Journal of Science and Technology - Smart Systems and Devices
Volume 36, Issue 1, January 2026, 009-017

UAV-to-Satellite Communication for 6G IoV Networks Using Beamforming Orthogonal Time Frequency Space: A Deep Q-Learning Approach

Nguyen Huu Trung
School of Electrical and Electronic Engineering, Hanoi University of Science and Technology, Ha Noi, Vietnam
**Corresponding author email: trung.nguyenhuu@hust.edu.vn*

Abstract

In the paper, we propose a novel beamforming-based Orthogonal Time Frequency Space (OTFS) transmission framework for UAV-to-Satellite Communication (U2SC) tailored for 6G-Enabled Internet of Vehicles (IoV) networks. To address the unique challenges of high Doppler shifts, long-range line-of-sight (LoS) links, and fast-moving Low Earth Orbit (LEO) satellites, we adopt OTFS modulation due to its inherent robustness against doubly dispersive channels. A Uniform Linear Array (ULA) is equipped on the UAV to enable highly directional transmission. Furthermore, we propose a Deep Q-Learning (DQL) framework for adaptive beamforming, in which the beam control problem is formulated as a Markov Decision Process (MDP). By leveraging DQL, the agent learns to dynamically steer the beam to align with the satellite's trajectory, optimizing both link quality and energy efficiency while minimizing misalignment. Simulation results demonstrate significant gains in signal robustness and beam alignment accuracy compared to conventional methods. In addition, future work will focus on building a hardware-in-the-loop (HIL) testbed using a UAV platform with phased-array antennas to validate the proposed model under real orbital satellite trajectories and Doppler conditions.

Keywords: 6G, Deep Q-Learning, IoV, LEO satellite, OTFS, UAV, U2SC.

1. Introduction

1.1. UAV-to-Satellite Communication

The sixth generation (6G) wireless systems-enabled Internet of Vehicles (IoV) will adopt space-air-ground-sea integrated communications to achieve ubiquitous connectivity. Therefore, satellite communication and Unmanned Aerial Vehicle (UAV) to Satellite communication (U2SC) play an important role. U2SC refers to the integration of UAV and satellites to enhance communication capabilities, especially in areas where terrestrial networks are limited or unreliable [1].

As UAVs become increasingly vital in applications such as disaster recovery, maritime monitoring, and intelligent transportation, the need for reliable long-range connectivity grows. While terrestrial communication networks provide high throughput, their limited coverage makes them unsuitable for remote or mobile aerial platforms. Low earth orbit satellites like Starlink offer global coverage and high-speed backhaul, making them ideal complements for UAV communications [2].

U2SC enables global 6G coverage but poses several challenges: high Doppler shifts from satellite motion, dynamic geometry, beam misalignment, and UAV power constraints. Addressing these issues requires modulation schemes and beam control methods that are resilient, adaptive, and energy efficient. In both Long-Term Evolution (LTE) and the fifth generation (5G) New Radio (NR) systems, Orthogonal Frequency Division Multiplexing (OFDM) and its variants have been widely adopted for high-speed wireless transmissions [3]. However, OFDM is inherently sensitive to Doppler effects, which can disrupt the orthogonality among subcarriers. This disruption leads to increased inter-carrier interference (ICI) and inter-symbol interference (ISI), severely degrading system performance in high-mobility scenarios [4].

1.2. Orthogonal Time Frequency Space Modulation

Orthogonal Time Frequency Space (OTFS) modulation is specifically designed to address the limitations of traditional schemes such as OFDM in rapidly time-varying and doubly dispersive wireless channels [5]. This transformation yields several key advantages for high-mobility communication. First, OTFS exhibits strong resilience to Doppler and delay spread, as each transmitted symbol is spread over the entire time-frequency plane, allowing it to experience the full diversity of the channel. This significantly improves robustness against both time selectivity (e.g., Doppler shift) and frequency selectivity (e.g., multipath delay) [6]. Second, OTFS effectively converts a sparse delay-Doppler channel into a well-conditioned channel matrix in the symbol domain, enhancing the reliability of symbol detection and decoding. Third, due to the inherent sparsity of the channel in the delay-Doppler domain, low-complexity equalization methods such as message passing or

p-ISSN 3093-3285
e-ISSN 3093-3315
https://doi.org/10.51316/jst.187.ssad.2026.36.1.2
Received: Jun 10, 2025; Revised: Dec 3, 2025;
Accepted: Dec 3, 2025; Online: Dec 20, 2025.

MMSE can be employed with near-optimal performance. Lastly, OTFS is highly compatible with multiple-input and multiple-output (MIMO) and beamforming techniques, enabling spatial diversity and multiplexing gains even under rapidly time-varying channel conditions [7].

These properties make OTFS a compelling candidate for U2SC, where the link is dominated by high mobility, large Doppler shifts, and strong line-of-sight (LOS) propagation. By leveraging delay-Doppler representation, OTFS enhances signal robustness and improves communication reliability under the dynamics of spaceborne systems [8].

In U2SC, maintaining a strong and stable link is challenging due to the rapid motion of Low Earth Orbit (LEO) satellites and the mobility of UAV platforms. Directional beamforming using Uniform Linear Arrays (ULA) at the UAV plays a vital role in focusing transmission energy to the satellite, thereby improving link quality and extending communication range [9]. However, the effectiveness of beamforming relies heavily on accurate and timely beam alignment with the moving satellite. Traditional beam steering techniques, such as exhaustive search or fixed-point tracking, are often computationally inefficient or slow to adapt to the rapidly changing geometry of U2SC links. Moreover, frequent misalignment leads to significant signal degradation and increased energy consumption due to retransmissions or power ramping [10].

To address this, Artificial Intelligence (AI)-based tracking approaches, particularly Deep Q-Learning (DQL), have emerged as promising solutions. DQL enables the UAV to learn an optimal beam control policy through interaction with the environment, dynamically adjusting the beam direction based on estimated satellite position, received signal strength, Doppler shift, or alignment error. Over time, the learning agent can generalize different trajectories and satellite velocities, offering robust performance under uncertainty and channel dynamics. When combined with OTFS modulation, AI-powered beamforming creates a synergistic framework: OTFS ensures modulation-level robustness in Doppler-rich environments, while DQL-driven beam control maintains spatial alignment. This integration is essential for enabling resilient, low-latency, and energy-efficient communication in future 6G U2SC systems [11].

1.3. Related Works

Satellite communication has become a key component in the development of 5G and future 6G networks, particularly in providing ubiquitous coverage in remote, rural, or disaster-stricken areas. Recent advances in LEO satellite constellations, such as Starlink and OneWeb, have enabled low-latency, high-throughput backhaul links that complement terrestrial networks.

In [12], the authors conceive a combination of Code Division Multiple Access (CDMA) and OTFS. This is an interesting study that combines the advantages of CDMA with OTFS, providing a robust candidate for high-Doppler and doubly dispersive channels. In [13], wavelet-aided orthogonal time-frequency space (W-OTFS) modulation is proposed as a novel approach for high-mobility vehicular communication, the advantage of discrete wavelet transforms improved performance over existing OTFS modulation.

Several works have demonstrated the superiority of OTFS over OFDM in high-mobility environments and satellite communication. For example, the work in [14] proposed OTFS with Non-Orthogonal Multiple Access (NOMA), the work in [15] proposed Intelligent Reflecting Surface (IRS)-Aided uplink OTFS-SCMA and the work in [16] proposed Wavelet-Based OTFS Scheme for Low Earth Orbit (LEO) Satellite Communication. OTFS maps data symbols in the delay-Doppler domain, allowing them to experience the full diversity of the channel, which improves reliability and robustness.

Beamforming has been widely applied in UAV systems to improve link quality, directional gain, and interference management. Both analog and digital beamforming architectures have been explored using Uniform Linear Arrays (ULA) or planar arrays mounted on UAVs [17]. Challenges include maintaining beam alignment during UAV motion, handling limited payload constraints, and dealing with fast angular variations due to the UAV's dynamic positioning.

Adaptive and intelligent beamforming is therefore crucial for UAV communication in long-range and dynamic environments such as UAV-to-satellite links. AI, especially Reinforcement Learning (RL), has recently gained traction in solving complex control problems in wireless communication, including beam tracking and dynamic beam selection. In [18], a Robust Beamforming Design for OTFS-NOMA is proposed for sharing the spectrum with multiple low-mobility NOMA users under uncertain and time-varying environments.

1.4. Motivation and Contributions of this Paper

In conventional non-terrestrial networks, direct communication between mobile users and LEO satellites requires high transmission power at the user equipment (UE) and suffers from significant propagation loss and Doppler-induced distortion. Motivated by the mentioned research gap, in this paper, we propose the use of a UAV as an aerial relay or flying Base Station, acting as an intelligent intermediary between ground UEs and satellites.

Our proposed relay-based U2SC architecture has the following advantages:

- Reduced UE transmit power, as the UAV provides a closer, lower-loss uplink target;

- Lower interference footprint, particularly in dense user environments, since the UAV can manage scheduling and power control locally;
- Improved link reliability and coverage, especially in obstructed or remote areas;
- Enhanced flexibility, as the UAV can reposition itself dynamically to optimize satellite visibility and user distribution.

By integrating this UAV relay concept with OTFS modulation and AI-based beamforming, our system significantly improves the overall spectrum and energy efficiency of U2SC in 6G-Enabled networks. To address the unique challenges of U2SC in 6G-Enabled IoV networks, we propose a comprehensive approach that integrates network architecture design, advanced modulation, and intelligent beamforming. Our research approaches include:

- *Towards an Edge BS architecture on UAV*, where the UAV acts as an aerial base station (BS) serving ground users and relaying aggregated data to the satellite, thereby reducing user transmit power, interference, and enabling centralized beam and resource management;

- *Towards OTFS-based uplink transmission*, enhancing robustness against Doppler and delay dispersion through full delay–Doppler diversity;

- *Towards DQL for beamforming control,* where beam steering is formulated as a Markov Decision Process (MDP) and a DQL agent adaptively aligns the beam using real-time feedback, e.g., Signal-to-Interference-plus-Noise Ratio (SINR), Doppler, ensuring precise and energy-efficient satellite tracking.

This paper proposes a robust solution combining OTFS modulation and beamforming with DQL to enable high-performance, energy-aware UAV-to-satellite communication in 6G-Enabled IoV networks.

Our contributions in this paper are summarized as follows:

1) A novel UAV-to-satellite edge architecture improving coverage, flexibility, and energy efficiency by employing the UAV as an aerial BS for LEO satellite relaying.

2) An OTFS-based delay–Doppler signal model capturing high-mobility Rician channel effects for robust transmission.

3) A DQL-based adaptive beamforming algorithm achieving accurate, low-power beam alignment and maintaining high link reliability under dynamic orbital motion.

Beyond simulation, this work provides a theoretical foundation for the implementation of a small-scale testbed. The testbed will use a UAV-mounted software-defined radio (SDR) and electronically steerable antenna array to validate beam tracking performance in a semi-realistic environment.

The rest of the paper is organized as follows. In Section 2, the proposed system model is shown. Section 3 presents proposed beam tracking with deep Q-Learning, Section 4 provides simulation results and performance of the proposed algorithm. Finally, the conclusion of this paper is to conclude remarks, and suggestions for further research.

Throughout the paper, bold uppercase letters are denoted for matrices, while bold lowercase letters are for vectors, $(.)^T$ and $(.)^H$ stand for transpose, Hermitian respectively.

2. System Model

2.1. Overall Architecture of Proposed U2SC System

Fig. 1. shows the proposed system model. The system consists of a single UAV flying at altitude h_{uav}, equipped with a N_t element uniform linear antenna array and a directional beamforming system capable of steering a transmission beam toward a LEO satellite at altitude H_{sat}. The UAV serves U ground users by collecting their uplink data and forwarding it to the satellite using OTFS-modulated beamformed transmission. At each time instant, the UAV aligns its beam with a single satellite. Additionally, the UAV maintains a directional backhaul link to a terrestrial Next Generation Node B (gNB), supporting control signaling and optional data offloading. This architecture allows the UAV to function as an aerial Base Station (BS), enabling reduced user transmit power, improved link robustness, and adaptive satellite connectivity in dynamic 6G-Enabled IoV environments.

2.2. OTFS Signal Model

We propose a system model in which the UAV acts as a relay BS between ground users and the satellite. Specifically, the i-th User Equipment (UE), located at a distance d_i, transmits a signal $s_i(t)$ to the UAV. Upon receiving this signal, the UAV performs necessary processing and forwards it to the satellite using directional beamforming and OTFS modulation to ensure robustness under high-mobility channel conditions.

Let $x_i[k,l] \in \mathbb{C}$ denote the complex data symbol of i-th UE located at delay index $k \in \{0,1,...,N-1\}$ and Doppler index $l \in \{0,1,...,M-1\}$ received from i-th UE. The OTFS modulation process consists of the following steps as follows.

Step 1. Inverse Symplectic Finite Fourier Transform (ISFFT): The data matrix $x_i[k,l]$ in the delay-Doppler domain is transformed into the time-frequency domain $X_i[n,m]$ via the ISFFT as

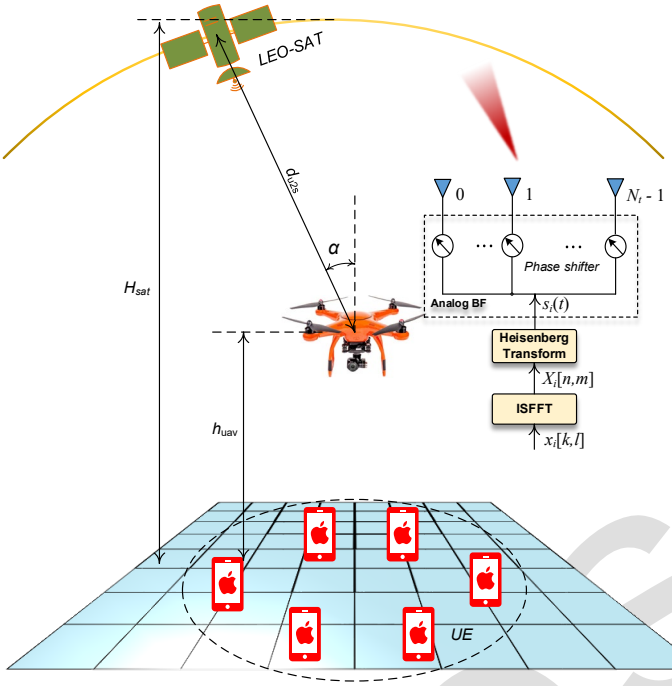

Fig. 1. Proposed beamforming OTFS for U2S system model

$$X_i[n,m] = \frac{1}{\sqrt{NM}} \sum_{k=0}^{N-1} \sum_{l=0}^{M-1} x_i[k,l] \cdot e^{j2\pi\left(\frac{nk}{N} - \frac{ml}{M}\right)}, \quad (1)$$

where $n = 0, 1, ..., N-1$, and $m = 0, 1, ..., M-1$ represent time and frequency indices, respectively.

Step 2. Heisenberg Transform at the transmitter: The time-frequency samples $X_i[n,m]$ are modulated into a continuous-time signal $s_i(t)$ using a transmit pulse $g_{tx}(t)$ as

$$s_i(t) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} X_i[n,m] \cdot g_{tx}(t - nT) \cdot e^{j2\pi m\Delta f(t-nT)}, \quad (2)$$

where T is time spacing between OTFS symbols, Δf is subcarrier spacing ($\Delta f = 1/T$), $g_{tx}(t)$ is transmit pulse (commonly rectangular or raised cosine).

Step 3. Wigner Transform and SFFT at the receiver: The receiver applies a matched filter $g_{rx}(t)$ to obtain time-frequency samples as follows

$$Y_i[n,m] = \int r_i(t) \cdot g_{rx}^*(t - nT) \cdot e^{-j2\pi m\Delta f(t-nT)} dt. \quad (3)$$

These samples are then transformed back to the delay-Doppler domain using the Symplectic Finite Fourier Transform (SFFT) which is presented as

$$\hat{x}_i[k,l] = \frac{1}{\sqrt{NM}} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} Y_i[n,m] \cdot e^{-j2\pi\left(\frac{nk}{N} - \frac{ml}{M}\right)}. \quad (4)$$

The OTFS signal model provides a robust end-to-end link capable of operating effectively under high Doppler and rapidly time-varying satellite channels. Detection can be performed using MMSE, message passing, or other equalization techniques in the delay-Doppler domain.

2.3. UAV-to-Satellite Communication Model

Suppose the satellite operates in LEO (Low Earth Orbit) with an altitude H_{sat} of approximately 550 km, the UAV flies at an altitude h_{uav} of approximately 1÷5 km. The straight-line distance between the UAV and the satellite at time t is

$$d_{u2s}(t) = \left\| \mathbf{p}_{sat}(t) - \mathbf{p}_{uav}(t) \right\|, \quad (5)$$

where, $\mathbf{p}_{sat}(t), \mathbf{p}_{uav}(t) \in \mathbb{R}^3$ are satellite coordinates (calculated according to orbit) and UAV coordinates (can be fixed or circling), respectively.

Free Space Path Loss (FSPL) is calculated as

$$PL_{dB}(t) = 20\log_{10} d_{u2s}(t) + 20\log_{10} f_c + 20\log_{10}\left(\frac{4\pi}{c}\right), \quad (6)$$

where $d_{u2s}(t)$ is UAV to satellite distance at time t, f_c is carrier frequency, and c is speed of light.

The UAV-to-Satellite Rician Fading Channel Model with Doppler shift is represented as

$$H(t) = \sqrt{\frac{K(t)}{K(t)+1}} \cdot h_{LOS}(t) + \sqrt{\frac{1}{K(t)+1}} \cdot h_{NLOS}(t), \quad (7)$$

where $h_{LOS}(t) = e^{j2\pi f_D(t)t}$ is Doppler phase of Line of Sight (LOS) component, $h_{NLOS}(t) \sim \mathcal{CN}(0,1)$ is random fading Non-Line of Sight (NLOS) component, $K(t)$ is Rician coefficient, characterizing the LOS/NLOS ratio at time t. The Doppler shift based on relative velocity between UAV and satellite is calculated as

$$f_D(t) = \frac{v_{ref}(t)}{\lambda} = \frac{1}{\lambda} \cdot \frac{d}{dt} d_{u2s}(t), \quad (8)$$

where $\lambda = c / f$ is wavelength, $v_{ref}(t)$ is relative velocity between UAV and satellite.

K-factor Model which depends on satellite distance is expressed as

$$K(t) = K_0 \cdot \left(\frac{d_0}{d_{\text{u2s}}(t)} \right)^{\gamma}, \quad (9)$$

where K_0 is LOS coefficient at standard distance d_0, γ is attenuation factor (usually 1.5-3). The signal received at the satellite is

$$r_i(t) = \sqrt{G(t)} \cdot H(t) \cdot s_i(t - \tau(t)) + w(t). \quad (10)$$

2.4. Satellite Mobility and Geometry Model

In the proposed system, we consider LEO satellites operating at altitudes ranging from 500 to 600 km, typical of commercial constellations such as Starlink. These satellites orbit the Earth at high speeds, approximately 7.5 km/s, completing a full revolution in about 90–100 minutes. As a result, the relative geometry between the satellite and the UAV changes rapidly over time, which directly affects the link distance, angle of elevation, and Doppler shift.

Satellite Position Model is as follows. Let the satellite follow a circular LEO orbit at altitude H_{sat}, with Earth radius R_e. The position of the satellite in a 2D plane (simplified case) at time *t* can be expressed as:

$$\mathbf{p}_{\text{sat}}(t) = (R_e + H_{\text{sat}}) \begin{bmatrix} \cos(\omega t) \\ \sin(\omega t) \\ 0 \end{bmatrix}, \quad (11)$$

The angular velocity of the satellite is $\omega = v_{\text{sat}} / (R_e + H_{\text{sat}})$, where v_{sat} is approximately of 7.5×10^3 m/s. The UAV is assumed to fly at a constant altitude h_{uav} with the position of $\mathbf{p}_{\text{uav}} = [x_u, y_u, h_{\text{uav}}]$. The instantaneous distance between UAV and satellite is

$$d_{\text{u2s}}(t) = \left\| \mathbf{p}_{\text{sat}}(t) - \mathbf{p}_{\text{uav}} \right\|, \quad (12)$$

The elevation angle $\alpha(t)$ from the UAV to the satellite is given by

$$\alpha(t) = \arcsin\left(\frac{H_{\text{sat}} - h_{\text{uav}}}{d_{\text{u2s}}(t)} \right). \quad (13)$$

This elevation angle plays a critical role in determining LOS availability, estimating the Rician *K*-factor in the channel model, and guiding the beamforming direction from the UAV toward the satellite. Due to orbital motion, the UAV can only maintain a connection with a given satellite within a limited visibility window. The contact duration T_c of the UAV with the satellite is

$$T_c \approx \frac{2R_e}{v_{\text{sat}}} \cdot \arccos\left(\frac{R_e}{R_e + H_{\text{sat}}} \right). \quad (14)$$

In practical systems, beam steering and handover between satellites must occur before the end of this duration.

2.5. ULA-Based Beamforming Model at UAV

The digital domain signal from one Radio Frequency (RF) chain is fed to *K* transmit antennas to perform transmit analog precoding. The analog precoder vector is expressed as

$$\overline{\mathbf{a}}_k = [a_{k0}, a_{k1}, a_{(kK-1)}]^T \in \mathbb{C}^{(K \times 1)}, k = 0,1,..., N_t - 1, \quad (15)$$

where $a_{ij} = A_{ij} e^{j\phi_{ij}}$, A_{ij} is attenuator factor and ϕ_{ij} is phase shift. Finally, every data symbol is transmitted by the sub-antenna array of N_t antennas.

The transmitted signal vector is $\mathbf{s}_u[k] = \left[s_0^{(u)}[k], s_1^{(u)}[k], ..., s_{N_t-1}^{(u)}[k] \right]$, where each component is

$$s_k^u[k] = y_k^u[k] \overline{\mathbf{a}}_k, k = 0,1,..., N_t - 1. \quad (16)$$

Assume the UAV collects data from *U* ground users and forwards all data to the satellite through a single uplink $R_{\text{uav-sat}}$. Then, the total system sum-rate is

$$R_{\text{sum}} = \sum_{i=1}^{U} R_i = R_{\text{uav-sat}}. \quad (17)$$

Assuming the UAV transmits with power P_t, the total channel gain is $G(t)$, and the total noise is $N_0 B$, the total sum-rate is

$$R_{\text{uav-sat}}(t) = B \cdot \log_2 \left(1 + \frac{P_t \cdot G(t)}{N_0 B} \right), \quad (18)$$

where *B* is channel bandwidth, N_0 is noise power spectral density (W/Hz), respectively.

The total channel gain $G(t)$ is

$$G(t) = \frac{\left\| \mathbf{H}(t) \right\|_F^2 G_{tx} G_{sat}}{PL(t)}, \quad (19)$$

where G_{tx} is transmit antenna gain, G_{sat} is satellite antenna gain, $PL(t)$ is total pathloss, $\mathbf{H}(t)$ is the small-scale fading matrix, which depends on the wavelength, relative velocity of UAV and satellite and Doppler spread, and $\|.\|_F^2$ denotes the matrix Frobenius norm. Since the UAV uses beamforming with N_t antennas, an additional beamforming gain G_{BF} equals N_t is included in the total sum-rate as [19]

$$R_{\text{uav-sat}}(t) = B \cdot \log_2 \left(1 + \frac{P_t \cdot N_t \cdot G(t)}{N_0 B} \right), \quad (20)$$

and the received power is

$$P_r = P_t \cdot N_t \cdot G(t). \quad (21)$$

3. Proposed Beam Tracking with Deep Q-Learning

3.1. Problem Formulation

Beam Alignment is defined as a Markov Decision Process as follows. The beam alignment task between the UAV and a moving satellite is modeled as a Markov

Decision Process (MDP), where the UAV must continuously adjust its beam direction to maintain alignment with the satellite's dynamic position. The environment evolves based on satellite motion and channel conditions, and the UAV agent learns to make sequential decisions to maximize long-term signal quality. At each step, the agent selects a beam direction that influences the next observed signal state and the reward.

3.2. State, Action, Reward Design

- **State** S_t includes observable features such as estimated elevation angle $\hat{\theta}(t)$, recent beam direction, received signal power or SINR, Doppler shift, and tracking error.
- **Action** A_t corresponds to discrete adjustments of the beam angle, e.g., $\hat{\theta}(t+1) = \hat{\theta}(t) + \Delta\theta$, where $\Delta\theta \in \{-\delta, 0, +\delta\}$.
- **Reward** R_t is designed to encourage precise tracking, e.g.,

$$R_t = \alpha \cdot \mathrm{SINR}_t - \beta \cdot |\theta(t) - \hat{\theta}(t)|\ , \tag{22}$$

where α, $\beta > 0$ control the trade-off between signal quality and angular error.

3.3. DQL Agent Architecture and Training

A Deep Q-Network (DQN) is used to approximate the optimal action-value function $Q(S_t, A_t)$. The network takes the current state as input and outputs the estimated value of each action. DQN is trained using experience replay and target network stabilization:

- **Network input**: State vector S_t.
- **Output**: Q-values for each beam adjustment action.
- **Loss function**:

$$\mathcal{L} = \mathbb{E}\left[\left(r_t + \gamma \max_{a'} Q_{\text{target}}(S_{t+1}, A^*) - Q(S_t, A_t)\right)^2\right]. \tag{23}$$

The agent is trained over simulated satellite trajectories with known channel conditions and then fine-tuned online.

We extend the DQL beam tracking policy to include energy awareness by modifying the reward function as

$$R_t = \alpha \cdot \mathrm{SINR}_t - \beta \cdot E_{\text{track}}(t) - \gamma \cdot E_b(t)\ , \tag{24}$$

where E_b denotes the energy consumption per successfully transmitted bit, and E_{track} is the energy per beam update as

$$E_b = \frac{P_{\text{total}} \cdot T_b}{\eta}, \quad E_{\text{track}} = P_{\text{ctrl}} \cdot T_{\text{align}}\ , \tag{25}$$

where T_b is time to transmit one bit, η is transmission success rate (accounts for retransmissions), and T_{align} is duration of beam realignment phase.

High-frequency alignment leads to more accurate tracking but increases E_{track}. Balancing alignment frequency is crucial for energy efficiency. The agent now balances beam accuracy and energy cost, learning a policy that adapts tracking frequency and action aggressiveness based on the UAV's remaining energy and satellite visibility window. This approach allows real-time, onboard learning for energy-optimal tracking under mobility and channel uncertainty. Finally, the proposed Beam Tracking with Deep Q-Learning is as follows.

Algorithm 1. Proposed Deep Q-Learning for Beam Tracking

Initialization:
Initialize the Q-Network $Q(S_t, A_t)$ for action A_t
Initialize replay memory D with capacity C
Initialize beam at random angle θ_0
end for
Learning:
while not convergence **do**
Initialize environment and observe initial state s_0
for iteration **do**
Obtain receive power level P_r, UE angle $\theta(t)$, beam error $\Delta\theta = \theta(t) - \hat{\theta}(t)$
if $c <= \epsilon$ **then**
Choose action A_t randomly (exploration), with probability ϵ.
else
select the action $A_t = \arg\max_A Q^*(S_t, A_t)$
end if
Execute action by applying vector beamforming in (15)
Get new signal power $P_{r,\,t+1}$
Calculate reward R_t according to (24)
Store the experience (S_t, A_t, R_r, S_{t+1}) into replay memory D
Update Q:
$Q(S_t, A_t) = Q(S_t, A_t) + \alpha\left[R_t + \gamma \max_A Q^*(S_{t+1}, A) - Q(S_t, A_t)\right]$
Perform gradient descent step on loss according to (23)
Update θ to minimize $\mathcal{L}$
Update current state $S_t \leftarrow S_{t+1}$
if satellite out of view **then**
break
end if
end for // iteration
end while //not convergence

4. Simulation results

In this section, the numerical results are presented. The Monte-Carlo simulation is used to evaluate the system performance to verify the efficiency of the proposed system in different simulation conditions. The simulation is used to evaluate the performance of the proposed system and estimate the impact of parameters on the performance of the whole system. Bit error rate (BER) with 5×10^6 channel realizations is used to evaluate the system performance. The simulation parameters are listed in Table 1. All the simulations run on a workstation with CPU Intel E5-1603 v3 @ 2.80GHz RAM 15.8 GB and GPU Nvidia GTX 1050 Ti (4 GB), using Python 3.12.

The simulation was conducted using the proposed method and compared with two baseline strategies as follows. The first baseline is the Fixed Offset approach, which assumes an ideal scenario where the UAV beam is perfectly aligned with the satellite direction (offset equals to 0). The second baseline is the Beam Switching method, which mimics conventional directional scanning: the entire angular domain is divided into discrete directions, and the UAV cyclically scans all directions to select the one with the highest received signal power.

In the first simulation, we investigated the impact of the number of transmit antennas $N_t \in \{8,16,32\}$ on the BER for the proposed OTFS-based algorithm and compared it with two baseline methods. The Rician K-factor is fixed at K equals 10 dB to reflect a realistic line-of-sight dominant scenario in UAV-to-Satellite communication. Fig. 2 demonstrates that the proposed method achieves consistently lower BER as the number of antennas increases, thanks to its beamforming adaptability and delay–Doppler channel structure exploitation. In contrast, baseline methods show less sensitivity to antenna scaling, highlighting the efficiency of the proposed approach under high spatial diversity conditions.

Table 1. Simulation Parameters.

Parameter	Symbol	Value
Carrier frequency	f_c	12 GHz
Bandwidth	B	10 MHz
UAV transmit power	P_t	-5÷35 dBm
UAV number of antennas	N_t	{8, 16, 32, 64}
UAV-SAT distance	$d_{u2s}(t)$	500 km
UAV altitude	h_{uav}	1 km
Satellite altitude	Hs_{at}	500 km
Satellite antenna gain	G_{sat}	35 dBi
Orbital velocity	v_{sat}	≈ 7.5 km/s
Rician K-factor	$K(t)$	{5, 10, 15} dB
Doppler shift	$f_D(t)$	100 ns
Noise spectral density	N_0	−174 dBm/Hz
Noise figure	NF	–9 dB
Pulse shaping	$g_{tx}(t)$	Raised Cosine
Number of subcarriers	M	64
Doppler bins	N	64

The hyperparameters for learning model is described in Table 2.

Table 2. Hyperparameters for learning models.

Parameter	Value
Learning rate	0.001
Replay memory buffer size	50000
Minibatch	64
Discount factor	0.99
Number of Episode	200

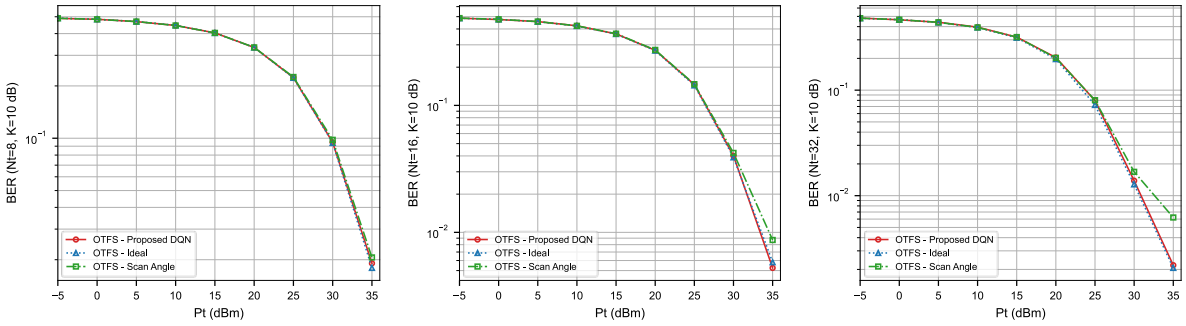

Fig. 2. BER vs number of transmit antennas comparison at K=10dB: (a) N_t=8; (b) N_t =16; and (c) N_t =32

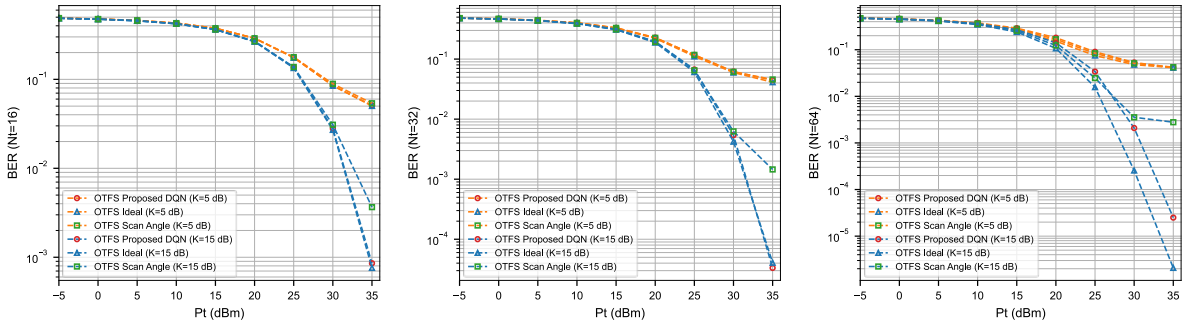

Fig. 3. BER performance comparison with various K values (5, 15 dB): (a) $N_t = 16$; (b) $N_t = 32$; and (c) $N_t = 64$

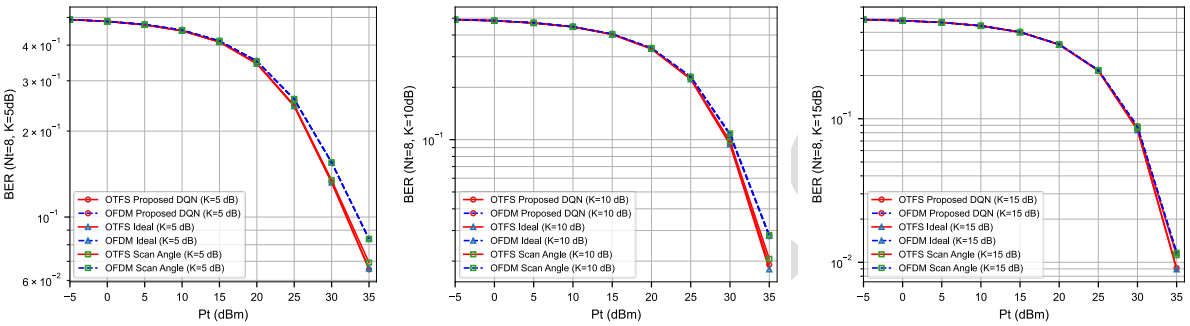

Fig. 4. OTFS vs OFDM BER comparison at $N_t = 8$ with various Rician fading: (a) K = 5 dB; (b) K = 10 dB; and (c) K = 15 dB

Next, we evaluated the impact of Rician K-factors on the performance of the proposed algorithm compared with baseline method. Specifically, we compare the bit error rate under different Rician fading conditions: K equals 5 dB (purely scattered), K equals 15 dB (moderate line-of-sight) with the number of transmit antennas at N_t equals 16, 32, and 64. Fig. 3 presents the results, showing that the proposed OTFS-based method benefits from increasing K-factors, due to improved channel predictability and robustness. This trend is particularly evident when compared to the baseline methods under harsh fading conditions.

Finally, we compare the performance of OTFS and OFDM modulation schemes under different Rician K-factors. The number of transmit antennas is fixed at low value of N_t equals 8, and the Rician fading conditions are varied by K equals 5, 10, and 15 dB to observe how LOS dominance affects system robustness. Fig. 4 illustrates that OTFS consistently outperforms OFDM, particularly in high mobility or severe fading conditions (low K), thanks to its delay-Doppler domain resilience. As the K-factor increases, the performance gap narrows due to improved channel coherence for both schemes.

The proposed DQL beamforming framework is qualitatively compared with Kalman filter–based and geometric beam tracking approaches. Unlike model-based methods, DQL learns beam control policies directly from interaction with the environment, providing more stable alignment and lower energy consumption under non-linear satellite motion. Compared with policy gradient and actor–critic reinforcement learning, DQL achieves faster convergence and lower computational complexity, making it suitable for real-time UAV implementation.

The control signaling overhead for beam updates between UAV and gNB is minimal (below 1% of total bandwidth). Under rain fade and ionospheric scintillation at 12 GHz, attenuation of 2–4 dB may occur, which can be mitigated by adaptive power control and beamwidth adjustment integrated into the learning policy. Overall, the proposed OTFS–DQL framework demonstrates robust and energy-efficient performance under realistic non-terrestrial 6G conditions.

5. Conclusion

In this paper, we have proposed a UAV-to-Satellite communication system using OTFS modulation combined with Deep Q-Learning algorithm to control beam direction. We simulated the proposed system in a Rician fading channel environment with a 12 GHz frequency band with a transmission distance of up to 500 km, fully simulating realistic factors such as background noise, propagation loss, and antenna gain. BER simulation results show that OTFS demonstrates effective transmission in complex channel environments, especially under large K-factor conditions, when compared with the traditional

Quadrature Phase Shift Keying (QPSK) modulation method. Integrating delay–Doppler channel estimation and OTFS simulation according to the standard pipeline has helped to accurately reproduce the transmission characteristics in the time–frequency domain, demonstrating the advantages against multipath interference and large delays.

The next research direction is to build a hardware-in-the-loop testbed using a UAV platform with phased-array antennas to validate the proposed model under real orbital satellite trajectories and Doppler conditions and integrate sensing and communication (ISAC) to simultaneously transmit data and monitor position and Doppler from UAVs.

Acknowledgments

This research is funded by the Vietnam Ministry of Science and Technology under grant number NĐT/CN/24/02 “Research and Development on New Generation Intelligent Internet of Vehicles Technologies”.

References

[1] Z. Haleed, X. Mao, J. Huang and C. -X. Wang, A novel 3D GBSM for 6G satellite-UAV-ground wireless communications, 2024 IEEE 99th Vehicular Technology Conference (VTC2024-Spring), Singapore, Singapore, 2024, pp. 1-5,

[2] M. Li, Y. Hong, C. Zeng, Y. Song and X. Zhang, Investigation on the UAV-To-Satellite optical communication systems, IEEE Journal on Selected Areas in Communications, vol. 36, no. 9, pp. 2128-2138, Sept. 2018, https://doi.org/ 10.1109/JSAC.2018.2864419.

[3] E. M. Mohamed, M. M. Fouda, OTFS-based proactive dynamic UAV positioning for high-speed train coverage, IEEE Open Journal of the Communications Society, vol. 5, pp. 5718-5734, 2024, https://doi.org/ 10.1109/OJCOMS.2024.3453906.

[4] Y. Su, M. Liwang, Z. Chen and X. Du, Toward optimal deployment of UAV relays in UAV-assisted IoV networks, IEEE Transactions on Vehicular Technology, https://doi.org/ 10.1109/TVT.2023.3272648.

[5] L. Gaudio, G. Colavolpe and G. Caire, OTFS vs. OFDM in the presence of sparsity: A fair comparison, IEEE Transactions on Wireless Communications, vol. 21, no. 6, pp. 4410-4423, June 2022, https://doi.org/ 10.1109/TWC.2021.3129975.

[6] M. Aldababsa, S. Özyurt, G. K. Kurt and O. Kucur, A survey on orthogonal time frequency space modulation, IEEE Open Journal of the Communications Society, vol. 5, pp. 4483-4518, 2024, https://doi.org/10.1109/OJCOMS.2024.3422801.

[7] G. Sharma and A. A. B. Raj, OTFS based SAR with low complexity receiver, IEEE Access, vol. 11, pp. 66194-66200, 2023, https://doi.org/10.1109/ACCESS.2023.3290322.

[8] A, C. K. Thomas and K. S, Enhanced spatial modulation based orthogonal time frequency space system, IEICE Transactions on Communications, vol. E107-B, no. 11, pp. 785-796, November 2024, https://doi.org/ 10.23919/transcom.2023EBP3206.

[9] M. Bayat and A. Farhang, Time and frequency synchronization for OTFS, IEEE Wireless Communications Letters, vol. 11, no. 12, pp. 2670-2674, Dec. 2022, https://doi.org/ 10.1109/LWC.2022.3214002.

[10] A. Kumar, S. Chakravarthy, N. Gaur and A. Nanthaamornphong, Hybrid approaches to PAPR, BER, and PSD optimization in 6G OTFS: Implications for healthcare, Journal of Communications and Networks, vol. 26, no. 3, pp. 308-320, June 2024, https://doi.org/ 10.23919/JCN.2024.000027.

[11] M. S. Khan, Y. J. Kim, Q. Sultan, J. Joung and Y. S. Cho, Downlink synchronization for OTFS-based cellular systems in high doppler environments, IEEE Access, vol. 9, pp. 73575-73589, 2021, https://doi.org/ 10.1109/ACCESS.2021.3079429.

[12] H. Hawkins, C. Xu, L. -L. Yang and L. Hanzo, CDMA/OTFS sensing outperforms pure OTFS at the same communication throughput, IEEE Open Journal of Vehicular Technology, vol. 6, pp. 502-519, 2025, https://doi.org/ 10.1109/OJVT.2025.3532848.

[13] M. H. Abid, I. A. Talin, M. A. Matin and M. I. Kadir, Wavelet-aided OTFS: A 2-D multicarrier modulation format for high-mobility communications, IEEE Open Journal of the Communications Society, vol. 6, pp. 2660-2678, 2025, https://doi.org/ 10.1109/OJCOMS.2025.3533027.

[14] Umakoglu, M. Namdar and A. Basgumus, Deep learning-assisted signal detection for OTFS-NOMA systems, IEEE Access, vol. 12, pp. 119105-119115, 2024, https://doi.org/ 10.1109/ACCESS.2024.3449812.

[15] Li and Y. Hong, Performance analysis and phase shift design of IRS-aided uplink OTFS-SCMA, IEEE Access, vol. 11, pp. 133059-133069, 2023, https://doi.org/ 10.1109/ACCESS.2023.3336677.

[16] S. R. Sabuj and H. -S. Jo, A complex wavelet-based OTFS scheme for LEO satellite communication, in IEEE Access, vol. 13, pp. 85807-85825, 2025, https://doi.org/ 10.1109/ACCESS.2025.3569789.

[17] T. Izydorczyk, G. Berardinelli, P. Mogensen, M. M. Ginard, J. Wigard and I. Z. Kovács, Achieving high UAV uplink throughput by using beamforming on board, in IEEE Access, vol. 8, pp. 82528-82538, 2020, https://doi.org/10.1109/ACCESS.2020.2991658.

[18] Z. Ding, Robust beamforming design for OTFS-NOMA, IEEE Open Journal of the Communications Society, vol. 1, pp. 33-40, 2020, https://doi.org/ 10.1109/OJCOMS.2019.2953574.

[19] N.H Trung, N.T Anh, N.M Duc, D.T Binh, L.T Tan, System theory based multiple beamforming, Vietnam Journal of Science and Technology, ISSN 2525-2518, Vol. 55, No. 5, Oct. 2017, pp. 653-665. https://doi.org/0.15625/2525-2518/55/5/9149.