Cross-Layer Secure Sum-Rate Maximization in IRS–NOMA O-RAN

Cross-Layer Secure Sum-Rate Maximization in IRS–NOMA O-RAN

Muhammad J. Shehab, *Member, IEEE,* Ahmed Badawy, *Member, IEEE,*
Mohamed Elsayed, *Member, IEEE,* Tamer Khattab, *Senior Member, IEEE,*
Mahmoud Barhamgi, *Senior Member, IEEE,* Saeed Salem, *Senior Member, IEEE,*
Carla Fabiana Chiasserini, *Fellow, IEEE*

***Abstract*—The integration of intelligent reflecting surfaces (IRS) and non-orthogonal multiple access (NOMA) within the Open Radio Access Network (O-RAN) offers significant opportunities for 6G Internet of Things (IoT) systems but also raises new challenges in security, reliability, and resource efficiency. In this paper, we propose a cross-layer secure sum-rate maximization framework that jointly addresses physical-layer secrecy and network-layer packet loss in IRS–NOMA O-RAN environments. We analytically derive *closed-form expressions* for secrecy rates and packet loss under a cascaded Rician fading model, and formulate a secrecy sum-rate optimization problem that accounts for IRS phase shifts, NOMA power allocation, and energy-harvesting constraints. The resulting problem is NP-hard due to non-convex coupling across layers. To overcome this, we develop SecureO-RAN-SAC, a deep reinforcement learning algorithm based on Soft Actor-Critic v2, which learns near-optimal policies in real time. Simulation results demonstrate that SecureO-RAN-SAC achieves *comparable or superior performance to grid-based search (GBS)* while requiring only $\sim$10% of its computational cost for a 64-element IRS. These findings highlight the scalability and efficiency of our approach, establishing a new paradigm for secure, resource-aware, and ML-driven cross-layer optimization in O-RAN-enabled IoT networks.**

***Index Terms*—Secure NOMA, Packet Loss, 6G, IRS, O-RAN, Machine Learning, DRL, Soft Actor-Critic v2.**

I. Introduction

The landscape of communication networks is on the cusp of change driven by two key forces. The first is the arising of Open Radio Access Network (O-RAN), a paradigm shift in network architecture and a software-defined approach to RANs, leveraging enhanced intelligence, cost efficiency, and optimized performance. At its core, O-RAN adopts a disaggregated architecture, separating the radio, distributed, and centralized units for greater flexibility. A key part of the O-RAN radio interface is the near-real-time (Near-RT) RAN intelligent controller (RIC), which can include intelligent network applications (xApps), i.e., machine learning (ML)-based applications that provide a range of functionalities enhancing network performance and intelligence [1], [2]. The second is the rapid growth of Internet of Things devices (IoTDs), along with the opportunities and challenges this brings, including the need for highly effective security protocols and efficient network solutions. IoT application scenarios such as smart healthcare, industrial automation, and critical infrastructure monitoring require enhanced security, stringent quality of service (QoS), and high reliability. To address these challenges and improve the performance of the IoTDs, there is a need to employ innovative technologies such as intelligent reflecting surfaces (IRS), non-orthogonal multiple access (NOMA), and energy harvesting (EH). IRS offers a solution by reconfiguring the wireless channel through phase tuning, extending coverage, optimizing network throughput, and enhancing energy efficiency. NOMA allows multiple IoTDs to share resources concurrently, increasing device capacity, reducing latency, and improving the QoS based on channel conditions. EH is a sustainable power source for IoTDs, reducing waste and battery costs [3]. Integrating these technologies and optimizing cross-layer design, including physical-layer security (PLS) and network-layer QoS, can address the constraints of traditional IoT architectures. This sets the stage for addressing the scenario of a secure IRS downlink NOMA system within the O-RAN architecture, where a radio unit transmits the signal to IoTDs, with a passive eavesdropper attempting to intercept the signal. We will consider a comprehensive cross-layer framework that aims to maximize the secure sum rate of the IoTDs at the physical-layer while transferring energy to the legitimate IoTDs and minimizing their packet loss at the network layer, thereby enhancing reliability.

A. Related Work

Several studies have investigated secure IRS–NOMA implementations, highlighting their ability to improve coverage, capacity, spectral efficiency, and connectivity [4]–[7].In particular, [4] explains the inherent limitations and security risks of power-domain NOMA and shows that IRS-enabled reconfigurable beamforming can reshape channel disparities and support secure transmission against both external and SIC-induced internal eavesdropping. Further, the works [5]–[7] explored secure IRS–NOMA scenarios with different objectives, ranging from maximizing sum-rate and improving secrecy performance to enhancing secrecy throughput and reducing eavesdropping efficiency. In contrast, [8] studied a non-cooperative NOMA scheme that applied artificial noise (AN) to secure

M. Shehab, A. Badawy, M. Barhamgi, and S. Salem are with the Department of Computer Science and Engineering, Qatar University, Doha, Qatar. M. Elsayed is with Electrical Engineering, Qatar University, Doha, Qatar. T. Khattab is with IIPL Lab, Electrical Engineering, Qatar University, Doha, Qatar. C. F. Chiasserini is with the Politecnico di Torino, Torino, Italy.
Research reported in this publication was supported by the Qatar Research Development and Innovation Council ARG01-0525-230339. The content is solely the responsibility of the authors and does not necessarily represent the official views of Qatar Research Development and Innovation Council.

downlink transmissions and transfer energy to legitimate users, though IRS was not included. These works [4]–[8] neither develop a cross-layer formulation that simultaneously captures physical and network-layer requirements, nor do they consider cascaded Rician fading. They focused only on the physical-layer IRS–NOMA or AN-based NOMA systems with direct links and small-scale optimization. However, we adopt a cross-layer O-RAN-compliant design that jointly considers physical-layer secrecy, network-layer packet loss, and energy-harvesting constraints. Furthermore, while prior studies rely on SDR-based alternating optimization or exhaustive grid-based search with rapidly increasing complexity, the proposed SecureO-RAN-SAC framework enables scalable optimization through a constraint-aware DRL approach, making it suitable for large-scale IRS deployments.

Moreover, research on O-RAN has examined various aspects of its architecture [9]–[13]. For instance, [9] addressed network slicing by proposing intelligent automated schemes with deep learning-based xApps. The authors in [10] implemented and assessed the open-RU architecture in noisy environments. Other works, such as [11], have investigated intelligent traffic steering, while [12] focused on adaptive slice allocation to meet the latency requirements of virtual reality (VR). In addition, [13] suggested ML-based methods for optimizing gNodeB handovers. However, these O-RAN studies did not incorporate IRS, NOMA, or cross-layer optimization.

Recent studies have also addressed cross-layer design in NOMA-enabled networks from different perspectives, focusing on the joint optimization of physical and network layers. For instance, cache-aided Fog-RANs integrate index coding with coded NOMA at the physical layer to achieve energy efficiency [14], while energy-efficient resource allocation in H-CRANs was studied in [15]. Interference management in PD-NOMA with D2D communications was proposed in [16], and hybrid precoding for energy-efficient NOMA-enabled C-RANs was explored in [17]. A layered approach combining packet-level fountain coding with channel coding and power allocation was presented in [18]. Long-term cross-layer optimization under imperfect CSI and dynamic traffic was developed in [19] using Lyapunov techniques. For video transmission, [20] introduced optimization and DL-based methods to jointly improve capacity and video quality by linking physical-layer metrics (outage, ergodic capacity) with application-layer metrics (PSNR) in 5G URLLC. Delay-sensitive communications were studied in [21], focusing on uplink latency through joint physical and network-layer scheduling. Furthermore, reinforcement learning has been proposed to enhance adaptive cross-layer security for 6G [22], while [23] connected physical-layer NOMA gains to user-centric QoE. Lastly, [24] designed a URLLC-oriented cross-layer scheduler that integrates queue-based scheduling with superposition coding for improved delay–power tradeoff.

Distinctiveness of the Proposed Approach. While existing works provide valuable foundations on secure IRS, NOMA, and cross-layer optimization, they do not address an O-RAN-compliant IRS–NOMA architecture that jointly accounts for physical-layer secrecy (secure sum-rate) and network-layer reliability (packet-loss). To fill this gap, we propose a unified cross-layer O-RAN framework that jointly optimizes the IRS phase shifts, NOMA power allocation, and energy-harvesting factors under coupled secrecy and reliability constraints. The proposed design is supported by closed-form analytical modeling and a scalable constraint-aware DRL solution, and it is further validated in large-scale deployments where exhaustive GBS becomes computationally infeasible.

B. Our Contribution

To address the challenges posed by the growing demands of IoT applications, this work leverages a novel cross-layer design that integrates IRS and NOMA within the O-RAN framework. This design bridges the physical and network layers, addressing the objective of securing communications from passive eavesdroppers, alongside network-layer enhancements that minimize packet loss. This cross-layer optimization establishes a cohesive framework capable of meeting the stringent requirements of modern communication systems, laying the foundation for our novel contributions.

Algorithmic novelty vs. application contribution: Our algorithmic contribution is a constraint-aware Soft Actor-Critic version 2 (SACv2)-based deep reinforcement learning (DRL) framework, named SecureO-RAN-SAC, that jointly optimizes IRS phase shifts, NOMA power allocation, and SWIPT power-splitting under coupled secrecy, packet-loss, and energy-harvesting constraints, to deal with the resulting non-convex NP-hard optimization problem. Our application-level contribution lies in integrating this learning-based solution within an O-RAN-compliant IRS-NOMA architecture for secure IoT communications, supported by closed-form analytical modeling and scalable performance evaluation.

Thus, our contributions can be summarized as follows:

- We derive closed-form expressions for the secrecy rate.
- Considering our IRS-NOMA scheme with a cascaded Rician channel, we derive a closed-form expression for the packet loss.
- We formulate a secrecy sum-rate maximization problem considering IRS phase shifts, NOMA power allocation, and EH factors while addressing constraints related to packet loss.
- As the problem turns out to be NP-hard, we introduce our scheme SecureO-RAN-SAC to find near-optimal parameters for secrecy sum-rate maximization.
- We conduct extensive simulations to derive near-optimal parameters using SecureO-RAN-SAC and evaluate its performance against a grid-based search (GBS) benchmark.
- We demonstrate that our proposed algorithm achieves performance comparable to or better than GBS while utilizing only a fraction of the computational resources.

C. Paper Organization

The subsequent sections of the paper are structured as follows. Section II explains the system model. The theoretical analysis of the proposed algorithm, including the derivations of secrecy rates, packet loss expressions, computational resources

analysis in O-RAN, optimization problem, and the proposed DRL scheme, is elaborated in Section III. Section IV demonstrates the numerical results, covering simulation parameters, O-RAN computational resources, performance metrics, and DRL outcomes. Section V concludes the paper.

II. System Model

A downlink IRS-NOMA scheme within the O-RAN architecture is represented in Figures 1, 2, and 3, where the radio unit (RU) with directional antennas is transmitting to two legitimate IoTDs using a non-cooperative downlink NOMA transmission scheme. The system includes a passive eavesdropper, denoted with eve e, which attempts to intercept the transmitted signal. In the considered indoor deployment, we assume that eve e is located outside the intended service area (e.g., behind walls), which naturally places it farther from the IRS than the legitimate far IoTD and is consistent with common indoor IoT threat models. Nevertheless, the proposed SecureO-RAN-SAC framework is not tied to this geometric assumption and can readily handle mobile eavesdropper locations through adaptive learning without modifying the algorithmic structure. Further, the signal from the RU is transmitted to the IRS, which consists of M reflecting semi-passive elements. The IRS adjusts the phase of the incoming signal and reflects it toward the legitimate IoTDs. In this scenario, the IRS is used for coverage extension; there exists no direct physical connection or link between the radio access point and the legitimate IoTDs; rather, the IRS acts as a semi-passive reflector. The legitimate IoTDs are divided into two clusters, namely, the near cluster and the far cluster, based on their respective distances from the IRS. In Fig.1, IoTD n is chosen from the near cluster, while IoTD f is selected from the far cluster. To ensure that the channel between the IRS and the eavesdropper, as well as the channel between the IRS and the legitimate IoTDs, are uncorrelated, the distances between IoTD n and e and between IoTD f and e are assumed to be longer than half a wavelength. This assumption is generally valid since IoT devices typically operate in frequency ranges where wavelengths are measured in centimeters. The distances between the IRS and IoTD n and between the IRS and IoTD f are denoted as d_n and d_f, respectively. The wireless channels are modeled using Rician fading [25]. Specifically, $\mathbf{h_t}\in\mathbb{C}^{1\times M}$ represents the transmitter channel between the RU and the IRS, $\mathbf{h}_n \in \mathbb{C}^{M\times 1}$ indicates the channel between the IRS and IoTD n, $\mathbf{h}_f \in \mathbb{C}^{M\times 1}$ represents the channel between the IRS and IoTD f, and $\mathbf{h_e} \in \mathbb{C}^{M\times 1}$ represents the channel between the IRS and the eavesdropper. It's important to note that the channel between the IRS and the eavesdropper follows a Rician distribution, indicating a worst-case scenario with a dominant Line-of-Sight (LOS) link in this particular channel. Furthermore, the eavesdropper lacks knowledge of the power allocation factors used by legitimate users; therefore, it is unable to decode the desired signals effectively.

$$\mathbf{h}_\varkappa = \sqrt{\frac{\kappa_\varkappa}{\kappa_\varkappa+1}}\overline{\mathbf{h}}_\varkappa + \sqrt{\frac{1}{\kappa_\varkappa+1}}\tilde{\mathbf{h}}_\varkappa, \quad \varkappa \in \{t,n,e,f\} \quad (1)$$

where $\varkappa\in\{t,n,e,f\}$ represents the RU, IoTD n, IoTD f and e, respectively. In our system model, the input bit stream is first

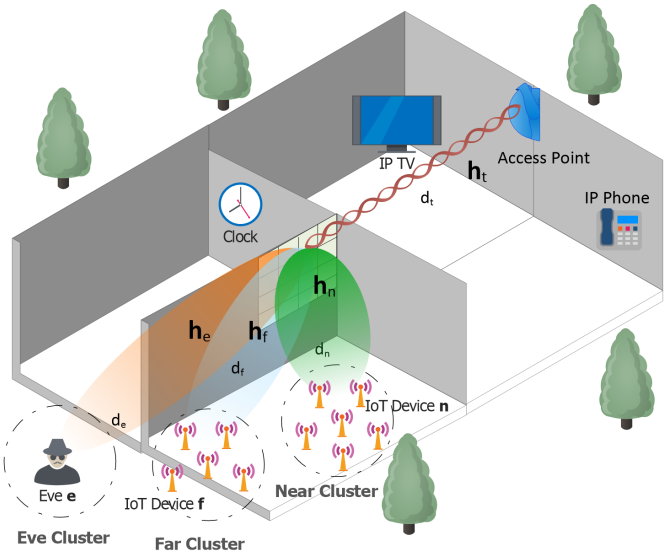

Figure 1. Downlink indoor coverage using IRS-NOMA network topology.

modulated employing 16-QAM modulation to enhance spectral efficiency and support the high data rate demands typical in 5G IoT scenarios. This modulation scheme balances spectral efficiency and power requirements, making it suitable for IoT devices operating under constrained energy and computational resources [26]. The signal transmitted from the RU point can be written as:

$$x = \sqrt{\alpha P_t}s_n + \sqrt{\overline{\alpha} P_t}s_f, \quad (2)$$

In the above expression, x is a scalar representing the transmitted signal for one sample, while s_n and s_f denote the signals intended for IoTDs n and f, respectively. Such signals have unit power, i.e., $\mathbb{E}[|s_n|^2]=1$ and $\mathbb{E}[|s_f|^2]=1$, where $\mathbb{E}[\cdot]$ denotes the expectation function. The power allocated to IoTD n is given by $P_n=\alpha P_t$, with α being the NOMA power allocation factor for IoTD n. Similarly, the power allocated to IoTD f is $P_f=\overline{\alpha}P_t$, where $\overline{\alpha}=1-\alpha$ is the power allocation factor for IoTD f. Here, $0\leq\alpha\leq 1$ represents the power allocation factor, and P_t corresponds to the total transmit power at the RU. The received signals for IoTDs n and f in the frequency domain can be expressed as:

$$\nu_w = c_w c_t \mathbf{h}_w^H \mathbf{\Phi_M} \mathbf{h}_t^H x + n_w, \quad w \in \{n,f\} \quad (3)$$

In (3), $c_t=(1+d_t^{\Xi_t})^{-\frac{1}{2}}$, $c_n=(1+d_n^{\Xi_n})^{-\frac{1}{2}}$, and $c_f=(1+d_f^{\Xi_f})^{-\frac{1}{2}}$ are channel coefficients that account for the path loss between the RU, IRS, and the IoTDs [27]. The parameters d_t, d_n, and d_f represent the distances between the radio access point and the IRS, the IRS and IoTD n, and the IRS and IoTD f, respectively. The symbols Ξ_t, Ξ_n, and Ξ_f designate the path loss exponents. The phase-shift matrix $\mathbf{\Phi}_M \triangleq \text{diag}(\phi_1,\phi_2,\ldots,\phi_M)$, where $\phi_m \triangleq e^{j\theta_m}$, satisfies the unit-modulus constraint $|\phi_m|^2 = 1$ for all $m \in \{1,\ldots,M\}$ since the IRS reflects the signal without amplification. The terms $n_n \sim \mathcal{CN}(0,\sigma_n^2)$ and $n_f \sim \mathcal{CN}(0,\sigma_f^2)$ denote the additive white Gaussian noise (AWGN) at IoTDs n and f, respectively. In the considered scenario, the legitimate IoTDs

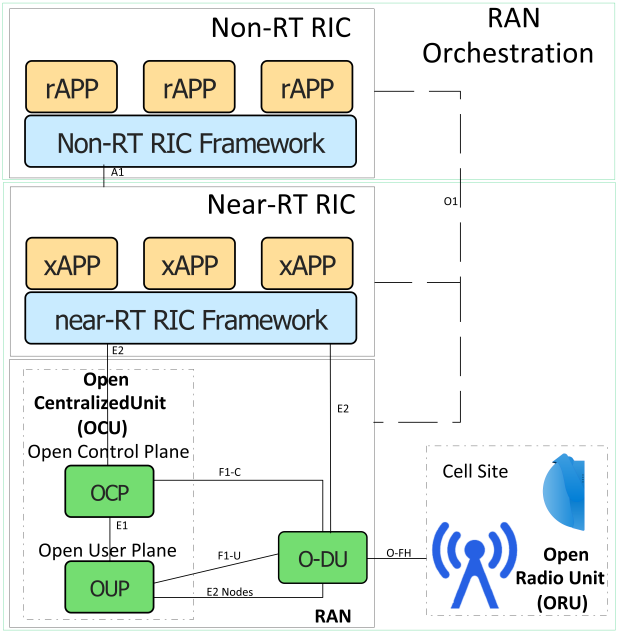

Figure 2. O-RAN Architecture.

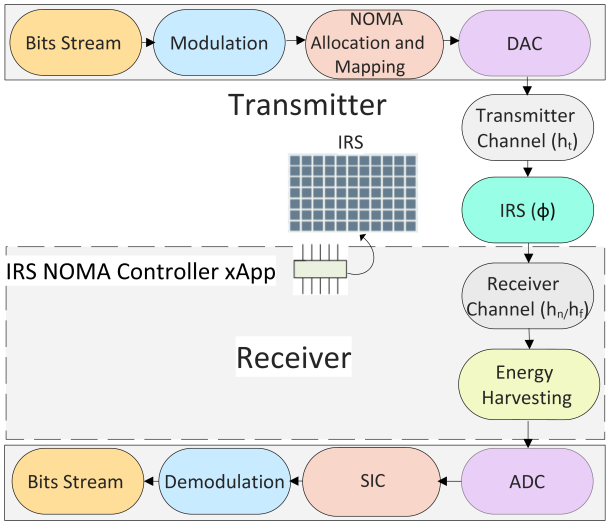

Figure 3. Diagram of an IRS-NOMA system.

are EH nodes. A power-splitting EH scheme is applied, where a portion of the received signal is used for EH, and the remaining portion is used for information decoding. The received signals after power splitting can be written as:

$$y_w = c_w c_t \sqrt{\beta_w} \mathbf{h}_w^H \mathbf{\Phi_M} \mathbf{h}_t^H x + n_w \quad w \in \{n, f\}, \tag{4}$$

Here, β_w represents the power splitting factors for IoTD w. This factor determines how the incoming signal power is divided between EH and information decoding. The term $\sqrt{\beta_w}$ scales the received signal accordingly. The signals received through the EH circuitry at both IoTDs n and f can be expressed as:

$$\upsilon_n = c_n c_t \sqrt{\overline{\beta}_n} \mathbf{h}_n^H \mathbf{\Phi_M} \mathbf{h}_t^H (\sqrt{\alpha P_t} s_n + \sqrt{\overline{\alpha} P_t} s_f), \tag{5}$$

$$\upsilon_f = c_f c_t \sqrt{\overline{\beta}_f} \mathbf{h}_f^H \mathbf{\Phi_M} \mathbf{h}_t^H (\sqrt{\alpha P_t} s_n + \sqrt{\overline{\alpha} P_t} s_f), \tag{6}$$

Here, $\overline{\beta}_n{=}1{-}\beta_n$, and $\overline{\beta}_f{=}1{-}\beta_f$. The harvested energy E_n and E_f at IoTDs n and f, respectively, can be obtained by taking the average harvested power $P_{H,w}$ at user $w \in \{n, f\}$, which is the squared magnitude of the received signals and considering the RF–to–DC efficiency factor η for RF energy conversion multiplied by the EH window T_H.

$$E_n = P_{H,n} T_H = \eta c_n^2 c_t^2 \overline{\beta}_n \left| \mathbf{h}_n^H \mathbf{\Phi_M} \mathbf{h}_t^H (\sqrt{\alpha P_t} s_n + \sqrt{\overline{\alpha} P_t} s_f) \right|^2 T_H, \tag{7}$$

$$E_f = P_{H,f} T_H = \eta c_f^2 c_t^2 \overline{\beta}_f \left| \mathbf{h}_f^H \mathbf{\Phi_M} \mathbf{h}_t^H (\sqrt{\alpha P_t} s_n + \sqrt{\overline{\alpha} P_t} s_f) \right|^2 T_H, \tag{8}$$

In the above expressions, $0{\leq}\eta{\leq}1$ denotes the efficiency factor of the RF energy conversion operation at the energy harvester circuit, i.e., it indicates the efficiency of converting the received RF signal into usable energy.

III. Theoretical Analysis of the Proposed Algorithm

In this section, we present the analytical analysis underlying the proposed optimization algorithm. Specifically, we derive the expressions for key performance metrics, including secrecy rates, packet loss probabilities, and harvested energy. These metrics serve as critical components in formulating the optimization problem addressed in the following section, where we propose a DRL algorithm to achieve near-optimal solutions efficiently.

A. Secrecy Rate Derivation

In the above non-cooperative downlink NOMA scheme, IoTD f, which is the device located farther from the IRS, decodes its signal in the presence of interference from the nearby IoTD n. IoTD f treats the signal coming from IoTD n as noise without removing it from the received signal. On the other hand, IoTD n performs successive interference cancellation (SIC) [1] by removing the message intended for IoTD f and subsequently decoding its signal. Consequently, the received SINR at IoTDs n and f can be expressed as follows:

$$\gamma_n = \left(\frac{c_n^2 c_t^2 \beta_n |\mathbf{h}_n^H \mathbf{\Phi_M} \mathbf{h}_t^H|^2 \alpha P_t}{\sigma_n^2} \right), \tag{9}$$

$$\gamma_f = \left(\frac{c_f^2 c_t^2 \beta_f |\mathbf{h}_f^H \mathbf{\Phi_M} \mathbf{h}_t^H|^2 \overline{\alpha} P_t}{c_f^2 c_t^2 \beta_f |\mathbf{h}_f^H \mathbf{\Phi_M} \mathbf{h}_t^H|^2 \alpha P_t + \sigma_f^2} \right), \tag{10}$$

[1] We assume perfect SIC. We follow the standard NOMA assumption that the strong user performs perfect SIC when decoding the weak user's signal. This assumption is widely adopted in IRS–NOMA secrecy studies and enables a clean comparison of power allocation and IRS control strategies. We note, however, that imperfect SIC would introduce residual interference and reduce the secrecy rate; this is discussed as an important extension in Sec. V.

The data rates for IoTDs n and f can be written as:

$$R_w = \log_2(1+\gamma_w), \quad w \in \{n, f\} \tag{11}$$

Further, the achievable secrecy rate is the rate at which the legitimate devices n and f can securely communicate in the presence of e. Therefore, the achievable secrecy rate at device n can be defined as the discrepancy between the rate of the main communication channel and the maximum rate achievable by eavesdropper e. This metric quantifies the level of secure communication for the legitimate device by considering the potential information leakage to e. If e is interested in decoding either n's or f's message, then it will consider the signal that it is not interested in as interference. Hence, if e is interested in decoding n's message only, then e's received SINR is given by:

$$\gamma_{e,n} = \left(\frac{c_e^2 c_t^2 |\mathbf{h}_e^H \mathbf{\Phi}_\mathbf{M} \mathbf{h}_t^H|^2 \alpha P_t}{c_e^2 c_t^2 |\mathbf{h}_e^H \mathbf{\Phi}_\mathbf{M} \mathbf{h}_t^H|^2 \overline{\alpha} P_t + \sigma_e^2} \right), \tag{12}$$

The variable c_e represents the path loss coefficient between the IRS and e, and it is computed as $c_e = (1 + d_e^{\Xi_e})^{-\frac{1}{2}}$. Here, d_e denotes the distance between the IRS and e, and it is required that $d_e > d_f$ to ensure that e is farther away from the IRS than IoTD f. Ξ_e indicates the path loss exponent[2]; finally, the term σ_e^2 denotes the variance of the AWGN at e. If e is interested in decoding f's message only, then e's received SINR can be written as:

$$\gamma_{e,f} = \left(\frac{c_e^2 c_t^2 |\mathbf{h}_e^H \mathbf{\Phi}_\mathbf{M} \mathbf{h}_t^H|^2 \overline{\alpha} P_t}{c_e^2 c_t^2 |\mathbf{h}_e^H \mathbf{\Phi}_\mathbf{M} \mathbf{h}_t^H|^2 \alpha P_t + \sigma_e^2} \right), \tag{13}$$

Instead, if e is interested in decoding only n's or f's messages, then the rate for e can be given as:

$$R_{e,w} = \log_2(1+\gamma_{e,w}), \quad w \in \{n, f\} \tag{14}$$

The secrecy rates, $R_{s,n}$ and $R_{s,f}$, for IoTDs n and f, respectively can then be expressed as:

$$R_{s,w} = \left[R_w - R_{e,w} \right]^+, \qquad w \in \{n, f\} \tag{15}$$

In (15), $[.]^+$ represents the positive part function, and it ensures that the per-user secrecy rate is non-negative. The total secrecy rate, R_s, is calculated by adding the secrecy rates for IoTDs n and f, and it is upper bounded by the following expression [5]:

$$R_s = R_{s,n} + R_{s,f} = \Big[R_n + R_f - (R_{e,n} + R_{e,f}) \Big]^+, \tag{16}$$

Both rates, $R_{s,n}$ and $R_{s,f}$, need to satisfy the constraints given in (15). Additionally, the total secrecy rate, R_s, should satisfy the upper bound constraint given in (16).

B. Packet Loss

To address the cross-layer aspect in our design, we aim to optimize the network-layer packet loss while enhancing the physical-layer security, as discussed earlier. Specifically, in what follows, we derive a closed-form expression for the packet loss probability for both the near and far devices, under a cascaded Rician fading channel model, thereby enabling analytical insights into our proposed IRS-NOMA cross-layer design. We adopt the two-state Gilbert-Elliott Markov model to model the packet loss in our wireless communication system. The model comprises two states: the bad state, which represents a high packet loss probability, and the good state, which represents a low packet loss probability. The steady-state probabilities of the channel being in the bad or good state are given by [28]:

$$\pi_{bad} = \frac{P}{1-\Omega+P} \qquad \pi_{good} = \frac{1-\Omega}{1-\Omega+P} \tag{17}$$

where $0 \leq P \leq 1$ denotes the probability of transitioning from the good to the bad state, and $0 \leq \Omega \leq 1$ denotes the probability of the channel remaining in the bad state. The values of Ω and P can be obtained by analyzing the packet loss statistics. We employ the K-State Markov chain model described in [29] with M-QAM modulation, where K denotes the number of states. In particular, we focus on a special case where $K = 2$, representing the good state and the bad state. The average packet loss probability is expressed as [30]:

$$P_{l_w} = e_{0,w}\ \pi_{bad} + e_{1,w}\ \pi_{good}, \quad w \in \{n, f\}. \tag{18}$$

where $e_{0,w}$ denotes the probability of packet error when the wireless channel is in a bad state, while $e_{1,w}$ represents the probability of packet error when the wireless channel is in a good state. The packet error probabilities can be calculated as:

$$e_{k,w} = 1 - (1 - b_{k,w})^L, \quad k \in \{0, 1\}, \quad w \in \{n, f\}, \tag{19}$$

where L refers to the packet size in bits, and $b_{k,w}$ is the average bit error probability for a specific state k (e.g., "good" or "bad"), and a specific IoTD w, over a defined SINR range $[A_k, A_{k+1}]$, and it is defined as:

$$b_{k,w} = \frac{\int_{A_k}^{A_{k+1}} f^{\text{cascaded}}(\gamma_w) P_b(\gamma_w)\, d\gamma_w}{\int_{A_k}^{A_{k+1}} f^{\text{cascaded}}(\gamma_w)\, d\gamma_w}, \tag{20}$$

where A_k and A_{k+1} are the SINR thresholds for state k. $A_0, \ldots, A_K$ represent the partition of the SINR range, where $0 = A_0 < A_1 < \ldots < A_K = \infty$. Each A_k specifies the threshold for the corresponding SINR range. In our case, $K=2$, and the thresholds $A_0=0$, A_1 are the pre-selected SINR thresholds that separate good and bad channel conditions, and $A_2=\infty$. Since we require bit error probability, we take γ_w in $P_b(\gamma_w)$ to be the per-bit SINR. Thus, $P_b(\gamma_w)$ represents the instantaneous BEP for 16-QAM, and is given by [31]:

$$P_b(\gamma_w) \approx \frac{3}{8} \operatorname{erfc}\Big(\sqrt{0.4\,\gamma_w}\Big), \tag{21}$$

Here, erfc(.) is the complementary error function, and γ_w represents the per-bit SINR for IoTD w. Note that the γ_n and γ_f for our IRS-NOMA system are given in (9) and (10), respectively. Further, $f^{\text{cascaded}}(\gamma_w)$ is the probability

[2]Practically, for the eavesdropper to be hidden from legitimate users, it has to be outside the considered indoor scenario

density function (PDF) of the cascaded Rician channel, and is expressed as [32]:

$$f^{\text{cascaded}}(\gamma_w) = \frac{1}{2}\sum_{\zeta_1=0}^{\infty}\sum_{\zeta_2=0}^{\infty}\mathcal{C}_2\,\gamma_w^{2\zeta_1+1}\left(\frac{1}{2\delta_2^2}\right)^{\zeta_1-\zeta_2} \times G_{0,2}^{2,0}\left(\frac{\gamma_w^2}{4\delta_1^2\delta_2^2}\,\middle|\,\zeta_2-\zeta_1,\,0\right), \tag{22}$$

where ζ_1 and ζ_2 are non-negative integer summation indices representing the orders of the series expansion for the first and second Rician fading channels in the cascaded model, respectively, and

$$\mathcal{C}_2 = \frac{1}{\zeta_1!\,\delta_1^2}\,\frac{1}{\zeta_2!\,\delta_2^2}\exp\!\left(\frac{-v_1^2}{2\delta_1^2}\right)\left(\frac{v_1^2}{2\delta_1^2}\right)^{2\zeta_1}\exp\!\left(\frac{-v_2^2}{2\delta_2^2}\right)\left(\frac{v_2^2}{2\delta_2^2}\right)^{2\zeta_2} \tag{23}$$

The Rician K-factors are given by $\kappa_1 = \frac{v_1^2}{2\delta_1^2}$ and $\kappa_2 = \frac{v_2^2}{2\delta_2^2}$, where v_1 and v_2 denote the amplitudes of the LOS components of the two channels, and δ_1 and δ_2 denote the standard deviations of the scattered (non-LOS) components. Moreover, $G_{0,2}^{2,0}(\cdot)$ denotes the Meijer G-function. Thus, the cumulative distribution function (CDF) of the cascaded Rician channel $F^{\text{cascaded}}(\gamma_w)$ is expressed as in [32].

$$F^{\text{cascaded}}(\gamma_w) = \frac{1}{4}\sum_{\zeta_1=0}^{\infty}\sum_{\zeta_2=0}^{\infty}\mathcal{C}_2\,\gamma_w^{2(\zeta_1+1)}\left(\frac{1}{2\delta_2^2}\right)^{\zeta_1-\zeta_2} \times G_{1,3}^{1,3}\left(\frac{\gamma_w^2}{4\delta_1^2\delta_2^2}\,\middle|\,-\zeta_1\,;\,\zeta_2-\zeta_1,\,0,\,-\zeta_1-1\right), \tag{24}$$

Hence, $b_{k,w}$ can be expressed in terms of $F^{\text{cascaded}}(\gamma_w)$. The denominator of $b_{k,w}$ is given by:

$$\int_{A_k}^{A_{k+1}} f_\gamma^{\text{cascaded}}(\gamma_w)\,d\gamma_w = F_\gamma^{\text{cascaded}}(A_{k+1}) - F_\gamma^{\text{cascaded}}(A_k), \tag{25}$$

By substituting the expression for the denominator into the definition of $b_{k,w}$, we get the expression for $b_{k,w}$ in terms of the CDF:

$$b_{k,w} = \frac{\int_{A_k}^{A_{k+1}} f_\gamma^{\text{cascaded}}(\gamma_w)\,P_b(\gamma_w)\,d\gamma_w}{F_\gamma^{\text{cascaded}}(A_{k+1}) - F_\gamma^{\text{cascaded}}(A_k)}, \tag{26}$$

By applying the transformation from Prudnikov, Brychkov, and Marichev (eq. (8.4.14.2)) [33], the erfc in terms of the Meijer G-function can be expressed as:

$$\text{erfc}(z) = \frac{1}{\sqrt{\pi}}\,G_{1,2}^{2,0}\left(z^2\middle|\begin{matrix}1\\0,\,\frac{1}{2}\end{matrix}\right) \qquad \text{substitute } z=\sqrt{0.4\,\gamma_w}$$

$$\frac{3}{8}\text{erfc}\left(\sqrt{0.4\,\gamma_w}\right) = \frac{3}{8\sqrt{\pi}}\,G_{1,2}^{2,0}\left(0.4\,\gamma_w\middle|\begin{matrix}1\\0,\,\frac{1}{2}\end{matrix}\right).$$

Thereby, $b_{k,w}$ can be written as:

$$b_{k,w} = \frac{\int_{A_k}^{A_{k+1}} \frac{3}{8\sqrt{\pi}} G_{1,2}^{2,0}\left(0.4\gamma_w\middle|\begin{matrix}1\\0,\,\frac{1}{2}\end{matrix}\right) f^{\text{cascaded}}(\gamma_w)\,d\gamma_w}{F_\gamma^{\text{cascaded}}(A_{k+1}) - F_\gamma^{\text{cascaded}}(A_k)}, \tag{27}$$

$$b_{k,w} = \frac{3}{16\sqrt{\pi}}\sum_{\zeta_1=0}^{\infty}\sum_{\zeta_2=0}^{\infty}\mathcal{C}_2\left(\frac{1}{2\delta_2^2}\right)^{\zeta_1-\zeta_2}\int_{A_k}^{A_{k+1}}\gamma_w^{2\zeta_1+1} \times \frac{G_{1,2}^{2,0}\left(0.4\,\gamma_w\middle|\begin{matrix}1\\0,\,\frac{1}{2}\end{matrix}\right)\,G_{0,2}^{2,0}\left(\frac{\gamma_w^2}{4\prod_{i=1}^{2}\delta_i^2}\,\middle|\,\zeta_2-\zeta_1,\,0\right)}{F_\gamma^{\text{cascaded}}(A_{k+1}) - F_\gamma^{\text{cascaded}}(A_k)}\,d\gamma_w, \tag{28}$$

$$\cdot\; b_{k,w} = \frac{3}{16\sqrt{\pi}}\,\frac{\sum_{\zeta_1=0}^{\infty}\sum_{\zeta_2=0}^{\infty}\mathcal{C}_2\left(\frac{1}{2\delta_2^2}\right)^{\zeta_1-\zeta_2}\left[\mathcal{J}(A_{k+1}^2)-\mathcal{J}(A_k^2)\right]}{F_\gamma^{\text{cascaded}}(A_{k+1}) - F_\gamma^{\text{cascaded}}(A_k)}, \tag{29}$$

$$\mathcal{J}(\varphi) = \frac{1}{2}\varphi^{\zeta_1+1}\,G_{4,4}^{2,4}\left(\frac{16}{25}\,\delta_1^2\delta_2^2\,\varphi\,\middle|\begin{matrix}-\zeta_2,\,-\zeta_1,\,\frac{1}{2},\,1\\0,\,\frac{1}{2},\,\frac{1}{4},\,\frac{3}{4}\end{matrix}\right),$$

Here, φ denotes the argument of the function $\mathcal{J}(\cdot)$. In the derivation, φ is later replaced by the quantities A_k^2 or A_{k+1}^2. However, the exact expression of $b_{k,w}$ in (29) involves nested special functions and infinite series, which makes closed-form manipulation and subsequent optimization impractical. For analytical tractability, we approximate the distribution of the aggregate cascaded baseband channel for IoTD w by a complex Gaussian via the central limit theorem. When many independent cascaded Rician components with comparable statistics (e.g., similar K-factors) contribute to the sum, their superposition is well approximated as Gaussian; hence $\varrho_w \sim \mathcal{CN}(\mu_{g_w}, \sigma_{g_w}^2)$ [34].

$$\varrho_w \triangleq \sum_{m=1}^{M}\mathbf{H}_{w_m}^*\,\phi_m\,\mathbf{H}_{t_m}, \qquad w\in\{n,f\} \tag{30}$$

For each IRS element $m \in \{1,\ldots,M\}$, the two hops are independent: $\mathbf{H}_{t_m} \perp\!\!\!\perp \mathbf{H}_{w_m}$. Across different IRS elements, channel coefficients are mutually independent: for any $m \neq m'$, $(\mathbf{H}_{t_m}, \mathbf{H}_{w_m}) \perp\!\!\!\perp (\mathbf{H}_{t_{m'}}, \mathbf{H}_{w_{m'}})$. The mean μ_{g_w} of this approximate Gaussian distribution is

$$\begin{aligned}\mu_{g_w} &= \sum_{m=1}^{M}|\phi_m|\,\mathbb{E}\big[\mathbf{H}_{w_m}^*\,\mathbf{H}_{t_m}\big]\\ &= \sum_{m=1}^{M}|\phi_m|\,\mathbb{E}\big[\mathbf{H}_{w_m}^*\big]\;\mathbb{E}[\mathbf{H}_{t_m}]\\ &= \sum_{m=1}^{M}\mathbb{E}\big[\mathbf{H}_{w_m}^*\big]\;\mathbb{E}[\mathbf{H}_{t_m}], \quad w\in\{n,f\}\end{aligned} \tag{31}$$

Here, $\mathbf{H}_{t_m}$ and $\mathbf{H}_{w_m}$ denote the small-scale fading coefficients for the AP→IRS and IRS→IoTD-w hops, respectively, for the m^{th} IRS element, and ϕ_m is the m^{th} IRS phase shift ($|\phi_m| = 1$). Further, the variance of the resulting approximate Gaussian

distribution for IoTD w can be expressed as follows:

$$\begin{aligned}\sigma_{g_w}^2 &= \mathrm{Var}\Bigg(\sum_{m=1}^{M} \mathbf{H}_{w_m}^{*}\,\phi_m\,\mathbf{H}_{t_m}\Bigg), \qquad w \in \{n, f\}\\ &= \sum_{m=1}^{M}\Bigg(\mathbb{E}\Big[\,\big|\mathbf{H}_{w_m}^{*}\,\phi_m\,\mathbf{H}_{t_m}\big|^2\Big] - \big|\mathbb{E}\big[\mathbf{H}_{w_m}^{*}\,\phi_m\,\mathbf{H}_{t_m}\big]\big|^2\Bigg)\\ &= \sum_{m=1}^{M}\Bigg(|\phi_m|^2\,\mathbb{E}\big[|\mathbf{H}_{w_m}|^2\big]\,\mathbb{E}\big[|\mathbf{H}_{t_m}|^2\big]\\ &\qquad\qquad - |\phi_m|^2\big|\mathbb{E}[\mathbf{H}_{w_m}^{*}]\,\mathbb{E}[\mathbf{H}_{t_m}]\big|^2\Bigg)\\ &= \sum_{m=1}^{M}\Big(\mathbb{E}\big[|\mathbf{H}_{w_m}|^2\big]\,\mathbb{E}\big[|\mathbf{H}_{t_m}|^2\big] - \big|\mathbb{E}[\mathbf{H}_{w_m}]\,\mathbb{E}[\mathbf{H}_{t_m}]\big|^2\Big),\end{aligned} \tag{32}$$

Therefore, the PDF of the approximated Gaussian-distributed composite channel linking the access point to IoTD w can be expressed as [35]:

$$f(\gamma_w) = \frac{1}{\sqrt{2\pi\sigma_{g_w}^2}}\exp\left(-\frac{(\gamma_w-\mu_{g_w})^2}{2\sigma_{g_w}^2}\right), \quad w \in \{n, f\} \tag{33}$$

Furthermore, the CDF of the approximated Gaussian-distributed composite channel can be expressed as: [35]:

$$F(\gamma_w) = \frac{1}{2}\times\left(1+\mathrm{erf}\left(\frac{\gamma_w-\mu_{g_w}}{\sqrt{2}\times\sigma_{g_w}}\right)\right), \quad w \in \{n, f\} \tag{34}$$

The Gaussian approximation simplifies the analysis, making it computationally efficient while preserving accuracy in performance evaluation. Therefore,(29) can be expressed as below:

$$b_{k,w} = \frac{\int_{A_k}^{A_{k+1}} \frac{1}{\sqrt{2\pi\sigma_{g_w}^2}}\exp\Big(-\frac{(\gamma_w-\mu_{g_w})^2}{2\sigma_{g_w}^2}\Big)\cdot\frac{3}{8}\,\mathrm{erfc}\big(\sqrt{0.4\gamma_w}\big)\,d\gamma_w}{F_\gamma^{\text{cascaded}}(A_{k+1}) - F_\gamma^{\text{cascaded}}(A_k)}, \tag{35}$$

After calculating $b_{k,w}$ from (35), and $e_{k,w}$ in (19), the packet loss probability P_{l_w} can be computed as defined in (18). Using a nonnegative exponential surrogate for the erfc,

$$\mathrm{erfc}\big(\sqrt{0.4\,\gamma_w}\big) \;\approx\; \sum_{i=1}^{\mathrm{O}} c_i\,e^{-\lambda_i\gamma_w}, \qquad c_i \geq 0,\ \lambda_i > 0. \tag{36}$$

Here, O is the number of exponential terms. A commonly adopted two-term exponential approximation, known as the Chiani-type closed-form, is given by:

$$\begin{aligned}\mathrm{erfc}\big(\sqrt{0.4\,\gamma_w}\big) &\approx \tfrac{1}{6}e^{-0.4\,\gamma_w} + \tfrac{1}{2}e^{-\frac{4}{3}\,0.4\,\gamma_w}\\ &\Rightarrow\ \ c_1 = \tfrac{1}{6},\ \lambda_1 = \tfrac{2}{5}, \quad c_2 = \tfrac{1}{2},\ \lambda_2 = \tfrac{8}{15}.\end{aligned} \tag{37}$$

By substituting (36) into (35) we get the below equation:

$$\begin{aligned}b_{k,w} &\approx \frac{\frac{3}{8}\sum_{i=1}^{O} c_i \int_{A_k}^{A_{k+1}} \frac{1}{\sqrt{2\pi\sigma_{g_w}^2}}\exp\Big(-\frac{(\gamma_w-\mu_{g_w})^2}{2\sigma_{g_w}^2} - \lambda_i\gamma_w\Big)\,d\gamma_w}{F_\gamma^{\text{cascaded}}(A_{k+1}) - F_\gamma^{\text{cascaded}}(A_k)}\\ &= \frac{\frac{3}{8}\sum_{i=1}^{O} c_i \int_{A_k}^{A_{k+1}} \frac{1}{\sqrt{2\pi\sigma_{g_w}^2}} e^{-\frac{(\gamma_w-\mu_{g_w})^2}{2\sigma_{g_w}^2}-\lambda_i\gamma_w}\,d\gamma_w}{F_\gamma^{\text{cascaded}}(A_{k+1}) - F_\gamma^{\text{cascaded}}(A_k)}.\end{aligned} \tag{38}$$

By completing the square in the exponent for each λ_i we get:

$$\begin{aligned}-\frac{(\gamma_w-\mu_{g_w})^2}{2\sigma_{g_w}^2} - \lambda_i\gamma_w = &-\frac{\big(\gamma_w-(\mu_{g_w}-\lambda_i\sigma_{g_w}^2)\big)^2}{2\sigma_{g_w}^2}\\ &- \lambda_i\mu_{g_w} + \tfrac{1}{2}\lambda_i^2\sigma_{g_w}^2.\end{aligned} \tag{39}$$

so that

$$\begin{aligned}&\int_{A_k}^{A_{k+1}} \frac{\exp\Big(-\frac{(\gamma_w-\mu_{g_w})^2}{2\sigma_{g_w}^2} - \lambda_i\gamma_w\Big)}{\sqrt{2\pi}\,\sigma_{g_w}}\,d\gamma_w =\\ &\tfrac{1}{2}\exp\big(-\lambda_i\mu_{g_w} + \tfrac{1}{2}\lambda_i^2\sigma_{g_w}^2\big)\left[\mathrm{erf}\Big(\frac{A_{k+1}-\mu_{g_w}+\lambda_i\sigma_{g_w}^2}{\sqrt{2}\,\sigma_{g_w}}\Big) - \mathrm{erf}\Big(\frac{A_k-\mu_{g_w}+\lambda_i\sigma_{g_w}^2}{\sqrt{2}\,\sigma_{g_w}}\Big)\right].\end{aligned} \tag{40}$$

Therefore, a closed-form expression is

$$\begin{aligned}b_{k,w} \approx\ &\frac{3}{16\big[F_\gamma^{\text{cascaded}}(A_{k+1}) - F_\gamma^{\text{cascaded}}(A_k)\big]}\\ &\sum_{i=1}^{O} c_i\,\exp\Big(-\lambda_i\mu_{g_w} + \tfrac{1}{2}\lambda_i^2\sigma_{g_w}^2\Big)\\ &\times\left[\mathrm{erf}\Big(\frac{A_{k+1}-\mu_{g_w}+\lambda_i\sigma_{g_w}^2}{\sqrt{2}\,\sigma_{g_w}}\Big) - \mathrm{erf}\Big(\frac{A_k-\mu_{g_w}+\lambda_i\sigma_{g_w}^2}{\sqrt{2}\,\sigma_{g_w}}\Big)\right].\end{aligned} \tag{41}$$

O is chosen to be two, with the coefficients in (37), then (41) reduces to

$$\begin{aligned}b_{k,w} \approx\ &\frac{3}{16\big[F_\gamma^{\text{cascaded}}(A_{k+1}) - F_\gamma^{\text{cascaded}}(A_k)\big]}\\ &\Bigg\{\tfrac{1}{6}\,e^{-\frac{2}{5}\mu_{g_w}+\frac{1}{2}(\frac{2}{5})^2\sigma_{g_w}^2}\\ &\quad\times\left[\mathrm{erf}\Big(\frac{A_{k+1}-\mu_{g_w}+\frac{2}{5}\sigma_{g_w}^2}{\sqrt{2}\,\sigma_{g_w}}\Big) - \mathrm{erf}\Big(\frac{A_k-\mu_{g_w}+\frac{2}{5}\sigma_{g_w}^2}{\sqrt{2}\,\sigma_{g_w}}\Big)\right]\\ &\quad+\tfrac{1}{2}\,e^{-\frac{8}{15}\mu_{g_w}+\frac{1}{2}(\frac{8}{15})^2\sigma_{g_w}^2}\\ &\quad\times\left[\mathrm{erf}\Big(\frac{A_{k+1}-\mu_{g_w}+\frac{8}{15}\sigma_{g_w}^2}{\sqrt{2}\,\sigma_{g_w}}\Big) - \mathrm{erf}\Big(\frac{A_k-\mu_{g_w}+\frac{8}{15}\sigma_{g_w}^2}{\sqrt{2}\,\sigma_{g_w}}\Big)\right]\Bigg\}.\end{aligned} \tag{42}$$

C. Computational Resources Analysis in O-RAN

We now provide a detailed analysis and offer valuable insights into the computational requirements of SecureO-RAN-SAC within the O-RAN framework. The use of computational resources in O-RAN, especially for xApps, is critical due to the shared nature of the computational infrastructure. The computational resources in the O-RAN ecosystem are shared among different xApps, and each xApp receives a specific allocation of these resources [1]. Consequently, understanding the computational complexity of our proposed SecureO-RAN-SAC algorithm and GBS is vital in assessing their feasibility and performance in real-world O-RAN scenarios. Our analysis includes the computational requirements of both the learning phase, which depends on the actor and critic-networks, and the steady-state phase of SecureO-RAN-SAC. Thus, the total complexity of SecureO-RAN-SAC equals the summation of both the learning and steady-state phases. The total computational complexity consists of two parts: the learning phase and the steady-state phase.

$$C_{\text{SecureO-RAN-SAC}} = C_{\text{Learning}} + C_{\text{Steady-State}}\,. \tag{43}$$

The learning phase depends on the number of iterations, the batch size, and the FLOPs of the actor and the twin critics:

$$C_{\text{Learning}} = I \times B \times \left(C_{\text{Actor}} + 2\,C_{\text{Critic}}\right) \times G, \tag{44}$$

where I is the number of learning iterations, B the batch size, and G the number of gradient updates per iteration. The complexity of the stochastic actor network (Gaussian policy with reparameterization and tanh squashing) is:

$$C_{\text{Actor}} = S \times K_{i_a} \;+\; H \times K_{i_a} \times K_{j_a} \;+\; K_{j_a} \times (2A), \tag{45}$$

where S is the number of state features, A the number of action dimensions, K_{i_a} and K_{j_a} are the input/output neuron counts per hidden layer, and H the number of hidden layers. The factor $(2A)$ accounts for the fact that the actor outputs both the mean and the log-standard-deviation for each action dimension. For each critic, which estimates the Q-value, the FLOPs are:

$$C_{\text{Critic}} = (S + A) \times K_{i_c} \;+\; H \times K_{i_c} \times K_{j_c} \;+\; K_{j_c}, \tag{46}$$

where K_{i_c}, K_{j_c} are the input/output neuron counts of the critic layers, and the final term corresponds to the scalar Q-value output. Because SACv2 uses two critics, the cost is doubled. An additional minor overhead stems from the entropy temperature update, but this is a scalar parameter and is negligible compared to actor/critic backpropagation. Once trained, the online complexity comes mainly from policy inference:

$$C_{\text{Steady-State}} = T \times \Big(S \times K_{i_a} + H \times K_{i_a} \times K_{j_a} + K_{j_a} \times (2A)\Big). \tag{47}$$

where T denotes the number of decision steps in deployment. This forward-pass cost is lightweight and suitable for O-RAN Near-RT operation. In contrast, the complexity of the exhaustive GBS scheme is given by [36]:

$$C_{\text{GBS}} = O\Big(Q \times U \times \left(1 + \tfrac{1}{\Delta\beta_n}\right) \times \left(1 + \tfrac{1}{\Delta\beta_f}\right) \times \left(1 + \tfrac{1}{\Delta\alpha}\right) \times \left(1 + \lfloor \tfrac{2\pi}{\Delta\Phi} \rfloor\right)^{M}\Big). \tag{48}$$

where Q is the FLOPs required to compute secrecy rate, harvested energy, and packet loss; U the number of IoTDs; M the number of IRS elements; and $\Delta\Phi$, $\Delta\beta_n$, $\Delta\beta_f$, and $\Delta\alpha$ the discretization steps for phase shifts, power-splitting, and power allocation. It is obvious that C_{GBS} grows exponentially with the number of IRS elements M, while the complexity of SecureO-RAN-SAC scales only linearly with network and action dimensions, making it far more practical for RT O-RAN scenarios.

D. Optimization Problem

Our goal is to maximize the NOMA IoTDs n and f secrecy sum-rate, acting on the following decision variables: (i) the phase shift matrix Φ_M, (ii) the power splitting factors β_n and β_f, (iii) the NOMA power allocation parameter α, and (iv) the amount of computational resources, denoted by ρ, that are allocated for both the GBS and DRL, where ι serves as the threshold. First, the phase shift matrix should satisfy the unit modulus constraint, ensuring that $|\phi_m|^2 = 1$ for all elements of the matrix. The power splitting factors and the NOMA power allocation factors are between 0 and 1. Additionally, the energy harvested at IoTDs n and f should meet their energy-harvesting thresholds ψ_n and ψ_f. The packet loss at the network-layer should not exceed the specified thresholds ϵ_n and ϵ_f. Finally, we need to ensure that ρ does not exceed the available computational resources, denoted by ι.

$$\begin{aligned} &\max_{\Phi_M,\beta_n,\beta_f,\alpha} : \{\mathbb{E}\{R_s\}\}, \\ \text{s.t.} \quad & 0 \le \alpha \le 1, \quad \text{and} \quad \rho \le \iota, \\ & |\phi_m|^2 = 1, \forall m \in \{1,2,...,M\}, \\ & 0 \le \beta_n \le 1, \quad \text{and} \quad 0 \le \beta_f \le 1, \\ & E_n \ge \psi_n, \quad \text{and} \quad E_f \ge \psi_f, \\ & P_{l_n} \le \epsilon_n, \quad \text{and} \quad P_{l_f} \le \epsilon_f, \\ & R_{s,f} \ge R_{s,f,min}, \end{aligned} \tag{49}$$

Here, $R_{s,f,min}$ is the minimum secure rate threshold for the far user. The optimization problem is NP-hard due to the non-convexity caused by the constant modulus restrictions at the IRS [37], [38]. Further, (49) represents a long-term optimization problem because the decision variables, such as the phase shift matrix and power allocation factors, are optimized over extended time scales to adapt to changing conditions. While the packet loss constraints are hard limits, the optimization variables indirectly influence packet loss by improving SINR, power allocation, energy harvesting, and computational efficiency. This ensures reliable communication while maximizing the secrecy sum-rate.

E. Feasibility of the Constraint Set

Although the optimization problem in (49) is NP-hard, the feasible set of constraints is non-empty under practical

operating conditions. In particular, the packet loss constraints can be satisfied by ensuring a minimum received SINR through appropriate IRS phase alignment and minimum power allocation to the legitimate IoTDs. Similarly, the energy-harvesting constraints are met by allocating a sufficient harvesting portion via the power-splitting factors β_n and β_f, which can be adjusted independently of the secrecy objective. For the considered channel statistics, transmit power levels, and IRS phases, a feasible operating region exists in which secrecy rate, packet loss, and energy harvesting constraints are simultaneously satisfied. This is verified later through the numerical results in Section IV, where feasible solutions persist across different numbers of IRS elements and legitimate users. These results confirm that the proposed constraint set is practically achievable and well-suited for learning-based optimization.

F. SACv2 for Secure Cross–Layer IRS–NOMA

To solve (49) under continuous controls and stochastic channels, we adopt SACv2: an off–policy, maximum–entropy RL method that (i) handles continuous actions, (ii) learns a stochastic policy via the reparameterization trick with tanh squashing, (iii) uses twin critics to mitigate overestimation bias, and (iv) tunes the entropy temperature automatically [39].

1) DRL Agent: The SecureO-RAN-SAC agent in our system corresponds to the SecureO-RAN-SAC controller implemented as an xApp in the Near-RT RIC. It consists of a stochastic actor network, twin critic networks, and supporting modules such as the replay buffer and entropy tuner. The agent observes cross-layer system states and outputs optimized control actions $(\Phi_M, \beta_n, \beta_f, \alpha)$ to maximize the secrecy sum-rate under O-RAN constraints. The SecureO-RAN-SAC agent runs at the Near-RT RIC and interacts with the IRS–NOMA environment to maximize the secrecy sum-rate while meeting energy–harvesting and packet-loss constraints.

2) State space: At decision epoch τ, we expose a 9-D cross-layer state measured at the beginning of epoch τ and performance metrics produced by the action taken at epoch $(\tau-1)$:

$$\begin{aligned}\mathbf{s}^{(\tau)} = \big[\, &\gamma_n^{(\tau-1)},\ \gamma_f^{(\tau-1)},\ R_{s,n}^{(\tau-1)},\ R_{s,f}^{(\tau-1)},\ R_s^{(\tau-1)},\\ &E_n^{(\tau-1)},\ E_f^{(\tau-1)},\ P_{l_n}^{(\tau-1)},\ P_{l_f}^{(\tau-1)}\,\big].\end{aligned} \tag{50}$$

Here the quantities $\gamma_n, \gamma_f, R_{s,n}, R_{s,f}, R_s, E_n, E_f, P_{l_n}, P_{l_f}$ are outcomes of the previous control, hence the $(\tau-1)$ index.

3) Action space: The continuous action of the SecureO-RAN-SAC algorithm at time step (τ) represents the set of actions that the DRL agent can take. It consists of four components: $\alpha^{(\tau)}$, $\beta_n^{(\tau)}$, $\beta_f^{(\tau)}$, and $\mathbf{\Phi}_{\boldsymbol{M}}^{(\tau)}$, which correspond to the NOMA power allocation factor, the EH factor of IoTD n, the EH factor of IoTD f, and the IRS phase shift matrix, respectively. The action-space $\mathbf{a}^{(\tau)}$ is then defined as:

$$\mathbf{a}^{(\tau)} = [\alpha^{(\tau)}, \beta_n^{(\tau)}, \beta_f^{(\tau)}, \mathbf{\Phi}_{\boldsymbol{M}}^{(\tau)}], \tag{51}$$

4) Reward: We use a reward function centered on secrecy performance, where the primary objective is to maximize the secrecy sum-rate. Additional cross-layer requirements, such as energy harvesting and packet-loss constraints, are incorporated through soft and hard penalty terms:

$$\begin{aligned} r^{(\tau)} = &\underbrace{w_{rs}\big(R_{s,n}^{(\tau)} + R_{s,f}^{(\tau)}\big)}_{\text{maximize secrecy}} + \underbrace{w_{eh}\big(E_n^{(\tau)} + E_f^{(\tau)}\big)}_{\text{encourage EH}} \\ &- \underbrace{w_{pl}\big(P_{l_n}^{(\tau)} + P_{l_f}^{(\tau)}\big)}_{\text{discourage loss}} \\ &- \Big(\chi_1(\psi_n - E_n^{(\tau)})_+ + \chi_2(\psi_f - E_f^{(\tau)})_+ \\ &\quad + \chi_3(E_n^{(\tau)} - E_{n,\max})_+ + \chi_4(E_f^{(\tau)} - E_{f,\max})_+ \\ &\quad + \chi_5(P_{l_n}^{(\tau)} - \epsilon_n)_+ + \chi_6(P_{l_f}^{(\tau)} - \epsilon_f)_+ \\ &\quad + \chi_7(P_{l_f}^{(\tau)} - P_{l_f,\max})\Big). \end{aligned} \tag{52}$$

In (52), the reward is structured to jointly encourage secrecy maximization, energy harvesting, and reliability, while penalizing violations of system constraints. The coefficients w_{rs}, w_{eh}, and w_{pl} are design weights that balance the primary objectives: w_{rs} promotes higher secrecy rates, w_{eh} rewards harvested energy, and w_{pl} penalizes packet loss. The terms χ_i represent penalty multipliers that enforce constraint satisfaction. Specifically, χ_1 and χ_2 penalize energy values falling below the minimum harvesting thresholds ψ_n and ψ_f, while χ_3 and χ_4 penalize exceeding the maximum harvesting limits $E_{n,\max}$ and $E_{f,\max}$. The coefficients χ_5 and χ_6 impose soft penalties on packet losses that exceed the tolerances ϵ_n and ϵ_f, respectively, while χ_7 enforces a hard cap on the far-user packet loss through the parameter $P_{l_f,\max}$. Here $(x)_+ = \max(0, x)$ represents the hinge operator that activates and applies penalties only when there is a violation of a constraint. This structure ensures that the agent is guided by secrecy performance while remaining compliant with energy harvesting and reliability constraints.

5) Feasibility Handling During Training: The formulated optimization problem is NP-hard, so infeasible actions may arise during exploration. To address this issue during training, the continuous action output of the SecureO-RAN-SAC agent is first projected onto the admissible action space to satisfy the constraints, including the NOMA power-allocation factors, power-splitting limits, and the constant-modulus constraint of the IRS. The remaining inequality constraints, namely packet-loss and energy-harvesting requirements, are enforced through the constraint-aware reward function in (52), which includes both soft and hard penalty terms. Specifically, soft constraint violations incur weighted penalties proportional to their magnitudes, while hard constraint violations (e.g., packet loss exceeding predefined thresholds) are strongly penalized. This mechanism guides the policy learning process toward the feasible region without interrupting exploration.

Table I
LIST OF SYMBOLS USED IN SECUREO-RAN-SAC FRAMEWORK

Symbol	Description	Symbol	Description
System and Channel Model Parameters			
$\mathcal{U} = \{n, f\}$	Set of legitimate IoT users: near user n and far user f	e	Passive eavesdropper
M	Number of IRS reflecting elements	$\mathbf{h}_t$	Channel vector from RU to IRS
$\mathbf{h}_n, \mathbf{h}_f, \mathbf{h}_e$	Channel vectors from IRS to near IoTD n, far IoTD f, and eavesdropper e	$\mathbf{h}_\varkappa$	Channel vector for link $\varkappa \in \{t, n, f, e\}$
$\kappa_\varkappa$	Rician K-factor for link $\varkappa$	$\overline{\mathbf{h}}_\varkappa, \tilde{\mathbf{h}}_\varkappa$	LOS and scattered components of the Rician channel
$\mathbf{\Phi}_M$	IRS phase-shift matrix	ϕ_m	Phase shift of the m-th IRS element
d_t, d_n, d_f, d_e	Distances RU–IRS, IRS–IoTDn, IRS–IoTDf, IRS–Eve$e$	$\Xi_t, \Xi_n, \Xi_f, \Xi_e$	Path-loss exponents of RU–IRS, IRS–IoTDn, IRS–IoTDf, IRS–Eve$e$ links
c_t, c_n, c_f, c_e	Large-scale path-loss coefficients of the corresponding links	x	Superposed NOMA transmit signal from the RU
s_n, s_f	Information symbols for near and far IoTDs	P_t	Total transmit power at the RU
P_n, P_f	Transmit power allocated to IoTDs n and f	$\alpha, \overline{\alpha}$	NOMA power allocation factors for IoTDs n and f
$\beta_w, \overline{\beta}_w$	Power-splitting factor and its complement at IoTD $w \in \{n, f\}$	η	RF-to-DC energy conversion efficiency at IoTDs
T_H	EH window duration per slot	y_w, ν_w	Received baseband signals at IoTD w after/before power splitting
n_w	AWGN at node $w \in \{n, f, e\}$	σ_w^2	Noise variance at node $w \in \{n, f, e\}$
E_n, E_f	Harvested energy at near and far IoTDs per slot	γ_n, γ_f	SINR at IoTDs n and f
$\gamma_{e,n}, \gamma_{e,f}$	SINR at the eavesdropper when decoding n's or f's signal	R_w	Achievable rate of IoTD $w \in \{n, f\}$
$R_{e,w}$	Achievable rate at the eavesdropper when decoding the signal of w	$R_{s,w}$	Secrecy rate of IoTD $w \in \{n, f\}$
R_s	Total secrecy sum-rate of IoTDs n and f	$R_{s,n}, R_{s,f}$	Secrecy rates of near and far IoTDs
Statistical and Packet-Loss Parameters			
L	Packet length (bits)	P_{l_w}	Packet-loss probability at IoTD $w \in \{n, f\}$
P, Ω	Gilbert–Elliott transition parameters (good→bad, bad→bad)	$\pi_{\text{bad}}, \pi_{\text{good}}$	Steady-state probabilities of bad and good channel states
$e_{k,w}$	Packet-error probability in state $k \in \{0, 1\}$ for user w	$b_{k,w}$	Average bit-error probability in state $k \in \{0, 1\}$ for user w
A_k	SINR threshold delimiting state k in the Markov model	$P_b(\gamma_w)$	Instantaneous bit-error probability for 16-QAM at SINR γ_w
$f^{\text{cascaded}}(\gamma_w)$	PDF of the cascaded Rician SINR at IoTD w	$F^{\text{cascaded}}(\gamma_w)$	CDF of the cascaded Rician SINR at IoTD w
K_1, K_2	Rician K-factors of the two cascaded links	v_1, v_2	LOS component amplitudes in cascaded Rician links
δ_1, δ_2	Standard deviations of scattered components in the two links	ζ_1, ζ_2	Summation indices in cascaded Rician series representation
$\mathcal{C}_2$	Coefficient in cascaded Rician PDF/CDF expressions	φ	Argument of $\mathcal{J}(\cdot)$, later set to A_k^2 or A_{k+1}^2
ϱ_w	Composite baseband channel sum for IoTD w	$\mu_{g_w}, \sigma_{g_w}^2$	Mean and variance of Gaussian-approximated composite channel for IoTD w
$f(\gamma_w)$	Gaussian-approximated PDF of composite channel SINR	$F(\gamma_w)$	Gaussian-approximated CDF of composite channel SINR
O	Number of exponential terms in the erfc approximation	c_i, λ_i	Non-negative coefficients and decay rates
ψ_n, ψ_f	Minimum EH thresholds at IoTDs n and f	ϵ_n, ϵ_f	Maximum allowable packet-loss probabilities at IoTDs n and f
$P_{l_f,\max}$	Upper limit on far-user packet-loss probability	$R_{s,f,\min}$	Minimum secrecy-rate requirement of far IoTD f
DRL and O-RAN Parameters			
$\mathbf{s}^{(\tau)}$	State vector at decision epoch τ	$\mathbf{a}^{(\tau)}$	Action vector at decision epoch τ
$\Phi_M^{(\tau)}$	IRS phase-shift matrix chosen at epoch τ	$\alpha^{(\tau)}, \beta_n^{(\tau)}, \beta_f^{(\tau)}$	NOMA power allocation and EH factors at epoch τ
$r^{(\tau)}$	Immediate reward at epoch τ	γ	Discount factor in SecureO-RAN-SAC
$\pi_\theta(\mathbf{a} \mid \mathbf{s})$	Stochastic policy with parameters θ	Q_{ϕ_1}, Q_{ϕ_2}	Twin critic networks with parameters ϕ_1, ϕ_2
$Q_{\phi_1'}, Q_{\phi_2'}$	Target critic networks	Λ	Entropy temperature parameter (auto-tuned)
$\mathcal{H}_{\text{tgt}}$	Target policy entropy	I	Number of learning iterations
B	Mini-batch size per update	G	Gradient updates per iteration
T	Deployment decision steps (steady-state)	S	Dimension of the state space
A	Dimension of the action space	H	Number of hidden layers in actor/critic networks
K_{i_a}, K_{j_a}	Actor hidden-layer input/output neuron counts	K_{i_c}, K_{j_c}	Critic hidden-layer input/output neuron counts
C	Replay buffer capacity (maximum stored transitions)	τ	Polyak averaging factor for target-network updates
ρ	Computational resources allocated to SecureO-RAN-SAC/GBS	ι	Total available computational resources in O-RAN
$C_{\text{SecureO-RAN-SAC}}$	Total complexity of SecureO-RAN-SAC	C_{Learning}	Complexity of the learning (training) phase
$C_{\text{Steady-State}}$	Complexity of the steady-state (inference) phase	C_{Actor}	FLOPs per forward/backward pass of the actor
C_{Critic}	FLOPs per forward/backward pass of one critic	C_{GBS}	Computational complexity of the grid-based search
Q	FLOPs to evaluate secrecy rate, EH, and packet loss	U	Number of IoTDs in the system
$\Delta\alpha$	Discretization step for α in GBS	$\Delta\beta_n, \Delta\beta_f$	Discretization steps for β_n, β_f in GBS
$\Delta\Phi$	Discretization step for IRS phase shifts in GBS	λ	Learning rate

6) SACv2 Algorithm: We adopt SACv2 with automatic entropy tuning. The agent is initialized with a replay buffer, a squashed Gaussian policy $\pi_\theta(\mathbf{a} \mid \mathbf{s})$, twin critics Q_{ϕ_1}, Q_{ϕ_2}, and corresponding target networks. During the exploration phase, actions are sampled uniformly; afterwards, they are drawn from the reparameterized stochastic policy. Transitions $(\mathbf{s}_t, \mathbf{a}_t, r_t, \mathbf{s}_{t+1}, d_t)$ are stored off-policy and used to update the critics via the clipped double-Q Bellman loss. The policy is then improved under the maximum-entropy objective, while the temperature Λ is adjusted online toward a target entropy $\mathcal{H}_{\text{tgt}} = -\dim(\mathcal{A})$, where $\dim(\mathcal{A})$ denotes the action dimension, i.e., the number of continuous action variables, to balance exploration and exploitation. Finally, target critics are updated by Polyak averaging for stability. This procedure is repeated for T steps, yielding the trained policy θ^π, critics $(\theta^{Q_1}, \theta^{Q_2})$, and the learned temperature parameter.

IV. Numerical Results

We conduct simulations to investigate how optimization parameters affect key metrics in our IRS–NOMA system, presenting graphical representations to illustrate their impact.

A. Simulation Parameters

We vary the transmit power P_t between $-30\,\text{dBm}$ and $30\,\text{dBm}$, and set the number of IRS reflecting elements M to 64 and the number of IoTDs to 2. The channels between the radio access point and the IRS, the IRS and IoTD n, the IRS and IoTD f, and the IRS and IoTD e are modeled as Rician with factor $K_1 = K_2 = 10$. The path-loss exponents, denoted by Ξ_t, Ξ_n, Ξ_f, and Ξ_e, are all fixed at 2.2, and the noise variance is $\sigma^2 = 10^{-3}\,\text{mW}$. The default distances between the radio access point and the IRS, the IRS and IoTD n, the IRS and IoTD f, and the IRS and IoTD e are $8\,\text{m}$, $10\,\text{m}$, $12\,\text{m}$, and $20\,\text{m}$, respectively. The probabilities P and Ω are set to 0.1 and 0.4, respectively. The packet length L is 1024 bits. The parameters α, β_n, and β_f range between 0 and 1; the EH efficiency is $\eta = 0.6$, and the far-user secrecy-rate minimum is $R_{s,f,\min} = 0.1$ b/s/Hz. The energy-harvesting window duration is set to $T_H = 1$ (one slot), so harvested energies are reported per slot. Results are averaged over $1,000$ Monte Carlo runs per configuration. All simulations are executed on a server equipped with an AMD EPYC 7V13 64-Core Processor (2.44 GHz), and 220 GB RAM running Windows Server 2022 Datacenter Azure Edition (version 21H2).

B. SecureO-RAN-SAC Hyperparameters

We employ SecureO-RAN-SAC with entropy regularization, consisting of *(i)* a stochastic Gaussian policy (actor), *(ii)* two independent Q-value networks (critics), and *(iii)* Polyak-averaged target critics for stable updates. The state dimension is 9, and the action dimension is 4. The actor is a tanh-squashed Gaussian policy with two hidden layers of size $(256, 256)$ and ReLU activations; it outputs mean and log-standard-deviation, samples actions via the reparameterization trick, and applies the exact $\tanh$ log-probability correction. Each critic takes the concatenated state–action input and

Algorithm 1 SecureO-RAN-SAC for IRS-NOMA

Input: Learning rate λ, discount factor γ, batch size B, replay buffer capacity C, target update rate τ, maximum steps T, start-steps T_{start}, gradient updates per step G.
Output: Trained policy parameters θ^π, twin critics $\theta^{Q_1}, \theta^{Q_2}$, and entropy temperature Λ.
Initialize: Replay buffer $\mathcal{D}$ with capacity C; stochastic policy $\pi_{\theta^\pi}(\mathbf{a} \mid \mathbf{s})$ (Gaussian, squashed by $\tanh$); critics $Q_{\theta^{Q_1}}, Q_{\theta^{Q_2}}$; target critics $\bar{Q}_{\theta^{Q_1}} \leftarrow Q_{\theta^{Q_1}}$, $\bar{Q}_{\theta^{Q_2}} \leftarrow Q_{\theta^{Q_2}}$.
Set target entropy $\mathcal{H}_{\text{tgt}} = -|\mathcal{A}|$. Initialize $\log \Lambda$ (trainable) for automatic temperature tuning.
Reset environment; observe initial state $\mathbf{s}_0$.
for $t = 0, 1, \ldots, T-1$ **do**
  if $t < T_{\text{start}}$ **then**
    Sample $\mathbf{a}_t$ uniformly from the action bounds (pure exploration).
  else
    Sample noise $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$.
    Compute pre-squash action $\mathbf{x}_t = \mu_{\theta^\pi}(\mathbf{s}_t) + \sigma_{\theta^\pi}(\mathbf{s}_t) \odot \boldsymbol{\xi}$.
    Apply squashing: $\mathbf{a}_t = \tanh(\mathbf{x}_t)$. ▷ $\mathbf{a}_t = [\alpha^{(t)}, \beta_n^{(t)}, \beta_f^{(t)}, \mathbf{\Phi}_M^{(t)}]$ in our system
  Execute $\mathbf{a}_t$ in the IRS-NOMA O-RAN environment; observe reward r_t (from (52)), next state $\mathbf{s}_{t+1}$, and terminal flag $d_t \in \{0, 1\}$.
  Store $(\mathbf{s}_t, \mathbf{a}_t, r_t, \mathbf{s}_{t+1}, d_t)$ into $\mathcal{D}$.
  if $|\mathcal{D}| \geq B$ **then**
    for $g = 1$ **to** G **do**
      Sample a minibatch $\{(\mathbf{s}_i, \mathbf{a}_i, r_i, \mathbf{s}'_i, d_i)\}_{i=1}^B$ from $\mathcal{D}$.
      Critic target: sample $\mathbf{a}'_i \sim \pi_{\theta^\pi}(\cdot \mid \mathbf{s}'_i)$, compute $\log \pi_{\theta^\pi}(\mathbf{a}'_i \mid \mathbf{s}'_i)$, and

$$y_i = r_i + \gamma(1-d_i)\left[\min_{k=1,2} \bar{Q}_{\theta^{Q_k}}(\mathbf{s}'_i, \mathbf{a}'_i) - \Lambda \log \pi_{\theta^\pi}(\mathbf{a}'_i \mid \mathbf{s}'_i)\right].$$

      Critic update: for $k \in \{1, 2\}$, minimize

$$L_{Q_k} = \frac{1}{B}\sum_{i=1}^{B} \big(Q_{\theta^{Q_k}}(\mathbf{s}_i, \mathbf{a}_i) - y_i\big)^2.$$

      Policy update: sample $\mathbf{a}_i \sim \pi_{\theta^\pi}(\cdot \mid \mathbf{s}_i)$ and minimize

$$J_\pi = \frac{1}{B}\sum_{i=1}^{B}\Big[\Lambda \log \pi_{\theta^\pi}(\mathbf{a}_i \mid \mathbf{s}_i) - \min_{k=1,2} Q_{\theta^{Q_k}}(\mathbf{s}_i, \mathbf{a}_i)\Big].$$

      Temperature update (auto-tuning):

$$J(\Lambda) = \frac{1}{B}\sum_{i=1}^{B} \Lambda\Big(-\log \pi_{\theta^\pi}(\mathbf{a}_i \mid \mathbf{s}_i) - \mathcal{H}_{\text{tgt}}\Big),$$

      and update $\log \Lambda$ using $\nabla_{\log \Lambda} J(\Lambda)$.
      Soft target update: for $k \in \{1, 2\}$,

$$\bar{\theta}^{Q_k} \leftarrow \tau\, \theta^{Q_k} + (1-\tau)\, \bar{\theta}^{Q_k}.$$

  if $d_t = 1$ **then**
    Reset environment; observe new initial state $\mathbf{s}_{t+1}$.
End.

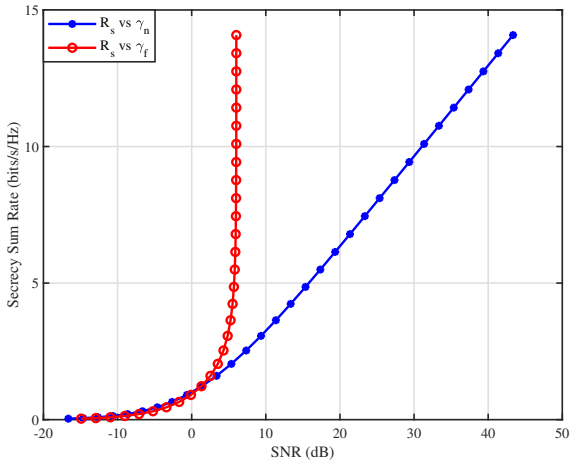

Figure 4. Secure sum-rate vs γ_n. $\beta_n = \beta_f = 0.5$, $\alpha = 0.2$.

uses two hidden layers $(256, 256)$ with ReLU activations, outputting a scalar Q-value. Two critics are trained in parallel, and the minimum value is used for policy improvement to reduce overestimation bias.

All networks are optimized with Adam at a learning rate of $3{\times}10^{-4}$, and gradient norms are clipped to 5. We set $\gamma = 0.99$, employ soft target updates with $\tau = 5{\times}10^{-3}$, a replay buffer capacity of $C = 10^6$, and mini-batch size $B = 256$. Further, the entropy temperature is initialized to $\alpha_0 = 0.2$ and tuned automatically during training to keep the policy entropy close to the target $\mathcal{H}_{\text{tgt}} = -\dim(\mathcal{A}) = -4$; thus, α becomes a learnable parameter that increases when the entropy falls below the target and decreases otherwise, typically remaining in the range $[0.01, 0.5]$ over training. The agent is trained for 1.6×10^5 environment steps, with 400 steps per episode, 7000 initial random exploration steps, and periodic evaluation runs to measure stability and convergence. A small sensitivity study with an alternative choices (learning rate in $[10^{-4}, 10^{-3}]$, $\tau \in \{10^{-3}, 5 \times 10^{-3}, 10^{-2}\}$, batch sizes 128-512, and different target entropy levels) showed that the selected parameters provide the best trade-off between convergence speed, numerical stability, and smooth secrecy-rate performance; other parameters either slowed convergence or led to higher variance in the learned secrecy sum-rate.

C. Performance Metrics Evaluation

To gain a deeper understanding of how different optimization parameters affect the system under study, we show their impact on the secure sum-rate, harvested energy, and packet loss rate (see Fig. 4–8). We emphasize that all figures in this subsection that use a fixed power allocation (e.g., $\alpha = 0.2$) are shown only as conventional baselines; the SecureO-RAN-SAC results reported in Sec. IV-F use parameters fully optimized by the DRL agent.

Secure Sum-rate. Fig. 4 depicts the secure sum-rate of IoTDs n and f, highlighting how the SINR of each IoTD affects this performance metric while maintaining constant values $\beta_n{=}\beta_f{=}0.5$ and $\alpha{=}0.2$. The secure sum-rate for IoTD n grows as the SINR increases, indicating that as the channel quality and received signal strength improve, IoTD n can achieve higher secure communication rates. On the other hand, the secure sum-rate for IoTD f initially increases with γ_f but saturates beyond 6 dB. This behavior is due to the interference caused by the nearby IoTD n in the downlink NOMA scheme: as the interference from IoTD n increases, it occupies a larger portion of the available channel capacity, leaving less capacity for IoTD f and ultimately degrading its secure rate. Hence, while the total secure sum-rate can still benefit from the strong link of IoTD n, the growing interference limits the achievable secrecy rate for IoTD f, highlighting the interference-management challenges in downlink NOMA systems.

Packet Loss Rate. To empirically validate the derived closed-form packet-loss expressions, we perform extensive Monte Carlo simulations under diverse channel conditions. Specifically, packet loss is evaluated using a time-domain simulation of a true two-state GE channel combined with instantaneous cascaded Rician fading across the IRS elements. For each SNR point, we generate a GE state sequence of length 2×10^5 samples and average the resulting packet loss outcomes over multiple Monte Carlo runs, yielding stable, low-variance estimates. Figures 5 and 6 compare the simulated curves with the analytical approximation and show a close match across a wide SNR range.

Fig. 5 shows the effect of the Rician factor K on the packet loss probability for IoTD n, with $K_1 = K_2 = K$. As the SNR increases, the packet-loss probability decreases, and a larger K further improves reliability due to the stronger LOS component. Fig. 6 reveals the impact of the GE transition probabilities (i.e., π_{bad} and π_{good}) while fixing $K_1 = K_2 = 10$; higher π_{bad} leads to increased packet loss due to longer occupancy of the bad state, whereas smaller π_{bad} improves reliability. Hence, the close alignment between analytical and simulated results in both figures provides empirical support that the Gaussian approximation in (33)–(34) accurately captures the effective behavior of the cascaded Rician link while substantially reducing computational complexity.

Harvested Energy. In Figures 7 and 8, we analyze the EH performance of the legitimate IoTDs for different values of the EH factors β_n and β_f, as a function of their respective SINRs (i.e., γ_n and γ_f). Such EH factors represent the power split ratios for IoTDs n and f, determining the portion of the received signal used for EH. As expected, as β_n and β_f decrease, the complementary power splitting factors, $\overline{\beta_n}$ and $\overline{\beta_f}$, increase. This implies that a larger part of the received signal is allocated for EH, whereas a smaller part is used for signal decoding. Consequently, the energy harvested by the IoTDs increases.

The results in Fig.7 demonstrate that the EH can increase by up to 9 orders of magnitude as β_n decreases from 0.9 to 0.1 (i.e., $\overline{\beta}_n$ increases from 0.1 to 0.9). This significant increase in EH is due to the higher power allocation for EH, allowing the IoTDs to capture more energy from the received signals. Furthermore, the results depicted in Fig. 8 reveal that, as the EH factor E_f increases, the SINR γ_f also grows until a certain point. Beyond that point, increasing E_f no longer leads

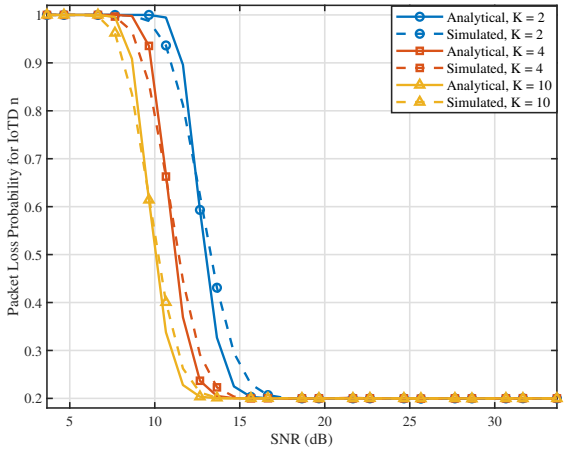

Figure 5. Simulated numerical vs. analytical approximation for various Rician factors K ($\pi_{bad} = 0.2$, $\pi_{good} = 0.8$).

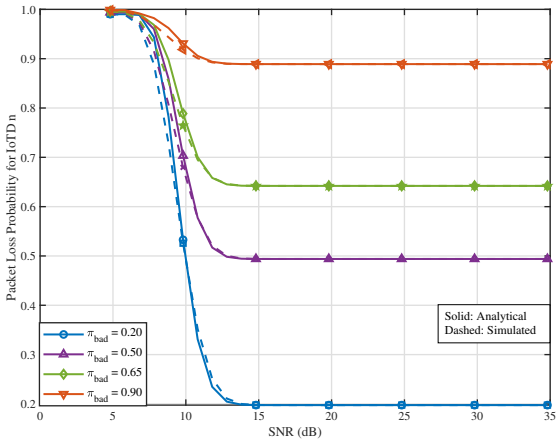

Figure 6. Simulated numerical vs. analytical approximation for various transition probabilities, ($K_1 = K_2 = 10$).

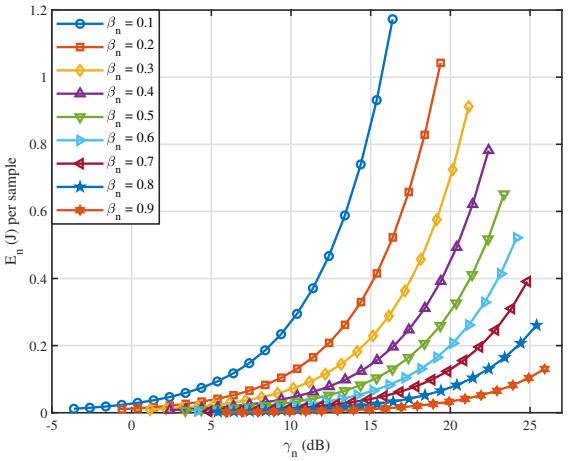

Figure 7. E_n vs. γ_n, for β_f=0.5, α=0.2, and $P_t \in [-10, 10]$ dBm.

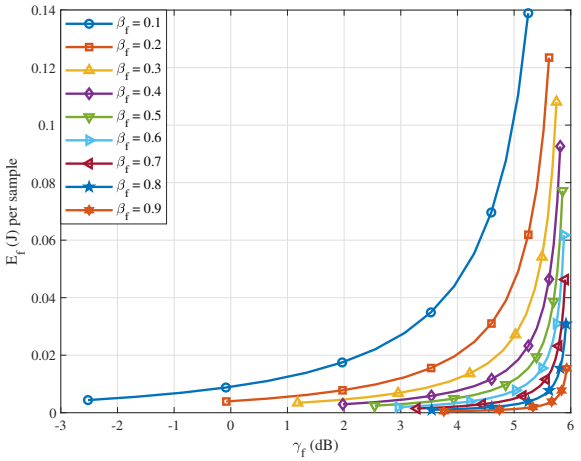

Figure 8. E_f vs. γ_f, for β_n=0.5, α=0.2, and $P_t \in [-10, 5]$ dBm.

to a further increase in γ_f. This phenomenon is, again, due to the interference caused by IoT device n on IoT device f. As E_f continues to increase, the interference from IoT device n becomes more significant, limiting the potential gains in SINR for IoT device f.

D. O-RAN's Computational Resources

Efficient utilization of computational resources is crucial for the scalability and performance of O-RAN systems. In this subsection, we compare the complexity of the proposed SecureO-RAN-SAC scheme with traditional approaches, highlighting its advantages in large-scale deployments of IRS. Fig. 9 illustrates the complexity and computational demands comparison between the suggested SecureO-RAN-SAC scheme and the traditional GBS approach. The figure shows that as the number of IRS elements increases, the DRL-based approach exhibits a near-linear growth in complexity, while the GBS approach experiences exponential growth. This highlights the scalability advantage of the DRL-based solution for large-scale IRS deployments.

Moreover, the behavior of GBS across varying IRS element counts M and phase search step ($\Delta\Phi$) is detailed in Fig. 10. Specifically, increasing the search step size leads to a noticeable reduction in GBS complexity, as fewer discrete phase values need to be evaluated per IRS element. Conversely, at finer step resolutions (i.e., more detailed search), the complexity of GBS increases exponentially. This can be attributed to the fact that a smaller search step demands a large number of search points to cover the total search space. Additionally, the figure illustrates that the complexity of GBS increases significantly with the increase in the number of IRS elements, as the search space grows exponentially with M. This highlights the computational limitations of the GBS approach for large-scale IRS systems. To ensure an equitable comparison, both techniques were evaluated under standardized complexity conditions. Furthermore, the complexity of the DRL algorithm as a function of the number of training

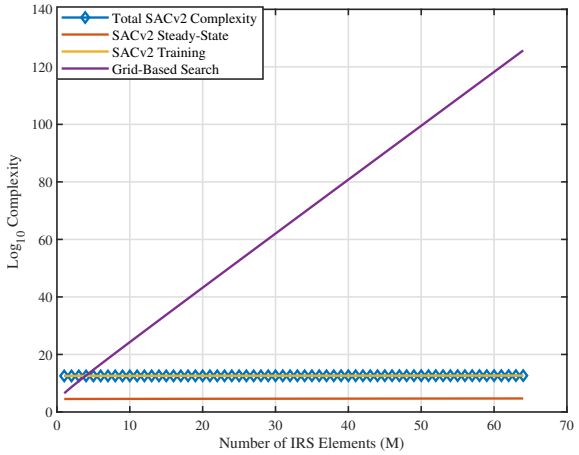

Figure 9. SecureO-RAN-SAC scheme's complexity vs. GBS.

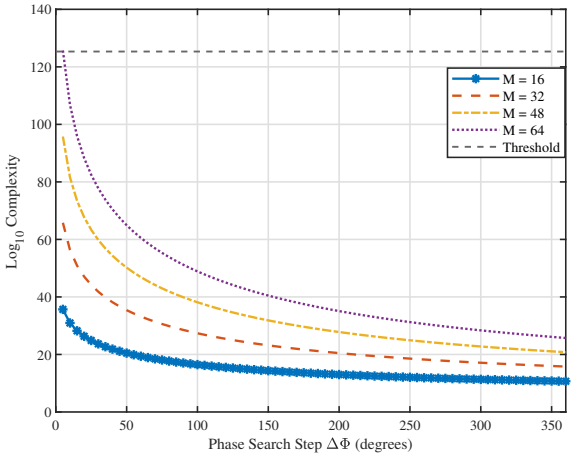

Figure 10. Complexity of GBS vs. IRS phase search step.

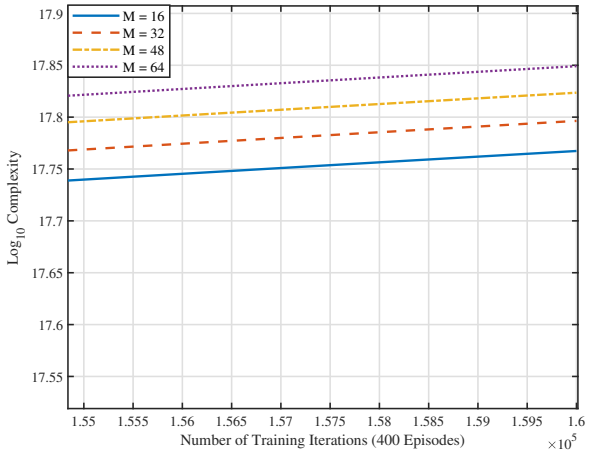

Figure 11. Complexity of SACv2 vs. number of episodes.

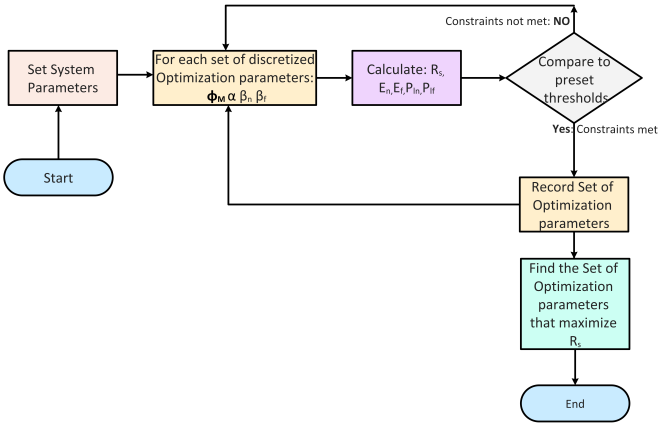

Figure 12. Flow chart of the grid-based methodology.

episodes for different IRS element counts M is shown in Fig. 11. The DRL complexity increases with the number of training episodes and IRS elements, reflecting its learning-based iterative nature. However, the DRL algorithm converges to a stable solution after a certain number of training episodes, demonstrating its efficiency in learning optimal policies for obtaining IRS phase shifts. For instance, under a logarithmic complexity threshold of 125.712 and an IRS with 64 elements (see Fig. 10), GBS can operate with a maximum resolution of $2\pi/72$ (i.e., $5°$), whereas DRL can process far more than 160000 iterations and 400 training episodes (Fig. 11) while consuming only a logarithmic complexity of 17.85. When scaling beyond 64 IRS elements, GBS requires larger search steps to remain within the same complexity threshold, which degrades performance and reduces the accuracy of the obtained solutions.

E. Grid-based Search Results

We use GBS as a benchmark for the SecureO-RAN-SAC performance. The GBS is employed in this work as a benchmark for small-scale scenarios to approximate the optimal solution. The discretization step sizes for the IRS phase shifts, power allocation, and energy-harvesting factors are selected to balance solution accuracy and computational feasibility. In particular, the step sizes are chosen sufficiently fine such that further refinement yields only marginal improvements in the achievable secrecy sum-rate, while incurring a prohibitive increase in computational complexity. This behavior ensures that the reported GBS results are representative of near-optimal performance. We emphasize that GBS is not intended as a scalable solution and is used only to provide a meaningful performance baseline for evaluating the proposed SecureO-RAN-SAC framework.

We thus address the optimization problem in (49) by using the approach depicted in Fig. 12. The process begins by configuring the system parameters, followed by establishing a search space and incremental steps for each optimization parameter. Specifically, the search space for Φ_M spans from 0 to $2\times\pi$, with an incremental step of $2\times\pi/72$, for α it ranges from 0 to 1 with increments of 0.02, and for β_n and β_n,

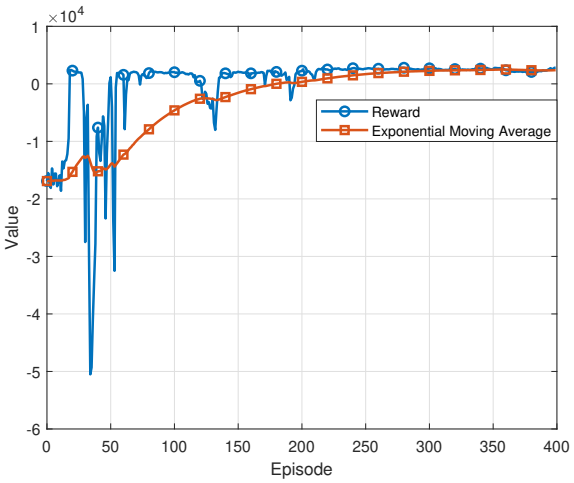

Figure 13. Convergence of the SecureO-RAN-SAC algorithm. Reward value vs episodes.

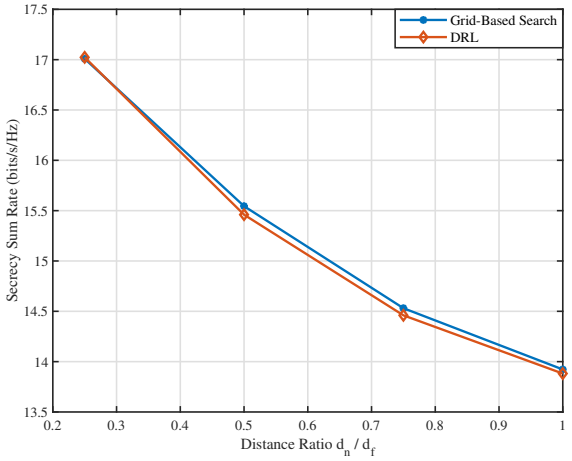

Figure 14. R_s for DRL versus GBS. d_t=2, $d_n \in [1,4]$ m, d_f=4 m, d_e=6 m.

the range is $[0,1]$ with increments of 0.05. The performance metrics, namely, R_s, E_n, E_f, P_{l_n}, and P_{l_f}, are computed for each set of discretized optimization parameters. Subsequently, such performance values are compared against predefined thresholds. Out of the sets of optimization parameters that satisfy all constraints, we select the one that provides the highest R_s, accompanied by its corresponding parameters E_n, E_f, P_{l_n}, and P_{l_f}. The GBS methodology systematically explores all feasible combinations of Φ_M, β_n, β_f, and α within the designated search space to identify the optimal solution. It is worth noting that, although the GBS guarantees the location of the optimal solution within the provided search space, its applicability to our system is limited due to its high complexity, especially as the number of IRS elements (M) or users increases.

F. DRL Results

The results obtained from the SecureO-RAN-SAC algorithm demonstrate the convergence of our scheme, as depicted in Fig.13. The plot shows that the rewards increase over time, indicating the successful learning process of the SecureO-RAN-SAC algorithm.

1) Comparison of SecureO-RAN-SAC Versus GBS: Fig.14 compares the secure sum-rates achieved using the SecureO-RAN-SAC algorithm and those obtained via GBS as a function of the distance ratio d_n/d_f. The two curves almost overlap, demonstrating that the proposed DRL method achieves performance very close to that of the exhaustive GBS benchmark. This confirms that DRL attains near-optimal secure sum-rates while offering a significant reduction in computational complexity. Further, this highlights the effectiveness of DRL in challenging scenarios with closely located users. It is worth noting that the SecureO-RAN-SAC consumes only 10.06% of the resources required by GBS for 64 IRS elements (see Fig.9), which highlights its efficiency in practical applications.

Moreover, Table II presents the optimized values for EH factors β_n and β_f, and power allocation factor α obtained using the SecureO-RAN-SAC algorithm. These values are

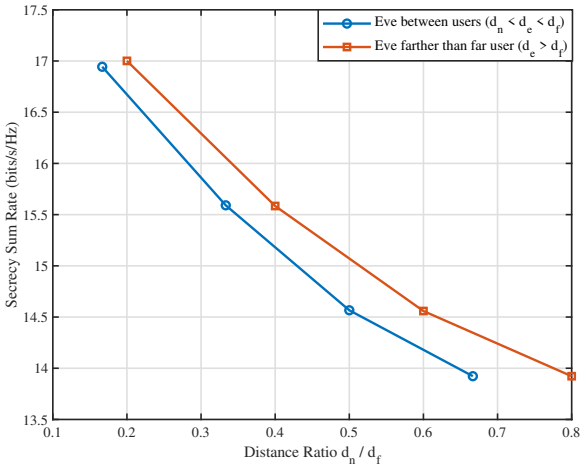

Figure 15. Secure sum-rate versus distance ratio d_n/d_f for different eavesdropper locations: Eve between the near and far users ($d_n < d_e < d_f$) and Eve beyond the far user ($d_e > d_f$).

derived for a specific combination of distances (d_t, d_n, d_f, and d_e) that maximizes the total secure rate for IoT devices while ensuring EH levels (E_n and E_f) above the thresholds for both near and far NOMA IoTDs, while also minimizing packet loss (P_{l_n}, and P_{l_f}) for both IoTDs. These values are obtained through the SecureO-RAN-SAC algorithm while satisfying the constraints defined in the optimization problem (49). The SecureO-RAN-SAC algorithm allows us to find the near-optimal solution by iteratively improving the values of the optimization parameters. By utilizing the rewards and feedback from the IoTDs, the algorithm adapts and adjusts the parameters $\mathbf{\Phi_M}$, β_n, β_f, and α to achieve the highest secure sum-rate, provide energy to the IoTDs, and minimize their packet loss. The optimized values obtained from the SecureO-RAN-SAC algorithm provide insights into the near-optimal configuration for maximizing the performance of the system in terms of secure communication and energy transfer.

Table II
OPTIMIZED VALUES FOR OPTIMIZATION PARAMETERS OBTAINED THROUGH SECUREO-RAN-SAC FOR P_t=0 dBm

d_t	d_n	d_f	d_e	α	β_n	β_f	R_s	E_n	E_f	P_{ℓ_n}	P_{ℓ_f}
2	1	4	6	0.476053	0.726873	0.654938	17.022998	119.677180	13.518936	0.000001	0.143316
2	1	6	7	0.451578	0.731842	0.584804	16.999733	117.499718	6.849316	0.000001	0.163402
2	2	5	7	0.490000	0.405747	0.502364	14.677743	93.081678	12.146267	0.002001	0.153277
2	2	6	8	0.478621	0.374108	0.521240	14.558575	98.037425	7.897908	0.002001	0.163308
3	2	5	7	0.490000	0.817715	0.502406	14.562742	13.081478	5.564396	0.002001	0.153277
3	4	7	9	0.490000	0.950000	0.503214	12.796558	0.907880	2.689537	0.006001	0.173277

Table III
OPTIMIZED VALUES FOR THE OPTIMIZATION PARAMETERS OBTAINED THROUGH GBS FOR P_t=0 dBm

d_t	d_n	d_f	d_e	α	β_n	β_f	R_s	E_n	E_f	P_{ℓ_n}	P_{ℓ_f}
2	1	4	6	0.490000	0.730000	0.500000	17.009121	118.306980	19.589163	0.000001	0.143277
2	1	6	7	0.490000	0.730000	0.500000	17.009107	118.306980	8.248291	0.000001	0.163277
2	2	5	7	0.490000	0.370000	0.500000	14.544691	98.680898	12.203959	0.002001	0.153277
2	2	6	8	0.490000	0.370000	0.500000	14.544711	98.680898	8.248291	0.002001	0.163277
3	2	5	7	0.490000	0.800000	0.500000	14.531145	14.352739	5.591302	0.002001	0.153277
3	4	7	9	0.490000	0.950000	0.500000	12.796558	0.907880	2.706938	0.006001	0.173277

Table III provides the optimized values of β_n, β_f, and α for various combinations of distances (d_t, d_n, d_f, and d_e) obtained through the discretized GBS scheme. These optimized values aim to maximize the total transmission sum-rate for legitimate IoT devices. The main objective of this table is to compare the secure sum-rates, EH levels, and packet loss rates achieved by the SecureO-RAN-SAC algorithm with those obtained through the ES method. This comparison highlights the effectiveness of the SecureO-RAN-SAC scheme, demonstrating that it achieves performance comparable to the GBS method while requiring significantly lower computational complexity. By employing both the SecureO-RAN-SAC technique and the GBS approach, we determine nearly optimal parameters that provide an approximation of the maximum secure sum-rate achievable subject to the constraints outlined in the optimization problem (49). The results and performance of the SecureO-RAN-SAC scheme demonstrate its usefulness in addressing the non-convex optimization problem. It effectively explores the continuous action space to deliver near-optimal performance with far lower complexity than the discretized GBS method. While we cannot guarantee global optimality, SecureO-RAN-SAC provides near-optimal solutions that significantly reduce computational complexity while maintaining performance levels comparable to the exhaustive GBS method.

2) Impact of Eavesdropper Location: The effect of the eavesdropper's spatial location on the secrecy performance is illustrated in Fig. 15, which depicts the secure sum-rate as a function of the distance ratio d_n/d_f under two representative eavesdropping scenarios. In this evaluation, the transmitter-to-IRS distance is fixed at $d_t = 2$, while the near-user distance d_n is gradually increased from 1 to 4, and the far-user distance d_f is kept constant at 6 for each scenario. Specifically, we consider the case where the eavesdropper is located between the near and far users ($d_n < d_e < d_f$) at $d_e = 5$, and the case where the eavesdropper is positioned farther than the far user ($d_e > d_f$) where $d_f = 5$t, and $d_e = 6$.

As shown in the figure, the secure sum-rate decreases monotonically with increasing d_n/d_f in both scenarios. This behavior is due to the reduction in channel disparity between the near and far NOMA users as d_n increases, which weakens the effectiveness of power-domain NOMA and reduces the achievable secrecy gain. Moreover, for a given distance ratio, the scenario where the eavesdropper is located beyond the far user consistently achieves a higher secure sum-rate compared to the case where the eavesdropper lies between the legitimate users. This can be attributed to the increased path loss experienced by the eavesdropper in the farther scenario, which significantly limits its interception capability. In contrast, placing the eavesdropper between the near and far users results in a stronger eavesdropping channel, leading to more secrecy degradation. These results confirm that the proposed SecureO-RAN-SAC framework can effectively adapt its resource allocation and IRS-assisted transmission strategy to mitigate eavesdropping threats under different distance scenarios.

3) Effect of Imperfect SIC on Secrecy Performance: To capture the effect of imperfect SIC, we adopt a residual-interference model that is widely used in the NOMA literature [37]. Specifically, after SIC at the near IoTD n, a fraction of the far-user signal remains uncancelled and appears as additional interference in the near-user decoding. Then, the resulting SINR at IoTD n under imperfect SIC is modeled as:

$$\gamma_n^{\text{imp}} = \frac{c_n^2 c_t^2 \beta_n \left|\mathbf{h}_n^H \mathbf{\Phi}_M \mathbf{h}_t^H\right|^2 \alpha P_t}{\sigma_n^2 + \varepsilon\, c_n^2 c_t^2 \beta_n \left|\mathbf{h}_n^H \mathbf{\Phi}_M \mathbf{h}_t^H\right|^2 \overline{\alpha} P_t}, \tag{53}$$

Here, $\varepsilon \in [0,1]$ denotes the SIC imperfection coefficient, where ε=0 corresponds to perfect SIC and ε>0 captures residual interference due to incomplete cancellation. Fig. 16 illustrates the secure sum-rate versus the distance ratio d_n/d_f for perfect SIC and representative imperfect-SIC levels (ε=10^{-6}, 10^{-5}, 10^{-4}). The results confirm that increasing ε consistently degrades the secure sum-rate across all distance

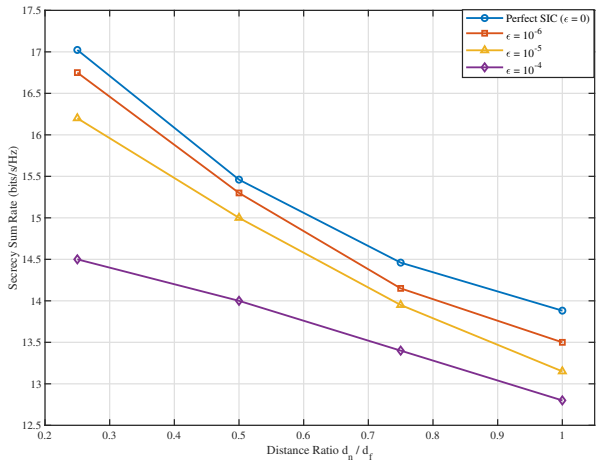

Figure 16. Secure sum rate versus distance ratio d_n/d_f under perfect and imperfect SIC. d_t=2 m, $d_n \in [1, 4]$ m, d_f=4 m, d_e=6 m.

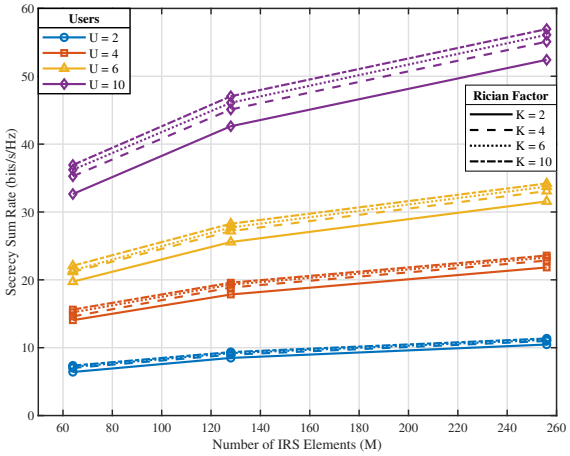

Figure 17. Secure Sum-Rate versus IRS elements for various users $\mathcal{U}$ and rician factors $\kappa_{\mathscr{H}}$.

ratios, since the residual term scales with the far-user received power and reduces the near-user decoding quality, which affects the overall secrecy performance. Further, all curves decrease monotonically with d_n/d_f, reflecting the reduced channel disparity between near and far users as their distances become comparable, which weakens the power-domain separability of NOMA and diminishes the secrecy gain. Also, the gap between perfect and imperfect SIC remains controlled for small residual levels, indicating that the proposed SecureO-RAN-SAC framework maintains robust secrecy performance under practical SIC imperfections.

4) Large-Scale Deployment: Fig. 17 evaluates the scalability of the proposed SecureO-RAN-SAC scheme in large-scale deployments where exhaustive GBS becomes computationally infeasible due to the exponential growth of the search space with the number of IRS elements and users. In this figure, the secure sum-rate performance is evaluated as a function of the number of IRS reflecting elements M (up to 256) for different numbers of legitimate NOMA IoTDs, namely $\mathcal{U} \in \{2, 4, 6, 10\}$. For all user configurations, the secure sum-rate increases monotonically with M, confirming that larger IRS deployments provide stronger gains, which enhance the effective legitimate channels and improve secrecy performance.

Moreover, increasing the number of users yields higher aggregate secure throughput at each M, demonstrating that SecureO-RAN-SAC can jointly optimize IRS phase shifts and resource allocation to efficiently accommodate more users while satisfying the system constraints.

In addition, Fig. 17 further illustrates the impact of different Rician factors $\kappa_{\mathscr{H}} \in \{2, 4, 6, 10\}$ on the secure sum-rate performance. As the Rician factor increases, the secure sum-rate consistently improves for all values of M and $\mathcal{U}$, owing to the stronger LOS components in the cascaded IRS-assisted channels. These results validate the effectiveness of the proposed SecureO-RAN-SAC framework and its robustness across a wide range of channel conditions, spanning moderate to strong LOS environments. Thus, it is capable of effectively exploiting favorable propagation characteristics in large-scale IRS-assisted O-RAN deployments.

5) Performance-Complexity Trade-off: Increasing the number of IRS reflecting elements introduces a fundamental performance-complexity trade-off in IRS–NOMA O-RAN systems. On the performance side, a larger number of IRS elements provides stronger gains, which enhance the effective legitimate channels and suppress information leakage toward the eavesdropper. As a result, higher secrecy rates, extended coverage, and improved energy efficiency can be achieved, as revealed by Fig. 17, where the secure sum-rate increases monotonically with the number of IRS elements for all considered numbers of IoTDs. However, these gains come at the cost of significantly increased optimization difficulty. Specifically, the dimensionality of the IRS phase-shift vector grows linearly with the number of elements, leading to an exponential growth of the search space for conventional optimization methods such as GBS, which quickly becomes computationally infeasible in large-scale deployments. Moreover, a larger number of IRS elements intensifies the interdependence among optimization variables, including IRS phase shifts, NOMA power allocation, and energy-harvesting factors, further complicating the optimization landscape. The proposed SecureO-RAN-SAC framework effectively addresses this trade-off by employing a learning-based strategy whose computational complexity scales gracefully with system dimensions. By exploiting long-term interactions with the environment, SecureO-RAN-SAC can learn near-optimal control policies without GBS, thereby enabling the system to harness the performance benefits of a larger number of IRS elements deployment while maintaining computational tractability in practical O-RAN-enabled IoT networks.

V. Conclusions

We focused on maximizing the secrecy rate of legitimate IoT devices while simultaneously enabling energy transfer to both near and far IoT nodes and improving overall reliability. To this end, IRS-assisted secure downlink NOMA was investigated within the O-RAN framework. The optimization problem aimed to determine near-optimal values for the NOMA

power allocation parameters, IRS phase shift matrix, EH factors, and packet-loss error rate. This problem is inherently non-convex, making it intractable with conventional mathematical optimization methods. To tackle this, the SecureO-RAN-SACv2 approach was employed, using a stochastic actor with twin critics and entropy tuning. The simulation results validate the effectiveness and applicability of SecureO-RAN-SAC in handling the complexity of the problem. Furthermore, the results show that for 64 IRS elements, the proposed DRL algorithm achieves performance comparable to the discretized GBS method while consuming only 10.06% of the computational resources, demonstrating its effectiveness in handling the difficult optimization task.

Several promising extensions can further improve the proposed SecureO-RAN-SAC scheme. First, future studies may consider more general eavesdropper locations, including mobile or randomly positioned adversaries, to capture realistic secrecy threats better. Second, including imperfect SIC and channel-state information uncertainty would provide a more robust evaluation of the system performance under practical constraints. Third, extending the analysis to multi-cell and multi-RU O-RAN deployments would enable the study of inter-cell interference, cooperative control, and distributed resource management. Finally, multi-agent DRL architectures could be explored in order to support a larger number of IoT devices and multiple coordinated RUs.

REFERENCES

[1] M. J. Shehab, Y. Aly, A. Badawy, A. Mohamed, M. Barhamgi, and S. Salem, "O-Cloud Security: A Comprehensive Survey of Threats, Mitigation Strategies, and Future Directions," in IEEE Open Journal of the Communications Society, doi: 10.1109/OJCOMS.2025.3600528.

[2] M. Polese, L. Bonati, S. D'Oro, S. Basagni and T. Melodia, "Understanding O-RAN: Architecture, Interfaces, Algorithms, Security, and Research Challenges," in IEEE Communications Surveys & Tutorials, vol. 25, no. 2, pp. 1376-1411, Secondquarter 2023, doi: 10.1109/COMST.2023.3239220.

[3] M. J. Shehab, I. Kassem, A. A. Kutty, M. Kucukvar, N. Onat and T. Khattab, "5G Networks Towards Smart and Sustainable Cities: A Review of Recent Developments, Applications and Future Perspectives," in IEEE Access, vol. 10, pp. 2987-3006, 2022, doi: 10.1109/ACCESS.2021.3139436.

[4] W. Wang et al., "Secure Beamforming for IRS-Enhanced NOMA Networks," in IEEE Wireless Communications, vol. 30, no. 1, pp. 134-140, February 2023, doi: 10.1109/MWC.012.2100639.

[5] Y. Qi and M. Vaezi, "IRS-Assisted Physical Layer Security in MIMO-NOMA Networks," in IEEE Communications Letters, vol. 27, no. 3, pp. 792-796, March 2023, doi: 10.1109/LCOMM.2023.3235722.

[6] H. Han, Y. Cao, M. Sheng, N. Zhao, J. Liu and D. Niyato, "IRS-Aided Secure NOMA Networks Against Internal and External Eavesdropping," in IEEE Transactions on Communications, vol. 70, no. 11, pp. 7536-7548, Nov. 2022, doi: 10.1109/TCOMM.2022.3208341.

[7] Z. Zhang, J. Chen, Q. Wu, Y. Liu, L. Lv and X. Su, "Securing NOMA Networks by Exploiting Intelligent Reflecting Surface," in IEEE Transactions on Communications, vol. 70, no. 2, pp. 1096-1111, Feb. 2022, doi: 10.1109/TCOMM.2021.3126636.

[8] A. Badawy and A. E. Shafie, "Securing OFDM-Based NOMA SWIPT Systems," in IEEE Transactions on Vehicular Technology, vol. 69, no. 10, pp. 12343-12347, Oct. 2020, doi: 10.1109/TVT.2020.3017570.

[9] S. -P. Yeh, S. Bhattacharya, R. Sharma and H. Moustafa, "Deep Learning for Intelligent and Automated Network Slicing in 5G Open RAN (O-RAN) Deployment," in IEEE Open Journal of the Communications Society, vol. 5, pp. 64-70, 2024, doi: 10.1109/OJCOMS.2023.3337854.

[10] S. J. Seelam, S. Andra and P. C. Jain, "Impact of Remote Radio Head on 5G Open-RAN Technology," 2022 8th International Conference on Signal Processing and Communication (ICSC), Noida, India, 2022, pp. 131-136, doi: 10.1109/ICSC56524.2022.10009237.

[11] F. Kavehmadavani, V. -D. Nguyen, T. X. Vu and S. Chatzinotas, "Intelligent Traffic Steering in Beyond 5G Open RAN Based on LSTM Traffic Prediction," in IEEE Transactions on Wireless Communications, vol. 22, no. 11, pp. 7727-7742, Nov. 2023, doi: 10.1109/TWC.2023.3254903.

[12] A. Casparsen, B. Soret, J. J. Nielsen and P. Popovski, "Near Real-Time Data-Driven Control of Virtual Reality Traffic in Open Radio Access Network," GLOBECOM 2023 - 2023 IEEE Global Communications Conference, Kuala Lumpur, Malaysia, 2023, pp. 3481-3486, doi: 10.1109/GLOBECOM54140.2023.10437947.

[13] H. Kumar, V. Sapru and S. K. Jaisawal, "O-RAN based proactive ANR optimization," 2020 IEEE Globecom Workshops (GC Wkshps, Taipei, Taiwan, 2020, pp. 1-4, doi: 10.1109/GCWkshps50303.2020.9367582.

[14] Y. Guo, C. W. Sung, S. Mostafa and J. Zou, "A Cross-Layer Optimization Framework for Index-Coded NOMA in Cache-Aided F-RANs," in IEEE Transactions on Communications, vol. 70, no. 11, pp. 7322-7336, Nov. 2022, doi: 10.1109/TCOMM.2022.3205955.

[15] A. Mokdad, P. Azmi, N. Mokari, M. Moltafet and M. Ghaffari-Miab, "Cross-Layer Energy Efficient Resource Allocation in PD-NOMA Based H-CRANs: Implementation via GPU," in IEEE Transactions on Mobile Computing, vol. 18, no. 6, pp. 1246-1259, 1 June 2019, doi: 10.1109/TMC.2018.2860985.

[16] I. Budhiraja, N. Kumar and S. Tyagi, "Cross-Layer Interference Management Scheme for D2D Mobile Users Using NOMA," in IEEE Systems Journal, vol. 15, no. 2, pp. 3109-3120, June 2021, doi: 10.1109/JSYST.2020.2997731.

[17] J. Tang et al., "Cross-Layer Optimization for Industrial Internet of Things in NOMA-Based C-RANs," in IEEE Internet of Things Journal, vol. 9, no. 18, pp. 16962-16975, 15 Sept.15, 2022, doi: 10.1109/JIOT.2021.3138461.

[18] L. Yuan, K. J. Kim and J. Zhang, "Cross-Layer Design for Fountain Coded Non-Orthogonal Multiple Access Transmission," ICC 2019 - 2019 IEEE International Conference on Communications (ICC), Shanghai, China, 2019, pp. 1-5, doi: 10.1109/ICC.2019.8761703.

[19] H. Ding and K. -C. Leung, "Cross-Layer Resource Allocation in HetNet NOMA Systems With Dynamic Traffic Arrivals," in IEEE Transactions on Communications, vol. 71, no. 3, pp. 1403-1415, March 2023, doi: 10.1109/TCOMM.2023.3239631.

[20] S. -M. Tseng, C. -S. Tsai and C. -Y. Yu, "Outage-Capacity-Based Cross Layer Resource Management for Downlink NOMA-OFDMA Video Communications: Non-Deep Learning and Deep Learning Approaches," in IEEE Access, vol. 8, pp. 140097-140107, 2020, doi: 10.1109/ACCESS.2020.3004865.

[21] X. Zhao and W. Chen, "Non-Orthogonal Multiple Access for Delay-Sensitive Communications: A Cross-Layer Approach," in IEEE Transactions on Communications, vol. 67, no. 7, pp. 5053-5068, July 2019, doi: 10.1109/TCOMM.2019.2904577.

[22] X. Lu et al., "Reinforcement Learning-Based Physical Cross-Layer Security and Privacy in 6G," in IEEE Communications Surveys & Tutorials, vol. 25, no. 1, pp. 425-466, Firstquarter 2023, doi: 10.1109/COMST.2022.3224279.

[23] W. Wang, Y. Liu, Z. Luo, T. Jiang, Q. Zhang and A. Nallanathan, "Toward Cross-Layer Design for Non-Orthogonal Multiple Access: A Quality-of-Experience Perspective," in IEEE Wireless Communications, vol. 25, no. 2, pp. 118-124, April 2018, doi: 10.1109/MWC.2018.1700081.

[24] Y. Liu and W. Chen, "Ultra Reliable and Low Latency Non-Orthogonal Multiple Access: A Cross-Layer Approach," ICC 2021 - IEEE International Conference on Communications, Montreal, QC, Canada, 2021, pp. 1-6, doi: 10.1109/ICC42927.2021.9500319.

[25] J. A. Roberts and J. M. Bargallo, "DPSK performance for indoor wireless Rician fading channels," in IEEE Transactions on Communications, vol. 42, no. 234, pp. 592-596, Feb-Apr 1994, doi: 10.1109/TCOMM.1994.577086.

[26] T. -H. Le, M. -C. Ho and L. -C. Nguyen, "Design of Compact and High Selective RF Front-end Module for Low-band 5G and IoT Applications," 2023 Photonics & Electromagnetics Research Symposium (PIERS), Prague, Czech Republic, 2023, pp. 1034-1039, doi: 10.1109/PIERS59004.2023.10221308.

[27] X. Li, Q. Wang, Y. Liu, T. A. Tsiftsis, Z. Ding, and A. Nallanathan, "UAV-Aided Multi-Way NOMA Networks With Residual Hardware Impairments," in IEEE Wireless Communications Letters, vol. 9, no. 9, pp. 1538-1542, Sept. 2020, doi: 10.1109/LWC.2020.2996782.

[28] G. Hasslinger and O. Hohlfeld, "The Gilbert-Elliott Model for Packet Loss in Real Time Services on the Internet," 14th GI/ITG Conference - Measurement, Modelling and Evaluation of Computer and Communication Systems, Dortmund, Germany, 2008, pp. 1-15.

[29] Jean Pierre Ebert, Andreas Willig, 1999. “A Gilbert-Elliott Bit Error Model and the Efficient Use in Packet Level Simulation”. Tec. Report. Technical University Berlin. Telecommunication Networks Group.
[30] R. Herrero, “Analysis of IoT mechanisms for media streaming,” Internet Things, vol. 9, Mar. 2020, Art. no. 100168.
[31] Jianhua Lu, K. B. Letaief, J. C. . -I. Chuang and M. L. Liou, "M-PSK and M-QAM BER computation using signal-space concepts," in IEEE Transactions on Communications, vol. 47, no. 2, pp. 181-184, Feb. 1999, doi: 10.1109/26.752121.
[32] Ghareeb, Ibrahim and Tashman, Deemah. (2018). Statistical Analysis of Cascaded Rician Fading Channels. International Journal of Electronics Letters. 8. 10.1080/21681724.2018.1545925.
[33] A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev, Integrals and Series, Vol. 3: More Special Functions. New York, NY, USA: Gordon and Breach Science Publishers, 1990.
[34] Laplace, P.-S. (1810), "Sur les approximations des formules qui sont fonctions de très grands nombres et sur leur application aux probabilités. Œuvres complètes", 12, 301–345.
[35] J. L. Devore, "Probability and Statistics for Engineering and the Sciences. Cengage Learning", 2015.
[36] M. Shehab, "On Applications of Machine Learning in IRS NOMA Architectures for 5G and Beyond," Ph.D. dissertation, Politecnico di Torino, Turin, Italy, 2023. Available: https://hdl.handle.net/11583/2987291
[37] M.Shehab, B. S. Ciftler, T. Khattab, M. M. Abdallah, and D. Trinchero, "Deep Reinforcement Learning Powered IRS-Assisted Downlink NOMA," in IEEE Open Journal of the Communications Society, vol. 3, pp. 729-739, 2022, doi: 10.1109/OJCOMS.2022.3165590.
[38] M. Shehab et al., "Terahertz Multiple Access: A Deep Reinforcement Learning Controlled Multihop IRS Topology," in IEEE Open Journal of the Communications Society, vol. 5, pp. 1072-1087, 2024, doi: 10.1109/OJCOMS.2024.3357701.
[39] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor," in Proc. 35th Int. Conf. on Machine Learning (ICML), vol. 80, pp. 1861-1870, Jul. 2018.