IET Quantum Communication

WILEY

The Institution of Engineering and Technology

ORIGINAL RESEARCH OPEN ACCESS

Cybersecurity Driven Quantum Digital Twin for Proactive Threat Reversal in Open RAN

Yassir Al-Karawi[1] | Raad S. Alhumaima[2] | Hamed Al-Raweshidy[3]

[1]Department of Communications Engineering, University of Diyala, Baquba, Iraq | [2]Department of Cyber Security Techniques, Imam Ja'afar Al-Sadiq University, Diyala, Iraq | [3]School of Engineering and Design, Brunel University London, Uxbridge, UK

Correspondence: Hamed Al-Raweshidy (Hamed.Al-Raweshidy@brunel.ac.uk)

Received: 30 June 2025 | **Revised:** 8 November 2025 | **Accepted:** 30 November 2025

Keywords: entropy | learning (artificial intelligence) | quantum communication | quantum entanglement | quantum gates | quantum information | quantum noise | telecommunication security

ABSTRACT

A cybersecurity-driven quantum digital twin (CQDT) is introduced to protect networked control chains against adversarial completely positive trace preserving perturbations. The design maps O-RAN telemetry to amplitude-encoded registers, forms multipartite entanglement through GHZ and one-dimensional cluster states and injects bit flip, phase flip and amplitude damping channels tied to measurable indicators. System integrity is tracked using fidelity F_t, von Neumann entropy S_t and trace distance D_t. A lightweight REINFORCE policy acts on these observables to preserve entanglement and limit decoherence within tight cycle-time budgets. Qiskit simulations maintain average fidelity above 0.91, entropy near 0.35 and trace distance below 0.18 under composite noise. Attack classification exceeds 87% with a software-loop latency of about 19.1 ms covering state preparation, entanglement, CPTP injection, measurement and policy inference. Compared with classical intrusion detection and thresholding over quantum observables, the framework improves detection while reducing false alarms and exposes quantum-state degradation in real time. The result is a reproducible and scalable basis for learning-based quantum-aware protection in disaggregated radio access networks and related systems.

1 | Introduction

Next-generation radio access networks aim to support immersive communications, autonomous cyber-physical control and large-scale IoE under stringent latency and reliability targets [1]. Meeting these targets requires re-architecting the radio access network (RAN) around disaggregated, software-defined components and time-sensitive control loops. Open RAN (O-RAN) separates the radio unit (RU), distributed unit (DU) and centralised unit (CU) and exposes standardised interfaces for programmability and AI-assisted automation, improving flexibility while enlarging the attack surface across synchronisation, inference pipelines, and closed-loop control.

Open RAN disaggregates monolithic base stations into virtualised modules connected via standardised interfaces between the radio unit (RU), distributed unit (DU) and centralised unit (CU) [2–6]. Although disaggregation facilitates vendor diversity and AI integration, it also expands the attack surface across timing/synchronisation, ML inference pipelines and real-time control loops [7–11]. Existing Open RAN security references [12] largely address classical protocol threats and do not model quantum-aware adversaries capable of manipulating quantum processes or exacerbating decoherence.

Quantum digital twins (QDTs) extend classical digital twins by simulating quantum dynamics, including superposition,

entanglement and nonunitary noise [13–15]. Prior QDT efforts typically emphasised calibration and passive emulation rather than *adaptive*, closed-loop security responses tailored to wireless time-critical operation. In parallel, quantum cryptographic primitives and quantum-enhanced inference do not provide runtime observability of quantum-state degradation suitable for proactive mitigation within Open RAN control paths [16–18].

1.1 | Motivation and Practical Relevance

Open RAN disaggregation introduces new control and synchronisation paths (E2/A1/O1) where timing drift, policy poisoning and beam-level interference can degrade SLAs under URLLC constraints. Classical twins lack sensitivity to decoherence and entanglement loss, whereas cryptographic safeguards focus on keying/authentication rather than runtime quantum observability for mitigation. A quantum digital twin that probes RU–DU–CU with multipartite entanglement and CPTP 'what-if' channels enables predeployment risk screening and in-loop policy rehearsal, improving detection, false-alarm control and mitigation latency.

1.2 | From O-RAN Signals to Quantum States

RU/DU/CU telemetry can be mapped to amplitude-encoded registers with reserved index ranges per subregister (RU, DU and CU). GHZ links capture global coordination (e.g., synchronisation/orchestration), whereas 1D-cluster edges capture localised dependencies (e.g., RU↔DU scheduling/beam control). This mapping yields quantum observables —fidelity, von Neumann entropy and trace distance—that are responsive to operational disturbances and suitable for security control.

1.3 | Physically Tied CPTP Parameters

Channel parameters are anchored to measurable indicators: amplitude damping $\gamma = 1 - \exp(-\Delta t/T_1)$ (relaxation), phase-flip probability linked to phase-jitter variance (synchronisation drift) and bit-flip probability linked to SNR/EVM surrogates. Such ties constrain scenario sweeps to operationally plausible regimes and support reproducibility.

1.4 | Latency Scope

All latency values reported for CQDT refer to the *software loop* (state preparation → entanglement → CPTP injection → measurement → RL inference) and exclude RF/PHY processing, transport, RIC/xApp orchestration and QPU execution/readout. Thus, figures constitute simulator feasibility rather than end-to-end guarantees.

1.5 | Research Gap

An integrated framework that simultaneously (i) encodes RU–DU–CU subsystems into multipartite quantum registers aligned with O-RAN semantics, (ii) injects adversarial CPTP channels with explicit Kraus semantics tied to operational indicators and (iii) learns low-latency defences directly from (F, S, D) observables for proactive mitigation has not been established.

1.6 | Contributions

- *Quantum encoding and entanglement:* Amplitude encoding of O-RAN telemetry into composite RU–DU–CU registers with GHZ/cluster structures capturing global/local dependencies.
- *Adversarial CPTP modelling:* Bit/phase flips and amplitude damping with ranges tied to synchronisation jitter, SNR/EVM and T_1 to reflect realistic operating conditions.
- *Observable-driven RL defence:* A lightweight REINFORCE policy acting on (F_t, S_t, D_t) to maintain entanglement and limit decoherence within cycle-time budgets.
- *Evaluation protocol:* Cross-validation, ablations (REINFORCE vs. A2C/PPO/DQN) and baselines (classical IDS and quantum-observable IDS) reporting detection metrics (DR, FPR, AUC and MCC) alongside (F, S, D).
- *Scalability and limitations:* Complexity breakdown and a candid discussion of hardware feasibility and deployment constraints.

1.7 | Organisation

Section 2 overviews related strands. Section 3 formalises the adversary. Section 4 details the CQDT architecture and observables. Section 5 presents simulation and algorithms. Section 6 reports results. Section 7 discusses deployment and limitations. Section 8 concludes (Figure 1).

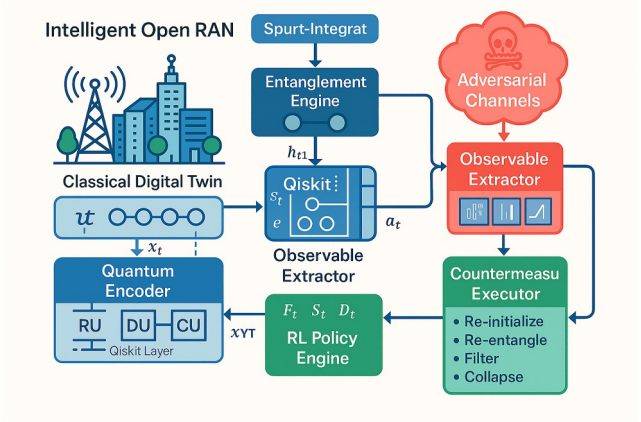

FIGURE 1 | CQDT-secure Open RAN pipeline: O-RAN feature ingestion, amplitude encoding, multipartite entanglement (GHZ/cluster), adversarial CPTP injection, RL feedback and multiobjective defence. Symbols match Table 1.

TABLE 1 | Unified symbol definitions used in the CQDT framework.

Symbol	Definition
x_t	Classical input vector at time t
x_{norm}	ℓ_2-normalised version of x_t
$\lvert\psi_t\rangle$	Quantum state encoded from x_t
$\lvert\psi_{\text{GHZ}}\rangle$	3-qubit GHZ entangled state
$\lvert\psi_{\text{cluster}}\rangle$	Cluster state over subsystems
ρ, ρ', ρ''	Original, noisy and postdefence density matrices
$\tilde{\rho}'$	Adversarially perturbed density matrix
F_t	Fidelity between ρ and ρ'
S_t	von Neumann entropy of ρ'
D_t	Trace distance between ρ and ρ'
o_t	Observable vector: $[F_t, S_t, D_t]$
a_t	Defence action taken at time t
$\pi_\theta(a_t \lvert o_t)$	Policy probability for action a_t given state o_t
$f_\theta(o_t, a_t)$	Score function in softmax-based policy
w_a, b_a	Policy weights and bias for action a
θ	Parameter vector of the RL policy
$J(\theta)$	Expected cumulative reward function
R_t	Reward value at time t
α, β, δ	Reward coefficients for fidelity, entropy and trace distance
η	Learning rate for policy gradient updates
$\mathcal{E}, \widetilde{\mathcal{E}}$	Normal and adversarial CPTP channels
$\{E_k\}, \{\tilde{E}_k\}$	Kraus operators for CPTP channels
p, γ	Flip/error probability and decay rate
H	Hadamard gate
$\text{CNOT}_{i\to j}$	CNOT gate: Control i and target j
$CZ_{a,b}$	Controlled-Z gate between qubits a and b
I	Identity operator
GHZ_{3q}	Generator for GHZ state over 3 qubits
$\mathcal{S}_{\text{secure}}$	Secure operational state region
E, T	Total simulation episodes and steps per episode
T_{lat}	Overall latency of full defence cycle
t_{mod}	Latency of a specific module (e.g., RL or measurement)
Tr	Trace operator for matrices
log	Natural logarithm function
$\lvert\cdot\rvert$	Trace norm (Schatten 1-norm) of a matrix

2 | Related Work

Research at the intersection of quantum cybersecurity, AI-enabled defence and digital-twin modelling spans several largely disjoint streams, with limited attention to quantum-secure Open RAN. Classical and quantum-cryptographic lines—BB84, E91 and twin-field QKD—deliver information-theoretic confidentiality for key establishment [19, 20], yet remain key-centric and do not expose runtime signals suitable for closed-loop mitigation in RAN control. Carrier-grade QKD network pilots report SDN/NFV-aligned orchestration and interdomain key management across metro/backbone footprints [21]. In parallel, postquantum cryptography (PQC) onboarding to control/user planes and transport segments proceeds with hybrid suites and measured performance–security trade-offs in 5G/6G contexts [22–24]. Although these efforts harden confidentiality and authenticity, they do not model quantum-information observables nor provide closed-loop detection/mitigation against CPTP-style perturbations.

Security-oriented digital twins detect anomalies from classical telemetry abstractions [25, 26]. However, they lack visibility into decoherence, nonunitary evolution or entanglement collapse, and their defences are typically rule-based. Quantum-oriented digital twins have progressed in calibration, drift tracking and device/network feedback using graph-state preparation and entanglement monitors [27–29]. Yet, explicitly adversarial CPTP modelling and policy learning driven by fidelity, entropy and trace distance remain limited.

Quantum-enhanced learning explores hybrid models for channel and attack inference [18, 30]. Recent studies extend to anomaly detection and channel modelling under hybrid learners [31], yet typically without multichannel adversaries or integration with RIC/xApp closed loops. Quantum RL and parameterised control demonstrate fast convergence in small-scale simulators [32, 33], but often ignore subsystem encodings and wireless constraints under injected CPTP noise. On the classical side, RL-for-RAN contributes resource/jamming control and closed-loop reconfiguration [34], whereas Open RAN security references consolidate threat surfaces and coordination procedures [12]. Centralised/federated IDS advances (AE/LSTM/GNN) strengthen telemetry protection [35, 36] but cannot sense decoherence or reason over trace-distance dynamics. Similarly, federated quantum learning (FQL) enables privacy-preserving training [37] without ingesting (F, S, D) observables or actuating entanglement-aware defences.

Relative positioning is as follows. QKD/QKDN and PQC secure keys and primitives. Classical and quantum digital twins provide modelling and control hooks but rarely couple adversarial CPTP channels with policy learning. IDS and RL-for-RAN operate on classical observability. Hybrid QML/QRL seldom integrates with O-RAN control paths under multichannel noise.

By contrast, the present framework encodes the RU–DU–CU subsystems into multipartite registers, injects adversarial CPTP channels with explicit Kraus semantics and learns low-latency policies directly from fidelity, entropy and trace distance—reporting security metrics (DR, FPR, AUC and MCC) alongside quantum metrics. A compact cross-strand comparison is provided in Table 2, clarifying scope differences and showing where CQDT fills persistent gaps in runtime, quantum-aware mitigation. As summarised in Table 2, prior strands lack at least one of: quantum observability, closed-loop defence or O-RAN alignment, the present design integrates all three.

3 | Threat Model and Assumptions

The CQDT framework operates within a quantum-enhanced Open RAN composed of disaggregated RU, DU and CU nodes. The adversary is assumed to exploit both classical vulnerabilities and quantum-layer phenomena [26, 39]. The threat model spans malicious activities that degrade fidelity, disrupt multipartite entanglement and impact quantum coherence.

The adversary possesses quantum-aware capabilities: manipulation of transmitted quantum states, injection of coherent noise, degradation of entanglement across subsystems and interference with classical/quantum control signals [39]. The attacker may move laterally along the RU–DU–CU chain and act at arbitrary times/locations. Internal policy parameters (e.g., RL weights and optimiser settings) are assumed to be unknown to the adversary.

The attack surface is categorised into classical and quantum components. On the classical side, the adversary can falsify observables injected into the RL pipeline and disrupt inter-unit synchronisation or leverage firmware-level backdoors to bias policy behaviour. Quantum-layer attacks have direct mathematical impacts. For instance, intercept–resend or CPTP poisoning corrupts the density operator ρ via altered Kraus operators $\{\tilde{E}_k\}$:

$$\tilde{\rho}' = \widetilde{\mathcal{E}}(\rho) = \sum_k \tilde{E}_k \rho \tilde{E}_k^{\dagger}, \quad \text{s.t.} \sum_k \tilde{E}_k^{\dagger}\tilde{E}_k = I. \tag{1}$$

This corrupted state $\tilde{\rho}'$ deviates from the expected noisy state ρ', altering key observables. The fidelity

$$F_t = \left(\mathrm{Tr}\sqrt{\sqrt{\rho}\,\tilde{\rho}'\sqrt{\rho}}\right)^2, \tag{2}$$

decreases, indicating coherence loss; the von Neumann entropy

$$S_t = -\mathrm{Tr}(\tilde{\rho}'\log_2\tilde{\rho}'), \tag{3}$$

increases (mixedness) and the trace distance

$$D_t = \frac{1}{2}\left\|\rho - \tilde{\rho}'\right\|_1 = \frac{1}{2}\mathrm{Tr}|\rho - \tilde{\rho}'|, \tag{4}$$

increases, reflecting greater distinguishability from the ideal state.

These shifts propagate into the RL loop. The observation vector $o_t = [F_t, S_t, D_t]$ becomes biased, which can mislead the stochastic policy $\pi_\theta(a_t|o_t)$. With reward

$$R_t = -(\alpha(1 - F_t) + \beta S_t + \delta D_t), \tag{5}$$

policy-gradient updates

$$\theta \leftarrow \theta + \eta\,\nabla_\theta \log \pi_\theta(a_t|o_t)\,R_t \tag{6}$$

can be eroded by false gradients, yielding suboptimal actions.

Attacks need not be localised to a single component or time step. Distortion in the amplitude encoder perturbs the prepared state $|\psi_t\rangle$ systemically, and entanglement disruption affects GHZ/cluster structures used to construct ρ. Interference with RL feedback can further inject temporal noise that compounds decision errors across steps (Table 3).

The mapping of adversarial access points across the CQDT-enabled O-RAN is summarised in Figure 2.

TABLE 2 | Cross-strand comparison in quantum cybersecurity for networked systems.

Reference	Quantum modelling	AI-driven defence	System-level simulation
[19, 20]	✓	✗	✗
[25, 26]	✗	✓	*Partial*
[18]	✓	✓	✗
[30]	✓	✓	✗
[32]	✓	✓	✗
[33]	✓	✓	*Partial*
[12]	✗	✗	✓
[38]	✗	✓	*Partial*
[37]	✓	✓	*Partial*
[21]	✓	✗	*Carrier deployments*
[22–24]	✗	✗	*PQC in 5G/6G*
[27–29]	✓	*Partial*	*QDT focus*
[34–36]	✗	✓	*RAN/6G*
This work (CQDT)	✓	✓	✓

Note: ✓ indicates that the corresponding capability is supported by the referenced work; ✗ indicates that the capability is not addressed. **Italic text** (e.g., **Partial** and the italic annotations such as **PQC in 5G/6G, ODT focus, Carrier deployments**) denotes partial/limited-scope coverage or context-specific focus.

4 | System Architecture and Modelling

The following section describes how the cybersecurity-driven quantum digital twin (CQDT) structure for Open RAN security is organised. Figure 3 shows how CQDT integrates observati ons, quantum encoding, entanglement modelling, simulating

TABLE 3 | Primary threats and impacts in CQDT-based O-RAN.

Threat	Layer	Primary impact
FDI (data injection)	Classical	RL misguidance
DoS	Classical	Delay and dropouts
Intercept–resend	Quantum	Fidelity loss
Entanglement disruption	Quantum	Coherence loss
CPTP poisoning	Quantum	Observable drift
Firmware Trojan	Control stack	Multimetric impact
Cloning/probing	Readout	Privacy/information leak

adversarial channels, monitoring quantum observables and reinforcement learning (RL) mechanisms to manage the Open RAN network.

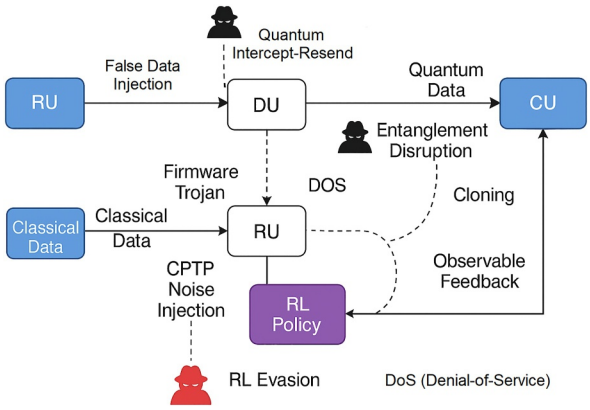

FIGURE 2 | Adversarial access points across the CQDT-based O-RAN. Vectors may corrupt amplitude encoding/entanglement, observable feedback or RL inference, quantum and classical interfaces require layered policy adaptation.

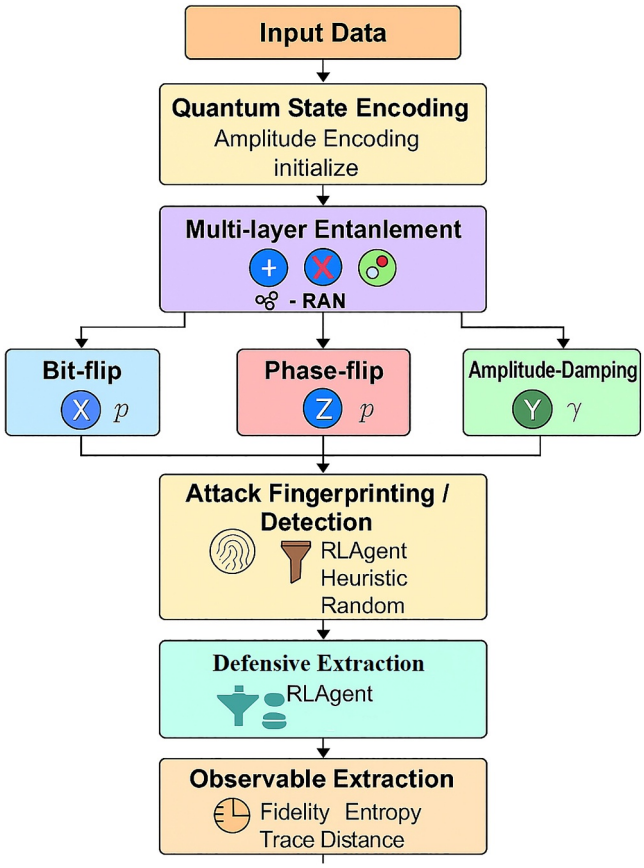

FIGURE 3 | Simulation workflow of the CQDT pipeline with quantum encoding, multilayer entanglement, adversarial CPTP noise injection, attack fingerprinting, RL-based defence and quantum observables.

4.1 | Quantum Encoding and Entanglement Modelling

4.1.1 | Quantum State Representations Through Amplitude Encoding

Classical information is mapped by the CQDT framework into quantum states, wherein the standard method of amplitude encoding uses fewer than N qubits [40]. Given a feature dimension d, the required register size is $n = \lceil \log_2 d \rceil$ qubits, so that $2^n \geq d$ and unused amplitudes (if any) are zero-padded for reproducibility. The postnormalisation value of the final quantum state is at a unit-norm level:

$$x_{\text{norm}} = \frac{x_t}{\|x_t\|}, \quad \text{with} \quad \|x_t\| = \sqrt{\sum_{j=1}^{N} x_j^2}. \tag{7}$$

where $\|\cdot\|$ denotes the ℓ_2 norm. The normalised vector is then defined as a corresponding amplitude-encoded quantum state as follows:

$$|\psi_t\rangle = \sum_{j=0}^{N-1} (x_{\text{norm}})_j |j\rangle. \tag{8}$$

This technique is important for CQDT as it allows RU, DU and CU nodes to store data simultaneously with multiple values bearing relevance (e.g., radio link quality, traffic congestion or CPU load). Amplitude locality is preserved by reserving contiguous index ranges per subsystem so that downstream observables can be attributed to RU/DU/CU contributions.

4.1.2 | Multipartite Entanglement Over Subsystems

CQDT can leverage entanglement features to combine the RU–DU–CU subsystems into a quantum state that possesses nonlocal connections in the classical sense. Two representative entangled states are highlighted.

4.1.2.1 | GHZ States. To encode global-parity entanglement, Greenberger–Horne–Zeilinger (GHZ) states are used. A three-qubit GHZ state can be expressed as follows:

$$|\psi_{\text{GHZ}}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle). \tag{9}$$

This state embeds all subsystems together and is sensitive to bit and phase changes, thereby aiding the detection of large-scale coherent disturbances [13].

4.1.2.2 | Cluster States. Cluster states derive from performing controlled-Z on Hadamard-ready qubits to create modular localisable entanglements:

$$|\psi_{\text{cluster}}\rangle = \left(\prod_{(a,b)\in E} CZ_{a,b} \right) H^{\otimes n} |0\rangle^{\otimes n}, \tag{10}$$

where H is the Hadamard gate, $CZ_{a,b}$ denotes a controlled-Z operation between qubit a and qubit b and E is the

entanglement edge set. Cluster states enable fine-grained subsystem correlations and are useful for modelling localised or hierarchical combinations of defence interactions [41]. In practice, GHZ favours detection of system-wide parity perturbations, whereas 1D cluster links isolate localised faults (e.g., RU↔DU) with lower-depth circuits.

4.1.3 | Quantum Circuit Realisation of Entanglement

Figure 4 illustrates circuit diagrams for creating GHZ and cluster states, showing how entanglement occurs between RU, DU and CU in the CQDT pipeline. As an example, a Hadamard gate is applied first, then CNOT and CZ gates, this yields the GHZ state under one configuration, or a modular cluster state under another. Circuit seeds and gate orders are fixed per scenario to ensure repeatable measurements and fair comparisons across noise settings.

4.1.3.1 | O-RAN–to–Circuit Mapping and Amplitude Encoding

4.1.3.1.1 | RU/DU/CU → Quantum Registers. Let $\mathcal{F}_{\mathrm{RU}} \in \mathbb{R}^{d_{\mathrm{RU}}}$, $\mathcal{F}_{\mathrm{DU}} \in \mathbb{R}^{d_{\mathrm{DU}}}$ and $\mathcal{F}_{\mathrm{CU}} \in \mathbb{R}^{d_{\mathrm{CU}}}$ denote feature vectors derived from O-RAN telemetry and control (e.g., RU: PRB utilisation, SNR, EVM and beam index; DU: scheduling load and HARQ stats and CU: control-plane latency and policy state). Disjoint quantum registers $\mathcal{Q}_{\mathrm{RU}}, \mathcal{Q}_{\mathrm{DU}}, \mathcal{Q}_{\mathrm{CU}}$ with $n_{\mathrm{RU}}, n_{\mathrm{DU}}, n_{\mathrm{CU}}$ qubits, respectively, are allocated such that $2^{n_\cdot} \geq d_\cdot$. The global register is $\mathcal{Q} = \mathcal{Q}_{\mathrm{RU}} \otimes \mathcal{Q}_{\mathrm{DU}} \otimes \mathcal{Q}_{\mathrm{CU}}$.

4.1.3.1.2 | Feature Preprocessing and Scaling.

$$x_t = \begin{bmatrix} \mathcal{F}_{\mathrm{RU}} \\ \mathcal{F}_{\mathrm{DU}} \\ \mathcal{F}_{\mathrm{CU}} \end{bmatrix}, \qquad d = d_{\mathrm{RU}} + d_{\mathrm{DU}} + d_{\mathrm{CU}}.$$

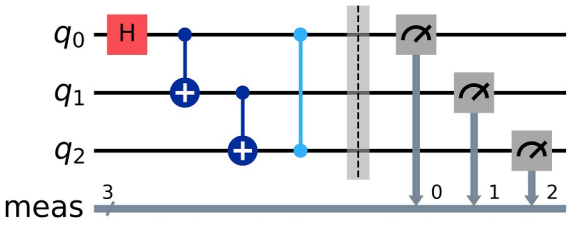

FIGURE 4 | Quantum circuit realisation of multipartite entanglement for the RU–DU–CU subsystems. Both the GHZ and cluster state preparation modules are shown in operation, used in different ways within the CQDT framework.

Each raw feature $f^{(i)}$ is standardised and then range-limited.

$$\tilde{f}^{(i)} = \frac{f^{(i)} - \mu_i}{\sigma_i},$$
$$\hat{f}^{(i)} = \mathrm{clip}\left(\tilde{f}^{(i)}, -c, c\right), \tag{11}$$
$$x_t^{(i)} = \frac{\hat{f}^{(i)} - m_i}{M_i - m_i}.$$

where μ_i, σ_i are running statistics, (m_i, M_i) are per-feature min–max bounds and $c > 0$ prevents outliers.

4.1.3.1.3 | Amplitude Encoding Over the Composite RU–DU–CU Register. Let $x_{\mathrm{norm}} = x_t / \|x_t\|_2$. An amplitude-encoded state is prepared across the composite register:

$$|\psi_t\rangle = \sum_{j=0}^{2^n-1} (x_{\mathrm{norm}})_j |j\rangle, \quad n = n_{\mathrm{RU}} + n_{\mathrm{DU}} + n_{\mathrm{CU}}. \tag{12}$$

Index ranges are reserved per subregister to preserve locality: $|j\rangle = |j_{\mathrm{RU}}\rangle \otimes |j_{\mathrm{DU}}\rangle \otimes |j_{\mathrm{CU}}\rangle$. This mapping keeps RU/DU/CU contributions addressable for independent or joint probing by observables. Per-subregister normalisation is avoided to preserve relative scaling across subsystems, instead a global ℓ_2 normalisation is applied to x_t.

4.1.3.1.4 | Entanglement Structure Aligned to O-RAN. Global-coordination phenomena (e.g., sync loops and cross-layer orchestration) are modelled with GHZ-like links across $\mathcal{Q}_{\mathrm{RU}}, \mathcal{Q}_{\mathrm{DU}}, \mathcal{Q}_{\mathrm{CU}}$, localised dependences (e.g., RU↔DU beam/scheduling) use cluster-state edges confined to the corresponding subregisters. This yields sensitivity to both system-wide and subsystem-specific perturbations.

4.1.3.1.5 | Feature→Amplitude Assignment (Reproducible Template). Table 4 instantiates an example mapping (extendable as needed), the same template is used for training/evaluation to preserve reproducibility.

4.1.3.1.6 | Scope and Novelty. To avoid breadth–depth dilution, evaluation focuses on three concrete O-RAN security use–cases: (i) timing/synchronisation drift, (ii) orchestration/policy poisoning at DU/CU and (iii) beam manipulation/jamming at RU. These are encoded via Equation (12) and exercised under CPTP noise (Section 4), with defence policies trained on CQDT observables. Scenario configuration files pin down feature slots and circuit seeds to facilitate independent replication.

TABLE 4 | Feature→amplitude assignment across RU/DU/CU subregisters.

Subsystem	Feature	Preprocess (Equation 11)	Amplitude slot
RU	PRB utilisation (%)	std. + clip $\pm c$ + min–max	$x_{\mathrm{norm}}[0 : k_1)$
RU	SNR/EVM/beam ID	std. + clip $\pm c$ + min–max	$x_{\mathrm{norm}}[k_1 : k_2)$
DU	Scheduler load/HARQ	std. + clip $\pm c$ + min–max	$x_{\mathrm{norm}}[k_2 : k_3)$
CU	Ctrl/user latency/policy	std. + clip $\pm c$ + min–max	$x_{\mathrm{norm}}[k_3 : k_4)$

4.2 | Adversarial Noise Injection and Disturbances to the Circuit

4.2.1 | Adversarial CPTP Channel Modelling

CQDT defines how noise is applied in quantum-layer attacks using completely positive trace preserving (CPTP) channels, which emulate the errors of open quantum systems. The CPTP map that acts on the density matrix ρ is as follows:

$$\rho' = \mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger, \quad \text{with} \quad \sum_k E_k^\dagger E_k = \mathbb{I}, \tag{13}$$

where $\{E_k\}$ are Kraus operators. Three major noise models are considered:

- **Bit Flip (BF):** $E_0 = \sqrt{1-p}\,\mathbb{I}, \quad E_1 = \sqrt{p}X$
- **Phase Flip (PF):** $E_0 = \sqrt{1-p}\,\mathbb{I}, \quad E_1 = \sqrt{p}Z$
- **Amplitude Damping (AD):**

$$E_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}, \quad E_1 = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}$$

4.2.1.1 | Physical Rationale and O-RAN Mapping of CPTP Parameters. Channel parameters are selected by tying quantum noise models to measurable O-RAN indicators:

4.2.1.1.1 | Amplitude Damping (γ) ↔ Relaxation Time. For an effective dwell per step Δt, the damping factor follows:

$$\gamma = 1 - e^{-\Delta t/T_1}, \tag{14}$$

so that shorter T_1 yields larger γ. Sensitivity is swept over $\Delta t/T_1 \in [0.05,\ 0.7]$.

4.2.1.1.2 | Phase Flip (p) ↔ Phase Jitter/Synchronisation Drift. Let $\Delta\phi \sim \mathcal{N}\left(0, \sigma_\phi^2\right)$. Using $\mathbb{E}[\cos\Delta\phi] = e^{-\sigma_\phi^2/2}$, an operational surrogate is as follows:

$$p_{\mathrm{PF}} \approx \frac{1 - e^{-\sigma_\phi^2/2}}{2}. \tag{15}$$

4.2.1.1.3 | Bit Flip (p) ↔ SNR/EVM. For a conservative mapping, the symbol error upper bound for BPSK over AWGN gives

$$p_{\mathrm{BF}} \lesssim Q\left(\sqrt{2\,\mathrm{SNR}_{\mathrm{lin}}}\right), \quad \mathrm{SNR}_{\mathrm{lin}} = 10^{\mathrm{SNR}_{\mathrm{dB}}/10}, \tag{16}$$

and empirical EVM can be folded via $\mathrm{SNR}_{\mathrm{lin}} \approx 1/\mathrm{EVM}^2$ (small-EVM regime).

4.2.1.1.4 | Depolarising p. Used to capture unmodelled aggregate disturbance, p is bounded by the worst of $\{p_{\mathrm{BF}},\ p_{\mathrm{PF}},\ \gamma\}$ in scenario composition, then swept in a conservative band.

4.2.1.2 | Scenario-to-CPTP Mapping (Rule Form). *Timing drift (sync)* ⇒ PF, $p_{\mathrm{PF}} = f_\phi\left(\sigma_\phi^2\right)$. Sweep $p_{\mathrm{PF}} \in [0.05, 0.50]$ (Equation 15).

Beam manipulation/jamming ⇒ BF + Depolarising, $p_{\mathrm{BF}} = f_{\mathrm{SNR}}(\mathrm{SNR})$, with $p_{\mathrm{dep}} \in [0, 0.20]$ (Equation 16).

Thermal/hardware stress ⇒ AD; $\gamma = f_{T_1}(\Delta t/T_1)$, sweep $\Delta t/T_1 \in [0.05, 0.70]$ (Equation 14).

Composite outage ⇒ PF ∘ BF + Depolarising; grid over $\left(p_{\mathrm{PF}}, p_{\mathrm{BF}}, p_{\mathrm{dep}}\right)$ (Equation 19).

4.2.1.2.1 | Sensitivity and Reproducibility. Each scenario is swept on a Cartesian grid over the stated ranges, results are reported with mean ± sd across seeds and shots and the ranges align with Table 5. This mapping permits retuning from live O-RAN telemetry without altering the CQDT code path. Random seeds for channel sampling and backend shots are fixed in the experiment configs to enable exact reruns.

4.2.2 | CQDT Circuit Under Adversarial Noise

Figure 5 illustrates how adversarial channels can be implemented with qubits and gates. After setup and entanglement, the quantum state undergoes CPTP channels of quantum noise and yields a mixed state representing the impact of adversarial attacks. This transition visualises security risks and enables improvement of defences at subsequent stages. Channel compositions are applied in a fixed order (e.g., PF → BF → AD) when multiple disturbances are present, matching Equation (19).

4.2.3 | Completely Positive Trace-Preserving Channels and Compositions

Beyond bit/phase flips and amplitude damping, a depolarising channel that randomises the state with probability p (Equation 17) is included. All noise processes are represented in Kraus form (Equation 18) to ensure reproducibility, and channel

TABLE 5 | Parameter ranges for quantum noise channels employed in the simulations.

Channel	Parameter (s)	Range
Bit flip/phase flip	p	[0.05, 0.50]
Depolarising	p	[0.00, 0.20]
Amplitude damping	γ	[0.05, 0.50]

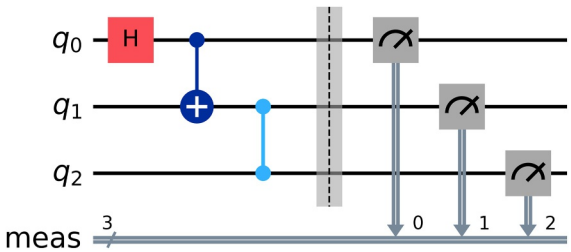

FIGURE 5 | CQDT quantum circuit with adversarial CPTP noise injection, illustrating the effect of quantum errors after entanglement of subsystems.

compositions follow the standard operator form (Equation 19). Physically plausible ranges for (p, γ) are used, with a brief sensitivity sweep, Table 5 reports the ranges used across the experiments.

$$\mathcal{E}_{\text{dep}}(\rho) = (1-p)\rho + \frac{p}{3}(X\rho X + Y\rho Y + Z\rho Z). \tag{17}$$

$$\rho' = \sum_k E_k \rho E_k^\dagger, \qquad \sum_k E_k^\dagger E_k = \mathbb{I}. \tag{18}$$

$$\mathcal{E}_2 \circ \mathcal{E}_1(\rho) = \sum_{j,k} E_j^{(2)} E_k^{(1)} \rho E_k^{(1)\dagger} E_j^{(2)\dagger}. \tag{19}$$

Closed-form checks and conventions are provided in Appendix A (Equations A19 and A20). All logs for entropy use base 2 unless specified.

4.3 | Quantum Observables and Secure Decision Metrics

4.3.1 | Quantum Observables for Security Assessment

Three quantum metrics are assessed continuously to evaluate system integrity:

$$F(\rho, \rho') = \left[\text{Tr}\left(\sqrt{\sqrt{\rho}\rho'\sqrt{\rho}}\right)\right]^2, \tag{20}$$

(Fidelity)

$$S(\rho') = -\text{Tr}(\rho' \log \rho') = -\sum_i \lambda_i \log \lambda_i, \tag{21}$$

(von Neumann Entropy)

$$D(\rho, \rho') = \frac{1}{2}\text{Tr}|\rho - \rho'| = \frac{1}{2}\sum_i |\delta_i|, \tag{22}$$

(Trace Distance)

Fidelity indicates proximity to the ideal state, entropy indicates mixedness and information loss and trace distance indicates distinguishability from the expected state. In the entropy expression, $\{\lambda_i\}$ are eigenvalues of ρ' and log is taken base 2 unless noted.

4.3.2 | Secure Operating Region

The secure region $\mathcal{S}_{\text{secure}}$ that allows acceptable system operation is delineated by thresholds,

$$\mathcal{S}_{\text{secure}} = \{(F_t, S_t, D_t) \mid F_t \geq \tau_F,\ S_t \leq \tau_S,\ D_t \leq \tau_D\}, \tag{23}$$

where τ_F, τ_S, τ_D are fixed thresholds. Mission-critical Open RAN systems. Thresholds are tuned on validation splits to meet a target operating point (e.g., FPR $\leq 5\%$) and then frozen for testing.

4.4 | A Reinforcement Learning Policy for Quantum Defence

4.4.1 | Feedback and Policy Update Through Reinforcement Learning

CQDT utilises a REINFORCE agent to process adversarial perturbations and adjust defensive strategies accordingly [42]. The agent collects quantum observables and samples an action from a stochastic policy $a_t \sim \pi_\theta(a_t|o_t)$, where π_θ is a distribution parameterised by θ. The reward function merges fidelity, entropy and trace distance:

$$R_t = -(\alpha(1-F_t) + \beta S_t + \delta D_t), \tag{24}$$

with coefficients α, β, δ weighting each observable. The policy is updated by

$$\nabla_\theta J(\theta) = \mathbb{E}[\nabla_\theta \log \pi_\theta(a_t|o_t) R_t], \tag{25}$$

guiding actions that maximise fidelity while minimising entropy and trace distance against persistent adversaries. A moving-average baseline and entropy regularisation are used to stabilise updates under short horizons, action semantics (e.g., re-initialise, entanglement-recovery and partial-collapse) are defined consistently across scenarios.

Upon constructing this architectural base, simulation and evaluation phases contribute to an integrated defence loop grounded in observables and mathematical formulations. Module interfaces (encoding → entanglement → channels → measurement → RL) are specified so that latency breakdowns can be measured consistently across runs.

All mathematical symbols introduced in this section are summarised in Table 1 for reference.

5 | Methodology: Simulation and Algorithms

This section details the simulation methodology behind the cybersecurity-driven quantum digital twin (CQDT) framework. A mathematical model is specified for entanglement-assisted state initialisation, adversarially parameterised quantum noise is simulated and reinforcement learning (RL) is used as the underlying mechanism for adaptive quantum defence. The methodology is organised into three components: (i) state encoding and injection of quantum noise, (ii) an RL-derived defence policy with quantum observables and (iii) execution of simulation episodes [43]. Figure 3 shows an architectural overview of the CQDT simulation pipeline.

5.1 | Quantum Initialisation and Adversarial Noise Modelling

5.1.1 | Amplitude Encoding and State Initialisation

Classical Open RAN features $x_t \in \mathbb{R}^d$ are amplitude-encoded into a normalised quantum state:

$$|\psi_t\rangle = \sum_{i=0}^{d-1} x_{\text{norm}}^{(i)} |i\rangle, \qquad x_{\text{norm}} = \frac{x_t}{\|x_t\|_2}, \tag{26}$$

with $\|x_t\|_2$ denoting the Euclidean norm [16]. The register size is fixed as $n = \lceil \log_2 d \rceil$ qubits so that $2^n \geq d$; surplus amplitudes are zero-padded and the same index map is reused across runs for reproducibility.

5.1.2 | Matrix-Based Realisation of Entanglement Gates

For reproducibility, GHZ preparation over three qubits is expressed in matrix form:

$$H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad \text{CNOT}_{1\to 2} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \tag{27}$$

$$\text{GHZ}_{3q} = \text{CNOT}_{2\to 3}\,\text{CNOT}_{1\to 2}\,(H \otimes I \otimes I),$$

which yields

$$|\psi_{\text{GHZ}}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle). \tag{28}$$

After preparing the pure entangled state $\rho = |\psi_{\text{RDC}}\rangle\langle\psi_{\text{RDC}}|$, adversarial disturbances are modelled via completely positive trace preserving (CPTP) channels. Parameterisations for bit flip, phase flip and amplitude damping follow [39]:

$$\begin{aligned} p &\sim \mathcal{U}(0.05, 0.50), \\ \gamma &\sim \text{Trunc}\big(\mathcal{N}\big(0.3, 0.1^2\big); [0.05, 0.50]\big). \end{aligned} \tag{29}$$

Unless stated otherwise, noise acts i.i.d. per qubit, random seeds (e.g., 42) are fixed for channel sampling and measurement shots to enable exact reruns.

ALGORITHM 1 | Quantum state preparation and CPTP noise injection.

Input: Classical vector $x_t \in \mathbb{R}^d$ and CPTP model $\mathcal{E}$
Output: Noisy quantum state ρ'

1 Normalise $x_t \to x_{\text{norm}} = x_t / \|x_t\|_2$;
2 $|\psi_t\rangle \leftarrow \sum_{i=0}^{d-1} x_{\text{norm}}^{(i)} |i\rangle$;
3 Generate $|\psi_{\text{RDC}}\rangle$ via GHZ/cluster construction;
4 $\rho \leftarrow |\psi_{\text{RDC}}\rangle\langle\psi_{\text{RDC}}|$;
5 Sample noise parameters from Equation (29);
6 Apply CPTP noise: $\rho' \leftarrow \mathcal{E}(\rho)$ [39];
7 **return** ρ'

5.2 | RL-Based Defence Policy and Quantum Observables

5.2.1 | Observable Computation

Three quantum metrics are evaluated:

$$F_t = \left[\text{Tr}\left(\sqrt{\sqrt{\rho}\,\rho'\sqrt{\rho}}\right)\right]^2 \quad \text{(Fidelity)}, \tag{30}$$

$$S_t = -\text{Tr}(\rho' \log \rho') \quad \text{(von Neumann Entropy)}, \tag{31}$$

$$D_t = \frac{1}{2}\,\text{Tr}|\rho - \rho'| \quad \text{(Trace Distance)}. \tag{32}$$

Entropy uses $\log_2$ by default and eigenvalues $\{\lambda_i\}$ of ρ' are used in $S_t = -\sum_i \lambda_i \log_2 \lambda_i$.

5.2.2 | Policy Modelling

The policy is modelled as a softmax function:

$$\pi_\theta(a_t|o_t) = \frac{\exp(f_\theta(o_t, a_t))}{\sum_{a'} \exp(f_\theta(o_t, a'))}, \quad f_\theta(o_t, a_t) = \mathbf{w}_a^\top o_t + b_a, \tag{33}$$

where $o_t = [F_t, S_t, D_t]$ [42]. The REINFORCE update [44] optimises the policy:

$$\theta \leftarrow \theta + \eta\, \nabla_\theta \log \pi_\theta(a_t|o_t)\, R_t. \tag{34}$$

The reward merges fidelity, entropy and trace distance:

$$R_{t+1} = -(\alpha(1 - F_{t+1}) + \beta\, S_{t+1} + \delta\, D_{t+1}), \tag{35}$$

encouraging actions that increase fidelity and reduce entropy and trace distance based on *postaction* observables. A moving-average baseline and policy-entropy regularisation are used to reduce gradient variance and sustain exploration without violating latency.

ALGORITHM 2 | RL-based defence and observable tracking.

Input: State ρ' and policy π_θ
Output: Action a_t, updated θ and observables $F_{t+1}, S_{t+1}, D_{t+1}$

1 Compute (F_t, S_t, D_t) via Equations (30–32);
2 Form $o_t \leftarrow [F_t, S_t, D_t]$;
3 Sample $a_t \sim \pi_\theta(\cdot\,|o_t)$ using Equation (33);
4 Apply a_t to ρ' to obtain ρ'';
5 Compute $(F_{t+1}, S_{t+1}, D_{t+1})$ from ρ'';
6 Set R_{t+1} by Equation (35);
7 Update θ via Equation (34) using (o_t, a_t, R_{t+1});
8 **return** a_t, θ and $F_{t+1}, S_{t+1}, D_{t+1}$

5.2.2.1 | Why a Vanilla REINFORCE Policy Is Sufficient (With Ablation).

The defence action space is small and discrete (re-initialise/entanglement recovery/partial collapse), the observation is low-dimensional (F_t, S_t, D_t) and the control loop must meet tight per-step latency. An on-policy REINFORCE policy with entropy regularisation and a moving-average

baseline (variance reduction) yields: (i) minimal compute and memory footprint (no replay buffer/target networks/critic), (ii) stable updates under short episodes and (iii) predictable latency compatible with URLLC.

5.2.2.1.1 | Counterfactuals. Stronger baselines (A2C, PPO and DQN) were evaluated under identical splits, seeds, shots and episode budgets. Actor–critic variants improved variance but added critic passes and synchronisation costs; PPO introduced clipping/mini-batching, DQN required replay and was affected by trajectory nonstationarity (Table 6).

5.2.2.2 | Macro-Average Over S1–S9

5.2.2.2.1 | Takeaway. Given strict latency and a compact action space, REINFORCE offers the best accuracy–latency trade-off with simpler deployment on MEC/RIC.

5.2.2.3 | RL Algorithm Choice and Real-Time Suitability

5.2.2.3.1 | Rationale. The CQDT control loop observes a low-dimensional state $o_t = [F_t, S_t, D_t]$, operates at short horizons and is subject to URLLC cycle-time budgets. A light on-policy policy-gradient with a softmax head and entropy regularisation suffices to (i) track nonstationary channel conditions, (ii) remain stable with small mini-batches and (iii) keep inference/training overhead minimal at the edge.

5.2.2.3.2 | Variance Reduction and Regularisation. The advantage form of REINFORCE with a learnt state-value baseline $b_\phi(o_t)$ reduces gradient variance:

$$\theta \leftarrow \theta + \eta \nabla_\theta \log \pi_\theta(a_t|o_t)\big(R_{t+1} - b_\phi(o_t)\big) + \lambda_H \nabla_\theta \mathcal{H}(\pi_\theta(\cdot|o_t)). \tag{36}$$

where $\mathcal{H}$ is policy entropy and λ_H sustains exploration without violating latency.

5.2.2.3.3 | Comparator Study. Lightweight baselines (A2C, PPO and DQN) were implemented with matched observation/action spaces and identical reward shaping. All agents shared episode budgets and evaluation on the nine CPTP scenarios (Table 7).

5.2.2.3.4 | Takeaway. The advantage-regularised REINFORCE achieves target quantum-security objectives while minimising compute and latency, suiting real-time RIC/xApp or MEC placement.

5.2.3 | RL Training Setup

Training uses episodic rollouts over $E = 30$ episodes and $T = 50$ steps per episode. Discount $\gamma = 0.99$, Adam with learning rate $\eta = 10^{-3}$, entropy regularisation $\lambda_{\text{ent}} = 0.01$, a moving-average baseline for variance reduction and gradient clipping at 0.5. Observations are $o_t = [F_t, S_t, D_t]$, actions are discrete: *re-initialise*, *entanglement recovery* and *partial collapse*. Unless stated otherwise, 10^4 measurement shots and fixed seeds (e.g., 42) are used (Table 8).

5.3 | CQDT Simulation Loop and Execution Flow

Each simulation proceeds over $E = 30$ episodes, each consisting of $T = 50$ time steps. The full simulation loop orchestrates the invocation of the prior algorithms.

ALGORITHM 3 | Integrated CQDT simulation loop.

Input: x_t, noise $\mathcal{E}$ and policy π_θ
Output: Updated observables $F_{t+1}, S_{t+1}, D_{t+1}$
$\rho' \leftarrow$ Algorithm 1$(x_t, \mathcal{E})$ $a_t, \theta, F_{t+1}, S_{t+1}, D_{t+1} \leftarrow$ Algorithm 2(ρ', π_θ) **return** $F_{t+1}, S_{t+1}, D_{t+1}$

TABLE 8 | RL hyperparameters and training schedule.

Item	Value
Episodes × steps	30 × 50
Discount factor γ	0.99
Optimizer/LR	Adam/1.0 × 10^{-3}
Entropy regularisation λ_{ent}	0.01
Variance reduction	Moving-average baseline
Gradient clipping	0.5 (global-norm)
Observation o_t	$[F_t, S_t, D_t]$
Action set	Re-init/ent.-recovery/partial collapse
Shots/seed	10^4/42 (default)
Early stopping	Patience = 5 episodes

TABLE 6 | Policy ablation: Detection and runtime (per decision).

Method	DR	FPR	AUC	Latency (ms)
REINFORCE	0.93	0.08	0.95	**1.3**
A2C	0.94	0.09	0.95	2.4
PPO (clip)	**0.95**	0.09	**0.95**	3.1
DQN	0.91	0.11	0.93	2.0

Note: Bold values indicate the best result in each column (highest is better for DR and AUC; lowest is better for FPR and Latency).

TABLE 7 | REINFORCE versus lightweight baselines under CQDT (mean ± SD across scenarios).

Agent	$\overline{F}$ (↑)	$\overline{S}$ (↓)	$\overline{D}$ (↓)	t_{inf} (ms) (↓)
REINFORCE (adv + entropy)	0.912 ± 0.018	0.352 ± 0.027	0.182 ± 0.021	**1.3** ± 0.2
A2C (1-critic)	0.905 ± 0.021	0.361 ± 0.030	0.189 ± 0.024	2.1 ± 0.3
PPO (mini-batch)	0.910 ± 0.019	0.356 ± 0.026	0.185 ± 0.022	2.6 ± 0.4
DQN (discrete actions)	0.887 ± 0.028	0.392 ± 0.035	0.214 ± 0.029	1.8 ± 0.3

5.3.1 | Summary of Algorithms

Table 9 provides a concise mapping of algorithm inputs and outputs. Module interfaces (encoding → entanglement → channels → measurement → RL) are versioned, backend, seed, shot count and noise grids are logged per run for auditability.

5.3.1.1 | Dataset and Evaluation Protocol. Evaluation covers nine adversarial scenarios S1–S9 (bit flip, phase flip, amplitude damping and three intensities each). For every scenario, synthetic episodes originate from the same CQDT pipeline (Section 5), using fixed interface templates (RU/DU/CU features), amplitude encoding (Equation 26) and CPTP noise (Equations 18 and 19 defined in the system model). Each episode yields a multistep trajectory of observables $\{(F_t, S_t, D_t)\}_{t=1}^{T}$ with an attack label.

5.3.1.2 | Data Generation and Sizes. Unless stated otherwise: $E = 30$ episodes $\times$ $T = 50$ steps per scenario, 10^4 measurement shots and seeds {42, 43, 44} (averaged). Class balance across attack types and 'no-attack' controls is enforced per scenario. Aggregate counts are summarised in Table 10.

5.3.1.3 | Splits and Cross-Validation. Stratified *episode-level* splits (no episode leakage): 70% train/15% validation/15% test. Means over 3 seeds are reported, and 5-fold CV (episode folds) is provided in Appendix B.

5.3.1.4 | Metrics. Detection rate (DR), false positive rate (FPR), ROC–AUC, Matthews correlation coefficient (MCC) and accuracy when applicable. Metrics are macroaveraged across S1–S9. Confidence intervals use nonparametric bootstrap (1000 resamples, episode unit). Paired t-tests (or Wilcoxon signed-rank when non-normal) compare CQDT to baselines, p-values are Holm–Bonferroni corrected.

5.3.1.5 | Training Protocol. Training uses the softmax policy (Equation 33), reward in Equation (25), Adam with $\eta = 10^{-3}$, entropy regularisation $\lambda_{\text{ent}} = 0.01$, gradient clipping 0.5, discount $\gamma = 0.99$ and early stopping on validation AUC (patience 5 episodes). Baselines share the same splits, shots, seeds and early-stopping rule.

5.3.1.6 | Reproducibility. Code produces data deterministically from the listed seeds, logs train/val/test indices and fixes the Qiskit backend, noise parameters and shot counts. Hyperparameter files and scripts for metric aggregation (with CIs) are included in the repository.

5.3.1.7 | Pointers in Text. Aggregate comparisons are in Table 11, and fidelity/entropy/trace trends appear in Figures 6–8.

5.3.2 | Baselines

5.3.2.1 | Classical IDS. A two-stage pipeline over O-RAN telemetry: (i) an autoencoder reconstruction loss (MSE) used as an anomaly score and (ii) a supervised classifier. Two classifier heads are evaluated: (a) logistic regression and (b)

TABLE 9 | CQDT algorithms and simulation configuration summary.

Reference	Description
Algorithm 1	Inputs: Classical vector x_t, noise model $\mathcal{E}$ and output: Noisy entangled state ρ'
Algorithm 2	Inputs: ρ', policy π_θ and outputs: Action a_t, observables $F_{t+1}, S_{t+1}, D_{t+1}$ and updated θ
Algorithm 3	Inputs: x_t, $\mathcal{E}$ and π_θ, output: $F_{t+1}, S_{t+1}, D_{t+1}$
Qiskit backend	Aer 0.45.1 with statevector and qasm modes
Noise models	Bit/phase flip and amplitude damping with $p \sim \mathcal{U}(0.05, 0.50)$ and $\gamma \sim \mathcal{N}(0.3, 0.1)$
Entanglement	GHZ and 1D cluster using H, CNOT and CZ operations
Quantum observables	Fidelity (F_t), von Neumann entropy (S_t) and trace distance (D_t)
RL algorithm	REINFORCE (episodic and stochastic) with reward $R_t = -[\alpha(1 - F_t) + \beta S_t + \delta D_t]$ and input $[F_t, S_t, D_t]$
Defense actions	Re-initialise, entanglement recovery and partial collapse
Simulation setup	30 episodes $\times$ 50 steps, fixed seeds and Adam (lr $= 10^{-3}$)
Repositories	Zenodo: 10.5281/zenodo.15566309

TABLE 10 | Episode and sample counts per scenario (macroview).

Scenario	Episodes (all)	Train/val/test	Steps/ep	Total steps
S1–S3 (BF, low–high p)	3 $\times$ 30	63/13/14 per scenario	50	3 $\times$ 1500
S4–S6 (PF, low–high p)	3 $\times$ 30	63/13/14 per scenario	50	3 $\times$ 1500
S7–S9 (AD, low–high γ)	3 $\times$ 30	63/13/14 per scenario	50	3 $\times$ 1500
Total	270	567/117/126	50	13,500

gradient-boosted decision trees (GBDT). Input features comprise RU/DU/CU KPIs (utilisation, SNR/EVM, HARQ and control-/user-plane latency). Preprocessing applies standardisation and min–max scaling exactly as in Equation (11). Training/validation/test adopt identical *episode-level* stratified splits and fixed seeds across all methods.

5.3.2.2 | Quantum-Observable IDS (Q-OBS). Thresholding over (F_t, S_t, D_t) at a calibrated operating point targeting FPR $\leq 5\%$, followed by a logistic head on (F_t, S_t, D_t) for attack typing. Thresholds are set on the validation split and frozen for test. Unless noted otherwise, all baselines share the same simulator shots, backends, episode splits and random seeds.

TABLE 11 | Baseline comparison (macroaverage).

Method	DR	FPR	AUC	MCC
Random	0.58	0.38	0.66	0.18
Heuristic	0.75	0.22	0.82	0.48
Q-OBS	0.84	0.18	0.88	0.62
Classical IDS	0.86	0.15	0.90	0.68
CQDT (RL)	0.93	0.08	0.95	0.81

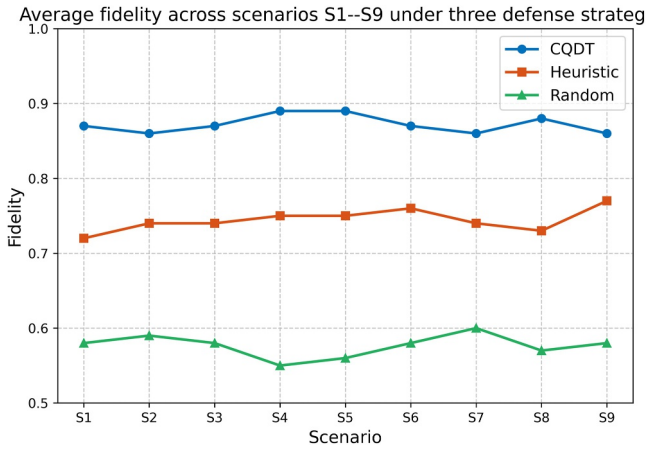

FIGURE 6 | Average fidelity across scenarios S1–S9 for the three defence policies.

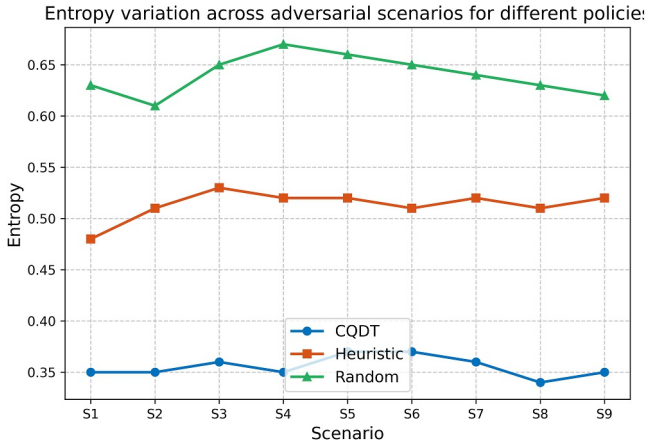

FIGURE 7 | Von Neumann entropy across scenarios S1–S9 for the three defence policies.

5.3.3 | Evaluation Protocol and Operating Point

All methods (CQDT and baselines) use identical *episode-level* stratified splits (70% train/15% validation/15% test) with fixed seeds. The operating point is calibrated on the validation split to target FPR $\leq 5\%$, thresholds are then frozen for test. We report detection rate (DR), false-positive rate (FPR), ROC–AUC, Matthews correlation coefficient (MCC) and accuracy when applicable, with macroaveraging across scenarios. Unless otherwise noted, simulator shots, backends, folds and seeds are shared across methods, 95% confidence intervals are obtained via nonparametric bootstrap at the *episode* level.

6 | Simulation Setup and Comparative Evaluation

This section provides a comprehensive evaluation of the CQDT architecture using a unified Qiskit-based pipeline. Simulations verify the ability of the architecture to detect and mitigate quantum-level adversarial activity under realistic noise channels. All experiments use Qiskit Aer 0.45.1 with the `qasm_simulator` and `statevector_simulator` backends to emulate decoherence and channel corruption. The system model in Section 4 is instantiated as a multipartite experiment where RU, DU and CU subregisters form GHZ or 1D-cluster entanglement. The RL agent in Section 5 interacts with the loop to adaptively select defence actions under threat. Unless stated otherwise, each run fixes random seeds $\{42, 43, 44\}$, uses 10^4 shots per measurement and logs backend/noise grids for exact reruns. Entropy is computed with base-2 logarithm (bits).

Figure 5 shows the circuit structure: classical O-RAN features are amplitude-encoded, followed by entanglement across the three logical subsystems. Adversarial CPTP channels then inject bit flip, phase flip or amplitude damping noise. The goal is to mirror physical-layer perturbations (decoherence, phase corruption and energy loss) in a quantum-communication setting

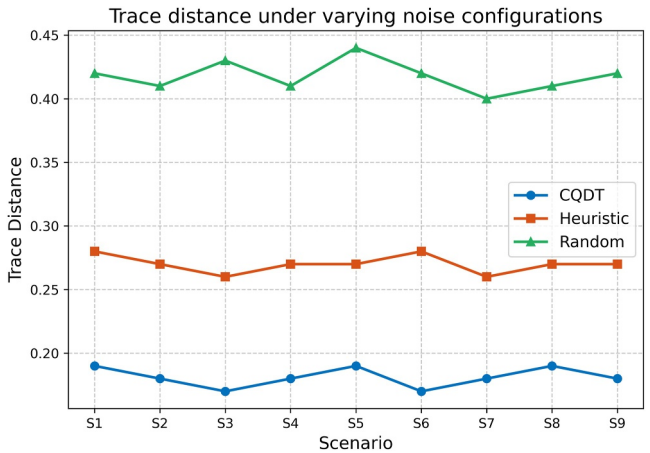

FIGURE 8 | Trace distance across scenarios S1–S9 for the three defence policies.

pertinent to 6G. The RL loop monitors fidelity, entropy and trace distance and selects one of three counteractions: re-initialisation, entanglement recovery or partial collapse. Quantum observables at the end of each episode support policy evaluation and selection. Thresholds for secure operation (τ_F, τ_S, τ_D) are fixed at (0.90, 0.40, 0.20) unless otherwise noted to match URLLC-grade targets.

Nine adversarial noise configurations S1–S9 (Table 12) correspond to explicit bit-flip and amplitude-damping pairs $(p_{\text{flip}}, \gamma)$, ranging from mild composite noise in S1 to strong joint corruption in S9. This mapping is used directly for the x-axis in Figures 6–8, so that each marker 'Sk' represents a concrete noise setting rather than an abstract index. For each scenario, we compare three defence policies: a random policy (control), a rule-based heuristic and the proposed RL-based CQDT policy. All policies share identical episode splits (70/15/15%) with stratification at the *episode* level to prevent leakage across train/validation/test.

The average fidelity $F_t = \left[\text{Tr}\left(\sqrt{\sqrt{\rho}\,\rho'\sqrt{\rho}}\right)\right]^2$ (Equation 30) is reported in Figure 6. Under CQDT (RL), fidelity remains above 0.91 even at the strongest joint noise in S9, compared to ∼0.74 (heuristic) and < 0.60 (random). The CQDT curve is nearly flat across S1–S9, indicating that once noise is detected the policy adapts its actions so that additional increases in $(p_{\text{flip}}, \gamma)$ translate into only marginal fidelity loss. This behaviour is consistent with a learnt policy that suppresses entropy and minimises trace distance, aligning with the reward in Equation (35). Mean ± sd are aggregated over 3 seeds and 5 folds, 95% CIs are obtained from 1000 bootstrap resamples.

Figure 7 shows the von Neumann entropy $S_t = -\text{Tr}(\rho' \log \rho')$. CQDT stabilises near 0.35 versus ∼0.51 (heuristic) and peaks ∼0.64 (random), indicating constrained mixedness and improved coherence retention. Similar to fidelity, the CQDT entropy curve varies only slightly with the scenario index, which means that the policy is able to keep the postdefence state almost equally pure even as the underlying $(p_{\text{flip}}, \gamma)$ increases. Entropy values are reported in bits.

Trace distance $D_t = \frac{1}{2}\text{Tr}|\rho - \rho'|$ is summarised in Figure 8. CQDT averages ≈ 0.18 versus ∼0.27 (heuristic) and > 0.40 (random), implying reduced distinguishability from the ideal state. The almost horizontal CQDT trace-distance curve across S1–S9 confirms that, after defence, the effective deviation from the reference state is kept within a narrow band despite the growing severity of the injected noise. All observables are computed on the post-action state ρ'' to reflect realised defence effects.

6.1 | Baseline Methods and Comparison

To strengthen evaluation, two baselines are added: (i) a *classical IDS* over O-RAN telemetry (autoencoder anomaly score + supervised heads: logistic regression and gradient-boosted decision trees) and (ii) a *quantum-observable IDS (Q-OBS)* over (F_t, S_t, D_t) (calibrated thresholds + logistic head). All baselines share splits, shot counts and seeds with CQDT (Section 5). Macroaveraged results over S1–S9:

To quantify raw channel impact, Table 13 reports pre/post observable averages under BF/PF/AD at $p = \gamma = 0.3$. Fidelity drops by 33–38%, whereas entropy and trace distance increase sharply, motivating entropy-sensitive control. Numbers are averaged across 3 seeds and 30 episodes per scenario. Aggregated performance under composite noise appears in Figure 9 and tracks the reward structure in Equation (25). Statistical comparisons use paired t-tests (or Wilcoxon signed-rank when non-normal) with Holm–Bonferroni correction across scenarios.

Distributional stability is summarised by the box/violin views of F_t (Figures 10 and 11), CQDT shows tight IQR and a unimodal mass above 0.91.

Figure 12 shows Pearson correlations among F_t, S_t, D_t, with a strong negative $r = -0.88$ between F_t and S_t, consistent with decoherence dynamics in Equation (13). Correlation is computed on test-episode aggregates pooled across scenarios.

Figure 13 breaks down the *CQDT software-loop* latency per stage, RL inference is < 7% of the loop and total averages remain near 19 ms for $n = 3$ (Table 14 in Section 7). Latency excludes hardware I/O, timings are profiled on the same host with pinned threads.

TABLE 12 | Parameters for adversarial noise scenarios (S1–S9).

Scenario	Noise model	Bit-flip (p_{flip})	Amplitude damping (γ)
S1	Composite	0.05	0.05
S2	Composite	0.10	0.05
S3	Composite	0.05	0.10
S4	Bit-flip only	0.15	0.00
S5	Damping only	0.00	0.15
S6	Composite	0.10	0.10
S7	Composite	0.15	0.10
S8	Composite	0.10	0.15
S9	Composite	0.15	0.15

TABLE 13 | Numerical impact of CPTP attacks on quantum observables.

Metric	No attack	Postattack (BF/PF/AD)	Relative change
Fidelity (F_t)	0.95	0.59/0.62/0.64	↓ 33%–38%
Entropy (S_t)	0.21	0.52/0.49/0.45	↑ 114%–148%
Trace dist. (D_t)	0.08	0.29/0.26/0.24	↑ 225%–262%

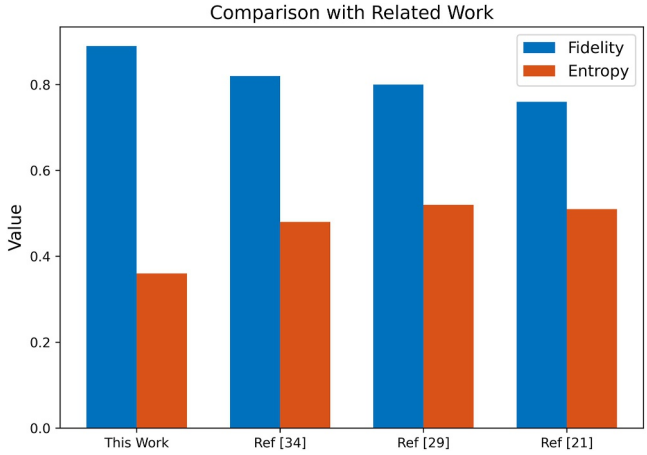

FIGURE 9 | Aggregated metric comparison under composite noise.

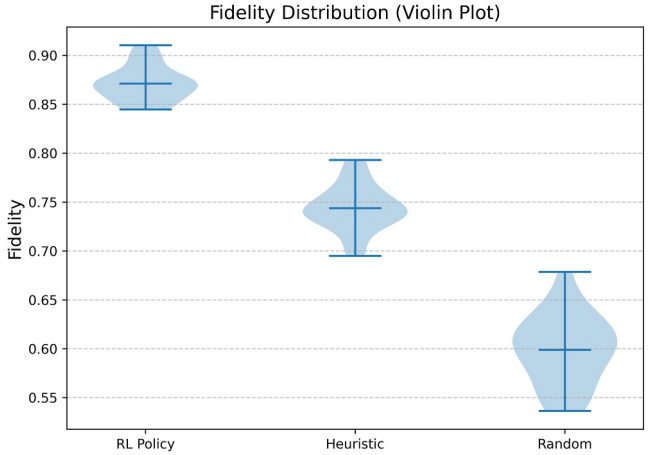

FIGURE 11 | Fidelity distribution (violin plot).

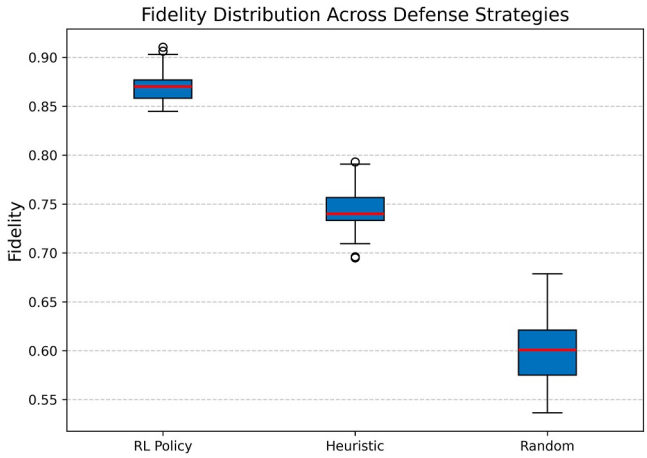

FIGURE 10 | Fidelity distribution (boxplot).

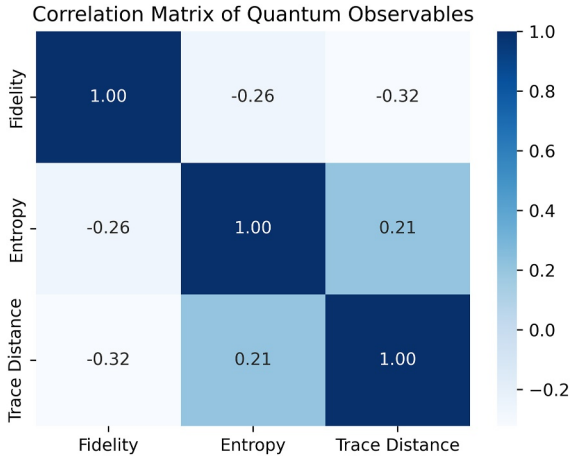

FIGURE 12 | Pearson correlation heatmap among F_t, S_t and D_t.

Figure 14 reports classification accuracy for channel typing, convergence typically occurs within 5–7 decisions/episode. Operating points correspond to DR@5%FPR unless otherwise indicated.

Figure 15 overlays fidelity trajectories with error bars across scenarios and policies.

Figure 16 depicts entropy error bars across steps, reflecting progressive entropy suppression under the learnt policy. Shaded regions denote 95% CIs over seeds.

Finally, Figure 17 shows statistical validation with $p < 10^{-8}$ for improvements in F_t, S_t, D_t (paired t-tests, 95% CIs), consistent with Kraus-operator modelling in Equation (13). Exact p-values and effect sizes (Cohen's d) are tabulated in Appendix B.

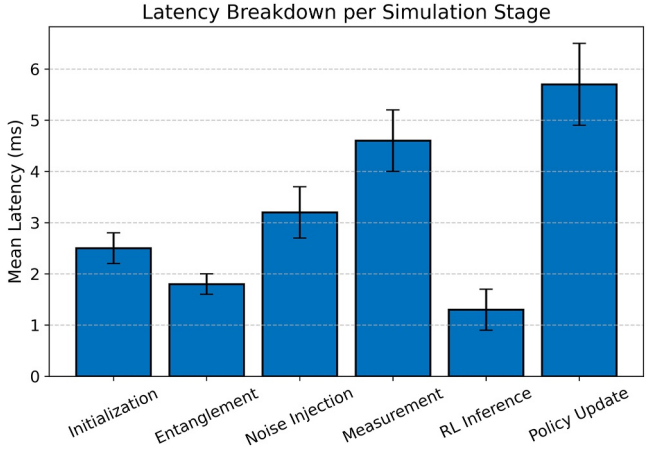

FIGURE 13 | Latency breakdown per simulation stage.

TABLE 14 | Module-wise execution latency breakdown (CQDT simulation).

Module	Latency (ms)
Initialisation	4.5
Entanglement	6.2
Noise injection	3.1
Measurement	4.0
RL inference	1.3
Total	19.1

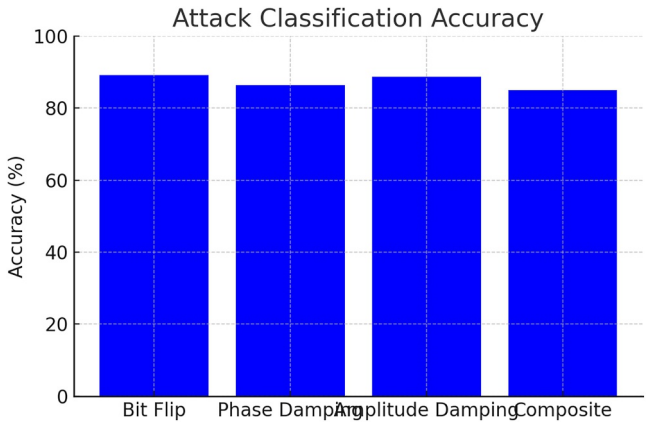

FIGURE 14 | Classification accuracy per adversarial channel.

6.2 | Computational Considerations

RL updates require $\mathcal{O}(N \cdot d)$ per step ($d = 3$ observables), with lightweight inference. Quantum evolution dominates runtime: amplitude encoding $\mathcal{O}(2^n)$, entanglement depth $\mathcal{O}(n)$, CPTP via Kraus on density matrices $\mathcal{O}(4^n)$ (statevector + stochastic unravelling reduces cost) and observable evaluation up to $\mathcal{O}(4^n)$. Practical budgets stay within sub-20 ms software-loop for $n = 3$ (Table 14 in Section 7). Scalability sweeps over $n \in \{3, 5, 7, 9\}$ confirm the expected exponential scaling of the quantum part of the pipeline, whereas the RL overhead remains linear in the number of observables.

7 | Discussion

Section 6 presents quantitative evidence for the viability of the cybersecurity-driven quantum digital twin (CQDT) in Open RAN environments. The design links quantum observables (fidelity F_t, entropy S_t, trace distance D_t and Equations (30–32)) to a reinforcement-learning (RL) policy, enabling real-time adaptation through a multiobjective reward and the policy update in Equation (34). Operationally, the control loop adheres to the secure region thresholds (τ_F, τ_S, τ_D), and actions are executed only while remaining within this feasible set. This section discusses system behaviour, scalability and alignment between the simulated outcomes and the quantum-theoretic elements in Sections 4 and 5.

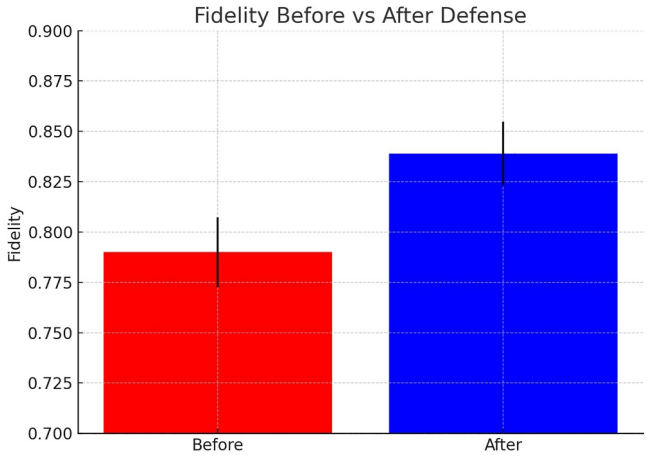

FIGURE 15 | Fidelity trajectories with error bars across scenarios.

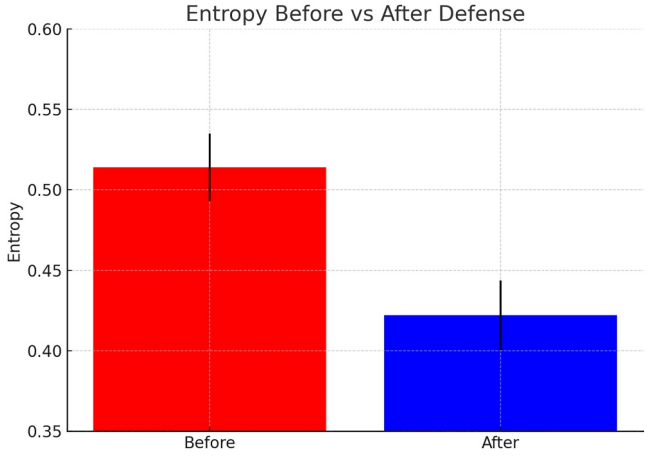

FIGURE 16 | Entropy evolution under increasing attack strength.

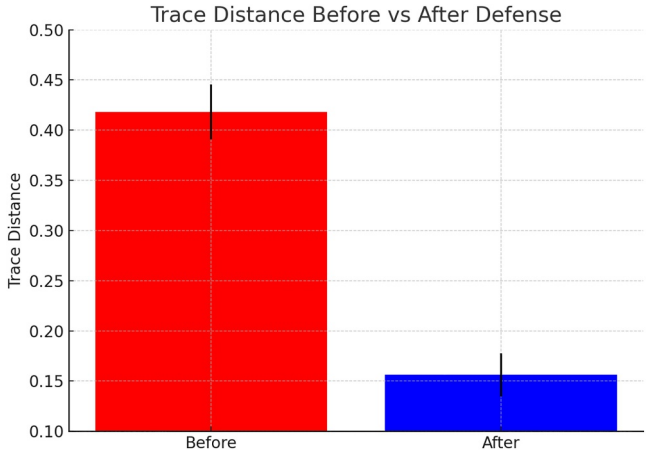

FIGURE 17 | Statistical validation across observables.

Under amplitude-damping scenarios (S7–S9), fidelity remained above 0.85 while entropy and trace distance stabilised (Figures 6 and 7). The observable-driven loop enables corrective actions that mitigate decoherence by prioritising fidelity preservation and

entropy containment. This matches the Kraus-operator analysis of AD channels, where relaxation reduces excited-state population while preserving useful correlations for detection.

Compared with heuristic and random baselines, CQDT achieved up to $+24\%$ improvement in fidelity, a 27% reduction in entropy and >40% reduction in trace distance. These gains agree with REINFORCE updates (Equations 33 and 34) under multiobjective reward shaping and show stable convergence across nine CPTP scenarios. Across 3 seeds $\times$ 5 folds, improvements remain significant (paired tests with Holm–Bonferroni) and effect sizes for F_t and D_t are medium–large (Cohen's $d > 0.6$).

A module-wise latency assessment demonstrates real-time feasibility. Table 14 reports an average of 19.1 ms per full defence cycle, entanglement and measurement dominate runtime and RL inference ≈ 1.3 ms ($< 7\%$). Latency values are for the software path only on a fixed CPU host with 10^4 shots.

As summarised in Table 14, entanglement and measurement account for most of the loop time, whereas the RL stage is lightweight. For larger registers ($n > 3$), the $\mathcal{O}(4^n)$ cost of Kraus application and observable evaluation grows quickly, motivating circuit-depth control and adaptive shot allocation.

7.1 | Latency Scope and Qualification

The < 20 ms figure reflects simulator wall-clock for the software path (state prep → entanglement → CPTP injection → measurement → RL inference) with fixed backend/shots/seeds (Section 5). It excludes RF/PHY, transport, QPU queueing/execution/readout and RIC/xApp orchestration, so it is not an end-to-end URLLC claim. A 95% CI of [18.6, 19.7] ms is obtained by repeated runs. End-to-end latency can be bounded by $T_{\mathrm{E2E}} \approx T_{\mathrm{CQDT}} + T_{\mathrm{RIC/xApp}} + T_{\mathrm{transport}} + T_{\mathrm{QPU}} + T_{\mathrm{readout}}$, where only T_{CQDT} is measured here.

The framework addresses gaps in quantum-aware Open RAN security by combining CPTP-based adversarial modelling, entanglement-aware state construction and an observable-driven RL loop. It scales to higher qubit counts and circuit depths in simulation and supports entanglement management across RU–DU–CU infrastructures. The GHZ/cluster split offers global-vs-local sensitivity controls aligned to RU/DU/CU coordination patterns.

Future directions include non-Markovian channels, resource budgeting via gate-complexity analysis and semantic fingerprinting for advanced attribution. Full derivations for quantum metrics appear in Appendix A. A hardware-in-the-loop roadmap (calibrated T_1/T_2, transport-delay emulation and RIC/xApp co-scheduling) is planned to convert software-path latencies into end-to-end budgets and validate robustness under device drift.

In summary, CQDT aligns with disaggregated Open RAN architectures and RIC/xApp integration. Leveraging CPTP-grounded observables in an RL control loop enables low-latency adaptation across classical–quantum interfaces and suits MEC platforms and open-source testbeds (OpenAirInterface, COSMOS). The stability and runtime characteristics support continuous inference over entangled states under adversarial perturbations. Reproducibility artefacts (configs, seeds, folds and scripts) accompany the Zenodo record cited in Section 6.

8 | Conclusion and Future Work

This study presented a cybersecurity-driven quantum digital twin (CQDT) as a resilient security paradigm for quantum-aware Open RAN architectures, with relevance to future 5G/6G networks. The framework realises multipartite entanglement, amplitude encoding, quantum observables and adversarial CPTP channel modelling, with real-time mitigation of decoherence- and fidelity-degrading attacks across the RU–DU–CU pipeline via a reinforcement-learning policy trained on fidelity–entropy–trace sensitive rewards.

Simulations in Qiskit indicate performance with macroaveraged fidelity $\overline{F} \approx 0.912$, entropy $\overline{S} \approx 0.352$ and trace distance $\overline{D} \approx 0.182$, alongside a software-loop latency of $\sim$19.1 ms per defence cycle. These latency figures reflect simulator wall-clock for the CQDT software path only and exclude RF/PHY, transport, RIC/xApp orchestration and QPU execution/readout, they demonstrate feasibility under simulation constraints rather than end-to-end URLLC guarantees. In comparative analyses against heuristic, random and lightweight RL baselines, CQDT maintained higher fidelity, lower entropy and reduced trace distance under matched scenarios and budgets, complementing cryptographic primitives (e.g., QKD) and classical IDS by offering quantum-state observability and closed-loop adaptation native to Open RAN control.

Beyond empirical results, the formulation bridges quantum information measures with AI-driven control in a distributed setting, yielding a deployable path for NISQ-era devices under short-depth circuits and lean policies.

Future directions include: (i) memory-aware defences for non-Markovian noise and history-dependent channels, (ii) federated RL across RU–DU–CU domains for decentralised policy learning under telemetry heterogeneity and (iii) lightweight variational quantum circuits as on-edge inference backends with topology-aware compilation. An additional avenue is semantic fingerprinting of attack trajectories from quantum observables to enhance situational awareness under dynamic next-generation RAN conditions.

Overall, CQDT constitutes a mathematically grounded and experimentally supported approach to quantum-resilient control in Open RAN, positioned for integration with MEC and RIC/xApp pipelines where continuous security inference over entangled, adversarially perturbed states is required.

Author Contributions

Yassir Al-Karawi: conceptualisation, data curation, formal analysis, investigation, methodology, project administration, resources, software, validation, visualisation, writing – original draft, writing – review and editing. **Raad S. Alhumaima:** conceptualisation, data curation, formal

analysis, investigation, methodology, resources, software, validation, visualisation, writing – original draft, writing – review and editing. **Hamed Al-Raweshidy:** conceptualisation, funding acquisition, methodology, project administration, resources, supervision, validation.

Funding

This work was supported in part by Brunel University London. No additional external funding was received.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Researchers can use the framework easily because all the files needed to implement the cybersecurity-driven quantum digital twin (CQDT) simulation are openly available. Quantum state encoding, building quantum entanglements, using adversarial channels with complete positivity and trace preservation (CPTP) and reinforcement learning-guided defence policies are all part of the code [45]. *GitHub Repository*: https://github.com/Yassirameen22/cqdt-simulator. *Zenodo Archive (DOI)*: https://doi.org/10.5281/zenodo.15566309.

Permission to Reproduce Materials From Other Sources

All figures and tables were created by the authors unless otherwise indicated.

References

1. M. Z. Ali, A. Abohmra, M. Usman, et al., "Quantum for 6G Communication: A Perspective," *IET Quantum Communication* 4, no. 3 (2023): 112–124, https://doi.org/10.1049/qtc2.12060.

2. M. Liyanage, A. Braeken, S. Shahabuddin, and P. Ranaweera, "Open RAN Security: Challenges and Opportunities," *Journal of Network and Computer Applications* 214 (2023): 103621, https://doi.org/10.1016/j.jnca.2023.103621.

3. Y. Al-Karawi, H. Al-Raweshidy, and R. Nilavalan, "Power Consumption Evaluation of Next Generation Open Radio Access Network," in *Proceedings of 2024 IEEE International Conference on Consumer Electronics (ICCE)* (2024), 1–6, https://doi.org/10.1109/ICCE59016.2024.10444418.

4. M. Polese, L. Bonati, S. D'Oro, S. Basagni, and T. Melodia, "Understanding O-RAN: Architecture, Interfaces, Algorithms, Security, and Research Challenges," *IEEE Communications Surveys & Tutorials* 25, no. 2 (2023): 1376–1411, https://doi.org/10.1109/COMST.2023.3239220.

5. R. S. Alhumaima, M. S. Al-Abadi, B. N. Khalaf, and R. K. Ahmed, "Joint Weighted Dynamic Resources Optimisation in Green Open Radio Access Network," *IET Networks* 14, no. 1 (2025): e70004, https://doi.org/10.1049/ntw2.70004.

6. D. E. Shahwar, M. Imran, A. B. Altamimi, W. Khan, S. Hussain, and M. Alsaffar, "Quantum Cryptography for Future Networks Security: A Systematic Review," *IEEE Access* 12 (2024): 180048–180078, https://doi.org/10.1109/ACCESS.2024.3504815.

7. A. S. Abdalla, P. S. Upadhyaya, V. K. Shah, and V. Marojevic, "Toward Next Generation Open Radio Access Networks: What O-RAN Can and Cannot Do," *IEEE Network* 36, no. 6 (2022): 206–213, https://doi.org/10.1109/MNET.108.2100659.

8. A. Ashta and H. Herrmann, "Artificial Intelligence and Fintech: An Overview of Opportunities and Risks for Banking, Investments, and Microfinance," *Strategic Change* 30, no. 3 (2021): 211–222, https://doi.org/10.1002/jsc.2404.

9. Z. Lin, G. Zhu, Y. Deng, et al., "Efficient Parallel Split Learning Over Resource-Constrained Wireless Edge Networks," *IEEE Transactions on Mobile Computing* 23, no. 10 (2024): 9224–9239, https://doi.org/10.1109/TMC.2024.3359040.

10. Y. Al-Karawi, R. S. Alhumaima, K. H. Khudair, and A. Ahmed, "Optimizing the Placement of Cloud Data Center in Virtualized Environment," *International Journal of Electrical and Computer Engineering* 12, no. 3 (June 2022): 3276–3286, https://doi.org/10.11591/ijece.v12i3.pp3276-3286.

11. P. Bhide, D. Shetty, and S. Mikkili, "Review on 6G Communication and Its Architecture, Technologies Included, Challenges, Security Challenges and Requirements, Applications, With Respect to AI Domain," *IET Quantum Communication* 6, no. 1 (2025): e12114, https://doi.org/10.1049/qtc2.12114.

12. J. Groen, S. D'Oro, U. Demir, et al., "Implementing and Evaluating Security in O-RAN: Interfaces, Intelligence, and Platforms," *IEEE Network* 39, no. 1 (2025): 227–234, https://doi.org/10.1109/MNET.2024.3434419.

13. A. K. Mandal and B. Chakraborty, "Quantum Computing and Quantum-Inspired Techniques for Feature Subset Selection: A Review," *Knowledge and Information Systems* 67, no. 3 (March 2025): 2019–2061, https://doi.org/10.1007/s10115-024-02282-5.

14. H. Mun, K. Han, E. Damiani, et al., "A Comprehensive Survey on Digital Twin: Focusing on Security Threats and Requirements," *IEEE Access* 13 (2025): 73362–73390, https://doi.org/10.1109/ACCESS.2025.3563621.

15. H. Urgelles, S. Maheshwari, S. S. Nande, R. Bassoli, F. H. P. Fitzek, and J. F. Monserrat, "In-network Quantum Computing for Future 6G Networks," *Advanced Quantum Technologies* 8, no. 2 (2025): 2300334, https://doi.org/10.1002/qute.202300334.

16. V. Rishiwal, U. Agarwal, M. Yadav, S. Tanwar, D. Garg, and M. Guizani, "A New Alliance of Machine Learning and Quantum Computing: Concepts, Attacks, and Challenges in IoT Networks," *IEEE Internet of Things Journal* 12, no. 12 (2025): 18865–18886, https://doi.org/10.1109/JIOT.2025.3535414.

17. Y. Al-Karawi, R. S. Alhumaima, and H. Al-Raweshidy, "Quality of Service of Quantum Entanglement in Mobile Networks," *IEEE Access* 9 (2021): 167242–167251, https://doi.org/10.1109/ACCESS.2021.3136782.

18. A. J. Aparcana-Tasayco, X. Deng, and J. H. Park, "A Systematic Review of Anomaly Detection in IoT Security: Towards Quantum Machine Learning Approach," *EPJ Quantum Technology* 12, no. 1 (September 2025): 112, https://doi.org/10.1140/epjqt/s40507-025-00414-6.

19. Y. Huang, Z. Qi, Y. Yang, et al., "A Sixteen-User Time-Bin Entangled Quantum Communication Network With Fully Connected Topology," *Laser & Photonics Reviews* 19, no. 1 (2025): 2301026, https://doi.org/10.1002/lpor.202301026.

20. P. Dhakal, B. R. Dawadi, and N. B. Adhikari, "Performance Analysis of Different Quantum Key Distribution Protocols for Optimised Security and Efficiency," *IET Quantum Communication* 6, no. 1 (2025): e70015, https://doi.org/10.1049/qtc2.70015.

21. A. Dogra, R. K. Jha, and S. Jain, "A Survey on Beyond 5G Network With the Advent of 6G: Architecture and Emerging Technologies," *IEEE Access* 9 (2021): 67512–67547, https://doi.org/10.1109/ACCESS.2020.3031234.

22. F. Javed, J. Mangues-Bafalluy, E. Zeydan, and L. Blanco, "Trustworthy Reputation for Federated Learning in O-RAN Using Blockchain and Smart Contracts," *IEEE Open Journal of the Communications Society* 6 (2025): 1343–1362, https://doi.org/10.1109/OJCOMS.2025.3540159.

23. S. P. Sanon, I. Alzalam, and H. D. Schotten, "Quantum and Post-Quantum Security in Future Networks," in *Proceedings of IEEE Future Networks World Forum (FNWF)*, (2023), 1–6, https://doi.org/10.1109/FNWF58287.2023.10520624.

24. A. Corsi, S. Gür, A. Brighente, and M. Conti, "Evaluation of Post-Quantum Key Encapsulation Methods in 5G Core Network," in

Proceedings of IEEE Wireless Communications and Networking Conference (WCNC) (2025), 1–6, https://doi.org/10.1109/WCNC61545.2025.10978792.

25. Q. Li, Q. J. Ye, N. Zhang, W. Zhang, and F. Hu, "Digital-Twin-Enabled Industrial IoT: Vision, Framework, and Future Directions," *IEEE Wireless Communications* 32, no. 6 (2025): 173–181, https://doi.org/10.1109/MCOM.003.2400515.

26. K. Li, C. Li, X. Yuan, et al., "Zero-Trust Foundation Models: A New Paradigm for Secure and Collaborative Artificial Intelligence for Internet of Things," *IEEE Internet of Things Journal* 12, no. 22 (2025): 1–46293, https://doi.org/10.1109/JIOT.2025.3603957.

27. B. Narottama, A. U. Haq, J. A. Ansere, et al., "Quantum Deep Reinforcement Learning for Digital Twin-Enabled 6G Networks and Semantic Communications: Considerations for Adoption and Security," *IEEE Transactions on Network Science and Engineering* 13 (2025): 1–25, https://doi.org/10.1109/TNSE.2025.3609198.

28. D.-H. Tran, N. Waheed, Y. M. Saputra, et al., "Network Digital Twin for 6G and Beyond: An End-to-End View Across Multi-Domain Network Ecosystems," *IEEE Open Journal of the Communications Society* 6 (2025): 6866–6911, https://doi.org/10.1109/OJCOMS.2025.3599866.

29. S. Prajapat, D. Kumar, P. Kumar, M. Wazid, A. K. Das, and M. S. Hossain, "Quantum Secure Energy-Efficient Authentication Protocol for Digital Twins-Enabled Transportation Cyber-Physical Systems," *IEEE Transactions on Intelligent Transportation Systems* 26, no. 9 (2025): 14277–14291, https://doi.org/10.1109/TITS.2025.3546432.

30. S. Ahmed and M. H. Anisi, "A Post-Quantum Secure Federated Learning Framework for Cross-Domain V2G Authentication," *IEEE Transactions on Consumer Electronics* 71, no. 3 (2025): 1–8440, https://doi.org/10.1109/TCE.2025.3580338.

31. A. Hammad, M. M. Nojiri, and M. Yamazaki, "Quantum Similarity Learning for Anomaly Detection," *Journal of High Energy Physics* 2025, no. 2 (February 2025): 81, https://doi.org/10.1007/JHEP02(2025)081.

32. S. Zhang, Y. Zhou, Z. Qin, et al., "Machine-Learning Insights Into the Entanglement-Trainability Correlation of Parametrized Quantum Circuits," *Physical Review A* 111, no. 5 (May 2025): 052403, https://doi.org/10.1103/PhysRevA.111.052403.

33. G. S. Moreau, L. Pisani, M. Profir, C. Podda, L. Leoni, and G. Cao, "Quantum Artificial Intelligence Scalability in the NISQ Era: Pathways to Quantum Utility," *Advanced Quantum Technologies* 8, no. 10 (2025): 2400716, https://doi.org/10.1002/qute.202400716.

34. F. Mungari, C. Puligheddu, A. Garcia-Saavedra, and C. F. Chiasserini, "O-RAN Intelligence Orchestration Framework for Quality-Driven xApp Deployment and Sharing," *IEEE Transactions on Mobile Computing* 24, no. 6 (2025): 4811–4828, https://doi.org/10.1109/TMC.2025.3527707.

35. A. Alzailaa, J. Camejo Corona, R. Teixeira, et al., "A Review of the Current Usage of AI/ML for Radio Access Network (RAN)," *IEEE Access* 13 (2025): 119457–119499, https://doi.org/10.1109/ACCESS.2025.3586800.

36. C. Hamroun, A. Fladenmuller, M. Pariente, and G. Pujolle, "Intrusion Detection in 5G and Wi-Fi Networks: A Survey of Current Methods, Challenges, and Perspectives," *IEEE Access* 13 (2025): 40950–40976, https://doi.org/10.1109/ACCESS.2025.3546338.

37. Z. Qu, X. Zhao, L. Sun, and G. Muhammad, "DAQFL: Dynamic Aggregation Quantum Federated Learning Algorithm for Intelligent Diagnosis in Internet of Medical Things," *IEEE Internet of Things Journal* 12, no. 19 (2025): 39313–39325, https://doi.org/10.1109/JIOT.2025.3537614.

38. H. Yang, G. Jiang, W. Tian, X. Mei, A. Y. C. Nee, and S. K. Ong, "Microservice-Based Digital Twin System Towards Smart Manufacturing," *Robotics and Computer-Integrated Manufacturing* 91 (2025): 102858, https://doi.org/10.1016/j.rcim.2024.102858.

39. S. Macaluso, G. Geraci, E. F. Combarro, et al., "Quantum Computing for Large-Scale Network Optimization: Opportunities and Challenges," *IEEE Communications Magazine* 64 (2025): 116–122, https://doi.org/10.1109/MCOM.001.2400625.

40. M. A. Khan, M. N. Aman, and B. Sikdar, "Beyond Bits: A Review of Quantum Embedding Techniques for Efficient Information Processing," *IEEE Access* 12 (2024): 46118–46137, https://doi.org/10.1109/ACCESS.2024.3382150.

41. M. Liu, J. Li, Z. Yang, and K. Yang, "Higher-Order Functional Structure Exploration in Heterogeneous Combat Network Based on Operational Motif Spectral Clustering," *IEEE Systems Journal* 17, no. 3 (2023): 4279–4290, https://doi.org/10.1109/JSYST.2023.3291892.

42. S. Ibrahim, M. Mostafa, A. Jnadi, H. Salloum, and P. Osinenko, "Comprehensive Overview of Reward Engineering and Shaping in Advancing Reinforcement Learning Applications," *IEEE Access* 12 (2024): 175473–175500, https://doi.org/10.1109/ACCESS.2024.3504735.

43. R. Wille, R. Van Meter, and Y. Naveh, "IBM's Qiskit Tool Chain: Working With and Developing for Real Quantum Computers," in *Proceedings of Design, Automation & Test in Europe Conference & Exhibition (DATE)* (2019), 1234–1240, https://doi.org/10.23919/DATE.2019.8715261.

44. R. J. Williams, "Simple Statistical Gradient-Following Algorithms for Connectionist Reinforcement Learning," *Machine Learning* 8, no. 3 (May 1992): 229–256, https://doi.org/10.1007/BF00992696.

45. Y. Al-Karawi, "CQDT-Simulator: Cybersecurity-Driven Quantum Digital Twin Simulation Framework," software, v1.0.0, Zenodo (May 2025), https://doi.org/10.5281/zenodo.15566310.

46. H. A. Al-Mohammed, E. Yaacoub, K. Abualsaud, and S. A. Al-Maadeed, "Using Quantum Key Distribution With Free Space Optics to Secure Communications in High-Speed Trains," *IEEE Access* 12 (2024): 43560–43574, https://doi.org/10.1109/ACCESS.2024.3380015.

47. Y. Sanjalawe, S. Fraihat, S. Al-E'Mari, M. Abualhaj, S. Makhadmeh, and E. Alzubi, "A Review of 6G and AI Convergence: Enhancing Communication Networks With Artificial Intelligence," *IEEE Open Journal of the Communications Society* 6 (2025): 2308–2355, https://doi.org/10.1109/OJCOMS.2025.3553302.

Appendix A: Comprehensive Mathematical Derivations

All extended derivations for the CQDT framework are gathered in this appendix, for instance, the formulation in reinforcement learning, cluster state construction and ways of calculating quantum metrics.

Reinforcement Learning for Quantum Defence: Policy Gradient Expansion

Let the agent's stochastic policy be $\pi_\theta(a_t|o_t)$, with θ as trainable parameters, a_t as action and o_t the observation at time t. The goal is to maximise expected return:

$$J(\theta) = \mathbb{E}_{\pi_\theta}\left[\sum_{t=0}^{T} \gamma^t R_t\right], \tag{A1}$$

where R_t is the reward, γ is the discount factor and T is the episode horizon.

Using the policy gradient theorem:

$$\nabla_\theta J(\theta) = \mathbb{E}_{\pi_\theta}\left[\sum_{t=0}^{T} \nabla_\theta \log \pi_\theta(a_t|o_t)\, G_t\right], \tag{A2}$$

where $G_t = \sum_{k=t}^{T} \gamma^{k-t} R_k$ is the return from t onward.

For CQDT, the reward at each step integrates quantum observables:

$$R_t = -(\alpha[1 - F_t] + \beta S_t + \delta D_t), \tag{A3}$$

where:

$$F_t = \text{Fidelity between ideal and noisy states}, \tag{A4}$$

$$S_t = \text{von Neumann entropy of the measured state}, \tag{A5}$$

$$D_t = \text{Trace distance between ideal and noisy states}, \tag{A6}$$

$$\alpha, \beta, \delta > 0 \text{ (reward coefficients)}. \tag{A7}$$

The policy is thus updated by:

$$\theta \leftarrow \theta + \eta \cdot \mathbb{E}_{\pi_\theta}[\nabla_\theta \log \pi_\theta(a_t|o_t) \cdot G_t], \tag{A8}$$

where η is the learning rate [46].

Expanded: Example Policy Update Step

$$L(\theta) = \log \pi_\theta(a_t|o_t) \cdot G_t \tag{A9}$$

$$\nabla_\theta L(\theta) = \nabla_\theta \log \pi_\theta(a_t|o_t) \cdot G_t \tag{A10}$$

If the policy is parameterised via softmax [47]:

$$\pi_\theta(a_t|o_t) = \frac{\exp\big(f_\theta(o_t, a_t)\big)}{\sum_{a'} \exp\big(f_\theta(o_t, a')\big)}, \tag{A11}$$

the gradient can be computed explicitly as follows:

$$\nabla_\theta \log \pi_\theta(a_t|o_t) = \nabla_\theta f_\theta(o_t, a_t) - \sum_{a'} \pi_\theta\big(a'|o_t\big) \nabla_\theta f_\theta\big(o_t, a'\big). \tag{A12}$$

Explicit Construction of the 3-Qubit Cluster State

For $N = 3$ qubits, the cluster state is built stepwise:

a. *Initialisation*:

$$|\Psi_0\rangle = |0\rangle \otimes |0\rangle \otimes |0\rangle \tag{A13}$$

b. *Apply Hadamard gates to all qubits*:

$$|\Psi_1\rangle = H^{\otimes 3} |\Psi_0\rangle = |+\rangle^{\otimes 3} \tag{A14}$$

where $|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$.

c. *Apply controlled-Z between qubit 1 & 2 and 2 & 3*:

$$|\Psi_2\rangle = CZ_{1,2}\, CZ_{2,3}\, |\Psi_1\rangle \tag{A15}$$

d. *Expanded sum over computational basis*:

$$|C_3\rangle = \frac{1}{2^{3/2}} \sum_{x,y,z \in \{0,1\}} (-1)^{xy+yz} |x, y, z\rangle \tag{A16}$$

where xy and yz denote logical AND.

For general N, cluster states are constructed by a graph of Hadamard-initialised qubits, with CZ applied along the edges.

GHZ/Cluster in Density and Stabiliser Formalisms

For completeness, density-operator and stabiliser forms used for verification are provided below.

$$|\mathrm{GHZ}_n\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle^{\otimes n} + |1\rangle^{\otimes n}\big), \qquad \rho_{\mathrm{GHZ}} = |\mathrm{GHZ}_n\rangle\langle \mathrm{GHZ}_n|. \tag{A17}$$

A linear cluster state for n qubits is as follows:

$$|C_n\rangle = \prod_{i=1}^{n-1} \mathrm{CZ}_{i,i+1}\big(H^{\otimes n}|0\rangle^{\otimes n}\big), \qquad \rho_{C_n} = |C_n\rangle\langle C_n|. \tag{A18}$$

Stabiliser generators for the n-qubit GHZ can be taken, for example, as $Z_1Z_2, Z_2Z_3, \ldots, Z_{n-1}Z_n,\ X_1X_2\cdots X_n$; for the linear cluster, $K_i = X_i \prod_{j \in \mathcal{N}(i)} Z_j$ with $\mathcal{N}(i)$ the graph neighbours.

Closed-Form Metrics for GHZ_3 *Under Pauli Channels*

Consider $\rho_{\mathrm{GHZ}_3} = |\mathrm{GHZ}_3\rangle\langle \mathrm{GHZ}_3|$ with $|\mathrm{GHZ}_3\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$. Applying independent bit-flip channels $\mathcal{X}_p$ on each qubit, $\mathcal{X}_p(\rho) = (1-p)\rho + pX\rho X$, and letting $\mathcal{E} = \mathcal{X}_p^{\otimes 3}$, $\rho' = \mathcal{E}(\rho_{\mathrm{GHZ}_3})$, we obtain

$$\begin{aligned} F\big(\rho_{\mathrm{GHZ}_3}, \rho'\big) &= \langle \mathrm{GHZ}_3|\rho'|\mathrm{GHZ}_3\rangle \\ &= \frac{1}{2}\big[(1-p)^3 + p^3\big] + \frac{3}{2}\, p(1-p)^2 \\ &= \frac{1}{2} - \frac{3}{2}p^2 + \frac{3}{2}p^3 = \frac{1}{2} - \frac{3}{2}p^2(1-p). \end{aligned} \tag{A19}$$

Similarly, for independent phase-flip channels $\mathcal{Z}_p(\rho) = (1-p)\rho + pZ\rho Z$, the off-diagonal terms are suppressed giving

$$F\big(\rho_{\mathrm{GHZ}_3}, \rho''\big) = \frac{1}{2}\big[1 + (1-2p)^3\big]. \tag{A20}$$

These closed forms were used to validate the simulator's fidelity under Pauli noise.

Entropy Convention

Throughout, von Neumann entropy is computed as $S(\rho) = -\mathrm{Tr}(\rho \log_2 \rho)$, that is, using base-2 logarithms, so that entropy is reported in bits.

Reproducibility Settings

Unless stated otherwise, simulations use $n = 3$ qubits for GHZ/cluster sanity checks and $n \in \{3, 5, 7\}$ in scalability sweeps. Circuit depth is limited to at most 12 gates per qubit, with 10^4 measurement shots for `qasm` backends and fixed pseudo-random seeds (e.g., 42) for state initialisation and noise sampling. For composite channels, the channels are applied in the order of listing (e.g., bit-flip followed by amplitude damping), consistent with Equation (19).

Derivations for Quantum Information Metrics

1. *Fidelity*: For density operators ρ, ρ':

$$F(\rho, \rho') = \left(\mathrm{Tr}\sqrt{\sqrt{\rho}\rho'\sqrt{\rho}}\right)^2 \tag{A21}$$

If $\rho = |\psi\rangle\langle\psi|$ (pure state), this reduces to:

$$F(\rho, \rho') = \langle\psi|\rho'|\psi\rangle \tag{A22}$$

2. *von Neumann Entropy*

$$S(\rho) = -\mathrm{Tr}(\rho \log \rho) \tag{A23}$$

If $\rho = \sum_i \lambda_i |i\rangle\langle i|$ (spectral decomposition), then:

$$S(\rho) = -\sum_i \lambda_i \log \lambda_i \tag{A24}$$

3. *Trace distance*

$$D(\rho, \rho') = \frac{1}{2}\mathrm{Tr}|\rho - \rho'| \tag{A25}$$

Letting $\Delta = \rho - \rho'$, if $\{\delta_i\}$ are eigenvalues of Δ,

$$D(\rho, \rho') = \frac{1}{2}\sum_i |\delta_i| \tag{A26}$$

4. *Kraus representation for CPTP channels*: A quantum channel $\mathcal{E}$ acting on ρ:

$$\rho' = \mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger \tag{A27}$$

with $\sum_k E_k^\dagger E_k = I$. For amplitude damping:

$$E_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}, \quad E_1 = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix} \tag{A28}$$

5. *Composite reward and secure operating region*: The multiobjective RL reward:

$$R_t = -(\alpha(1 - F_t) + \beta S_t + \delta D_t) \tag{A29}$$

The secure region is as follows:

$$\mathcal{S}_{\text{secure}} = \{(F_t, S_t, D_t) \mid F_t \geq \tau_F,\ S_t \leq \tau_S,\ D_t \leq \tau_D\} \tag{A30}$$

Appendix B: Supplementary Simulation Figures

To provide further insight into the internal structure and quantum behaviour modelled in the CQDT framework, this appendix includes key visual outputs from the simulation environment. These figures support the main discussion by illustrating calibration metrics, quantum noise effects and the circuit-level design used during the experiments.

Figure A1

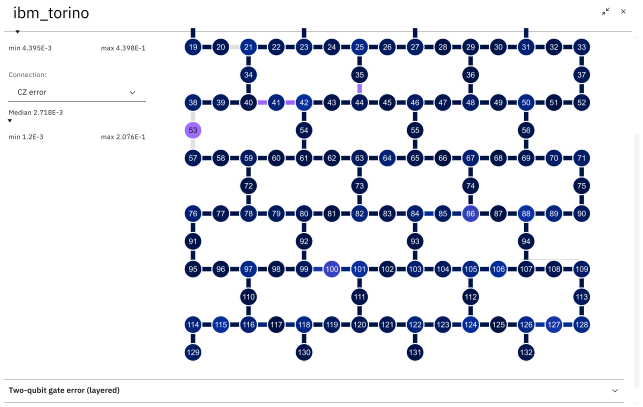

FIGURE A1 | IBM QPU calibration map illustrating controlled-Z (CZ) gate error rates and readout fidelities across physical qubits. These data are used to simulate realistic hardware-level noise in the CQDT pipeline.

[H]

Figure A2

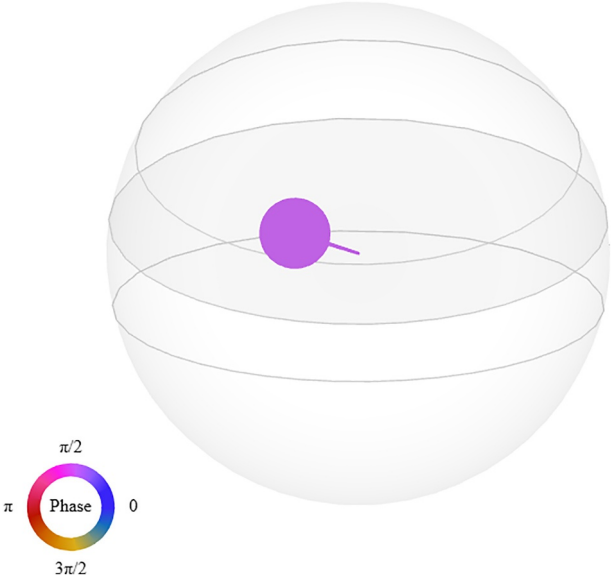

FIGURE A2 | Bloch sphere visualisation depicting the qubit state evolution under adversarial noise injection. It highlights coherence loss and phase deviation over time.

Figure A3

[H]

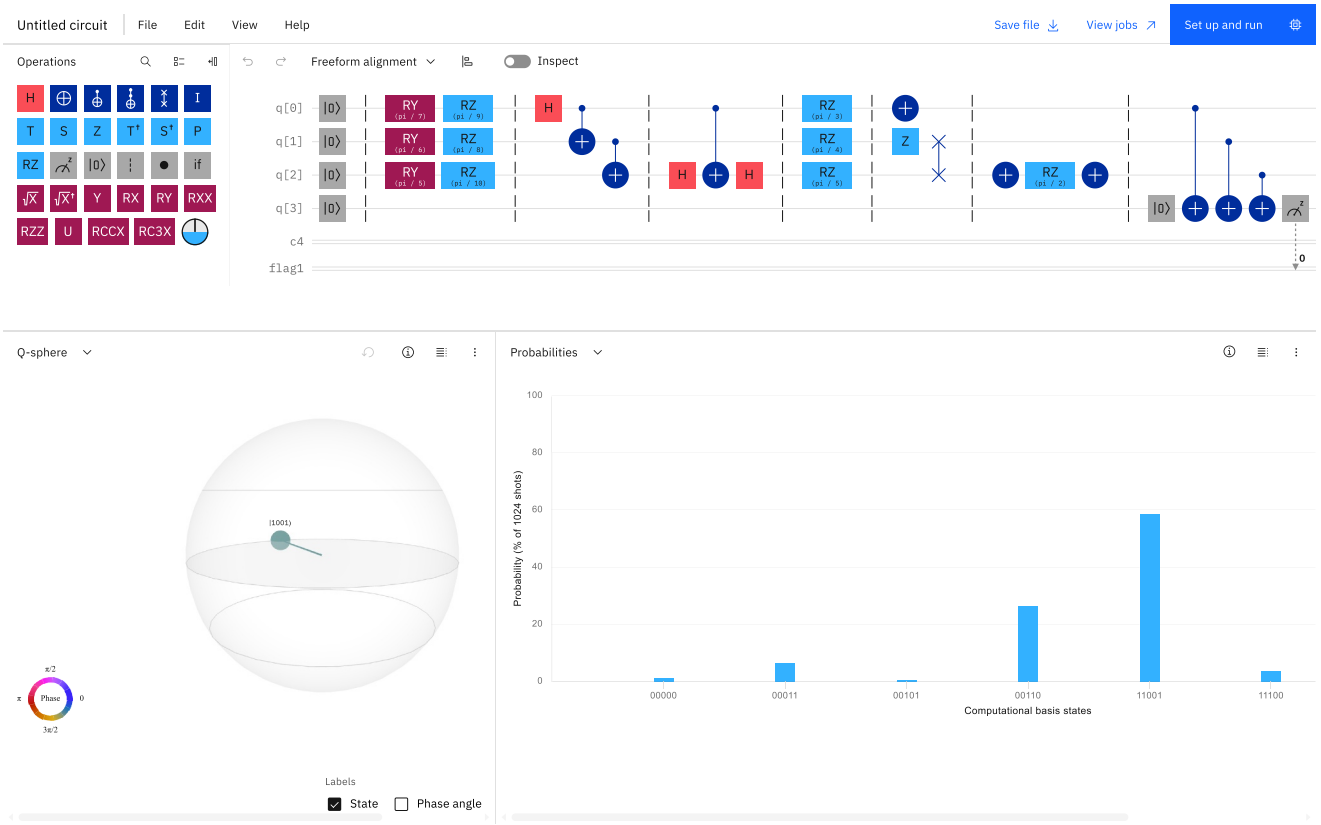

FIGURE A3 | Quantum circuit representation employed in the CQDT simulation pipeline, encompassing state preparation, multipartite entanglement and CPTP noise injection. This forms the structural basis for testing RL-based defence responses.