborative interference avoidance technology in GEO-LEO co-existing satellite system," *Int. J. Satell. Commun. Network.*, vol. 42, no. 4, pp. 257–272, 2022.

[13] D. Yan, Y. He, and H. Fu, "Interference analysis of NGSO constellation to GEO satellite communication system based on spatio–temporal slices," *IEEE Internet Things J.*, vol. 10, no. 18, pp. 16 605–16 616, 2023.

[14] B. Li, J. Park, A. Al-Hourani, S. R. Pokhrel, and J. Choi, "A novel frequency reuse model for co-existing LEO and GEO satellites," *IEEE Wireless Commun. Lett.*, vol. 13, no. 4, pp. 1024–1028, 2024.

[15] Q. T. Ngo, B. Jayawickrama, Y. He, and E. Dutkiewicz, "A novel satellite-based REM construction in cognitive GEO-LEO satellite IoT networks," *IEEE Internet Things J.*, vol. 12, no. 6, pp. 7532–7548, 2025.

[16] T. Lillicrap, J. Hunt, A. Pritzel, and N. Heess, "Continuous control with deep reinforcement learning," in *Proc. Int. Conf. Learning Representations*, 2016, pp. 1–6.

[17] Q. T. Ngo, K. T. Phan, A. Mahmood, and W. Xiang, "Physical layer security in IRS-assisted cache-enabled satellite communication networks," *IEEE Trans. Green Commun. Netw.*, vol. 7, no. 4, pp. 1920–1931, 2023.

[18] Q. T. Ngo, B. Jayawickrama, and et.al., "Optimizing spectrum sensing in cognitive GEO-LEO satellite networks: Overcoming challenges for effective spectrum utilization," *IEEE Veh. Technol. Mag.*, pp. 1–1, 2025.

[19] International Civil Aviation Organization, "Manual for ICAO Aeronautical Mobile Satellite (Route) Service, Part 2-Iridium, Draft v4.0," Mar 2007. [Online]. Available: www.icao.int/ACP-WG-M-Iridium-8

[20] R. H. Y. Perdana, T.-V. Nguyen, and B. An, "Adaptive user pairing in multi-IRS-aided massive MIMO-NOMA networks: Spectral efficiency maximization and deep learning design," *IEEE Trans. Commun.*, vol. 71, no. 7, pp. 4377–4390, 2023.

[21] Q. T. Ngo, K. Phan, A. Mahmood, and W. Xiang, "Hybrid IRS-assisted secure satellite downlink communications: A fast deep reinforcement learning approach," *IEEE Trans. Emerg. Top. Comput. Intell.*, vol. 8, no. 4, pp. 2858–2869, 2024.

[22] Q. T. Ngo, B. A. Jayawickrama, Y. He, and E. Dutkiewicz, "Multi-agent DRL-based RIS-assisted spectrum sensing in cognitive satellite–terrestrial networks," *IEEE Wireless Commun. Lett.*, vol. 12, no. 12, pp. 2213–2217, 2023.

[23] R. R. O. Al-Nima, T. Han, S. A. M. Al-Sumaidaee, T. Chen, and W. L. Woo, "Robustness and performance of deep reinforcement learning," *Applied Soft Computing*, vol. 105, p. 107295, 2021.

[24] K. Weththasinghe, Q. T. Ngo, Y. He, and B. Jayawickrama, "Optimising beam size in multibeam LEO satellite networks: Addressing interbeam interference, doppler shift, and frequency reuse," *IEEE Trans. Aerospace and Electronics Syst.*, pp. 1–15, 2024.

[25] A. Abdi, W. C. Lau, M. Alouini, and M. Kaveh, "A new simple model for land mobile satellite channels: first- and second-order statistics," *IEEE Trans. Wireless Commun.*, vol. 2, no. 3, pp. 519–528, May 2003.

[26] S. Han, W. Shin, and J.-H. Kim, "Evaluation for elevation angle-dependent beam size in NR NTN systems," in *Proc. Int. Conf. Information Commun. Technol. Convergence (ICTC)*, 2023, pp. 321–323.

[27] S. Han, W. Lee, W. Shin, and J.-H. Kim, "Adaptive beam size design for leo satellite networks with doppler shift compensation," in *Proc. IEEE VTS Asia Pacific Wireless Commun. Symposium (APWCS)*, 2022, pp. 26–30.

[28] P. Gu, R. Li, C. Hua, and R. Tafazolli, "Dynamic cooperative spectrum sharing in a multi-beam LEO-GEO co-existing satellite system," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1170–1182, 2022.

[29] NASA, "State-of-the-art of small spacecraft technology," March 2025. [Online]. Available: https://www.nasa.gov/smallsat-institute/sst-soa/small-spacecraft-avionics/.

QUYNH TU NGO (Senior Member, IEEE) received a B.Sc. in Electrical Engineering (Magna Cum Laude) from California State University Los Angeles, USA, in 2013; an M.Sc. in Telecommunications from Vietnam National University - University of Sciences, Vietnam, in 2016; and a Ph.D. in Computer Science from La Trobe University, Australia, in 2023. She is currently a Postdoctoral Research Fellow at the School of Electrical and Data Engineering, University of Technology Sydney, Australia. Her research interests include satellite communications, IoT networks, intelligent non-terrestrial networks, and machine learning in wireless communications and networking.

YING HE (Senior Member, IEEE) received the B.Eng. degree in Telecommunications Engineering from Beijing University of Posts and Telecommunications, China, in 2009, and the Ph.D. degree in Telecommunications Engineering from the University of Technology Sydney, Australia, in 2017. She is currently a Senior Lecturer with the School of Electrical and Data Engineering, University of Technology Sydney. Her research interests are physical layer algorithms in wireless communication with machine learning, vehicular communication, spectrum sharing and satellite communication.

BEESHANGA JAYAWICKRAMA (Senior Member, IEEE) received the B.E. degree (Hons. I) in Telecommunications Engineering and the Ph.D. degree in Electronic Engineering from Macquarie University, Australia, in 2011 and 2015, respectively. He is currently affiliated as a Visiting Fellow at University of Technology Sydney, Australia. He was extensively involved in spectrum sensing and interference mitigation research for spectrum access systems. His research interests include non-terrestrial networks, 5G/6G, cognitive radio, and signal processing.

ERYK DUTKIEWICZ (Senior Member, IEEE) received his B.E. degree in Electrical and Electronic Engineering in 1988 and his M.Sc. degree in Applied Mathematics in 1992 from the University of Adelaide, Australia, and his Ph.D. in Telecommunications from the University of Wollongong, Australia, in 1996. His industry experience includes management of the Wireless Research Laboratory at Motorola in early 2000's. Prof. Dutkiewicz is currently Associate Dean International in the Faculty of Engineering and IT at University of Technology Sydney, Australia. He also holds a professorial appointment at Hokkaido University in Japan. His current research interests cover 5G/6G and IoT networks.