ASIAN BULLETIN OF BIG DATA MANAGMENT
Vol. 4. Issue 2 (2024)
https://doi.org/ 10.62019/abbdm.v4i02.69.

ASIAN BULLETIN OF BIG DATA MANAGEMENT
ISSN (Print): 2959-0795
http://abbdm.com/
ISSN (online): 2959-0809

# Enhancing NOMA Network Security with RIS-UAV Integration: Exploring PPO Technique

Syed Zain Ul Abideen*, Abdul Wahid, Mian Muhammad Kamal, Touseef Hussain, Naveed Jan

## Chronicle

**Article history**

**Received**: May 4, 2024
**Received in the revised format:** May 18, 2024
**Accepted:** May 19, 2024
**Available online:** May 22, 2024

**Syed Zain Ul Abideen, Abdul Wahid & Touseef Hussain** are currently affiliated with College of Computer Science and Technology, Qingdao University, Qingdao, Shandong, China.
**Email:** zain208shah@gmail.com,
**Email:** wahidjan999@gmail.com
**Email:** touseefhussain098@gmail.com

**Mian Muhammad Kamal** is currently affiliated with the Joint International Research Laboratory of Information Display and Visualization, Southeast University, Nanjing, People's Republic of China.
**Email:** mmkamal@seu.edu.cn

**Naveed Jan** is currently affiliated with the Department of Information Engineering Technology, University of Technology Nowshera, 24100, KPK, Pakistan.
**Email:** naveed.jan@uotnowshera.edu.pk

**Corresponding Author***

## Abstract

In order to improve the security aspects of a nonorthogonal multiple access (NOMA) downlink network, we investigate the use of the proximal policy optimization (PPO) technique in this study. This network is uniquely augmented with reconfi- gurable intelligent surfaces (RIS) mounted on unmanned aerial vehicles (UAVs). The main objective of this article is to prevent possible eves from accessing the network while still assuring continuous and secure connectivity for legitimate users. The network configuration includes a UAV outfitted with a RIS, which plays an important role in signal propagation. This work aims to increase the secrecy rates of all user communications in eavesdropping prone locations. By adaptively modifying the phase shifts and power distributions through the RIS the proposed PPO algorithm not only shows adaptability and effectiveness in securing wireless communications, but also highlights the significant advances made in communication system security through this technology.

**Keywords:** Index Terms—Reconfigurable intelligent surface, Secrecy, UAV, DRL.



# INTRODUCTION

The development of fifth-generation (5G) wireless technology aims to enable seamless, universal connectivity, presenting significant challenges for traditional terrestrial cellular networks. Advanced technologies such as massive multiple-input-multiple-output, which equips base stations with a large array of active antennas, have played a critical role in improving spectral efficiency in modern network systems. Furthermore, reconfigurable intelligent surfaces, also known as intelligent reflecting surfaces or large intelligent surfaces (Pan et al., 2021), consist of numerous reflective elements that adapt to incoming signals. RISs are emerging as an important step towards what could be termed massive MIMO 2.0 and are receiving considerable attention for their potential to enhance network capacity and reach (Pan et al., 2020). The typical static deployment of RIS has its limitations as fixed positions may not always align with the dynamic needs of a network environment (Chen et al., 2022a), (Chen et al., 2022b). Introducing Unmanned Aerial Vehicles (UAVs) equipped with RIS units introduces a mobile dimension to the technology allowing for an adaptive response to fluctuations in network demand and environmental conditions. This adaptation is particularly crucial in scenarios where traditional communication infrastructures fail or are insufficient (Chen et al., 2022c), (Mu et al., 2021). Unmanned aerial vehicles have shown their effectiveness in addressing the limitations of conventional terrestrial networks. They can be rapidly deployed as aerial base stations in areas with temporary increases in tele-traffic, such as during political gatherings, sports events, or after natural disasters (Jeon et al., 2022). UAVs are crucial for maintaining uninterrupted and extensive connectivity. However, to fully optimize the integration of UAVs into wireless networks, it is essential to overcome certain limitations, such as restricted coverage areas and limited power supply. The integration of RISs with UAV-supported networks enables the creation of continuous virtual line-of-sight links through passive reflection of signals. This enhances coverage and minimizes UAV displacement (You et al., 2021).

However, the inherent openness of wireless systems, especially those utilizing UAVs escalates the risk of eavesdropping. Physical-layer security technologies are thus crucial in safeguarding private communications within these networks. While there has been progress in enhancing network privacy and security using centralized RIS reflection strategies these often do not leverage advanced spectrum utilization technologies such as NOMA, which can further enhance the efficiency of the spectrum (Yu et al., 2020). Research into RIS-assisted NOMA networks has begun to address these gaps though the unique challenges of UAV-based networks need more focused attention (Han et al., 2022). Cellular-connected UAVs offer significant advantages, but they are limited by the constrained spectrum resources within cellular networks. To optimize spectrum usage, power-domain non-orthogonal multiple access is recognized as an impactful strategy. This method enhances spectrum utilization and supports improved connectivity by allowing multiple users to share a single resource block simultaneously (Song et al., 2017). NOMA improves network efficiency by enabling simultaneous access of several users in the power domain, using techniques such as superposition coding at the transmitter and successive interference cancellation at the receiver (Ding et al., 2017).

Reconfigurable Intelligent Surfaces have recently been recognized as a valuable technology for enhancing physical layer security in wireless communication. A two-dimensional array full of several inexpensive, passive components makes up a RIS, every one capable of modifying its electromagnetic response through the use of embedded positive intrinsic-negative diodes (Wang & Swindlehurst, 2023). The incident signals on the RIS can be finely tuned by an intelligent controller that adjusts the phase response of each element to amplify desired signals and reduce interference or signals aimed at unintended receivers. Unlike traditional strategies such as deploying artificial noise or utilizing multi-antenna beamforming, which require costly radio frequency chains, RISs operate by passively reflecting signals, offering a more efficient and cost-effective method for improving secrecy performance (Almohamad et al., 2020). Moreover, integrating RISs into current wireless networks is straightforward since they can be installed on various structures like roadside billboards, building facades, windows, or even wearable garments (Bae et al., 2024).

## LITERATURE REVIEW

From our review, most studies in this field primarily aim to boost security performance by adjusting critical communication network parameters, such as the base station's active beamforming and the phase shift matrix of RIS. Only a limited number of studies focus on analyzing the performance of physical layer security (Wang et al., 2021), where they establish a closed-form expression for the secrecy outage probability

(SOP). Aligning with this main research direction, our project also emphasizes the optimization of communication network parameters. In (Yu et al., 2019), the presence of obstacles blocking the direct path between the transmitter and receiver led to the use of a RIS to boost security performance. Here, the transmitter's beamformer and the RIS's phase shift matrix were optimized in turns through the block coordinate descent (BCD) method. Study (Jiang et al., 2020) examined a multiple-input multiple-output multiple-antenna eavesdropper (MIMOME) communication network, where the transmit signal covariance matrix and the RIS's phase shift matrix were optimized via an alternating algorithm. In (Lu et al., 2020), the focus was on maximizing the worst-case achievable secrecy rate, addressing scenarios with both colluding and non-colluding eves, and considering a realistic range of channel state information (CSI) for the eves.

Research presented in (Yu et al., 2020) incorporated artificial noise (AN) to secure a multiple RIS-assisted communication system involving multiple legitimate users and eves. An effective algorithm was developed to maximize the sum-rate while limiting confidential signal leakage, using techniques such as alternating optimization (AO), successive convex approximation (SCA), and semidefinite relaxation (SDR) to optimize the transmit signal covariance matrix and phase shifters. Additionally, (Li et al., 2021) explored the deployment of a UAV as an aerial base station, enhancing continuous communication services for terrestrial nodes with RIS support. Here, a weighted secrecy rate that combined uplink and downlink transmissions was maximized through AO.

The collaborative use of RIS and UAV technologies is transforming 5G and future networks, significantly enhancing the services provided to users (Ranjha & Kaddoum, 2021), (Yang et al., 2020). This integration leverages UAVs' mobility to create direct line-of-sight (LoS) connections with ground users and uses RISs to refine their reflective elements for effective passive beamforming. For instance, UAVs can act as mobile base stations, connecting with ground devices with the support of one or several RISs. Furthermore, this combination helps mitigate issues such as LoS obstructions and energy consumption in UAVs while improving the overall quality of wireless channels.

Research into this combined approach has yielded convincing numerical evidence, supporting its practical implementation in areas such as energy optimization and physical layer security (PLS). Particularly, configurations where UAVs are outfitted with RISs have proven effective for securing communications. In applications utilizing deep learning-based algorithms, the arrangement of RIS's passive beamforming and the UAV's navigation path can be finely tuned to optimize data rates and ensure equitable resource distribution. Investigations into single-antenna UAV systems have focused on optimizing resource allocation and system trajectory (Sun et al., 2019), (Li et al., 2020). While equipping UAVs with multiple antennas can enhance device performance (Zheng & Tse, 2003), it introduces challenges like increased weight, size, and power consumption. Conversely, RISs can replicate the multi-input multi-output (MIMO) functionality of multiple antennas using economical meta-surfaces (Hu et al., 2018), (Nadeem et al., 2020), achieving higher gains through phase shift adjustments without the need for multiple antennas on UAVs. Moreover, incorporating RIS into UAV systems can streamline the UAV's flight path, as users nearer to the RIS can directly receive signals, minimizing the need for extensive UAV travel. This integration facilitates improved signal transmission to distant users through beamforming, ensuring they receive data at satisfactory rates.

## Motivation and Contribution

The research is motivated by the important contribution of Reconfigurable Intelligent Surfaces to the development of 6G wireless communication technologies. As we move into a new phase in wireless communications, Better spectrum management, cost-effectiveness, and energy efficiency are becoming increasingly important. RIS addresses these challenges by effectively altering electromagnetic wave propagation, providing significant benefits in areas where direct communication links are blocked. The increasing occurrence of security risks in wireless networks highlights the importance of implementing robust physical layer security protocols. Reflective Intelligent Surface technology presents a fresh strategy for protecting data transfer from unauthorized access and monitoring, as well as enhancing service performance through advanced non-line-of-sight links. Through precise manipulation of RIS components, it becomes feasible to improve Quality of Service, reinforce security measures, and leverage RIS's distinct spatial selectivity to maintain signal privacy for authorized users while deterring possible eves.

The combination of NOMA and RIS, and possibly UAVs highlights the significant potential for these technologies to greatly impact the creation of a more flexible, effective, and secure wireless communication environment. This research is motivated by the ability of these advanced technologies to address the intricate challenges found in dynamic multi-user scenarios where traditional optimization approaches are not successful due to their non-linear nature. Our study uses the algorithm of PPO to simultaneously improve UAV positioning, phase shifting, and power distribution. This approach enables quick adjustments to changing channel conditions and user requirements. By integrating this optimization approach, we guarantee effective real-time allocation of wireless resources that enhances spectrum efficiency and upholds peak performance in diverse operational situations. Additionally, the system consistently refines its techniques to adapt to changing network dynamics, Consequently, with sophisticated AI-driven learning algorithms, resource consumption in wireless communication contexts is improved.

Our research concentrates on creating more advanced strategies for securing the physical layer, aiming to improve overall system performance and protect against potential eavesdropping threats. Through the incorporation of secure beamforming and complex transmission techniques into our optimization framework, we aim to enhance the security and reliability of wireless transmissions, thereby decreasing the risk of unauthorized access and data interception. Our study emphasizes the significant importance and real-world applicability of proactive security measures in wireless communication networks that handle sensitive data or function in adversarial environments. We aim to verify the efficiency of our PPO-based approach through comprehensive computer simulations. Through the use of RIS, NOMA, and UAV technologies, these simulations will demonstrate the capacity to strengthen the physical layer to prevent security breaches and enhance network performance. The future development of secure and effective wireless communication networks is what this research aims to achieve.

# SYSTEM MODEL

We investigate a NOMA network of communications with UAV assistance to improve confidentiality. A set of U end users receive data transmissions from a base station (BS) within the system, denoted by $U = \{1,2\ldots.,U\}$ This system utilizes a UAV equipped with a RIS containing N passive elements to reflect and direct signals efficiently. Operating autonomously, the UAV maintains a specified altitude over area A and starts its mission from a predefined charging station. Designed for flat-fading channels with quasi-static frequencies, the system assumes perfect CSI at both the BS and UAV-mounted RIS, simplifying by ignoring energy consumption and operation duration. The system is susceptible to clandestine surveillance by a collection of eves, labeled as E = {1,2…, E} who seek to covertly intercept communications. The transmission link from the BS to the UAV-mounted RIS is represented by the matrix $F \in C^{NX1}$. Concurrently,

In parallel, the routes of communication that connect each RIS to each u-th end user and e-th eve are captured through $G_{r,u} \in C^{NX1}$ and $G_{r,e} \in C^{NX1}$ correspondingly. The signal received by each u-th user is mathematically formulated as:

$$y_u = \left(\mathbf{G}_{r,u}^H \mathbf{\Phi} \mathbf{F}\right) \sum_{i=1}^{U} \sqrt{\rho_i} s_i + n_u,\ u \in \mathcal{U} \tag{1}$$

where $\mathbf{\Phi} = \text{diag}\left(e^{j\theta_1}, \dots, e^{j\theta_N}\right)$ accounts for the RIS's phase shift. The power allocation $p_i$ for each user's signal $s_i$ is restricted within [0.1], summing to 1 across all users. $s_i$ has an expected power of 1 denoted as $\mathbb{E}\left[s_i^2\right] = 1$, and $n_U$ is the additive white Gaussian noise with a mean and variance of zero $\sigma^2$. The position of RIS on an UAV, is situated at $v(x, y)$ with a height $h_I$, whereas the source is the BS (0,0) with height $h_B$. The horizontal position of each $u$-th user is given by $u_u(x_u, y_u)$. The distance from the RIS and BS is calculated as $d_{BI} = \sqrt{x^2 + y^2 + (h_B - h_I)^2}$, and the distance between the RIS and the $u$-th user is $d_{Iu_u} = \sqrt{(x - x_u)^2 + (y - y_u)^2 + h_I^2}$. The signal received at Eve:

$$y_e = \left(\mathbf{G}_{r,e}^H \mathbf{\Phi} \mathbf{F}\right) \sum_{i=1}^{U} \sqrt{p_i} s_i + n_e,\ e \in \mathcal{E} \tag{2}$$

The channel gains for each legitimate user u and eavesdropper e, considering path loss are expressed

$$\Gamma_U = \frac{\boldsymbol{G}_{rk}^H \boldsymbol{\Phi} \boldsymbol{F}}{(d_{BI} d_{Iu_u})^\alpha} \tag{3}$$

$$\Gamma_e = \frac{\boldsymbol{G}_{re}^H \boldsymbol{\Phi} \boldsymbol{F}}{(d_{BI} d_{Ie})^\alpha} \tag{4}$$

where $\alpha$ represents the path loss exponent, and, $d_{BI}$ and $d_{Iu_u}$ and $d_{Ie}$ denote The distances between user u or eavesdropper e and the RIS, and between the RIS and the BS, respectively.

The SINR required for NOMA users to use successive interference cancellation (SIC), and accounting for potential interception by eves, is calculated as:

$$\Upsilon_{t \to u} = \frac{|\Gamma_u|^2 P_{\max} \rho_t}{\sum_{i=t+1}^{u} |\Gamma_u|^2 P_{\max} \rho_i + \sigma^2} \tag{5}$$

$$\Upsilon_{t \to e} = \frac{|\Gamma_e|^2 P_{\max} \rho_t}{\sum_{i=t+1}^{u} |\Gamma_e|^2 P_{\max} \rho_i + \sigma^2} \tag{6}$$

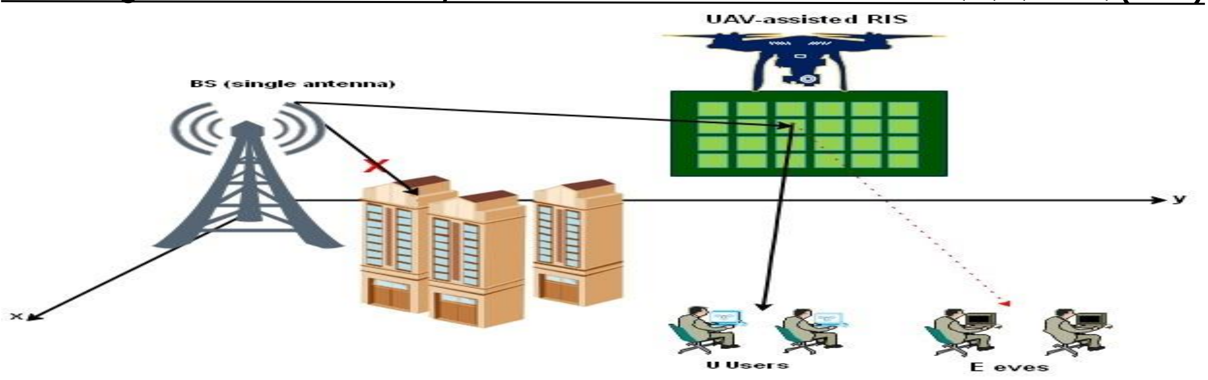


**Figure 1.**
**UAV-Assisted Reconfigurable intelligent surface NOMA Downlink System**

Lastly, the secure transmission rate is quantified by defining the secrecy rate for each u-th user.

$$R_{s,u} = \left[R_u - \max_{e \in \mathcal{E}} R_{e,u}\right]^+ \tag{7}$$

where $R_{e,u}$ is the possible data rate that the e-th eve can obtain, and $R_u$ is the acceptable transmission rate to the u-th user. These rates are expressed as follows:

$$R_u = \log_2\left(1 + \frac{|\mathbf{G}_{r,u}^H \mathbf{\Phi} \mathbf{F}|^2 p_u}{\sigma^2 + \sum_{i \neq u} |\mathbf{G}_{r,u}^H \mathbf{\Phi} \mathbf{F}|^2 p_i}\right) \tag{8}$$

$$R_{e,u} = \log_2\left(1 + \frac{|\mathbf{G}_{r,e}^H \mathbf{\Phi} \mathbf{F}|^2 p_u}{\sigma^2 + \sum_{i \neq u} |\mathbf{G}_{r,e}^H \mathbf{\Phi} \mathbf{F}|^2 p_i}\right) \tag{9}$$

In order to ensure robust secure communication against eavesdropping threats, this detailed model considers communication paths that are possibly compromised as well as those that are secure. Optimizing the total secrecy rates of all users is the ultimate objective.

## PROBLEM FORMULATION

The main aim of this advanced communication framework is to maximize the total secrecy rates for all users, which is essential for ensuring secure communications in the presence of eavesdropping threats. Our all-inclusive RIS-UAV-NOMA downlink network model optimizes the total sum rate and modifies critical parameters to increase security in order to improve both physical layer security and communication efficiency. These variables include the power allocation $\rho_i$ at the Base Station (BS), the phase adjustment $\Phi$ of the RIS, and the horizontal coordinates $v(x, y)$ of the UAV. The coordination of these factors creates a complicated optimization task that can be defined in the following manner:

The aim of this setup is to strengthen the physical layer's security by reducing potential eavesdropping risks $R_{s,t}$. represents the secrecy rate for user t. Constraint (10b) ensures that each user t's secrecy rate exceeds a minimum value, thereby ensuring Quality of Service for all users. The implementation of Successive Interference Cancellation, which is essential for effective decoding of NOMA signals, relies on Constraint (10c). To achieve optimal power management, constraint (10d) controls the overall transmission power of the Base Station. The operational zone of the UAV is defined by constraint (10e), requiring it to remain within area $\mathcal{A}$. Finally, constraint (10f) governs the phase shifts carried out by RIS, restricting $\theta_n$ each element to be between 0 and $2\pi$.

$$\max_{\{\Theta,\rho,v\}} \sum_{t=1}^{U} R_{s,t} \tag{10a}$$
$$\text{s.t. } R_{s,t} \geq R_{\min}, \forall t \in U \tag{10b}$$
$$R_{t\to j} \geq R_{t\to t}, \forall t, j \in U, t > j \tag{10c}$$
$$\sum_{u=1}^{K} \rho_u \leq 1 \tag{10d}$$
$$v(x,y) \in \mathcal{A} \tag{10e}$$
$$0 \leq \theta_n \leq 2\pi, n = 1, \cdots, N \tag{10f}$$

Finding the most efficient global solution to this optimization problem presents a significant challenge due to its non-linear characteristics, influenced by intricate interconnections among variables $\{\Theta, \rho, v\}$. We propose an extensive and successful method for addressing these challenges through Deep Reinforcement Learning with PPO. This approach aims to manage the problem's complex nature and high-dimensional aspects, providing a dependable and effective means of attaining nearly optimal solutions while enhancing the security and efficiency of communication networks.

## DRL-BASED APPROACH FOR SECURING NOMA COMMUNICATIONS

In this section we detail the DRL strategy utilized to optimize the UAV-assisted NOMA communication system which is aimed at enhancing confidentiality through intelligent and dynamic control of the RIS-equipped UAV. Notably, DRL excels in environments requiring complex decision-making by learning optimal state-action mappings directly from interactions with the environment without the need for predefined training, (Li et al., 2020). DRL algorithms are distinguished into two primary types: On-policy and Off-policy methods. On-policy algorithms such as PPO Asynchronous Advantage Actor-Critic (A3C) and Phasic Policy Gradient (PPG), (Gu et al., 2017) evaluate and improve the same policy used to make decisions in the environment. These methods maintain the exploration and exploitation balance by updating the policy incrementally using a function known as the clipped surrogate objective which ensures that the new policy does not deviate significantly from the existing policy.

Conversely, Off-policy algorithms like Deep Deterministic Policy Gradient (DDPG) operate independently of the agent policy, allowing the exploration of new strategies by learning from actions that are outside the current policy. This characteristic can be advantageous in environments where acquiring new data is costly or risky. Given the UAV-assisted NOMA system's setup, where a base station communicates with multiple users via a UAV-mounted RIS, both On-policy and Off-policy methods are applicable for optimizing signal reflections and directing transmissions securely away from potential eves. However, for our specific implementation challenges characterized by the need for rapid adaptation to dynamic communication channels and the imperative of minimizing computational overhead On-policy methods are preferable. Thus, we adopt PPO, an On-policy algorithm known for its effectiveness in continuous action spaces and its computational efficiency. PPO will be used to dynamically adjust the RIS's phase shifts ($\mathbf{\Phi}$) to maximize the secure transmission rates while minimizing potential interceptions by eves. The algorithm will interact with the environment to maximize the expected sum of secrecy rates across all communications, which is mathematically defined for each user $u$ in the system as:

$$R_{s,u} = \left[R_u - \max_{e \in \mathcal{E}} R_{e,u}\right]^+ \quad (11)$$

Where $R_u$ and $R_{e,u}$ represent the transmission rates to the legitimate users and the potential eves, respectively. The aim of our study is to optimize the configuration of the RIS to maximize $R_u$ and minimize $R_{e,u}$ across all channels thereby enhancing the NOMA communication system's overall security. The subsequent subsections will provide an overview of the DRL formulation and an in-depth analysis of the PPO algorithm that has been customized for our particular application. In the following subsections, we will outline the DRL formulation and delve into the specifics of the PPO algorithm tailored for our application.

# PROPOSED PPO

The implementation of the PPO algorithm is essential within the integrated system that combines RIS technology with UAV-assisted NOMA downlink communications, aiming to bolster Physical Layer Security (PLS). Initially, this section provides a concise introduction to PPO before examining the modifications made to the PPO framework to address the particular optimization challenges presented by this system. PPO as opposed to conventional algorithms such as Deep Q-Networks (DQN), which are designed for discrete action spaces is specifically optimized for environments characterized by continuous action spaces. PPO being a model-free on-policy algorithm utilizes a stochastic policy gradient approach that is improved by cropped probability ratios in the objective function. By establishing a robust equilibrium between exploration and exploitation this design empowers PPO to efficiently traverse environments characterized by a multitude of continuous actions.

## The PPO Framework

PPO employs an actor-critic structure where policy and value functions are distinct. Its efficacy stems from the equilibrium it maintains between exploration of new strategies and exploitation of known rewards. This balance is achieved by constraining the magnitude of updates to the policy, a process regulated by a clipping mechanism in the objective function that curtails excessively large and potentially disruptive updates. The equation you've provided can be condensed into two lines as follows:

$$\text{Objective} = \min\left(\frac{\pi_\theta(a|s)}{\pi_{\theta_{old}}(a|s)} A^{\pi_{\theta_{old}}}(s,a),\ \text{clip}\left(\frac{\pi_\theta(a|s)}{\pi_{\theta_{old}}(a|s)}, 1-\epsilon, 1+\epsilon\right) A^{\pi_{\theta_{old}}}(s,a)\right) \quad (12)$$

where $\epsilon$ is a small number like 0.1 or 0.2, which defines the clipping range to moderate updates, thus stabilizing the training process.

## The PPO Processing

The state space delineates the complete set of observable variables at each discrete time step undefined, providing a full representation of the system's current condition. It includes the previous secrecy metrics (undefined) for each user undefined, the most recent RIS phase modulations and power distribution along with the current coordinates of the UAV. These elements collectively form the state space, offering the necessary information to inform subsequent decisions.

## PPO ALGORITHM FOR RIS-UAV-NOMA COMMUNICATION SYSTEM

The state space ($s$) encompasses critical elements such as the UAV's position, represented by coordinates $v(x, y)$, the phase shifts of the RIS denoted as $\Phi$, and the channel state information (CSI) pertinent to all users and eves. This collection of data points forms the foundation for decision-making within the system. The action space ($a$) includes several key adjustments crucial for system optimization: the coefficients of power allocation ($\rho_k$) for each user, the ($\Phi$) is phase shift matrix of RIS, and the future positioning of the UAV, given by $v'(x, y)$. These elements allow for dynamic control over the system's operational parameters to meet performance and security objectives effectively. The reward ($r$) is calculated by evaluating the secrecy rates of users, with the objective of optimizing both the network performance and security. This ensures that the system not only functions efficiently but also maintain stringent security standards, reflecting a dual-focus approach in the optimization process.

## PROPOSED PPO-BASED ALGORITHM INITIALIZATION:

Randomly initialize the policy network $\pi(a|s;\ \theta^{\pi})$ and the value network $V(s;\ \theta^{V})$ with their respective weights $\theta\pi$ and $\theta V$.

Prepare the experience replay buffer $\boldsymbol{\mathcal{D}}$ with capacity C.

Set learning parameters: learning rate β, discount factor γ, clipping parameter ε, and minibatch size $NB$.
Procedure:
For each iteration *i* = 1 to I = do
For each episode *j* = 1 to *J* = do
Initialize the UAV's position $v(x, y)$, RIS phase shifts $\Phi$, and obtain the channel state information $G(j)$ and $h_{rk}^{(j)}$ for all users and eves.
Calculate the initial state $S_1$.
For each timestep t = 1 to T do:

During each timestep of the simulation, the process unfolds as follows: An action $a_t$ is selected based on the current policy $\pi(a|s_t;\theta^{\pi})$ This command controls the RIS phase changes, power distributions, and UAV placement. Subsequently, the new secrecy rates $R_{s,k}^{(t)}$ for each user are calculated to determine the total secrecy rate, which is used to compute the reward $r_t$. The system then captures the updated state $s_{t+1}$, including changes in the UAV's position, RIS settings, and channel state information. This transition data current state, action, reward, and new state is stored in the replay buffer *D* aiding the learning process by providing historical data for future analysis.

Sample a minibatch of $N_B$ transitions from *D* for training.

For each sampled transition, compute advantages using the difference between returns and value estimates.

$$L(\theta^{\pi}) = \hat{\mathbb{E}}[\min(\text{ratio} \times \text{advantage}, \text{clip}(\text{ratio}, 1-\epsilon, 1+\epsilon) \times \text{advantage})],$$

Where ratio $= \frac{\pi(a|s;\theta^{\pi})}{\pi_{\text{old}}(a|s)}$

Update the value network by minimizing the loss function:

$$L(\theta^{V}) = \hat{\mathbb{E}}[(V(s;\theta^{V}) - \text{Returns})^2].$$

# SIMULATION RESULTS

In our research, we employ the PPO method within a RIS-enabled UAV-NOMA communication system to assess its effectiveness in enhancing both security and performance. In our simulations, the BS is strategically positioned at the beginning, while the UAV outfitted with RIS starts at the coordinates (40,0). The area designated for users is defined by the vertices (35,35), (45,35), (45,45), and (35,45), with user locations remaining constant in each simulation run.

The system uses a Rician fading model to maintain Line-of-Sight (LoS) communication between the RIS and the users, as well as between the RIS and the BS.

$$\boldsymbol{G} = \sqrt{\frac{\Omega}{\Omega+1}}\overline{\boldsymbol{H}} + \sqrt{\frac{1}{\Omega+1}}\boldsymbol{H}_{\text{Rayleigh}} \quad (13)$$

In this model, $\overline{\boldsymbol{H}}$ denotes the component of line-of-sight, $\boldsymbol{H}_{\text{Rayleigh}}$ the component of non-line-of-sight influenced by Rician K-factor, and the Rayleigh fading, Ω, is fixed at 8. A path loss exponent ($\alpha$) of 2 is used, with channel conditions being predetermined at the beginning of each episode. The UAV is positioned at an altitude of 20 meters, matching the elevation of the Base Station (BS). The system's noise power is configured to $\sigma^2 = -60$ dB, and it requires a minimum user rate of $R_{\min} = 1.2$ bps/Hz. The policy (actor) and value (critic) networks within the PPO framework are structured using dual-layer fully connected neural networks, To guarantee effective gradient propagation, the first layer's ReLU activation functions and the output layer's tanh.

Key hyperparameters guiding the simulation include a lr ($\beta$) of 0.0001, a $\gamma$ (discount factor) of 0.95, a $\tau$ (soft update rate ) of 0.004, and replay buffer with a capacity of 40,000. The simulation unfolds over 400 episodes, with each episode consi- sting of 200 steps, and utilizes a minibatch size of 16. Exploration within the system is driven by noise characterized by a complicated Gaussian distribution with 0.1 variance and zero mean, promoting varied policy exploration. To account for potential security threats, the simulation incorporates the RIS-to-Eve channel, $\mathbf{h}_{r,e}$, and modifies Eve's channel gain relative to that of Authentic users, especially tailored to unique location. This inclusion is essential for realistically assessing and mitigating the risk of eavesdropping in the communication system.

$$\Gamma_e = \frac{\mathbf{G}_{r,e}^H \mathbf{\Phi} \mathbf{F}}{(d_{B,I} \cdot d_{I,e})^\alpha}$$

This configuration allows the PPO algorithm to dynamically tune power allocations ($\rho_i$) and phase adjustments ($\theta_n$) to continuously optimize the security and efficiency of communications in response to evolving threats and communication dynamics.

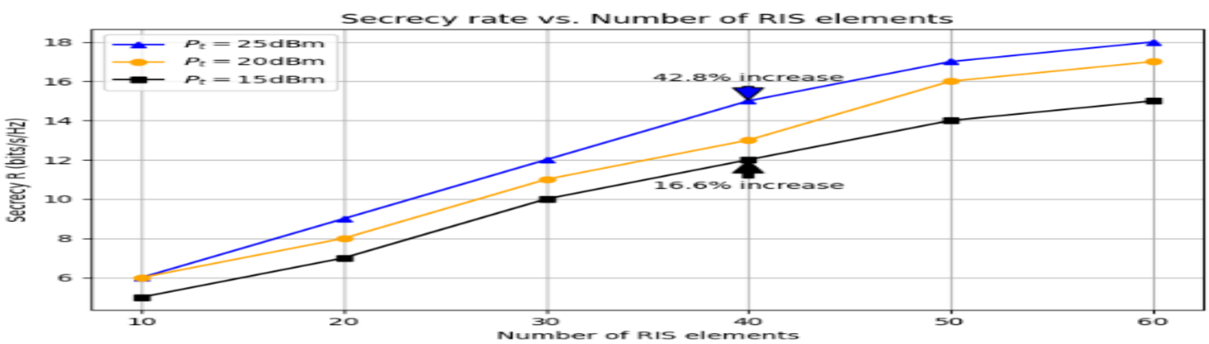


**Figure 2.**
**Number of RIS elements**

In Fig 2, the line graph illustrates how the secrecy rate, measured in bps/Hz, escalates with an increase in the number of RIS elements within a communication system. Three distinct curves represent various transmit power levels—specifically 15 dBm, 20 dBm, and 25 dBm. Each curve trends upwards, clearly showing that the addition of RIS elements is positively correlated with the improvement of the system's secrecy rate. Notably, at the point where the system employs 40 RIS elements, there are marked percentage increases in secrecy rates: a 16.6 percent rise at a transmit power of 15 dBm, and an even more significant 42.8 percent surge at 25 dBm. These figures suggest substantial gains in secrecy performance due to increasing the number of RIS elements, with higher transmit powers yielding more pronounced benefits. While the graph does not specify the baseline from which these percentages are derived, it is evident that incorporating a greater number of RIS elements can greatly enhance the secure transmission capabilities of the system, particularly when operated at higher power levels.

Fig 3 illustrates a comparison of secrecy rates, measured in bps/Hz, as a function of varying transmit powers at the BS, denoted in dBm. The configurations under examination include systems with 40 RIS elements at a user-to-eavesdropper ratio of 6, as well as systems with 30 RIS elements at user-to-eavesdropper ratios of 6 and 4. A consistent upward trend across these scenarios reveals that higher transmit powers correlate with increased secrecy rates. The highest secrecy rate is observed in the system boasting the greatest number of RIS elements coupled with the largest user-to-eavesdropper ratio, highlighting the significant roles both factors play in the enhancement of secure communications.

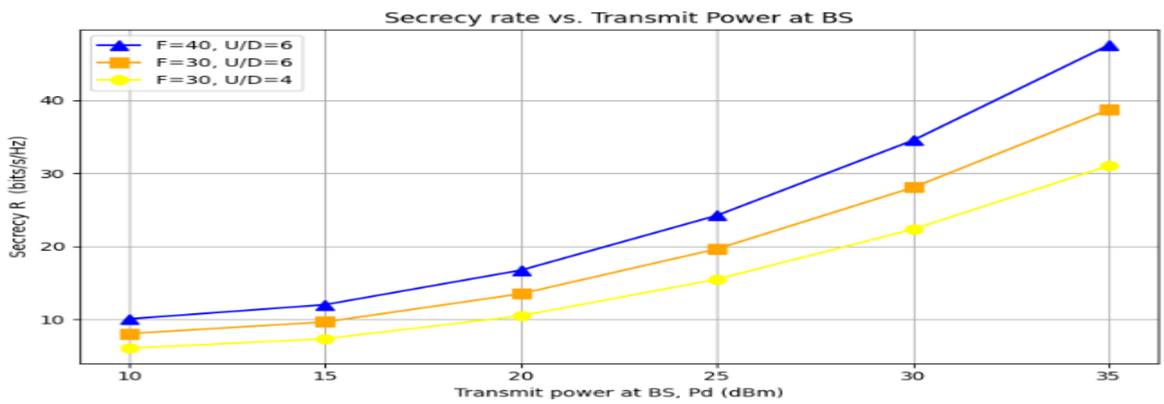


**Figure 3.**
**Transmit power**

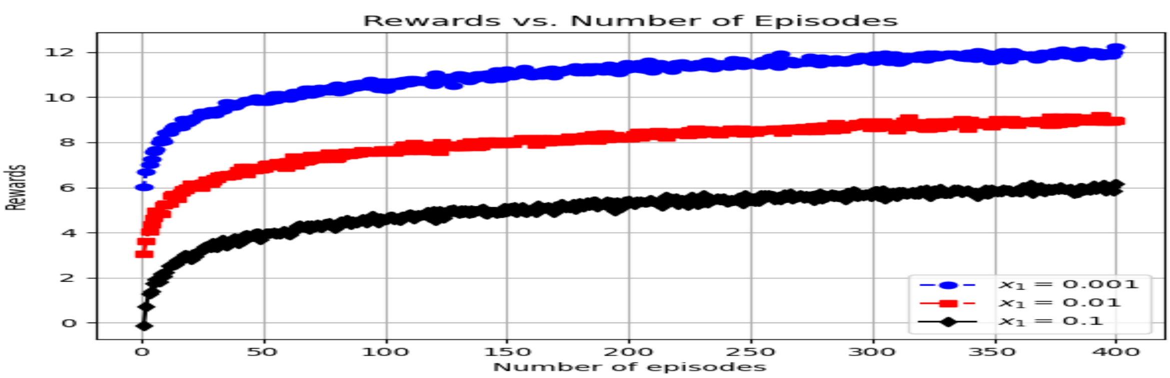


**Figure 4.**
**Number of episodes**

In Fig 4, the plotted data illustrates the correlation between accumulated rewards and the number of completed episodes for three different configurations of the parameter $x_1$. These configurations are represented by the blue, red, and black lines, corresponding to $x_1$ values of 0.001, 0.01, and 0.1, respectively. Each line shows a reward increase as the episode count climbs to 400. The blue line exhibits a notable ascent and eventually levels off above a reward value of 12. The red line presents a steady, moderate incline, reaching stability around a reward value of 8. The black line, on the other hand, after a promising start, plateaus at the lowest final reward value, slightly above 4. This visualization underscores the influence of the $x_1$ parameter on the learning trajectory, where a setting of $x_1 = 0.001$ appears to lead to the highest reward attainment after 400 episodes.

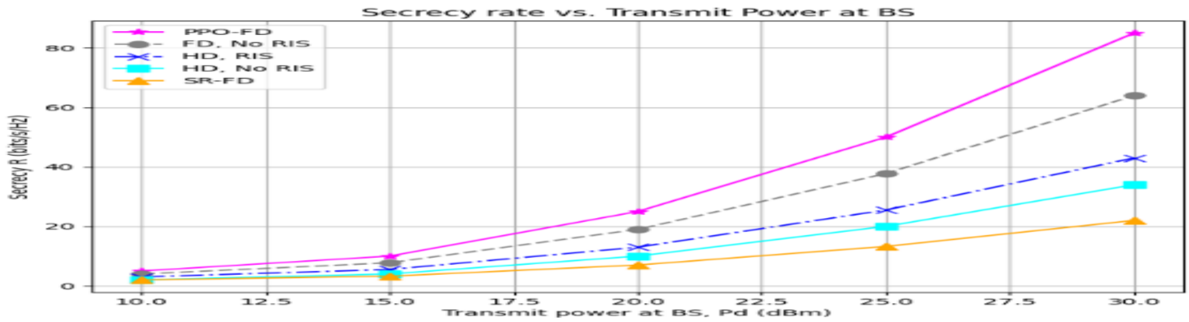


**Figure 5.**
**Impact of various transmit power at BS on Secrecy rate**

Fig 5 charts the relationship between secrecy rates in bps/Hz and the base station's transmit power, displayed in dBm, across a range of communication techniques. The graph reveals that the PPO-FD method, featuring Full Duplex with Proximal Policy Optimization and illustrated by pink stars, significantly leads in performance, climbing sharply with increased power. Full Duplex without the addition of RIS, depicted with grey circles, shows a consistent but less dramatic rise. Half Duplex setups, both with RIS (blue crosses) and without (orange triangles), demonstrate only slight gains, with the latter being the least improved. The SR-FD strategy, represented by cyan squares, exhibits moderate progress, better than Half Duplex without RIS but not as effective as Full Duplex approaches. This visualization underscores the efficacy of combining PPO and RIS in Full Duplex systems for superior secrecy rate enhancements as transmit power is amped up.

# CONCLUSION

This study investigates the potential of using an Intelligent RIS together with a UAV to protect data transmission in a downlink NOMA network. The approach utilizes the PPO algorithm to precisely adjust the IRS phase shifts allocate power to the BS, and position the UAV spatially. To ensure SIC confidentiality while increasing overall data rate, a decoding sequence adaptable to fluctuating channel conditions is necessary. Simulation results demonstrate that this PPO-centric approach enhances network performance as a whole and effectively handles variations in IRS configurations and user counts. Importantly, it implements heightened preventive measures against unauthorized entry, strengthening network security against surveillance while guaranteeing unrestricted access for authorized users; thus, representing significant improvement in wireless network security.

# DECLARATIONS

**Acknowledgement:** We appreciate the generous support from all the supervisors and their different affiliations.
**Funding:** No funding body in the public, private, or nonprofit sectors provided a particular grant for this research.
**Availability of data and material:** In the approach, the data sources for the variables are stated.
**Authors' contributions:** Each author participated equally to the creation of this work.
Conflicts of Interests: The authors declare no conflict of interest.
**Consent to Participate:** Yes
**Consent for publication and Ethical approval:** Because this study does not include human or animal data, ethical approval is not required for publication. All authors have given their consent.

# REFERENCES

Almohamad, A., Tahir, A., Al-Kababji, A., Furqan, H M., Khattab, T., Hasna, M O., & Arslan, H. (2020, January 1). Smart and Secure Wireless Communications via Reflecting Intelligent Surfaces: A Short Survey. IEEE Communications Society, 1, 1442-1456. https://doi.org/10.1109/ojcoms.2020.3023731

Bae, J., Khalid, W., Lee, A., Lee, H., Noh, S., & Yu, H. (2024, March 1). Overview of RIS-enabled secure transmission in 6G wireless networks. KeAi. https://doi.org/10.1016/j.dcan.2024.02.005

Chen, J., Guo, L., Jia, J., Shang, J., & Wang, X. (2022, April 1). Resource Allocation for IRS Assisted SGF NOMA Transmission: A MADRL Approach. Institute of Electrical and Electronics Engineers, 40(4), 1302-1316. https://doi.org/10.1109/jsac.2022.3144726

Chen, J., Ma, Z., Zou, Y., Jia, J., & Wang, X. (2022, December 4). DRL-based Energy Efficient Resource Allocation for STAR-RIS Assisted Coordinated Multi-cell Networks. https://doi.org/10.1109/globecom48099.2022.10001595

Chen, J., Xie, Y., Mu, X., Jia, J., Liu, Y., & Wang, X. (2022, July 1). Energy Efficient Resource Allocation for IRS Assisted CoMP Systems. Institute of Electrical and Electronics Engineers, 21(7), 5688-5702. https://doi.org/10.1109/twc.2022.3142784

Ding, Z., Liu, Y., Choi, J., Sun, Q., Elkashlan, M., Chih-Lin, I., & Poor, H V. (2017, February 1). Application of Non-Orthogonal Multiple Access in LTE and 5G Networks. Institute of Electrical and Electronics Engineers, 55(2), 185-191. https://doi.org/10.1109/mcom.2017.1500657cm

Gu, S. S., Lillicrap, T., Turner, R. E., Ghahramani, Z., Schölkopf, B., & Levine, S. (2017). Interpolated policy gradient: Merging on-policy and off-policy gradient estimation for deep reinforcement learning. Advances in Neural Information Processing Systems, 30.

Han, Y., Li, N., Liu, Y., Zhang, T., & Tao, X. (2022, June 1). Artificial Noise Aided Secure NOMA Communications in STAR-RIS Networks. Institute of Electrical and Electronics Engineers, 11(6), 1191-1195. https://doi.org/10.1109/lwc.2022.3161020

Hu, S., Rusek, F., & Edfors, O. (2018, May 15). Beyond Massive MIMO: The Potential of Data Transmission With Large Intelligent Surfaces. Institute of Electrical and Electronics Engineers, 66(10), 2746-2758. https://doi.org/10.1109/tsp.2018.2816577

Jeon, H., Park, S., Park, J., Huang, K., & Chae, C. (2022, August 1). An Energy-Efficient Aerial Backhaul System With Reconfigurable Intelligent Surface. Institute of Electrical and Electronics Engineers, 21(8), 6478-6494. https://doi.org/10.1109/twc.2022.3149903

Jiang, W., Zhang, Y., Wu, J., Wang, F., & Jin, Y. (2020, January 1). Intelligent Reflecting Surface Assisted Secure Wireless Communications With Multiple- Transmit and Multiple-Receive Antennas. Institute of Electrical and Electronics Engineers, 8, 86659-86673. https://doi.org/10.1109/access.2020.2992613

Li, R., Wei, Z., Yang, L., Ng, D W K., Yuan, J., & An, J. (2020, July 1). Resource Allocation for Secure Multi-UAV Communication Systems With Multi-Eavesdropper. IEEE Communications Society, 68(7), 4490-4506. https://doi.org/10.1109/tcomm.2020.2983040

Li, S., Duo, B., Renzo, M D., Tao, M., & Yuan, X. (2021, October 1). Robust Secure UAV Communications With the Aid of Reconfigurable Intelligent Surfaces. Institute of Electrical and Electronics Engineers, 20(10), 6402-6417. https://doi.org/10.1109/twc.2021.3073746

Li, Y., Han, W., & Wang, Y. (2020). Deep reinforcement learning with application to air confrontation intelligent decision making of manned/unmanned aerial vehicle cooperative system. IEEE Access, 8, 67887–67898.

Lu, X., Yang, W., Guan, X., Wu, Q., & Cai, Y. (2020, December 1). Robust and Secure Beamforming for Intelligent Reflecting Surface Aided mmWave MISO Systems. Institute of Electrical and Electronics Engineers, 9(12), 2068-2072. https://doi.org/10.1109/lwc.2020.3012664

Mu, X., Liu, Y., Guo, L., Lin, J., & Poor, H V. (2021, October 1). Intelligent Reflecting Surface Enhanced Multi-UAV NOMA Networks. Institute of Electrical and Electronics Engineers, 39(10), 3051-3066. https://doi.org/10.1109/jsac.2021.3088679

Nadeem, Q., Kammoun, A., Chaaban, A., Debbah, M., & Alouini, M. (2020, December 1). Asymptotic Max-Min SINR Analysis of Reconfigurable Intelligent Surface Assisted MISO Systems. Institute of Electrical and Electronics Engineers, 19(12), 7748-7764. https://doi.org/10.1109/twc.2020.2986438

Pan, C., Ren, H., Wang, K., Kolb, J F., Elkashlan, M., Chen, M., Renzo, M D., Hao, Y., Wang, J., Swindlehurst, A L., You, X., & Hanzo, L. (2021, June 1). Reconfigurable Intelligent Surfaces for 6G Systems: Principles, Applications, and Research Directions. Institute of Electrical and Electronics Engineers, 59(6), 14-20. https://doi.org/10.1109/mcom.001.2001076

Pan, C., Ren, H., Wang, K., Xu, W., Elkashlan, M., Nallanathan, A., & Hanzo, L. (2020, August 1). Multicell MIMO Communications Relying on Intelligent Reflecting Surfaces. Institute of Electrical and Electronics Engineers, 19(8), 5218-5233. https://doi.org/10.1109/twc.2020.2990766

Ranjha, A., & Kaddoum, G. (2021, March 15). URLLC Facilitated by Mobile UAV Relay and RIS: A Joint Design of Passive Beamforming, Blocklength, and UAV Positioning. Institute of Electrical and Electronics Engineers, 8(6), 4618-4627. https://doi.org/10.1109/jiot.2020.3027149

Song, L., Li, Y., Ding, Z., & Poor, H V. (2017, July 1). Resource Management in Non-Orthogonal Multiple Access Networks for 5G and Beyond. Institute of Electrical and Electronics Engineers, 31(4), 8-14. https://doi.org/10.1109/mnet.2017.1600287

Sun, Y., Xu, D., Ng, D W K., Dai, L., & Schober, R. (2019, June 1). Optimal 3D-Trajectory Design and Resource Allocation for Solar-Powered UAV Communication Systems. IEEE Communications Society, 67(6), 4281-4298. https://doi.org/10.1109/tcomm.2019.2900630

Wang, F., & Swindlehurst, A L. (2023, January 1). Applications of Absorptive Reconfigurable Intelligent Surfaces in Interference Mitigation and Physical Layer Security. Institute of Electrical and Electronics Engineers, 1-1. https://doi.org/10.1109/twc.2023.3312693

Wang, W., Tian, H., & Ni, W. (2021, December 1). Secrecy Performance Analysis of IRS-Aided UAV Relay System. Institute of Electrical and Electronics Engineers, 10(12), 2693-2697. https://doi.org/10.1109/lwc.2021.3112752

Yang, L., Meng, F., Zhang, J., Hasna, M O., & Renzo, M D. (2020, September 1). On the Performance of RIS-Assisted Dual-Hop UAV Communication Systems. Institute of Electrical and Electronics Engineers, 69(9), 10385-10390. https://doi.org/10.1109/tvt.2020.3004598

You, C., Kang, Z., Zeng, Y., & Zhang, R. (2021, December 1). Enabling Smart Reflection in Integrated Air-Ground Wireless Network: IRS Meets UAV. Institute of Electrical and Electronics Engineers, 28(6), 138-144. https://doi.org/10.1109/mwc.001.2100148

Yu, X., Xu, D., & Schober, R. (2019, December 1). Enabling Secure Wireless Communications via Intelligent Reflecting Surfaces. https://doi.org/10.1109/globecom38437.2019.9014322

Yu, X., Xu, D., Sun, Y., Ng, D W K., & Schober, R. (2020, November 1). Robust and Secure Wireless Communications via Intelligent Reflecting Surfaces. Institute of Electrical and Electronics Engineers, 38(11), 2637-2652. https://doi.org/10.1109/jsac.2020.3007043

Yu, X., Xu, D., Sun, Y., Ng, D W K., & Schober, R. (2020, November 1). Robust and Secure Wireless Communications via Intelligent Reflecting Surfaces. Institute of Electrical and Electronics Engineers, 38(11), 2637-2652. https://doi.org/10.1109/jsac.2020.3007043
Zheng, L., & Tse, D. (2003, May 1). Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels. Institute of Electrical and Electronics Engineers, 49(5), 1073-1096. https://doi.org/10.1109/tit.2003.810646