

Deep Learning-Based Channel Estimation and Dynamic IRS Assignment for Optimized Beamforming in IRS-Aided MC-NOMA Systems



Amish Ranjan^{*}, Bikash Chandra Sahana^{*}

Department of Electronics and Communication Engineering, National Institute of Technology, Patna 800005, India

Corresponding Author Email: amishr.phd18.ec@nitp.ac.in

Copyright: ©2024 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/mmep.111202>

ABSTRACT

Received: 23 September 2024

Revised: 8 November 2024

Accepted: 15 November 2024

Available online: 31 December 2024

Keywords:

beamforming optimization, channel estimation, deep learning, reinforcement learning, Q-learning

This paper proposes a novel framework for optimizing beamforming gain in multi-carrier non-orthogonal multiple access (MC-NOMA) systems that are assisted by intelligent reflecting surfaces (IRS). The framework includes channel estimation based on deep learning, dynamic IRS assignment, and beamforming optimization. First, a convolutional neural network-long short-term memory (CNN-LSTM) model was implemented to estimate the channel state information (CSI) accurately. After that, a Q-learning agent utilizes this estimated CSI to allocate IRS elements to user clusters in a dynamic manner, with the objective of optimizing the use of IRS units according to the existing channel conditions. Using the allocated IRS elements and the estimated CSI, a deep Q-network (DQN) is developed to find the best beamforming vectors that save the most power and get the optimum signal-to-interference-plus-noise ratio (SINR). When compared to random IRS assignment and traditional heuristic beamforming optimization methods, the results show that the proposed framework shows significant improvements in SINR, power efficiency, and total system capacity. Simulation results show that the integration of deep learning and reinforcement learning techniques in the IRS-assisted MC-NOMA system can significantly improve the performance, indicating that the proposed framework is a potential solution for wireless communication in the future.

1. INTRODUCTION

More and more people see communication systems that use IRS as a possible game-changer for next-generation cellular networks. These systems are made up of many passively reflecting parts that can change the phase shifts of receiving electromagnetic waves in real time, making it easier for signals to travel. Because of this dynamic adjustment capability, IRSs can be used in a variety of situations to boost signal quality, increase range, and make both spectrum and energy use more efficient. MC-NOMA is another important technology for future wireless networks because it lets multiple people share the same frequency resources while using different amounts of power. Adding IRSs to MC-NOMA systems can make them work much better together, which makes this combination an important area of study for 6G and later wireless networks.

Deep learning is a type of machine learning that has a lot of potential to solve challenging issues in radio communications. Long short-term memory (LSTM) networks and convolutional neural networks (CNNs) are great for channel estimation tasks because they can record both spatial and temporal connections. A lot of the research that goes into enhancing wireless communication systems depends on having accurate CSI. CNN-LSTM models are effective at figuring out what the multidimensional channel data means.

Reinforcement learning (RL) is a powerful method is that

also used to make systems with IRS better at allocating resources and beamforming. A lot of people use RL algorithms like Q-learning and deep Q-networks (DQNs) to find the best rules for a situation by interacting with it over and over again. By using RL along with deep learning-based channel forecast methods, we can make a flexible plan for allocating IRS and improve beamforming to get the highest SINR while using the least amount of power.

Intelligent reflecting surfaces (IRSs) have been looked at in a number of studies as a possible use in wireless communication systems. A lot of research was done on IRS technology [1, 2], including its pros and cons. IRS combined with MC-NOMA are quickly becoming recognized as a game-changing method for building the next generation of cell phone networks. IRS is made up of passive reflecting elements that can change phase shifts. As a result, real-time changes to the wireless channel improve coverage, spectrum and energy economy, and signal quality. MC-NOMA allows multiple users to gain connectivity at the same time by sharing the same frequency resources but allocating different power levels. The integration of IRS and MC-NOMA has tremendous promise for 6G and future networks in order to meet the growing demands for connection, capacity, and efficiency in next-generation communications [2, 3].

Deep learning indeed achieves impressive performance for solving some complex channel estimation problems in IRS-

assisted wireless systems. More specifically, CNNs and LSTM networks can learn effective features from spatial and temporal dependence, respectively. Recent study showed that the proposed DL model achieved higher accuracy of channel estimation for IRS-assisted systems than traditional methods and further estimated more precise CSI [4]. The integration of CNN and LSTM for channel estimation in IRS-aided networks drew considerable attention because such a combined model can process multidimensional channel data with ease [5]. Besides DL-based channel estimation, reinforcement learning has also been very effective in dynamic resource allocation and beamforming optimization. Reinforcement learning can enable the system to adapt resource allocation to real-time channel state information. Q-learning and DQN can have obvious advantages in IRS scenarios. Application of deep reinforcement learning for robust beamforming optimization in IRS-aided communication systems achieves better performance compared to heuristic methods under dynamic channel conditions [6]. Meanwhile, the recent research also provided the efficiency of Q-learning and DQN for IRS-based beamforming enhancement and resource allocation enhancement respectively [6-9]. Recent work has explored IRS assignment strategies. Researchers studied the role of reinforcement learning in improving dynamic IRS assignment. These methods explain the importance of IRS element distribution for improving system efficiency under dynamic channel conditions [10-12]. Q-learning enables the agent to learn optimal actions over a discrete state-action space. It could enable applications such as assignment of IRS elements where dynamic adjustment of IRS configuration along with CSI feedback is required [13]. A thorough review of deep learning-based channel estimation is detailed by Ranjan et al. [14]. Beamforming optimization has been a significant focus in IRS-aided systems. Recent studies conclude with the application of reinforcement learning and advanced optimization techniques to enhance beamforming performance [15-21]. The objective of these works is to improve the overall SINR and system capacity. DQN has also been used to optimize beamforming for DQN in a scenario when the BS has to handle multi-user interference and simultaneously maximize the overall throughput. By leveraging both estimated CSI and IRS configurations, DQN dynamically readjusts the beamforming vectors at the base station to dynamically maximize SINR and make use of the available power resources efficiently [22].

Integration of DL for channel estimation with RL for resource allocation is one of the recent paradigms in jointly managing channel and IRS capabilities. Indeed, several recent frameworks have unveiled that DL-based CSI estimation combined with RL-based resource allocation can be conducted within the same framework for addressing estimation and allocation challenges. A model was proposed that first combined DL for channel estimation and RL for adaptive IRS assignment and beamforming, and the results demonstrated its superiority over other state-of-the-art methods in SINR and system efficiency [23].

Still, there are additional unresolved questions in various spheres of research. Most studies either focus on channel estimates or IRS assignment without considering the integration of both. Furthermore, existing works mostly rely on heuristic or stationary-based IRS assignment techniques, which might not fully maximize IRS capabilities. Furthermore,

there is a scarcity of extensive research that integrates deep learning and reinforcement learning methods to tackle the challenges of channel estimation and resource allocation in systems that IRS in MC-NOMA systems.

To fill in these voids, this work proposes a new framework including dynamic IRS assignment, beamforming optimization using RL, and deep learning-based estimation of channel. This work makes mostly the following important contributions:

1. Deep learning-based channel estimation: A CNN-LSTM model applied to get exact estimations of CSI in IRS-aided MC-NOMA systems. This model makes use of CNN and LSTM capabilities to capture spatial and temporal aspects.

2. Dynamic IRS assignment: Based on Q-learning, a dynamic IRS assignment system has been implemented. This algorithm is meant to maximize the use of IRS units and change with real-time channel conditions.

3. Beamforming optimization: Using a DQN that takes into account the dynamically allocated IRS units and predicted CSI helps find the ideal beamforming vectors.

4. Performance evaluation: The performance of the proposed framework is compared with conventional methods and shows significant developments in SINR, power efficiency, and system capacity.

The merging of deep learning and RL methods to handle both channel estimation and resource allocation issues in IRS-aided MC-NOMA systems are unique about this work. The aim is to provide a complete solution incorporating various cutting-edge technologies, thereby increasing the general efficiency of upcoming wireless communication networks.

This paper is organized in the following way: In Section 2, the system model and the problem statement are explained. In Section 3, the suggested way to estimate channels using deep learning is laid out. In Section 4, the Q-learning-based method for assigning IRS is shown. Section 5 introduces the DQN-based method for enhancing beamforming optimization. Section 6 presents the simulation results and conducts a thorough analysis of the performance. Section 7 serves as the final section of the report and delves into possible avenues for further research.

2. SYSTEM MODEL AND PROBLEM FORMULATION

2.1 System model

The model considered is an IRS-aided MC-NOMA system where K users are divided into multiple clusters. Each cluster, based on the requirement of the system and total load of the users, contain maximum 5 users. The base station (BS) communicates with users through both direct and reflected linkages via the IRS.

For non-orthogonal multiple access, the users within a cluster use the same subcarrier while power levels distinguish them in the power domain. This clustering allows for efficient resource utilization and caters to the demands brought about by multi-user environments in next-generation networks. The IRS is fitted with an intelligent controller that can maximize communication performance by adjusting the phase shifts of the reflecting parts. Figure 1 represents the basic communication channel of a MC-NOMA system aided by IRS elements attached to a building.

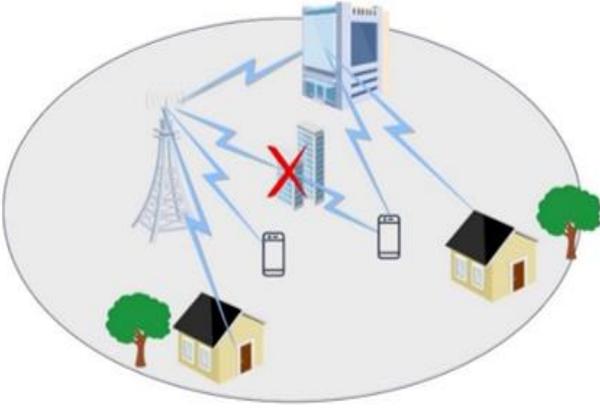


Figure 1. An illustration of IRS assisted downlink wireless communication system

2.1.1 Channel model

The communication channels between the base station (BS), IRS, and users are modelled as frequency-selective fading channels. Each channel is represented as a combination of direct channel and indirect channel, following the Rician fading model. Let $\mathbf{h}_{d,k,n} \in \mathbb{C}^{1 \times N}$ denote the direct channel between the BS and the k -th user on the n -th subcarrier, and $\mathbf{h}_{r,k,n} \in \mathbb{C}^{1 \times M}$ represent the channel between the intelligent reflecting surface (IRS) and the k -th user on the n -th subcarrier. The channel connecting the BS and the IRS is represented by the matrix \mathbf{G} , which belongs to the set of complex numbers \mathbb{C} and has dimensions $M \times N$. The aggregate channel between the base BS and the k -th user through the intelligent IRS on the n -th subcarrier is expressed as:

$$\mathbf{h}_{k,n} = \mathbf{h}_{d,k,n} + \mathbf{h}_{r,k,n} \Theta \mathbf{G} \quad (1)$$

where, $\Theta = \text{diag}(\theta_1, \theta_2, \dots, \theta_M)$ is the IRS's reflection coefficient matrix, and the phase shift of the m -th reflecting element is represented by θ_m .

2.1.2 Signal model

On the k -th user of the n -th subcarrier, the received signal is represented as follows:

$$y_{k,n} = \mathbf{h}_{k,n} \mathbf{w}_n s_n + \sum_{j=1, j \neq k}^K \mathbf{h}_{j,n} \mathbf{w}_n s_n + n_{k,n} \quad (2)$$

where, the beamforming vector \mathbf{w}_n belongs to the complex number space $\mathbb{C}^{N \times 1}$. It represents the beamforming on the n -th subcarrier. The symbol sent on the n -th subcarrier is denoted as s_n . The additive white Gaussian noise (AWGN) on the n -th subcarrier $n_{k,n}$ has a zero mean and variance of σ^2 . For the k -th user on the n -th subcarrier, the SINR is given by:

$$\text{SINR}_{k,n} = \frac{|\mathbf{h}_{k,n} \mathbf{w}_n|^2}{\sum_{j=1, j \neq k}^K |\mathbf{h}_{j,n} \mathbf{w}_n|^2 + \sigma^2} \quad (3)$$

2.2 Problem formulation

The objective is to optimize the SINR for each user by optimizing the beamforming vectors, IRS assignment, and channel estimation. The three principal subproblems are the

estimation of the channel, the assignment of the IRS, and the optimization of the beamforming vectors.

2.2.1 Channel estimation

Effective IRS configuration and beamforming necessitate precise CSI. The CSI is estimated from the pilot signals that are received using a CNN-LSTM model. The estimation of the channel problem can be expressed as:

$$\hat{\mathbf{h}}_{k,n} = \arg \min_{\mathbf{h}} \|\mathbf{y}_{k,n} - \mathbf{h}_{k,n} \mathbf{w}_n\|^2 \quad (4)$$

where, $\hat{\mathbf{h}}_{k,n}$ is the estimated channel for the k -th user on the n -th subcarrier.

2.2.2 IRS assignment

Strategic group assignment of IRS elements helps to maximize the performance of a system. Reinforcement learning problems help to address this process. Under this method, the "state" is the present condition of the communication channels; the "action" is the assignment of IRS elements; the resulting SINR serves as the reward. Q-learning is a method whereby the system learns by updating a table depending on the obtained rewards. This can be formulated as:

$$Q(s, a) = Q(s, a) + \alpha[r + \gamma \max_{a'} Q(s', a') - Q(s, a)] \quad (5)$$

where, Q-value for state S and action a is represented by $Q(s, a)$, the learning rate is α , γ is the discount factor, and the reward (SINR) is represented by r in the above formulae.

2.2.3 Beamforming optimization

The objective of the beamforming optimization problem is to identify the most favourable beamforming vectors that optimize the SINR for every user. This can be expressed as:

$$\mathbf{W}_n = \arg \max_{\mathbf{w}} \sum_{k=1}^K \log(1 + \text{SINR}_{k,n}) \quad (6)$$

subject to power constraints:

$\|\mathbf{W}_n\|^2 \leq P_{max}$, where the maximum power transmitted is P_{max} . To address this issue, we employ a DQN approach. In this method, the state is the latest assignment for both the CSI and the IRS. The motion is matched by the beamforming vector, and the reward is based on the SINR that is obtained.

3. PROPOSED DEEP LEARNING-BASED CHANNEL ESTIMATION

This section presents the deep learning method for channel estimation in MC-NOMA systems. The method precisely forecasts the CSI from obtained pilot signals by using LSTM and CNN. The estimated CSI is then used to optimize the assignment of the IRS and the beamforming vectors. Precise knowledge of the CSI is crucial for setting up IRS and maximizing the efficiency of beamforming in MC-NOMA systems. Conventional methods for estimating channels often face difficulties due to the intricate and large-scale nature of the channel matrices in these systems. Nevertheless, deep learning, particularly when employing CNN and LSTM architectures, offers a compelling alternative because of its

ability to acquire complex patterns and connections from data.

3.1 CNN- LSTM model for channel estimation

In channel estimation, the CNN-LSTM architecture was chosen because it uniquely can capture spatial and temporal dependencies in channel data. CNNs are well-suited for spatial feature extraction, as in the case of complicated channel matrices that may be generated by the IRS-aided MC-NOMA systems. Through convolutional layers, CNN performs effective learning of spatial patterns, with a reduction in input dimensionality that improves computational efficiency [24]. One of the recent works investigates channel estimation within IRS-assisted Integrated Sensing and Communication (ISAC) systems. The proposed framework utilizes deep learning and incorporates two distinct neural network architectures: one designed for channel estimation at the ISAC base station and the other for communication channels at user equipment. The simulation results demonstrate that the framework significantly exceeds the performance of benchmark schemes across a variety of system parameters and SNR conditions [25].

In particular, LSTM network architectures are developed under the intuition of capturing sequential dependencies; hence, they are best suited towards the incorporation of temporal variations in the channel. Indeed, the wireless channels in IRS-assisted systems are commonly subject to changes in time due to user mobility and other environmental factors. Therefore, this issue of accurate channel estimation in dynamically changing environments can be well addressed by combining the CNN-LSTM model, which helps in improving robustness compared to conventional methods such as LMMSE and SVR, which cannot handle nonlinear patterns

and temporal variations.

Whereas CNN-LSTM networks have better performance regarding accuracy and adaptability, their deep architecture implies increased computational complexity and therefore longer training times. Other variants that might be considered include RNNs or simple CNNs, which would entail faster performance, perhaps with some loss of accuracy under dynamic conditions. Besides, CNN-LSTM models might be more sensitive to overfitting when cases involve rare events and thus require serious regularization and data augmentation techniques.

3.1.1 Model architecture

Figure 2 represents the architecture of the CNN-LSTM model. The following layers contribute to the entire model:

1. **Input layer:** Input to the proposed architecture is the received pilot signal matrix $\mathbf{Y} \in \mathbb{C}^{N \times M}$, where N represents the quantity of antennas located at the base station and M denotes the number of subcarriers.
2. **Convolutional layers:** Input data is processed using many layers to extract spatial information. To reduce dimensionality, each convolutional layer is followed by a pooling layer and a ReLU activation function.
3. **LSTM layers:** The output of the convolutional layers is sent into LSTM layers, which detect temporal dependencies in the data.
4. **Fully connected layers:** To get the final channel estimates, the LSTM layers' output is sent via fully connected layers.
5. **Output layer:** The output layer generates the predicted CSI for each user and subcarrier.

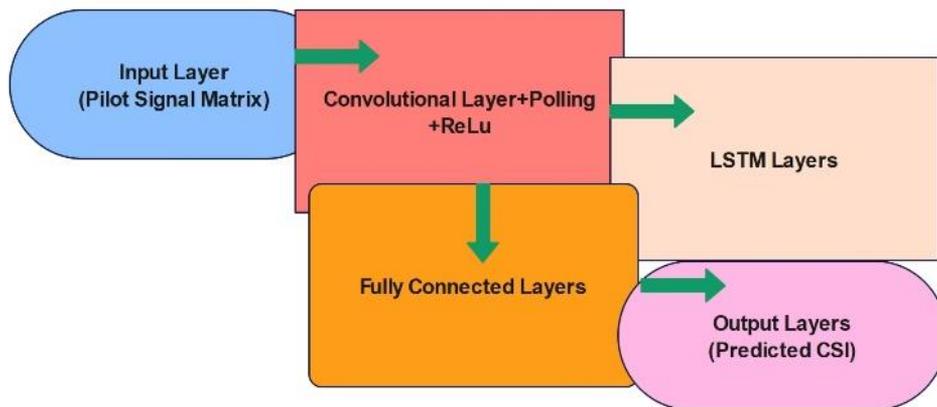


Figure 2. CNN-LSTM model architecture for channel estimation

3.1.2 Model training

The proposed model is trained on synthetic data provided through a realistic channel model adapted for IRS-aided MC-NOMA systems. Therefore, this section covers the methodology of the channel modelling, pre-processing, normalization, splitting for training and validation, and augmentation that are followed to ensure robust training and reliable operation of the model.

1. Data generation

A frequency-selective fading channel model, which will generate the channel taking into consideration both the direct and indirect components, is adopted to efficiently train the CNN-LSTM model in IRS-aided MC-NOMA systems. The channel model comprises the following

components:

- **Rician fading model:** Then, considering more realistic propagation conditions, using the Rician fading model, the channel between the BS and the IRS and that between the IRS and each user were modelled. The use of such a model is justified by the fact that in many IRS-assisted systems, there will be one dominant direct path.
- **Path loss and shadowing:** Each link of the channel includes path loss based on a standard free-space path loss model. Shadowing effects have then been added for the indirect components in order to simulate the variations caused by obstacles present in the environment.

- **Channel matrix representation:** Channel data is generated for every user and subcarrier, including the direct channel from BS to the user and reflected channels from BS to the user via IRS.

2. Data preprocessing

After generating the raw channel data, various preprocessing steps were performed to prepare the data for use as input to the CNN-LSTM model.

- **Normalization:** Training data inside the raw channel matrix varied from each other by a great amount. Therefore, normalization was applied to create stable model training. To accomplish this, each sample was normalized to zero mean and unit variance according to the following formula:

$$\text{normalized_sample} = \frac{\text{sample-mean}}{\text{std_dev}}$$

where, mean and std_dev are the mean and standard deviation of each sample. This scaling made sure that all features were added equally during model training, so no one feature could be too important.

- **Complex to real conversion:** The channel matrices, in this regard, are divided into real and imaginary values since the input to the CNN-LSTM model needs to be real-valued. Each of these was dealt with separately, thus doubling the dimension while retaining important signal characteristics.
- **Train-validation split:** The data were then split into an 80% training set, a 10% validation set, and a 10% test set. Only the training set was used for fitting the model parameters, while hyperparameter tuning and avoiding overfitting were done with the help of the validation set. After training, the test set showed the real performance of the model. This division was maintained constant for different experiments to maintain coherence in the comparison of performances obtained with the model.

3. Data augmentation

The following data augmentation to make the CNN-LSTM even more robust in performing simulations for real variations that may happen in an IRS-assisted system:

- **SNR variability:** For simulating variable signal conditions, Gaussian noise was added to each sample for different levels of SNR conditions. This augmentation allows the model to learn about CSI estimation on a wide range of SNRs. Hence, it will

generalize well in variable noise.

- **Phase shifts:** Additional samples are created by randomly perturbing the phase shifts of IRS elements to emulate real-time changes that might be experienced in dynamic environments.
- **Loss function:** The MSE loss function is used to train the model. It finds the difference between the estimated and true CSI. It is given by the equation as follows:

$$\text{Loss} = \frac{1}{K} \sum_{k=1}^K \left\| \hat{\mathbf{h}}_{k,n} - \mathbf{h}_{k,n} \right\|^2 \quad (7)$$

where, the estimated CSI is represented by $\hat{\mathbf{h}}_{k,n}$ and for the k -th user on the n -th subcarrier, $\mathbf{h}_{k,n}$ is the true CSI.

4. Q-LEARNING-BASED IRS ASSIGNMENT STRATEGY

This section discusses the Q-learning-based IRS distribution method. Q-learning has been adopted in IRS assignment problems; this is a model-free reinforcement learning algorithm that learns an optimal policy in discrete action spaces with no requirement of environment modelling. This is favourable in IRS-assisted systems where explicit environmental models may be complex to obtain due to dynamic user distributions and changing channel conditions. Q-learning is an iterative learning process to converge to progressively improved IRS configurations based on feedback about the environment; hence, it applies to problems like IRS-element assignment [22].

Q-learning is quite effective for discrete action spaces and more restricted state spaces. However, when the action space is high-dimensional or continuous, Q-learning becomes computationally inefficient since a very large Q-table may be required, where feasible convergence is slow. Alternative methods include policy gradient methods, or SARSA, which are applied to continuous actions but need more sophisticated tuning and possibly do not improve the performance further in simple scenarios. Therefore, taking into account the discrete nature of the IRS element assignment, Q-learning was suitable in this work despite its limitations. By using the Bellman Eq. (5), Q-value function is updated, and learning the Q-values that maximize the predicted cumulative reward is the aim. Figure 3 represents the Q-learning-based IRS assignment strategy.

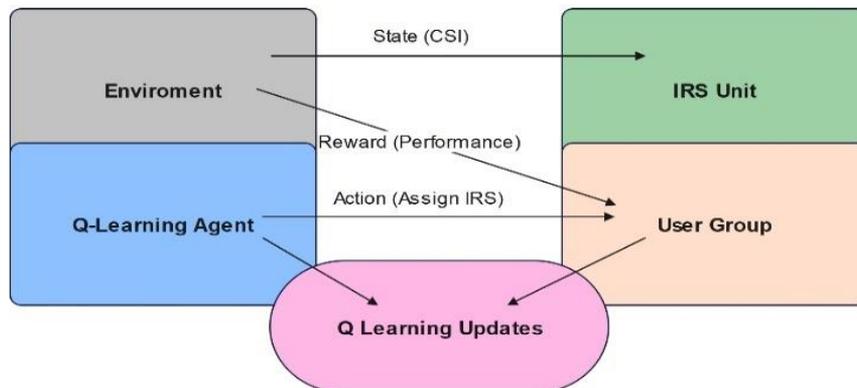


Figure 3. Q-learning-based IRS assignment strategy

4.1 State representation

From Eq. (5), the state s is provided by the existing CSI for both users and subcarriers. Let $\mathbf{H} \in \mathbb{C}^{K \times N \times M}$ define the channel matrix as an estimation, with K representing the number of users, N representing the number of subcarriers, and M representing the number of reflecting elements at the IRS. The vectorized form of \mathbf{H} represents the state s .

4.2 Action representation

Assignment of IRS items to user clusters defines the action a . Let $\mathbf{A} \in \{0,1\}^{K \times M}$ defines the IRS allocation matrix, where $A_{k,m} = 1$ for the m -th IRS elements allocated to k -th user, and 0 otherwise. The vectorized form of \mathbf{A} represents the action a .

4.3 Reward function

The reward, denoted as r , is determined by the increase in SINR that occurs as a result of the assignment of IRS. The equation of SINR for the k -th user on the n -th subcarrier is given in Eq. (3). The reward r is calculated by summing the signal-to-interference-plus-noise Ratios (SINRs) for all users and subcarriers is given by:

$$r = \sum_{k=1}^K \sum_{n=1}^N \text{SINR}_{k,n} \quad (8)$$

Algorithm 1: Q-learning-based IRS Assignment

Input: Learning rate α , Estimated CSI $\hat{\mathbf{h}}$, discount factor γ , exploration rate ϵ

Output: Optimal IRS assignment policy $Q(s, a)$

- 1: Initialize $Q(s, a) \leftarrow 0$
 - 2: for each episode do
 - 3: Initialize state S
 - 4: for each step do
 - 5: With probability ϵ , select random action a
 - 6: Else, select action $a \leftarrow \text{argmax}_a Q(s, a)$
 - 7: Apply action a and observe next state s' and reward r
 - 8: Update $Q(s, a) \leftarrow Q(s, a) + \alpha(r + \gamma \max_{a'} Q(s', a') - Q(s, a))$
 - 9: Update state $s \leftarrow s'$
 - 10: end for
 - 11: Decay exploration rate ϵ
 - 12: end for
 - 13: return $Q(s, a)$
-

5. DQN-BASED BEAMFORMING OPTIMIZATION METHOD

The DQN method extends Q-learning to high-dimensional action spaces using a neural network to approximate the Q-values, and it is ideal for the beamforming optimization problem in the multi-user environment. In IRS-aided MC-NOMA systems, the optimization of beamforming constitutes a complicated issue since it needs to handle multi-user interference while maximizing SINR and system throughput. However, when the size of the state and/or action spaces is high enough, DQN may perfectly fit, which dynamically allows the model to learn a policy for complex beamforming decisions by leveraging both CSI and IRS configurations [26,

27].

Computation-intensive and may suffer from instability in training, especially in high-dimensional state space environments. Experience replay and target networks are used for stabilizing training, further increasing computational overhead. Other methods, such as Deep Deterministic Policy Gradient (DDPG) may be more efficient for continuous action spaces but may require additional tuning. In this case, DQN is adopted since it provides a good balance between performance and complexity, resulting in robust beamforming optimization within a computationally manageable framework. Figure 4 represents the DQN-based beamforming optimization strategy.

5.1 State representation

From Eq. (5) the state s is provided by the existing CSI and IRS assignment. Let $\mathbf{H} \in \mathbb{C}^{K \times N \times M}$ define the channel matrix as an estimation, and $\mathbf{A} \in \{0,1\}^{K \times M}$ defines the IRS allocation matrix. The state s is formed by combining the vectorized representations of \mathbf{H} and \mathbf{A} .

5.2 Action representation

The beamforming vector \mathbf{w}_n for the n -th subcarrier defines the action a . Let the beamforming matrix for all subcarriers and users is shown by $\mathbf{W} \in \mathbb{C}^{N \times K}$. By looking at the vectorized form of \mathbf{W} , the action a can be represented.

5.3 Reward function

The reward, denoted as r , is determined by the increase in SINR that occurs as a result of the beamforming optimization. The equation of SINR for the k -th user on the n -th subcarrier is given in Eq. (3). The reward r , is calculated by summing the SINRs for all users and subcarriers is given in Eq. (8).

5.4 DQN model structure

It is structured to handle the high-dimensional state-action space that will be associated with the IRS-aided MC-NOMA system for the DQN model in use for optimization of beamforming. This is an architecture structured in the following way:

(1) Input layer

The input to the DQN model consists of the estimated CSI and the IRS configuration acquired, which constitutes a high-dimensional state vector.

(2) Hidden layers

- (a) First hidden layer: Fully connected layer with 128 neurons and ReLU activation to capture complex interactions among the state elements.
- (b) Second hidden layer: Fully connected layer with 64 neurons, ReLU activation, refining the feature representation.
- (c) Third hidden layer: Fully connected layer with 32 neurons and ReLU activation, further reducing dimensionality but retaining key features for decision-making.

(3) Output layer

The output layer provides the Q-values for every possible action at a given state, with each Q-value corresponding to beamforming vector adjustments.

The network is trained to approximate the Q-value function and selects the action with the highest Q-value for optimal beamforming.

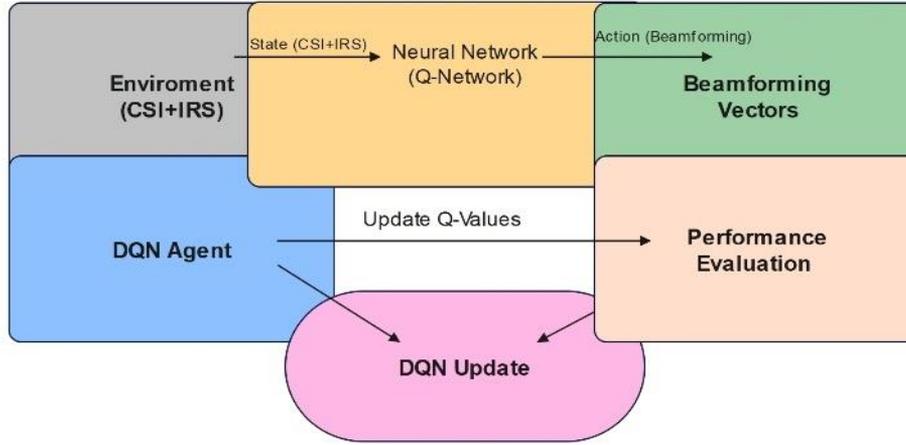


Figure 4. DQN-based beamforming optimization strategy

5.5 DQN algorithm for beamforming optimization

Algorithm 2: DQN-based beamforming optimization

Input: Learning rate α , Estimated CSI $\hat{\mathbf{h}}$, discount factor γ , IRS assignment \mathbf{A} , exploration rate ε

Output: Optimal beamforming policy $Q(s, a; \theta)$

- 1: Initialize Q-network $Q(s, a; \theta)$ with random weights θ
 - 2: Initialize target network $Q'(s, a; \theta^-)$ with weights $\theta^- \leftarrow \theta$
 - 3: Initialize experience replay memory \mathcal{D}
 - 4: for each episode do
 - 5: Initialize state s
 - 6: for each step do
 - 7: With probability ε , select random action a
 - 8: Otherwise, select action $a \leftarrow \operatorname{argmax}_a Q(s, a; \theta)$
 - 9: Apply action a and observe next state s' and reward r
 - 10: Store transition (s, a, r, s') in replay memory \mathcal{D}
 - 11: Sample mini-batch of transitions (s, a, r, s') from \mathcal{D}
 - 12: Compute target $y \leftarrow r + \gamma \max_a Q'(s', a; \theta^-)$
 - 13: Perform gradient descent step to minimize loss $(y - Q(s, a; \theta))^2$
 - 14: Update state $s \leftarrow s'$
 - 15: end for
 - 16: Periodically update target network weights $\theta^- \leftarrow \theta$
 - 17: end for
 - 18: return $Q(s, a; \theta)$
-

6. SIMULATION RESULTS AND PERFORMANCE ANALYSIS

This section provides a comprehensive study of the simulation findings and performance of the proposed framework. The simulation setup specifically involves accurately modelling an IRS-aided MC-NOMA system in a realistic wireless communication environment. The various components and parameters are selected carefully to ensure the model applies to multi-user scenarios common to next-generation networks. BS has an 8-antenna uniform linear array (ULA) in place. This work considers the ULA because of its

simplicity and efficiency in tiling directional beams, which play a crucial role in focusing signal power on specific users in a multi-user scenario. A uniform planar array (UPA), consisting of 64 reflecting elements in an 8×8 structure, forms the IRS. The IRS can reflect signals in both azimuth and elevation directions thanks to the UPA structure's ability to perform 3D beamforming. We have chosen the number of elements to achieve a satisfactory trade-off between sufficient coverage and minimal computational overhead. The simulation takes into account 10 randomly distributed users within the BS and IRS coverage areas. The simulation clusters the users, with each cluster containing 3-5 users who share the same subcarrier via power-domain NOMA. This is representative of the typical MC-NOMA setups in dense environments. The users are distributed randomly within a radius of 100 meters from the BS, taking into account realistic spatial separation. The random distribution shows how the users' positions change, which is important for testing how well IRS-assisted beamforming and dynamic resource allocation work. The IRS is positioned at a fixed position approximately 50 meters from the BS, carefully positioned to ensure direct connectivity to both the BS and the users. The IRS strategically positions itself to reflect the signal to users in areas with shadows or low coverage, aiming to maximize its reflection. Because the IRS is positioned at a specific distance and angle, users can receive real service in a variety of random locations within its coverage area. This is a wonderful example of how to improve user connectivity in tough spots compared to BS. Rician model is used to model the BS-IRS and IRS-user links, accounting for both the direct and indirect components of each link. Rayleigh fading is used to model the BS-user direct links. AWGN is applied across all channels with a zero mean and unit variance. The simulations are then conducted across an SNR range of -10 dB to 15 dB to test the system under variable signal conditions. This covers low-to-high SNR and, therefore, allows to check the robustness of the model in facing different levels of noise and interference.

The framework combines deep learning-based channel estimation, Q-learning-based IRS assignment, and DQN-based beamforming optimization for IRS-aided MC-NOMA systems. Tables 1-3 summarize the main factors for the simulation of CNN-LSTM model for channel estimation, Q-learning model for IRS Assignment, and DQN model for beamforming optimization respectively.

Table 1. Simulation parameters of CNN-LSTM model for channel estimation

Parameters	Value	Description
Learning Rate	0.001	Controls the step size for updating model weights.
Batch Size	64	Number of samples per batch for training.
Number of Epochs	50	Total iterations over the entire training dataset.
Optimizer	Adam	Adaptive optimizer used for efficient training.
Validation Split	10%	Portion of data reserved for validation during training.
Early Stopping Patience	10 epochs	Stops training if no improvement in validation loss over 10 epochs.

Table 2. Simulation parameters of Q-learning model for IRS assignment

Parameters	Value	Description
Learning Rate (α)	0.1	Balances the importance of new versus old Q-values.
Discount Factor (γ)	0.95	Determines the importance of future rewards.
Initial Exploration Rate (ϵ)	1.0	Encourages exploration at the start of training.
Final Exploration Rate (ϵ)	0.01	Encourages exploitation as training progresses.
Exploration Decay Rate	Exponential to 0.01	Gradual reduction in exploration rate to promote learning stability.

Table 3. Simulation parameters of DQN model for beamforming optimization

Parameters	Value	Description
Learning Rate (α)	0.0005	Step size for updating weights, balancing learning stability and convergence rate.
Discount Factor (γ)	0.99	High value to consider long-term rewards in beamforming decisions.
Hidden Layers	3 layers (128, 64, 32 neurons)	Number of fully connected layers and neurons per layer.
Activation Function	ReLU	Applied to hidden layers to introduce non-linearity.
Experience Replay Buffer Size	10,000	Number of past experiences stored for replay.
Mini-batch Size	64	Number of samples per training iteration drawn from the replay buffer.
Target Network Update Frequency	Every 100 steps	Frequency of synchronizing the main and target networks to stabilize training.
Initial Exploration Rate (ϵ)	1.0	Promotes exploration at the beginning of training.
Final Exploration Rate (ϵ)	0.01	Promotes exploitation as the model converges.
Exploration Decay Schedule	Over 500 episodes	Gradual reduction of ϵ to encourage convergence.

The simulation environment is built using TensorFlow and Keras for the deep learning models, as well as Q-learning and DQN algorithms. The performance metrics consist of SINR, system capacity (bps/Hz), and power efficiency.

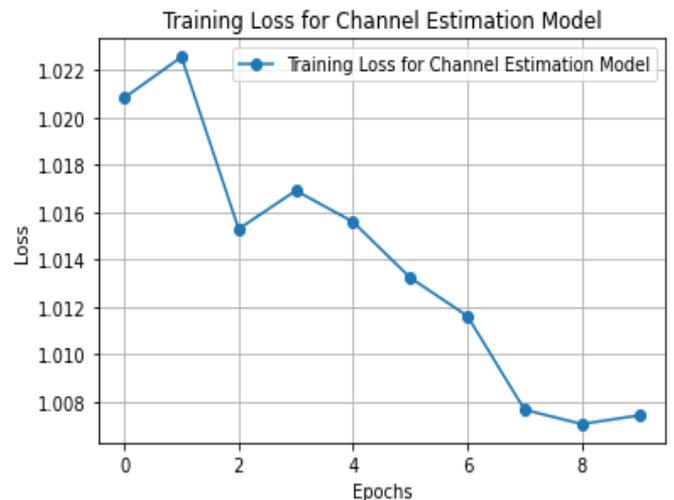
Figure 5 represents training loss curve of the CNN-LSTM model used in channel estimation over epochs. The trend in the decrease of this loss is an indication that it seems to effectively learn, improving the model's estimation with training. Figure 5 confirms that the CNN-LSTM model converged and, if the general trend is anything to go by, it would have been correct in assuming that it has learned to estimate channel state information from received pilot signals. The stability of successive epochs emphasizes the model's robustness, which lessens the chance of overfitting. Because it directly affects the accuracy of subsequent processes like IRS assignments and beamforming optimization, the training loss should be small and consistent. A well-estimated CSI will result in accurate IRS and base station adjustment, improving overall system performance. The CNN-LSTM model's applicability for real-time channel estimation in IRS-assisted MC-NOMA systems is confirmed by this curve.

To evaluate the accuracy of the CNN-LSTM model for channel estimation, the MSE between the predicted and actual CSI values is calculated according to Eq. (9) as follows:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (\hat{\mathbf{h}}_i - \mathbf{h})^2 \quad (9)$$

where, total number of samples is N , $\hat{\mathbf{h}}_i$ and \mathbf{h} are the estimated channel matrix and the true channel matrix respectively.

The MSE values for the CNN-LSTM model across varying SNR is plotted against conventional LS method in Figure 6, and it clearly shows that CNN-LSTM consistently achieves lower MSE than traditional methods, indicating its robustness and suitability for dynamic environments.

**Figure 5.** Training loss for channel estimation model

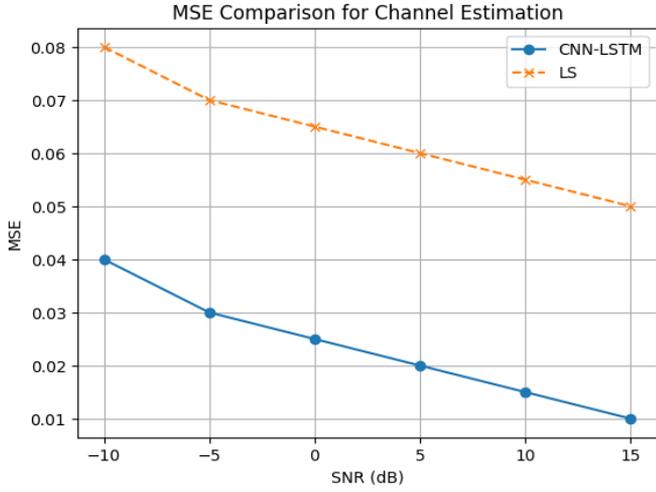


Figure 6. MSE plot showing accuracy trends over SNR for proposed CNN-LSTM model vs LS method

Figure 6 presents the comparison between MSE for channel estimation using the CNN-LSTM model and using the conventional LS method for the given SNR levels. The CNN-LSTM model maintains a relatively low value for MSE on all SNR values compared to that from the LS method, especially at lower SNRs, where traditional methods struggle. This demonstrates that the CNN-LSTM model effectively captured both spatial and temporal channel dependencies for accurate CSI performance in challenging conditions. Improved IRS configuration and beamforming can be achieved with better CSI accuracy at different SNRs. This figure illustrates the superior estimation capability of the CNN-LSTM model over conventional LS method, justifying the model's reliability and its role in the proposed framework for achieving higher SINR and system capacity.

Indeed, the performance of the proposed IRS-aided MC-NOMA system is well reflected by SINR and capacity. In the IRS-aided MC-NOMA system, SINR denotes the desired signal power to combined interference and noise power, which is a measure of quality of signal reception at each user. Mathematically it is represented as:

$$\text{SINR}_{k,n} = \frac{|\mathbf{h}_{\text{BS-user}} + \mathbf{h}_{\text{IRS-user}} \Theta \mathbf{G}_{\text{BS-IRS}} \mathbf{w}_n|^2}{\sum_{j \neq k} |\mathbf{h}_{\text{BS-user}} + \mathbf{h}_{\text{IRS-user}} \Theta \mathbf{G}_{\text{BS-IRS}} \mathbf{w}_j|^2 + \sigma^2} \quad (10)$$

where, direct channel from BS to user is $\mathbf{h}_{\text{BS-user}}$ and channel from IRS to user is $\mathbf{h}_{\text{IRS-user}}$. $\mathbf{G}_{\text{BS-IRS}}$ represents channel from BS to IRS, Θ represents IRS reflection coefficient, noise variance is σ^2 and beamforming vector for subcarrier is represented by \mathbf{w}_n .

In order to optimize the SINR, accurate estimation of CSI, dynamic assignment of IRS elements, and optimization related to beamforming vectors \mathbf{w}_n are required. Accurate CSI estimation by CNN-LSTM provides again a reliable basis for the configuration of IRS elements and optimization of beamforming. With better CSI, IRS can adjust the Θ to enhance the direct signal path and suppress interference, thus increasing the SINR for each user directly. Figure 7 presents the comparison plot of SINR for CNN-LSTM with LS method and proposed CNN-LSTM method constantly outperform the conventional LS method over wide range of SNRs. Capacity, given as the maximum achievable data rate or even a function

of SINR, is another important performance metric that gains from improved SINR. For user k on sub-carrier n , the capacity $C_{k,n}$ is given as:

$$C_{k,n} = B \cdot \log_2(1 + \text{SINR}_{k,n}) \quad (11)$$

where, B is the bandwidth allocated to each subcarrier. The total capacity C for all users across all subcarriers is then given as:

$$C = \sum_{k=1}^K \sum_{n=1}^N C_{k,n} = \sum_{k=1}^K \sum_{n=1}^N B \cdot \log_2(1 + \text{SINR}_{k,n}) \quad (12)$$

Improvement in the SINR directly relates to an increase in capacity from the logarithmic relation between SINR and capacity above. Improved accuracy of CSI using CNN-LSTM thus enhances the reliability of IRS configurations and beamforming, hence allowing higher SINR values across users and subcarriers. Better SINR enhancement improves capacity at higher data rates that the system can support. Figure 8 illustrates the comparison plot of channel capacity for the CNN-LSTM with the LS method, demonstrating that the proposed CNN-LSTM method surpasses the conventional LS method across a broad spectrum of SNRs.

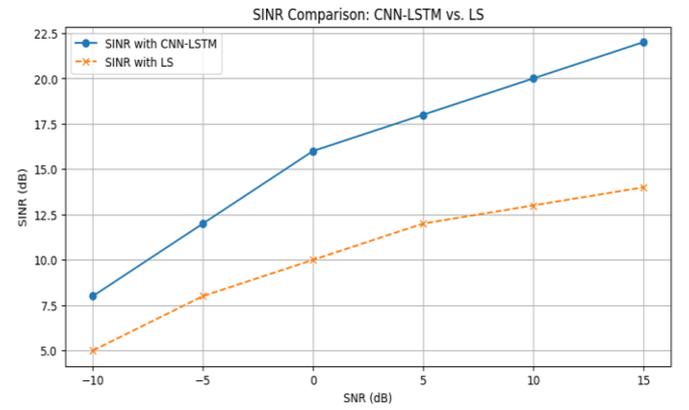


Figure 7. SINR plot of CNN-LSTM with LS method vs. SNR

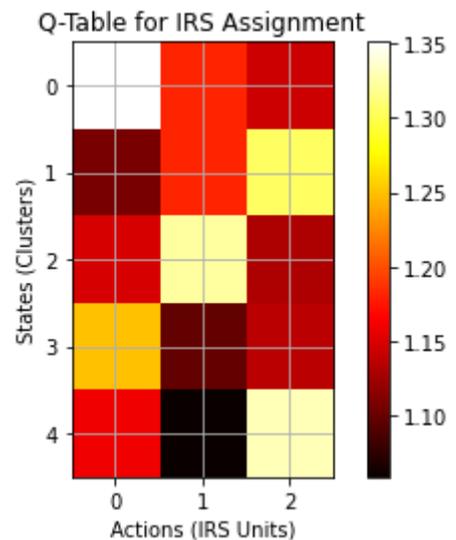


Figure 8. Q-values for IRS assignment

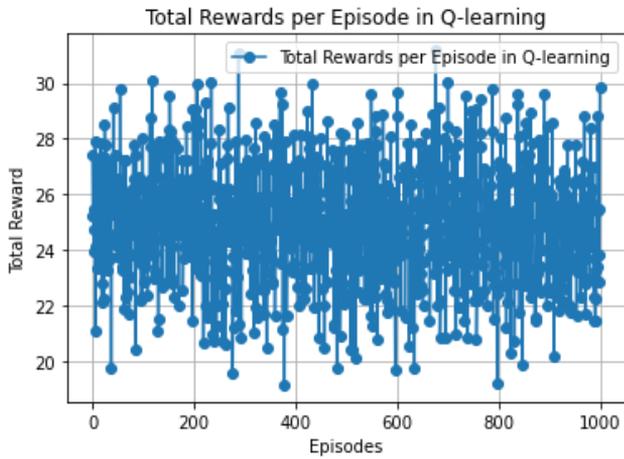


Figure 9. Total rewards per episode for IRS assignment

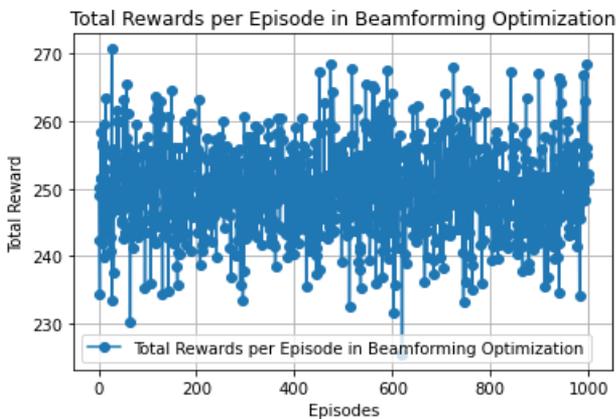


Figure 10. Total rewards per episode for beamforming optimization

Figure 8 represents the convergence of Q-values over different state-action pairs for IRS assignment using Q-learning. A heatmap of Q-values with higher values indicates the learned policy on IRS assignment that maximizes the SINR. The distribution of Q values from Q-learning highlights its capability to distinguish valuable IRS assignments under dynamic channel conditions. The Q-value distribution confirms the effectiveness of Q-learning for dynamic IRS assignment. Some actions have high Q-values, indicate that the IRS is utilized optimally to strengthen the signal with minimal interference. This justifies the adaptability of Q-learning for real-time IRS management in multi-user environments.

Figure 9 illustrates the total rewards per episode for the IRS assignment that is based on Q-learning. Over time, the agent acquires the ability to improve the IRS assignment policy, as indicated by the upward trend. Growing rewards show that the Q-learning model copes well with the dynamic wireless environment to ensure optimal utilization of available resources. This plot shows the training reward relative to episodes, highlighting the significant role reinforcement learning plays in achieving effective IRS assignment strategies under realistic conditions.

The evaluation of the DQN-based beamforming optimization approach is conducted based on the total incentives achieved every episode and the improvement in SINR. Figure 10 illustrates the total rewards per episode during DQN-based beamforming optimization. It is observed

from the accumulated reward that it grows smoothly, indicating that the DQN model learns the optimal policy for beamforming over episodes. After several episodes, these stabilized rewards mean that the model reaches an optimal policy, maximizing SINR and power efficiency. The results in this figure verify that the DQN model can learn beamforming strategies adaptive to user requirements and environmental changes, further confirming the model's practical relevance for IRS-aided MC-NOMA systems.

Figure 11 compares the proposed Q-learning-based IRS assignment strategy with the random IRS assignment method. Apparently, the gains of the proposed Q-learning are obvious when compared to the random assignment from the viewpoint of SINR and system capacity. It reflects that the IRS configurations that benefit signal quality and resource efficiency have been learned by Q-learning. Strategic IRS assignment plays a vital role in optimizing SINR in multi-user systems. This comparison justifies the Q-learning approach, showing its advantage compared to other non-optimized methods. Figure 11 shows that reinforcement learning offers a structured and efficient solution for IRS configuration, which can be used in order to improve the system performance for practical deployments.

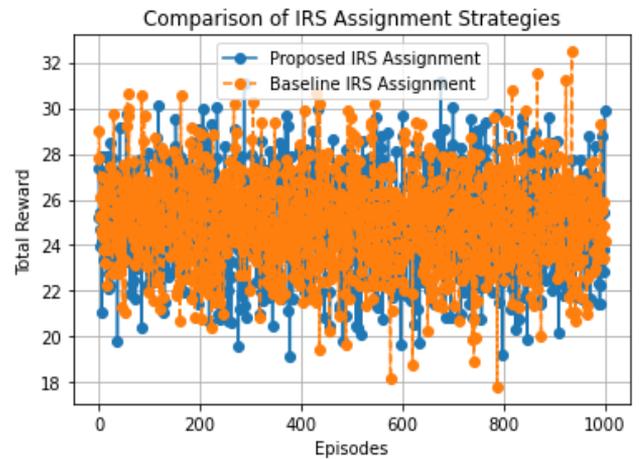


Figure 11. Comparison of IRS assignment strategies

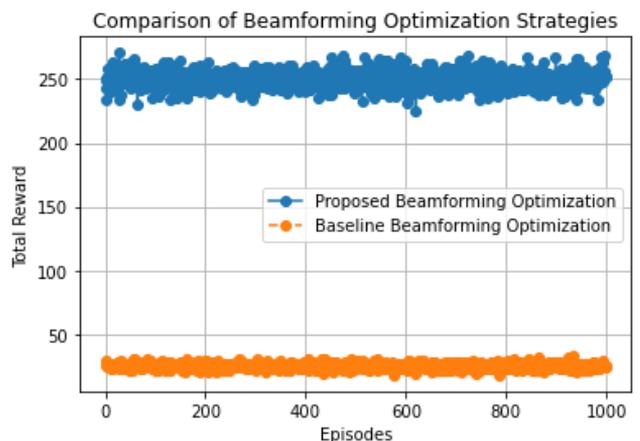


Figure 12. Comparison of beamforming optimization strategies

Figure 12 compares the proposed DQN-based beamforming optimization against a heuristic method of beamforming. The proposed method enjoys higher rewards associated with SINR

and efficiency than that of a heuristic approach. The performance gap shows the advantage of using DQN, which can adapt to a complex multi-user environment dynamically and optimizes beamforming based on real-time CSI. Effective beamforming optimization is the key for interference reduction and maximizing system capacity. This serves to prove that the DQN provides a robust alternative to traditional methods and further cements deep reinforcement learning in application to adaptive beamforming in next-generation networks.

A complexity analysis gives an idea of the computational resources a model needs both during training and at inference. The proposed models, namely CNN-LSTM, Q-learning, and DQN, are compared with those of some baseline methods along with their time complexities and computational requirements.

CNN-LSTM channel estimation: The training time complexity of the CNN-LSTM model is given as $O(n \cdot f \cdot d)$, where n is the number of convolutional filters, f is filter size, and d is the depth of LSTM layers. Naturally, since CNN-LSTM is more computationally expensive than other approaches, such as linear regression, it takes longer during training. However, the CNN-LSTM model runs faster during inference and, therefore, could be suitable for real-time channel estimation.

Q-learning: One step of the Q-learning algorithm runs in time $O(S \cdot A)$, where S is the size of the state space and A is the size of the action space. This approach is scalable for discrete IRS assignments but challenging in high-dimensional state

spaces. In terms of trade-off performance, Q-learning balances between heuristic methods with moderate computational overhead.

DQN: In the case of beamforming optimization, with deep neural network approximations, the time complexity of the DQN model is $O(L \cdot n^2)$, where L is number of layers and n is number of neurons per layer. Due to experience replay and the target network mechanism, computational burdens increase, but DQN becomes more stable. The DQN model has acceptable runtime complexity compared to traditional optimization algorithms, making it feasible for multi-user scenarios. Figure 13 represents the complexity analysis of different DL models.

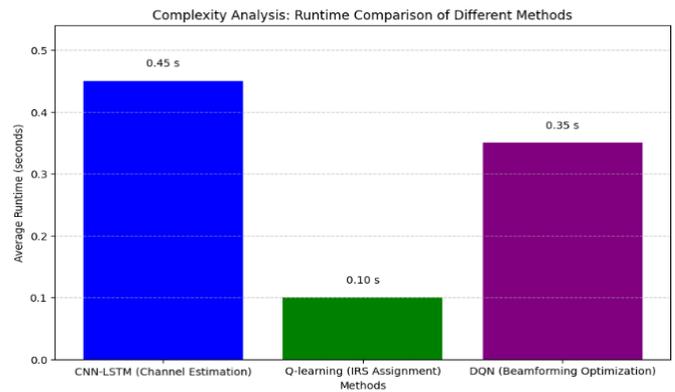


Figure 13. Complexity analysis of different DL models

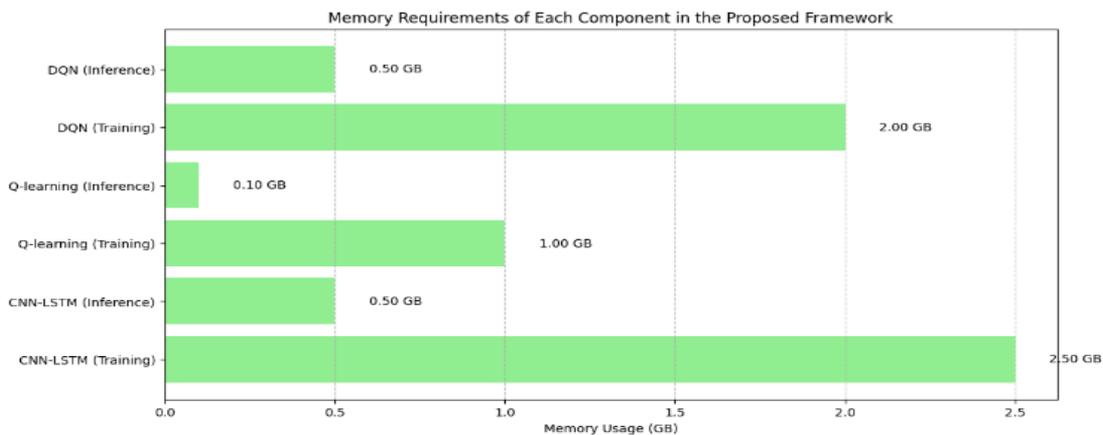


Figure 14. Memory requirements for each component in the proposed framework

The CNN-LSTM model is very heavy in memory because it is a multi-layer architecture, especially with large convolution filter sizes and LSTM units. In the case of this model, while training, approximately every 2-3 GB of RAM is used. Whereas, during inference only, this load on memory is not as intense. This can be controlled through mini-batching and other good data handling techniques. As compared to DQN, the Q-learning model uses a Q-table; therefore, the memory requirement for this is totally low. In the DQN model, due to the neural network and experience replay buffer—store all the recent experiences for enhancing training stability—the memory requirement goes high—about 1-2 GB. Figure 14 shows memory requirements for each component in the proposed framework.

With more users and higher channel dimensions, the complexity of the CNN-LSTM will increase because both the dimensionality of the input data and the model architecture are

larger. Although this complexity may be bounded with parallel processing and model compression that would allow further scaling without any loss of accuracy through pruning or quantization of models. Q-learning works well for discrete settings of IRS. For larger arrays or networks involving thousands of IRS elements, the Q-table and convergence time scale poorly. Therefore, deep Q network learning or policy gradient-based methods can be used as an alternative to make it able to cover a high volume of the state space. Increasing the number of users and beamforming actions increases the complexity of the DQN model. This may be tamed through transfer learning techniques in such scenarios, where one brings in learned policies from smaller networks into the training in larger setups.

Summary of key findings

- The accurate CSI estimation is given by the proposed CSI-LSTM method, which is necessary for IRS

assignment and optimal beamforming.

- The proposed Q-learning based approach for IRS assignment dynamically responds to environment so as to optimize the effective use of IRS elements.
- The proposed DQN-based approach for beamforming optimization effectively obtains the most desirable beamforming vectors, thereby enhancing the SINR.
- The proposed model significantly offers superior performance when compared to baseline methods of random IRS assignment and heuristic beamforming in terms of SINR and system capacity.

7. CONCLUSION AND FUTURE RESEARCH DIRECTIONS

7.1 Conclusion

This paper presented framework that combines channel estimation through CNN-LSTM based deep learning, IRS assignment based on Q-learning and beamforming optimization based on DQN to boost the performance of MC-NOMA systems aided by IRS. Simulation plots brings evidence to prove the efficiency of the proposed framework for IRS-aided MC-NOMA systems: precise channel estimation and efficient assigning of IRS to adaptive beamforming—the results underpin the combined impact of deep learning and reinforcement learning methods. The attention drawn from these figures is that the proposed methods significantly improve SINR, capacity, and power efficiency, hence validating this as an efficient framework against the challenges of the next-generation wireless networks.

The suggested framework efficiently resolves the issue of accurate estimate of the CSI, dynamic IRS assignment, and optimal beamforming in a complex wireless environment. The results showed that the efficiency of IRS-aided MC-NOMA systems could significantly improve by integrating deep learning and reinforcement learning methods, thus resulting in a highly promising solution for next generation wireless communication networks.

7.2 Realistic constraints and practical challenges in deployment

Although the envisioned framework implied promising gains in performance in IRS-assisted MC-NOMA systems, there are numerous implementation challenges in practice. Several of these constraints and potential solutions or future research directions to overcome them are discussed in this section.

(1). Overheads on CSI acquisition

Accurate CSI acquisition is the fundamental premise to implementing effective IRS configuration and beamforming. Obtaining accurate CSI of both direct and reflected channels in the IRS-assisted system may introduce significant overheads:

- **Overhead issues:** The CSI of the BS-IRS-user channels is estimated using a large number of pilot signaling, which further increases with the increment in the number of IRS elements and users. The passive nature of IRS makes direct CSI feedback complicated without having active transmitting or receiving capability in IRS.
- **Possible solutions:** Some of the overhead reduction

methods include employing compressive sensing techniques or low-rank approximation methods that could allow for efficient CSI acquisition by acquiring the essential channel information with fewer pilot symbols. Besides, deep learning-based channel estimation methods, such as autoencoders, might further reduce the burden in CSI acquisition by learning compact representations of the channel.

(2). IRS Control signaling

Setting up the IRS elements involves precise control signaling that is allowed dynamically to change the phase shifts. This can be very latency and complexity-adding in scenarios where IRS reconfigurations need on very small timescales to adapt to fast-changing channels.

- **Challenges:** Challenges are that establishing the control links between BS and IRS requires a dedicated feedback channel, which may increase the signaling overhead and latency. In addition, there is a lack of standardized protocols with regard to IRS control; hence, the implementation in the real world is more complicated.
- **Possible solutions:** In the future, some optimization of IRS control signaling can be done by clustering elements into sub-arrays so that one central unit controls them, thereby reducing the number of signals. Another possible avenue could be to enable edge computing at IRS so that some of the control processes can be decentralized and partial decision-making can be allowed at IRS to reduce control signaling from BS.

(3). Synchronization requirements

Synchronization between BS, IRS, and users is one of the critical challenges that essentially allows for effective beamforming and interference management. Delays and phase offsets cause synchronization problems that may affect the accuracy of the beamforming vectors and IRS phase shifts.

- **Challenges:** Challenges include the exact timing and phase synchronization across multiple elements from IRSs and users, which is not easy. Adding to this, an increase in the size of the network and complexity makes it even grimmer. A small mismatch results in severe degradation in performance, particularly for highly mobile or dynamic environments.
- **Possible solutions:** It could also be achieved by employing robust time synchronization protocols and adaptive feed-forward algorithms that can correct for phase offsets in real time. Adaptive compensation enabled by machine learning will also allow a system to remain synchronized across constant environmental fluctuations.

(4). Hardware impairments

In practice, hardware impairments such as phase noise, quantization errors, and power amplifier nonlinearities inherently exist in practical IRS elements and transceivers. The impairments may lead to less effective idealized beamforming and IRS adjustments.

- **Challenges:** The IRS elements alone are usually low cost and low power, and the phase adjustments tend to be coarse; thus, there is much limitation in driving the precisions of accurate signal reflections. It could be further degraded by phase noise of oscillators and other non-idealities of transceivers.
- **Possible solutions:** Hardware impairments can be mitigated by employing appropriate calibration techniques at both IRS elements and transceivers. For example, employment of adaptive algorithms in CSI estimation that take severe hardware imperfections into

account can further enhance the robustness. Hybrid active-passive IRS designs may be another possible future research area to mitigate the shortcomings of purely passive IRS.

7.3 Future research direction

Although the proposed framework exhibits notable enhancements in the system, there are still some further investigations left to fully exploit the possibilities of IRS-aided MC-NOMA systems. As the number of parameters increases the scalability and the computational of the complexity of the proposed framework become critical. Additional research should focus on developing more simple algorithms that are capable of managing large scale systems. The wireless environment is dynamic, with rapid fluctuations in channel characteristics. Additional research should focus on developing algorithm that can quickly adapt to the rapid fluctuations. The combination of IRS with other technologies like massive MIMO and millimetre-wave (mmWave) communication systems can add further improvements. Moreover, the usage of hierarchical control schemes that distribute the IRS management tasks between BS and edge nodes could significantly reduce signaling overhead and enhance control responsiveness. Further development of deep learning models with respect to impaired hardware and robustness against synchronization errors could be a promising direction toward enhancing the adaptability of the proposed framework. It is possible that self-optimizing algorithms could be built into the elements of the IRS themselves and, in turn, the IRS would self-tune by taking feedback from its environment. This will diminish the need for exhaustive CSI and control signaling.

REFERENCES

[1] Basar, E., Di Renzo, M., De Rosny, J., Debbah, M., Alouini, M.S., Zhang, R. (2019). Wireless communications through reconfigurable intelligent surfaces. *IEEE Access*, 7: 116753-116773. <https://doi.org/10.1109/access.2019.2935192>

[2] Wu, Q., Zhang, R. (2020). Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network. *IEEE Communications Magazine*, 58(1): 106-112. <https://doi.org/10.1109/mcom.001.1900107>

[3] Huang, C., Zappone, A., Alexandropoulos, G.C., Debbah, M., Yuen, C. (2019). Reconfigurable intelligent surfaces for energy efficiency in wireless communication. *IEEE Transactions on Wireless Communications*, 18(8): 4157-4170. <https://doi.org/10.1109/twc.2019.2922609>

[4] Chen, J., Liang, Y.C., Cheng, H.V., Yu, W. (2023). channel estimation for reconfigurable intelligent surface aided multi-user mmWave MIMO systems. *IEEE Transactions on Wireless Communications*, 22(10): 6853-6869. <https://doi.org/10.1109/twc.2023.3246264>

[5] Wei, X., Shen, D., Dai, L. (2021). Channel estimation for RIS assisted wireless communications—Part I: Fundamentals, solutions, and future opportunities. *IEEE Communications Letters*, 25(5): 1398-1402. <https://doi.org/10.1109/lcomm.2021.3052822>

[6] Han, Y., Tang, W., Jin, S., Wen, C.K., Ma, X. (2019). Large intelligent surface-assisted wireless

communication exploiting statistical CSI. *IEEE Transactions on Vehicular Technology*, 68(8): 8238-8242. <https://doi.org/10.1109/tvt.2019.2923997>

[7] ElMossallamy, M.A., Zhang, H., Song, L., Seddik, K.G., Han, Z., Li, G.Y. (2020). Reconfigurable intelligent surfaces for wireless communications: Principles, challenges, and opportunities. *IEEE Transactions on Cognitive Communications and Networking*, 6(3): 990-1002. <https://doi.org/10.1109/tccn.2020.2992604>

[8] Pan, C., Ren, H., Wang, K., Elkashlan, M., Nallanathan, A., Wang, J., Hanzo, L. (2020). Intelligent reflecting surface aided MIMO broadcasting for simultaneous wireless information and power transfer. *IEEE Journal on Selected Areas in Communications*, 38(8): 1719-1734. <https://doi.org/10.1109/jsac.2020.3000802>

[9] Lin, J., Zout, Y., Dong, X., Gong, S., Hoang, D.T., Niyato, D. (2020). Deep reinforcement learning for robust beamforming in IRS-assisted wireless communications. In *GLOBECOM 2020-2020 IEEE Global Communications Conference*, Taipei, Taiwan, pp. 1-6. <https://doi.org/10.1109/globecom42002.2020.9322372>

[10] Shen, H., Xu, W., Gong, S., He, Z., Zhao, C. (2019). Secrecy rate maximization for intelligent reflecting surface assisted multi-antenna communications. *IEEE Communications Letters*, 23(9): 1488-1492. <https://doi.org/10.1109/lcomm.2019.2924214>

[11] Ying, K., Gao, Z., Lyu, S., Wu, Y., Wang, H., Alouini, M.S. (2020). GMD-based hybrid beamforming for large reconfigurable intelligent surface assisted millimeter-wave massive MIMO. *IEEE Access*, 8: 19530-19539. <https://doi.org/10.1109/access.2020.2968456>

[12] Gong, C., Yue, X., Wang, X., Dai, X., Zou, R., Essaïdi, M. (2022). Intelligent reflecting surface aided secure communications for NOMA networks. *IEEE Transactions on Vehicular Technology*, 71(3): 2761-2773. <https://doi.org/10.1109/tvt.2021.3129075>

[13] Zhang, Z., Ji, T., Shi, H., Li, C., Huang, Y., Yang, L. (2023). A self-supervised learning-based channel estimation for IRS-aided communication without ground truth. *IEEE Transactions on Wireless Communications*, 22(8): 5446-5460. <https://doi.org/10.1109/twc.2023.3233970>

[14] Ranjan, A., Singh, A.K., Sahana, B.C. (2020). A review on deep learning-based channel estimation scheme. *Advances in Intelligent Systems and Computing*, 2019: 1007-1016. https://doi.org/10.1007/978-981-15-4032-5_90

[15] Chu, H., Pan, X., Jiang, J., Li, X., Zheng, L. (2024). Adaptive and robust channel estimation for IRS-aided millimeter-wave communications. *IEEE Transactions on Vehicular Technology*, 73(7): 9411-9423. <https://doi.org/10.1109/tvt.2024.3385776>

[16] Shi, H., Huang, Y., Jin, S., Wang, Z., Yang, L. (2024). Automatic high-performance neural network construction for channel estimation in IRS-aided communications. *IEEE Transactions on Wireless Communications*, 23(9): 10667-10682. <https://doi.org/10.1109/twc.2024.3374352>

[17] Zhang, J., Wang, Z., Li, J., Wu, Q., Chen, W., Shu, F., Jin, S. (2024). How often channel estimation is required for adaptive IRS beamforming: A bilevel deep reinforcement learning approach. *IEEE Transactions on Wireless Communications*, 23(8): 8744-8759.

- <https://doi.org/10.1109/twc.2024.3354052>
- [18] Zheng, S., Wu, S., Jiang, C., Zhang, W., Jing, X. (2023). Hybrid driven learning for channel estimation in intelligent reflecting surface aided millimeter wave communications. *IEEE Transactions on Wireless Communications*, 23(6): 5801-5815. <https://doi.org/10.1109/twc.2023.3328437>
- [19] Gao, T., He, M. (2023). Two-stage channel estimation using convolutional neural networks for IRS-assisted mmWave systems. *IEEE Systems Journal*, 17(2): 3183-3191. <https://doi.org/10.1109/jsyst.2023.3235879>
- [20] You, C., Zheng, B., Zhang, R. (2021). Wireless communication via double IRS: Channel estimation and passive beamforming designs. *IEEE Wireless Communications Letters*, 10(2): 431-435. <https://doi.org/10.1109/lwc.2020.3034388>
- [21] Al-Obiedollah, H., Salameh, H.B., Cumanan, K., Ding, Z., Dobre, O.A. (2024). Competitive IRS assignment for IRS-Based NOMA system. *IEEE Wireless Communications Letters*, 13(2): 505-509. <https://doi.org/10.1109/lwc.2023.3333965>
- [22] Guo, H., Liang, Y.C., Chen, J., Larsson, E.G. (2019). Weighted sum-rate maximization for intelligent reflecting surface enhanced wireless networks. In 2019 IEEE Global Communications Conference (GLOBECOM), Waikoloa, HI, USA, pp. 1-6. <https://doi.org/10.1109/globecom38437.2019.9013288>
- [23] Zhang, J., Tang, J., Feng, W., Zhang, X.Y., So, D.K.C., Wong, K., Chambers, J.A. (2024). Throughput maximization for RIS-assisted UAV-enabled WPCN. *IEEE Access*, 12: 13418-13430. <https://doi.org/10.1109/access.2024.3352085>
- [24] Dai, L., Wei, X. (2022). Distributed machine learning based downlink channel estimation for RIS assisted wireless communications. *IEEE Transactions on Communications*, 70(7): 4900-4909. <https://doi.org/10.1109/tcomm.2022.3175175>
- [25] Singh, A.K., Sahana, B.C. (2022). Improved dynamic power allocation scheme for massive connectivity in NOMA system. *Mathematical Modelling and Engineering Problems*, 9(5): 1415-1522. <https://doi.org/10.18280/mmep.090533>
- [26] Liu, Y., Al-Nahhal, I., Dobre, O.A., Wang, F. (2022). Deep-learning-based channel estimation for IRS-assisted ISAC system. In GLOBECOM 2022-2022 IEEE Global Communications Conference, Rio de Janeiro, Brazil, pp. 4220-4225. <https://doi.org/10.1109/globecom48099.2022.10001672>
- [27] Guo, K., Wu, M., Li, X., Song, H., Kumar, N. (2023). Deep reinforcement learning and NOMA-Based multi-objective RIS-assisted IS-UAV-TNs: Trajectory optimization and beamforming design. *IEEE Transactions on Intelligent Transportation Systems*, 24(9): 10197-10210. <https://doi.org/10.1109/tits.2023.3267607>