This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3520796

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Secure Satellite Downlink with Hybrid RIS and AI-Based Optimization

Mian Muhammad Kamal[1*], Syed Zain Ul Abideen[2*], Syed Sikandar Shah[3], Nabila Sehito[4], Salahuddin Khan[5], Bal S. Virdee[6], Mohammad Alibakhshikenari[7*], Patrizia Livreri[8]

[1]Joint International Research Laboratory of Information Display and Visualization, School of Electronic Science and Engineering, Southeast University, Nanjing 210018, China. (Email : mmkamal@seu.edu.cn)
[2]College of Computer Science and Technology, Qingdao University, Qingdao 266071, China. (Email: zain208shah@gmail.com)
[3]School of optical Electrical and Computer Science Engineering, University of Shanghai Science and Technology (USST). (Email : sikandershah2020@gmail.com)
[4]Department of Computer Science,ILMA university, Karachi. (Email : dr.nabila@ilmauniversity.edu.pk)
[5]College of Engineering, King Saud University, P.O.Box 800, Riyadh 11421, Saudi Arabia (Email: drskhan@ksu.edu.sa)
[6]Center for Communications Technology, London Metropolitan University, London N7 8DB, United Kingdom (Email: b.virdee@londonmet.ac.uk)
[7]Electronics Engineering Department, University of Rome "Tor Vergata", 00133 Rome, Italy (Email: alibakhshikenari@ing.uniroma2.it)
[8]Department of Engineering, University of Palermo, viale delle Scienze BLDG 9, Palermo, IT 90128, Sicily, Italy (Email: patrizia.livreri@unipa.it)

Corresponding author: Syed Zain Ul Abideeen (zain208shah@gmail.com), Mian Muhammad Kamal (mmkamal@seu.edu.cn), , Mohammad Alibakhshikenari (alibakhshikenari@ing.uniroma2.it)

The authors appreciate funding from Researchers Supporting Project number (RSP2024R58), King Saud University, Riyadh, Saudi Arabia.

ABSTRACT In this paper, we explore a secure multiuser multiple-input single-output (MISO) satellite downlink communication system, enhanced by the integration of a hybrid reconfigurable intelligent surface (RIS). The study formulates a robust joint design for satellite and RIS beamforming, aimed at maximizing the secrecy rate of the overall system. Both the active and passive elements of the RIS are optimized, taking into account practical models that reflect real-world constraints, such as outdated channel state information (CSI) and the power consumption of the system. To address the highly complex, dynamic, and multidimensional nature of the beamforming design problem, deep reinforcement learning (DRL) techniques are employed. Simulation results demonstrate the effectiveness of the proposed beamforming strategy, highlighting significant performance improvements when utilizing hybrid-RIS compared to traditional passive RIS solutions in wireless communication systems.

INDEX TERMS Hybrid reconfigurable intelligent surface (RIS), Deep reinforcement learning (DRL), Active and passive RIS elements, Physical Layer Security (PLS).

I. INTRODUCTION

Satellite communication systems are pivotal for enabling global connectivity, particularly in areas where terrestrial infrastructure is either impractical or economically prohibitive [1]. With the surge in demand for high-data-rate services such as video streaming, internet access, and remote sensing, satellite networks are facing increasing pressure to deliver more efficient and secure communication solutions [2], [3]. Traditionally, satellites have relied on single-input single-output (SISO) or multiple-input single-output (MISO) configurations to serve users in various regions. MISO systems, in particular, offer enhanced performance by enabling beamforming, which allows satellites to direct their transmission energy more effectively to specific users, improving signal quality and system capacity [4].

However, as satellite networks evolve to support an ever-growing number of users, new challenges arise, particularly concerning security. The open nature of wireless communication makes satellite links vulnerable to eavesdropping and interception by unauthorized parties. Given the critical importance of satellite communications in sectors such as defense, finance, and disaster management, securing these links has become a top priority. PLS techniques, which exploit the unique characteristics of wireless channels to safeguard transmissions, have emerged as a promising solution to address these concerns [5]. A key metric in PLS is the secrecy rate, which quantifies the maximum rate at which information can be securely transmitted without being intercepted by an eavesdropper [6]. While traditional beamforming techniques can enhance the secrecy rate in MISO systems, they face limitations in dynamic and complex environments, such as satellite downlinks with multiuser scenarios and rapidly changing channel conditions. Furthermore, the presence of channel impairments, such as outdated CSI due to propagation delays, introduces additional challenges in maintaining reliable and secure communications [7]. To overcome these limitations, recent advancements have focused on the integration of RIS into satellite communication systems [8].

RIS technology consists of large surfaces composed of numerous low-cost, passive reflecting elements that can dynamically adjust their reflection coefficients to control the propagation of electromagnetic waves [9], [10]. By intelligently manipulating the incident signals, RIS can enhance the received signal quality, extend coverage, and mitigate interference, all while consuming minimal power. This has made RIS an attractive solution for enhancing wireless communication networks, particularly in challenging environments such as satellite systems, where the distance between the satellite and users is significant, and power consumption is a critical concern [11]. The traditional RIS consists of purely passive elements, which limits its ability to actively control the signal amplitude. To overcome this limitation, hybrid RIS, which incorporates both passive and active elements, has been proposed [12], [13]. The active components of a hybrid RIS can amplify the reflected signals, thereby compensating for the significant path loss inherent in satellite communication links. However, the inclusion of active elements introduces new design challenges, including increased power consumption and hardware complexity, necessitating careful consideration in the overall system design [14].

One of the most critical challenges in deploying RIS in satellite systems is the joint optimization of satellite beamforming and RIS configuration. This requires solving a highly complex, non-convex optimization problem that involves determining both the optimal transmit beamforming vectors at the satellite and the reflection coefficients at the RIS [15], [16]. The goal of this joint design is to maximize the secrecy rate of the system while adhering to practical constraints such as power consumption limits and imperfect CSI. This problem becomes even more challenging when outdated CSI is considered, as is often the case in satellite communications, where large propagation delays can result in mismatches between the actual and estimated channel states [17].

To address this problem, traditional optimization techniques such as convex optimization or iterative algorithms have been employed in the past. However, these approaches often suffer from high computational complexity and may not be suitable for real-time implementation, especially in dynamic environments where the channel conditions change rapidly. As a result, recent research has explored the application of machine learning (ML) techniques, particularly DRL, to solve the joint beamforming problem in RIS-assisted communication systems [18].

DRL a subset of reinforcement learning (RL), uses deep neural networks to approximate optimal policies in environments with high-dimensional state and action spaces [19], [20]. In the context of RIS-assisted satellite communications, DRL can be used to learn an optimal beamforming strategy that adapts to changing channel conditions in real-time. By interacting with the environment and receiving feedback in the form of rewards (e.g., secrecy rate), the DRL agent gradually improves its decision-making process, leading to an efficient and adaptive solution to the joint beamforming problem [21].

In addition to its ability to handle dynamic environments, DRL offers several advantages over traditional optimization methods. First, it can learn directly from data, reducing the need for accurate mathematical models of the channel and system parameters. This is particularly beneficial in satellite communication systems, where the propagation environment is complex and difficult to model accurately. Second, once trained, the DRL model can make decisions in real-time, making it suitable for practical implementations in satellite systems where computational resources and time are often limited [22].

The application of DRL in RIS-assisted satellite systems represents a significant step forward in addressing the challenges associated with secure and efficient multiuser communications. By jointly optimizing the satellite beamforming and RIS configuration, DRL-based techniques can significantly improve the secrecy rate of the system, even in the presence of practical constraints such as outdated CSI and power limitations. Moreover, the integration of hybrid RIS further enhances the system's flexibility and performance, making it a promising technology for future satellite communication networks [23].

The integration of hybrid RIS into satellite communication systems, combined with advanced DRL techniques for beamforming optimization, is expected to play a crucial role in the development of secure, high-performance satellite networks for next-generation applications. As the demand for satellite-based services continues to grow, the ability to ensure secure and reliable communication will be critical in maintaining the integrity of global communication infrastructures [24].

While this work primarily focuses on enhancing PLS by addressing hostile eavesdropping in satellite communications, the proposed hybrid RIS framework and DRL-based optimization techniques can also be extended to mitigate interference caused by legitimate users in multi-satellite or multi-user systems operating on overlapping frequencies. In such scenarios, "passive eavesdroppers," or legitimate users inadvertently causing interference, represent a critical challenge for efficient spectrum management and resource allocation. The integration of hybrid RIS and intelligent optimization methods presents an opportunity to address these challenges, making this approach versatile for both secure communications and interference mitigation. This broader applicability highlights the potential of hybrid RIS-aided systems to advance both the security and efficiency of future satellite networks.

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3520796

A. LITERATURE REVIEW

The integration of RIS into wireless communication systems has garnered significant attention due to its potential to enhance system performance in terms of coverage, capacity, and security [8], [9]. In the context of satellite communications, RIS technology offers promising solutions to overcome challenges such as signal attenuation over long distances and limited beamforming capabilities inherent in satellite systems [11], [17].

Several studies have explored the use of RIS to improve satellite communication systems. In [25], the authors investigate the application of RIS in secure satellite-terrestrial transmissions. They demonstrate that RIS can enhance PLS by intelligently reflecting signals to legitimate users while minimizing leakage to eavesdroppers. Similarly, examines the performance of RIS-assisted satellite communications, showing improvements in signal quality and coverage.

Hybrid RIS, which combines both passive and active elements, has been proposed to overcome the limitations of purely passive RIS by providing additional control over the amplitude and phase of the reflected signals [14], [26]. The active components in a hybrid RIS can amplify the reflected signals, which is particularly beneficial in satellite communications where the path loss is significant. Research on hybrid RIS in satellite communications, however, is still in its infancy, and more studies are needed to fully understand and exploit its potential benefits. Physical layer security is a critical aspect of satellite communications due to the broadcast nature of satellite signals and the increasing concern over eavesdropping and unauthorized access [6], [27]. Traditional methods to enhance security include encryption and secure key distribution, but these can be computationally intensive and may not be suitable for all applications. As an alternative, PLS techniques, such as secure beamforming and artificial noise generation, have been proposed to improve the secrecy rate without relying solely on cryptographic methods [7], [24]. The application of RIS to enhance PLS has been explored in terrestrial communication systems. For instance, [28] proposes a secrecy rate maximization scheme for RIS-assisted multi-antenna communications, demonstrating significant improvements in secure transmission. In [15], the authors address the secrecy rate optimization in an RIS-assisted MISO system, considering the presence of multiple eavesdroppers.

DRL has emerged as a powerful tool for solving complex optimization problems in wireless communications due to its ability to handle high-dimensional and dynamic environments [18], [19]. DRL has been applied to resource allocation, beamforming design, and network management, showing promising results in terms of performance and adaptability. In the context of RIS-assisted systems, DRL has been used to optimize beamforming strategies to enhance system performance [21], [29]. Specifically, in , the authors employ DRL to design [21] beamforming for an RIS-assisted massive MIMO system, achieving energy-efficient resource allocation. In [22], a DRL-based approach is proposed for secure beamforming design in RIS-assisted MISO communication systems, demonstrating improvements in secrecy rate under imperfect CSI.

However, the combination of hybrid RIS, secure satellite downlink communications, and DRL remains relatively unexplored. Most existing studies focus on terrestrial systems or consider only passive RIS. The work in [30] addresses the robust beamforming design for RIS-assisted secure communications with imperfect CSI but does not consider hybrid RIS or satellite systems. Similarly, [11] investigates channel estimation and passive beamforming for RIS-assisted satellite communication but does not address security aspects or the use of DRL.

While most studies in RIS-aided communication systems focus on combating hostile eavesdropping, interference mitigation remains another critical challenge, particularly in multi-satellite systems sharing similar frequency bands [31]. For example, passive eavesdropping, where legitimate users unintentionally interfere with each other, has been addressed in terrestrial communication systems [32], but its application in satellite networks is underexplored. The integration of hybrid RIS, with its ability to amplify and manipulate signals, offers a promising solution for mitigating interference in such scenarios. Moreover, DRL-based techniques, which have been successfully applied in resource allocation and interference management in terrestrial systems, can be adapted to handle these challenges in satellite communications.While most studies in RIS-aided communication systems focus on combating hostile eavesdropping, interference mitigation remains another critical challenge, particularly in multi-satellite systems sharing similar frequency bands [31]. For example, passive eavesdropping, where legitimate users unintentionally interfere with each other, has been addressed in terrestrial communication systems [32], but its application in satellite networks is underexplored. The integration of hybrid RIS, with its ability to amplify and manipulate signals, offers a promising solution for mitigating interference in such scenarios. Moreover, DRL-based techniques, which have been successfully applied in resource allocation and interference management in terrestrial systems, can be adapted to handle these challenges in satellite communications.

In addition to enhancing PLS against eavesdropping, interference mitigation is a key challenge in satellite systems, especially in multi-user and multi-satellite setups with overlapping frequencies. Hybrid RIS enables precise control of signal reflection and amplification, reducing interference while maintaining energy efficiency. DRL-based optimization frameworks excel in adapting beamforming strategies to dynamic environments, improving both security and interference management [33]–[35]. This joint optimization of satellite and RIS configurations makes the proposed system versatile for modern satellite networks with dense user deployments.

TABLE 1. Comparison of Existing Works and Proposed Contribution

Ref	Focus	Methodology	RIS Configuration	Key Contributions
[8]	RIS in wireless systems	General overview of RIS	Not Applicable (N/A)	Enhancing wireless performance for coverage, capacity, and security
[9]	Wireless RIS systems	Enhancing wireless signal propagation	Passive RIS	Improves capacity and security in wireless networks
[11]	Active RIS in satellite systems	Passive RIS beamforming in satellite communication	Passive RIS	Addressing long-distance signal attenuation in satellite systems
[25]	Satellite-terrestrial transmission	PLS using beamforming	Passive RIS	Improving PLS via signal reflection for legitimate users
[26]	Hybrid RIS for wireless systems	Joint control of active and passive RIS elements	Hybrid RIS (Active + Passive)	Optimizes signal reflection in wireless systems
[14]	Reconfigurable hybrid RIS	Hybrid RIS for enhanced security	Hybrid RIS (Active + Passive)	Mitigates path loss and enhances security in terrestrial systems
[27]	PLS in satellite systems	Secure beamforming with artificial noise generation	Not Applicable (N/A)	Enhancing secrecy rate in satellite systems
[28]	Secrecy rate maximization with RIS	Secrecy rate optimization for multi-antenna systems	Passive RIS	Improves secure transmission in RIS-assisted systems
[15]	RIS-assisted MISO communications	Secrecy rate optimization with multiple eavesdroppers	Passive RIS	Handles secrecy optimization for multi-eavesdropper systems
[18]	DRL for wireless systems	Resource allocation with DRL	Not Applicable (N/A)	DRL-based beamforming for resource efficiency in wireless networks
[21]	DRL-based beamforming for RIS	Energy-efficient beamforming design using DRL	Passive RIS	DRL-based optimization for RIS-MIMO systems under resource constraints
Proposed Solution	Secure satellite downlink with hybrid RIS	Joint DRL-based beamforming optimization with DDPG	Hybrid RIS (Active + Passive)	Maximizes secrecy rate and energy efficiency in satellite systems with hybrid RIS

B. MOTIVATION AND CONTRIBUTIONS

The increasing reliance on satellite communication systems for global connectivity has introduced critical challenges, particularly concerning security. Due to the open nature of satellite signals, they are highly susceptible to interception and eavesdropping, making secure transmission a vital concern. PLS has emerged as a promising solution to mitigate these risks, using inherent properties of the communication channel to protect transmissions. Recently, RIS have been employed in wireless systems to enhance PLS by intelligently controlling signal propagation, thus improving security for legitimate users while degrading signal quality for potential eavesdroppers.

Despite the promise of RIS, traditional optimization techniques used for PLS face scalability challenges when applied to large-scale RIS-assisted systems, especially when dealing with the complexity introduced by outdated CSI. These systems are inherently dynamic, requiring more advanced and flexible optimization methods. Additionally, while most current research focuses on passive RIS, there remains significant untapped potential in Hybrid RIS, which integrates both active and passive elements. Hybrid RIS can offer improved performance while maintaining energy efficiency, making it a compelling solution for satellite communications.

This paper makes the following key contributions:

- Innovative System Model: A hybrid RIS-assisted multiuser MISO satellite downlink communication system is proposed, where a portion of the RIS elements are active and capable of signal amplification. This hybrid approach demonstrates significant improvements in secrecy performance over systems that rely solely on passive RIS.
- Joint Beamforming Optimization: A robust joint design for satellite and hybrid RIS beamforming is developed, aimed at maximizing the system's secrecy rate. The design takes into account real-world constraints, such as outdated CSI and power consumption, ensuring the model's practical relevance.
- DRL-Based Optimization Framework: To address the high complexity and dynamic nature of the system, we employ a DRL approach. Specifically, we use the Deep Deterministic Policy Gradient (DDPG) algorithm to efficiently optimize beamforming strategies, overcoming the limitations of traditional optimization methods.
- Extensive Performance Evaluation: Through comprehensive simulations, we evaluate the effectiveness of the proposed DRL-based beamforming approach. The results highlight significant improvements in secrecy rates when using hybrid RIS, demonstrating that incorporating a small fraction of active RIS elements can enhance the security of satellite communication systems over conventional passive RIS configurations.
- Versatility in Application: Although this study focuses on combating hostile eavesdroppers, the proposed framework can also address interference management in multi-satellite or multi-user communication systems. By leveraging hybrid RIS and DRL, interference between legitimate users operating on shared frequencies can be mitigated, paving the way for improved spectrum coexistence and resource utilization.

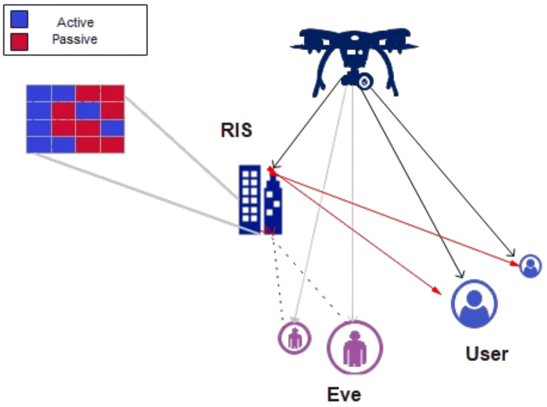


FIGURE 1. Hybrid RIS-assisted secure satellite downlink with deep reinforcement learning for optimized security

In addition to its primary focus on enhancing PLS against hostile eavesdroppers, the proposed hybrid RIS framework offers significant potential for mitigating interference from non-hostile or passive eavesdroppers in multi-satellite or multi-user systems. This contribution expands the applicability of the hybrid RIS to address challenges in spectrum sharing and interference management, making it a versatile solution for future satellite communication networks. These contributions offer a forward-thinking and practical solution for improving the security and efficiency of satellite communications using hybrid RIS technology, pushing the boundaries of current approaches in the field.

II. SYSTEM MODEL

A. NETWORK MODEL

In this work, we consider a hybrid-RIS-aided multiple-input single-output (MISO) satellite communication system. The system architecture comprises a geostationary satellite equipped with multiple antennas that serves a group of legitimate users on the ground, each equipped with a single antenna. Additionally, the system is safeguarded against eavesdropping by using a hybrid RIS to manipulate the propagation environment.The RIS is positioned on the ground between the satellite and users, optimizing signal quality. For multi-beam phased array satellites, a single RIS can support multiple beams by dynamically adjusting its reflection coefficients.

Let S represent the satellite with L_S antennas, which serves U legitimate users denoted by the set $\mathcal{L} = \{1, 2, \ldots, U\}$, each having a single antenna. Furthermore, there are E eavesdroppers, denoted by $\mathcal{E} = \{1, 2, \ldots, E\}$, attempting to intercept the communication. To improve the system's performance and security, a hybrid-RIS, represented by R, with M_R reflecting elements is introduced. Of these M_R elements, M_A are active, and the remaining $M_R - M_A$ are passive.

The hybrid-RIS can manipulate incoming signals by adjusting both their phase and amplitude. Active RIS elements have the ability to amplify signals, while passive elements can only adjust the phase of reflected signals. The configuration of the RIS can be described by the reflection coefficients of its elements, where active elements contribute both phase shifts and amplitude amplification. We represent the position of active elements using the binary vector $\mathbf{a} = \{0, \ldots, 1, \ldots, 0\}$, where $a_i = 1$ indicates that the i-th element is active, and $a_i = 0$ denotes a passive element. The reflection coefficient for each RIS element is denoted by $\gamma_i = (\alpha_i)^{a_i} e^{j\theta_i}$, where $\alpha_i \in (1, \alpha_{\max})$ is the amplitude amplification factor for active elements, and $\theta_i \in [0, 2\pi]$ is the phase shift. For passive elements, we have $\gamma_i = e^{j\theta_i}$.

The RIS interaction matrix is thus defined as:

$$\mathbf{\Phi} = \text{diag}(\gamma_1, \ldots, \gamma_{M_R}), \tag{1}$$

while the matrix for active element interaction is given by:

$$\mathbf{\Psi} = \text{diag}(0, \ldots, \gamma_i, \ldots, 0), \tag{2}$$

where the non-zero entries correspond to the active elements' reflection coefficients.

In this model, the communication between the satellite and ground users is represented by the set of channel coefficients. Let $\mathbf{h}_{SL_u} \in \mathbb{C}^{L_S \times 1}$ represent the channel between the satellite and user u, $\mathbf{h}_{SE_e} \in \mathbb{C}^{L_S \times 1}$ the channel between the satellite and eavesdropper e, and $\mathbf{H}_{SR} \in \mathbb{C}^{M_R \times L_S}$ the channel between the satellite and the RIS. Additionally, the channels between the RIS and user u and eavesdropper e are denoted by $\mathbf{h}_{RL_u} \in \mathbb{C}^{M_R \times 1}$ and $\mathbf{h}_{RE_e} \in \mathbb{C}^{M_R \times 1}$, respectively.

The satellite-to-user and satellite-to-eavesdropper channels are modeled using a Shadowed-Rician fading model, which accounts for the fact that satellite signals experience shadowing and fading as they propagate through space and the atmosphere. On the other hand, the RIS-to-user and RIS-to-eavesdropper channels follow a Rayleigh fading model, assuming the environment around the RIS is rich in scatterers.

The satellite's beamforming matrix is represented by:

$$\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_U] \in \mathbb{C}^{L_S \times U}, \tag{3}$$

where $\mathbf{w}_u \in \mathbb{C}^{L_S \times 1}$ is the beamforming vector for the u-th legitimate user. When the satellite transmits symbol s_u (where $s_u \sim \mathcal{CN}(0, 1)$ represents a complex normal distribution with zero mean and unit variance), the received signal at user u can be expressed as:

$$\begin{aligned} y_u =& (\mathbf{h}_{SL_u}^H + \mathbf{h}_{RL_u}^H \mathbf{\Phi} \mathbf{H}_{SR}) \mathbf{w}_u s_u \\ &+ \sum_{u' \neq u} (\mathbf{h}_{SL_u}^H + \mathbf{h}_{RL_u}^H \mathbf{\Phi} \mathbf{H}_{SR}) \mathbf{w}_{u'} s_{u'} + n_u, \end{aligned} \tag{4}$$

where $(\cdot)^H$ represents the Hermitian transpose, and n_u is the total noise at the user u, including the noise generated by the RIS active elements, denoted as $n_G \in \mathbb{C}^{M_R \times 1}$, and the complex noise at user u, $n_{L_u} \sim \mathcal{CN}(0, \sigma_{L_u}^2)$, with noise power $\sigma_{L_u}^2$.

The signal received by an eavesdropper e, trying to intercept the transmission to user u, is given by:

$$y_e = (\mathbf{h}_{SE_e}^H + \mathbf{h}_{RE_e}^H \mathbf{\Phi} \mathbf{H}_{SR}) \sum_{u \in \mathcal{L}} \mathbf{w}_u s_u + n_e, \tag{5}$$

where $n_e = \mathbf{h}_{RE_e}^H \mathbf{\Psi} n_G + n_{E_e}$, with $n_{E_e} \sim \mathcal{CN}(0, \sigma_{E_e}^2)$ being the noise experienced by the eavesdropper.

In practical systems, the time delay between when the satellite and RIS obtain the CSI and when they transmit signals may result in **outdated CSI**. Since the RIS reflection beamforming heavily depends on the CSI, the outdated CSI model must be considered, which is described as:

$$\mathbf{h}_{qp} = \rho \mathbf{h}'_{qp} + \sqrt{1-\rho^2} \Delta \mathbf{h}_{qp}, \tag{6}$$

where $qp \in \{SL_u, SE_e, RL_u, RE_e\}$, $\rho \in [0, 1]$ is the outdated CSI correlation coefficient, $\mathbf{h}'_{qp}$ is the estimated CSI, and $\Delta \mathbf{h}_{qp}$ represents the CSI error, constrained to an ellipsoid defined by $E_{qp} = \{\Delta \mathbf{h}_{qp} | \Delta \mathbf{h}_{qp}^H Q_{qp} \Delta \mathbf{h}_{qp} \leq 1\}$, where Q_{qp} determines the size and shape of the ellipsoid.

The SINR for user u is:

$$\text{SINR}_u = \frac{|(\mathbf{h}_{SL_u}^H + \mathbf{h}_{RL_u}^H \mathbf{\Phi} \mathbf{H}_{SR}) \mathbf{w}_u|^2}{\begin{array}{c}\sum_{u' \neq u} |(\mathbf{h}_{SL_u}^H + \mathbf{h}_{RL_u}^H \mathbf{\Phi} \mathbf{H}_{SR}) \mathbf{w}_{u'}|^2 \\ + (\mathbf{h}_{RL_u}^H \mathbf{\Psi} \mathbf{\Psi}^H \mathbf{h}_{RL_u}) \sigma_G^2 + \sigma_{L_u}^2\end{array}}. \tag{7}$$

The **achievable rate** for user u is:

$$R_u = \log_2(1 + \text{SINR}_u). \tag{8}$$

For an eavesdropper e attempting to intercept user u's transmission, the SINR is:

$$\text{SINR}_{e,u} = \frac{|(\mathbf{h}_{SE_e}^H + \mathbf{h}_{RE_e}^H \mathbf{\Phi} \mathbf{H}_{SR}) \mathbf{w}_u|^2}{\begin{array}{c}\sum_{u' \neq u} |(\mathbf{h}_{SE_e}^H + \mathbf{h}_{RE_e}^H \mathbf{\Phi} \mathbf{H}_{SR}) \mathbf{w}_{u'}|^2 \\ + (\mathbf{h}_{RE_e}^H \mathbf{\Psi} \mathbf{\Psi}^H \mathbf{h}_{RE_e}) \sigma_G^2 + \sigma_{E_e}^2\end{array}}. \tag{9}$$

The eavesdropper's achievable rate when intercepting user u is:

$$R_{e,u} = \log_2(1 + \text{SINR}_{e,u}). \tag{10}$$

The secrecy rate for the transmission from the satellite to user u is the difference between the user's achievable rate and the maximum achievable rate by any eavesdropper:

$$R_{\text{sec},u} = \max(0, R_u - \max_{e \in \mathcal{E}} R_{e,u}). \tag{11}$$

Finally, the total secrecy rate of the system is:

$$R_{\text{sec}} = \sum_{u \in \mathcal{L}} R_{\text{sec},u}. \tag{12}$$

B. HYBRID-RIS POWER CONSUMPTION MODEL

The total power consumption of the hybrid-RIS consists of the power consumed by both passive and active elements. Passive RIS elements require power for the RF circuitry and control systems, modeled as a constant power consumption P_C per element. Hence, if the RIS contains $M_p = M_R - M_A$ passive elements, the total power consumption for these elements is:

$$P_{R,p} = M_p P_C. \tag{13}$$

Active RIS elements, on the other hand, require additional power for signal amplification. The power needed for signal amplification is proportional to the input signal power and is described as:

$$P_{\text{amp},i} = \frac{\alpha_i^2 P_{\text{in},i}}{\epsilon}, \tag{14}$$

where $P_{\text{in},i}$ is the incoming signal power at the i-th active element, and ϵ is the efficiency of the amplifier. Therefore, the total power consumption of all active elements is:

$$\begin{aligned} P_{R,a} = & M_A P_C + M_A P_A \\ & + \frac{1}{\epsilon} \text{Tr}(\mathbf{\Psi} (\mathbf{H}_{SR} \mathbf{W} \mathbf{W}^H \mathbf{H}_{SR}^H + \sigma_G^2 \mathbf{I}_{M_R}) \mathbf{\Psi}^H). \end{aligned} \tag{15}$$

The hybrid-RIS is subject to a total power budget, $P_{R,\max}$, which must satisfy the following constraint:

$$P_{R,p} + P_{R,a} \leq P_{R,\max}. \tag{16}$$

This power constraint ensures that the system operates within practical power limits while leveraging both passive and active elements of the hybrid-RIS.

The proposed hybrid RIS and DRL framework is applicable to non-geostationary satellites (NGS), which face challenges such as Doppler shifts and frequent beam handovers. DRL's adaptability allows real-time optimization, making it suitable for dynamic NGS environments. Future work will address these specific complexities.

III. PROBLEM FORMULATION AND PROPOSED DRL SOLUTION

A. PROBLEM FORMULATION

The objective of this system is to maximize the overall secrecy rate of the hybrid-RIS-aided MISO satellite downlink communication system. Specifically, the goal is to jointly optimize the beamforming matrix $\mathbf{W}$ at the satellite and the reflection coefficients matrix $\mathbf{\Phi}$ at the RIS, including the active and passive elements. This joint optimization problem is subject to power constraints at both the satellite and the RIS, as well as secrecy rate requirements for the legitimate users.

The overall secrecy rate R_{sec}, which was derived in the system model, can be expressed as:

$$R_{\text{sec}} = \sum_{u \in \mathcal{L}} R_{\text{sec},u} = \sum_{u \in \mathcal{L}} \max(0, R_u - \max_{e \in \mathcal{E}} R_{e,u}), \tag{17}$$

where R_u is the achievable rate for legitimate user u, and $R_{e,u}$ is the achievable rate for eavesdropper e trying to wiretap user u.

The optimization problem can be formulated as follows:

$$\max_{\mathbf{W}, \mathbf{\Phi}}, R_{\text{sec}} = \max_{\mathbf{W}, \mathbf{\Phi}} \sum_{u \in \mathcal{L}} \max(0, R_u - \max_{e \in \mathcal{E}} R_{e,u}) \tag{18}$$

$$R_{\text{sec},u} \geq R_{\text{sec},u}^{\min}, \quad \forall u \in \mathcal{L} \tag{19}$$

This ensures that each legitimate user's secrecy rate meets a minimum threshold for secure communication.

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2024.3520796

$$\mathrm{Tr}(\mathbf{W}^H\mathbf{W}) \leq P_S^{\max} \tag{20}$$

where $P_S^{\max}$ is the maximum power available at the satellite.

$$P_{R,p} + P_{R,a} \leq P_{R,\max} \tag{21}$$

This ensures that the total power consumption of the hybrid-RIS (passive and active elements combined) does not exceed the predefined power budget $P_{R,\max}$.

$$\alpha_i \in (1, \alpha_{\max}), \quad \forall i \in M_A \tag{22}$$

where $\alpha_{\max}$ is the maximum allowable amplitude gain for the active RIS elements.

$$\theta_i \in [0, 2\pi], \quad \forall i \in M_R \tag{23}$$

This constraint ensures that the phase shifts applied by both the active and passive elements of the RIS are within the allowable range.

Worst-case Secrecy Rate Optimization: Since the system operates under outdated CSI, we consider the worst-case scenario to guarantee a robust design. The problem is thus formulated to account for the worst-case secrecy rate. The CSI errors are modeled using the ellipsoid uncertainty model, and the secrecy rate is maximized over the worst-case CSI, i.e.,

$$\max_{\Delta\mathbf{h}SL_u, \Delta\mathbf{h}RL_u} \min_{\Delta\mathbf{h}SE_e, \Delta\mathbf{h}RE_e} \sum_{u\in\mathcal{L}} R_{\text{sec},u} \tag{24}$$

where $\Delta\mathbf{h}_{qp}$ represents the CSI error vectors, which are constrained to lie within the predefined ellipsoid set as detailed in the system model.

Equation (24) ensures secure communication by maximizing the secrecy rate under worst-case CSI conditions. Reformulated as an MDP, the DRL framework iteratively optimizes beamforming and RIS configurations to balance secrecy and power constraints.

B. PROPOSED DRL SOLUTION

The problem of maximizing the secrecy rate in the hybrid-RIS-aided MISO satellite communication system involves highly complex, dynamic, and multidimensional optimization. Traditional optimization techniques struggle to handle the non-convex nature of the joint beamforming design for both the satellite and RIS under practical constraints like outdated CSI and limited power budgets. To overcome these challenges, DRL is leveraged to intelligently learn the optimal beamforming policies by interacting with the system environment over time.

In this approach, the beamforming problem is reformulated as a MDP, where the satellite and RIS act as decision-making agents. The goal of these agents is to learn a policy that maximizes the system-wide secrecy rate by adjusting their beamforming vectors and RIS configurations based on the observed environment (e.g., channel conditions). The process involves the following key components:

State Space The state space $\mathcal{S}$ includes all the information required to describe the environment at any given time. In the context of the hybrid-RIS-aided secure communication system, the state at time step t, denoted by s_t, consists of: CSI for legitimate users and eavesdroppers at time t. Since outdated CSI is considered, the state also includes the estimated and error components of the CSI. Previous beamforming vectors $\mathbf{w}_{u,t-1}$ from the satellite for each legitimate user u. Previous RIS reflection settings, including the amplitude gains $\alpha_{i,t-1}$ and phase shifts $\theta_{i,t-1}$ for each RIS element i.

Thus, the state vector at time t is represented as:

$$s_t = \{\mathbf{h}_t, \mathbf{w}_{u,t-1}, \alpha_{i,t-1}, \theta_{i,t-1}\}, \tag{25}$$

where $\mathbf{h}_t$ denotes the CSI of users and eavesdroppers at time t, including both the legitimate and outdated CSI components.

Action Space The action space $\mathcal{A}$ defines the set of possible actions the agent can take to affect the environment. In this problem, an action at time t, denoted by a_t, involves the following adjustments: Beamforming vector $\mathbf{w}_u$ for each legitimate user u, which dictates how the satellite directs its transmission to that user. Amplitude gain α_i for each active RIS element i. The amplitude gain modifies the strength of the signal reflected by the active RIS elements. Phase shift θ_i for both active and passive RIS elements, which changes the phase of the reflected signals.

Thus, the action at time t is represented as:

$$a_t = \{\mathbf{w}_u, \alpha_i, \theta_i\}, \tag{26}$$

where $\mathbf{w}_u \in \mathbb{C}^{L_S\times 1}$, $\alpha_i \in (1, \alpha_{\max})$, and $\theta_i \in [0, 2\pi]$ for the RIS elements.

The agent selects these actions in order to maximize the long-term secrecy performance of the system.

Reward Function The reward function quantifies the immediate benefit of taking a particular action in a given state. In this problem, the reward at each time step is based on the secrecy rate achieved by the system. After taking action a_t, the environment provides a reward r_t, which reflects how effective the beamforming and RIS configuration are at increasing the secrecy rate. The reward is defined as:

$$r_t = R_{\text{sec},t}, \tag{27}$$

where $R_{\text{sec},t}$ is the overall system secrecy rate at time step t, calculated based on the beamforming vectors and RIS settings chosen in action a_t.

This reward encourages the agent to select actions that improve the secrecy rate for legitimate users while minimizing the information intercepted by eavesdroppers.

Policy The policy π is a strategy that the agent follows to select actions based on the observed state. In the context of reinforcement learning, the policy can be deterministic or stochastic. For this problem, the policy aims to map states to actions in a way that maximizes the accumulated reward over time.

The policy is often parameterized by a neural network π_θ, where θ represents the parameters of the neural network. The agent updates these parameters during training to improve its decision-making capability. The policy takes the current state s_t as input and outputs the action a_t that the agent should take.

Discounted Accumulated Reward The goal of reinforcement learning is to maximize the discounted accumulated reward over time. The discounted reward at time t is given by:

$$R_t = \sum_{i=t}^{\infty} \gamma^{i-t} r_i, \tag{28}$$

where $\gamma \in (0, 1)$ is the discount factor that determines the importance of future rewards. A higher γ places more emphasis on long-term rewards, whereas a lower γ focuses more on immediate rewards.

The agent's objective is to maximize the expected accumulated reward R_t, which corresponds to improving the long-term secrecy rate of the system.

1) Computational Complexity

The computational complexity of the proposed DDPG-based secure beamforming framework for hybrid-RIS-aided MISO satellite communication systems can be analyzed in two distinct phases: the offline training stage and the online deployment stage.

During the offline training stage, the complexity arises primarily from the structure and operations of the actor and critic neural networks. For each training step, the actor network computes the optimal actions, including beamforming vectors, RIS amplitude gains, and phase shifts, while the critic network evaluates the quality of these actions using the state-action value function (Q-value). The computational complexity of the actor network for a forward pass through its layers can be expressed as:

$$\mathcal{O}\left(Z_0^a Z_1 + \sum_{g=1}^{G_a} Z_g Z_{g+1} \right), \tag{29}$$

where Z_0^a represents the number of neurons in the input layer, and Z_g denotes the number of neurons in the g-th hidden layer. Similarly, the critic network's complexity follows the same structure as the actor network but includes additional action evaluation. The complexity of the critic network is given by:

$$\mathcal{O}\left(Z_0^c Z_1 + \sum_{g=1}^{G_c} Z_g Z_{g+1} \right), \tag{30}$$

where Z_0^c represents the number of input neurons in the critic network. Both networks are updated using mini-batches, and the overall training complexity per episode is determined by the number of episodes N_{ep}, time steps per episode T, mini-batch size $|B|$, and the dimensions of the state ($|\mathcal{S}|$) and action spaces ($|\mathcal{A}|$). The total training complexity per episode can be written as:

Algorithm 1 DDPG-based secure beamforming and hybrid-RIS configuration optimization for MISO satellite downlink communications.

Input: Current CSI $s_0 = \{\mathbf{h}_{SL_u}, \mathbf{h}_{SE_e}, \mathbf{H}_{SR}, \mathbf{h}_{RL_u}, \mathbf{h}_{RE_e}, \mathbf{w}_u, \alpha_i, \theta_i\}$, maximum satellite power $P_S^{\max}$, maximum RIS power $P_{R,\max}$

Output: Optimized beamforming vector $\mathbf{w}_u^*$, optimized RIS configuration α_i^*, θ_i^*, maximized Secrecy Rate.

Initialization:

Initialize actor network $\mu(s|\Theta^\mu)$ and critic network $Q(s, a|\Theta^Q)$ with random weights Θ^μ, Θ^Q Initialize target networks μ' and Q' with weights $\Theta^{\mu'} \leftarrow \Theta^\mu$, $\Theta^{Q'} \leftarrow \Theta^Q$ Initialize replay buffer $\mathcal{B}$ Initialize exploration noise $\mathcal{N}$

For each episode:

Initialize state $s_0 = \{\mathbf{h}_{SL_u,0}, \mathbf{h}_{SE_e,0}, \mathbf{H}_{SR,0}, \mathbf{h}_{RL_u,0}, \mathbf{h}_{RE_e,0}, \mathbf{w}_{u,0}, \alpha_{i,0}, \theta_{i,0}\}$

for *each time step* $t = 1$ *to* T_{max} **do**

Select action $a_t = \mu(s_t|\Theta^\mu) + \mathcal{N}_t$, where $a_t = \{\mathbf{w}_{u,t}, \alpha_{i,t}, \theta_{i,t}\}$ Apply action a_t in the environment, observe reward r_t and next state s_{t+1} Store transition (s_t, a_t, r_t, s_{t+1}) in the replay buffer $\mathcal{B}$

if *replay buffer* $\mathcal{B}$ *is sufficiently filled* **then**

Sample a random mini-batch of transitions (s_i, a_i, r_i, s_{i+1}) from the buffer $\mathcal{B}$ Compute target $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1}|\Theta^{\mu'})|\Theta^{Q'})$

Update the critic network by minimizing the loss: $\mathcal{L} = \frac{1}{|B|}\sum_i (Q(s_i, a_i|\Theta^Q) - y_i)^2$ Update actor network using policy gradient:

$\nabla_{\Theta^\mu} J \approx \frac{1}{|B|}\sum_i \nabla_a Q(s_i, a|\Theta^Q)|_{a=\mu(s_i)} \nabla_{\Theta^\mu}\mu(s_i|\Theta^\mu)$

Update target networks:

$\Theta^{Q'} \leftarrow \tau\Theta^Q + (1-\tau)\Theta^{Q'}$

$\Theta^{\mu'} \leftarrow \tau\Theta^\mu + (1-\tau)\Theta^{\mu'}$

end

end

End episode when convergence or maximum time step is reached Return optimized beamforming vector $\mathbf{w}_u^*$ and RIS configuration α_i^*, θ_i^* that maximize the secrecy rate.

$$\mathcal{O}\Big(N_{\text{ep}} T |B| \Big(|\mathcal{S}| \cdot |\mathcal{A}| + 2\left(|\mathcal{S}| Z_1 + \sum_{g=1}^{G_a} Z_g Z_{g+1} \right) + 2\left(|\mathcal{A}| Z_1 + \sum_{g=1}^{G_c} Z_g Z_{g+1} \right) \Big) \Big). \tag{31}$$

The training stage is computationally expensive but is performed offline on ground-based stations or data centers equipped with high-performance computing infrastructure. This ensures that the heavy computational demands of training neural networks, including managing replay buffers and iterative updates, do not burden the satellite system. Conducting training on the ground also allows for iterative refinement and avoids the hardware and energy constraints faced by satellites.

In contrast, the online deployment stage is computationally efficient and involves executing the trained actor network to make real-time decisions for beamforming and RIS configurations. This phase only requires a forward pass through the actor network, with complexity:

$$\mathcal{O}\left(Z_0^a Z_1 + \sum_{g=1}^{L_a} Z_g^a Z_{g+1}^a\right), \tag{32}$$

where Z_0^a and Z_g^a represent the input and hidden layer neurons of the actor network, and L_a is the total number of layers in the network. The reduced computational load during this phase makes it feasible for real-time operation on the satellite.

Online computations can be efficiently performed onboard the satellite using modern AI hardware, such as edge AI processors or low-power GPUs, which are specifically designed for lightweight neural network inference. The reduced complexity ensures that optimal beamforming and RIS configurations can be computed in real-time without significant latency or resource constraints. This approach also eliminates the need for frequent communication with ground stations, enabling the system to dynamically adapt to changing channel conditions and ensure secure, low-latency satellite communications.

IV. SIMULATION RESULTS

Figure 2 and Figure 3 illustrate the effects of varying the learning rate and the decaying rate, respectively, on the average reward over 2000 episodes in a DRL-based hybrid RIS-aided satellite communication system. Figure 2 focuses on the impact of different learning rates, demonstrating that a learning rate of 0.001 achieves the best trade-off between stability and convergence speed, with smoother learning curves and higher average rewards compared to higher rates (e.g., 0.01, which causes instability) and lower rates (e.g., 1e-05, which leads to slower convergence). In contrast, Figure 3 highlights the effect of varying the decaying rate (e.g., exploration decay or learning rate decay), where a decaying rate of 0.01 achieves the highest rewards by optimally balancing exploration and exploitation. Lower decaying rates, such as 1e-05, result in slower learning and underperformance. While both figures evaluate average rewards, they focus on different hyperparameters, underscoring the importance of jointly optimizing learning rate and decaying rate to enhance system performance.

Figure 4 illustrates the secrecy sum-rate (in bits/s/Hz) as a function of the satellite power budget P_s, max (in dBm) for two different configurations of system parameters. The blue curve represents a system with $L = N = K = 4$ and $MM = 102$, while the red curve corresponds to a system with $L = N = K = 10$ and $MM = 204$. As the satellite power budget increases, both configurations demonstrate an improvement in the secrecy sum-rate. The red curve shows that the system with a larger number of antennas, users, and

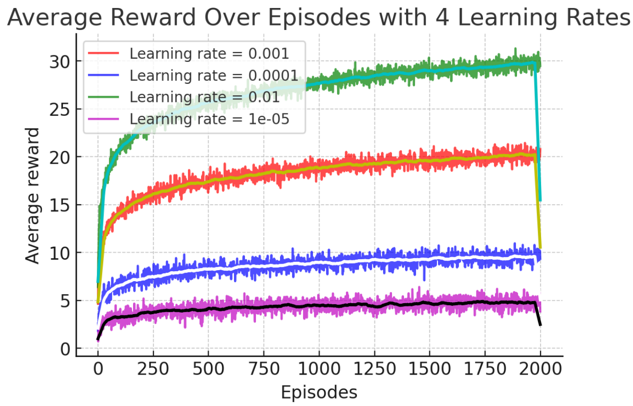


FIGURE 2. Variation in Learning Rate

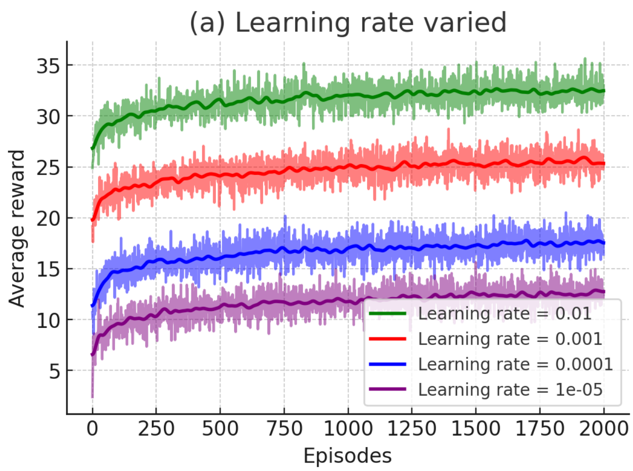


FIGURE 3. Variation in Decaying Rate

reflecting elements achieves significantly higher secrecy rates compared to the blue curve, indicating that increasing these parameters results in improved security performance. Specifically, for P_s, max $= 50\,\text{dBm}$, the secrecy rate for the larger system reaches around 27.5 bits/s/Hz, while the smaller system achieves approximately 20 bits/s/Hz. This demonstrates the effectiveness of scaling up system resources, such as the number of users and reflecting elements, in enhancing secure communication in hybrid RIS-aided satellite systems.

Figure 5 shows the secrecy sum-rate (in bits/s/Hz) as a function of the satellite power budget P_s, max (in dBm) for two different configurations, similar to Figure 4. The blue curve corresponds to a system with $L = N = K = 4$ and $MM = 102$, while the red curve represents a system with $L = N = K = 10$ and $MM = 204$. As in the previous figure, the secrecy sum-rate increases with the satellite power budget for both system configurations. However, the system with larger parameters ($L = N = K = 10$) consistently outperforms the smaller configuration, achieving a secrecy sum-rate of nearly 28 bits/s/Hz at P_s, max $= 50\,\text{dBm}$, compared to around 20 bits/s/Hz for the smaller configuration. This highlights the significant impact of system scaling

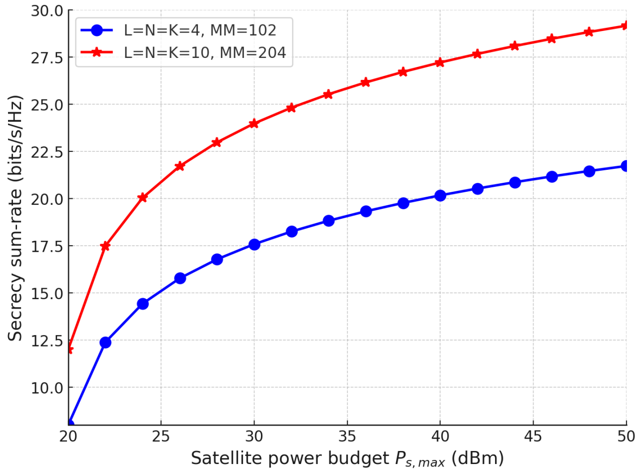


FIGURE 4. Secrecy Performance Across Different Satellite Power Budgets

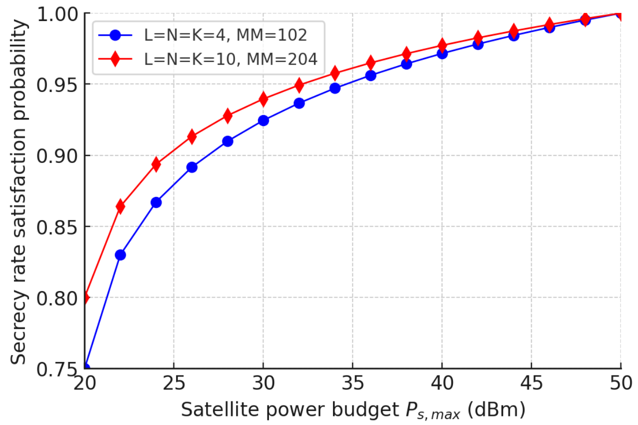


FIGURE 5. Secrecy Rate Satisfaction Probability

(i.e., increasing the number of antennas, users, and reflecting elements) on enhancing the secrecy rate. As the satellite power budget increases, the larger system's ability to secure communications becomes more pronounced, confirming the benefits of scaling resources in secure satellite communication scenarios.

Figure 6 illustrates the secrecy sum-rate as a function of the number of RIS elements for hybrid RIS and passive RIS configurations. The plot indicates that the secrecy sum-rate increases with the number of RIS elements for both setups. The hybrid RIS, with a portion of its elements active, consistently demonstrates superior performance. At 50 RIS elements, the hybrid RIS achieves approximately 60 bits/s/Hz, significantly outperforming the passive RIS, which achieves around 40 bits/s/Hz. The baseline case without RIS remains constant at roughly 25 bits/s/Hz, emphasizing the substantial gains offered by RIS technology, particularly with hybrid configurations that utilize active elements for signal enhancement

The simulation results not only validate the effectiveness of hybrid RIS in enhancing secrecy rates but also suggest

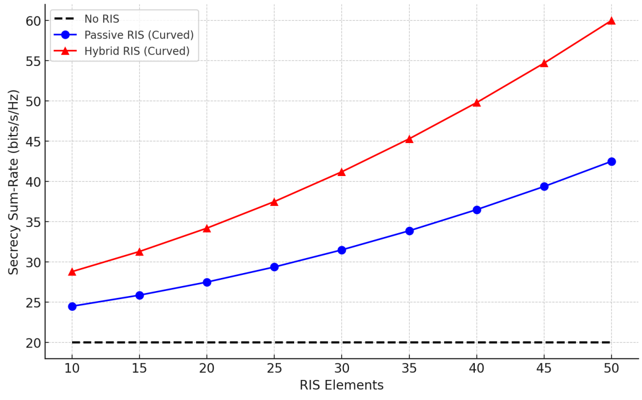


FIGURE 6. Secrecy Performance Across Different RIS Configurations

its potential for mitigating interference in shared-spectrum scenarios. By intelligently controlling signal reflection and amplification, hybrid RIS can suppress interference caused by overlapping communication zones or shared satellite resources. This capability highlights the framework's utility beyond secure communications, particularly in managing non-hostile interference.

V. CONCLUSIONS

This paper investigates the enhancement of PLS in a hybrid-RIS-assisted multiuser MISO satellite downlink communication system. By leveraging hybrid RIS with both active and passive elements, we address the challenges of outdated CSI and power limitations. A novel DRL-based optimization framework using the DDPG algorithm is proposed for joint optimization of satellite and RIS beamforming. Simulation results demonstrate significant improvements in secrecy rates compared to conventional passive RIS systems, confirming the potential of hybrid RIS to strengthen security in satellite communications.

REFERENCES

[1] M. Kang, S. Park, and Y. Lee, "A survey on satellite communication system security," Sensors, vol. 24, no. 9, p. 2897, 2024.

[2] S. Plastras, D. Tsoumatidis, D. N. Skoutas, A. Rouskas, G. Kormentzas, and C. Skianis, "Non-terrestrial networks for energy-efficient connectivity of remote iot devices in the 6g era: A survey," Sensors, vol. 24, no. 4, p. 1227, 2024.

[3] A. O.-O. Esho, T. D. Iluyomade, T. M. Olatunde, and O. P. Igbinenikaro, "A comprehensive review of energy-efficient design in satellite communication systems," International Journal of Engineering Research Updates, vol. 6, no. 02, pp. 013–025, 2024.

[4] L. You, K.-X. Li, J. Wang, X. Gao, X.-G. Xia, and B. Ottersten, "Massive mimo transmission for leo satellite communications," IEEE Journal on Selected Areas in Communications, vol. 38, no. 8, pp. 1851–1865, 2020.

[5] L. Song, Y. Zhang, and X. S. Zhou, Physical Layer Security in Wireless Communications. Taylor & Francis/CRC Press, 2014.

[6] M. Bloch and J. Barros, Physical-layer security: from information theory to security engineering. Cambridge University Press, 2011.

[7] J. Zhang, T. Q. Duong, R. Woods, and A. Marshall, "Securing wireless communications of the internet of things from the physical layer, an overview," Entropy, vol. 19, no. 8, p. 420, 2017.

[8] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," IEEE communications magazine, vol. 58, no. 1, pp. 106–112, 2019.

[9] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," IEEE access, vol. 7, pp. 116 753–116 773, 2019.

[10] M. Zhang, X. Ding, Y. Tang, S. Wu, and K. Xu, "Star-ris assisted secrecy communication with deep reinforcement learning," IEEE Transactions on Green Communications and Networking, 2024.

[11] Y. Ge and J. Fan, "Active reconfigurable intelligent surface assisted secure and robust cooperative beamforming for cognitive satellite-terrestrial networks," IEEE Transactions on Vehicular Technology, vol. 72, no. 3, pp. 4108–4113, 2022.

[12] B. Di, H. Zhang, L. Li, L. Song, Y. Li, and Z. Han, "Practical hybrid beamforming with finite-resolution phase shifters for reconfigurable intelligent surface based multi-user communications," IEEE Transactions on Vehicular Technology, vol. 69, no. 4, pp. 4565–4570, 2020.

[13] Q. Ngo, T. K. Phan, A. Mahmood, and W. Xiang, "Hybrid irs-assisted secure satellite-terrestrial communications: A fast deep reinforcement learning approach," Authorea Preprints, 2023.

[14] S. Zeng, H. Zhang, B. Di, Y. Tan, Z. Han, H. V. Poor, and L. Song, "Reconfigurable intelligent surfaces in 6g: Reflective, transmissive, or both?" IEEE Communications Letters, vol. 25, no. 6, pp. 2063–2067, 2021.

[15] Z. Chu, W. Hao, P. Xiao, D. Mi, Z. Liu, M. Khalily, J. R. Kelly, and A. P. Feresidis, "Secrecy rate optimization for intelligent reflecting surface assisted mimo system," IEEE transactions on information forensics and security, vol. 16, pp. 1655–1669, 2020.

[16] Q. T. Ngo, K. T. Phan, A. Mahmood, and W. Xiang, "Drl-based secure beamforming for hybrid-ris aided satellite downlink communications," in 2022 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT). IEEE, 2022, pp. 432–437.

[17] S. Lin, B. Zheng, G. C. Alexandropoulos, M. Wen, and F. Chen, "Progressive channel estimation and passive beamforming for ris-assisted ofdm systems," in GLOBECOM 2020-2020 IEEE Global Communications Conference. IEEE, 2020, pp. 01–06.

[18] H. Ye and G. Y. Li, "Deep reinforcement learning for resource allocation in v2v communications. in 2018 ieee icc," 2018.

[19] L. Liang, H. Ye, and G. Y. Li, "Spectrum sharing in vehicular networks based on multi-agent reinforcement learning," IEEE Journal on Selected Areas in Communications, vol. 37, no. 10, pp. 2282–2292, 2019.

[20] K. Qi, Q. Wu, P. Fan, N. Cheng, W. Chen, J. Wang, and K. B. Letaief, "Deep-reinforcement-learning-based aoi-aware resource allocation for ris-aided iov networks," arXiv preprint arXiv:2406.11245, 2024.

[21] S. Amin, J. A. Sheikh, B. A. Malik et al., "A deep reinforcement learning for energy efficient resource allocation intelligent reflecting surface (irs) driven non-orthogonal multiple access beamforming (noma-bf)," Physical Communication, vol. 60, p. 102148, 2023.

[22] H. Yang, Z. Xiong, J. Zhao, D. Niyato, L. Xiao, and Q. Wu, "Deep reinforcement learning-based intelligent reflecting surface for secure wireless communications," IEEE Transactions on Wireless Communications, vol. 20, no. 1, pp. 375–388, 2020.

[23] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design," in 2018 IEEE global communications conference (GLOBECOM). IEEE, 2018, pp. 1–6.

[24] M. Cui, G. Zhang, and R. Zhang, "Secure wireless communication via intelligent reflecting surface," IEEE Wireless Communications Letters, vol. 8, no. 5, pp. 1410–1414, 2019.

[25] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," IEEE transactions on wireless communications, vol. 18, no. 8, pp. 4157–4170, 2019.

[26] B. Di, H. Zhang, L. Song, Y. Li, Z. Han, and H. V. Poor, "Hybrid beamforming for reconfigurable intelligent surface based multi-user communications: Achievable rates with limited discrete phase shifts," IEEE Journal on Selected Areas in Communications, vol. 38, no. 8, pp. 1809–1822, 2020.

[27] X. Zhou, L. Song, and Y. Zhang, Physical layer security in wireless communications. Crc Press, 2013.

[28] H. Shen, W. Xu, S. Gong, Z. He, and C. Zhao, "Secrecy rate maximization for intelligent reflecting surface assisted multi-antenna communications," IEEE Communications Letters, vol. 23, no. 9, pp. 1488–1492, 2019.

[29] K. Singh, H. Albinsaid, S. K. Singh, C. Pan, and S. Biswas, "Drl-based beamforming design in ris-aided multi-user wireless networks," in 2023 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS). IEEE, 2023, pp. 224–229.

[30] Q. Zhang, J. Liu, Z. Gao, Z. Li, Z. Peng, Z. Dong, and H. Xu, "Robust beamforming design for ris-aided noma secure networks with transceiver hardware impairments," IEEE Transactions on Communications, vol. 71, no. 6, pp. 3637–3649, 2023.

[31] W. Shi, J. Xu, W. Xu, C. Yuen, A. L. Swindlehurst, and C. Zhao, "On secrecy performance of ris-assisted miso systems over rician channels with spatially random eavesdroppers," IEEE Transactions on Wireless Communications, 2024.

[32] H. Niu, Z. Lin, K. An, X. Liang, Y. Hu, D. Li, and G. Zheng, "Active ris-assisted secure transmission for cognitive satellite terrestrial networks," IEEE Transactions on Vehicular Technology, vol. 72, no. 2, pp. 2609–2614, 2022.

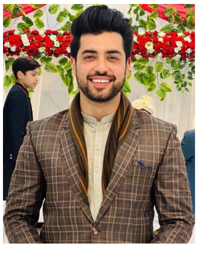
[33] P. Zhang, L. Li, K. Niu, Y. Li, G. Lu, and Z. Wang, "An intelligent wireless transmission toward 6g," Intelligent and Converged Networks, vol. 2, no. 3, pp. 244–257, 2021.

[34] W. Mei, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface-aided wireless networks: From single-reflection to multireflection design and optimization," Proceedings of the IEEE, vol. 110, no. 9, pp. 1380–1400, 2022.

[35] Y. Ju, S. Gong, H. Liu, C. Xing, J. An, and Y. Li, "Beamforming optimization for hybrid active-passive ris assisted wireless communications: A rate-maximization perspective," IEEE Transactions on Communications, 2024.

MIAN MUHAMMAD KAMAL received the M.S. degree from Northwestern Polytechnical University, in 2019, and the Ph.D. degree from the Zhengzhou University in 2023. He was a part of the Next Generation Ubiquitous Network Laboratory, College of Electronic and Information, Northwestern Polytechnical University, China, from Sep 2016 to June 2019. He is currently a Postdoctoral Fellow at Southeast University. His research interests include MIMO antenna systems, smartphone antennas, slot antennas, 5G antennas, millimeter-wave, terahertz components, origami and kirigami antennas, multiband antennas, and metasurfaces.

SYED ZAIN UL ABIDEEN received his M.S. degree in Network Security from the College of Computer Science and Technology, Qingdao University, Qingdao, China. Since 2018, he has been working as a software developer, gaining extensive experience in designing and developing applications across various domains. His professional background complements his academic focus on cutting-edge research areas such as wireless communication, machine learning, deep learning, physical layer security, reconfigurable intelligent surfaces (RIS), and backscatter communication. His combined expertise in software development and advanced communication technologies positions him to contribute significantly to the evolving field of 6G and IoT networks. E-mail: zain208shah@gmail.com



SYED SIKANDAR SHAH received the M.S. degree in Motor and Electrical Appliances from Yangzhou University, China. He is currently pursuing his Ph.D. degree at the School of Optical Electrical and Computer Science Engineering, University of Shanghai for Science and Technology (USST), China. His research interests include motor design, electrical appliances, electrical engineering, and power electronics.

BAL S. VIRDEE)(SM'08) received the B.Sc. and MPhil degrees in Communications Engineering from the University of Leeds, UK, and his Ph.D. in Electronic Engineering from the University of London, UK. He has worked in industry for various companies including Philips (UK) as an R&D-engineer and Teledyne Defence and Space as a future products developer in RF/microwave communications. He has taught at several academic institutions before joining London Metropolitan University where he is a Senior Professor of Communications Technology in the School of Computing and Digital Media where he is Head the Communications Technology Research Center. His research, in collaboration with industry and academia, is in wireless communications encompassing mobile-phones to satellite-technology. Prof. Virdee has chaired technical sessions at IEEE international conferences and published numerous research papers. He is Executive Member of IET's Technical and Professional Network Committee on RF/Microwave-Technology. He is a Fellow of IET and a Senior-Member of IEEE.

MOHAMMAD ALIBAKHSHIKENARI (Member, IEEE) was born in Mazandaran, Iran, in 1988. He received the Ph.D. degree with the Excellent Cum Laude in electronics engineering from the University of Rome "Tor Vergata" (UNITOV), Italy, in 2020. In 2018, he was a Ph.D. visiting researcher at the Chalmers University of Technology, Gothenburg, Sweden, for eight months. His training during this visit included a research stage in the Swedish Company "Gap Waves AB". During his Ph.D. program (Nov. 2016 – Feb. 2020), he has attended thirteen European Doctoral Schools organized by various universities, and he successfully achieved all the credits leading him to obtain his Ph.D. degree with the European label. In November 2019 he was winner of a two-years postdoctoral research fund awarded by the UNITOV. Between July 2021 to August 2024, he was with the Department of Signal Theory and Communications, Universidad Carlos III de Madrid (uc3m), Spain, as a Principal Investigator of the CONEX (CONnecting EXcellence)-Plus Talent Training Program and Marie Skłodowska-Curie Actions. He had two secondments at the (i) Microwave Engineering Center for Space Applications (MECSA), Rome, Italy, between April to August 2024, and (ii) SARAS Technology Ltd Company, Leeds, UK, between December 2022 to May 2023. In addition, he had some research visits at the (i) University of Catania, Italy, in May 2024 along with an invited lecture entitled "Terahertz Antennas based on Metasurface and SIW" for master and PhD students, and postdoctoral researchers; (ii) University of Messina, Italy, in May 2024; (iii) University of Bradford, UK, in May 2023; (iv) Edinburgh Napier University, UK, in April 2023. For academic year 2021–2022, he received the "Teaching Excellent Acknowledgement" certificate for the course of "electromagnetic fields" from the Vice-Rector of studies of uc3m. In September 2024, he joined UNITOV, as a senior researcher. His main research interests are "electromagnetic systems, antennas and wave-propagations, metamaterials and metasurfaces, sensors, synthetic aperture radars (SAR), 5G and beyond wireless communications, multiple input multiple output (MIMO) systems, RFID tag antennas, substrate integrated waveguides (SIWs), impedance matching circuits, microwave components, millimeter-waves and terahertz integrated circuits, gap waveguide technology, beamforming matrix, and reconfigurable intelligent surfaces (RIS)", leading him to publish eight book chapters; more than 150 research articles in peer-reviewed scientific journals; and more than 100 research papers in peer-reviewed conference proceedings including 47 in-person presentations in 29 conferences and 26 on-line presentations in 12 conferences. His research works received more than 6600 citations with H-index above 50 reported by Scopus, Google Scholar, and ResearchGate. According to Stanford University yearly analysis, since 2020 he is among the World's Top 2% Scientists of highly cited scientific authors. He is listed in both Career Long and Single Year Impact in Stanford's list. He was a recipient of the two Young Engineer Awards of the 47th and 48th European Microwave Conferences held in Nuremberg, Germany, in 2017, and in Madrid, Spain, in 2018, respectively. In April 2020, his research article entitled "High-Gain Metasurface in Polyimide On-Chip Antenna Based on CRLH-TL for Sub Terahertz Integrated Circuits, Scientific Reports, volume 10, Article number 4298 (2020)" was awarded as the Best Month Paper at the University of Bradford, UK. He is the editor of the book project entitled "Ultra-Wideband Technologies - Diverse Techniques and Applications" to be published by IntechOpen. In addition, he serves the role of an Associate Editor for two scientific journals of (i) Radio Science, and (ii) IET Journal of Engineering. He acts as a referee in several journals and session chair in several international conferences. In addition, he is member of the reviewer panel for (i) The Dutch Research Council (NWO); (ii) The UK Research and Innovation (UKRI) Funding Service; and an External examiner of PhD dissertations for various universities.

PATRIZIA LIVRERI (M'90, SM'2023) PhD, is a Professor with the Department of Engineering, University of Palermo, and a Visiting Professor with the San Diego State University. She received her "Laurea degree" in Electronics Engineering with honors in 1986 and her Ph.D. in Electronics and Communications Engineering in 1992, both from the University of Palermo, Italy. From 1993 to 1994, she was a researcher at CNR. Since 1995, she has been serving as the scientific director for the "Microwave Instruments and Measurements Lab" of the Engineering Department at the University of Palermo. In 2020, she also joined the CNIT National Laboratory for Radar and Surveillance Systems RaSS in Pisa. Her research interests are in microwave and millimeter vacuum high power (TWT, Klystron) and solid-state power amplifiers for radar applications; high power microwave source (virtual cathode oscillator, magnetically insulated transmission line oscillator); microwave and optical antennas, radar, and microwave quantum radar. She is the principal investigator of the "Microwave Quantum Radar" project, funded by the Ministry of Defense in 2021. She is the supervisor of many funded project and the author of more than 200 published papers.