IEICE Communications Express, Vol.12, No.1, 1–4

ComEX

LETTER

DRL-Based Power Optimization for Hybrid FSO/THz-Enabled UAV Systems Using IR-HARQ

Tien H. Do[1], **Thang V. Nguyen**[1], **Hoang D. Le**[2], **and Ngoc T. Dang**[1]

Abstract This letter addresses the transmit power minimization challenge in unmanned aerial vehicle (UAV)-assisted hybrid free-space optical (FSO)/terahertz (THz) systems integrated with incremental redundancy hybrid automatic repeat request (IR-HARQ), a cornerstone for 6G's ultra-reliable low-latency communications (URLLC). We propose a deep reinforcement learning (DRL)-driven framework, leveraging proximal policy optimization (PPO), to adapt power allocation across retransmissions via an agent-learned policy dynamically. This ensures reliable packet delivery under stringent delay bounds while accounting for channel impairments, including atmospheric attenuation, scintillation fading, and beam pointing errors. The system model incorporates SNR-based FSO/THz switching, with FSO as the primary link and THz as backup, evaluated through outage probabilities tailored to IR-HARQ, chase combining HARQ (CC-HARQ), and automatic repeat request (ARQ). Simulations across diverse environmental conditions reveal the proposed DRL-IR-HARQ hybrid achieves up to 0.7 dBm savings over THz-only baselines and conventional HARQ protocols, underscoring its robustness for energy-efficient 6G aerial backhauls and disaster-resilient networks.

Keywords: IR-HARQ, Hybrid FSO/THz link, UAV, and DRL.
Classification: Wireless Communication Technologies.

1. Introduction

In the pursuit of sixth-generation (6G) networks, unmanned aerial vehicle (UAV)-assisted communications have emerged as pivotal enablers, offering unparalleled flexibility, rapid deployment, and cost-effectiveness to support ultra-reliable low-latency communications (URLLC), terabit-per-second rates, and integrated sensing and communication (ISAC) paradigms [1]. On the other hand, free-space optical (FSO) and terahertz (THz) technologies have gotten a lot of attention because they have a lot of spectrum resources and can handle data rates of more than 100 Gbps, which will help with the spectral congestion that is expected in 6G [2], [3]. Combining FSO and THz into hybrid links makes them much more resistant to changes in the atmosphere, like turbulence and molecular absorption, while also increasing overall throughput and coverage. As a result, UAV-enabled hybrid FSO/THz architectures are poised to underpin mission-critical 6G applications, including real-time disaster response, ad-hoc temporary networks, and high-capacity aerial backhauls [4].

Notwithstanding these advancements, UAV-assisted FSO/THz systems grapple with formidable impediments, including beam misalignment, severe path attenuation, and stochastic fading, which exacerbate outage risks and energy demands in dynamic aerial environments. Mitigating these requires sophisticated error-control mechanisms like hybrid automatic repeat request (HARQ) protocols, long-established in wireless standards for balancing reliability and efficiency. Nevertheless, the energy-constrained nature of UAVs—limited by battery life and payload restrictions—necessitates intelligent power allocation schemes that uphold 6G's stringent URLLC targets (e.g., < 1 ms latency, $> 99.99999\%$ reliability) without compromising endurance. Recent works have advanced HARQ for optical links, such as [5] devises a power-minimal HARQ strategy for FSO under turbulence, curbing retransmission energy while preserving decoding fidelity, and [6] proposes non-orthogonal HARQ for UAV-THz relays, boosting spectral utilization amid absorption losses. However, these efforts predominantly scrutinize single-modality links (FSO or THz) and overlook adaptive power optimization in hybrid setups, particularly under joint delay-reliability constraints.

Bridging this gap in the literature, this study introduces a pioneering deep reinforcement learning (DRL)-driven framework for transmit power orchestration in UAV-assisted hybrid FSO/THz systems leveraging incremental redundancy HARQ (IR-HARQ). To our knowledge, this proposal represents the inaugural investigation to holistically integrate dynamic power minimization with outage-constrained packet delivery in time-varying FSO/THz channels, explicitly tailored to 6G's energy-sustainable aerial paradigms. The design uses a Markov decision process (MDP) to capture the main channel distortions, such as attenuation, scintillation, and pointing errors. A proximal policy optimization (PPO)-trained agent dynamically adjusts power across retransmissions to lower average cycle energy while making sure that decoded throughput meets URLLC thresholds. Extensive simulations across diverse atmospheric regimes validate the framework's superiority, yielding up to 0.7 dBm power savings over baseline HARQ variants, thus affirming its viability for robust, green 6G UAV deployments.

2. System Descriptions

Figure 1(a) features a UAV transmitting data to a ground BS via hybrid FSO/THz with HARQ. A cycle is divided into static slots t, allowing M retransmissions for N-bit packets.

[1] Wireless Systems and Applications Lab., Posts and Telecommunications Institute of Technology, 10000, Vietnam
[2] School of Computer Science and Engineering, The University of Aizu, Aizuwakamatsu 965-0006, Japan

©IEICE 2026
DOI: 10.23919/comex.2025XBL0136
Received October 22, 2025
Accepted November 19, 2025

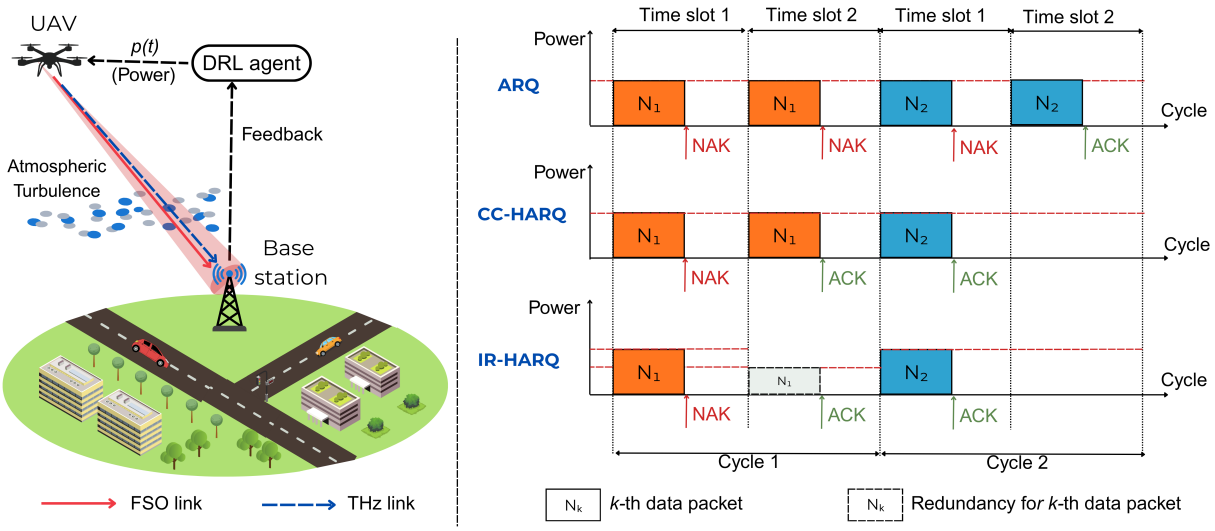

Fig. 1 (a) System model with hybrid FSO/THz link and (b) An illustration of transmit power for the operational differences between ARQ, CC-HARQ, and IR-HARQ.

2.1 FSO Channel Modeling

For the FSO link, the major impairments include atmospheric loss h_F^l, atmospheric turbulence h_F^t, and pointing error loss h_F^p. The composite FSO channel coefficient is given as $h_F = h_F^l h_F^t h_F^p$.

As for the atmospheric turbulence, it is modeled by the Log-Normal distribution. The probability density function (PDF) of h_F^t can be described as [7, Eq. (14)]

$$f_{h_F^t}\left(h_F^t\right) = \frac{1}{\sqrt{2\pi}\sigma_R h_F^t} \exp\left[\frac{-1}{2}\left(\frac{\ln\left(h_F^t\right) + \frac{\sigma_R^2}{2}}{\sigma_R}\right)^2\right], \quad (1)$$

where σ_R^2 is the Rytov variance, which determines the strength of the turbulence. It can be given as

$$\sigma_R^2 = 2.25k^{7/6}\sec^{11/6}(\xi)\int_{h_B}^{h_U} C_n^2(h)\,(h-h_B)^{5/6}dh, \quad (2)$$

where $k = 2\pi/\lambda_F$ with λ_F is the FSO wavelength, ξ is the zenith angle of the UAV, h_U and h_B are the height of the UAV and BS, respectively, $C_n^2(h)$ is the variation of the refractive index structure parameter described by the Hufnagel-Valley model [8], and sec(.) is the secant function.

The atmospheric loss represents the power reduction along the propagation path, which can be expressed as $h_F^l = 10^{\frac{-\kappa L}{10}}$ where $L = (h_U - h_B)/\cos(\xi)$ is the transmit distance, and κ is the weather-dependent attenuation coefficient.

The pointing error loss with the beam profile follows a Gaussian distribution, and the detecting aperture is circular in shape, which can be expressed as [9, Eq. (9)]

$$h_F^p = A_0^F \exp\left(-\frac{2(\rho_F)^2}{w_{eq}^2}\right), \quad (3)$$

where ρ_F is the radial distance between the beam and the lens center. $A_0^F = \mathrm{erf}[v_F]$ denotes the maximum optical power at the $\rho_F = 0$, $v_F = (\sqrt{\pi}a_F)/(\sqrt{2}w_d)$. a_F is the radius of the BS detector, $w_{eq}^2 = w_d^2 \frac{\sqrt{\pi}\mathrm{erf}(v_F)}{2v_F \exp(-v_F^2)}$, and w_d is the beam width at the distance L can be found in [9].

2.2 THz Channel Modeling

Major impairments considered for the THz link include path loss h_T^l, fading h_T^f, and pointing misalignment h_T^p. The composite THz channel coefficient is given as $h_T = h_T^l h_T^f h_T^p$.

Regarding path loss, it contains free-space loss and molecular attenuation, which is expressed as [6, Eq. (2)]

$$h_T^l = \frac{c\sqrt{G_t G_r}}{4\pi f_c^T L}\exp\left(-\frac{1}{2}\kappa_\alpha^T(f_c^T)L\right), \quad (4)$$

where c is the speed of light, f_c^T is the THz carrier frequency, $\kappa_\alpha^T(f_c^T)$ is the absorption coefficient found in [10] and G_t and G_r are the gain of the transmitter and receiver antennas, respectively.

The fading effect and the pointing misalignment $h_T^{fp} = x = h_T^f h_t^p$, its PDF can be expressed as [6, Eq. (3)]. The α-μ distribution models the fading effect, and the pointing misalignment is modeled by the Rayleigh distribution.

$$f_{h_T^{fp}}(x) = \rho^2 A_0^{-\rho^2} \frac{\mu^{\frac{\rho^2}{\alpha}}}{\Omega^\alpha \Gamma(\mu)} x^{\rho^2-1}\Gamma\left(\frac{\alpha\mu-\rho^2}{\alpha}, \mu\frac{x^\alpha}{\Omega^\alpha}A_0^{-\alpha}\right), \quad (5)$$

where $\alpha > 0$ describes the nonlinearity of the propagation medium, $\mu > 0$ is the inverse of the normalized variance of $(h_T^f)^\alpha$, Ω is the α-root mean value of the fading channel envelop, $A_0 = [\mathrm{erf}(u)]^2$ with $u = (\sqrt{\pi}a_T)/(\sqrt{2}w_d)$ where w_d is the beam waist at the propagation distance, and a_T is the detector radius of the BS, $\rho = w_d/(2\sigma_s)$ with $\sigma_s$is the spatial jitter standard deviation. $\Gamma(.)$ and $\Gamma(.,.)$ are the Gamma and the upper incomplete Gamma functions.

2.3 Switching Scheme for Hybrid FSO/THz links

In this system, the FSO link is considered the primary link and the THz link as the backup link. Based on channel condition, the FSO link or the THz link is determined to be active in a given time slot t. The switching threshold for the hybrid FSO/THz link can be calculated as $\gamma_{th} = 2^N - 1$, where $\gamma_F = [\mathcal{R}P_F(t)h_F]^2/\sigma_F^2$ and $\gamma_T = P_T(t)(h_T)^2/\sigma_T^2$ present the instantaneous received SNR of the FSO link and THz link. Where $\mathcal{R}$ is the responsivity of the FSO link, σ_F^2,

σ_T^2. $P_F(t)$, $P_T(t)$ are the noise variance and transmit power for the FSO and THz links, respectively. When $\gamma_F < \gamma_{\text{th}}$, the BS sends a one-bit feedback signal to the UAV to activate the backup THz link [1]. Conversely, the FSO link remains active whenever $\gamma_F \geq \gamma_{\text{th}}$.

Let $b_F(t)$ and $b_T(t)$ denote the flag variable of the FSo and THz link, respectively. It is used to indicate which link is used at time slot t, where $b_F(t) = 1$ and $b_T(t) = 0$ indicate the FSO link to be used at the time slot t and otherwise. The transmit power of the hybrid FSO/THz link at the time slot t can be calculated as $p(t) = b_F(t)P_F(t) + b_T(t)P_T(t)$.

2.4 HARQ with Power Control

In the considered system, the time interval considered is one cycle, and each cycle is divided into time slots t whose time is considered static. Erroneous data packets are retransmitted with a maximum of M transmissions allowed, and each packet has N bits.

Example: Fig. 1(b) depicts the transmit power of different HARQ protocols with maximum transmission $M = 2$ time slots. In ARQ, a packet is discarded after the maximum number of retransmissions if decoding fails (e.g., packet N_1 in cycle 1). The required transmit power in ARQ is comparable to that of CC-HARQ since the entire packet is resent in each retransmission. However, CC-HARQ combines previously received packets to enhance decoding reliability. In contrast, IR-HARQ retransmits only redundancy bits, thereby reducing transmit power (as in the time slot 2 of cycle 1).

For ARQ, CC-HARQ, and IR-HARQ schemes, the outage probability of a packet after the t-th transmission round can be respectively expressed as

$$P_t^{\text{ARQ}} = \Pr(I_1 < N, , I_2 < N, , \ldots, , I_t < N), \quad (6)$$

$$P_t^{\text{IR,CC}} = \Pr(I_1 < N, , I_2^{\Sigma} < N, , \ldots, , I_t^{\Sigma} < N), \quad (7)$$

where n is the channel uses, and $I_i = n\log_2(1+\gamma_i)$, $I_t^{\Sigma} = \sum_{i=0}^{t} I_i$ are the mutual information at the i-the transmission with $i \in \{1, 2, \ldots, t\}$.

3. PPO-Based Transmit Power Optimization

Our goal is to minimize the average transmit power of the UAV per cycle while ensuring the successful transmission of N bits with at maximum M transmissions. To do so, the transmit power $p(t)$ is adjusted at each transmission.

3.1 Problem Formulation

Let L be the successful decoded bits at time slot t. $P_F^{\max}$ and $P_T^{\max}$ are the maximum transmit power of FSO and THz links, respectively. The power optimization problem is formulated as

$$\mathcal{P}: \min \quad \frac{1}{M}\sum_{t=1}^{M} p(t) \quad (8a)$$

$$\text{s.t.} \quad p(t) = b_F(t)P_F(t) + b_T(t)P_T(t), \quad (8b)$$

$$L \geq N | t \leq M, \quad (8c)$$

$$0 \leq P_F(t) \leq P_F^{\max}, 0 \leq P_T(t) \leq P_T^{\max}. \quad (8d)$$

[1] The feedback channel is assumed to be error-free and have negligible latency

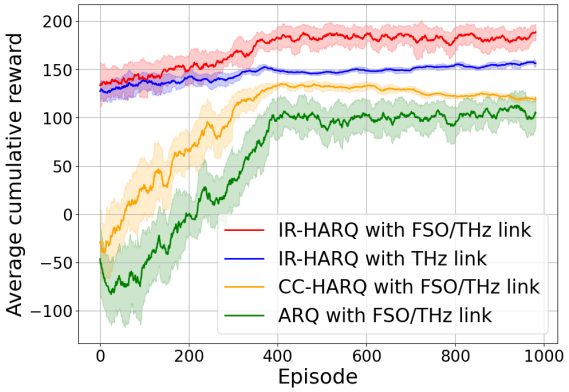

Fig. 2 Average cumulative reward convergence versus training episodes for different HARQ protocols with $N = 400$ bytes, $M = 2$ retransmissions.

3.2 MDP Reformulation

To proceed with the DRL-based algorithm, $\mathcal{P}$ is reformulated by the Markov Decision Process (MDP) as follows.

State Space: At time slot t, the state space $\mathcal{S}_t$ is treated as the input of the neural network to determine the subsequent action, i.e., $\mathcal{S}_t = \{t, N - L\}$.

Action Space: The action space $\mathcal{A}_t$ consists of the transmit power $\mathcal{A}_t = \{p(t)\}$.

Reward Function: It is formulated to minimize the average transmit power subject to the satisfaction of target performance constraints

$$r(t) = \exp\left[\beta \max\left(0, \frac{L-N}{N}\right) - \alpha p(t)\right] - \delta I(t > M), \quad (9)$$

3.3 PPO-Based Algorithm

To solve the formulated MDP, the PPO algorithm is employed [11]. PPO consists of two neural networks: a policy network and a value network. The value network, parameterized by ϕ, estimates the valure state $V_\phi(\mathcal{S}_t)$ by minimizing the loss $L_t^V(\phi) = \mathbb{E}_t\left[\left(V_\phi(\mathcal{S}_t) - \hat{R}_t\right)^2\right]$ where $\hat{R}_t = r_t + \gamma_{\text{DIS}} r_{t+1} + \ldots + \gamma_{\text{DIS}}^{T-t+1} r_{T-1} + \gamma_{\text{DIS}}^{T-t} V(\mathcal{S}_{T+1})$ is rewards-to-go and γ_{DIS} is the discount factor. The policy network, parameterized by θ, is optimized using the clipped surrogate objective $L_t^P(\theta) = \min\left\{\text{clip}(\rho_t(\theta), 1-\varepsilon, 1+\varepsilon)\hat{A}_t, \rho_t(\theta)\hat{A}_t\right\}$ where $\rho_t(\theta) = \frac{\pi_\theta(\mathcal{A}_t|\mathcal{S}_t)}{\pi_{\theta_{\text{old}}}(\mathcal{A}_t|\mathcal{S}_t)}$ and $\hat{A}_t$ is the advantage estimated via generalized advantage estimation (GAE) [12]. The loss of the policy network can be expressed as $L_t^{P+S}(\theta) = \mathbb{E}_t\left[L_t^P(\theta) + cS_{\pi_0}(\mathcal{S}_t)\right]$ where $S_{\pi_0}(\mathcal{S}_t)$ is the policy entropy and c is a weighting coefficient.

4. Numerical Results and Discussions

In our considered system, the parameters used in the analysis are as follows. For the FSO channel, wavelength $\lambda_F = 1550$ nm, zenith angle $\xi = 60°$, radius $a_F = 0.1$ m, esponsivity $\mathcal{R} = 0.9$, noise variance $\sigma_F^2 = -60$ dBm, and power $P_F^{\max} = 30$ dBm. For the THz channel, frequency $f_c^T = 0.3$ THz, $G_t = G_r = 55$ dBi, radius $a_T = 0.1$ m, fading parameters $(\alpha, \mu) = (2, 1)$, noise variance $\sigma_T^2 = -60$ dBm, and power $P_T^{\max} = 30$ dBm. Other parameters, $c = 3 \times 10^8$ m/s, $h_U = 100$ m, $h_B = 20$ m and channel use $n = 100$. For PPO, we use the PPO clip variant with a clip ratio $\epsilon = 0.2$,

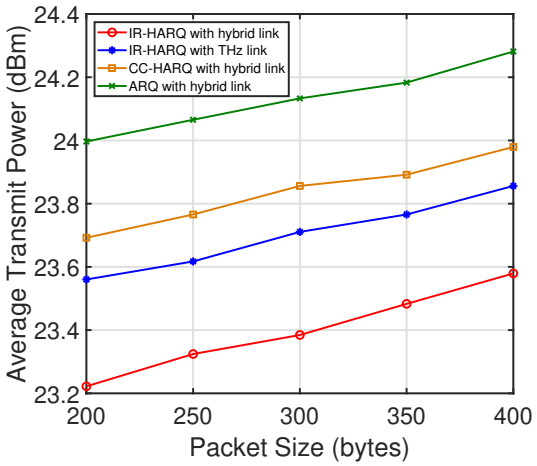

Fig. 3 Average transmit power per transmission cycle versus packet size for different HARQ protocols with $M = 2$ retransmissions

$\lambda = 0.95$, $\gamma_{\text{DIS}} = 0.99$, $\alpha = 2$, $\beta = 0.5$, and $\delta = 10$. The value and policy networks are trained with a learning rate of 4×10^{-4} and 2×10^{-4}, respectively. The batch size is 2048, while the minibatch size is 128. The hidden layers of both networks have 128 neurons, and the Tanh function is used as the activation function and each episode have 1000 steps.

Figure 2 illustrates the performance comparison of the HARQ protocols optimized by the proposed method for $N = 400$ bytes and $M = 2$ time slots. As observed, all protocols achieve convergence after approximately 400 training episodes, confirming the stability of the learning process. The IR-HARQ scheme attains the highest reward (≈ 200), clearly outperforming CC-HARQ (≈ 130) and ARQ (≈ 100). Moreover, the hybrid FSO/THz system achieves a higher reward compared to the single THz link (≈ 150), demonstrating the effectiveness of the proposed hybrid configuration in reducing transmission power while maintaining reliable communication performance.

Figure 3 presents the average transmit power per transmission cycle for different HARQ mechanisms and compares the hybrid FSO/THz system with the system employing only the THz link. The results indicate that the IR-HARQ protocol achieves the lowest average transmit power among the three HARQ protocols. Specifically, for $N = 400$ bytes, the average transmit power required by the IR-HARQ protocol is approximately 23.6 dBm, while CC-HARQ and ARQ require 24 dBm and 24.3 dBm, respectively. Moreover, the IR-HARQ system operating solely over the THz link demonstrates better power efficiency compared to the hybrid FSO/THz system when using CC-HARQ and ARQ protocols. However, when the hybrid system employs the IR-HARQ mechanism, it achieves a significantly lower transmit power—around 23.6 dBm for $N = 400$ bytes—highlighting its superior performance. These results confirm that the proposed hybrid FSO/THz system combined with the IR-HARQ protocol can effectively minimize transmission power while maintaining reliable data delivery.

5. Conclusion

This letter advances 6G-enabling UAV communications by introducing a PPO-DRL framework for dynamic transmit power optimization in hybrid FSO/THz systems augmented with IR-HARQ. By formulating the problem as an MDP that encapsulates attenuation, fading, and misalignment impairments, the proposed agent achieves adaptive power scaling across retransmissions and SNR-based FSO/THz switching, with FSO as the main link and THz as a backup. This is tested using outage probabilities that are specific to heterogeneous atmospheric regimes, underscoring the framework's robustness and affirming hybrid IR-HARQ's pivotal role in fostering energy-efficient, resilient aerial architectures for disaster-resilient backhauls and ISAC paradigms. Future extensions may incorporate multi-UAV swarms and real-time beam tracking to further propel sustainable 6G deployments.

References

[1] A. Masaracchia, Y. Li, K. K. Nguyen, C. Yin, S. R. Khosravirad, D. B. D. Costa, and T. Q. Duong, "Uav-enabled ultra-reliable low-latency communications for 6g: A comprehensive survey," *IEEE Access*, vol. 9, pp. 137 338–137 352, 2021.

[2] S. A. Al-Gailani, M. F. Mohd Salleh, A. A. Salem, R. Q. Shaddad, U. U. Sheikh, N. A. Algeelani, and T. A. Almohamad, "A survey of free space optics (fso) communication systems, links, and networks," *IEEE Access*, vol. 9, pp. 7353–7373, 2021.

[3] W. Jiang, Q. Zhou, J. He, M. A. Habibi, S. Melnyk, M. El-Absi, B. Han, M. D. Renzo, H. D. Schotten, F.-L. Luo, T. S. El-Bawab, M. Juntti, M. Debbah, and V. C. M. Leung, "Terahertz communications and sensing for 6g and beyond: A comprehensive review," *IEEE Commun. Surveys & Tutorials*, vol. 26, no. 4, pp. 2326–2381, 2024.

[4] H. D. Le, C. T. Nguyen, T. K. Nguyen, and A. T. Pham, "Hybrid fso/sub-thz-based vertical networks for internet of vehicles," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 60, no. 2, pp. 1865–1881, 2024.

[5] G. D. Chondrogiannis, N. A. Mitsiou, N. D. Chatzidiamantis, A.-A. A. Boulogeorgos, and G. K. Karagiannidis, "Power-optimal harq protocol for reliable free space optical communication," in *2023 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2023, pp. 1765–1770.

[6] T. D. Nguyen, C. T. Nguyen, H. D. le, and A. T. Pham, "Non-orthogonal harq for uav-aided thz systems: Design and analysis," *IEICE Communications Express*, vol. 14, no. 4, pp. 151–154, 2025.

[7] S. Song, M. Choi, D.-E. Ko, and J.-M. Chung, "Multi-uav trajectory optimization considering collisions in fso communication networks," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 11, pp. 3378–3394, 2021.

[8] H. Kaushal and G. Kaddoum, "Optical communication in space: Challenges and mitigation techniques," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 57–96, 2017.

[9] T. V. Nguyen, H. D. Le, N. T. Dang, and A. T. Pham, "On the design of rate adaptation for relay-assisted satellite hybrid fso/rf systems," *IEEE Photonics Journal*, vol. 14, no. 1, pp. 1–11, 2022.

[10] O. S. Badarneh, F. E. Bouanani, and F. Almehmadi, "A general framework for uav-aided thz communications subject to generalized geometric loss," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 11, pp. 14 589–14 600, 2023.

[11] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *ArXiv*, vol. abs/1707.06347, 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:28695052

[12] J. Schulman, P. Moritz, S. Levine, M. Jordan, and P. Abbeel, "High-dimensional continuous control using generalized advantage estimation," 2018. [Online]. Available: https://arxiv.org/abs/1506.02438