Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Physics-Informed Multi AI-Agent Localization-Assisted THz--OWC L3-Aware MPTCP-Inspired Dual-Subflow Cross-Layer Framework for Extreme-Band 6G Networks

Emmanuel U. Ogbodo

ogbodoeu@ieee.org

National Institute of Telecommunications

Luciano L. Mendes

National Institute of Telecommunications

Article

Keywords: 6G, terahertz, optical wireless communication, L3-aware routing, MPTCP-inspired, dual subflow, localization, physics-informed artificial intelligence, Double Deep Q-Network, graph attention network, soft actor-critic, Sionna ray tracing, ns-3, 5G-LENA

Posted Date: September 14th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10851492/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Physics-Informed Multi AI-Agent Localization-Assisted THz–OWC L3-Aware MPTCP-Inspired Dual-Subflow Cross-Layer Framework for Extreme-Band 6G Networks

Emmanuel U. Ogbodo[1*] and Luciano L. Mendes[1]

[1*]xGMobile, National Institute of Telecommunications (Inatel), Santa Rita do Sapucai, Minas Gerais, Brazil.

*Corresponding author(s). E-mail(s): ogbodoeu@ieee.org;
Contributing authors: lucianol@inatel.br;

Abstract

Terahertz (THz) and optical wireless communication (OWC) are complementary technologies for sixth-generation (6G) networks, but their practical integration is challenged by severe THz propagation loss and blockage, OWC geometric and field-of-view constraints, mobility, and localization uncertainty. These effects influence not only physical-layer link quality but also Layer 3 (L3) path selection and transport-layer traffic distribution. Existing studies commonly address these functions separately, leaving a gap in reproducible cross-layer frameworks that jointly connect heterogeneous physical states and localization information with L3 path intelligence, multipath transport adaptation, and packet-level network performance.

This work proposes a physics-informed (PI) multi-agent artificial intelligence (AI) framework with a Multipath TCP (MPTCP)-inspired THz–OWC dual-subflow architecture. The intelligent PI-DDQN–GAT–SAC strategy combines a Double Deep Q-Network (DDQN) for physics-constrained access and stability decisions, a Graph Attention Network (GAT) for heterogeneous THz–OWC L3 path ranking and selection, and Soft Actor-Critic (SAC) for continuous transport-layer subflow allocation. A complementary PI-Static 50/50 strategy provides a deterministic non-learning dual-path reference for persistent parallel utilization. The framework combines Sionna ray-tracing states for the THz–OWC environment, with the OWC branch additionally anchored to calibrated optical-channel parameters, together with localization information. The resulting cross-layer decisions are evaluated through packet-level ns-3/5G-LENA simulations.

The results demonstrate coordinated adaptation across the network and transport layers: as THz propagation conditions deteriorate, PI-GAT increasingly favors suitable OWC access paths, while PI-SAC correspondingly shifts a larger fraction of transport traffic toward the OWC subflow. Under the highest evaluated traffic load, PI-DDQN–GAT–SAC improves throughput by approximately 10–12% over strongest-link selection, which serves as the principal conventional adaptive baseline. THz-only operation is limited by exclusive dependence on the terahertz radio link and its susceptibility to severe blockage and propagation loss, whereas OWC-only operation is constrained by optical geometry, receiver field-of-view, orientation, and blockage. Strongest-link selection adapts between THz and OWC but remains a single-path strategy and therefore cannot exploit simultaneous THz–OWC dual-subflow transmission. PI-Static 50/50 achieves the highest raw parallel capacity when both links remain persistently feasible, revealing a complementary trade-off between fixed dual-path utilization and adaptive L3-aware transport intelligence.

Keywords: 6G, terahertz, optical wireless communication, L3-aware routing, MPTCP-inspired, dual subflow, localization, physics-informed artificial intelligence, Double Deep Q-Network, graph attention network, soft actor-critic, Sionna ray tracing, ns-3, 5G-LENA

1 Introduction

Future sixth-generation (6G) systems are expected to combine very large bandwidths, directional links, spatial awareness, and intelligent control. Terahertz (THz) and optical wireless communication (OWC) are prominent candidates in the extreme-band [1], but neither is uniformly superior: THz is strongly affected by spreading LOS, molecular absorption and blockage[2–4], whereas OWC depends on line-of-sight (LOS) geometry, field of view, orientation, and receiver conditions[5–8]. Their complementary behavior motivates coordinated use rather than independent link optimization.

The coordination problem extends beyond the physical layer. Changes in THz or OWC quality alter queueing, network-layer or Layer 3 (L3) access-path ranking, and transport scheduling. Heterogeneous multihoming can improve service diversity, but requires both a preferred network path and a traffic split rule across dissimilar links[9]. Multipath Transmission Control Protocol (MPTCP) motivates the use of multiple transport subflows[10, 11]; conventional scheduling, however, is not designed to consume ray-traced THz indicators, optical quality, localization uncertainty, and extreme-band feasibility jointly.

Physics-informed (PI) artificial intelligence (AI) provides a structured way to couple these decisions while enforcing hard feasibility. A Double Deep Q-Network (DDQN), Graph Attention Network (GAT), and Soft Actor-Critic (SAC) naturally match discrete access/stability preference, graph-structured L3 ranking, and continuous subflow allocation, respectively[12–14]. Related work on AI-enabled communication using 5G Reduced Capability (RedCap) devices demonstrates the value of learning for heterogeneous connectivity[15]. However, the Layer 3 (L3) bridge

between calibrated extreme-band states and multipath packet-level execution remains insufficiently quantified.

The research gap is therefore the lack of an implementation-driven framework that links the THz–OWC physical states and the uncertainty of the localization with physics-constrained access decisions, the heterogeneous Layer 3 (L3) path ranking, the continuous dual-subflow allocation, and the evaluation of the performance at the packet-level under increasing traffic load. We address this gap through an integrated physical-layer (PHY)/localization–L3–transport workflow, with end-to-end performance assessed using ns-3/5G-LENA key performance indicators (KPIs).

The contributions are:

- Develop a localization-assisted THz–OWC state and hard-feasibility architecture that connects physical information to the heterogeneous Layer 3 (L3) and transport decisions;
- Propose two dual-subflow strategies: PI-Static 50/50, a deterministic non-learning equal-split policy, and PI-DDQN–GAT–SAC, a physics-informed multi-agent strategy for adaptive L3-aware THz–OWC subflow control;
- Implement an MPTCP-inspired ns-3.48/5G-LENA New Radio (NR) v5.0 packet-level framework with exact per-state packet-budget conservation and separate THz-oriented and OWC-equivalent logical bandwidth-part impairments; and
- Evaluate the framework through a five-seed, 450-run campaign with 95% Student-t confidence intervals, including dedicated L3 analysis of THz–OWC medium preference and OWC access-node ranking.

2 Proposed physics-informed dual-subflow framework

2.1 End-to-end architecture

Figure 1 shows the end-to-end PHY/localization–L3–transport trace-to-packet workflow. The physical inputs comprise fixed Sionna RT THz states, calibrated 30-luminaire OWC optical states on the same mobility/localization trajectory, and localization error/uncertainty. After fusion into a canonical 300-state trace, a physics-feasibility gateway exposes the same feasible state to the fixed PI-Static 50/50 branch and the intelligent PI-DDQN–GAT–SAC branch. PI-DDQN forms the access/stability prior, PI-GAT performs the topology-aware L3 ranking, and PI-SAC converts physical/L3 information into continuous transport-subflow shares. Both branches are resolved into MPTCP-inspired THz–OWC packet actions and executed in ns-3/5G-LENA.

Let the fused physical state at time t be

$$\mathbf{s}_t = [\mathbf{s}_t^{\mathrm{THz}}, \mathbf{s}_t^{\mathrm{OWC}}, \mathbf{s}_t^{\mathrm{loc}}, \Delta\mathbf{s}_t, \mathbf{m}_t], \tag{1}$$

where $\mathbf{s}_t^{\mathrm{THz}}$ contains received-power, path-count, LOS-path, and earliest-delay information; $\mathbf{s}_t^{\mathrm{OWC}}$ contains OWC signal-to-noise ratio (SNR) and photocurrent information; $\mathbf{s}_t^{\mathrm{loc}}$ contains position error and uncertainty; $\Delta\mathbf{s}_t$ contains temporal changes; and $\mathbf{m}_t$ contains the physics-availability mask. Scenario labels are retained for campaign organization but are not privileged controller inputs.

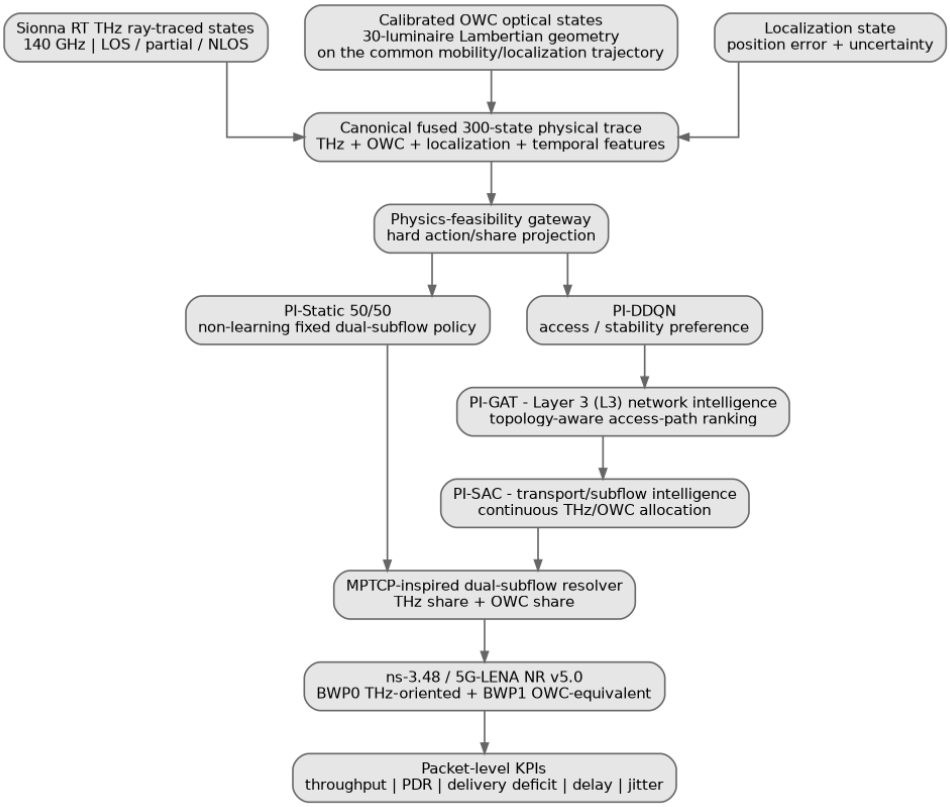


Fig. 1 Revised end-to-end architecture of the physics-informed localization-assisted THz–OWC L3-aware dual-subflow framework.

2.2 Two complementary proposed control strategies

The first control strategy, PI-Static 50/50, provides a fixed dual-path reference for persistent parallel utilization. When both links are physically feasible,

$$(\omega_{\mathrm{THz}}, \omega_{\mathrm{OWC}}) = (0.5, 0.5). \tag{2}$$

The fixed pair is passed through the same hard physics projection as all other actions. Thus, if only one medium is feasible, its share becomes one, and the unavailable medium receives zero. PI-Static 50/50 contains no learned intelligence and no channel-dependent weighting; its purpose is to quantify the benefit of continuous equal parallel use when both logical paths remain available.

The second strategy, PI-DDQN–GAT–SAC, is the intelligent multi-agent policy. As summarized in Fig. 2, DDQN provides the discrete access/stability prior, PI-GAT jointly ranks the THz access and OWC access-point (AP) tier at Layer 3 (L3), and PI-SAC determines the continuous THz transport share, with the OWC share as its

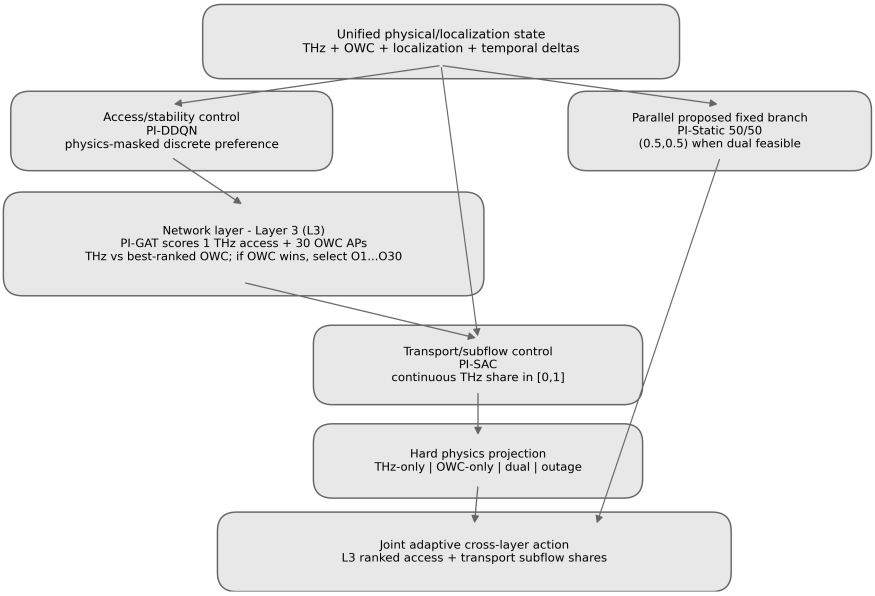


Fig. 2 Internal decision pipeline for PI-DDQN–GAT–SAC and the parallel PI-Static 50/50 branch.

complement when both media are feasible. PI-GAT reports the decision hierarchically as THz versus the best-ranked OWC candidate, followed by intra-OWC AP selection when OWC is preferred. Hard projection prevents allocation to an unavailable medium.

3 Methods

3.1 THz ray-traced physical states

The THz branch uses a fixed Sionna RT campaign[16, 17] at 140 GHz with 5 GHz bandwidth and 10 dBm transmit power in a $10 \times 7 \times 3.2$ m indoor environment. The transmitter is located at $(1.0, 3.5, 2.6)$ m and the receiver traverses 20 positions at 0.25 s intervals. Five seeds $(11, 17, 23, 29, 37)$ and three propagation conditions [LOS, partial blockage, and non-line-of-sight (NLOS)] produce $5 \times 3 \times 20 = 300$ physical states. The THz trace configuration is fixed before controller training.

For packet-level coupling, the ray-traced path loss is converted into a state-dependent excess loss rather than replacing the 5G-LENA propagation chain. For episode state t,

$$L_{\mathrm{THz,ex}}(t) = \max\left(0, L_{\mathrm{RT}}(t) - L_{\mathrm{ref}}\right), \tag{3}$$

where L_{ref} is the strongest finite LOS state when LOS exists. For a genuine NLOS episode with no LOS state, the strongest finite state from that same NLOS episode is used as the internal reference. This preserves the original ray-traced gains and avoids constructing an artificial LOS condition.

3.2 Calibrated OWC optical-state generation on the common mobility trajectory

The OWC branch uses the same receiver mobility/orientation trajectory that is fused with the Sionna RT THz campaign, but the optical channel itself is evaluated with the fixed calibrated Lambertian OWC implementation. It uses a regular 6×5 ceiling grid of 30 luminaires at $z = 3.1$ m. The fixed configuration uses a 27.13° half-power angle, a 13 mm^2 detector area, and a 100° receiver field of view. For an LOS optical transmitter i, the direct-current (DC) channel gain follows the standard Lambertian form[18, 19]

$$H_i(0) = \frac{(m+1)A}{2\pi d_i^2} \cos^m(\phi_i) T_s(\psi_i) g(\psi_i) \cos(\psi_i), \tag{4}$$

for $0 \le \psi_i \le \Psi_c$, and zero otherwise. Here $m = -\ln 2 / \ln(\cos \Phi_{1/2})$ is the Lambertian order, A is the detector area, d_i is separation, ϕ_i and ψ_i are the irradiance and incidence angles, T_s is optical filter gain, g is concentrator gain, and Ψ_c is the field of view.

The resulting calibrated optical state supplies the fixed OWC photocurrent and SNR inputs used by the controller and packet abstraction. A state-dependent equivalent loss is formed as

$$L_{\mathrm{OWC,eq}}(t) = \max\left(0, \gamma_{\mathrm{ref}} - \gamma_{\mathrm{OWC}}(t)\right), \tag{5}$$

where γ_{ref} is the strongest finite OWC SNR in the selected episode with 20-states. This preserves equal packet generation across controllers while allowing the OWC-equivalent packet path to respond to the fixed optical state.

3.3 Localization state and physics feasibility

Localization is represented by position error and position uncertainty derived from the trace-generation process[20, 21]. The uncertainty term is included because extreme-band connectivity depends not only on estimated position but also on confidence in geometry-sensitive path selection. A binary availability vector

$$\mathbf{m}_t = [m_t^{\mathrm{THz}}, m_t^{\mathrm{OWC}}, m_t^{\mathrm{dual}}] \tag{6}$$

is used to construct hard action masks and share projections. These constraints are applied during both learning and inference.

3.4 Physics-informed DDQN access/stability agent

The Double Deep Q-Network (DDQN) agent uses a 19-dimensional state formed from 15 prepared physical/localization features plus the previous-action one-hot vector. The discrete action space contains THz-only, OWC-only, dual, and outage actions. Physics-invalid actions are masked before selection. The online network chooses the next maximum action, and the target network evaluates it, following Double-DQN learning[12]. Two hidden layers of 128 units are used with $\gamma = 0.97$, learning rate 3×10^{-4}, replay capacity 10,000, batch size 64, and target synchronization every

100 updates. The training, validation and held-out partitions contain 180, 60, and 60 states, respectively; seed 37 is excluded from checkpoint selection.

The DDQN output is interpreted as an access/stability prior rather than a downstream hard prohibition. In the final fixed model, its 300-state action counts are 164 OWC-only and 136 THz-only, while the held-out seed contains 33 OWC-only and 27 THz-only selections.

3.5 PI-GAT Layer-3 topology-aware access-path ranking

The physics-informed Graph Attention Network (PI-GAT) operates on a 32-node graph comprising one user-equipment (UE) context node, one THz access node, and 30 spatially distributed OWC access-point (AP) nodes. The single THz access node represents one L3 access candidate informed by multiple ray-traced THz propagation components; it does not represent a single THz ray. PI-GAT jointly assigns a scalar L3 score to the 31 access candidates using 11 features per node and a UE-centered bidirectional star with self-loops. The first attention layer uses four heads with 16 hidden features per head, and the second uses two heads with 32 hidden features, followed by a scalar path scorer[13].

For interpretation, the joint argmax is decomposed exactly into two steps without changing the trained model or its decisions: first compare the THz score with the maximum score among the 30 OWC nodes (THz versus best-ranked OWC); if OWC wins, select the OWC AP attaining that maximum. This separates the medium-level preference from the intra-OWC AP ranking while preserving the original decision of 31-candidates. The physics-availability mask remains the hard L3 feasibility constraint, and DDQN enters as an access/stability prior. The trained PI-GAT reaches 98.33% top-1 agreement on the validation and held-out test states, with held-out mean regret 1.51×10^{-4}.

3.6 PI-SAC continuous dual-subflow allocation

The Soft Actor-Critic (SAC) agent receives 22 prepared input features containing normalized THz/OWC physical indicators, localization information, DDQN Q-values, L3 PI-GAT route indicators, and the PI-GAT predicted path score. Its internal state dimension is 23. The actor maps its Gaussian policy through a hyperbolic tangent and then to the interval $[0, 1]$ to obtain the requested THz share,

$$\omega_{\mathrm{THz}}^{\mathrm{raw}} = \frac{\tanh(z) + 1}{2}, \tag{7}$$

$$\omega_{\mathrm{OWC}}^{\mathrm{raw}} = 1 - \omega_{\mathrm{THz}}^{\mathrm{raw}}, \tag{8}$$

where z is the continuous output of the Soft Actor-Critic (SAC) policy before conversion to a bounded traffic-allocation share. The terms $\omega_{\mathrm{THz}}^{\mathrm{raw}}$ and $\omega_{\mathrm{OWC}}^{\mathrm{raw}}$ denote the preliminary fractions of traffic assigned to the THz and OWC subflows, respectively. The superscript "raw" indicates that these shares are obtained before the physics-based feasibility projection is applied. The requested action is then projected onto the physics-feasible set. If only THz is available the shares are $(1, 0)$; if only OWC

is available they are $(0, 1)$; if neither is available they become $(0, 0)$; and when both are available the learned complementary shares are retained. The SAC implementation uses twin critics, target critics, automatic entropy temperature, two 128-unit hidden layers, $\gamma = 0.97$, $\tau = 0.005$, and learning rate 3×10^{-4}[14]. The held-out mean reward is 0.70834, and the held-out THz share spans 0.0247–0.9781. These rewards are controller-training surrogates; final communication KPIs are measured only in the packet-level campaign.

This decomposition is consistent with the broader use of AI to coordinate multiple network layers and heterogeneous links[22], but the present implementation restricts the executed agent chain to the three validated DDQN, GAT, and SAC modules.

3.7 L3-informed MPTCP-inspired dual-subflow packet execution

The final packet implementation is built on ns-3.48 with 5G-LENA New Radio (NR) v5.0[23]. L3-ranked access information and transport shares are converted into two logical downlink subflows mapped to separate bandwidth parts (BWPs): BWP0 is THz-oriented and BWP1 is OWC-equivalent. The standard 5G-LENA channels are retained and augmented with the state-dependent losses in Eqs. (3) and (5). BWP0 follows the existing urban-microcell (UMi)/Third Generation Partnership Project (3GPP) propagation chain plus Sionna excess loss, whereas BWP1 follows its NR chain plus OWC-equivalent SNR-deficit loss. The BWPs therefore act as model-equivalent packet carriers for cross-layer impairment and scheduling experiments.

Traffic is generated through independent bounded User Datagram Protocol (UDP) windows, one window per 0.25 s physical/controller state. For offered packet rate λ and state duration Δt, the aggregate budget is

$$N_t = \text{round}(\lambda \Delta t). \tag{9}$$

The THz and OWC budgets are

$$N_{\text{THz},t} = \text{round}\left(N_t \omega_{\text{THz},t}\right), \tag{10}$$

$$N_{\text{OWC},t} = N_t - N_{\text{THz},t}, \tag{11}$$

where N_t is the total integer packet budget assigned to controller state t, $\omega_{\text{THz},t}$ is the final THz traffic-allocation share after feasibility projection, and $N_{\text{THz},t}$ and $N_{\text{OWC},t}$ are the corresponding numbers of packets assigned to the THz and OWC subflows. The operator round$(\cdot)$ rounds its argument to the nearest integer.

This formulation guarantees exact packet-budget conservation,

$$N_{\text{THz},t} + N_{\text{OWC},t} = N_t, \tag{12}$$

for every controller state. It is particularly important for PI-Static 50/50 when N_t is odd, because the two subflow budgets cannot both be exactly equal integers.

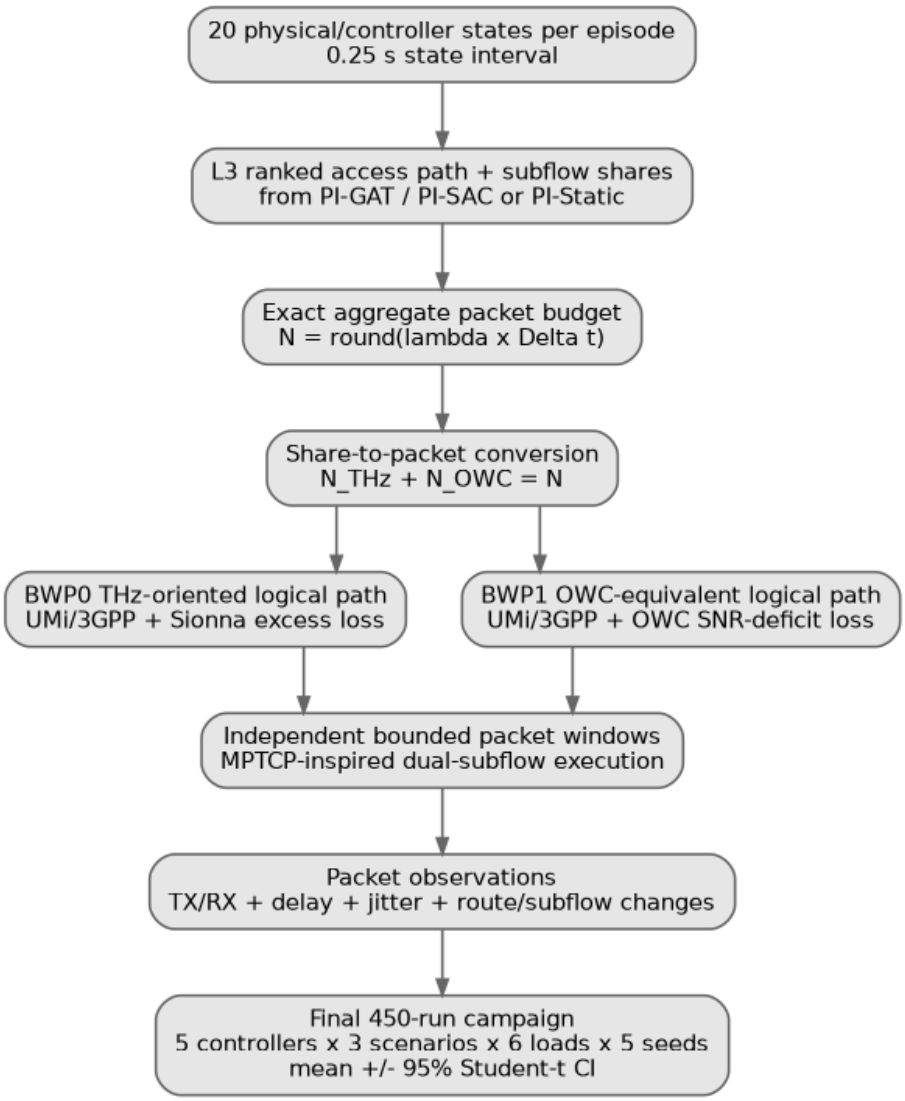


Fig. 3 Trace-to-packet implementation and campaign design.

Figure 3 summarizes the trace-to-packet bridge. Traffic is generated for 5.0 s, followed by a drain period of 0.6 s within the simulation of 6.0 s. A simulation-stop sensitivity analysis at 5.5, 6.0, 7.0, and 8.0 s confirmed that stressed receive counts stabilize by 6.0 s. For each controller state, an exact integer packet budget is partitioned between the bandwidth part 0 (BWP0) and the bandwidth part 1 (BWP1) before packet execution.

Algorithm 1 Physics-informed THz–OWC L3-aware dual-subflow control and packet execution

Require: trained THz trace $\mathcal{T}$, OWC trace $\mathcal{O}$, localization trace $\mathcal{L}$, controller mode c, packet rate λ

Ensure: Per-state THz/OWC packet allocation and packet-level performance record

Fuse $\mathcal{T}$, $\mathcal{O}$, and $\mathcal{L}$ into canonical states $\{\mathbf{s}_t\}_{t=1}^{20}$
Build physics mask $\mathbf{m}_t$ for every state
for each state t **do**
    if $c =$ PI-Static 50/50 **then**
        $(\omega_T, \omega_O) \leftarrow \Pi_{\mathcal{F}(\mathbf{m}_t)}(0.5, 0.5)$
    else if $c =$ PI-DDQN–GAT–SAC **then**
        Obtain physics-masked DDQN access/stability prior
        PI-GAT scores the THz access and all feasible OWC APs
        Resolve L3 medium as THz versus the best-ranked OWC AP; if OWC wins, retain that AP
        Form SAC state from physical, DDQN, and L3 PI-GAT features
        $\omega_T \leftarrow \pi_{\mathrm{SAC}}(\mathbf{s}_t)$; $\omega_O \leftarrow 1 - \omega_T$
        $(\omega_T, \omega_O) \leftarrow \Pi_{\mathcal{F}(\mathbf{m}_t)}(\omega_T, \omega_O)$
    else
        Resolve THz-only, OWC-only, or standardized strongest-link baseline
    end if
    $N_t \leftarrow \mathrm{round}(\lambda \Delta t)$
    $N_T \leftarrow \mathrm{round}(N_t \omega_T)$; $N_O \leftarrow N_t - N_T$
    Update Sionna excess loss on BWP0 and OWC-equivalent loss on BWP1
    Install bounded THz and OWC packet windows with N_T and N_O packets
end for
Run ns-3/5G-LENA until the fixed post-traffic drain expires
Export throughput, packet delivery ratio, delivery deficit, delay, jitter, route changes, and subflow changes

Algorithm 1 gives the complete controller and packet-execution procedure used in the final experiments.

3.8 Experimental design, baselines, and statistics

Table 1 lists the main trained parameters. Five controller modes are compared. THz-only, OWC-only, and standardized strongest-link are conventional comparison baselines, while PI-Static 50/50 and PI-DDQN–GAT–SAC are the two proposed strategies. Strongest-link selects one medium at each state by comparing the training-standardized THz received-power feature and OWC signal-to-noise ratio (SNR) feature after physical-availability checks. It therefore serves as a practical heuristic baseline based on instantaneous link-quality indicators, without assuming perfect knowledge of the achievable data rate of each medium.

Table 1 Implementation and campaign parameters.

Parameter	Implementation value
THz ray-trace carrier / bandwidth	140 GHz / 5 GHz
THz transmit power	10 dBm
Room / trace step	$10 \times 7 \times 3.2$ m / 0.25 s
States per episode	20
Propagation scenarios	LOS, partial blockage, NLOS
Campaign seeds	11, 17, 23, 29, 37
OWC deployment	30 luminaires, regular 6×5 ceiling grid
OWC half-power angle / detector / field of view	27.13° / 13 mm^2 / 100°
DDQN state / hidden layers	19 / 128×128
L3 PI-GAT graph / node features	32 nodes (1 UE context + 1 THz access + 30 OWC APs); 11 features/node
SAC input features / internal state	22 / 23
Packet simulator	ns-3.48 + 5G-LENA NR v5.0
Logical packet paths	BWP0 THz-oriented; BWP1 OWC-equivalent
Packet size	1200 bytes payload
Offered packet rates	2,500; 5,000; 10,000; 15,000; 20,000; 25,000 packets/s
Nominal payload loads	24; 48; 96; 144; 192; 240 Mb/s
Traffic duration / simulation time	5.0 s / 6.0 s
Post-traffic drain	0.6 s
Final campaign size	5 controllers $\times$3 scenarios $\times$6 loads $\times$5 seeds = 450 runs

For each experimental cell, the five-seed mean is reported with a two-sided 95% Student-t confidence interval (CI),

$$\mathrm{CI}_{95} = \bar{x} \pm t_{0.975,4} \frac{s}{\sqrt{5}}, \qquad t_{0.975,4} = 2.776445. \tag{13}$$

Seeds 11, 17, and 23 are used for controller training, seed 29 for validation, and seed 37 as the strict held-out test seed. The five-seed confidence intervals are therefore interpreted as full-campaign robustness summaries, while seed 37 provides the out-of-sample generalization check.

The principal packet key performance indicators (KPIs) are

$$R = \frac{8B_{\mathrm{rx}}}{T}, \tag{14}$$

$$\mathrm{PDR} = \frac{N_{\mathrm{rx}}}{N_{\mathrm{tx}}}, \tag{15}$$

$$D_{\mathrm{def}} = 1 - \mathrm{PDR}, \tag{16}$$

where R is the throughput in bits per second, B_{rx} is the total number of successfully received bytes, and T is the controlled traffic observation interval. The packet delivery ratio (PDR) is the fraction of transmitted packets successfully received, where N_{tx} and N_{rx} denote the numbers of transmitted and received packets, respectively. The delivery-deficit ratio, D_{def}, represents the fraction of transmitted packets not received within the observation interval.

Mean delay is the average end-to-end time between packet transmission and successful reception. Jitter is the variation in packet delay over successive received packets and therefore indicates the temporal stability of packet delivery. A Layer 3 (L3) route change occurs when the ranked or selected access path changes between consecutive controller states, while a subflow change occurs when the transport-layer traffic-allocation decision changes.

FlowMonitor-classified lost packets are recorded separately from the delivery-deficit ratio. This distinction is necessary because a packet that has not been received by the end of the observation interval is counted in D_{def}, but may not yet be classified by FlowMonitor as a lost packet.

4 Results

4.1 Multi-agent and heterogeneous L3 validation

The trained agents were evaluated using the predefined training, validation, and held-out test sets. The training set comprises seeds 11, 17, and 23, the validation set uses seed 29, and seed 37 is reserved as the held-out test seed. DDQN achieved mean rewards of 0.76007, 0.76021, and 0.76009 on the training, validation, and held-out test sets, respectively. PI-GAT achieved 98.33% held-out top-1 path-ranking agreement with a mean regret of 1.51×10^{-4}, while PI-SAC achieved a held-out reward of 0.70834. These values assess controller performance, while the final communication performance is evaluated using packet-level results.

Figure 4 presents the PI-GAT decisions at both the medium and the OWC access-point levels. Panel (a) shows that the THz–OWC L3 preference changes from 49/51% under line-of-sight (LOS) conditions to 30/70% under partial blockage and 10/90% under non-line-of-sight (NLOS) conditions. Panel (b) shows the selected OWC access points (APs) when OWC is preferred. O9 is selected most frequently in LOS (29.4%) and partial blockage (21.4%), while O9, O16 and O23 each represent 16.7% of the NLOS selections. The corresponding scenario-average top-1 agreements are 99%, 94%, and 100%.

Panel (c) compares the L3 medium preference with the continuous transport-layer allocation. PI-SAC increases the OWC traffic share from 37.2% in LOS to 50.9% under partial blockage and 77.8% in NLOS. Across the 15 observational sequences, the preference for PI-GAT OWC and the traffic share of PI-SAC OWC exhibit a

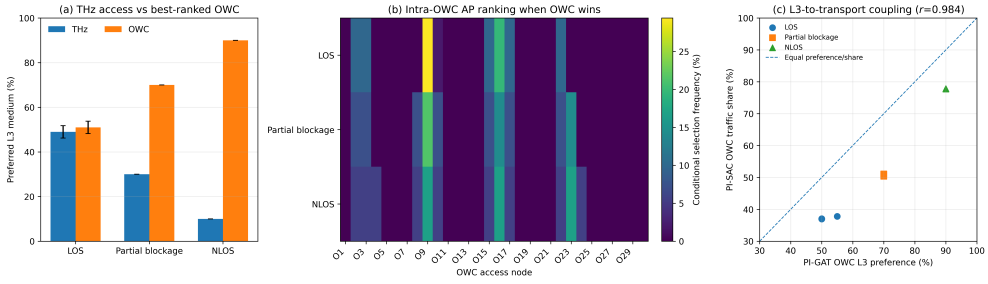


Fig. 4 Heterogeneous L3-to-transport validation. (a) Exact hierarchical view of the joint PI-GAT argmax: THz access versus the best-ranked OWC candidate, with five-seed 95% CIs. (b) Conditional intra-OWC ranking over the 30 APs when OWC wins. (c) Per-seed relationship between PI-GAT OWC L3 preference and PI-SAC OWC traffic share; the dashed line denotes equal preference and allocation.

strong Pearson correlation of $r = 0.984$. The two values need not be identical because PI-GAT ranks the preferred L3 path, whereas PI-SAC determines the proportion of traffic assigned to each subflow.

4.2 Throughput and delivery reliability

Delivery reliability is quantified using the PDR, defined as the fraction of transmitted packets successfully received. Figures 5 and 6 show that all strategies operate in an uncongested region up to 15,000 packets/s, achieving the offered throughput with PDR = 1.

At 20,000 packets/s, PI-DDQN–GAT–SAC achieves approximately 194–195 Mb/s across line-of-sight (LOS), partial-blockage, and non-line-of-sight (NLOS) conditions, with PDR values of about 0.99. PI-Static 50/50 provides slightly higher raw throughput at this transition because both logical paths remain feasible throughout the evaluated campaign.

At the highest load of 25,000 packets/s, PI-DDQN–GAT–SAC achieves 222.09, 219.52, and 210.63 Mb/s under LOS, partial blockage, and NLOS, respectively. These results correspond to throughput gains of 10.99%, 11.81%, and 9.88% over strongest-link selection, the principal conventional adaptive baseline, and gains of 21.27%, 19.94%, and 14.60% over THz-only operation. Delivery reliability also remains higher than for the single-link and strongest-link alternatives. PI-Static 50/50 retains the highest raw parallel capacity where both links remain persistently feasible, highlighting the trade-off between trained dual-path utilization and adaptive cross-layer control.

The $1 - \text{PDR}$ is distinct from FlowMonitor-classified packet loss.

4.3 Delay and jitter

Figure 7 presents delay and jitter together because both metrics characterize the timing performance of packet delivery.

At 20,000 packets/s, PI-DDQN–GAT–SAC maintains low mean delay across all propagation conditions, substantially below OWC-only and strongest-link selection. At

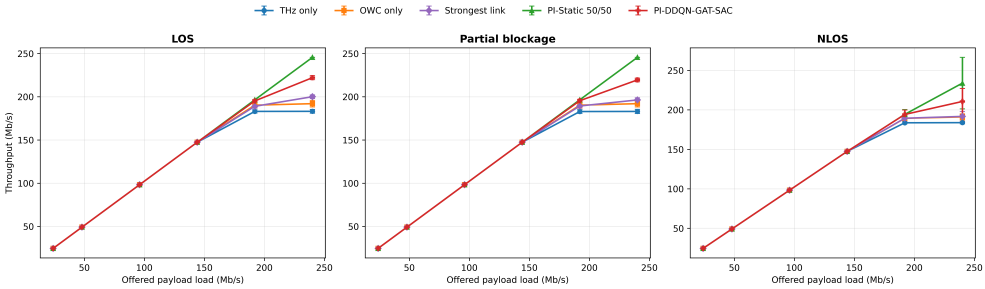


Fig. 5 Throughput versus offered payload load. Curves show five-seed means with 95% Student-t confidence intervals.

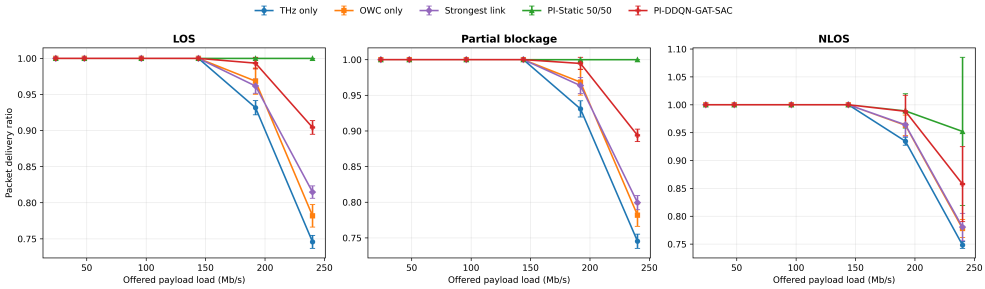


Fig. 6 Packet-delivery ratio versus offered payload load, reported as five-seed mean ± 95% confidence interval.

25,000 packets/s, the mean delay increases with worsening propagation, reaching 13.11, 22.02, and 42.42 ms under line-of-sight (LOS), partial blockage, and non-line-of-sight (NLOS), respectively. Despite the higher offered load, jitter remains below 0.07 ms for PI-DDQN–GAT–SAC in all three scenarios, indicating stable packet timing.

THz-only can exhibit lower jitter for the subset of packets it successfully delivers; therefore, jitter should be interpreted together with packet delivery ratio (PDR) rather than as an isolated performance measure.

4.4 Relative gain and held-out generalization

Relative gain measures the percentage change in throughput of PI-DDQN–GAT–SAC compared with a reference strategy. Figure 8 shows that the gains are negligible at low load, where all strategies are source-limited, but become positive over THz-only, OWC-only, and strongest-link selection as traffic increases. The negative gain relative to PI-Static 50/50 at the highest load reflects the raw-capacity advantage of persistent equal splitting when both logical paths remain feasible.

Seed 37 is retained as the strict held-out test seed to assess out-of-sample generalization. At 25,000 packets/s, PI-DDQN–GAT–SAC clearly outperforms strongest-link in the LOS and partial-blockage cases. The held-out NLOS case is more challenging,

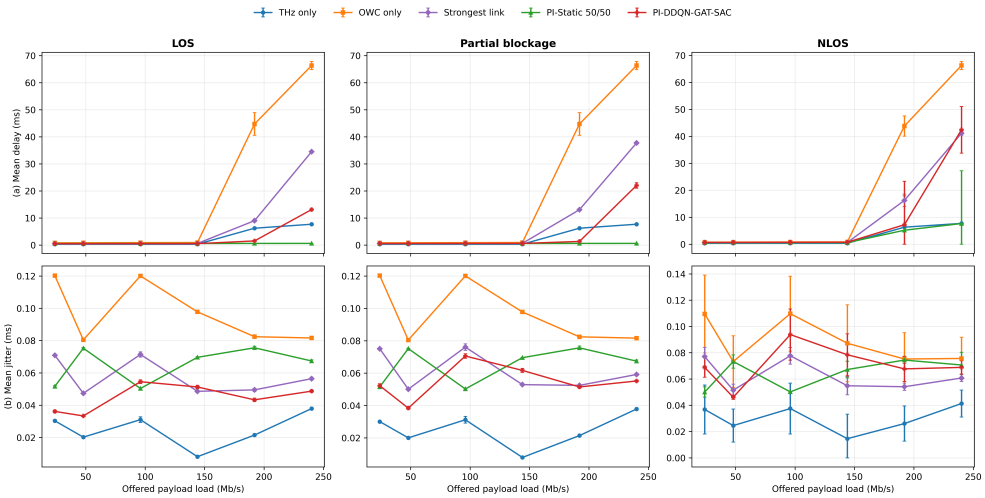


Fig. 7 Temporal packet performance: (a) mean end-to-end delay and (b) mean jitter versus offered payload load for LOS, partial blockage, and NLOS. Error bars show five-seed 95% confidence intervals.

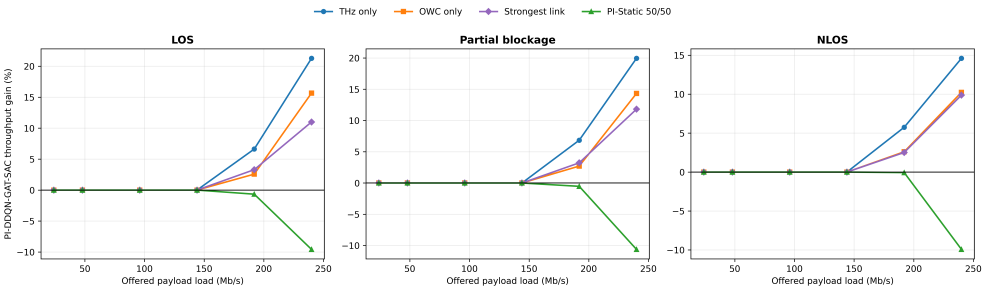


Fig. 8 Relative throughput gain of PI-DDQN–GAT–SAC over the comparison strategies. Positive values indicate higher throughput for the intelligent controller; negative values indicate higher throughput for the comparison strategy.

where PI-DDQN–GAT–SAC and PI-Static 50/50 achieve nearly identical throughput and OWC-only performs slightly better. Retaining this case provides a more realistic distinction between overall multi-seed robustness and strict held-out generalization.

5 Discussion

The results expose two complementary operating principles. PI-Static 50/50 is a strong fixed parallel-capacity reference when both paths remain feasible, whereas PI-DDQN–GAT–SAC provides a cross-layer response: PI-GAT moves the L3 decision from near-balanced THz versus best-OWC preference in LOS toward OWC in NLOS, then resolves the winning OWC tier among 30 APs; PI-SAC independently shifts continuous traffic allocation in the same direction.

The intelligent controller is therefore not expected to dominate the fixed equal split in every raw-capacity condition. Its main contribution is physics-constrained adaptation across access, Layer 3 (L3), and transport while outperforming single-link and strongest-link strategies in stressed regimes. This trade-off motivates a PI-Hybrid extension in which PI-GAT continues to provide a heterogeneous L3 path ranking, while a higher-level controller dynamically selects between adaptive subflow allocation of PI-SAC, 50/50 parallel transmission of PI-Static, and single-path fallback according to current link and traffic conditions.

6 Limitations and future work

The evaluation leverages fixed Sionna RT states for the THz–OWC environment, with the OWC branch additionally anchored to calibrated optical-channel parameters. However, both media remain feasible throughout the 300-state campaign, which favors persistent dual-path operation and does not capture complete link outages and subsequent recovery.

Future work will extend the physical campaign to include the temporary disappearance of the THz and OWC link, recovery, and more severe mobility and blockage conditions. The framework will also be evaluated through hardware in-the-loop (HIL) experiments and subsequently on real THz/sub-THz and OWC hardware interfaces using native MPTCP. The native MPTCP implementation will include actual subflow establishment, data-sequence mapping, congestion-control interaction, packet reinjection and recovery, and receiver-side reordering. A subsequent physics-informed hybrid controller (PI-Hybrid) will retain PI-GAT for heterogeneous Layer 3 (L3) path ranking while dynamically selecting between adaptive subflow allocation of PI-SAC, PI-Static 50/50 parallel transmission and single-path fallback according to the prevailing physical-link, L3 and traffic-load conditions, consistent with modular heterogeneous broadband IoT evolution[24].

7 Conclusion

This work develops an integrated cross-layer framework for extreme-band THz–OWC networks by linking physical-state information and localization uncertainty with physics-informed access decisions, heterogeneous L3 path ranking, continuous transport-layer subflow allocation, and packet-level evaluation in ns-3/5G-LENA. The proposed PI-DDQN–GAT–SAC framework combines physics-masked Double Deep Q-Network (DDQN) access and stability decisions, Graph Attention Network (GAT)-based THz–OWC L3 path ranking, and Soft Actor-Critic (SAC)-based dual-subflow allocation. The results show that as THz propagation conditions deteriorate from LOS through partial blockage to non-NLOS, PI-GAT increasingly favors suitable OWC access paths, while PI-SAC shifts a greater proportion of transport traffic toward the OWC subflow. This shows a coordinated but distinct adaptation across the network and transport layers.

Across 450 packet-level simulations, all evaluated strategies maintained full packet delivery in the uncongested operating region, while clear performance differences emerged at higher offered loads. At the highest load evaluated, PI-DDQN–GAT–SAC

improved throughput by approximately 10–12% over selection of the strongest-link, the principal conventional adaptive baseline, while maintaining stronger delivery reliability than single-link and strongest-link strategies. The complementary PI-Static 50/50 policy achieved the highest raw parallel capacity when both links remained persistently feasible, revealing a trade-off between fixed dual-path capacity utilization and propagation-aware L3/transport intelligence.

In general, the framework addresses the identified research gap by providing a reproducible connection between heterogeneous awareness of the physical-state of THz-OWC, L3 decision making assisted by localization, adaptive dual-subflow transport control and packet-level performance evaluation. The findings also motivate a future PI-Hybrid controller, in which PI-GAT continues to provide a heterogeneous L3 path ranking, while a higher-level controller dynamically selects among adaptive allocation of PI-SAC, 50/50 parallel transmission of PI-Static, and single-path fallback according to the prevailing physical-link and traffic-load conditions. Future work will further evaluate the framework through HIL experiments and real THz/sub-THz and OWC hardware interfaces using native MPTCP, including actual subflow establishment, congestion-control interaction, packet reinjection and recovery, and receiver-side reordering.

Data availability

The final packet-level KPI data, five-seed confidence-interval summaries, controller summaries, and configuration records have been archived by the authors and are available from the corresponding author upon reasonable request.

Code availability

The implementation code used in this study is available from the corresponding author and will be provided on a reasonable request. The package includes the Sionna RT THz configuration, calibrated Lambertian OWC geometry/channel and campaign scripts, ns-3/5G-LENA packet target, KPI accounting, and the scripts/data required to reproduce the heterogeneous analysis of L3, confidence intervals, and final figures.

Funding

This work has been funded by the following research projects: Brasil 6G Project with support from RNP/MCTI (Grant 01245.010604/2020-14), xGMobile Project code XGM-AFCCT-2026-7-10-2 with resources from EMBRAPII/MCTI (Grant 052/2023 PPI IoT/Manufatura 4.0) and FAPEMIG (Grant PPE-00124-23), SEMEAR Project supported by FAPESP (Grant No. 22/09319-9) and Scholarship Process No. 2023/15030-4), Ciência por Elas with resources from FAPEMIG (Grant APQ-04523-23), Fomento à Internacionalização das ICTMGs with resources from FAPEMIG (Grant APQ-05305-23), Programa de Apoio a Instalações Multiusuários with resources from FAPEMIG (Grant APQ-01558-24), and Redes Estruturantes, de Pesquisa Científica ou de Desenvolvimento Tecnológico with resources from FAPEMIG (Grant RED-00194-23). This work has also been supported by CNPq-Brasil.

Acknowledgements

Authors acknowledged the resources provided by the underlisted funding bodies and xGMobile at the National Institute of Telecommunication, Santa Rita do Sapucai, MG in administrative and technical support of this research.

Author contributions

E.U.O. conceived the framework, developed the physics-informed multi-agent formulation, implemented the trace-fusion, L3 PI-GAT, transport-controller workflow, developed the packet-level evaluation, analysed the results, and drafted the manuscript. L.L.M. contributed research supervision, technical direction, and manuscript review. Both authors reviewed and approved the manuscript.

Competing interests

The authors declare no competing interests.

References

[1] Liu, M., Kazemi, H., Safari, M., Tavakkolnia, I. & Haas, H. A comprehensive comparison between Terahertz and optical wireless communications. *npj Wireless Technology* **1**, 3 (2025). https://doi.org/10.1038/s44459-025-00002-1

[2] Jornet, J. M. & Akyildiz, I. F. Channel modeling and capacity analysis for electromagnetic wireless nanonetworks in the Terahertz band. *IEEE Trans. Wireless Commun.* **10**, 3211–3221 (2011).

[3] Akyildiz, I. F., Jornet, J. M. & Han, C. Terahertz band: next frontier for wireless communications. *Phys. Commun.* **12**, 16–32 (2014).

[4] Han, C., Bicen, A. O. & Akyildiz, I. F. Multi-ray channel modeling and wideband characterization for wireless communications in the Terahertz band. *IEEE Trans. Wireless Commun.* **14**, 2402–2412 (2015).

[5] Khalighi, M.-A. & Uysal, M. Survey on free space optical communication: a communication theory perspective. *IEEE Commun. Surv. Tutor.* **16**, 2231–2258 (2014).

[6] Haas, H., Yin, L., Wang, Y. & Chen, C. What is LiFi? *J. Lightwave Technol.* **34**, 1533–1544 (2016).

[7] Koonen, T. Indoor optical wireless systems: technology, trends, and applications. *J. Lightwave Technol.* **36**, 1459–1467 (2018).

[8] IEEE Standards Association. IEEE Std 802.11bb-2023: Amendment 6: Light Communications (2023).

[9] Ogbodo, E. U., Abu-Mahfouz, A. M. & Kurien, A. M. Enabling LPWANs for coexistence and diverse IoT applications in smart cities using lightweight heterogenous multihomed network model. *J. Sens. Actuator Netw.* **11**, 87 (2022). https://doi.org/10.3390/jsan11040087

[10] Ford, A., Raiciu, C., Handley, M., Bonaventure, O. & Paasch, C. TCP extensions for multipath operation with multiple addresses. RFC 8684, Internet Engineering Task Force (2020). https://doi.org/10.17487/RFC8684

[11] Raiciu, C. *et al.* How hard can it be? Designing and implementing a deployable multipath TCP. In *Proc. 9th USENIX Symposium on Networked Systems Design and Implementation (NSDI)* 399–412 (2012).

[12] van Hasselt, H., Guez, A. & Silver, D. Deep reinforcement learning with Double Q-learning. In *Proc. AAAI Conf. Artif. Intell.* **30**, 2094–2100 (2016).

[13] Velickovic, P. *et al.* Graph attention networks. In *International Conference on Learning Representations (ICLR)* (2018).

[14] Haarnoja, T., Zhou, A., Abbeel, P. & Levine, S. Soft Actor-Critic: off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *Proc. 35th International Conference on Machine Learning* 1861–1870 (2018).

[15] Ogbodo, E. U., Abu-Mahfouz, A. M. & Kurien, A. A. 5G RedCap enhancement towards improved cellular LPWAN/5G-IoT for smart cities and industrial IoT using genetic algorithm-based neural network. *Int. J. Sens. Wirel. Commun. Control* **15**, 14–29 (2025). https://doi.org/10.2174/0122103279312219240520053431

[16] Hoydis, J. *et al.* Sionna: an open-source library for next-generation physical layer research. *arXiv preprint* arXiv:2203.11854 (2022).

[17] Nimier-David, M., Hoydis, J. & Keller, A. Sionna RT: differentiable ray tracing for radio propagation modeling. *arXiv preprint* arXiv:2303.11103 (2023).

[18] Kahn, J. M. & Barry, J. R. Wireless infrared communications. *Proc. IEEE* **85**, 265–298 (1997).

[19] Komine, T. & Nakagawa, M. Fundamental analysis for visible-light communication system using LED lights. *IEEE Trans. Consum. Electron.* **50**, 100–107 (2004).

[20] Patwari, N. *et al.* Locating the nodes: cooperative localization in wireless sensor networks. *IEEE Signal Process. Mag.* **22**, 54–69 (2005).

[21] Wymeersch, H., Lien, J. & Win, M. Z. Cooperative localization in wireless networks. *Proc. IEEE* **97**, 427–450 (2009).

[22] Ogbodo, E. U., Renno, V. M. & Mendes, L. L. AI-driven NTN and 5G edge integration for cost-effective, secure, and sustainable smart agriculture in rural areas: a comparative study with traditional approaches. *IEEE Access* (2026). https://doi.org/10.1109/ACCESS.2026.3691027

[23] Koutlia, K., Bojovic, B., Ali, Z. & Lagen, S. Calibration of the 5G-LENA system level simulator in 3GPP reference scenarios. *Simul. Model. Pract. Theory* **119**, 102580 (2022). https://doi.org/10.1016/j.simpat.2022.102580

[24] Ogbodo, E. U., Renno, V. M. & Mendes, L. L. Broadband IoT for digital agriculture in rural and remote areas: field-level connectivity, coverage, throughput, and emerging technologies. *Electronics* **15**, 2908 (2026). https://doi.org/10.3390/electronics15132908