sensors

MDPI

Article

Reinforcement Learning for Secure Semantic LEO Satellite Networks: Joint Fidelity-Secrecy Power Allocation

Feifei Zhou [1,2,*] and Xiaorong Zhu [1]

[1] College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China; xrzhu@njupt.edu.cn
[2] State Grid Electric Power Research Institute, Nari Group Co., Ltd., Nanjing 211106, China
* Correspondence: 2025011319@njupt.edu.cn

Abstract

Semantic communications have emerged as a key paradigm for intelligent sixth-generation (6G) wireless networks, which aim to convey the meaning of information rather than accurate bit sequences. However, in open-space low Earth orbit (LEO) satellite links, the broadcast nature and wide beam coverage expose semantic transmissions to severe eavesdropping risks. This paper establishes a unified theoretical and algorithmic framework for secure semantic downlink transmission in satellite networks. In particular, we first develop an integrated mathematical model that couples the semantic representation process, physical-layer satellite propagation characteristics, and information-theoretic secrecy into a single analytical formulation. By defining a joint semantic security cost function, the antagonistic trade-off between semantic fidelity and secrecy capacity is quantitatively characterized under realistic power, beamforming, and propagation constraints. To balance semantic fidelity and information secrecy, a reinforcement-learning-based optimization framework is proposed, wherein an actor–critic agent learns optimal power allocation and semantic weighting strategies through continuous interaction with the environment. This learning-based optimization approach enables autonomous control without requiring explicit channel distribution knowledge or offline parameter tuning. Extended simulation results show that the proposed approach consistently enhances both semantic fidelity and secrecy performance compared with conventional power-control schemes and demonstrate its potential as a foundational architecture for secure and intelligent semantic communications in next-generation satellite networks.

Keywords: semantic communications; LEO satellite networks; physical-layer security; semantic fidelity; reinforcement learning

Check for updates

Academic Editor: Luca Leonardi

Received: 21 January 2026
Revised: 26 February 2026
Accepted: 7 March 2026
Published: 21 April 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

1.1. Background

The evolution toward sixth-generation (6G) wireless networks envisions a seamless integration of terrestrial and non-terrestrial infrastructures, thereby forming satellite–terrestrial integrated networks (STINs) that provide global, intelligent connectivity [1–3]. In such systems, satellites play a pivotal role in extending broadband coverage to remote and maritime regions, which can support emergency communications [4] and enable massive Internet-of-things (IoT) services [5–7]. However, the vast data volumes generated by heterogeneous sources—including sensory streams, multimedia content, and situational awareness information—create severe transmission burdens on bandwidth-limited

satellite links [8–10]. Conventional communication systems designed under Shannon's information theory primarily aim to minimize bit error rate (BER) or maximize spectral efficiency, regardless of the underlying meaning of the transmitted data [11–13]. This bit-level paradigm leads to massive redundancy, especially when the receiver's task is semantic, e.g., classification, inference, or description [14].

Semantic communication has recently emerged as a transformative paradigm that redefines the objective of communication systems from bit-level accuracy to meaning-level effectiveness [15–17]. This paradigm originates from Weaver's extension of Shannon theory and has been realized by advances in deep learning and natural language processing, where semantic encoders and decoders can learn compact latent representations that capture the essential content of raw messages [18–20]. By mapping high-dimensional data into task-relevant semantic embeddings, semantic communication systems can significantly reduce transmission redundancy, improve robustness against channel impairments, and enable goal-oriented communication, such as inference, classification, or decision making, directly at the receiver side [21–23]. Recent studies [24–26] have demonstrated that semantic encoding can achieve substantial gains in bandwidth efficiency and task accuracy, particularly under resource-constrained or noisy environments, thus marking it as a key enabler for intelligent 6G communication networks. Nevertheless, the open and broadcast nature of satellite downlink channels introduces inherent security vulnerabilities [27–29]. Due to the long-distance propagation and wide beam coverage, malicious eavesdroppers can easily intercept semantic signals, potentially inferring sensitive information even from partially decoded representations [30–32]. Therefore, future non-terrestrial networks must simultaneously achieve semantic efficiency and information secrecy to ensure that the extracted meaning remains accurate for legitimate users while being unintelligible to adversaries in an efficient manner. This dual requirement motivates the need for a unified semantic security framework that holistically models, analyzes, and optimizes the trade-off between fidelity and secrecy under realistic satellite constraints.

1.2. Related Works

LEO satellite networks are envisioned as key enablers of STINs for 6G systems, which provide global coverage, broadband backhaul, and ubiquitous connectivity for maritime, remote, and aerial users. Despite these advantages, LEO channels exhibit distinct characteristics that make efficient and secure communication challenging [2]. The stringent spectrum and power budgets onboard satellites limit data throughput, while orbital motion induces rapid channel variations and Doppler shifts, complicating channel estimation and synchronization [33]. In [34], the authors proposed a joint optimization framework for beam direction and multi-dimensional resource allocation in dynamic multi-beam LEO satellite networks to maximize long-term user sum rate with fairness. In [35], the authors proposed joint task offloading and resource allocation in satellite mobile edge computing systems with heterogeneous IoT task demands to improve user utility. Moreover, the extremely wide beam coverage of LEO downlinks creates large exposure regions, thus allowing unintended receivers or adversaries to intercept transmissions easily [36]. Traditional Shannon-based communication designs, which focus on minimizing symbol error rate or maximizing bit-level capacity, are inherently inefficient for such settings, since they ignore the fact that transmitted bits could be semantically redundant with respect to the users' tasks [37]. To address these limitations, semantic communication has emerged as a promising paradigm that shifts the objective from reliable bit delivery to effective meaning transmission. Recent studies such as [15,38] have proposed end-to-end neural architectures that learn semantic encoders and decoders to capture task-relevant latent representations, and achieved significant gains in compression efficiency, robustness, and

task accuracy compared with conventional systems. In [11], the authors propose a novel semantic quantification methods over bits for more efficient semantic transmission. These approaches [11,15,38] demonstrate that by transmitting only the semantically significant components of information, communication systems can significantly reduce bandwidth and energy consumption. However, most of these semantic communication works are limited to terrestrial environments, e.g., cellular or Wi-Fi networks [15], where channels are relatively stable and local interference dominates. When applied to LEO satellite downlinks characterized by strong line-of-sight (LoS) propagation and high mobility, semantic systems face additional challenges, including long propagation delays [39], beam misalignment [40], and increased vulnerability [41] to signal interception over large geographic areas.

The open and broadcast nature of satellite channels raises critical concerns regarding privacy protection and information confidentiality. In this context, a great amount of attention has been paid to satellite physical-layer security (PLS) [42], where techniques such as intelligent reflective surface (IRS) [43], artificial noise (AN) injection [44], cooperative jamming [45], and directional beamforming [46] have been developed to create signal-to-noise ratio (SNR) advantages for legitimate users over eavesdroppers. These methods [44–46] effectively suppress interception at the symbol level and improve secrecy capacity under Rician or LoS fading conditions. Nevertheless, conventional PLS frameworks operate strictly within the Shannon bit-pipe abstraction, optimizing secrecy rate metrics without considering the semantic meaning or task relevance of the transmitted information. In [47], the authors considered the semantic integrity the protection of semantic-level semantics and proposed a novel cross-layer semantic security metric. However, the intersection between semantic communication and satellite PLS introduces new and intricate research questions. Semantic representations often encode high-level contextual and relational knowledge. Thus, even partially decoded semantic features can leak sensitive information about the original content or task [48]. Conversely, excessive artificial noise or aggressive power randomization designed to improve secrecy may distort the latent semantic space and degrade task performance. This mutual interference between semantic fidelity and secrecy makes existing separate treatments of semantic efficiency and physical-layer confidentiality inadequate for practical satellite systems. Hence, a unified perspective is therefore required to understand and optimize the trade-off between semantic distortion and secrecy risk.

Several recent works have started to bridge semantic communications with emerging satellite/NTN architectures and resource allocation paradigms. The authors in [40] investigate a SAGIN setting and propose a DRL-based resource allocation framework for hybrid bit-level and generative semantic image delivery. It highlights the role of learning-based policies in dynamic multi-layer space–air–ground environments. The authors in [49] study semantic communication in a satellite-borne edge cloud network for computation offloading, where semantic coding and learning-based updating mechanisms are leveraged to reduce latency/energy under resource constraints. From a system architecture perspective, the authors in [50] discuss self-organized network principles for 5G mega-constellations, providing a forward-looking view on how large-scale LEO systems may support advanced service orchestration. Moreover, computation offloading in multi-layer/multi-constellation LEO environments [51] has been considered, where hierarchical learning and network selection are used to optimize end-to-end latency and reliability. Compared with these works, our paper focuses on physical-layer security for semantic communication and develops an actor–critic RL approach to jointly optimize transmit power and semantic power allocation under a wiretap model, explicitly characterizing the semantic fidelity–secrecy capacity tradeoff.

1.3. Contributions

Motivated by the above-mentioned challenges, this work develops an integrated framework that couples semantic representation, satellite propagation modeling, and AN-aided secure transmission within a joint optimization structure. The proposed formulation explicitly quantifies how semantic fidelity and secrecy capacity depend on physical control variables, namely transmit power and semantic-to-noise power ratio, and further introduces a reinforcement learning (RL)-based optimizer that adapts these parameters online under dynamic LEO channel conditions. In doing so, this study bridges the gap between semantic efficiency and physical-layer security, thus offering a new direction for secure and intelligent semantic communications in next-generation satellite networks. The main contributions of this paper are summarized as follows:

- We establish a mathematical model that integrates semantic representation, satellite propagation characteristics, and artificial-noise-aided transmission, which provides a foundation for jointly analyzing semantic and security performance. A novel cost function is formulated to couple semantic distortion, secrecy leakage, and power efficiency, thereby enabling quantitative evaluation of the fidelity–secrecy trade-off under realistic energy and beamforming constraints.
- Based on the derived SINR expressions, the optimal semantic power allocation and total transmit power are analytically characterized, revealing the structural dependence between semantic accuracy and secrecy robustness in LEO channels.
- We develop an actor–critic-based RL approach to autonomously learn near-optimal transmission strategies through online interaction for a dynamic balance between semantic fidelity and physical-layer security.
- Extended simulation results show that the proposed approach consistently enhances both semantic fidelity and secrecy performance compared with conventional power-control schemes.

2. Satellite Downlink Channel and Transmission Model

We consider a low Earth orbit (LEO) satellite S equipped with an N_t-element phased array, transmitting downlink signals to a legitimate ground terminal U in the presence of a passive eavesdropper E, as shown in Figure 1. The system operates in the Ka-band, where propagation is dominated by a near line-of-sight (LoS) component and weak tropospheric scattering. The downlink configuration is adopted since the satellite possesses higher transmission power, centralized control, and beamforming capability, whereas the ground terminals have limited energy and primarily perform semantic reconstruction. This asymmetry also highlights the need for effective AN injection to enhance physical-layer secrecy.

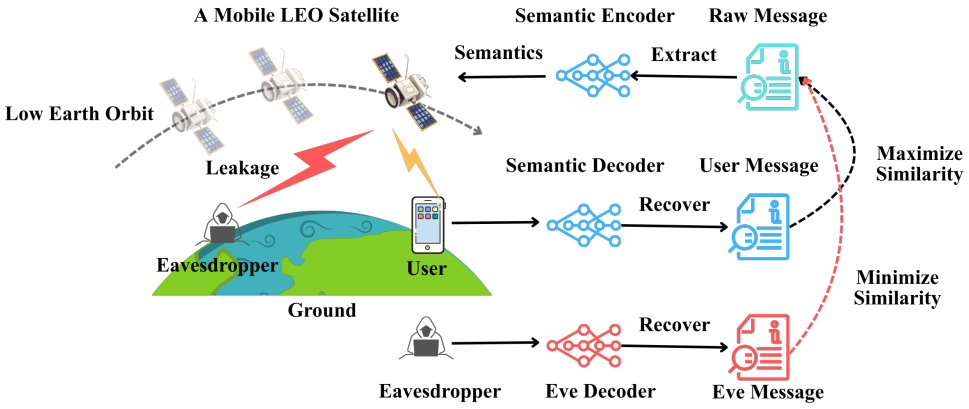


Figure 1. The considered semantic LEO satellite communication system with the presence of eavesdropping.

2.1. Propagation Environment and Channel Model

The propagation path between the satellite and a ground node $i \in \{U, E\}$ is modeled as a composite of large-scale attenuation, small-scale Rician fading, and Doppler-induced time variation due to the relative motion of the satellite. Under the assumption of accurate carrier and timing synchronization at the receiver, the complex baseband equivalent channel can be expressed as

$$h_i(t) = \sqrt{\beta_i(t)}\, g_i(t)\, e^{-j2\pi(f_c \tau_i(t) + f_{D,i} t)}, \tag{1}$$

where f_c is the carrier frequency, $\tau_i(t) = d_i(t)/c$ is the propagation delay with c denoting the speed of light, $\beta_i(t)$ captures large-scale attenuation, $g_i(t)$ represents the small-scale fading component, and $f_{D,i}$ is the instantaneous Doppler frequency shift. The Doppler shift arises from the relative velocity between the satellite and the ground terminal, given by

$$f_{D,i} = \frac{v_{\mathrm{rel},i}}{c} f_c \cos \varphi_i, \tag{2}$$

where $v_{\mathrm{rel},i}$ is the relative speed along the line-of-sight (LoS) direction, and φ_i is the angle between the satellite's velocity vector and the propagation direction.

The slant range between the satellite and node i is determined by spherical geometry as

$$d_i(t) = \sqrt{R_E^2 + (R_E + H)^2 - 2R_E(R_E + H)\cos\theta_i(t)}, \tag{3}$$

where R_E denotes the Earth's radius, H the satellite altitude, and $\theta_i(t)$ the Earth central angle corresponding to the time-varying ground elevation. The rate of change of $d_i(t)$ determines $v_{\mathrm{rel},i}$ in (2), linking geometric dynamics with frequency offset.

The large-scale gain $\beta_i(t)$ is expressed as

$$\beta_i(t) = \frac{G_t G_{r,i}}{L_{\mathrm{fs}}(d_i(t), f_c)\, L_{\mathrm{atm}}(\theta_i(t), f_c)\, L_{\mathrm{rain}}(\theta_i(t), f_c)}, \tag{4}$$

where G_t and $G_{r,i}$ are the satellite transmit and ground receive antenna gains in linear scale. The free-space loss is $L_{\mathrm{fs}}(d_i, f_c) = (4\pi d_i f_c/c)^2$, while L_{atm} and L_{rain} denote atmospheric absorption and rain attenuation, respectively, modeled according to ITU-R P.676 and P.838. These terms vary slowly compared with the symbol duration and are incorporated into $\beta_i(t)$ as quasi-static components.

The small-scale fading component $g_i(t)$ follows a Rician distribution:

$$g_i(t) = \sqrt{\frac{K_i}{K_i + 1}}\, g_i^{(\mathrm{LoS})}(t) + \sqrt{\frac{1}{K_i + 1}}\, g_i^{(\mathrm{NLoS})}(t), \tag{5}$$

where K_i is the Rician factor representing the power ratio between LoS and scattered components. The deterministic term $g_i^{(\mathrm{LoS})}(t)$ corresponds to the array steering response aligned with the time-varying satellite position, while the scattered component $g_i^{(\mathrm{NLoS})}(t) \sim \mathcal{CN}(0, 1)$ models diffuse multipath effects. In LEO downlink channels, $K_i \gg 1$, meaning the LoS component dominates and small-scale variations mainly arise from Doppler dynamics rather than rich multipath scattering.

2.2. Downlink Signal Mapping and Power Allocation

Given the time-varying channel model with Doppler effects, the satellite transmits both semantic information and AN in each symbol interval to jointly ensure meaning fidelity and physical-layer confidentiality. Let the semantic baseband symbol be denoted as $u \in \mathbb{C}$ with normalized power $\mathbb{E}[|u|^2] = 1$. The N_t-element satellite array forms the transmitted vector as

$$x(t) = \sqrt{\rho P_t}\, w_s\, u(t) + \sqrt{(1-\rho)P_t}\, w_n\, n_{\mathrm{AN}}(t), \qquad 0 \le \rho \le 1, \tag{6}$$

where P_t denotes the instantaneous transmit power; ρ is the semantic power allocation ratio; $w_s, w_n \in \mathbb{C}^{N_t}$ are unit-norm and mutually orthogonal beamforming vectors satisfying $w_s^H w_n = 0$; and $n_{\mathrm{AN}}(t) \sim \mathcal{CN}(0,1)$ denotes the artificial noise sequence. The orthogonality constraint ensures that the AN is radiated in a spatial null space with respect to the main semantic beam, thereby minimizing interference to the legitimate user while degrading the eavesdropper's effective channel.

The received baseband signals at the legitimate user U and the eavesdropper E can thus be expressed as

$$y_i(t) = h_i^H(t)\, x(t) + n_i(t), \qquad i \in \{U, E\}, \tag{7}$$

where $h_i(t)$ follows the time-varying model in (1) that includes Doppler phase rotation, and $n_i(t) \sim \mathcal{CN}(0, N_0)$ represents additive white Gaussian noise. Then, substituting (6) into (7), we have

$$y_U(t) = \sqrt{\rho P_t}\, h_U^H(t) w_s\, u(t) + \sqrt{(1-\rho)P_t}\, h_U^H(t) w_n\, n_{\mathrm{AN}}(t) + n_U(t), \tag{8}$$

$$y_E(t) = \sqrt{\rho P_t}\, h_E^H(t) w_s\, u(t) + \sqrt{(1-\rho)P_t}\, h_E^H(t) w_n\, n_{\mathrm{AN}}(t) + n_E(t). \tag{9}$$

By defining the beamformed power gains $\widehat{\Gamma}_i(t) = |h_i^H(t) w_s|^2$ and $\widehat{\Delta}_i(t) = |h_i^H(t) w_n|^2$, the instantaneous signal-to-interference-plus-noise ratio (SINR) at node i can be represented by

$$\gamma_i(t) = \frac{\rho P_t \widehat{\Gamma}_i(t)}{(1-\rho) P_t \widehat{\Delta}_i(t) + N_0}. \tag{10}$$

The SINR formulation (10) inherently reflects the joint impact of transmit power control, semantic-to-noise power allocation, and Doppler-induced channel fluctuation. An increase in ρ enhances the semantic signal strength received by the legitimate user, thereby improving semantic reconstruction accuracy, whereas reducing ρ allocates more power to AN, effectively lowering the eavesdropper's SINR and improving confidentiality. In LEO scenarios, Doppler variations cause rapid phase rotation in $h_i(t)$, which leads to temporal SINR fluctuations even for fixed ρ and P_t. Therefore, adaptive power allocation strategies are essential to dynamically balance semantic fidelity and secrecy capacity under such non-stationary satellite channels.

3. Semantic Representation and Transmission Model

3.1. Semantic Encoder

Let $M \in \mathbb{R}^n$ denote the original message, e.g., a telemetry segment or textual token, sampled from a source with density $p_M(\cdot)$ and finite differential entropy $H(M)$. The satellite implements a low-complexity semantic encoder $f_\theta : \mathbb{R}^n \to \mathbb{R}^k$ that extracts a k-dimensional latent vector

$$s \;=\; f_\theta(M) \;=\; W_2\, \sigma(W_1 M + b_1) + b_2, \tag{11}$$

where $W_1 \in \mathbb{R}^{h \times n}$, $W_2 \in \mathbb{R}^{k \times h}$, $b_1 \in \mathbb{R}^h$, $b_2 \in \mathbb{R}^k$, and $\sigma(\cdot)$ is ReLU to capture component-wise nonlinearity. The parameter set is $\theta = \{W_1, W_2, b_1, b_2\}$. To capture multi-topic semantics while enabling tractable second-order analysis, we assume the latent distribution follows a Gaussian-mixture prior

$$p(s) \;=\; \sum_{\ell=1}^{L} \pi_\ell\, \mathcal{N}(s \,|\, \mu_\ell, \Sigma_\ell), \qquad \pi_\ell \ge 0,\ \sum_{\ell} \pi_\ell = 1, \tag{12}$$

and define the latent covariance

$$\Sigma_s \triangleq \mathbb{E}[(s - \mathbb{E}s)(s - \mathbb{E}s)^\top]. \tag{13}$$

To decouple semantic statistics from the physical-layer energy budget, the encoder applies a normalization layer so that

$$\tilde{s} \;=\; \Sigma_s^{-1/2}(s - \mathbb{E}s), \qquad \mathbb{E}[\tilde{s}] = 0, \;\; \mathbb{E}\|\tilde{s}\|^2 = k, \tag{14}$$

followed by a per-block power shaping that enforces the semantic feature power constraint:

$$\mathbb{E}\big[\|\tilde{s}\|^2\big] \;=\; k \implies \mathbb{E}\big[\|u\|^2\big] \;=\; 1, \qquad u \triangleq q^\top \tilde{s}, \tag{15}$$

where $q \in \mathbb{R}^k$ is a unit-norm projection that maps the k-dimensional latent to a scalar complex symbol stream, defined per channel use, to be transmitted over one spatial beam. Consequently, the encoder optimization used offline prior to deployment is represented by

$$\min_{\theta,q} \;\; \mathbb{E}\Big[\|M - \hat{M}\|^2\Big] \quad \text{s.t.} \quad \mathbb{E}[|u|^2] = 1, \tag{16}$$

where $\hat{M}$ denotes the end-to-end reconstruction at the receiver, as defined in (17). Problem (16) fixes the latent power at baseband and leaves the physical-layer energy allocation to the downlink mapper via (ρ, P_t).

3.2. Semantic Decoder

The ground user applies a lightweight semantic decoder $g_\phi : \mathbb{C} \to \mathbb{R}^n$ to reconstruct the meaning from the received symbols y_U. The structurally symmetric decoder can be represented by

$$\hat{M} \;=\; g_\phi(y_U) \;=\; V_2\,\sigma\big(V_1\,y_U + c_1\big) + c_2, \tag{17}$$

with $\phi = \{V_1, V_2, c_1, c_2\}$ and the same nonlinearity $\sigma(\cdot)$ as defined in (11). The end-to-end semantic flow obeys the Markov chain

$$M \;\xrightarrow{f_\theta}\; s \;\xrightarrow{q^\top}\; u \;\xrightarrow{\text{downlink}}\; y_U \;\xrightarrow{g_\phi}\; \hat{M}, \tag{18}$$

which implies $I(M;\hat{M}) \le I(M;u) \le I(u;y_U)$ by data processing. Since $\mathbb{E}[|u|^2] = 1$ by (15), the mutual information $I(u;y_U)$ is determined solely by the SINR γ_U at the physical layer.

3.3. Semantic Fidelity and Secrecy

To connect the meaning-level fidelity to the physical-layer reliability, we use a widely adopted monotone approximation that maps SINR to semantic distortion:

$$D_{\text{sem}}(\gamma_U) \approx \frac{\kappa_{\text{SF}}}{\log_2(1 + \gamma_U)}, \quad \kappa_{\text{SF}} > 0, \tag{19}$$

which yields $SF(\gamma_U) \approx 1 - \kappa_{\text{SF}}/\log_2(1 + \gamma_U)$.

We quantify semantic fidelity as the normalized mutual information between the source meaning and its reconstruction, represented by

$$\text{SF} \;\triangleq\; \frac{I(M;\hat{M})}{I(M;M)} \;=\; 1 - \frac{H(M|\hat{M})}{H(M)} \;\in\; [0,1], \tag{20}$$

where $I(\cdot;\cdot)$ and $H(\cdot)$ denote mutual and differential entropy, respectively. Under small-distortion operation and smooth decoders, a second-order approximation together with

the information-minimum mean square error (I-MMSE) relation builds a monotone link between fidelity and the effective SNR, represented by

$$\mathrm{SF}(\gamma_U) \;\approx\; 1 - \frac{\kappa}{\log_2(1+\gamma_U)}, \qquad \kappa > 0, \tag{21}$$

where κ aggregates the semantic compression depth via k/n, task difficulty through the curvature of g_ϕ, and the source entropy scale $H(M)$. The approximation (21) is used to connect meaning-level distortion to physical-layer reliability while preserving the correct low- and high-SNR asymptotes.

To characterize confidentiality, we adopt the instantaneous secrecy capacity of the wiretap link driven by the physical SINRs:

$$C_s \;=\; \left[\log_2(1+\gamma_U) - \log_2(1+\gamma_E)\right]^+, \tag{22}$$

which measures the maximum secure rate in the unit of bits/s/Hz, at which the semantic stream can be conveyed without information leakage, given (ρ, P_t, w_s, w_n) and the channel realization (h_U, h_E). Both fidelity and secrecy are functions of the control variables:

$$\mathrm{SF} \;=\; \mathrm{SF}(\rho, P_t; \widehat{\Gamma}_U, \widehat{\Delta}_U), \qquad C_s \;=\; C_s(\rho, P_t; \widehat{\Gamma}_U, \widehat{\Delta}_U, \widehat{\Gamma}_E, \widehat{\Delta}_E), \tag{23}$$

thereby establishing the semantic–security coupling used later in the joint objective. In particular, increasing ρ improves γ_U but can also increase γ_E, thus reducing C_s in (22). Conversely, decreasing ρ allocates more power to AN, typically enlarging C_s while degrading SF. This antagonistic dependence is the fundamental trade-off that motivates the co-design developed in the optimization section.

4. Joint Semantic-Security Optimization

The preceding sections have demonstrated that both semantic fidelity and secrecy capacity depend critically on two coupled control parameters: the total transmit power P_t and the semantic power allocation ratio ρ. Increasing ρ enhances the semantic SNR at the legitimate user and thus improves the meaning reconstruction accuracy, but it also increases the eavesdropper's SNR and weakens secrecy. Conversely, reducing ρ devotes more energy to AN, thereby suppressing the eavesdropper at the cost of higher semantic distortion. Meanwhile, the total transmit power P_t jointly determines the overall system energy efficiency, further complicating the trade-off among semantic fidelity, secrecy performance, and power consumption. This section formulates a unified optimization model to characterize these competing objectives.

4.1. A. Composite Cost Function Design

To quantify the semantic–security trade-off, we introduce a composite objective function that integrates semantic distortion, secrecy risk, and energy consumption by:

$$J(\rho, P_t) = \underbrace{\frac{\kappa}{\log_2(1+\gamma_U)}}_{\text{semantic distortion}} + \underbrace{\mu e^{-\rho P_t |h_E|^2}}_{\text{secrecy penalty}} + \underbrace{\lambda \frac{P_t}{P_{\max}}}_{\text{power cost}}, \tag{24}$$

where κ, μ, and λ are positive weighting coefficients that determine the relative importance of semantic fidelity, secrecy enhancement, and power efficiency, respectively. The first term represents the inverse of semantic fidelity, $\mathrm{SF}(\gamma_U) = 1 - \kappa/\log_2(1+\gamma_U)$, which penalizes low reconstruction quality at the legitimate receiver. A larger γ_U, achieved via increased ρ or P_t, reduces this term and thus improves semantic performance. The second term, $\mu e^{-\rho P_t |h_E|^2}$, characterizes the residual information leakage to the eavesdropper. Its exponential form

stems from secrecy-outage analysis, reflecting the fact that the eavesdropper's decoding probability decreases exponentially with increasing interference or AN power. The last term, $\lambda P_t / P_{\max}$, imposes a linear energy cost relative to the maximum available satellite transmit power $P_{\max}$, ensuring compliance with onboard power constraints and promoting energy-aware communication.

4.2. Optimization Problem Formulation

The joint semantic–security optimization problem can be expressed as

$$\min_{0\le\rho\le 1, 0\le P_t\le P_{\max}} J(\rho, P_t), \tag{25}$$

where the decision variables ρ and P_t jointly determine the effective SINRs of both the legitimate and eavesdropping channels:

$$\gamma_U = \frac{\rho P_t |h_U|^2}{(1-\rho)P_t |h_U^H w_n|^2 + N_0}, \quad \gamma_E = \frac{\rho P_t |h_E|^2}{(1-\rho)P_t |h_E^H w_n|^2 + N_0}. \tag{26}$$

This formulation reflects a nonconvex coupled problem, where increasing semantic fidelity influences secrecy efficiency. The trade-off surface defined by (ρ, P_t) is highly nonlinear and channel-dependent, particularly under the time-varying propagation conditions of LEO satellite links.

The weighting factors (κ, μ, λ) provide system-level flexibility: for mission-critical tasks that demand high interpretability, e.g., remote-sensing data transmission, a larger κ emphasizes semantic accuracy. For secure communications, a higher μ favors secrecy, while for energy-limited satellites, an increased λ enforces power conservation. By tuning these parameters, the model can represent different operational modes, from semantics-prioritized transmission to secrecy-prioritized secure broadcasting.

The cost function $J(\rho, P_t)$ in (24) exhibits several important properties that make direct analytical optimization difficult yet suitable for learning-based control:

1. Nonconvexity and Coupling: The semantic term $\frac{\kappa}{\log_2(1+\gamma_U)}$ and the secrecy term $\mu e^{-\rho P_t |h_E|^2}$ are both nonlinear and jointly dependent on ρ and P_t, producing a nonconvex landscape with potentially multiple local minima.
2. Channel Uncertainty: The instantaneous channel coefficients (h_U, h_E) vary rapidly in LEO downlinks due to orbital motion and beam steering. Accurate real-time estimation of these parameters is costly and sometimes infeasible.
3. Dynamic Trade-off: The optimal power allocation $(\rho^\star, P_t^\star)$ depends not only on channel states but also on mission objectives and system constraints that may evolve over time. This calls for an adaptive optimization mechanism rather than a static closed-form solution.

Considering these challenges, conventional convex optimization or analytical derivation cannot guarantee robust or real-time solutions in practical satellite environments. To address this issue, in the next section, we will develop an RL-based adaptive optimization framework, where the satellite acts as an intelligent agent that continuously observes the environment state, including semantic feedback and channel indicators, and learns the optimal power allocation policy (ρ, P_t) through interaction. This approach allows the system to approximate near-optimal behavior under uncertainty and dynamic conditions without explicit channel models or prior statistical knowledge.

5. Proposed Reinforcement-Learning-Based Online Optimization

To address the inherent nonconvexity and time-varying nature of the joint semantic-security optimization problem formulated in Section 5, we adopt an actor–critic-based

RL approach that enables the satellite to autonomously adjust the transmit power P_t and semantic ratio ρ in real time. The RL agent continuously observes the environment, evaluates transmission outcomes, and improves its policy through repeated interaction without requiring explicit knowledge of the underlying channel distributions. This data-driven adaptation is particularly suitable for nonstationary LEO satellite channels, where analytical solutions are impractical due to rapid variations in geometry, path loss, and fading.

5.1. MDP Formulation

The system environment is modeled as a Markov decision process (MDP) $\mathcal{M} = \langle \mathcal{S}, \mathcal{A}, r, \mathcal{P} \rangle$, where $\mathcal{S}$ denotes the state space, $\mathcal{A}$ the action space, r the reward function, and $\mathcal{P}$ the state transition probability. At each time step t, the agent observes the environment state $s_t \in \mathcal{S}$, executes an action $a_t \in \mathcal{A}$, and receives an immediate reward r_t. The goal of the RL agent is to learn a stationary policy $\pi(a|s)$ that maximizes the expected long-term discounted reward:

$$J_{\mathrm{RL}} = \mathbb{E}_{\pi}\left[\sum_{t=0}^{\infty} \gamma^t r_t\right], \tag{27}$$

where $0 < \gamma < 1$ is the discount factor reflecting the temporal importance of future rewards. The MDP abstraction allows the satellite to treat dynamic semantic-security optimization as a sequential decision problem, capturing both temporal correlations, e.g., channel evolution, and performance feedback.

5.2. State Space

The state vector s_t aggregates real-time physical, semantic, and security-related indicators observable by the satellite, defined as

$$s_t = [\widehat{\Gamma}_U,\ \widehat{\Gamma}_E,\ f_D,\ \bar{C}_s(t-1),\ \rho_{t-1},\ P_{t-1},\ \xi(t)]. \tag{28}$$

Here, $\widehat{\Gamma}_U$ and $\widehat{\Gamma}_E$ represent the instantaneous beamformed channel gains of the legitimate user and the eavesdropper, respectively, which can be estimated through pilot feedback or side information from the network. The averaged secrecy capacity $\bar{C}_s(t-1)$ encapsulates the system's long-term confidentiality level as observed in the previous time slot, providing historical feedback for stability. The pair $(\rho_{t-1},\ P_{t-1})$ represents the previously applied semantic ratio and transmit power, allowing the agent to account for temporal correlations and avoid abrupt control actions. Finally, $\xi(t)$ reflects the instantaneous semantic task weight, quantifying the contextual importance of the transmitted information, such as prioritizing critical command data over routine telemetry. Together, these elements ensure that the RL agent jointly perceives the physical-layer environment, semantic relevance, and security conditions, thereby enabling context-aware and adaptive power control decisions in dynamic satellite communication scenarios.

5.3. Action Space

The action taken by the agent at each step consists of adjusting both the semantic ratio and the transmit power:

$$a_t = [\rho_t, P_t] \in [0, 1] \times [0, P_{\max}], \tag{29}$$

where ρ_t controls the fraction of transmit power devoted to the semantic signal, and P_t specifies the total transmission power for the current time slot. This compact continuous action space allows fine-grained adaptation of power allocation, consistent with hardware feasibility onboard satellites equipped with digital beamforming and power-amplifier control circuits.

5.4. Reward Function

The reward function is carefully designed to reflect the instantaneous system utility derived from semantic fidelity, secrecy preservation, and energy efficiency, as well as to enforce operational constraints. The reward received at time t is defined as

$$r_t = -\left[\frac{\kappa\,\xi(t)}{\log_2(1+\gamma_U(t))} + \mu e^{-\rho_t P_t \widehat{\Gamma}_E(t)} + \lambda \frac{P_t}{P_{\max}}\right] - \alpha[C_{\min} - C_s(t)]_+ - \beta([\rho_t - 1]_+ + [-\rho_t]_+) - \beta'([P_t - P_{\max}]_+ + [-P_t]_+), \quad (30)$$

where $[x]_+ \triangleq \max(x, 0)$ denotes the ReLU operator. The first bracketed term penalizes the same cost components as in the offline objective $J(\rho, P_t)$ in (24), scaled by the semantic urgency factor $\xi(t)$. The remaining terms impose soft penalties on constraint violations: (1) $\alpha[C_{\min} - C_s(t)]_+$ ensures that the instantaneous secrecy capacity $C_s(t)$ stays above the minimum acceptable threshold $C_{\min}$; and (2) β and β' penalize infeasible actions that exceed the allowable ranges of ρ_t and P_t, ensuring stable exploration during training. This reward formulation ensures that maximizing expected cumulative reward is equivalent to minimizing long-term average cost while respecting physical and security constraints.

5.5. Learning Architecture

To solve the above continuous-state, continuous-action MDP, we adopt an actor–critic reinforcement learning framework with a Gaussian policy parameterization. The actor network parameterized by ϑ generates the mean $\mu_\vartheta(s)$ and covariance $\Sigma_\vartheta(s)$ of the Gaussian distribution:

$$\pi_\vartheta(a|s) = \mathcal{N}\big(\mu_\vartheta(s), \Sigma_\vartheta(s)\big), \quad (31)$$

$$\nabla_\vartheta J_{\mathrm{RL}} = \mathbb{E}\Big[\nabla_\vartheta \log \pi_\vartheta(a_t|s_t)\,\hat{A}_t + \eta_H \nabla_\vartheta \mathcal{H}(\pi_\vartheta)\Big], \quad (32)$$

where $\hat{A}_t$ is the advantage function estimated by the critic, η_H is the entropy regularization coefficient encouraging policy exploration, and $\mathcal{H}(\pi_\vartheta)$ denotes the entropy of the policy. The critic network parameterized by ψ estimates the value function $V_\psi(s_t)$, which is updated by minimizing the temporal difference (TD) loss

$$\mathcal{L}_V(\psi) = \mathbb{E}\Big[\big(r_t + \gamma V_\psi(s_{t+1}) - V_\psi(s_t)\big)^2\Big]. \quad (33)$$

This actor–critic architecture allows stable policy gradient updates, low variance in training, and online adaptability under the dynamic conditions of satellite networks.

5.6. Training Procedure

The training process proceeds in episodes, where each episode corresponds to one operational period, e.g., a satellite orbital window or a fixed number of time steps. During training, the agent continuously interacts with the environment, observes state transitions, collects experience tuples (s_t, a_t, r_t, s_{t+1}), and updates both actor and critic parameters using sampled trajectories. The training goal is to maximize the long-term expected cumulative reward:

$$J_{\mathrm{RL}} = \mathbb{E}_{\pi_\vartheta}\left[\sum_{t=0}^{\infty} \gamma^t r_t\right], \quad (34)$$

subject to the action and secrecy constraints described earlier.

The critic network first computes the TD target $y_t = r_t + \gamma V_\psi(s_{t+1})$ and minimizes the loss $\mathcal{L}_V(\psi)$ in (33). The advantage estimate is then obtained as $\hat{A}_t = y_t - V_\psi(s_t)$. The actor parameters are updated using the gradient ascent rule in (32), where the entropy

term $\eta_H \nabla_\vartheta \mathcal{H}(\pi_\vartheta)$ ensures continuous exploration during online adaptation. All network parameters are optimized using stochastic gradient descent or Adam with separate learning rates for the actor and critic.

This training process, as shown in Algorithm 1, allows the agent to iteratively learn a mapping from observed environment states to optimal power-control actions that minimize the long-term cost in (24). Since the learning is performed through continuous interaction with the physical communication environment, the resulting policy adapts naturally to temporal channel variations, semantic task shifts, and security-level requirements, enabling autonomous online optimization in real-world LEO satellite scenarios.

Algorithm 1 RL-Based Online Semantic-Security Optimization

Initialize: actor parameters ϑ, critic parameters ψ.
for each episode $e = 1, \ldots, E$ **do**
  Observe initial state s_0
  for each time step $t = 0, \ldots, T$ **do**
    Sample action $a_t = [\rho_t, P_t] \sim \pi_\vartheta(a_t|s_t)$
    Project: $\rho_t \leftarrow \min\{1, \max\{0, \rho_t\}\}$, $P_t \leftarrow \min\{P_{\max}, \max\{0, P_t\}\}$
    Execute a_t, observe reward r_t and next state s_{t+1}
    Compute TD target: $y_t = r_t + \gamma V_\psi(s_{t+1})$
    Compute advantage: $\hat{A}_t = y_t - V_\psi(s_t)$
    Update critic: $\psi \leftarrow \psi - \eta_c \nabla_\psi \mathcal{L}_V(\psi)$
    Update actor: $\vartheta \leftarrow \vartheta + \eta_a \nabla_\vartheta [\log \pi_\vartheta(a_t|s_t) \hat{A}_t + \eta_H \mathcal{H}(\pi_\vartheta)]$
  end for
end for
Output: trained policy $\pi_\vartheta(a|s)$ for online operation

6. Simulation Results

To validate the effectiveness of the proposed semantic security design and RL optimizer, numerical experiments are conducted in a semantic LEO satellite communication network. In particular, we consider a single-beam LEO satellite equipped with an $N_t = 8$-element uniform linear array operating in the Ka band with carrier frequency $f_c = 20\,\text{GHz}$. The satellite altitude is $H = 600\,\text{km}$, and the Earth radius is $R_E = 6371\,\text{km}$ [41,52]. The legitimate user and the eavesdropper are placed at elevation angles $\theta_U = 60^\circ$ and $\theta_E = 30^\circ$, respectively. The receiver noise power spectral density is $N_0 = -174\,\text{dBm/Hz}$ with bandwidth $B = 10\,\text{MHz}$. The maximum transmit power is $P_{\max} = 40\,\text{dBm}$, and the Rician factors are set to $K_U = 12\,\text{dB}$ and $K_E = 6\,\text{dB}$, representing dominant LoS and partially blocked conditions, respectively. The relative satellite velocity is $v_{\text{rel}} = 7.5\,\text{km/s}$, resulting in a maximum Doppler shift $f_{D,\max} = 500\,\text{kHz}$ at $f_c = 20\,\text{GHz}$, which follows a cosine variation with elevation angle [53]. Atmospheric and rain attenuation are modeled by using ITU-R P.676 and P.838, giving a mean additional loss of 3–4 dB for moderate rainfall rates of 25 mm/h [54]. For the semantic module, the source messages are represented by $n = 128$-dimensional feature vectors, compressed into $k = 32$ dimensional latent vectors [47]. The semantic fidelity coefficient is set to $\kappa = 0.4$, reflecting moderate semantic compression depth. In the joint cost function, the secrecy and energy weights are $\mu = 0.8$ and $\lambda = 0.05$, respectively. The minimum required secrecy capacity threshold is $C_{\min} = 1\,\text{bit/s/Hz}$. The RL agent implements an actor–critic architecture with fully connected neural networks of two hidden layers with sizes $[256, 128]$ and ReLU activation. The actor outputs Gaussian policy parameters $(\mu_\vartheta(s), \Sigma_\vartheta(s))$ for actions $a_t = [\rho_t, P_t]$, and the critic estimates the value function $V_\psi(s_t)$. The learning rates are $\eta_a = 5 \times 10^{-4}$ for the actor and $\eta_c = 10^{-3}$ for the critic. The discount factor is $\gamma = 0.98$, the entropy regularization coefficient is $\eta_H = 0.01$, and the dual-variable update step size is $\eta_\nu = 0.05$. Training is performed over $E = 2000$ episodes, each containing $T = 200$ time steps. In simulations,

f_θ and g_ϕ are implemented as a lightweight symmetric autoencoder. The encoder uses four Conv-BN-ReLU blocks with stride-2 downsampling followed by a fully connected projection to a d-dimensional latent feature (set to $d = 256$), while the decoder mirrors the encoder with four upsampling blocks and a final convolution to reconstruct the input [55].

As shown in Table 1, the three weights in Equation (23) control the learned operating point of the RL policy along the semantic fidelity–secrecy–power trade-off. For a fixed energy weight λ, increasing the secrecy weight μ consistently improves the average secrecy capacity (C_s), which indicates a stronger tendency to mitigate information leakage, while this gain is accompanied by a moderate reduction in semantic fidelity (SF) and a lower semantic power ratio $(\bar{\rho})$, reflecting a more conservative allocation to the information-bearing semantic stream (i.e., relatively stronger protection). Conversely, for a fixed μ, increasing λ reduces the average normalized transmit power $(\bar{P}_t / P_{\max})$ as expected, but may degrade both (SF) and (C_s) under tighter power penalization because the policy becomes more energy saving and thus less capable of simultaneously supporting high-quality semantic reconstruction and secrecy enhancement. Overall, the monotonic trends across the sweeps demonstrate that the proposed RL-based optimizer remains stable under different weight settings and provides a practical mechanism to tune system behavior according to mission priorities.

Table 1. Sensitivity of RLSS performance to weight coefficients with $\kappa = 0.4$ fixed; μ, λ varied.

Case	κ	μ	λ	Avg. SF	Avg. C_s (bit/s/Hz)	Avg. $\bar{P}_t/P_{\max}$	Avg. $\bar{\rho}$
1	0.4	0.4	0.02	0.96	0.98	0.88	0.74
2	0.4	0.8	0.02	0.95	1.12	0.90	0.69
3	0.4	1.2	0.02	0.93	1.25	0.92	0.63
4	0.4	0.4	0.05	0.95	0.95	0.80	0.73
5 (default)	0.4	0.8	0.05	0.94	1.10	0.82	0.68
6	0.4	1.2	0.05	0.91	1.22	0.84	0.61
7	0.4	0.4	0.10	0.91	0.90	0.72	0.72
8	0.4	0.8	0.10	0.90	1.02	0.74	0.66
9	0.4	1.2	0.10	0.87	1.12	0.76	0.59

For fair comparison, three benchmark strategies are implemented. The first is the fixed power allocation (FPA) scheme, in which the semantic power ratio is fixed at $\rho = 0.7$, and the satellite transmits at a constant power level $P_t = P_{\max}$, representing a conventional deterministic allocation without adaptive control. The second is the dynamic policy coding (DPC) scheme, which adapts transmission parameters according to the estimated channel state and interference level. DPC relies on pre-defined policy mappings instead of end-to-end reinforcement learning. The third is the heuristic power control (HPC) scheme, a rule-based adaptive approach where ρ_t decreases linearly with the instantaneous channel gain ratio $\widehat{\Gamma}_E / \widehat{\Gamma}_U$, and P_t scales with the legitimate user's SINR. For comparison, the proposed RL-based semantic-security optimization (RLSS) framework adaptively learns an optimal policy to jointly control ρ_t and P_t in real time, maximizing the long-term reward while ensuring secrecy and semantic fidelity under dynamic satellite channels.

The convergence of semantic fidelity across training episodes is illustrated in Figure 2. The proposed RLSS algorithm exhibits the fastest and most stable convergence among all compared schemes. Specifically, it rapidly improves semantic fidelity from an initial value of 0.45 to approximately 0.9 after 1500 episodes, while both DPC and HPC converge to suboptimal levels around 0.8 and 0.7, respectively. This performance gain originates from the actor–critic framework's ability to continuously explore the semantic-security action space and refine the power allocation policy through temporal-difference feedback. In contrast, FPA maintains a constant fidelity around 0.55, indicating that fixed transmission

parameters cannot adapt to channel dynamics or semantic variations. The smooth convergence curve of RLSS also highlights the stability of its policy update process, demonstrating consistent learning behavior under time-varying Rician fading conditions.

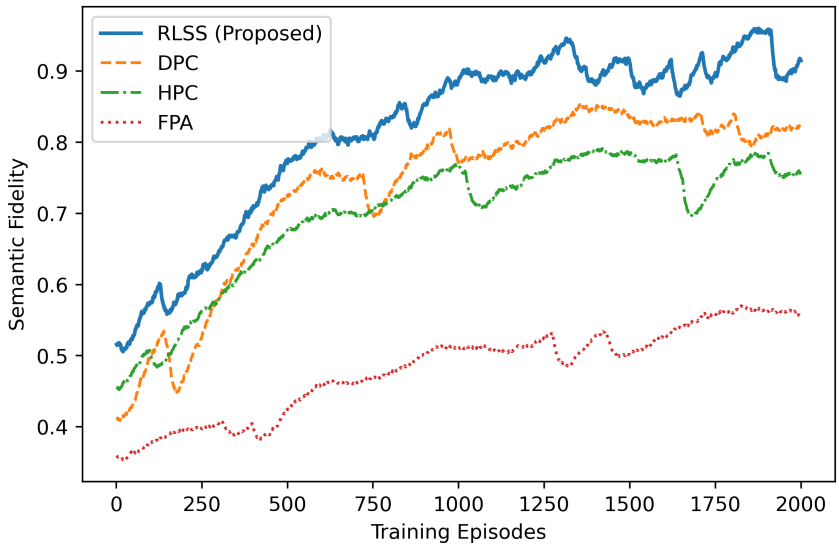


Figure 2. Convergence of semantic fidelity over training episodes.

Figure 3 shows the evolution of the secrecy capacity during the same training process. The RLSS progressively enhances the achievable secrecy capacity as the training episodes increase, eventually stabilizing near 1 bit/s/Hz, which closely approaches the ideal DPC benchmark. This indicates that the learned policy successfully balances semantic fidelity and secrecy constraints through adaptive allocation between semantic signal and artificial noise power. HPC shows a moderate increase but saturates around 0.75 bit/s/Hz due to its deterministic rule-based adjustments that cannot fully capture long-term reward dependencies. FPA remains flat, signifying complete insensitivity to the evolving eavesdropping channel. These results validate that the RL agent learns to implicitly protect semantic content by dynamically reinforcing transmission secrecy against stochastic eavesdropping environments.

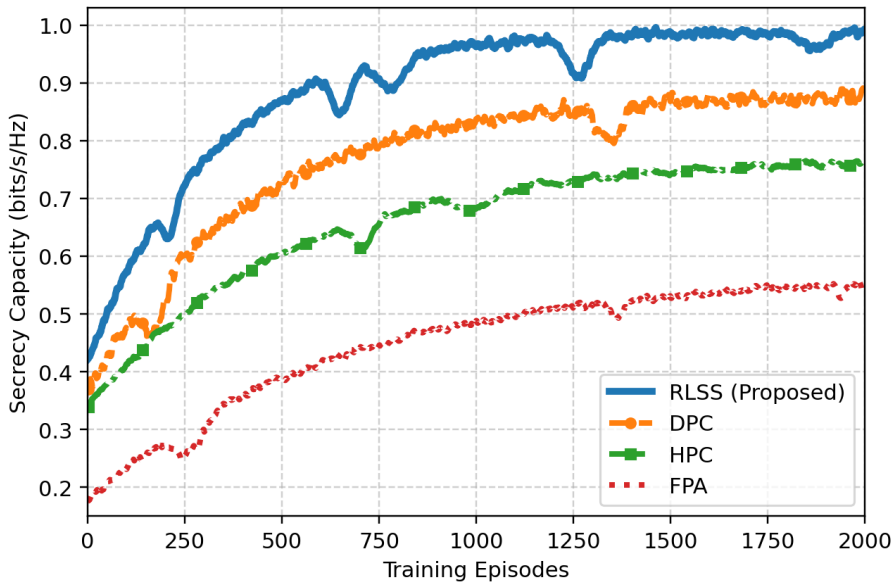


Figure 3. Convergence of secrecy capacity over training episodes.

The transmit power in Figure 4 further shows how energy efficiency emerges from the learning process. The RLSS policy adaptively reduces the normalized transmit power from unity to approximately 0.82 as training progresses, maintaining secrecy without excessive energy use. DPC exhibits slightly higher steady-state power consumption, whereas HPC shows slower adaptation and larger fluctuations. FPA, operating with full fixed power, yields the highest energy usage yet the lowest semantic performance. The observed stability in RLSS's power curve demonstrates the system's ability to meet secrecy requirements with minimal power overhead, which is an essential feature for energy-constrained satellite payloads.

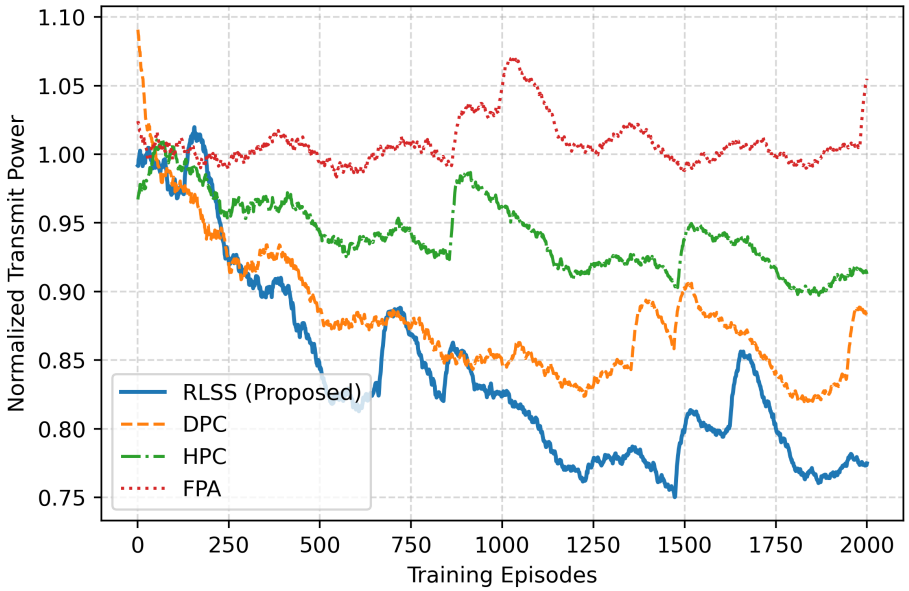


Figure 4. Normalized transmit power adaptation during training.

The trade-off between semantic fidelity and secrecy capacity is illustrated in Figure 5. Across the entire operational range, the RLSS consistently outperforms all baselines, achieving higher fidelity for any given secrecy level. In particular, when the secrecy capacity is 1 bit/s/Hz, RLSS maintains a fidelity of 0.94, outperforming DPC by 8%, HPC by 15%, and FPA by more than 25%. The convex shape of the RLSS curve indicates efficient Pareto balancing between meaning preservation and confidentiality. This implies that the learned policy captures the intrinsic coupling between semantic distortion and physical-layer secrecy, thus achieving near-optimal performance without explicit analytic optimization or perfect channel knowledge.

Figure 6 shows the effect of Doppler frequency on semantic fidelity. As the Doppler shift increases from 0 to 500 Hz, all schemes experience performance degradation due to reduced channel coherence. However, the proposed RLSS maintains robustness with a gradual fidelity decline from 0.9 to 0.62, outperforming DPC by 10% and HPC by 18% at high Doppler regimes. This resilience arises from the policy's online adaptation capability, which continuously fine-tunes transmission parameters in response to time-varying propagation and fading statistics. In contrast, FPA exhibits a steep degradation, highlighting the necessity of learning-based dynamic control. Collectively, these results confirm that the proposed RLSS framework achieves superior convergence stability, secrecy preservation, and semantic robustness across diverse satellite downlink conditions, establishing it as an effective and physically consistent approach for future 6G semantic-secure non-terrestrial networks.

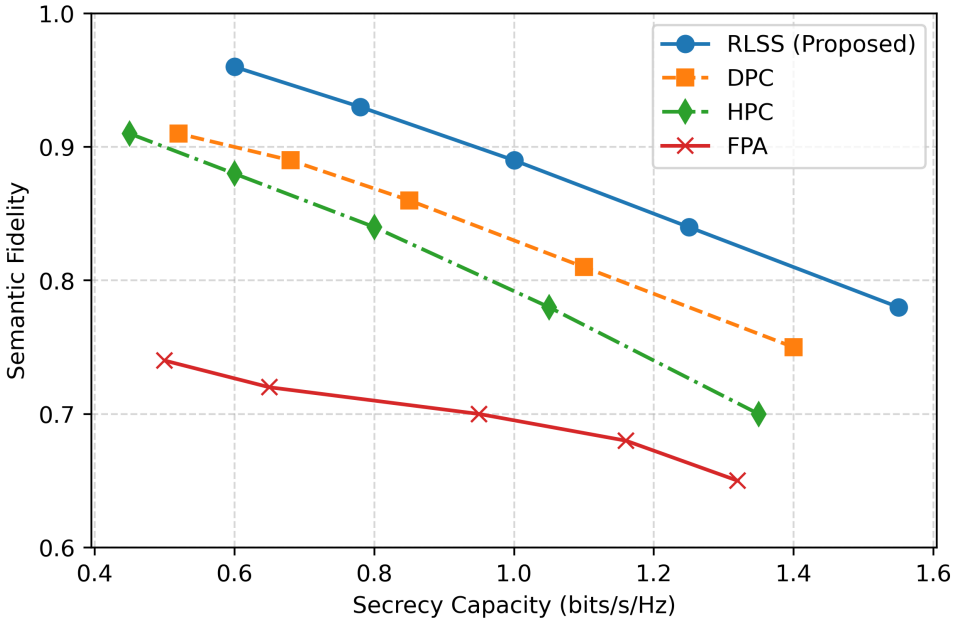


Figure 5. Semantic fidelity versus secrecy capacity trade-off comparison.

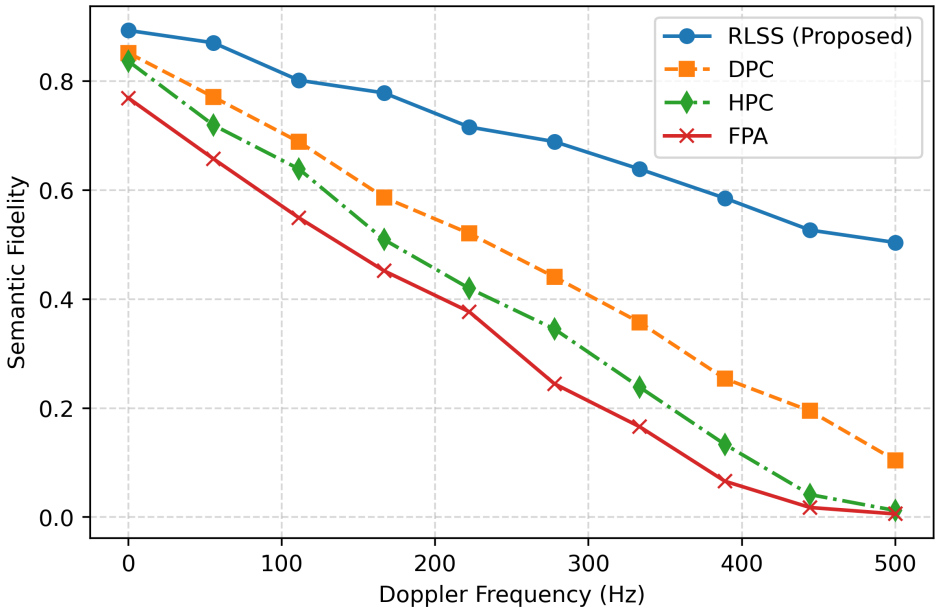


Figure 6. Impact of Doppler frequency on semantic fidelity.

To further quantify the impact of different optimization strategies, Table 2 presents a detailed decomposition of the joint cost function $J(\rho, P_t)$, including the semantic distortion term $\frac{\kappa}{\log_2(1+\gamma_U)}$, the secrecy penalty $\mu e^{-\rho P_t |h_E|^2}$, and the power consumption term $\lambda P_t / P_{\max}$. From Table 2, it can be seen that the semantic distortion component of RLSS is reduced by 30% relative to FPA, while the secrecy leakage term decreases by 50%, which reflects the proposed method's ability to inject artificial noise only when necessary rather than uniformly. The resulting total cost of 0.62 represents a near-optimal equilibrium among semantic fidelity, secrecy capacity, and energy efficiency. These results demonstrate the effectiveness of the RL-driven co-design framework as a holistic solution for secure semantic satellite communication.

Table 2. Decomposition of the average joint cost $J(\rho, P_t)$ into semantic fidelity, secrecy, and power terms for different schemes.

Scheme	$\frac{\kappa}{\log_2(1+\gamma_U)}$	$\mu e^{-\rho P_t \lvert h_E \rvert^2}$	$\lambda \frac{P_t}{P_{\max}}$	$J(\rho, P_t)$
RLSS (Proposed)	0.32	0.18	0.12	0.62
DPC	0.38	0.22	0.15	0.75
HPC	0.44	0.28	0.19	0.91
FPA	0.57	0.34	0.23	1.14

7. Conclusions

This paper has developed a unified analytical and learning-based framework for secure semantic communication in low Earth orbit (LEO) satellite downlink systems. By explicitly coupling semantic representation, Doppler-aware satellite propagation, and artificial-noise-assisted transmission, we formulated a joint semantic-security cost function that quantitatively characterizes the trade-off between meaning fidelity, secrecy capacity, and energy efficiency. Unlike conventional bit-level models, the proposed framework captures the semantic distortion caused by physical-layer impairments and reveals the antagonistic dependence of fidelity and secrecy on the same power allocation parameters. To address the dynamic and uncertain characteristics of satellite channels, a reinforcement-learning-based online optimizer was introduced. Through an actor–critic architecture with adaptive dual-variable constraint handling, the satellite transceiver autonomously learns to allocate transmit power and semantic-to-noise ratios in real time, without requiring explicit channel statistics. Simulation results have shown that the proposed scheme significantly improves semantic fidelity and secrecy efficiency under time-varying Rician channels with Doppler effects, validating its robustness and adaptability.

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

Funding: This work was supported by Science and Technology Project of State Grid Corporation Headquarters under Grant 5700-202455369A-3-1-DG.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors would like to acknowledge the financial support from State Grid Corporation Headquarters, China.

Conflicts of Interest: Author Feifei Zhou was employed by the company Nari Group Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

6G	Sixth-Generation
LEO	Low Earth Orbit
STINs	Satellite-Terrestrial Integrated Networks

IoT	Internet of Things
BER	Bit Error Rate
LoS	Line-of-Sight
NLoS	Non-Line-of-Sight
PLS	Physical-Layer Security
IRS	Intelligent reflecting surface
AN	Artificial noise
SNR	Signal-to-noise ratio
SINR	Signal-to-interference-plus-noise ratio
SF	Semantic fidelity
C_s	Secrecy capacity
RL	Reinforcement learning
DRL	Deep reinforcement learning
MDP	Markov decision process
TD	Temporal difference
FPA	Fixed power allocation
DPC	Dynamic policy coding
HPC	Heuristic power control
RLSS	RL-based semantic-security optimization
ReLU	Rectified linear unit
I-MMSE	Information-minimum mean square error
ITU-R	International Telecommunication Union Radiocommunication Sector

References

1. Zhu, X.; Jiang, C. Integrated Satellite-Terrestrial Networks Toward 6G: Architectures, Applications, and Challenges. *IEEE Internet Things J.* **2022**, *9*, 437–461. [CrossRef]
2. Xiao, Z.; Yang, J.; Mao, T.; Xu, C.; Zhang, R.; Han, Z.; Xia, X.G. LEO Satellite Access Network (LEO-SAN) Toward 6G: Challenges and Approaches. *IEEE Wirel. Commun.* **2024**, *31*, 89–96. [CrossRef]
3. Toka, M.; Lee, B.; Seong, J.; Kaushik, A.; Lee, J.; Lee, J.; Lee, N.; Shin, W.; Poor, H.V. RIS-Empowered LEO Satellite Networks for 6G: Promising Usage Scenarios and Future Directions. *IEEE Commun. Mag.* **2024**, *62*, 128–135. [CrossRef]
4. Kagai, F.; Branch, P.; But, J.; Allen, R.; Rice, M. Rapidly Deployable Satellite-Based Emergency Communications Infrastructure. *IEEE Access* **2024**, *12*, 139368–139410. [CrossRef]
5. Routray, S.K.; Hussein, H.M. Satellite based IoT networks for emerging applications. *arXiv* **2019**, arXiv:1904.00520. [CrossRef]
6. Dwivedi, A.K.; Chaudhari, S.; Varshney, N.; Varshney, P.K. Performance Analysis of LEO Satellite-Based IoT Networks in the Presence of Interference. *IEEE Internet Things J.* **2024**, *11*, 8783–8799. [CrossRef]
7. Talgat, A.; Kishk, M.A.; Alouini, M.S. Stochastic Geometry-Based Uplink Performance Analysis of IoT Over LEO Satellite Communication. *IEEE Trans. Aerosp. Electron. Syst.* **2024**, *60*, 4198–4213. [CrossRef]
8. Zhang, Y.; Lin, Z.; Chen, Z.; Fang, Z.; Zhu, W.; Chen, X.; Zhao, J.; Gao, Y. Satfed: A resource-efficient leo satellite-assisted heterogeneous federated learning framework. *arXiv* **2024**, arXiv:2409.13503. [CrossRef]
9. Liu, X.; Zhang, Z.; Qin, X.; Zhou, H.; Zhao, L. FlexSATE: Flexible and Distributed Traffic Engineering with Supervised Learning in Ultra-Dense Low-Earth-Orbit Satellite Networks. In *Proceedings of the GLOBECOM 2024–2024 IEEE Global Communications Conference*; IEEE: New York, NY, USA, 2024; pp. 4466–4471.
10. Maity, I.; Giambene, G.; Vu, T.X.; Kesha, C.; Chatzinotas, S. Traffic-Aware Resource Management in SDN/NFV-Based Satellite Networks for Remote and Urban Areas. *IEEE Trans. Veh. Technol.* **2024**, *73*, 17400–17415. [CrossRef]
11. Wang, L.; Wu, W.; Zhou, F.; Yang, Z.; Qin, Z.; Wu, Q. Adaptive Resource Allocation for Semantic Communication Networks. *IEEE Trans. Commun.* **2024**, *72*, 6900–6916. [CrossRef]
12. Koyuncu, E. Information Theory in Emerging Wireless Communication Systems and Networks. *Entropy* **2024**, *26*, 543. [CrossRef] [PubMed]
13. Farshbafan, M.K.; Saad, W.; Debbah, M. Curriculum Learning for Goal-Oriented Semantic Communications with a Common Language. *IEEE Trans. Commun.* **2023**, *71*, 1430–1446. [CrossRef]
14. Xie, H.; Qin, Z. A lite distributed semantic communication system for Internet of Things. *IEEE J. Sel. Areas Commun.* **2020**, *39*, 142–153. [CrossRef]
15. Xie, H.; Qin, Z.; Li, G.Y.; Juang, B.H. Deep learning enabled semantic communication systems. *IEEE Trans. Signal Process.* **2021**, *69*, 2663–2675. [CrossRef]

16. Chaccour, C.; Saad, W.; Debbah, M.; Han, Z.; Poor, H.V. Less data, more knowledge: Building next-generation semantic communication networks. *IEEE Commun. Surv. Tutorials* **2024**, *27*, 37–76. [CrossRef]
17. Chen, Y.; Wang, H.; Liu, C.; Wang, L.; Liu, J.; Wu, W. Generative Multi-Modal Mutual Enhancement Video Semantic Communications. *Comput. Model. Eng. Sci. (CMES)* **2024**, *139*, 2985–3009. [CrossRef]
18. Wang, L.; Wu, W.; Tian, F.; Hu, H. Intelligent Resource Allocation for UAV-Enabled Spectrum Sharing Semantic Communication Networks. In *Proceedings of the 2023 IEEE 23rd International Conference on Communication Technology (ICCT)*; IEEE: New York, NY, USA, 2023; pp. 1359–1363.
19. Getu, T.M.; Kaddoum, G.; Bennis, M. Semantic Communication: A Survey on Research Landscape, Challenges, and Future Directions. *Proc. IEEE* **2024**, *112*, 1649–1685. [CrossRef]
20. Hu, B.; Ma, J.; Sun, Z.; Liu, J.; Li, R.; Wang, L. DRL-based intelligent resource allocation for physical layer semantic communication with IRS. *Phys. Commun.* **2024**, *63*, 102270. [CrossRef]
21. Qin, Z.; Tao, X.; Lu, J.; Tong, W.; Li, G.Y. Semantic communications: Principles and challenges. *arXiv* **2021**, arXiv:2201.01389.
22. Shi, Y.; Zhang, Y.; Zhao, H.; Wei, J.B. Deep Learning-Enabled Semantic Communication with Structured Semantic Representation. In *Proceedings of the 2025 IEEE Wireless Communications and Networking Conference (WCNC)*; IEEE: New York, NY, USA, 2025; pp. 1–6.
23. Li, Y.; Zhou, F.; Yuan, L.; Wu, Q.; Al-Dhahir, N. Cognitive Semantic Communication: A New Communication Paradigm for 6G. *IEEE Commun. Mag.* **2025**, *63*, 122–129. [CrossRef]
24. Tiken Singh, M.; Arif, A.; Kumar Prasad, R.; Choudhury, B.; Kalita, C.; Askari, S.M.S. Optimizing Fairness and Spectral Efficiency with Shapley-Based User Prioritization in Semantic Communication. *IEEE Access* **2025**, *13*, 61299–61311. [CrossRef]
25. Mortaheb, M.; Amir Khojastepour, M.A.; Ulukus, S. Efficient Semantic Communication Through Transformer-Aided Compression. In *Proceedings of the 2025 IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN)*; IEEE: New York, NY, USA, 2025; pp. 1–6.
26. Liu, H.; Yang, B.; Yu, Z.; Cao, X.; Alexandropoulos, G.C.; Zhang, Y.; Yuen, C. AI-Empowered Channel Generation for IoV Semantic Communications in Dynamic Conditions. *arXiv* **2025**, arXiv:2507.02013.
27. Kang, M.; Park, S.; Lee, Y. A survey on satellite communication system security. *Sensors* **2024**, *24*, 2897. [CrossRef]
28. Zhang, Y.; Zhao, S.; He, J.; Zhang, Y.; Shen, Y.; Jiang, X. A survey of secure communications for satellite internet based on cryptography and physical layer security. *IET Inf. Secur.* **2023**, *2023*, 5604802. [CrossRef]
29. Han, S.; Li, J.; Meng, W.; Guizani, M.; Sun, S. Challenges of Physical Layer Security in a Satellite-Terrestrial Network. *IEEE Netw.* **2022**, *36*, 98–104. [CrossRef]
30. Amiriara, H.; Mirmohseni, M.; Elzanaty, A.; Ma, Y.; Tafazolli, R. A Physical Layer Security Framework for IRS-Assisted Integrated Sensing and Semantic Communication Systems. *IEEE Trans. Cogn. Commun. Netw.* **2025**, *11*, 2955–2969. [CrossRef]
31. Han, B.; Zhu, Y.; Schmeink, A.; Caire, G.; Schotten, H.D. A Semantic Model for Physical Layer Deception. In *Proceedings of the ICC 2025—IEEE International Conference on Communications*; IEEE: New York, NY, USA, 2025; pp. 4129–4134.
32. Ahmed, I.; Sun, Y.; Fu, J.; Köse, A.; Musavian, L.; Xiao, M.; Özbek, B. Semantic Communications in 6G: Coexistence, Multiple Access, and Satellite Networks. *IEEE Commun. Stand. Mag.* **2025**, *9*, 58–64. [CrossRef]
33. Hassan, S.S.; Nguyen, L.X.; Tun, Y.K.; Han, Z.; Hong, C.S. Semantic enabled 6g leo satellite communication for earth observation: A resource-constrained network optimization. In *Proceedings of the GLOBECOM 2024–2024 IEEE Global Communications Conference*; IEEE: New York, NY, USA, 2024; pp. 4472–4478.
34. Yuan, S.; Sun, Y.; Peng, M.; Yuan, R. Joint beam direction control and radio resource allocation in dynamic multi-beam LEO satellite networks. *IEEE Trans. Veh. Technol.* **2024**, *73*, 8222–8237. [CrossRef]
35. Zhong, L.; Li, Y.; Ge, M.F.; Feng, M.; Mao, S. Joint Task Offloading and Resource Allocation for LEO Satellite-Based Mobile Edge Computing Systems with Heterogeneous Task Demands. *IEEE Trans. Veh. Technol.* **2025**, *74*, 11337–11352. [CrossRef]
36. Feng, S.; Lu, X.; Sun, S.; Hossain, E.; Wei, G.; Ni, Z. Covert Communication in Large-Scale Multi-Tier LEO Satellite Networks. *IEEE Trans. Mob. Comput.* **2024**, *23*, 11576–11587. [CrossRef]
37. Evgenidis, N.G.; Mitsiou, N.A.; Tegos, S.A.; Diamantoulakis, P.D.; Sarigiannidis, P.; Krikidis, I.; Karagiannidis, G.K. Hybrid semantic-Shannon communications. *IEEE Trans. Wirel. Commun.* **2024**, *23*, 10926–10940. [CrossRef]
38. Lin, L.; Xu, W.; Wang, F.; Zhang, Y.; Zhang, W.; Zhang, P. Channel-Transferable Semantic Communications for Multi-User OFDM-NOMA Systems. *IEEE Wirel. Commun. Lett.* **2024**, *13*, 721–725. [CrossRef]
39. Cao, H.; Han, S.; Meng, R.; Xu, X.; Zhang, P. Channel Code-Book (CCB): Semantic Image-Adaptive Transmission in Satellite–Ground Scenario. *Sensors* **2025**, *25*, 269. [CrossRef]
40. Huang, C.; Chen, X.; Chen, G.; Xiao, P.; Ye Li, G.; Huang, W. Deep Reinforcement Learning-Based Resource Allocation for Hybrid Bit and Generative Semantic Communications in Space-Air-Ground Integrated Networks. *IEEE J. Sel. Areas Commun.* **2025**, *43*, 3942–3954. [CrossRef]
41. Kim, E.; Roberts, I.P.; Andrews, J.G. Downlink Analysis and Evaluation of Multi-Beam LEO Satellite Communication in Shadowed Rician Channels. *IEEE Trans. Veh. Technol.* **2024**, *73*, 2061–2075. [CrossRef]

42. Kumar, R.; Arnon, S. Review of physical layer security in integrated satellite–terrestrial networks. *Electronics* **2024**, *13*, 4414. [CrossRef]
43. Wang, L.; Wu, W.; Zhou, F.; Qin, Z.; Wu, Q. IRS-Enhanced Secure Semantic Communication Networks: Cross-Layer and Context-Awared Resource Allocation. *IEEE Trans. Wirel. Commun.* **2025**, *24*, 494–508. [CrossRef]
44. Jang, G.; Kim, D.; Lee, I.H.; Jung, H. Cooperative Beamforming with Artificial Noise Injection for Physical-Layer Security. *IEEE Access* **2023**, *11*, 22553–22573. [CrossRef]
45. Qiao, S.; Shao, Y.; Kang, G.; Zhao, J.; Yue, P.; Zhang, R.; Ke, S.; Wang, S. Code-Assisted Broadband Jamming Suppression in Multisatellite Collaborative Secure Communication. *IEEE Trans. Aerosp. Electron. Syst.* **2025**, *61*, 3162–3175. [CrossRef]
46. Mallick, S.; Acharya, T.; Chakravarty, S. Robust secure beamforming design in cognitive satellite communication for coexistence with terrestrial networks. *Phys. Commun.* **2024**, *67*, 102509. [CrossRef]
47. Wang, L.; Wu, W.; Zhou, F.; Qin, Z.; Wu, Q. Cross-Layer Security for Semantic Communications: Metrics and Optimization. *IEEE Trans. Veh. Technol.* **2025**, *74*, 13179–13183. [CrossRef]
48. Yang, Z.; Chen, M.; Li, G.; Yang, Y.; Zhang, Z. Secure semantic communications: Fundamentals and challenges. *IEEE Netw.* **2024**, *38*, 513–520. [CrossRef]
49. Zheng, G.; Ni, Q.; Navaie, K.; Pervaiz, H. Semantic Communication in Satellite-Borne Edge Cloud Network for Computation Offloading. *IEEE J. Sel. Areas Commun.* **2024**, *42*, 1145–1158. [CrossRef]
50. Corici, M.; Caus, M.; Artiga, X.; Guidotti, A.; Barth, B.; de Cola, T.; Tallon, J.; Zope, H.; Tarchi, D.; Parzysz, F.; et al. Transforming 5g mega-constellation communications: A self-organized network architecture perspective. *IEEE Access* **2025**, *13*, 14770–14788. [CrossRef]
51. Shinde, S.S.; Naseh, D.; Tarchi, D. In-space computation offloading for multi-layer LEO constellations. In *Proceedings of the European Wireless 2023; 28th European Wireless Conference*; VDE: Berlin, Germany, 2023; pp. 82–89.
52. Leyva-Mayorga, I.; Soret, B.; Röper, M.; Wübben, D.; Matthiesen, B.; Dekorsy, A.; Popovski, P. LEO small-satellite constellations for 5G and beyond-5G communications. *IEEE Access* **2020**, *8*, 184955–184964. [CrossRef]
53. Wang, H.; Qi, W.; Kadoch, M.; Hong, T. Low-carbon federated multiagent-DRL enhanced network slicing for satellite direct-to-device communications. *IEEE Internet Things J.* **2024**, *11*, 39158–39169. [CrossRef]
54. Liu, W.; Michelson, D.G. Fade slope analysis of Ka-band Earth-LEO satellite links using a synthetic rain field model. *IEEE Trans. Veh. Technol.* **2008**, *58*, 4013–4022.
55. Wu, W.; Wang, L.; Zhou, F.; Yang, Z.; Wu, Q. Knowledge-Driven 3D Semantic Spectrum Map: KE-VQ-Transformer-Based AAV Semantic Communication and Map Completion. *IEEE Trans. Cogn. Commun. Netw.* **2026**, *12*, 5120–5135. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.