Incorporating Machine Learning Algorithms for Channel Prediction and Power Optimization Scheme in Power Domain Non-Orthogonal Multiple Access System

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

by

Mohamed Gaballa

Department of Electronic and Electrical Engineering

College of Engineering, Design and Physical Sciences

Brunel University London

United Kingdom

August 2023

Abstract

With the fast increase in the publicity of the Internet of Things (IoT) and cloud computing, the requirement for massive connectivity and highly reliable data rates is increasing day by day for communication networks. IoT can establish the connections among many types of smart devices, such as smart sensors, robots, and mobile devices. To satisfy these demands and such massive connectivity, three main services have been presented to the communication networks. These key services consist of massive machine type communication (mMTC) that permits massive connections between IoT terminals, enhanced mobile broadband (eMBB) that delivers a high data rate for mobile devices, and ultra-reliable and low-latency communication (URLLC) that confirms reliability and minimum latency for critical and sensible applications. These services are characterized by their quality of service (QoS), where URLLC has a stringent QoS policy for high reliability and low latency application, eMBB service is categorized by a moderate QoS policy, while mMTC has no precise QoS policy. These types of QoS are usually difficult to realize with the traditional orthogonal multiple access (OMA) due to limited spectrum resources, and delays. To satisfy and enhance these diverse QoS requirements, many potential multiple access schemes have been introduced into communication network. Among them, is the non-orthogonal multiple access (NOMA) scheme that has a achieved a popularity because it can support massive connectivity with limited resources, tolerable transmission delays, and high spectral efficiency. The key feature of NOMA is that multiple user devices can be served from the same radio resource block, such as time, frequency, and code. NOMA scheme applies superposition coding to combine signals related to multiple users at the transmitter side and implements successive interference cancellation procedure to differentiate and recover the signals of multiple devices at the receiver side.

There are some challenges related to resource allocation in NOMA system, such as power allocation and channel estimation. Machine learning has obtained publicity over the past several years, and many machine learning models and algorithms have been industrialized. Deep learning is a subset of machine learning, and it has distinct advantages over traditional machine learning methods, such as being capable of working on huge volumes of data in complex networks. Furthermore, reinforcement learning also is a type of machine learning, and the main aim of reinforcement learning is to train an agent to carry out a certain task within an

uncertain environment. Deep learning and reinforcement learning approaches can be investigated and inspected to be one of the candidate's algorithms for resource allocation and channel estimation in NOMA system. In this thesis, simulation results clearly indicates that deep learning and reinforcement learning algorithms can provide a superior improvement in terms of diverse performance metrics when Rayleigh and Rician fading channels are considered. Also, in this thesis, a benchmark schemes are also simulated to highlight how much enhancement has been achieved in the system performance when our proposed machine learning models are applied compared to the results obtained by the benchmark schemes.

Publications Based on this Research

1. Gaballa, Mohamed, and Maysam Abbod. 2023. "Simplified Deep Reinforcement Learning Approach for Channel Prediction in Power Domain NOMA System" *Sensors* 23, no. 21: 9010. https://doi.org/10.3390/s23219010

2. Gaballa, Mohamed, Maysam Abbod, and Ammar Aldallal. 2023. "A Study on the Impact of Integrating Reinforcement Learning for Channel Prediction and Power Allocation Scheme in MISO-NOMA System" Sensors 23, no. 3: 1383. https://doi.org/10.3390/s23031383

3. Gaballa, Mohamed, Maysam Abbod, and Ammar Aldallal. 2022. "Investigating the Combination of Deep Learning for Channel Estimation and Power Optimization in a Non-Orthogonal Multiple Access System" Sensors 22, no. 10: 3666. https://doi.org/10.3390/s22103666

4. M. Gaballa, M. Abbod and S. Alnasur, "Hybrid Deep Learning for Channel Estimation and Power Allocation for MISO-NOMA System," 2022 IEEE Future Networks World Forum (FNWF), Montreal, QC, Canada, 2022, pp. 361-366, doi: 10.1109/FNWF55208.2022.00070.

5. M. Gaballa, M. Abbod and A. Aldallal, "Deep Learning and Power Allocation Analysis in NOMA System," 2022 Thirteenth International Conference on Ubiquitous and Future Networks (ICUFN), Barcelona, Spain, 2022, pp. 196-201, doi: 10.1109/ICUFN55119.2022.9829643.

6. M. Gaballa, M. Abbod and A. Jameel, "Power Optimization Analysis using Throughput Maximization in MISO Non-Orthogonal Multiple Access System," 2021 IEEE Globecom Workshops (GC Wkshps), Madrid, Spain, 2021, pp. 1-6, doi: 10.1109/GCWkshps52748.2021.9682080.

7. M. Gaballa, M. Abbod, A. Jameel and N. Khaled, "Throughput Maximization & Power Optimization Analysis in Non-Orthogonal Multiple Access System," 2021 IEEE 4th 5G World Forum (5GWF), Montreal, QC, Canada, 2021, pp. 82-87, doi: 10.1109/5GWF52925.2021.00022.

8. M. Gaballa, M. Abbod and M. Albasman, "Power Allocation & MRC Analysis for Single Input Multi Output Non-Orthogonal Multiple Access System," 2021 IEEE International Conferences on Internet of Things (iThings) and IEEE Green Computing & Communications (GreenCom) and IEEE Cyber, Physical & Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics), Melbourne, Australia, 2021, pp. 168-173, doi: 10.1109/iThings-GreenCom-CPSCom-SmartData-Cybermatics53846.2021.00038.

Declaration

It is hereby declared that the thesis in focus is the author's own work and is submitted for the first time to the Post Graduate Research Office. The study was originated, composed, and reviewed by the mentioned author in the Department of Electronic and Electrical Engineering, College of Engineering, Design and Physical Sciences, Brunel University London, UK. All the information derived from other works has been properly referenced and acknowledged.

Mohamed Gaballa

August 2023

London, UK

Acknowledgements

Firstly, I would like to express my sincere appreciation and gratitude to all those who made this thesis possible. This work would not have been possible without help, support and strong motivation of my principal supervisor, Prof. Maysam Abbod, who gave me the encouragement, guidance, and support from the beginning to the end made my work much easier and more enjoyable.

I would also like to thank my second supervisor Dr. Ammar Aldallal for his fascinating discussions about the area and the work undertaken.

Lastly, I thank my parents and my family for their support, love, and encouragement throughout the period. Also, I express my sincere thanks to all my teachers, relatives, colleagues, elders, and all thosefrom whom I have learnt and gained knowledge.

Thesis Contents

X

List of Figures

XIII

List of Tables

List of Abbreviations

6G	Sixth Generation
AWGN	Additive White Gaussian Noise
BER	Bit error rate
BS	Base station
CSI	Channel state information
DL	Deep Learning
DNN	Deep Neural Network
DRL	Deep Reinforcement Learning
eMBB	enhanced mobile broadband
FPA	Fixed Power allocation
GRU	Gated Recurrent unit
HetNet	Heterogeneous Networks
IoT	Internet of Things
ITU	International Telecommunication Union
ITU	International Telecommunication Union
KKT	Karush-Kuhn–Tucker
LSTM	Long-short term memory
LTE	Long term evolution
MEC	Mobile Edge Computing
MIMO	Multi Input – Multi Output
MISO	Multi-input single output
ML	Machine learning
MMSE	Minimum mean square error
mMTC	Massive machine type communication

MSE	Mean square error
MUD	Multiuser detection
OFDM	Orthogonal Frequency Division Multiplexing
OPS	Optimized power scheme
PD-NOMA	Power domain Non-Orthogonal Multiple Access
QoS	Quality of service
RL	Reinforcement learning
RNN	Recurrent Neural Networks
SARSA	State-Action-Reward-State-Action
SDMA	Space Division Multiple Access
SIC	Successive interference cancellation
SIMO	Single-input multi-output
URLLC	ultra-reliable and low-latency communication

Chapter 1

Introduction

1.1 Multiple Access System and Machine Learning Vision

Non-orthogonal multiple access (NOMA) system is classified as an inspiring multiple access scheme in forthcoming wireless networks toward enhancing the spectral efficiency and system capacity. NOMA can exploit the radio resource block (RRB) by characterizing the fading channels for each user in NOMA cell, to deliver a diverse quality of service (QoS) demands for the examined users in the cell. NOMA system enables various users to get simultaneous access to same time-frequency resource block based on the principle of superposition different signals from different users via the power domain or code domain [1]. In order that the spectrum can be effectively exploited, NOMA concept mainly depends on that, user with bad or weak channel condition and a user that has a good channel condition can both share the same subcarrier at same time slot. In NOMA scheme, each user device can receive the superposition or the multiplexing of signals from all users in the cell, therefore the exclusion of interference from non desired users come to be essential to achieve a managed decoding. Regularly, multiuser (MU) detection in NOMA system can be achieved via the successive interference cancellation (SIC) procedure that can be performed in power domain [2]. In SIC technique, signals from different users can be decoded sequentially on the basis of the availability of the channel parameters and the assigned power. On the other hand, comprehensive knowledge of the channel parameters for each user in NOMA cell is challenging because pilot symbols that mainly employed in channel prediction might interfere with signals from other users, hence disturbing the performance of conventional channel estimation procedures, such as minimum mean square error (MMSE).

In order to realize a high spectral efficiency and enormous connectivity in next generation of wireless networks, Power domain non-orthogonal multiple access (PD-NOMA) is one of the multiple access candidates' schemes, since it can deal with signals that have noticeable distinction in power levels. PD-NOMA also depends on SIC procedure to split up the high-

energy signal at the receiver side and then remove it to leave only the desired signal [3]. The incorporation of NOMA scheme and multiple antenna techniques is considered as a appropriate way towards characterizing the meaningful improvement in the system's data rates [4].

Machine learning (ML) has obtained a public interest over the past several years, and numerous machine learning approaches have been industrialized [5]. One of the popular ML models that used for prediction and forecasting, is artificial neural networks (ANN). Artificial neural network is mainly considered as an information processing structure that is inspired from the methodology that is used by the brain to process the information [6]. Principally, ANN can be considered as a smaller scale model for the neural structure of the brain. Neurons are the elementary components of an ANN that can take two or more inputs and give one output. The output of the neurons is calculated based on a set of weights and a bias applied to the input and these weights and biases can modify the output. When several neurons are organized into one layer, and these layers are connected such that the output of one-layer feeds into the input of the succeeding layer, and this structure can be described as a feedforward ANN.

ANN can find out how to resolve a complex problems based on the experience obtained during the training process for that ANN structure. Throughout the training process, ANN is given a set of certain inputs for processing, and the output from the ANN is then compared with the anticipated output, and then the error function is calculated. Throughout the training iterations, and based on the calculated error function, a backpropagation algorithm will be used to update the weights and biases in the ANN using gradient descent algorithm to reduce the output error [7]. This training process can be repeated many times until the error function is minimized, or the ANN starts to show a good performance. Once the ANN is efficiently trained, the ANN can deal or manage a new set of data and predict the output with relatively high precision. Normally, the ANN can run well when the training process follows to reliable settings or policy. To improve the convergence process in the training process, all the training data firstly need to be processed so that there are no substantial time gaps, and it is more suitable to normalize all the input training data, in order to accelerate the training process.

The utilization of deep neural networks (DNN) or deep learning (DL) in NOMA system, is greatly explored over the last few years. Various kinds of learning algorithms have been suggested and implemented in different types of communication networks to enhance the

system performance. Overall, there are three types of ML algorithms that are mainly discussed in the literature, such as supervised, unsupervised, and reinforcement learning. Supervised learning is extensively exploited to estimate the channel state information (CSI), and localization. Unsupervised learning is frequently used in user clustering. Reinforcement learning is a different branch of ML, and it is considered to have a remarkable impact in wireless networks.

Recently several RL algorithms have been proposed to obviously handle the issues associated with channel state information (CSI), beamforming, and power allocation. Typically, RL is developed on the basis of a Markov Decision Process (MDP) design, with the following basic elements [8]: a state space S, which is the set of all states in the environment and these states can be observed by the agent. An action space A, which is the set of actions that can be decided by the agent at each state. An immediate reward R, which is the direct reward that is given to the agent after selecting an action $a \in A$ to move to the new state $s \in S$. Policy P represents the framework to move from the current state to the new state based on the action that will be decided by the agent. Another essential component in the RL process is the State-action value function $Q(s, a)$, which is mathematically described as the expectation or the average of the cumulative rewards when a certain action $a \in A$ is selected by an agent in the state $s \in S$ when a certain policy is applied. Based on the above, we can consider RL as a method of understanding the agent's interaction in a stochastic environment, and the main aim of the reinforcement learning is to train the agent how to carry out a certain task within an uncertain environment [9].

1.2 Motivations

Many deep neural networks (DNN) approaches have been proposed in the literature to explicitly address the issues associated with resource allocation, channel state information, channel assignment, and signal detection. To the best of my knowledge, there has been no study that has investigated the combination between deep learning based channel estimation process and the optimal power allocation scheme for multiuser detection in a downlink non-orthogonal multiple access (NOMA) system in fading channels. Most of the current works are managing the issues of power optimization and deep learning based channel estimation separately [10].

Furthermore, it is worth mentioning that many of the proposed schemes in the literature that consider ML for channel estimation task are mainly focused on implementing several deep neural networks (DNN), which in turn leads to an increase in the number of hidden layers with a massive number of neurons in each layer. So, one of the significances of this research is to illuminate that we can also eliminate the need for such DNN approaches, and instead, we can adopt the RL based developed Q-learning algorithm to predict the channel coefficients for each user device in the NOMA cell, and at the same time, a notable improvement in system performance and network convergence is realized. It is also worth mentioning that unlike deep learning algorithms, that mainly depend on learning from a training data set, the proposed Q-learning algorithm in our research is developed to dynamically enhance the system performance and adjust to the variations in the channel based on the feedback and interaction with the environment and cell parameters [11].

In addition, several RL algorithms have been suggested in the literature to handle the issues associated with channel state information (CSI), beamforming, and power allocation, and to the best of our knowledge, there is no study that investigates the incorporation between the RL based Q learning algorithm for channel estimation and the optimum power allocation policy, in order that this integrated framework can be exploited for multiuser detection in downlink NOMA system in fading channels.

Recently, deep reinforcement learning based Q network have achieved a remarkable interest by authors `in various fields, and because of that, in our work we also manage to apply the deep reinforcement learning (DRL) based deep Q networks (DQN) as a tool in the channel estimation task. Then, we plan to compare between RL based Q algorithm and DRL based DQN when channel estimation problem is considered, and power optimization policy is also applied.

1.3 Aim of the Research

In this research, initially we aim to discuss and analyse the downlink NOMA system based on sum rate maximization problem with respect to the constraints of total transmitted power budget and minimum transmission rate. To satisfy this aim, we have elaborated and structured a mathematical analysis to formulate the objective function, then Lagrange function and optimality conditions are applied to derive a closed form expression for the optimal power

coefficients that can be assigned for users in the examined NOMA cell [4]. Fading channels may scatter the transmitted signal and initiate dependencies between the scattered samples, this might impact the channel estimation process and therefore influence the reliability of decoding procedure and the signal detection. Therefore, in our research, we also aim to investigate the impact of utilizing different ML approaches such as, DNN, RL based Q algorithm, and DRL based DQN in estimating the channel parameters for users in NOMA cell. In our research scenarios, we will incorporate the DNN, Q algorithm, and DQN approaches into NOMA cell to predict the channel taps that will be utilized by the receiver to recover the desired symbols [10].

Most of the works in the literature are managing the issues of power optimization and deep learning based channel estimation separately. So, in this research will explore how the channel estimation based on different ML approaches such as DNN, Q algorithm, and DQN can be integrated with the power optimization process for multi-user detection in a PD-NOMA system in fading channels [11][12].

1.4 Design Objectives
The objectives of this research can be outlined as follows:

- A structured mathematical analysis will be presented to derive an analytical expression for the optimum power coefficient for each user in the examined NOMA system based on maximizing the sum rates for all users in the NOMA cell.

- In this research, different machine learning approaches will be introduced and discussed as follows: DNN based LSTM network, RL based Q learning algorithm, and DRL based DQN structure. All of these ML approaches will be separately developed based on maximizing the sum rates for all users in a NOMA system. These proposed ML approaches will be employed dynamically to approximate the channel parameters for each user in the examined NOMA cell.

- The efficiency of the proposed machine learning algorithms: DNN based LSTM model, RL based Q-algorithm, and DRL based DQN will be inspected in approximating the channel parameters when different fading channels models such as Rayleigh fading channel and Rician fading channel are applied.

- A new design structure will be proposed and applied at the receiver side to investigate how the channel estimation process based on LSTM or Q algorithm, or DQN can be combined with the derived power allocation policy, in order that this new joint structure can be utilized for multiuser detection in the PD-NOMA system.

- The optimized power allocation scheme and the fixed power allocation scheme, both will be implemented separately, evaluated, and compared when any of the proposed ML procedures is applied as a channel estimator in the examined NOMA cell.

- As a benchmark comparison, different simulation environments related to different works in the literature will be conducted, in order that we can compare the results of these benchmark schemes with the outcomes of our developed machine learning algorithms. Moreover, different performance metrics will be examined to emphasize that reliability can be guaranteed by our proposed machine learning approaches when being utilized for channel estimation process.

1.5 Contributions to Knowledge

Based on the current works in the literature, many deep neural networks approaches have been proposed to clearly address different issues in NOMA system such as power allocation, channel assignment, and signal detection. In addition, most of the proposed machine learning approaches in the literature are managing the issues of power optimization and channel estimation separately. To the best of our knowledge, there is no research that has investigated the integration between machine learning approaches that can be used to estimate the channel parameters and the optimal power allocation policy, in order that this integrated model can be used for multiuser detection in a downlink PD-NOMA system in different fading channels.

In addition, there have been several works that adopt ML approach to handle the channel estimation task in wireless communication networks. However, most of this research that employ ML for channel prediction in NOMA systems is mainly discussed via a specific model of deep neural networks. To the best of my knowledge, currently, there is no study that manages the channel approximation task using RL based Q-learning algorithm in a multi-input single-output non-orthogonal multiple access (MISO-NOMA) system. The RL based Q-table procedure is developed based on maximizing the sum rates for all users in the network. Moreover, as a channel estimator schemes, RL based Q algorithm and DRL based DQN

6

structure are developed, implemented, and compared when the optimized power scheme is considered.

Furthermore, in this thesis, a structured mathematical analysis is introduced to formulate the analytical power allocation expression for user devices in the examined NOMA system based on maximizing the sum rate in the system while considering the constraints of the total power budget in the system, and the QoS requirements for each user.

In this research, the main contributions can be summed up as follows:

- Different ML models are proposed (DNN based LSTM, Q-algorithm, DQN) to illuminate how these ML models can be developed and utilized dynamically to predict the channel parameters for each user in NOMA cell when the sum rate maximization problem is considered.

- Evaluate the beneficial impact of cooperatively integrating the developed ML algorithms (DNN based LSTM, Q-algorithm, DQN) for channel prediction and the derived power allocation scheme to be used in multiuser recognition in the PD-NOMA system.

- The developed DNN based LSTM structure, the developed RL based Q algorithm, and the developed DQN model are all investigated using different types of fading channels such as Rayleigh or Rician fading channels, to validate the efficacy of these algorithms in estimating the channel parameters for each user.

- The optimized power allocation scheme and the fixed power allocation scheme are both compared when the developed ML approaches are implemented as a channel estimator.

- As a benchmark comparison, additional simulation environments are established as follows: the standard minimum mean square error (MMSE) procedure for channel estimation [13], the DNN model based on LSTM network for channel prediction [10], the RL based actor-critic technique for channel prediction [14], and the fourth simulation environment is mainly dependent on RL based State-Action-Reward-State-Action (SARSA) procedure [15]. The simulation outcomes of these benchmark environments will be compared with the results of our proposed ML approaches, and simulation outcomes have emphasized that dependability can be assured by our proposed ML schemes in predicting channel parameters even when the number of devices in the cell is increased.

1.6 Thesis Outline

This thesis in total consists of seven chapters, starting with an introductory chapter to outline the motivation behind the research, discuss the aims and objectives and then highlight the contributions of the research. The remaining chapters are organized to start with an introduction and end with a summary for each chapter while the information related to the research are discussed in each chapter based on the purpose of the chapter as below:

Chapter 2 "Background and Outlooks": The first section of this chapter provides an overview of PD-NOMA system and then discuss the basic components that involve, superposition coding signal (SC) and successive interference cancellation (SIC) procedure. Then, a literature survey for the power optimization techniques for SISO-NOMA, MISO-NOMA, and SIMO-NOMA systems is presented. The last section of this chapter is mainly dedicated toward introducing the various designs and the functionality of DNN and how it could be adopted in NOMA system.

Chapter 3 "Mathematical System Modelling": System model for the NOMA system that includes the superposition coding signal, the received signal, and the expected user rate are all formulated in the first section of this chapter. Then, based on the introduced objective function and the constraints, the power optimization problem is analysed, and the optimum power coefficients are derived for SISO-NOMA, MISO-NOMA, and SIMO-NOMA systems. Finally, simulation environments are conducted, and results are generated and discussed for different performance metrics.

Chapter 4 "Deep Neural Networks Based Long-Short Term Memory (LSTM) Architecture and Framework": The basic concept of Recurrent neural networks (RNN) is introduced at the start of this chapter. Then, the main design architecture for the proposed DNN based long-short term memory (LSTM) is presented and discussed. Based on the channel estimation task, dataset generation and network parameters are described and the DNN algorithm is provided to clarify how the developed DNN structure is utilized to estimate the channel parameters for each user in NOMA cell. Finally, simulations are generated, and benchmarks and results are discussed.

Chapter 5 "Reinforcement Learning Based Q-learning Algorithm for Channel Estimation in MISO-NOMA System ": Reinforcement learning theory and Framework is mainly introduced in the first part of this chapter. Then, based on the developed Q learning algorithm

the design architecture for the channel prediction scenario is illustrated and discussed. Dataset preparation and Q learning model are presented to clarify how the Q algorithm is initialized and updated in order to predict the channel parameters for the users in the NOMA cell. Simulation environment, benchmark comparisons, and results are also generated and discussed in this chapter.

Chapter 6 "Deep Reinforcement Learning Framework in NOMA System": The system model and the basic components of DRL based DQN procedure is introduced in the first section of this chapter. The proposed DQN based LSTM network architecture is illustrated and the DQN training phase is clarified and outlined. DQN policy and dataset generation are also designated to maximize the rewards and minimize the error. Finally, simulation scenarios are created, and benchmark comparisons are also generated to characterize the impact of DQN when being used as a channel estimator in different simulation scenarios.

Finally, Chapter 7 "Conclusion and Future Work": This chapter present a global summary for the research with conclusions. Future research outlook is also highlighted based on the estimated challenges that may arise with the deployment of more complicated machine learning algorithms in the upcoming communication networks.

Chapter 2

Background and Outlook

2.1 Introduction

Fifth generation (5G) or sixth generation (6G) cellular systems are not just an outspread form of the preceding generation because of the growing request of network data traffic, but the forthcoming generation must support the new emerged technologies such as the internet of things (IoT) devices, and the web-based artificial intelligence (AI) applications. Furthermore, the upcoming generation should have the capability to enhance the massive connectivity and develop the spectral efficiency [16]. On average, Orthogonal multiple access (OMA) scheme can be considered as the standard multiple access scheme but in limited communication environment, hence OMA scheme may not be satisfactory to support huge networks that may require a diverse quality of services (QoS). Also, OMA may suffer from the matter of restricted degrees of freedom (DoF), where user with a good channel condition is served first with respect to user with weak channel status who need to postpone for channel access [17]. In order to realise the requirements of QoS and DoF that needed for wireless communication networks, non-orthogonal multiple access (NOMA) was introduced as a new multiple access scheme to enhance the system capacity.

2.2 Power Domain Non-Orthogonal Multiple Access

As shown in Figure 2.1, in power domain non-orthogonal multiple access (PD-NOMA) system, multiplexing for the different user's signals is carried out in the power domain via the superposition coding (SC) scheme at the transmitter side, and these power differences between user's signals are exploited to assist in the interference cancellation at the receiver side [18]. Because PD-NOMA deals with signals that have differences in the allocated power levels, so PD-NOMA has the capability to decode and split up the interference signal at the receiver and then remove it to leave only the desired signal. By adopting superposition coding (SC) at the transmitter side and applying SIC procedure for interference cancellation at the receiver side, this will enable PD-NOMA scheme to distribute DoF between devices in a fair way.

Figure 2.1 can be clearly described as follows: In NOMA cell, numerous user devices can be served via the same resource block (RB) by adopting the power domain in both uplink and downlink transmissions. In Figure 2.1, we are considering a downlink NOMA cell, where the BS is serving three distinct types of users or devices at same time via different fading channels. In NOMA cell, the BS can assign one carrier to every set of user devices, and the signals of that devices can be superimposed using unique power levels. Then, each user device will receive the desired signal beside the undesirable signals related to other devices in same channel and these unwanted signals can be either considered as interference or noise. The undesirable received signals will be considered as noise if the power level of the desired signal is high, otherwise, these unwanted signals will be regarded as interference. To decode the desired signal for each user, each user device can use the successive interference cancelation (SIC) procedure. The SIC technique will decode the signal with the highest power level and then subtracts that signal from the principal signal until the desired signal is decoded. As shown in Figure 2.1, near user usually has a good channel condition along with the BS, therefore the near user is usually assigned low power. Therefore, at near user receiver side when SIC is applied, firstly an immediate decoding for far user signal is accomplished, then removing it from the composite signal. Next, decoding the middle user signal, and removing it from the remaining signal and finally the near user desired signal will be decoded.
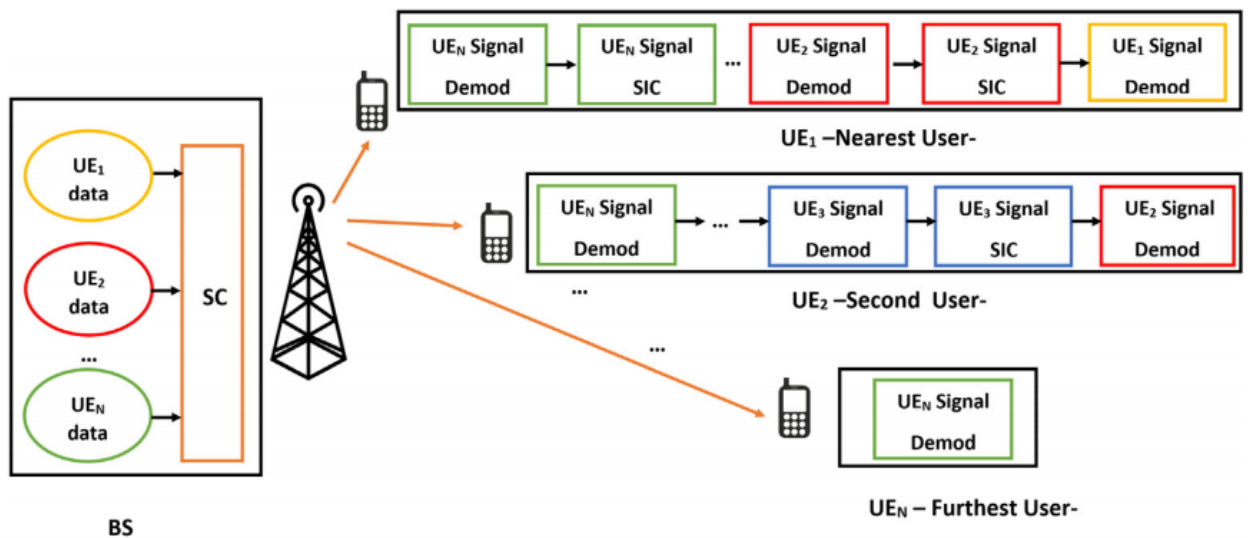
Figure 2.1 Basic block diagram for PD-NOMA System.

As mentioned above, when applying the superposition coding (SC) technique, the base station (BS) can transmit the superposition coded signal for all users in the NOMA cell. Then, at the receiver side for each user, the desired signal can be recovered based on the channel gains for all users in the cell and the power allocated for each user.

The main characteristics of NOMA can be described as follows [19]:

- Bandwidth efficiency: NOMA can demonstrate a high bandwidth efficiency and hence can improve the system throughput. This enhancement is related to the fact that NOMA system allows each resource block (RB) to be exploited by multiple user devices (time - Frequency).

- Fairness: A key feature of NOMA is that it can allocate more power to the weak user device, and by doing so, NOMA can realize the fairness among users in terms of their throughput. There are multiple techniques for maintaining fairness in NOMA system, such as the cooperative NOMA scheme.

- Connectivity: The next generation of wireless networks is expected to support a connection of billions of smart IoT devices [20], hence, the presence of NOMA can offer an encouraging design alternative to efficiently enhance this feature by exploiting its non-orthogonal characteristics. More specifically, in contrast to classical OMA, which requires the same number of frequency-time resource blocks as the number of user terminals, NOMA system on the other hand is capable of serving them by using less resource blocks.

- Compatibility: NOMA can be considered as an "add-on" procedure for any existing OMA techniques, such as TDMA/FDMA/CDMA, due to the fact that it exploits a new dimension, which is the power domain. Furthermore, based on the characteristics provided by superposition coding (SC) and successive interference cancellation (SIC) techniques, NOMA can be incorporated with the traditional multiple access techniques.

- Flexibility: based on the fact that NOMA is relying on sharing multiple users in a single resource block, NOMA can be considered as conceptually appealing and can provide an affordable complexity design.

2.3 Superposition Coding Signal

Non-Orthogonal Multiple Access (NOMA) can be considered as a powerful candidate multiple access scheme for future wireless communication networks. The fact that NOMA design permits multiple users to transmit and receive using the same time-frequency block is considered as attractive feature. The two key operations that make NOMA possible are the superposition coding (SC) which must be done at the transmitter side and successive interference cancellation (SIC) that implemented at the receiver side [4][18][19]. To explain the concept behind the SC process, assume we have two users in the cellular network and both users need to send their data simultaneously using the same time-frequency block.

Let x_1 denotes the data for user 1 and x_2 denotes the data for user 2. For simplicity, let us assume that data for each user can be represented by 4 bits as follows:

$x_1 = [1010]$ and $x_2 = [0110]$ are both represented graphically as shown in Figure 2.2.
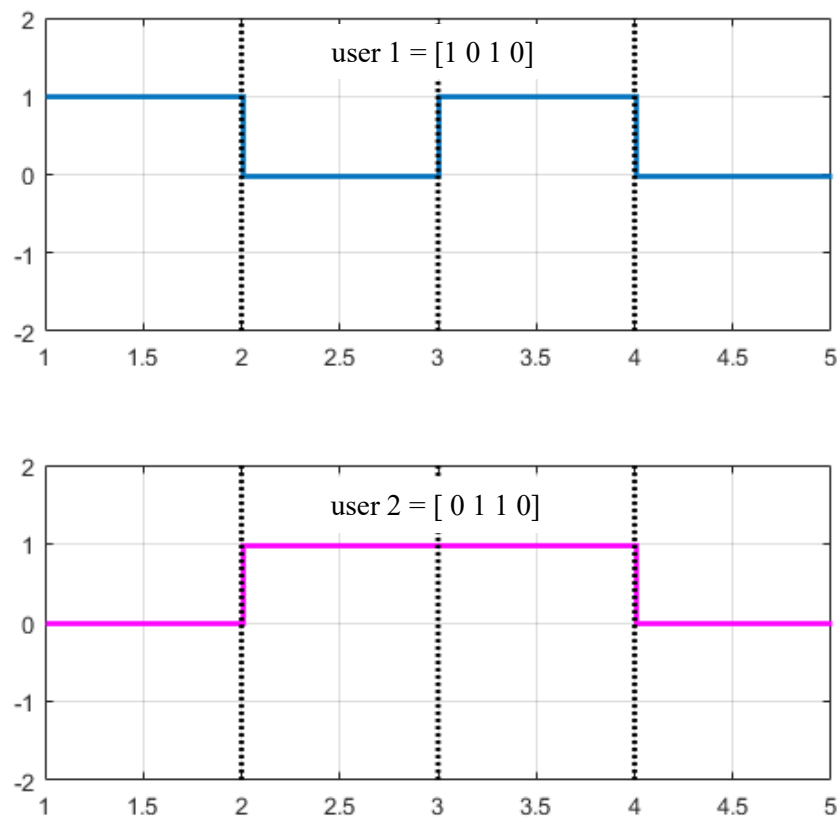
Figure 2.2 User 1 and user 2 data representation [19].

13

x_1 and x_2 must be modulated before transmitted via the channel, therefore a digital modulation scheme can be employed such as binary phase shift keying (BPSK). Basically, BPSK maps 0 to -1 and 1 to +1. After BPSK modulation,

x_1 is mapped to $x_1 = \begin{bmatrix} +1 & -1 & +1 & -1 \end{bmatrix}$

x_2 is mapped to $x_2 = \begin{bmatrix} -1 & +1 & +1 & -1 \end{bmatrix}$

and the modulated sequences can be presented graphically as shown in Figure 2.3.
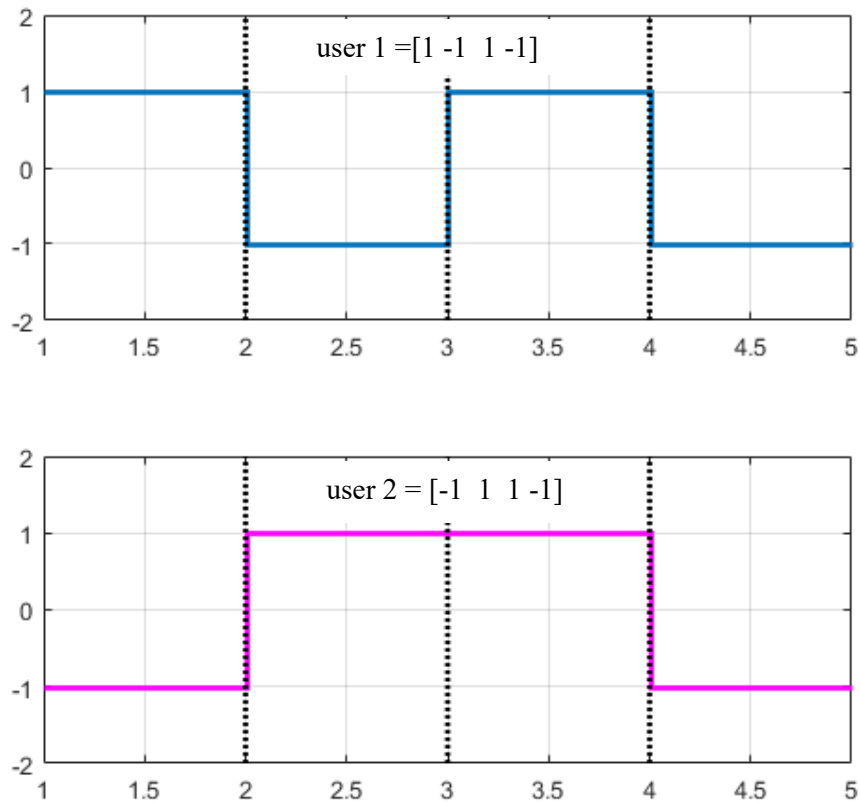
Figure 2.3 Modulated data representation for user 1 and user 2 [19].

Superposition coding is an essential stage in power domain NOMA. To superpose means to add, so, we need to add x_1 and x_2 together. But before doing so, we need to multiply the modulated data for each user with the allocated power factors, then, we can add them together [4][11][12]. From Figure 2.3, we can see that both x_1 and x_2 have a peak amplitude of ±1, which means that they both have unit power. We can allocate the power factors as follows: a_1=0.75 to user 1 and a_1=0.25 to user 2. A rule to obey here is that a_1 and a_2 must sum up to 1 which reflects the total power assigned by the base station. Since a_1 and a_2 denote the power

scaling factors, we need to apply square root for the power coefficients to transfer them to amplitude. So, we can scale x_1 and x_2 with $\sqrt{a_1}$ and $\sqrt{a_2}$ respectively. The amplitude of x_1 will be scaled by $\sqrt{a_1} = \sqrt{0.75} = 0.866$ and the amplitude of x_2 will be scaled by $\sqrt{a_2} = \sqrt{0.25} = 0.5$. Hence, the amplitude scaled version of the modulated data for each user, can be shown as follows: for user 1 $\sqrt{a_1}x_1 = [0.866 \; -0.866 \; 0.866 \; -0.866]$ and for user 2 $\sqrt{a_2}x_2 = [-0.5 \; 0.5 \; 0.5 \; -0.5]$

After scaling, the modulated data for each user can be shown graphically as shown in Figure 2.4.
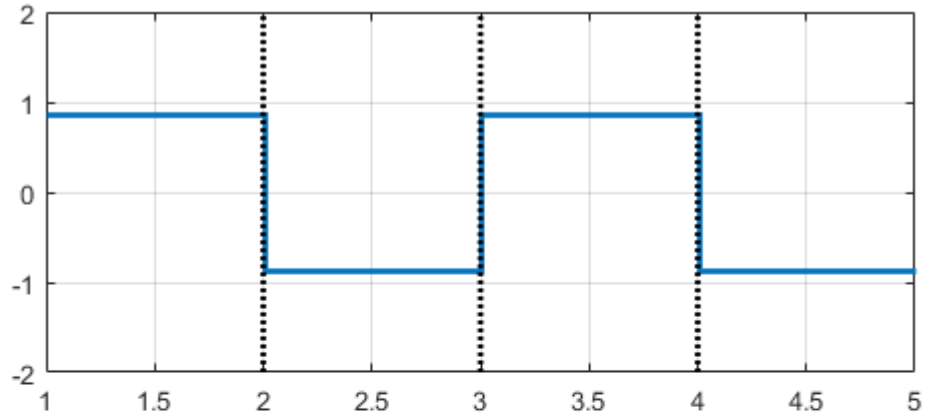
a) modulated data of user 1 after power scaling
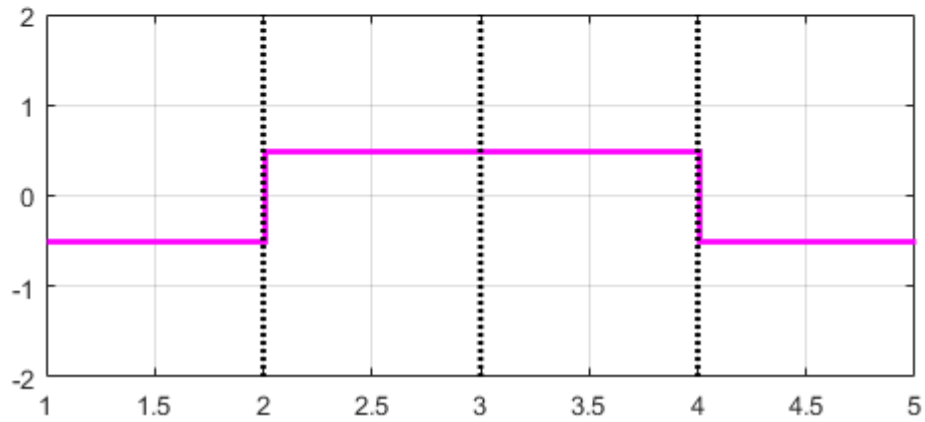
b) modulated data of user 2 after power scaling

Figure 2.4 Modulated data for user 1 and user 2 after power scaling [19].

Now we need to add the modulated scaled data for user1 and user 2, then the resulting signal is called superposition coded signal x that can be denoted as in equation 2.1.

$$x = \sqrt{a_1}x_1 + \sqrt{a_2}x_2 \qquad (2.1)$$

15

The resultant superposed signal can be represented as a values vector form as follows: $x = [0.366 - 0.366\ 1.366 - 1.366]$ and the graphical representation of this superposed signal is shown in Figure 2.5.
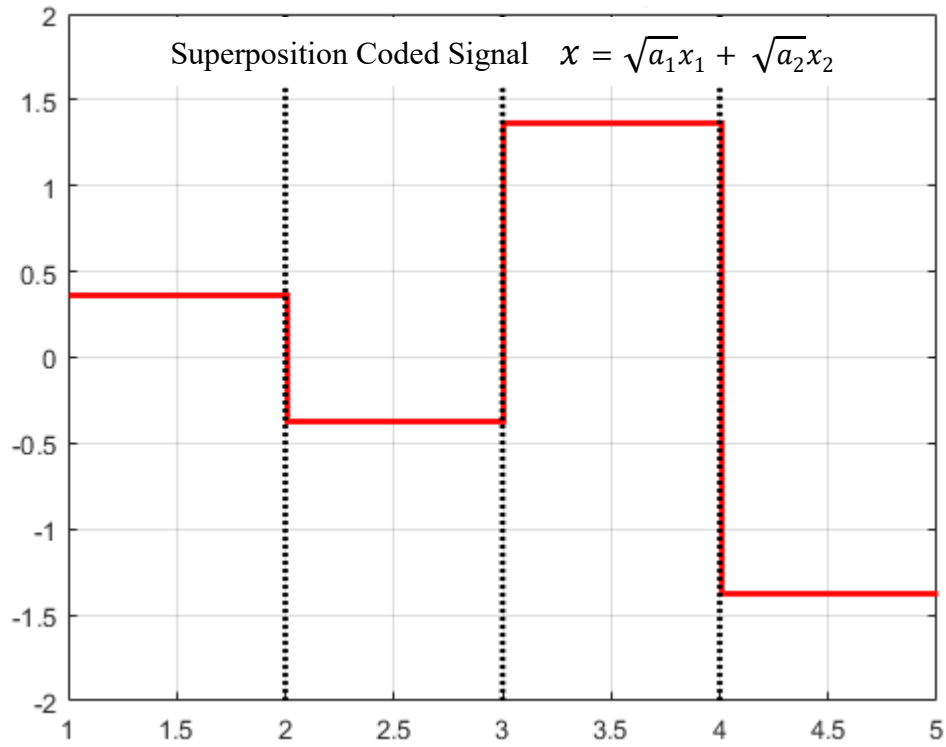
Figure 2.5 Superimposed coding signal x [19].

The resultant signal x can also be represented as a linear combination of x_1 and x_2 and this signal x is actually transmitted from BS towards all users in NOMA cell in downlink scenario.

2.4 Successive Interference Cancellation

PD-NOMA scheme depends on the power domain for multiplexing user's signals in same time and frequency resource block, and as explained in the previous section, this is accomplished by applying SC at the transmitter side. Successive interference cancellation (SIC) is performed at the receiver side, and it is considered as an iterative procedure where data is decoded in the order of decreasing power levels. The signal or the data corresponding to the user who is given the highest power is decoded first, then the data corresponding to the user who is given the next highest power is then decoded. This procedure will be continued till we have decoded all user's data. Based on the details given in the previous section, we will continue consider the

16

two users case study in a NOMA cell, and the steps involved in SIC can be clarified as follows:

Step 1: directly decode the superposition coded signal x to get the original signal for user that allocated the highest power. Therefore, based on our abovementioned analysis, where x_1 is given more power proportion $(a_1 > a_2)$, direct decoding of x will recover the original signal x_1.

Step 2: re-modulate x_1 and multiply the remodulated signal in step 1 by its corresponding power factor and subtract it from x, as follows $x - \sqrt{a_1}x_1$.

Step 3: decode the signal obtained in step 2, $(x - \sqrt{a_1}x_1)$ to recover the original signal x_2.

Since we have applied BPSK modulation scheme at the transmitter side, so we need to apply direct decoding at the receiver by applying BPSK demodulation directly to x. BPSK demodulation is basically, a simple thresholding [12][13][15].

So, simply we can assign a threshold level as zero, hence if the amplitude exceeds zero, the decoding symbol is 1, and 0 otherwise. As indicated in Figure 2.6, threshold level is denoted as a solid black horizontal line. It can be observed that the first and third symbols sit above the threshold level, so, we take a decision that the first and third bits are ones. The second and fourth symbols sit below the threshold level, so we take a decision that the second and fourth transmitted bits are zeros. Thus, the decoded sequence is [1010] which is the same as the original sequence x_1. We have recovered the original signal x_1 by directly performing BPSK demodulation on x and ignoring the fact that the superposed signal x has another component related to x_2 . This was possible because, we have allocated a higher power factor to x_1.
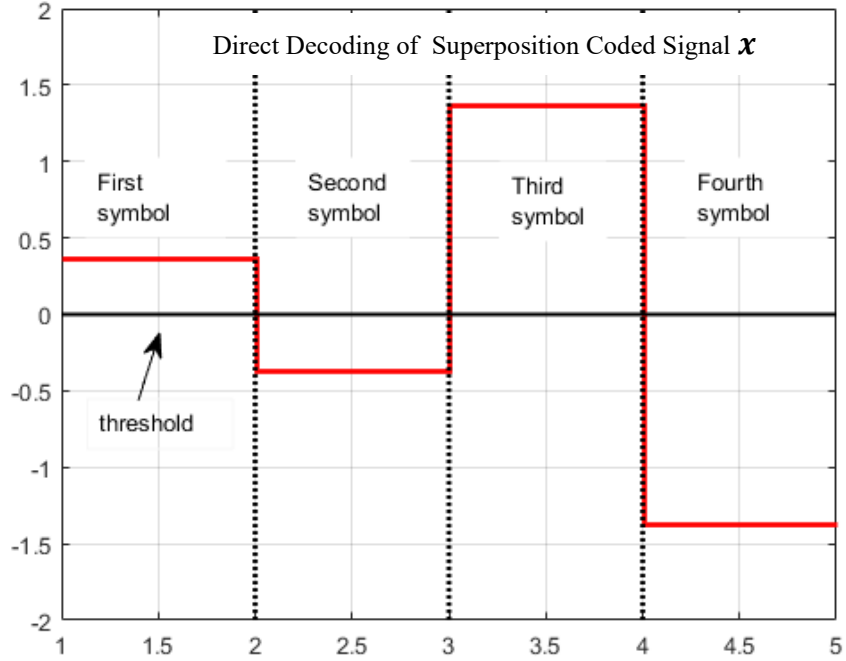
Figure 2.6 Direct decoding for signal x_1 [19].

It has been shown before that the coded signal x is analytically expressed as follows: $x = \sqrt{a_1}x_1 + \sqrt{a_2}x_2$. We obtained x_1 by following step 1, then to recover x_2 we need to subtract $\sqrt{a_1}x_1$ from x, then we will be left with $\sqrt{a_2}x_2$. Recall that we have recovered here x_1 by direct decoding and the result come to be [1010], but x_1 is present in x as a BPSK modulated form, So we need to reproduce x_1 in BPSK form again as [1 -1 1 -1], and scaled it by $\sqrt{a_1}$, then we need to subtract this BPSK modulated scaled version of x_1 from x as follows:

$$x - \sqrt{a_1}x_1 = [0.366 - 0.366\ 1.366 - 1.366] - [0.866 - 0.866\ 0.866 - 0.866]$$

$$x - \sqrt{a_1}x_1 = [-0.5\ 0.5\ 0.5\ -0.5]$$

After subtraction, the graph can be illustrated as shown in Figure 2.7.
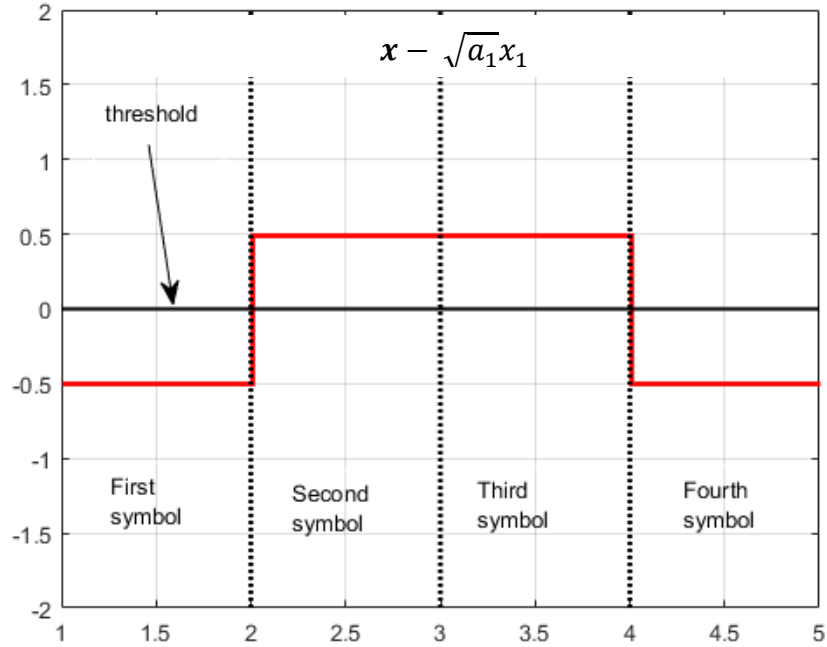
Figure 2.7 Decoding for signal x_2 [19].

After subtraction, we can simply apply the same BPSK demodulation rule to demodulate the resulting signal $(x - \sqrt{a_1}x_1)$. As shown in Figure 2.7, it is clearly shown that the first and fourth symbol can be demodulated as zeros and the other symbols should be demodulated as ones. Thus, the decoded sequence for user 2 will be [0110], which is matching with the original signal for user 2 generated in the transmitter side. Therefore, two different user's data have multiplexed in the power domain and successfully recovered based on SIC procedure.

2.5 NOMA System Based Power Optimization Literature Survey

In [21], authors have proposed the sum rate maximization problem based on certain constraints for uplink NOMA system and a closed form for the optimal power coefficients is derived. Numerical results revealed that the sum rates for NOMA system with the suggested power distribution scheme are more sensible than its OMA counterpart.

Power allocation and user clustering were described in [22], where the sum throughput maximization for uplink and downlink NOMA systems was studied under several constraints. The authors declared that user clustering based on distinctive channel conditions can provide a noteworthy gain in NOMA system compared to OMA system. Authors also went to conclude that with large NOMA clusters, the error transmission in SIC procedure will negatively impact the performance of NOMA system.

The upper bound for the sum rate were analysed in [23], based on a power threshold in distinct fading channels. Sum rate upper bound is derived, and simulation outcomes showed that system dependability is improving when the target received power is considered, whilst the system reliability is shown to decrease when power threshold is included. Authors also went to conclude that the applied fading distribution has a major role in allocating the maximum and minimum power values that needed to satisfy the communication requirements.

A dynamic power allocation for NOMA users with mobility was introduced in [24], where user device can adjust his position in a random way with reasonable speed and at same time, the suggested dynamic power scheme can deliver a remarkable sum rate. On the other hand, it has been shown that static power scheme is not satisfactory when user mobility is considered due to unstable channel condition.

Performance of a downlink NOMA cell in terms of bit error rate (BER) is examined in [25]. BER expressions were derived for NOMA scheme when Nakagami fading channel and imperfect SIC are considered and two or three users are assumed in the cell. Optimum power coefficients that can minimize the BER were also deduced to manage the fairness between users. Simulation outcomes demonstrated that the power coefficients should be determined carefully to avoid a huge BER difference between user devices.

In [26], optimum power allocation with channel assignment is investigated to maximize the users' rate in NOMA system when varying channel gains are considered. Quality of service (QoS) and downlink power budget constraints and are examined in the power allocation task. Simulation results have indicated the improvement in system performance when the system throughput is maximized.

The integration of NOMA and multiple antenna techniques [4] is investigated as a proper scheme to boost system performance. Hence, analysing multiple-input single-output (MISO) based NOMA system can be an encouraging towards highlighting the significant improvements in the achievable system rate.

 In downlink MISO-NOMA system, authors in [27] have suggested a closed form lower bound for the channel capacity and also optimal power level assigned to every user is found given that user rate constraints should be satisfied. An algorithm is introduced to determine the maximum number of users in the cell while the user rate constraint is considered. Authors went

to conclude that the number of users in the system will decrease when either the minimum rate requirement is increased or the channel gain gap between users is raised.

Power consumption and clustering in a multi cluster MISO-NOMA system were explored in [28]. In order to lower the total transmitted power, a two-layer user clustering algorithm is introduced to group users into clusters and each cluster contains 2 users. In the first layer, the algorithm is focusing in deciding the cluster heads according to condition of the channel gains, whereas in the second layer the procedure is responsible to couple each cluster head with a tail user. Numerical results are generated to validate the superiority of the proposed clustering algorithm in MISO-NOMA system.

An iterative power distribution algorithm is proposed in [29], for a power minimization in a multi cell MISO-NOMA system. Throughout each iteration, base station (BS) is accountable for adjusting the power transmitted for each user according to the channel status and the inter-cell interference plus noise at user devices. The algorithm showed performance improvements and the simulation outcomes emphasized that the considered algorithm with MISO-NOMA system can considerably outperform OMA technique.

Sum rate maximization in MISO-NOMA downlink system based simultaneous wireless information and power transfer (SWIPT) is investigated in [30]. Authors proposed a user pairing beamforming procedure where two user devices with distinct channel gains are selected to realize user pairing to enhance the sum rate and accordingly improve the energy harvesting capability. Based on convex approximation technique, authors have utilized low complexity iterative algorithm to solve the formulated sum throughput problem. Results confirmed the dominance of the proposed scheme over the non-user-pairing schemes.

Under the assumption of statistical channel state information (CSI), authors in [31], have formulated a closed form expression for the outage probability in downlink MISO-NOMA system. Outage formulas have been deduced based on probability density function (PDF) and cumulative distribution function (CDF) for the beamforming ratio. Outage probability is simulated against numerous power allocation scenarios to justify the reliability of the outage probability results.

In [32], energy efficiency maximization task-based clustering in a MISO-NOMA system is investigated based on imperfect channel state information, when total power transmitted, and

minimum user rate constraints are considered. Authors have utilized an iterative algorithm that use of Taylor series method to translate the resultant nonlinear optimization problem into a subtractive form. Also, two approaches of zero forcing (ZF) scheme were introduced to relieve the inter cluster interference. Simulation results illustrated that increasing the number of clusters will also expand the inter-cluster interference while zero forcing scheme verified to be a suitable candidate to provide a remarkable performance when energy efficiency is assessed.

Based on arbitrary path beamforming, authors in [33] investigated the single cell millimetre wave MISO-NOMA system, where users are grouped into multiple NOMA clusters based on their antenna's angles of arrival. Authors have developed the lowest achievable signal to interference plus noise ratio (SINR) associated with power allocation scheme. In addition, they propose an iterative algorithm to manage the power allocation factors in order to achieve the maximum-minimum SINR (maximizing the min. SINR to improve performance) Simulation outcomes have shown that the proposed algorithm can achieve a better performance than conventional orthogonal multiple access scheme in terms of user fairness.

The power minimization task for a standard multi cell MISO-NOMA system is explored in [34]. The joint power management and beamforming scheme is discussed as a non-convex dilemma, which is handled by the near optimal zero forcing beamforming technique. Numerical results showed that the suggested approach can act better than various standard MISO-NOMA schemes.

In [35], the main aim was to establish a feasible framework to handle the following tasks: power allocation, beamforming, and user clustering in MISO-NOMA system by controlling the transmit power and user's rates. Authors have manged the superposition coding signal for clustering two users and process the signals in space division multiple access via beamforming. A closed form expression is formulated to manage the beamforming with optimal power distribution. Based on the results, it was shown that the proposed approach outperforms the conventional counterpart and can provide noticeable gains in terms of power transmitted.

In [36], a cooperative PD-NOMA with energy harvesting procedure using multiple antenna system is explored. Authors suggest a framework to improve the bit error rate (BER), spectral efficiency and capacity compared to conventional NOMA system. Two users were considered in the cell, where one user is near to the BS, while the other user is far from the BS and both

users are allocated different power levels and share the same spectrum. Simulation results of this cooperative MISO-NOMA system with energy harvesting verified a reasonable enhancement compared to traditional system when different performance metrics is inspected.

2.6 Machine Learning and Artificial Neural Network

Over the past several years, machine learning (ML) has achieved popularity, and numerous machine learning approaches have been industrialized. Artificial neural networks (ANNs) are considered as one of the most popular models for ML, that principally exploited for prediction and forecasting [37]. ANN is an information processing structure that is inspired from the biological nervous system that manages how the brain is processing information. The most elementary component of an ANN is the neuron, which can take two or more inputs and give one output. The output of the neurons is calculated based on a set of weights applied to the inputs plus a bias value that can adjust the output. Various neurons are organized into a layer and these layers can be connected such that the output of each layer feeds into the input of the next layer, and this is called a feedforward ANN that can be employed to solve problems such as classification and signal detection [38].

2.7 Deep Learning Based NOMA System Related Works

In the literature, the encouraging procedure of machine learning (ML) has been jointed with NOMA system to manage different types of tasks such as channel assignment, beamforming, and power allocation. Also, the deep learning (DL) concept, which was proposed in 2006, is considered as a dynamic tool for handling huge data sets and solving complex nonlinear problems [39]. Deep learning influence in MIMO system where channel coding is considered is investigated in [40]. A deep learning based sparse code multiple access (SCMA) is developed in [41], to reduce the BER, and to satisfy a smaller computation time compared to the typical schemes. Approaches that incorporate the deep learning into the orthogonal frequency division multiplexing (OFDM) system also has been introduced in [42], and its satisfactory performance has been discussed in terms of channel estimation and signal detection. Furthermore, deep learning has been implemented in traffic systems, where the performance enhancement has been shown in [43]. A deep reinforcement learning based power management has been introduced in [44] to improve the system performance by identifying the power that should be allocated to each user to ensure the fairness between users. In addition, authors also suggest a dual DNN models to handle the noisiness problem in training data. In [45], authors have

presented a co-operative NOMA structure based on deep learning model to enhance the BER performance. Authors have introduced a cascaded DNN models to characterize the cooperative NOMA system. Loss functions were utilized to evaluate the BER performance, and a two-stage training process was also adopted to resolve training problem. Simulation outcomes have illustrated the merits of the proposed DL scheme over the orthogonal multiple access (OMA) scheme and the conventional NOMA scheme in different channel environments.

2.8 Artificial Neural Networks and Deep Learning Basic Theory

Throughout the training process, an ANN is given a set of input data for processing, the output from the ANN (predicted) is then compared to the target output, and an error can be calculated. The network weights and biases can be adjusted at each set of training iterations using gradient descent algorithm [46] to minimize the error. This training process can be repeated several times till the ANN starts to show a good performance or the error function converges. Once the ANN efficiently trained, it can receive a new data set and predict the output with reasonable precision. Usually, the ANN can perform well when the training model follows to certain policy. Commonly, all the input data need to be pre-processed to make sure that there are no substantial time gaps in the data sequence. Furthermore, it is desirable to normalize the training data, in order to be within the same range and to speed up the network convergence [10][46].

Overall, three types of ML algorithms are mainly discussed in the literature, such as supervised, unsupervised, and reinforcement learning. Supervised learning is extensively exploited to estimate the channel state information (CSI), spectrum sensing, and localization [47] [48]. Unsupervised learning is frequently used in user clustering and congestion control [49]. Reinforcement learning is a different branch of ML, and it is considered to have a remarkable impact in wireless networks. The most recognized application of reinforcement learning in communication networks is related to resource allocation [50].

The deployment of deep neural networks (DNN) or deep learning (DL) in NOMA system, is greatly explored over the last few years. Various kinds of learning algorithms have been suggested and implemented in different types of communication networks to enhance the system performance. Typically, deep learning can be adopted as a useful tool for complex data processing and also can be exploited in different applications such as estimating channel features. A typical deep Learning model is shown in Figure. 2.8, where the hidden layers are

particularly equipped with multiple neurons. Online training and offline training are considered as the two major approaches for training in DNN that can be adopted in NOMA scheme [11][51].

When we consider the channel parameters estimation task, a random sequence of the input data can be extensively trained using the channel parameters related to different environments collected from simulations. This type of training is described as offline training.

Alternatively, the other type of training is called online learning, that depends on the pilot symbols to carry the channel state information and the input signals can be trained using DNN with the help of that pilot sequences [52].
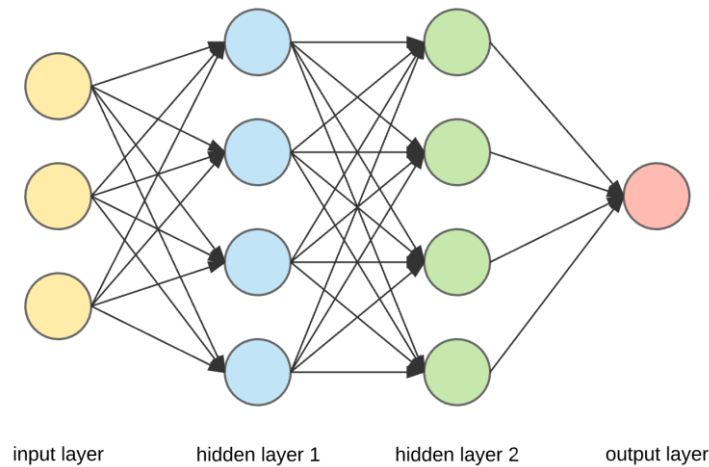
Figure 2.8 Deep learning basic architecture.

Sometimes, in NOMA system it can be difficult to apply the SIC procedure to decode the data perfectly when the number of user devices is huge, besides that the SIC technique might be also affected by the propagation error. Thus, the deep neural network can be a practical alternative to recover a desired sequence from a weaken received signal.

2.9 Recurrent Neural Network

Recently, many works have proposed different approaches for resource optimization based on maximizing the data rates, but these approaches need high computational complexity due to the nonlinear optimization problem. Traditionally, the channel characteristics in multiuser NOMA system are described in complex form and the conventional estimation procedures are not effective and reliable enough in capturing the variation of the channel parameters in real time. Accordingly, the channel state information estimation process is interrupted, and NOMA

system performance is affected. Many DNN approaches have been proposed for further enhancing the NOMA performance based on estimating the channel features [10][53][54].

Recurrent neural network (RNN) is considered as one of the candidates DL models that fits the CSI and channel estimation tasks. Recurrent neural network (RNN) is classified as one of the prevalent DL schemes that can be utilized for predicting time series data [55]. At each timestep, the RNN receives not only the new input data but also its own output from the previous time step, known as the hidden state, the underlying form of RNN structure is clearly illustrated in Figure 2.9.
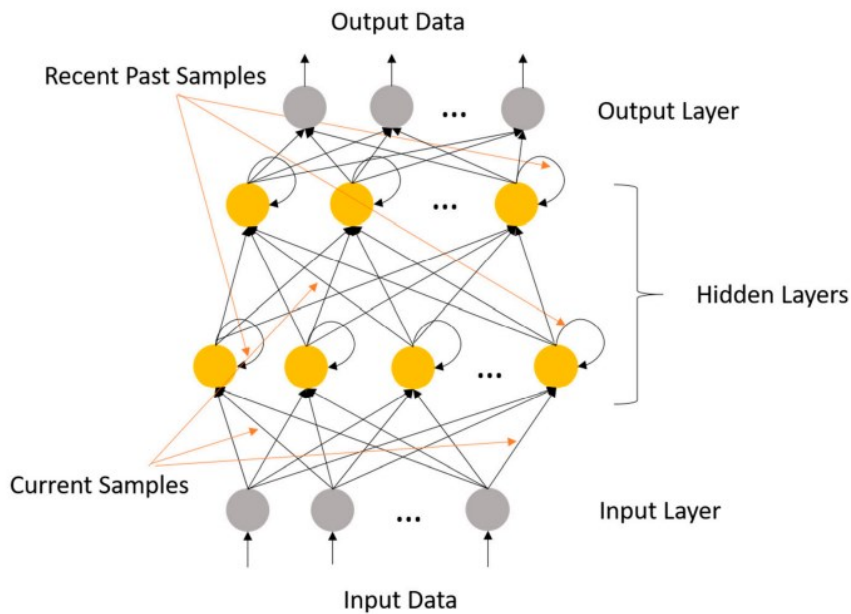
Figure 2.9 Recurrent neural networks fundamental construction.

Recurrent neural network (RNN) is regarded as a form of supervised learning algorithm where it can exploit the sequential information for estimation and recognition. This type of network consists of convoluted hidden layers made of artificial neurons with feedback. As shown in Figure 2.9, RNN have two inputs, the new input to each neuron and the recent previous which is the output that loops back into the network. Thus, the hidden layers are able to function as memory for the network at a certain instant, which is dependent on its previous state. This structure of RNN permits it to save and process the previous complex data for a prolonged period of time [56]. Likewise, RNN has the ability to map a certain input to the output sequence during the current time interval and estimate this sequence during the subsequent time slots.

2.10 Long Short-Term Memory Neural Networks

long-short term memory (LSTM) is mainly an artificial neural network that used in the fields of deep learning and artificial intelligence. Different from standard feedforward neural networks and similar to RNN, LSTM has a feedback connections and can process not only single data points, but also entire sequences of data. In LSTM, the hidden layers can act as a memory for the network state, which particularly permits the DNN to save, recall, and process the preceding complex data [57].

Transmitting the signal through a fading channel may result in a dispersion signal with long-term dependencies among different samples. These dependencies differ from one signal to another and do not necessarily follow a specified pattern [58]. Using a conventional neural network to represent such long-term dependencies may require a high-dimensional hidden layers with a huge number of neurons, which may lead to over-fitting [59]. Alternatively, RNN can model time dependencies with a lowered number of neurons, due to their feedback recurrent style. The traditional RNN using backpropagation through time is usually affected by the vanishing gradient problem and slow learning. In order to resolve this fundamental issues, long-short term memory (LSTM) neural network is proposed to be able to model selective dependencies between different portions of the received signal with a small number of neurons and with moderate learning rates. LSTM can accept a vector of complex data at each time step; therefore, it can include the magnitude and phase parts of the received signal simultaneously [60]. LSTM is considered as modified form of RNN. The LSTM module is classically referred to as cell rather than neurons and contain a sequence of gates. An illustration of an LSTM cell can be seen in Figure 2.10.

The standard LSTM cell includes a cell, an input gate, an output gate and a forget gate. The cell remembers values over arbitrary time periods and the three gates control the flow of data into and out of the cell. A forget gate examines the new input and the hidden state and decides which information in the cell state should be carefully forgotten. The input gate then determines what information from the new input should be delivered to the cell state to be remembered. The output gate gathers data from the cell state, input, and hidden state and produces the output for the current time step. By this procedure, LSTM has the power to remember information passed in different timesteps, making it ideal candidate for finding longer term decencies in data [10][59][60].
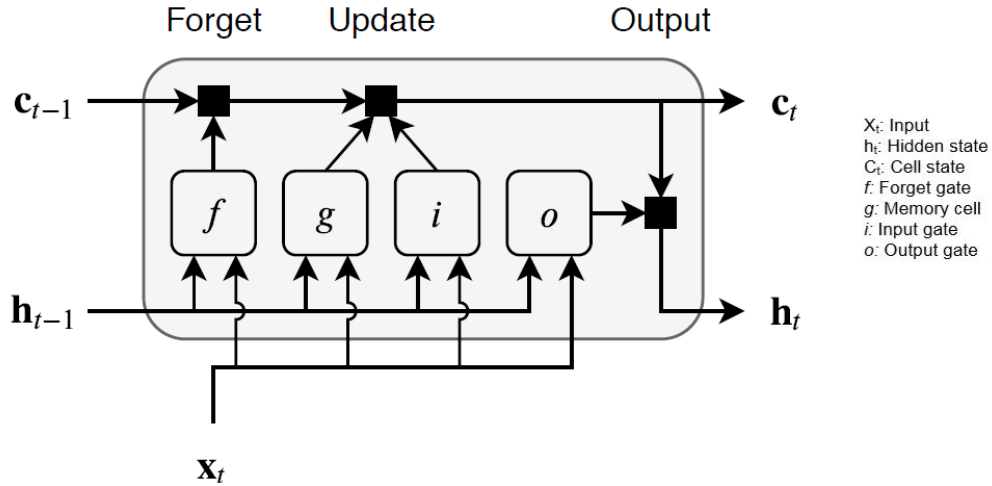
Figure 2.10 LSTM cell structure [10].

2.11 Modulation Scheme

The Quadrature Phase Shift Keying (QPSK) modulation scheme is a variation of a Binary Phase Shift Keying (BPSK), which sends two bits of digital information at a time. Instead of the conversion of digital bits into a series of digital stream, it converts them into bit pairs, this decreases the data bit rate to half, which allows space for the other users. Quadrature Phase Shift Keying is a modulation scheme that utilizes four different phase shifts to represent the binary values of data. It is called "Quadrature" because it uses two carriers that are 90 degrees out of phase with each other.

QPSK Modulator

The QPSK Modulator shown Figure 2.11 consists of the following: a bit-splitter hat contains a 2-bit serial to parallel converter, two multipliers with local oscillator, and a summer circuit.
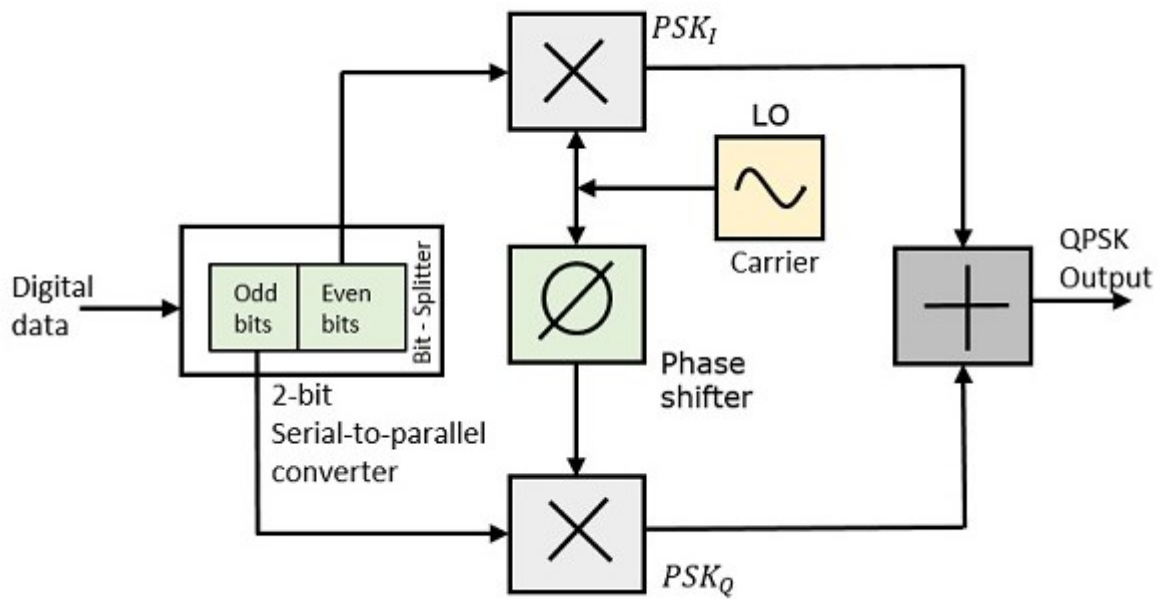
Figure 2.11 QPSK Modulator structure.

At the modulator's input, the message signal's even bits and odd bits are separated by the bits splitter and are multiplied with the same carrier to generate in phase BPSK component (called PSK_I) and out of phase BPSK component (called as PSK_Q). The PSK_Q signal is phase shifted by 90° before being modulated. The QPSK waveform for two-bits input is shown in Figure 2.12, which illustrates the modulated waveforms for different instances of binary inputs.
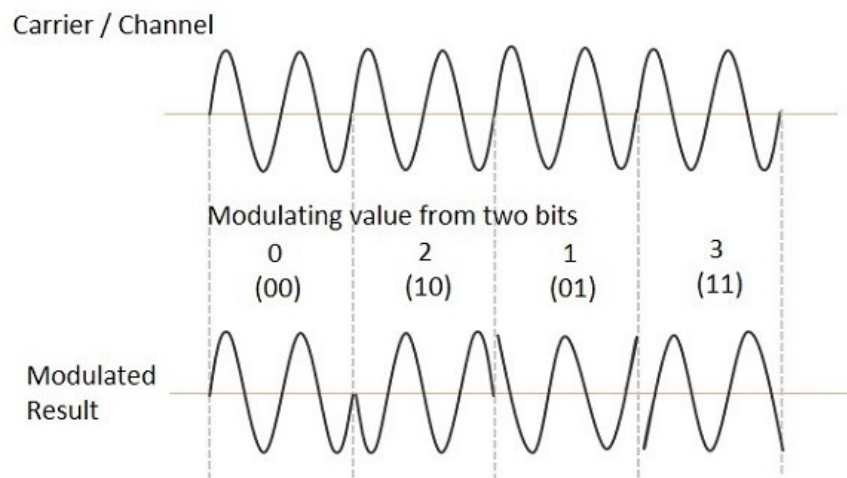
Figure 2.12 QPSK modulated waveforms.

QPSK Demodulator

The QPSK Demodulator shown in Figure 2.13, uses two product demodulator circuits with local oscillator, two band pass filters, two integrator circuits, and a 2-bit parallel to serial converter. The two product detectors at the input of demodulator simultaneously demodulate the two BPSK signals. These signals after processing, are passed to the parallel to serial converter.
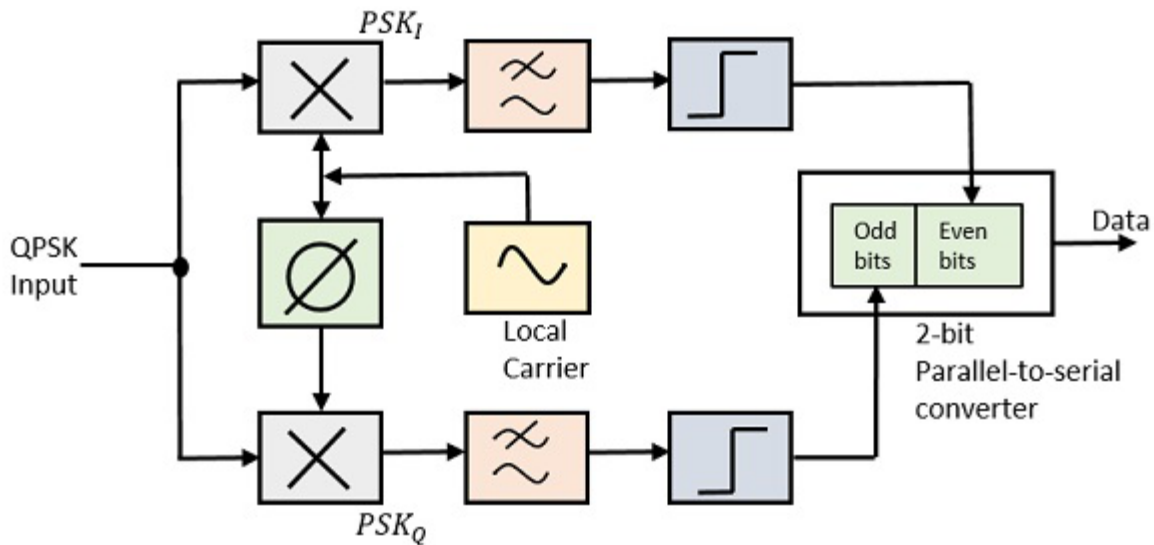
Figure 2.13 QPSK Demodulator structure.

2.12 Summary

OMA scheme may not be suitable to support huge networks that may require a diverse QoS, also OMA suffers from restricted DoF, where user with a good channel condition is served first with respect to user with weak channel status. In order to realise the requirements of QoS and DoF that needed for future communication networks, NOMA scheme is introduced as a new multiple access scheme to enhance these restrictions. It is clearly noticed that PD-NOMA can deal with signals that have differences in the allocated power levels, hence PD-NOMA has the capability to distribute a DoF between user devices in a fair way via adopting superposition coding at the transmitter side and applying SIC procedure at the receiver side for interference cancellation and signal decoding. In the next chapter, we will consider PD-NOMA system for the sum rate maximization problem, based on certain constraints and the optimum power coefficients for user devices will be derived.

Chapter 3

Mathematical System Modelling

3.1 Introduction

Non-orthogonal multiple access (NOMA) scheme has been classified as encouraging multiple access structure for forthcoming wireless communication networks to enhance system throughput and spectral efficiency. NOMA could exploit the available resources more efficiently by resourcefully benefiting from the users' channel status to provide several users with diverse quality of service (QoS) demands. NOMA enables several user devices to attain simultaneous arrival to the same time-frequency resource block by super-positioning the different signals in the code or power domains [61]. The concept of power domain non-orthogonal multiple access (PD-NOMA) is mainly depending on that the user device with a weak channel condition and the user with a good channel condition can both share the same time slot and the same subcarrier to ensure that the bandwidth block is well exploited.

In NOMA cell, the receiver at each user terminal will receive the multiplexing of signals from different users that are sharing the same subcarrier, therefore, removal of interference related to other users is essential for coordinated decoding. In PD-NOMA, multiuser detection (MUD) can be managed by applying the SIC procedure, where signals related to several users are decoded sequentially based on the availability of the channel parameters and the assigned power for each user in the NOMA system [10][11][61]. On the other hand, complete realization of channel parameters for each user device is challenging. Traditionally, pilot symbols can be basically utilized to estimate the channel parameters for users in wireless communication system, but this pilot symbols may also interfere with signals related to other users in the cell, thus affecting the performance of the standard channel estimation procedures, such as minimum mean square error (MMSE) [62]. Therefore, before decoding the desired signal for

each user we need first to carefully estimate the channel parameters for each user in the examined NOMA cell. Second, we may need to optimize the power factor allocated for each user device in the examined NOMA cell.

In this chapter, we will focus on presenting the framework that clarify how the allocated Power factors for user devices in the examined PD-NOMA cell are optimized based on maximizing the sum rate of the users while the entire power transmitted, and the Quality of service (QoS) constraints are considered. An investigation for the optimization problem is shown where Lagrange function and Karush–Kuhn–Tucker (KKT) optimality conditions are applied to derive the optimum power coefficients. Simulations will be conducted in terms of different performance metrics to explore the improvement achieved when the optimized power policy is applied compared to the users with fixed power scenario.

3.2 System Model

The downlink NOMA system is introduced in this section, where the path between the base station (BS) and user devices is represented by distinct fading channels. The PD-NOMA cell is considered, where a BS with a single antenna is assumed to serve three users concurrently and every user device also has a single antenna. As indicated in the above-mentioned discussion, each user device will receive the combined signal sent from BS that consists of a desired signal and the interfering signals and all these signals are sent through the same time-frequency block. Thus, multiplexing the signals using distinct power levels is critical to distinguish the signals and to strengthen the SIC procedure at the receiver side [10][11][61]. In PD-NOMA, users that are described by good channel condition are usually allocated low power level, while user devices with poor channel conditions can be assigned high power levels.

Typically, users are recognized according to their distance from BS and the fading channel distribution. In the proposed examined NOMA cell, the nearby device is indicated as near user while the user equipment at the edge of the cell is viewed as far user. In this chapter, we assume that we have three user devices in the cell and Rayleigh fading channels are considered with zero mean. Accordingly, the fading channel for each user device can be mathematically characterized as follows, for near user $h_n \sim (0, d_n^{-k})$, for middle user $h_m \sim (0, d_m^{-k})$, and for far user $h_f \sim \left(0, d_f^{-k}\right)$, where h_i indicates the fading channel between the BS and the user device i and k represents the path loss [63].

In this work, Additive White Gaussian Noise (AWGN) is assumed at the receiver side of each user, and the noise power is denoted as σ^2. Without loss of generality, the channel gains for the examined user devices can described as follows: $|h_n|^2 > |h_m|^2 > |h_f|^2$. Total power transferred from BS to all devices in the cell is specified as P_t. In NOMA system, the receiver at each user has the ability to carry out SIC to remove the signals related to other users with poor channel environments. In contrast, signals from devices with good channel conditions could not be removed and treated as interference. In downlink NOMA scenario, the BS can send the superposition coded signal x that can be expressed as follows [10][11][61]:

$$x = \sqrt{P_t}\left(\sqrt{\alpha_n}x_n + \sqrt{\alpha_m}x_m + \sqrt{\alpha_f}x_f\right) \qquad (3.1)$$

where α_n, α_m, and α_f are the power coefficients allocated to the near user, middle user, and far user respectively. Likewise, x_n, x_m, and x_f denote the target signals related to near, middle, and far users, respectively. The signal received at far user can be shown as follows:

$$y_f = xh_f + z_f \qquad (3.2)$$

where h_f represents the fading channel among BS and far user, while z_f represents AWGN noise component at far user with zero mean and variance σ^2. Far user device is usually described by poor channel condition, and its signal x_f can be assigned additional power level by BS compared to other users. Thus, based on SIC procedure, the far user can directly decode its own signal x_f from received signal y_f. The expression for the received signal at far user device can be detailed as follows:

$$y_f = \sqrt{P_t}\left(\sqrt{\alpha_n}x_n + \sqrt{\alpha_m}x_m + \sqrt{\alpha_f}x_f\right)h_f + z_f \qquad (3.3)$$

$$y_f = \sqrt{P_t\alpha_f}x_fh_f + \sqrt{P_t}\left(\sqrt{\alpha_m}x_m + \sqrt{\alpha_n}x_n\right)h_f + z_f \qquad (3.4)$$

The first term in Equation (3.4) represents the desired signal for a far user, while the second term denotes the interference term from middle and near users. Based on Equation 3.4, the possible rate for far user could be expressed as shown [64]:

$$R_f = \log_2\left(1 + \frac{|h_f|^2 P_t \alpha_f}{|h_f|^2 P_t(\alpha_n + \alpha_m) + \sigma^2}\right) \qquad (3.5)$$

Likewise, the attainable rate for middle user can also be expressed as follows:

$$R_m = \log_2\left(1 + \frac{|h_m|^2 P_t \alpha_m}{|h_m|^2 P_t(\alpha_n) + \sigma^2}\right) \qquad (3.6)$$

In same manner, the near user device can be characterized by a good channel condition along with BS, thus, the near user signal x_n can be allocated a low power level, and hence, the received signal at the near user can be shown as:

$$y_n = x h_n + z_n \qquad (3.7)$$

$$y_n = \sqrt{P_t \alpha_n} x_n h_n + \sqrt{P_t}\left(\sqrt{\alpha_m} x_m + \sqrt{\alpha_f} x_f\right) h_n + z_n \quad (3.8)$$

Similarly, the first term in Equation 3.8 represents the near user target signal, while the second term denotes the interfering term from middle and far devices. On the other hand, it can be observed from Equation 3.8 that the interfering term can be major due to the additional power that could be assigned to the far user. Therefore, at the near user side, SIC should be performed, where far user signal x_f is recovered by immediate decoding, then it will be removed from the composite signal. Next, the middle user signal x_m should also decoded and removed from the remaining signal. Finally, the near user achieved rate R_n can be expressed as follows:

$$R_n = \log_2\left(1 + \frac{|h_n|^2 P_t \alpha_n}{\sigma^2}\right) \qquad (3.9)$$

3.3 Optimization Problem

The aim here is to maximize the sum rate for user devices in NOMA system based on optimizing the power factors for each user in accordance with the channel conditions. Based on the above-mentioned analysis, the possible sum rates for N users in downlink NOMA system can be written as follows [65]:

$$R_{sum} = \sum_{k=1}^{N} \log_2\left(1 + \frac{|h_k|^2 P_t \alpha_k}{|h_k|^2 \sum_{j=1}^{k-1} P_t \alpha_j + \sigma^2}\right) \qquad (3.10)$$

In our examined system, the constraints and the objective function considered for the optimization problem can be discussed as follows:

3.4 Power Constraints

The power assigned for each user device in the cell is a percentage of the total power P_t transmitted from base station; thus, the allocated power level for each user device must conforms with the following expression [10][66]:

$$\sum_{x=1}^{N} \alpha_x \leq 1 \qquad (3.11)$$

where α_x is the power assigned for the xth user in NOMA cell, and N represents number of users in the cell.

3.5 QoS Constraints

To support user fairness, it is assumed that the weak user in the examined NOMA cell must satisfy the QoS constraint, which indicates that a minimum rate R_{min} needs to be satisfied and this constraint can be expressed as follows [64]:

$$\log_2(1 + SINR_n) \geq R_{min} \qquad (3.12)$$

where $SINR_n$ is the signal to interference plus noise ratio (SINR) for n^{th} user and R_{min} is the minimum transmission rate required in the system [66]. This constraint can be simplified in many ways, suppose we have $R_{m \to k}$ which is the rate of user k to detect the signal of user m, where $1 \leq k \leq m$ and $R_m = R_{min}$. When user k is not able to detect the message of user m with rate R_{min}, this can be indicated as $R_{m \to k} < R_{min}$, the complement of this event can be formulated as follows:

$$\frac{|h_k|^2 P_t \alpha_m}{|h_k|^2 P_t \sum_{i=1}^{m-1} \alpha_i + \sigma^2} > (2^{R_{min}} - 1) \qquad (3.13)$$

where α_i is the power factor for i^{th} user in the system. By dividing both the numerator and

35

denominator of left-hand side of Equation 3.13 by the noise power σ^2, the expression in Equation 3.13 can be reformulated as:

$$|h_k|^2 \rho \left(\alpha_m - (2^{R_{min}} - 1) \sum_{i=1}^{m-1} \alpha_i \right) > (2^{R_{min}} - 1) \quad (3.14)$$

where ρ represent the ratio between signal power and noise power (SNR). Based on the aforementioned discussion and Equation 3.14, we can deduce an expression for the outage probability performance metric, which simply can be defined as the probability that the received information rate is less than the required threshold information rate. Hence, it is worth mentioning that based on Equation 3.14, we can conclude that in order to satisfy the minimum transmission rate and avoid that k^{th} user being in outage, the following condition for the power factor must be satisfied as follows [65][66]:

$$\alpha_k > (2^{R_{min}} - 1) \sum_{i=1}^{k-1} \alpha_i \quad (3.15)$$

3.6 Sum Rate optimization problem

Based on the constraints in Equation 3.11 and Equation 3.12 and sum rate representation, the standard optimization problem can be normally expressed as follows [10][66][67]:

$$\max_{\alpha} R_{sum} = \sum_{k=1}^{N} \log_2 \left(\frac{|h_k|^2 P_t \sum_{j=1}^{k-1} \alpha_j + \sigma^2 + |h_k|^2 P_t \alpha_k}{|h_k|^2 P_t \sum_{j=1}^{k-1} \alpha_j + \sigma^2} \right) \quad (3.16)$$

such that:

$$\sum_{x=1}^{N} \alpha_x \leq 1 \quad (3.17)$$

$$\log_2(1 + SINR_n) \geq R_{min} \quad (3.18)$$

$$\alpha_k \geq 0 \ \forall k = 1,2,\dots,N$$

3.7 Optimization Analysis

In this section, the optimization analysis is realized where three users are considered in the NOMA system, hence the objective function and the constraints can be simply reformulated as

follows [68]:

$$\max_{\alpha} \quad R_{Sum} = R_n + R_m + R_f \tag{3.19}$$

$S.t.$

$$(2^{R_{min}} - 1) - |h_x|^2 \rho \left(\alpha_x - (2^{R_{min}} - 1) \sum_{i=1}^{x-1} \alpha_i \right) \leq 0 \tag{3.20}$$

$$\alpha_n + \alpha_m + \alpha_f - 1 \leq 0 \tag{3.21}$$

$$\alpha_n, \alpha_m, \alpha_f \geq 0$$

where α_x is the power assigned for the xth and R_{min} is the minimum rate required in the system. According to the above analysis, the constraints can also be represented as follows:

$$C_1(\alpha) = \alpha_n + \alpha_m + \alpha_f - 1 \tag{3.22}$$

$$C_2(\alpha) = (2^{R_{min}} - 1) - \rho |h_f|^2 (\alpha_f - (2^{R_{min}} - 1)(\alpha_n + \alpha_m)) \tag{3.23}$$

$$C_3(\alpha) = (2^{R_{min}} - 1) - \rho |h_m|^2 (\alpha_m - (2^{R_{min}} - 1)(\alpha_n)) \tag{3.24}$$

The constraints $C_1(\alpha), C_2(\alpha)$ & $C_3(\alpha)$ are linear in terms of α, hence the constraints are convex. Therefore, $\frac{\partial R_{Sum}(\alpha)}{\partial \alpha}$ and $\frac{\partial^2 R_{Sum}(\alpha)}{\partial \alpha}$ can be calculated [69]. We can find the first derivative for $R_{Sum}(\alpha)$ with respect to each of the power coefficients α_n, α_m, and α_f. After some mathematical manipulation, $\frac{\partial R_{Sum}(\alpha)}{\partial \alpha}$ can be expressed as follows [10][68][69]:

$$\frac{\partial R_{Sum}}{\partial \alpha_n}$$

$$= \frac{1}{\ln 2} \left(\frac{|h_n|^2 P_t}{|h_n|^2 P_t \alpha_n + \sigma^2} - \frac{(|h_m|^2 P_t)^2 \alpha_m}{(|h_m|^2 P_t(\alpha_n + \alpha_m) + \sigma^2)(|h_m|^2 P_t(\alpha_n) + \sigma^2)} \right.$$

$$\left. - \frac{\left(|h_f|^2 P_t\right)^2 \alpha_f}{\left(|h_f|^2 P_t(\alpha_n + \alpha_m + \alpha_f) + \sigma^2\right)\left(|h_f|^2 P_t(\alpha_{n+}\alpha_m) + \sigma^2\right)} \right) \tag{3.25}$$

$$\frac{\partial R_{Sum}}{\partial \alpha_m}$$

$$= \frac{1}{ln2} \left(\frac{|h_m|^2 P_t}{|h_m|^2 P_t (\alpha_n + \alpha_m) + \sigma^2} \right.$$

$$\left. - \frac{\left(|h_f|^2 P_t\right)^2 \alpha_f}{\left(|h_f|^2 P_t (\alpha_n + \alpha_m + \alpha_f) + \sigma^2\right)\left(|h_f|^2 P_t (\alpha_{n+} \alpha_m) + \sigma^2\right)} \right) \quad (3.26)$$

$$\frac{\partial R_{Sum}}{\partial \alpha_f} = \frac{1}{ln2} \left(\frac{|h_f|^2 P_t}{|h_f|^2 P_t (\alpha_n + \alpha_m + \alpha_f) + \sigma^2} \right) \quad (3.27)$$

At this point, a general formula can be derived for the first derivative of the objective function in terms of α as follows:

$$\frac{\partial R_{Sum}(\alpha)}{\partial \alpha_i} = \frac{1}{ln2} \left(\frac{|h_i|^2 P_t}{|h_i|^2 P_t \sum_{j=1}^{i} \alpha_j + \sigma^2} \right)$$

$$- \frac{1}{ln2} \sum_{k=1}^{N-i} \left\{ \left(\frac{\left(|h_{(i+k)}|^2 P_t\right)^2 \alpha_{i+k}}{\left(|h_{(i+k)}|^2 P_t \sum_{j=1}^{i+k} \alpha_j + \sigma^2\right)} \right) \right.$$

$$\left. \times \left(\frac{1}{\left(|h_{(i+k)}|^2 P_t \sum_{j=1}^{i+k-1} \alpha_j + \sigma^2\right)} \right) \right\} \quad (3.28)$$

The second derivative of the objective function $\frac{\partial^2 R_{Sum}(\alpha)}{\partial \alpha}$ with respect to each of the power coefficients α_n, α_m, and α_f can also be derived as follows [10][68][69]:

$$\frac{\partial^2 R_{Sum}}{\partial \alpha_n^2}$$

$$= -\frac{1}{ln2}\left\{\left(\frac{(|h_n|^2 P_t)^2}{(|h_n|^2 P_t \alpha_n + \sigma^2)^2}\right)\right.$$

$$-\left(\frac{(|h_m|^2 P_t)^3 \alpha_m \left[2(|h_m|^2 P_t \alpha_n + \sigma^2) + |h_m|^2 P_t \alpha_m\right]}{(|h_m|^2 P_t(\alpha_n + \alpha_m) + \sigma^2)^2 \ (|h_m|^2 P_t(\alpha_n) + \sigma^2)^2}\right)$$

$$\left.-\left(\frac{\left(|h_f|^2 P_t\right)^3 \alpha_f \left[2\left(|h_f|^2 P_t(\alpha_n + \alpha_m) + \sigma^2\right) + |h_f|^2 P_t \alpha_f\right]}{\left(|h_f|^2 P_t(\alpha_n + \alpha_m + \alpha_f) + \sigma^2\right)^2 \ \left(|h_f|^2 P_t(\alpha_n + \alpha_m) + \sigma^2\right)^2}\right)\right\} \quad (3.29)$$

$$\frac{\partial^2 R_{Sum}}{\partial \alpha_m^2}$$

$$= -\frac{1}{ln2}\left\{\left(\frac{(|h_m|^2 P_t)^2}{(|h_m|^2 P_t(\alpha_n + \alpha_m) + \sigma^2)^2}\right)\right.$$

$$\left.-\left(\frac{\left(|h_f|^2 P_t\right)^3 \alpha_f \left[2\left(|h_f|^2 P_t(\alpha_n + \alpha_m) + \sigma^2\right) + |h_f|^2 P_t \alpha_f\right]}{\left(|h_f|^2 P_t(\alpha_n + \alpha_m + \alpha_f) + \sigma^2\right)^2 \ \left(|h_f|^2 P_t(\alpha_n + \alpha_m) + \sigma^2\right)^2}\right)\right\} \quad (3.30)$$

$$\frac{\partial^2 R_{Sum}}{\partial \alpha_f^2} = -\frac{1}{ln2}\left\{\left(\frac{\left(|h_f|^2 P_t\right)^2}{\left(|h_f|^2 P_t(\alpha_n + \alpha_m + \alpha_f) + \sigma^2\right)^2}\right)\right\} \quad (3.31)$$

A general formula can also be found for the second derivative of the objective function in terms of α as follows:

$$\frac{\partial^2 R_{Sum}(\alpha)}{\partial \alpha_i^2} = -\frac{1}{ln2}\left\{\left(\frac{(|h_i|^2 P_t)^2}{\left(|h_i|^2 P_t \sum_{j=1}^{i} \alpha_j + \sigma^2\right)^2}\right)\right.$$

$$-\sum_{k=1}^{N-i}\left\{\left(\frac{\left(|h_{(i+k)}|^2 P_t\right)^3 \alpha_{i+k} \left[2\left(|h_{(i+k)}|^2 P_t \sum_{j=1}^{k+i-1} \alpha_j + \sigma^2\right) + |h_{(i+k)}|^2 P_t \alpha_{i+k}\right]}{\left(|h_{(i+k)}|^2 P_t \sum_{j=1}^{i+k} \alpha_j + \sigma^2\right)^2}\right)\right.$$

$$\left.\left.\times\left(\frac{1}{\left(|h_{(i+k)}|^2 P_t \sum_{j=1}^{i+k-1} \alpha_j + \sigma^2\right)^2}\right)\right\}\right\} \quad (3.32)$$

3.8 Lagrange Function and Optimality Condition

Based on the objective function, and $\frac{\partial R_{Sum}(\alpha)}{\partial \alpha_i}$ and $\frac{\partial^2 R_{Sum}(\alpha)}{\partial \alpha_i^2}$, it can be proved that the objective function is concave and has a unique global maximum [4][10][12][69]. The Lagrange function and the KKT necessary conditions could be applied to achieve optimal power factors as follows [70]:

$$\mathcal{L}(\alpha_n, \alpha_m, \alpha_f, \mu_1, \mu_2, \mu_3) = R_{Sum} - \mu_1 C_1(\alpha) - \mu_2 C_2(\alpha) - \mu_3 C_3(\alpha) \quad (3.33)$$

where $\mu_1, \mu_2,$ and μ_3 represent Lagrange multipliers for the three user scenarios, and the optimality conditions can be written as follows [4][10][12][68]:

$$\frac{\partial R_{Sum}}{\partial \alpha_n} - \mu_1 \frac{\partial C_1(\alpha)}{\partial \alpha_n} - \mu_2 \frac{\partial C_2(\alpha)}{\partial \alpha_n} - \mu_3 \frac{\partial C_3(\alpha)}{\partial \alpha_n} = 0 \quad (3.34)$$

$$\frac{\partial R_{Sum}}{\partial \alpha_m} - \mu_1 \frac{\partial C_1(\alpha)}{\partial \alpha_m} - \mu_2 \frac{\partial C_2(\alpha)}{\partial \alpha_m} - \mu_3 \frac{\partial C_3(\alpha)}{\partial \alpha_m} = 0 \quad (3.35)$$

$$\frac{\partial R_{Sum}}{\partial \alpha_f} - \mu_1 \frac{\partial C_1(\alpha)}{\partial \alpha_f} - \mu_2 \frac{\partial C_2(\alpha)}{\partial \alpha_f} - \mu_3 \frac{\partial C_3(\alpha)}{\partial \alpha_f} = 0 \quad (3.36)$$

Slackness conditions can also be expressed as follows:

$$\mu_1(\alpha_n + \alpha_m + \alpha_f - 1) = 0 \quad (3.37)$$

$$\mu_2\left((2^{R_{min}} - 1) - \rho|h_f|^2(\alpha_f - (2^{R_{min}} - 1)(\alpha_n + \alpha_m))\right) = 0 \quad (3.38)$$

$$\mu_3\left((2^{R_{min}} - 1) - \rho|h_m|^2(\alpha_m - (2^{R_{min}} - 1)(\alpha_n))\right) = 0 \quad (3.39)$$

Lagrange multipliers also need to satisfy the following:

$$\mu_1 \geq 0, \quad \mu_2 \geq 0, \quad \mu_3 \geq 0 \quad (3.40)$$

In the subsequent steps, Lagrange multipliers should be proved to be positive. This could be accomplished as follows:

$$\frac{\partial C_1(\alpha)}{\partial \alpha_n} = \frac{\partial C_1(\alpha)}{\partial \alpha_m} = \frac{\partial C_1(\alpha)}{\partial \alpha_f} = 1 \tag{3.41}$$

$$\frac{\partial C_2(\alpha)}{\partial \alpha_n} = \frac{\partial C_2(\alpha)}{\partial \alpha_m} = \rho|h_f|^2(2^{R_{min}} - 1) \tag{3.42}$$

$$\frac{\partial C_3(\alpha)}{\partial \alpha_n} = \rho|h_m|^2(2^{R_{min}} - 1) \tag{3.43}$$

$$\frac{\partial C_2(\alpha)}{\partial \alpha_f} = -\rho|h_f|^2 \tag{3.44}$$

$$\frac{\partial C_3(\alpha)}{\partial \alpha_m} = -\rho|h_m|^2 \tag{3.45}$$

Based on Equations 3.41 up to 3.45, this can be substituted in the optimality conditions for Lagrange and obtain the following:

$$\frac{\partial R_{Sum}}{\partial \alpha_n} - \mu_1(1) - \mu_2\rho|h_f|^2(2^{R_{min}} - 1) - \mu_3\rho|h_m|^2(2^{R_{min}} - 1) = 0 \tag{3.46}$$

$$\frac{\partial R_{Sum}}{\partial \alpha_m} - \mu_1(1) - \mu_2\rho|h_f|^2(2^{R_{min}} - 1) - \mu_3(-\rho|h_m|^2) = 0 \tag{3.47}$$

$$\frac{\partial R_{Sum}}{\partial \alpha_f} - \mu_1(1) - \mu_2(-\rho|h_f|^2) - \mu_3(0) = 0 \tag{3.48}$$

To simplify the above expressions, Let $\beta 1 = \rho|h_f|^2(2^{R_{min}} - 1)$, $\beta 2 = \rho|h_m|^2(2^{R_{min}} - 1)$, $\gamma 1 = (-\rho|h_f|^2)$, and $\gamma 2 = (-\rho|h_m|^2)$. Therefore, the above written optimality conditions for Lagrange can be rewritten as:

$$\frac{\partial R_{Sum}}{\partial \alpha_n} - \mu_1 - \mu_2\,\beta 1 - \mu_3\,\beta 2 = 0 \tag{3.49}$$

$$\frac{\partial R_{Sum}}{\partial \alpha_m} - \mu_1 - \mu_2\,\beta 1 - \mu_3\,\gamma 2 = 0 \tag{3.50}$$

$$\frac{\partial R_{Sum}}{\partial \alpha_f} - \mu_1 - \mu_2\,\gamma 1 = 0 \tag{3.51}$$

After a few mathematical substitutions, the above written optimality conditions can be solved and rewritten as follows:

$$\left(\frac{\partial R_{Sum}}{\partial \alpha_m} - \frac{\partial R_{Sum}}{\partial \alpha_f}\right) - \left(\frac{\partial R_{Sum}}{\partial \alpha_n} - \frac{\partial R_{Sum}}{\partial \alpha_f}\right)\left(\frac{\gamma 2}{\beta 2}\right)$$

$$= \mu_2\left(-\gamma 1 + \beta 1 + (\gamma 1 - \beta 1)\left(\frac{\gamma 1}{\beta 1}\right)\right) \qquad (3.52)$$

Based on the fact that $|h_n|^2 > |h_m|^2 > |h_f|^2$, we can simply prove that $\left(\frac{\partial R_{Sum}}{\partial \alpha_m} - \frac{\partial R_{Sum}}{\partial \alpha_f}\right)$ and $\left(\frac{\partial R_{Sum}}{\partial \alpha_n} - \frac{\partial R_{Sum}}{\partial \alpha_f}\right)$ are positive and the left-hand side of Equation 3.52 is positive. Furthermore, since $\left(\frac{\gamma 1}{\beta 1}\right)$ are negative scalar, the right-hand side Equation 3.52 can be proved to be also positive, which concludes that μ_2 is positive.

Additionally, Equation 3.51 can be reformulated as follows:

$$\frac{\partial R_{Sum}}{\partial \alpha_f} - \mu_2\,\gamma 1 = \mu_1 \qquad (3.53)$$

where $\frac{\partial R_{Sum}}{\partial \alpha_f}$ is positive by inspection and $(\mu_2\gamma 1)$ is negative quantity; therefore, the left-hand side of Equation 3.53 must be positive, which implies that μ_1 is a positive quantity. Similarly, μ_3 can be proved to be positive value. In accordance with the above-mentioned analysis, the examined constraints are feasible, and the closed form representation for the power factors can be determined from the slackness conditions as follows [4][10][68]:

$$\alpha_n + \alpha_m + \alpha_f = 1 \qquad (3.54)$$

$$(2^{R_{min}} - 1) = \rho|h_f|^2\left(\alpha_f - (2^{R_{min}} - 1)(\alpha_n + \alpha_m)\right) \qquad (3.55)$$

$$(2^{R_{min}} - 1) = \rho|h_m|^2(\alpha_m - (2^{R_{min}} - 1)(\alpha_n)) \qquad (3.56)$$

To simplify the derivation, it can be assumed that $A_1 = (2^{R_{min}} - 1)$ and $A_2 = \rho|h_f|^2$, $A_3 = \rho|h_m|^2$, After performing some mathematical substitutions and arrangements, the closed

form expression for each of the power coefficients for the user devices can be derived as follows:

$$\alpha_f = \left(\frac{A_1}{A_2}\right)\left(\frac{1+A_2}{1+A_1}\right) = \left(\frac{(2^{R_{min}}-1)}{2^{R_{min}}}\right)\left(1+\frac{1}{\rho|h_f|^2}\right) \tag{3.57}$$

$$\alpha_m = \left(\frac{1+\frac{1}{A_3}-\alpha_f}{1+\frac{1}{A_1}}\right)$$

$$\alpha_m = \left(\left(\frac{(2^{R_{min}}-1)}{2^{R_{min}}}\right)\left(1+\frac{1}{\rho|h_m|^2}\right) - \left(\frac{2^{R_{min}}-1}{2^{R_{min}}}\right)^2\left(1+\frac{1}{\rho|h_f|^2}\right)\right) \tag{3.58}$$

$$\alpha_n = 1 - (\alpha_m + \alpha_f)$$

$$\alpha_n = \frac{\left(\frac{1}{A_2}\right)\left(\frac{1+A_2}{1+A_1}\right) - \left(\frac{A_3-A_1A_2}{A_2A_3}\right)}{1+A_1}$$

$$\alpha_n = \frac{1}{(2^{R_{min}})}\left(\left(\frac{1+\rho|h_f|^2}{(2^{R_{min}})\rho|h_f|^2}\right) + \left(\frac{(2^{R_{min}}-1)}{\rho|h_m|^2} - \frac{1}{\rho|h_f|^2}\right)\right) \tag{3.59}$$

3.9 MISO-NOMA System

Due to the QoS demands that critically required for future wireless networks, non-orthogonal multiple access (NOMA) supported by practical successive interference cancellation (SIC) at the receiver side can be considered as a motivating multiple access technique for sixth generation (6G) networks. Typically, power domain NOMA (PD-NOMA) is designed to handle signals that have significant difference in power levels, therefore, PD-NOMA is observed as one of the multiple access candidate's schemes that can achieve high spectral efficiency and huge connectivity in upcoming wireless networks. The integration of NOMA and multiple antenna system is also an appropriate scheme to improve the system performance, therefore analyzing NOMA system with multiple antenna structure can be inspiring towards identifying the substantial improvement that can be achieved in the achievable user's data rates [4][71].

3.9.1 Multiuser Environment

In this section, a multiuser environment with a one Base Station (BS) and multiple user devices (UDs) is inspected. The BS is supplied with N antennas and all the UDs are provided with a single antenna. The network is assumed to work with a specified time slots and each time slot involves one transmission, that includes either uplink or downlink transmissions. Generally, the pilot-assisted channel prediction is considered in our work, where pilot symbols can be identified by BS and UDs [4][11][72]. Usually, each user device initially transmits its pilot symbols to BS via an uplink channel, then, prior to downlink data transmission, the BS can inspect the pilot symbols and the accessible network information to facilitate estimating the downlink CSI. The main aim of this section is to derive and manage the power allocation scheme when N antennas are considered at the BS. Assume we have one BS with N antennas, we can represent the matrix of channel coefficients between BS and UD i as follows:

$$H_i = [h_{1i}; \; h_{2i}; \dots; h_{Ni}] \qquad (3.60)$$

where h_{ji} represents the vector of channel parameters from jth antenna at BS to the ith UD, with $j \in [1, 2, \dots, N]$ and $i \in [1, 2, \dots, M]$, where N is the number of antennas at BS and M is the number of users in MISO-NOMA cell. Furthermore, we can denote the data signal transmitted to UD i as

$$s_i = [s_{i1}, s_{i2}, \dots, s_{iK}] \qquad (3.61)$$

where K is the length of the signal. Then, the matrix that represents all the UD's sequences can be shown as

$$S = [s_1; s_2; \dots; s_M] \qquad (3.62)$$

The received kth signal at jth UD can be denoted as:

$$y_{kj} = \sum_{i=1}^{N} h_{ij} s_{ki} + z_{kj} \qquad (3.63)$$

where z_{kj} denotes the AWGN with zero mean and variance σ^2 at jth UD through kth signal duration. The received kth symbol at all UDs is:

$$Y_k = \sum_{i=1}^{N} h_i \, s_{ki} + z_k \qquad (3.64)$$

where

$$Y_k = [y_{k1}; y_{k2}; \, \dots \, ; y_{kM}] \qquad (3.65)$$

$$z_k = [z_{k1}; z_{k2}; \, \dots ; z_{kM}] \qquad (3.66)$$

Many of the current research count on pilot symbols to approximate the uplink channel parameters and then utilize reciprocity of the channel to realize the prediction of the downlink channel weights [72]. These such schemes for predicting the CSI may not be reliable, especially in circumstances of inadequate channel reciprocity owing to hardware limitations. Besides, this kind of estimator may introduce estimation errors in case the uplink and downlink channel parameters are not stationary within a certain transmission.

As stated before, in our work, we manage to get assistance from pilot symbols, and network information to explicitly predict the downlink channel parameters. The set of estimated channel coefficients between the BS and M user devices can be shown as

$$\hat{\mathbf{H}} = \left[\hat{\mathbf{H}}_1; \ \hat{\mathbf{H}}_2; \dots; \hat{\mathbf{H}}_M\right] \qquad (3.67)$$

where $\hat{\mathbf{H}}_i$ is the predicted channel coefficients between the ith UD and the BS that contains N antennas and $\hat{\mathbf{H}}_i$ can be expressed as follows:

$$\hat{\mathbf{H}}_i = \left[\hat{h}_{1i}; \ \hat{h}_{2i}; \dots; \hat{h}_{Ni}\right] \qquad (3.68)$$

where $\hat{h}_{ji}$ represents the vector of predicted channel parameters between j^{th} antenna at BS and the ith UD.

3.9.2 MISO-NOMA System Model

The fundamental concept of NOMA system is to provide a non-orthogonal resource allocation between user devices while increasing the processing at the receiver side. With non-orthogonal resource allocation, NOMA can achieve a massive connectivity and accomplish high spectral efficacy. Existing research on NOMA schemes mainly focuses on the code domain and power domain. In the code domain NOMA, featured spread spectrum codes are designated for

different users and then these users can be multiplexed over the same time-frequency resource block. In the power domain NOMA, the transmitter can superimpose signals with different power levels to be sent to several users on the shared spectrum. At the receiver, each user equipment can decode his own desired signal by means of successive interference cancellation (SIC) [72].

In this section, the downlink MISO-NOMA system is analysed where each user device is linked by different fading channels to N antennas at the BS. To simplify our analysis, NOMA cell is considered with one BS, and the BS has two antennas, while each user device has one antenna. As stated before, in PD-NOMA, each user device will receive the superimposed signal sent from BS which involves the target and interfering signals sent through the same resources. Thus, combining different signals supported by unique power levels is critical to distinguish these signals and strengthen the successive interference cancellation (SIC) procedure. The system structure for the basic components that describe the examined MISO-NOMA system is shown in Figure 3.1.
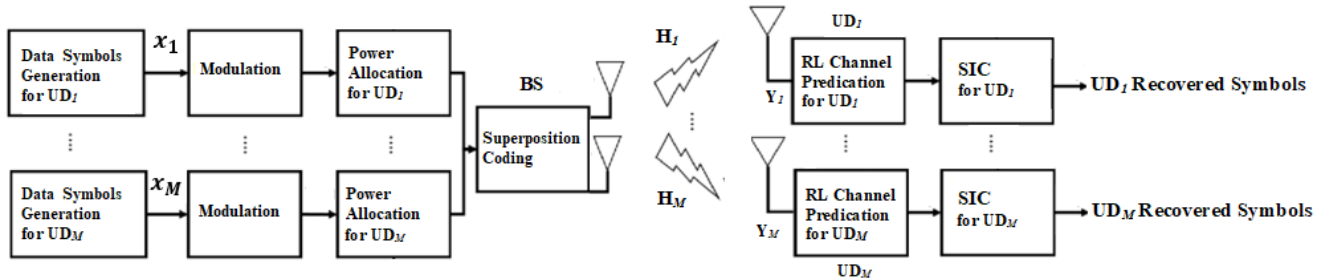
Figure 3.1 MISO-NOMA system basic Structure.

In this section, MISO-NOMA cell will be analysed, with three user devices are considered in the cell. The examined user devices can be characterized based on their fading channels models and the distances from BS. Fading channels with Rayleigh distribution are adopted to describe the channel model for every user in the MISO-NOMA cell. The user equipment at the boundary of the cell is realized as a far user, while the nearest user equipment is specified as a near user. The examined MISO-NOMA cell includes three user devices and the fading path for every user can be observed as follows [4][11]: $\boldsymbol{h}_n \sim (0, d_n^{-k})$ for near users, $\boldsymbol{h}_m \sim (0, d_m^{-k})$ for the middle user, and $\boldsymbol{h}_f \sim (0, d_f^{-k})$ for the user at the edge of the cell, where $\boldsymbol{h}_i$ implies the vector that represents the fading channel coefficients among BS and user i, k represents the path loss

exponent, and AWGN is considered with noise power indicated as σ^2. In terms of channel gains, the relation between user devices can also be indicated as $|h_n|^2 > |h_m|^2 > |h_f|^2$ [4][20], and the overall power transmitted from BS to all users in the cell is labelled as P_t. Each user device has a receiver that can activate the SIC process to get rid of signals related to other devices with bad channel characteristics. In contrast, signals related to user devices with good link conditions may not be eliminated and regarded as interference. According to the aforementioned assumptions, the superposition-coded signal x sent from BS can be stated as follows:

$$x = \sqrt{P_t}\left(\sqrt{\eta_n}x_n + \sqrt{\eta_m}x_m + \sqrt{\eta_f}x_f\right) \qquad (3.69)$$

where η_f, η_m and η_n represent power factors given for a far user device, middle user device, and near user device separately. Furthermore, x_f, x_m and x_n refer to the information signals related to far, middle, and near users respectively. The downlink received signal at a far user device in the MISO-NOMA cell can be written as:

$$y_f = xh_{f1} + xh_{f2} + z_f \qquad (3.70)$$

where h_{f1} represents the fading channel among far user device and the 1st antenna at BS, h_{f2} represents the fading channel between the far user device and 2nd antenna at BS and z_f is the AWGN noise component at the far user device with mean zero and variance σ^2. The far user is denoted by weak channel condition, therefore signal x_f is usually given additional power level by BS where $\eta_f > \eta_m > \eta_n$. The received signal at a far user device can be shown in detail as follows:

$$y_f = \sqrt{P_t\eta_f}x_f\left(h_{f1} + h_{f2}\right) + \left(\sqrt{P_t\eta_m}x_m + \sqrt{P_t\eta_n}x_n\right)\left(h_{f1} + h_{f2}\right) + z_f \quad (3.71)$$

The 1st term in Equation 3.71 implies the target signal for far user and the 2nd term indicates to the interference term coming from other user devices. Based on the above-mentioned analysis, the possible bit rate for a far user in MISO-NOMA system could be written as follows [4][11][72]:

$$R_f = log_2 \left(1 + \frac{|\boldsymbol{h}_{f1} + \boldsymbol{h}_{f2}|^2 P_t \eta_f}{|\boldsymbol{h}_{f1} + \boldsymbol{h}_{f2}|^2 P_t (\eta_n + \eta_m) + \sigma^2} \right) \qquad (3.72)$$

Similarly, the near user equipment has a good channel status alongside BS, therefore, a low power level can be assigned to x_n, and the received signal for the near user can be stated as:

$$y_n = \sqrt{P_t \eta_n} x_n (\boldsymbol{h}_{n1} + \boldsymbol{h}_{n2}) + \left(\sqrt{P_t \eta_m} x_m + \sqrt{P_t \eta_f} x_f \right) (\boldsymbol{h}_{n1} + \boldsymbol{h}_{n2}) + z_n \quad (3.73)$$

In Equation 3.73, the 1st term represents the expected near user signal, and the 2nd term implies interference from other user devices. It can be noted from Equation 3.73, that the interference can affect the received signal, since that the far user may be assigned a further power level. To recover the desired signal at the near user device, SIC is activated, where direct decoding for the far user signal x_f is implemented first, then eliminated the decoded signal from the aggregate signal. Afterwards, the middle device signal x_m is decoded and remove it from the resultant signal. Finally, the potential rate for the near user R_n can be shown as:

$$R_n = log_2 \left(1 + \frac{|\boldsymbol{h}_{n1} + \boldsymbol{h}_{n2}|^2 P_t \eta_n}{\sigma^2} \right) \qquad (3.74)$$

3.10 Optimization Problem Characterization

The key point here is to maximize the sum rates for the user devices in the examined MISO-NOMA cell. Sum rate maximization will be considered based on optimizing the power allocation factor for each user device while take in consideration the total power and the minimum rate constraints. In the examined downlink MISO-NOMA, the objective function or the sum rates for M user devices can be expressed as [4][11][66][67]:

$$R_{sum} = \sum_{i=1}^{M} log_2 \left(1 + \frac{|\boldsymbol{h}_{i1} + \boldsymbol{h}_{i2}|^2 P_t \eta_i}{|\boldsymbol{h}_{i1} + \boldsymbol{h}_{i2}|^2 P_t \sum_{j=1}^{i-1} \eta_j + \sigma^2} \right) \qquad (3.75)$$

In the subsequent analysis, the constraints will be presented as follows:

3.10.1 Power Constraint

The power given for every user device in the MISO-NOMA cell is a portion of the whole power

P_t sent from the BS, therefore the power factor for each device must conform with:

$$\sum_{i=1}^{M} \eta_i \leq 1 \qquad (3.76)$$

where η_i is the power percentage allocated for the ith user.

3.10.2 QoS Constraints

In our analysis, we assume that all the user devices in the examined MISO-NOMA cell need to realize a QoS requirement where the minimum rate R_{min} is required to be fulfilled in the system, this constraint can be expressed as follows:

$$Log_2(1 + SINR_i) \geq R_{min} \qquad (3.77)$$

where $SINR_i$ is the signal to interference plus noise ratio for ith user and R_{min} is the minimum required transmission rate in the examined MISO-NOMA cell. Equation 3.77 can be reformulated as follows [4][11][67][72]:

$$|\boldsymbol{h}_{i1} + \boldsymbol{h}_{i2}|^2 \rho \left(\eta_i - (2^{R_{min}} - 1) \sum_{j=1}^{i-1} \eta_j \right) > (2^{R_{min}} - 1) \qquad (3.78)$$

where ρ represents the SNR.

3.11 Optimization Framework

The primary goals in this section can be listed as follows: (1) introduce the objective function and the constraints in a standard form, (2) derive a general expression for the 1st and 2nd derivative for the objective function, (3) based on the mathematical analysis and the derived formulas, we can inspect that $\frac{\partial^2 R_{Sum}}{\partial \eta_i^2}$ is a negative, which supports that the objective function is a concave with a distinctive global maximum (4) finally, we can deduce the optimal power factors for each user in the MISO-NOMA cell based on applying the Lagrange function and the optimality conditions [4][11][69][70].

Based on the objective function in Equation 3.75 and the constraints in Equations 3.76 and 3.77, the standard optimization problem can be generally reformulated as follows:

$$\max_{\eta} R_{sum} = \sum_{i=1}^{M} log_2 \left(\frac{|\boldsymbol{h}_{i1} + \boldsymbol{h}_{i2}|^2 P_t \sum_{j=1}^{i-1} \eta_j + \sigma^2 + |\boldsymbol{h}_{i1} + \boldsymbol{h}_{i2}|^2 P_t \eta_i}{|\boldsymbol{h}_{i1} + \boldsymbol{h}_{i2}|^2 P_t \sum_{j=1}^{i-1} \eta_j + \sigma^2} \right) \quad (3.79)$$

such that

$$\sum_{j=1}^{M} \eta_j \leq 1 \quad (3.80)$$

$$(2^{Rmin} - 1) - \rho |\boldsymbol{h}_{i1} + \boldsymbol{h}_{i2}|^2 \left(\eta_i - (2^{Rmin} - 1) \sum_{j=1}^{i-1} \eta_j \right) \leq 0 \quad (3.81)$$

$$\eta_i \geq 0 \quad \forall i = 1, 2, \dots, M$$

Similarly, in this part, the power optimization framework for the MISO-NOMA cell is accomplished with regards to three user devices. Therefore, the examined constraints can be also represented as shown:

$$\psi_1(\eta) = \eta_n + \eta_m + \eta_f - 1 \quad (3.82)$$

$$\psi_2(\eta) = (2^{Rmin} - 1) - \rho |\boldsymbol{h}_{f1} + \boldsymbol{h}_{f2}|^2 (\eta_f - (2^{Rmin} - 1)(\eta_m + \eta_n) \quad (3.83)$$

$$\psi_3(\eta) = (2^{Rmin} - 1) - \rho |\boldsymbol{h}_{m1} + \boldsymbol{h}_{m2}|^2 (\eta_m - (2^{Rmin} - 1)(\eta_n) \quad (3.84)$$

Since the constraints $\psi_1(\eta), \psi_2(\eta)$ & $\psi_3(\eta)$ are linear in terms of η, they are considered convex. One way to prove that the objective function R_{Sum} is concave with a distinctive global maximum, is to find a general expression for the first derivative $\frac{\partial R_{Sum}}{\partial \eta_i}$ and the second derivative $\frac{\partial^2 R_{Sum}}{\partial \eta_i^2}$ of the objective function. The first derivative of the objective function can be derived in general form as follows [4][11][69][70]:

$$\frac{\partial R_{Sum}}{\partial \eta_i}$$

$$= \frac{1}{\ln 2}\left(\frac{|\boldsymbol{h}_{i1} + \boldsymbol{h}_{i2}|^2 P_t}{|\boldsymbol{h}_{i1} + \boldsymbol{h}_{i2}|^2 P_t \sum_{j=1}^{i}\eta_j + \sigma^2}\right)$$

$$-\frac{1}{\ln 2}\sum_{k=1}^{M-i}\left\{\left(\frac{\left(|\boldsymbol{h}_{(i+k)1} + \boldsymbol{h}_{(i+k)2}|^2 P_t\right)^2 \eta_{i+k}}{\left(|\boldsymbol{h}_{(i+k)1} + \boldsymbol{h}_{(i+k)2}|^2 P_t \sum_{j=1}^{i+k}\eta_j + \sigma^2\right)}\right)\right.$$

$$\left.\times\left(\frac{1}{\left(|\boldsymbol{h}_{(i+k)1} + \boldsymbol{h}_{(i+k)2}|^2 P_t \sum_{j=1}^{i+k-1}\eta_j + \sigma^2\right)}\right)\right\} \qquad (3.85)$$

Similarly, the second derivative of the objective function can be derived in general form as follows:

$$\frac{\partial^2 R_{Sum}}{\partial \eta_i^2} = -\frac{1}{\ln 2}\left\{\left(\frac{(|\boldsymbol{h}_{i1} + \boldsymbol{h}_{i2}|^2 P_t)^2}{(|\boldsymbol{h}_{i1} + \boldsymbol{h}_{i2}|^2 P_t \sum_{j=1}^{i}\eta_j + \sigma^2)^2}\right)\right.$$

$$-\sum_{k=1}^{M-i}\left\{\left(|\boldsymbol{h}_{(i+k)1} + \boldsymbol{h}_{(i+k)2}|^2 P_t\right)^3 \eta_{i+k}\right.$$

$$\times\left(\frac{\left[2\left(|\boldsymbol{h}_{(i+k)1} + \boldsymbol{h}_{(i+k)2}|^2 P_t \sum_{j=1}^{k+i-1}\eta_j + \sigma^2\right) + |\boldsymbol{h}_{(i+k)1} + \boldsymbol{h}_{(i+k)2}|^2 P_t \eta_{i+k}\right]}{\left(|\boldsymbol{h}_{(i+k)1} + \boldsymbol{h}_{(i+k)2}|^2 P_t \sum_{j=1}^{i+k}\eta_j + \sigma^2\right)^2}\right)$$

$$\left.\left.\times\left(\frac{1}{\left(|\boldsymbol{h}_{(i+k)1} + \boldsymbol{h}_{(i+k)2}|^2 P_t \sum_{j=1}^{i+k-1}\eta_j + \sigma^2\right)^2}\right)\right\}\right\} \qquad (3.86)$$

Based on the above mathematical analysis and the derived formulas, we can check that $\frac{\partial^2 R_{Sum}}{\partial \eta_i^2}$ is a negative function, which verifies that the objective function is a concave with a distinctive maximum [4][69][70]. As demonstrated in the previous section, to derive the optimal power factors, the Lagrange function and the KKT optimality conditions can be applied [4][11][69][70]:

$$\mathcal{L}(\eta_n, \eta_m, \eta_f, \mu_1, \mu_2, \mu_3) = R_{Sum} - \mu_1\psi_1(\eta) - \mu_2\psi_2(\eta) - \mu_3\psi_3(\eta) \qquad (3.87)$$

where μ_1, μ_2, and μ_3 represent Lagrange multipliers for the 3 users' scenario. Optimality

conditions can be written as follows:

$$\frac{\partial R_{Sum}}{\partial \eta_n} - \mu_1 \frac{\partial \psi_1(\eta)}{\partial \eta_n} - \mu_2 \frac{\partial \psi_2(\eta)}{\partial \eta_n} - \mu_3 \frac{\partial \psi_3(\eta)}{\partial \eta_n} = 0 \qquad (3.88)$$

$$\frac{\partial R_{Sum}}{\partial \eta_m} - \mu_1 \frac{\partial \psi_1(\eta)}{\partial \eta_m} - \mu_2 \frac{\partial \psi_2(\eta)}{\partial \eta_m} - \mu_3 \frac{\partial \psi_3(\eta)}{\partial \eta_m} = 0 \qquad (3.89)$$

$$\frac{\partial R_{Sum}}{\partial \eta_f} - \mu_1 \frac{\partial \psi_1(\eta)}{\partial \eta_f} - \mu_2 \frac{\partial \psi_2(\eta)}{\partial \eta_f} - \mu_3 \frac{\partial \psi_3(\eta)}{\partial \eta_f} = 0 \qquad (3.90)$$

Following the same mathematical analysis and manipulations performed in previous section, and given the fact that $|\boldsymbol{h}_n|^2 > |\boldsymbol{h}_m|^2 > |\boldsymbol{h}_f|^2$, we can demonstrate that the analyzed constraints are feasible and after a few further mathematical simplifications, the closed form expression for the power factors η_f, η_m, and η_n can also be derived and deduced as follows [4][11][70]:

$$\eta_f = \left(\frac{(2^{R_{min}} - 1)}{2^{R_{min}}} \right) \left(1 + \frac{1}{\rho |\boldsymbol{h}_{f1} + \boldsymbol{h}_{f2}|^2} \right) \qquad (3.91)$$

$$\eta_m = \left(\left(\frac{(2^{R_{min}} - 1)}{2^{R_{min}}} \right) \left(1 + \frac{1}{\rho |\boldsymbol{h}_{m1} + \boldsymbol{h}_{m2}|^2} \right) - \left(\frac{2^{R_{min}} - 1}{2^{R_{min}}} \right)^2 \left(1 + \frac{1}{\rho |\boldsymbol{h}_{f1} + \boldsymbol{h}_{f2}|^2} \right) \right) \qquad (3.92)$$

$$\eta_n = 1 - (\eta_m + \eta_f)$$

$$\eta_n = \frac{1}{(2^{R_{min}})} \left(\left(\frac{1 + \rho |\boldsymbol{h}_{f1} + \boldsymbol{h}_{f2}|^2}{(2^{R_{min}})\rho |\boldsymbol{h}_{f1} + \boldsymbol{h}_{f2}|^2} \right) + \left(\frac{(2^{R_{min}} - 1)}{\rho |\boldsymbol{h}_{m1} + \boldsymbol{h}_{m2}|^2} - \frac{1}{\rho |\boldsymbol{h}_{f1} + \boldsymbol{h}_{f2}|^2} \right) \right) \qquad (3.93)$$

3.12 Simulation Results and Discussion

In this section, system simulation is conducted to evaluate the efficiency of incorporating the optimum power coefficients that derived in section 3.11 and compare the results for that optimized scheme with the fixed power scheme scenario. Simulation files are considering a single base station with two antennas serving three users and each user is equipped with one antenna and the simulation is accomplished over 10^6 random channel generations. In fixed power scenario, the power factors are allocated as follows: $\alpha_n = 0.1$, $\alpha_m = 0.2$ while

$\alpha_f = 0.7$. The fading channels between each user and the BS is generated based on flat Rayleigh fading distribution, and the channel parameters are changing at each transmission time slot. The transmitted signals are modulated using Quadrature phase shift keying (QPSK) and the path loss is assigned to 4. Bandwidth of the system is $B = 1$ MHz, the noise spectral density is $N_0 = -174$ dBm/Hz and R_{min} is specified to be 1 bps. Figure 3.2 shows results for two distinct simulation environments: the first environment primarily relies on applying a fixed power scheme for each user device in the system. The second environment depends on employing the optimized power scheme that is derived in section 3.11 in this chapter. Simulation outcomes for far and middle users prove the superiority of the optimized power policy over the fixed power scheme in terms of BER versus transmitted power. For the near user results, the fixed power scheme provides some sort of improvement in terms of the received bits error over the optimized power scheme, and this could be clarified that for the near user environment, the stable channel condition is more valuable than the allocated power.
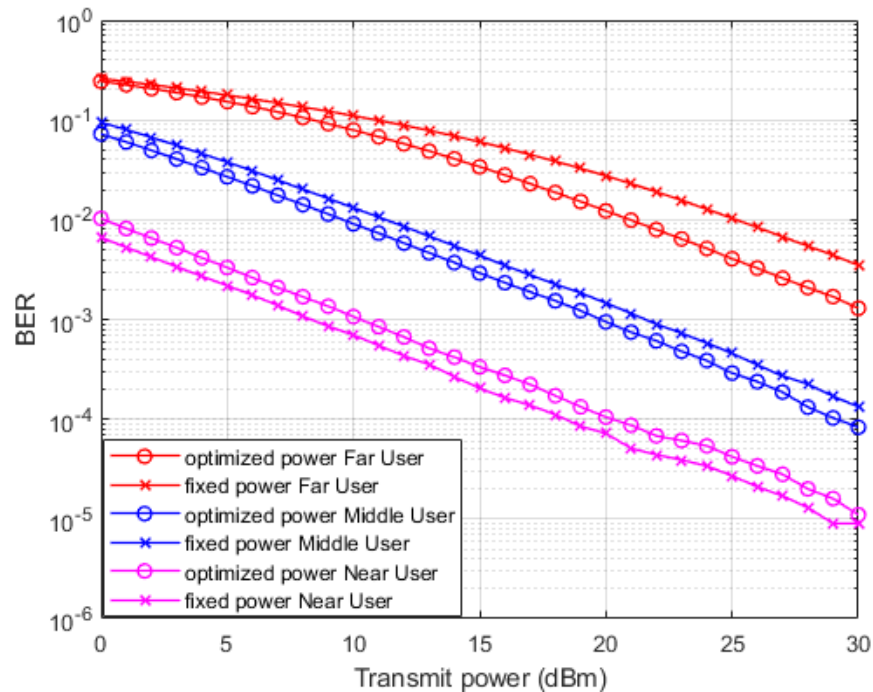
Figure 3.2 BER vs power (optimized power - fixed power).

Figure 3.3 shows the outage probability results versus the power for far, middle, and near users when the optimized power and the fixed power scenarios are implemented. Far user outcomes clearly show an improvement in outage probability and also a power saving is recorded to

approximately 5 dBm when the optimized power scheme is employed compared to the fixed power scheme. Likewise, for the middle user case, the optimized power scheme provides an obvious enhancement in the outage probability with a power saving nearly 3 dBm. Otherwise, the near user with the fixed power scenario, shows a noticeable enhancement in outage compared to the optimized power case. These outcomes also approve the results attained for the BER performance metric in Figure 3.2, which confirm that fixed power scheme jointly with high channel gain, are more adequate for near user than the optimized power scheme.
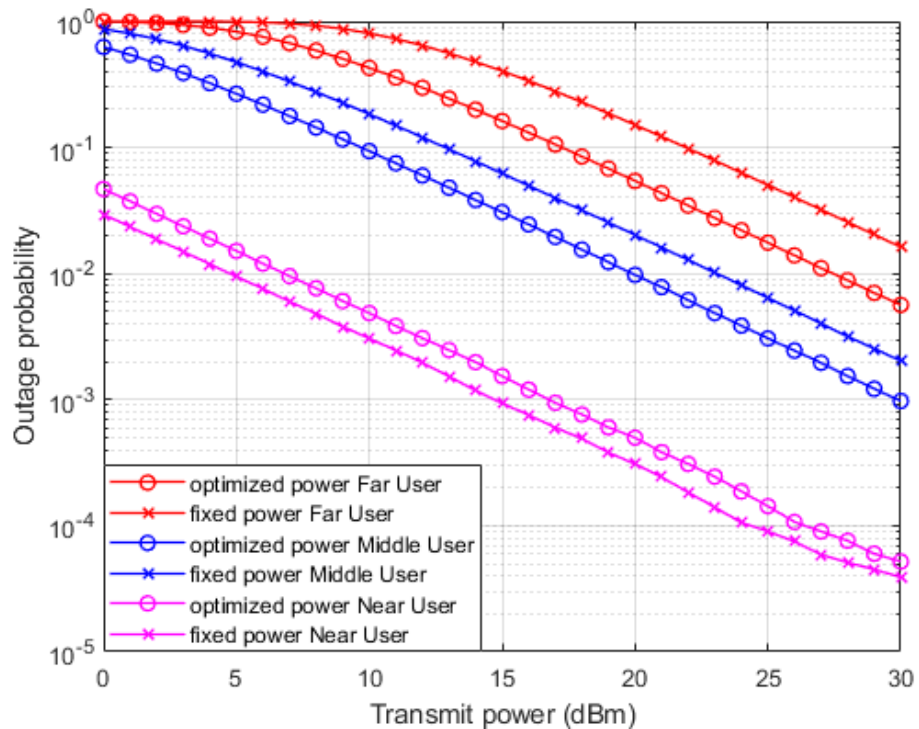
Figure 3.3 Outage probability vs power (optimized power - fixed power).

In Figure 3.4, the simulation outcomes for the sum rate for the examined users in the inspected NOMA cell are shown versus transmitted power. Both of the optimized power scheme and the fixed power scheme is incorporated with NOMA parameters prior to calculate the rate for each user device. Based on the simulation results, it can be observed that the examined NOMA cell combined with the optimized power scheme indicate a little improvement in the sum rate compared to the fixed power scenario when the applied power level is low. Starting from 15 dBm, both the fixed power scheme and the optimized power structure provide a comparable sum rate at high transmitted power levels.
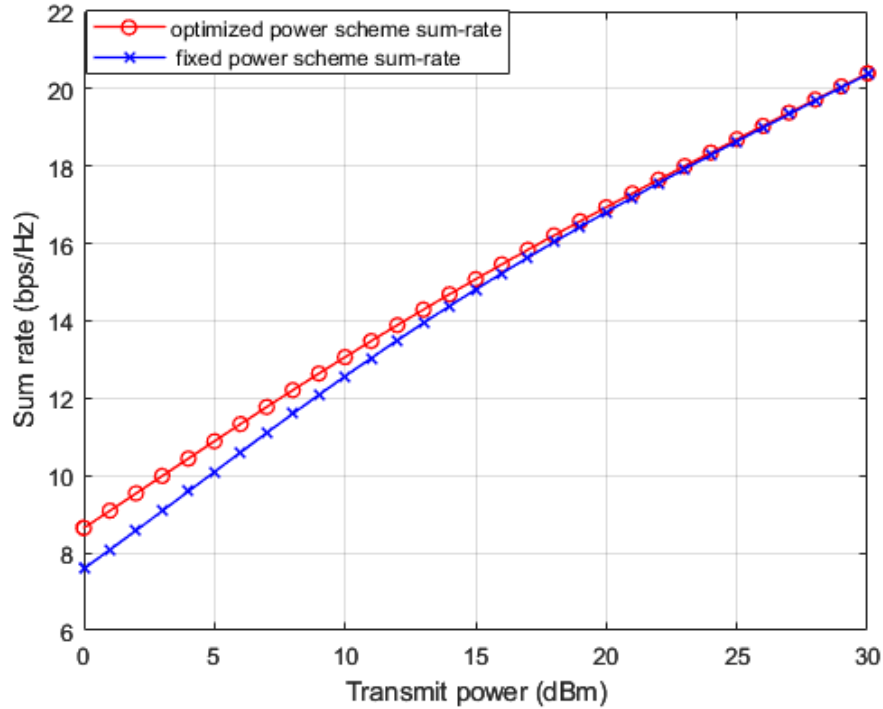
Figure 3.4. Sum rate vs power (optimized power - fixed power).

3.13 SIMO-NOMA System Model

In this section, downlink SIMO-NOMA system is investigated, where all users encounter diverse channel gains. SIMO-NOMA cell is examined where a single base station with one antenna is considered to serve three users concurrently on the same time-frequency block and each user device is equipped with two antennas. In NOMA system, each user receives the joint signal sent by the BS, and this combined signal includes the desired and interfering signals transmitted via same resource block, thus multiplexing numerous signals using various power levels is demanding to enable discriminating the signals and to enhance the successive interference cancellation (SIC) process at the receiver side [73]. In downlink NOMA system, users that labelled by high channel conditions are frequently given low power level while user devices with low channel gains are allocated high power levels. At each transmission, each user can be characterised according to its distance from BS. The nearest user is indicated as near user and the outmost user or user at the cell edge is regarded as far user. In this section, also Rayleigh fading channel with zero mean is assumed for the communication links between the antenna at the BS and user's antennas at the receiver side.

The fading channel between the BS and each user device can be mathematically represented

as $h_i \sim (0, d_i^{-k})$ where d_i is the distance between the BS and the ith user and k represents the path loss exponent. Additive white Gaussian noise (AWGN) is considered with zero mean and σ^2 variance. Typically, and without loss of generality, it can be assumed that $|h_n|^2 > |h_m|^2 > |h_f|^2$, where h_n, h_m and h_f are the fading channels amongst the BS and near user, middle user and far user respectively. It also assumed that the channel state information (CSI) is known and entire transmitted power from BS to users is indicated as P_t. As showed before, each receiver has the capability to perform SIC to eliminate signals related to other users with weak channel conditions. Whereas signals from users with robust channel gains cannot be taken away and can be managed as interference. In this section, it is assumed that there is one antenna at the BS and two antennas at each user equipment at the receiver side [73] [74].

Based on the above-mentioned analysis, we can introduce the following terms as follows:h_{1f} denotes the fading channel between the BS and the first receiving antenna at far user device. Similarly, h_{2f} denotes the fading channel between the BS and the second receiving antenna at far user device. The antenna at the BS can transmit the superposition coded signal x which can be written as [10][11][68][73]:

$$x = \sqrt{P_t}\left(\sqrt{\alpha_f}x_f + \sqrt{\alpha_m}x_m + \sqrt{\alpha_n}x_n\right) \qquad (3.94)$$

where α_f, α_m and α_n are the power factors for far, middle, and near users respectively. Also x_f, x_m and x_n represent the target symbols related to far, middle, and near users respectively. Therefore, the received signal at the far user can be formed as follows [73]:

$$y_{1f} = h_{1f}x + n_{1f} \qquad (3.95)$$

$$y_{2f} = h_{2f}x + n_{2f} \qquad (3.96)$$

where y_{1f} and y_{2f} are the received signals at the far user side via two dissimilar fading paths and n_{1f}, n_{2f} represent the corresponding noise components at the far user side. Noise components at each receiver are considered as a gaussian in nature with zero mean and σ^2 variance. Furthermore, it is assumed that $E\{n_{1f}n_{2f}\} = 0$ which basically implies that these noise components are uncorrelated. According to our examined SIMO-NOMA system, the received signals at the far user can be also represented as a vector form as follows:

$$\begin{bmatrix} y_{1f} \\ y_{2f} \end{bmatrix} = \begin{bmatrix} h_{1f} \\ h_{2f} \end{bmatrix} x + \begin{bmatrix} n_{1f} \\ n_{2f} \end{bmatrix} \tag{3.97}$$

$$\overline{y_f} = \overline{h_f} x + \overline{n_f} \tag{3.98}$$

Similarly, the vector representation for the received signals at the near user can be shown as follows [73]:

$$\begin{bmatrix} y_{1n} \\ y_{2n} \end{bmatrix} = \begin{bmatrix} h_{1n} \\ h_{2n} \end{bmatrix} x + \begin{bmatrix} n_{1n} \\ n_{2n} \end{bmatrix} \tag{3.99}$$

$$\overline{y_n} = \overline{h_n} x + \overline{n_n} \tag{3.100}$$

Based on the aforementioned analysis, there are two received signals y_1 and y_2 at any user receiver, and these two received signals can be combined into a new resultant signal represented by new variable $\tilde{y}$ and mathematically can be represented as follows [73] [74]:

$$\tilde{y} = w_1^* y_1 + w_2^* y_2 \tag{3.101}$$

where w_1^* and w_2^* are called complex combining weights, and in vector form the composite resultant signal $\tilde{y}$ can be shown as in (3.102).

$$\tilde{y} = [w_1^* \quad w_2^*] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \overline{w}^H \overline{y} \tag{3.102}$$

where $\overline{w}^H$ is the Hermitian combining weight vector. By substituting vector $\overline{y}$ from (3.100) into (3.102), the following is obtained:

$$\tilde{y} = \overline{w}^H (\overline{h} x + \overline{n}) = \overline{w}^H \overline{h} x + \overline{w}^H \overline{n} \tag{3.103}$$

The first term on the right-hand side of (3.103) represents the signal component while the second term represents the noise component. Hence, the signal to noise ratio (SNR) at the receiver side for any user in the proposed SIMO-NOMA system can be shown as [73][75]

$$SNR = \frac{\left| \overline{w}^H \overline{h} \right|^2 P_t \alpha}{E\left\{ \left| \overline{w}^H \overline{n} \right|^2 \right\}} \tag{3.104}$$

where P_t is the entire transmitted power from BS and α is the assigned power factor for each user in the cell. The denominator of (3.104) can be simplified as follows:

$$E\left\{ \left| \overline{w}^H \overline{n} \right|^2 \right\} = E\{(w_1^* n_1 + w_2^* n_2)^2\}$$

$$= E\{(w_1^{*2}n_1{}^2 + w_2^{*2}n_2{}^2 + 2w_1^*w_2^*n_1n_2)\}$$

$$= \{(w_1^{*2}E(n_1{}^2) + w_2^{*2}E(n_2{}^2))\} = (w_1^{*2}\sigma^2 + w_2^{*2}\sigma^2)$$

$$= \sigma^2(w_1^{*2} + w_2^{*2}) = \sigma^2 \left\|\overline{w}^H\right\|^2 \qquad (3.105)$$

where $\|\overline{w}\|$ represents the norm for the combining vector $\overline{w}$. Substituting (3.105) into (3.104) yields a new expression for SNR which can be written as:

$$SNR = \frac{\left|\overline{w}^H\overline{h}\right|^2 P_t\alpha}{E\left\{\left|\overline{w}^H\overline{n}\right|^2\right\}} = \frac{P_t\alpha\left|\overline{w}^H\overline{h}\right|^2}{\sigma^2 \left\|\overline{w}^H\right\|^2} \qquad (3.106)$$

Also, $\overline{w}^H\overline{h}$ can be factorized as follows:

$$\overline{w}^H\overline{h} = w_1^*h_1 + w_2^*h_2 = \overline{w}^H.\overline{h} \qquad (3.107)$$

where $\overline{w}^H.\overline{h}$ is the dot product between $\overline{w}^H$ and $\overline{h}$, and according to that, $\left|\overline{w}^H.\overline{h}\right|^2$ can be factorized as:

$$\left|\overline{w}^H.\overline{h}\right|^2 = \left\|\overline{w}^H\right\|^2 \|\overline{h}\|^2 cos^2\theta \qquad (3.108)$$

where θ is the angle between the vectors $\overline{w}$ and $\overline{h}$. Substituting (3.108) into (3.106) and the SNR can now be expressed as follows:

$$SNR = \frac{P_t\alpha\left|\overline{w}^H\overline{h}\right|^2}{\sigma^2 \left\|\overline{w}^H\right\|^2} = \frac{P_t\alpha\left\|\overline{w}^H\right\|^2 \|\overline{h}\|^2 cos^2\theta}{\sigma^2 \left\|\overline{w}^H\right\|^2} \qquad (3.109)$$

$$SNR = \frac{P_t\alpha\|\overline{h}\|^2 cos^2\theta}{\sigma^2} \qquad (3.110)$$

The signal to noise ratio SNR in (3.110) can be maximized when $\theta = 0$, which means that $\overline{w}$ has to be pointed along the direction of $\overline{h}$, also SNR can be maximized based on the optimal weight combining vector which can be written as [76].

$$\overline{w}_{opt} = \frac{\overline{h}}{\|\overline{h}\|} \qquad (3.111)$$

When the optimum weight vector $\overline{w}_{opt}$ is also adopted to maximize the SNR at the receiver

side of each user, this is called maximum ratio combining (MRC). The SNR based on MRC can also be reformulated as follows [73][76]:

$$SNR_{max} = \frac{P_t \alpha \|\overline{h}\|^2}{\sigma^2} = \frac{P_t \alpha(|h_1|^2 + |h_2|^2)}{\sigma^2} \qquad (3.112)$$

Based on aforementioned analysis, the achievable channel capacity for far user can be written as follows [73]:

$$R_f = \log_2\left(1 + \frac{P_t \alpha_f \left(|h_{1f}|^2 + |h_{2f}|^2\right)}{\left(|h_{1f}|^2 + |h_{2f}|^2\right)P_t(\alpha_n + \alpha_m) + \sigma^2}\right) \qquad (3.113)$$

Similarly, the rate for the middle user can be written as [73]

$$R_m = \log_2\left(1 + \frac{P_t \alpha_m(|h_{1m}|^2 + |h_{2m}|^2)}{(|h_{1m}|^2 + |h_{2m}|^2)P_t(\alpha_n) + \sigma^2}\right) \qquad (3.114)$$

Conventionally, near user has good channel status with the BS, hence the transmitted signal to near user is assigned less power $\alpha_n < \alpha_m < \alpha_f$. After implementing the SIC procedure, the expected near user rate to decode the required signal can be shown as:

$$R_n = \log_2\left(1 + \frac{P_t \alpha_n(|h_{1n}|^2 + |h_{2n}|^2)}{\sigma^2}\right) \qquad (3.115)$$

3.14 Optimization Problem

Based on the aforementioned analysis applied in the previous sections for SISO-NOMA and MISO-NOMA, in this section we also aim is to maximize the sum throughput and allocating the optimum power factors for each user according to the inspected constraints and the offered channel model. In the proposed scenario, we also consider three users classified as far, middle, and near users in a SIMO-NOMA system, therefore the sum of the possible rates can be written as shown:

$$R_{sum} = R_n + R_m + R_f \qquad (3.116)$$

The constraints and objective function accounted for our optimization problem in the SIMO-NOMA system can also be introduced as follows [66][73]:

- **Total power**

The power percentage given for each user device should follows:

$$\sum_{x=1}^{N} \alpha_x \leq 1 \tag{3.117}$$

where α_x is the power sharing fraction for the xth user in SIMO-NOMA network with N users.

- **Minimum transmission rate**

$$\log_2(1 + \delta_n) \geq R_{min} \tag{3.118}$$

where R_{min} is the minimum transmission rate in the SIMO-NOMA cell, and δ_n is the SINR for n^{th} user. According to the analysis shown in the previous systems, the expression in (3.118) can be rewritten as follows [66][67][73]:

$$(|h_{1k}|^2 + |h_{2k}|^2)\rho\left(\alpha_k - (2^{R_{min}} - 1)\sum_{i=1}^{k-1} \alpha_i\right) > (2^{R_{min}} - 1) \tag{3.119}$$

where ρ is the SNR and R_{min} is the minimum target rate for any user in the SIMO-NOMA cell and h_{1k} and h_{2k} are the first and second fading channels between the BS and user k.

- **Objective function**

According to the constraints in (3.117) & (3.118) and the sum rate expression and the fact that there is only 1 antenna at the base station and two receiver antennas at each user device, the objective function can be generally defined as [73]:

$$\max_{\alpha} \quad R_{sum} = \sum_{k=1}^{N} \log_2\left(1 + \frac{(|h_{1k}|^2 + |h_{2k}|^2)P_t\alpha_k}{(|h_{1k}|^2 + |h_{2k}|^2)\sum_{j=1}^{k-1} P_t\alpha_j + \sigma^2}\right) \tag{3.120}$$

3.15 Optimization Problem Analysis

Similar to SISO-NOMA, and MISO-NOMA systems earlier discussed in the previous sections, the optimization analysis in this section will be also discussed with respect to three users as follows [10][11][73]

$$\max_{\alpha} \quad R_{Sum} = R_n + R_m + R_f \tag{3.121}$$

Subject to:

$$\alpha_n + \alpha_m + \alpha_f - 1 \leq 0 \tag{3.122}$$

$$\alpha_n, \alpha_m, \alpha_f \geq 0 \tag{3.123}$$

$$(|h_{1k}|^2 + |h_{2k}|^2)\rho\left(\alpha_k - (2^{R_{min}} - 1)\sum_{i=1}^{k-1}\alpha_i\right) > (2^{R_{min}} - 1) \quad (3.124)$$

where R_{min} is the minimum rate required in the system. The sum of the achievable rates for the three users SIMO -NOMA system can be formulated as follows:

$$R_{sum} = \sum_{k=1}^{3} \log_2\left(\frac{(|h_{1k}|^2 + |h_{2k}|^2)P_t\sum_{j=1}^{k-1}\alpha_j + \sigma^2 + (|h_{1k}|^2 + |h_{2k}|^2)P_t\alpha_k}{(|h_{1k}|^2 + |h_{2k}|^2)P_t\sum_{j=1}^{k-1}\alpha_j + \sigma^2}\right) \quad (3.125)$$

The constraints can also be expressed as follows:

$$C_1(\alpha) = \alpha_n + \alpha_m + \alpha_f - 1 \quad (3.126)$$

$$C_2(\alpha) = (2^{R_{min}} - 1) - \rho(|h_{1k}|^2 + |h_{2k}|^2)\left(\alpha_f - (2^{R_{min}} - 1)(\alpha_n + \alpha_m)\right) \quad (3.127)$$

$$C_3(\alpha) = (2^{R_{min}} - 1) - \rho(|h_{1k}|^2 + |h_{2k}|^2)(\alpha_m - (2^{R_{min}} - 1)(\alpha_n)) \quad (3.128)$$

where α_f, α_m and α_n are the power factors for far, middle, and near users respectively. Since both $C_1(\alpha), C_2(\alpha)$ and $C_3(\alpha)$ are linear in terms of α , then the constraints are convex. The objective function in Equation (3.125) is considered as nonlinear optimization problem, hence $\frac{\partial R_{Sum}}{\partial \alpha_i}$ and $\frac{\partial^2 R_{Sum}}{\partial \alpha_i^2}$ need to be derived to be used to check if the objective function either convex or concave [69][70]. After some mathematical handling, a general formula can be derived for the first derivative for the objective function in terms of α_i as follows [73]:

$$\frac{\partial R_{Sum}}{\partial \alpha_i} = \frac{1}{ln2}\left(\frac{(|h_{i1}|^2 + |h_{i2}|^2)P_t}{(|h_{i1}|^2 + |h_{i2}|^2)P_t\sum_{j=1}^{i}\alpha_j + \sigma^2}\right)$$

$$-\frac{1}{ln2}\sum_{k=1}^{N-i}\left\{\left(\frac{\left((|h_{(i+k)1}|^2 + |h_{(i+k)2}|^2)P_t\right)^2\alpha_{i+k}}{\left((|h_{(i+k)1}|^2 + |h_{(i+k)2}|^2)P_t\sum_{j=1}^{i+k}\alpha_j + \sigma^2\right)}\right) \times \right.$$

$$\left.\left(\frac{1}{\left((|h_{(i+k)1}|^2 + |h_{(i+k)2}|^2)P_t\sum_{j=1}^{i+k-1}\alpha_j + \sigma^2\right)}\right)\right\} \quad (3.129)$$

In addition, the second derivative for objective function in terms of α_i can also be derived and formulated in general form as follows:

$$\frac{\partial^2 R_{Sum}}{\partial \alpha_i^2} = -\frac{1}{ln2}\left\{\left(\frac{((|h_{i1}|^2 + |h_{i2}|^2)P_t)^2}{\left((|h_{i1}|^2 + |h_{i2}|^2)P_t \sum_{j=1}^{i} \alpha_j + \sigma^2\right)^2}\right)\right.$$

$$-\sum_{k=1}^{N-i}\left\{\left(\frac{\left((|h_{(i+k)1}|^2 + |h_{(i+k)2}|^2)P_t\right)^3 \alpha_{i+k}}{\left((|h_{(i+k)1}|^2 + |h_{(i+k)2}|^2)P_t \sum_{j=1}^{i+k} \alpha_j + \sigma^2\right)^2}\right)\right.$$

$$\times\left(\left[2\left((|h_{(i+k)1}|^2 + |h_{(i+k)2}|^2)P_t \sum_{j=1}^{k+i-1} \alpha_j + \sigma^2\right) + (|h_{(i+k)1}|^2 + |h_{(i+k)2}|^2)P_t\alpha_{i+k}\right]\right)$$

$$\left.\left.\times\left(\frac{1}{\left((|h_{(i+k)1}|^2 + |h_{(i+k)2}|^2)P_t \sum_{j=1}^{i+k-1} \alpha_j + \sigma^2\right)^2}\right)\right\}\right\}$$
(3.130)

Based on the objective function and $\frac{\partial^2 R_{Sum}}{\partial \alpha_i^2}$, it can be proved that the objective function is negative and consequently it is strictly concave and has a unique global maximum. Lagrange function and the KKT optimality conditions can be applied to obtain optimum power coefficients [69][70][73]:

$$\mathcal{L}(\alpha_n, \alpha_m, \alpha_f, \mu_1, \mu_2, \mu_3) = R_{Sum} - \mu_1 C_1(\alpha) - \mu_2 C_2(\alpha) - \mu_3 C_3(\alpha) \quad (3.131)$$

where μ_1, μ_2, and μ_3 are Lagrange multipliers for the three users' scenario. Optimality condition can be formulated as follows:

$$\frac{\partial R_{Sum}}{\partial \alpha_n} - \mu_1 \frac{\partial C_1(\alpha)}{\partial \alpha_n} - \mu_2 \frac{\partial C_2(\alpha)}{\partial \alpha_n} - \mu_3 \frac{\partial C_3(\alpha)}{\partial \alpha_n} = 0 \quad (3.132)$$

$$\frac{\partial R_{Sum}}{\partial \alpha_m} - \mu_1 \frac{\partial C_1(\alpha)}{\partial \alpha_m} - \mu_2 \frac{\partial C_2(\alpha)}{\partial \alpha_m} - \mu_3 \frac{\partial C_3(\alpha)}{\partial \alpha_m} = 0 \quad (3.133)$$

$$\frac{\partial R_{Sum}}{\partial \alpha_f} - \mu_1 \frac{\partial C_1(\alpha)}{\partial \alpha_f} - \mu_2 \frac{\partial C_2(\alpha)}{\partial \alpha_f} - \mu_3 \frac{\partial C_3(\alpha)}{\partial \alpha_f} = 0 \quad (3.134)$$

Slackness conditions can be represented as follows:

$$\mu_1(\alpha_n + \alpha_m + \alpha_f - 1) = 0 \quad (3.135)$$

$$\mu_2\left((2^{R_{min}} - 1) - \rho(|h_{1k}|^2 + |h_{2k}|^2)(\alpha_f - (2^{R_{min}} - 1)(\alpha_n + \alpha_m))\right) = 0 \quad (3.136)$$

$$\mu_3\left((2^{R_{min}} - 1) - \rho(|h_{1k}|^2 + |h_{2k}|^2)(\alpha_m - (2^{R_{min}} - 1)(\alpha_n))\right) = 0 \quad (3.137)$$

Lagrange multipliers need to satisfy the following.

$$\mu_1 \geq 0, \qquad \mu_2 \geq 0, \qquad \mu_3 \geq 0 \tag{3.138}$$

In the following steps, the Lagrange multipliers should be proved to be positive. This can be achieved as follows:

$$\frac{\partial C_1(\alpha)}{\partial \alpha_n} = \frac{\partial C_1(\alpha)}{\partial \alpha_m} = \frac{\partial C_1(\alpha)}{\partial \alpha_f} = 1 \tag{3.139}$$

$$\frac{\partial C_2(\alpha)}{\partial \alpha_n} = \frac{\partial C_2(\alpha)}{\partial \alpha_m} = \frac{\partial C_3(\alpha)}{\partial \alpha_n} = \rho(|h_{1k}|^2 + |h_{2k}|^2)(2^{R_{min}} - 1) \tag{3.140}$$

$$\frac{\partial C_2(\alpha)}{\partial \alpha_f} = \frac{\partial C_3(\alpha)}{\partial \alpha_m} = -\rho(|h_{1k}|^2 + |h_{2k}|^2) \tag{3.141}$$

After some mathematical manipulations and based on the fact that

$$|h_n|^2 > |h_m|^2 > |h_f|^2 \tag{3.142}$$

$$\left(\frac{\partial R_{Sum}}{\partial \alpha_m} - \frac{\partial R_{Sum}}{\partial \alpha_f} \right) > 0 \tag{3.143}$$

$$\left(\frac{\partial R_{Sum}}{\partial \alpha_n} - \frac{\partial R_{Sum}}{\partial \alpha_f} \right) > 0 \tag{3.144}$$

It can be proven that μ_1, μ_2 and μ_3 are positive [70][73], and according to the aforementioned analysis, the examined constraints are feasible, and the closed form expression for the power coefficient for each ser device in SIMO-NOMA can be deduced from the slackness conditions. After performing some mathematical manipulations and substitutions, the closed form expression for the optimum power allocation for each user device in SIMO-NOMA cell can be expressed as follow [70][73][76]:

$$\alpha_n = \frac{1}{(2^{R_{min}})^2} + \frac{1}{\rho(|h_{1n}|^2 + |h_{2n}|^2)} \left(\frac{1}{(2^{R_{min}})^2} - 1 \right) \tag{3.145}$$

$$\alpha_m = \left(\frac{(2^{R_{min}} - 1)}{(2^{R_{min}})^2} \right) \left(1 + \frac{1}{\rho(|h_{1m}|^2 + |h_{2m}|^2)} \right) \tag{3.146}$$

$$\alpha_f = \left(\frac{(2^{R_{min}} - 1)}{2^{R_{min}}} \right) \left(1 + \frac{1}{\rho\left(|h_{1f}|^2 + |h_{2f}|^2\right)} \right) \tag{3.147}$$

63

3.16 Simulation Results and Discussions

In this section, system simulation is conducted to evaluate the efficiency of incorporating the power optimized scheme that derived earlier in section 3.15 for the examined SIMO-NOMA system and compare the results for that optimized scheme with the fixed power scheme scenario. Simulation files are considering a single base station with 1 antenna serving three users and each user is equipped with two antennas and the simulation is accomplished over 10^6 random channel generations. In fixed power scenario, the power factors are allocated as follows: $\alpha_n = 0.1$, $\alpha_m = 0.2$ while $\alpha_f = 0.7$. The fading channels between each user and the BS is generated based on flat Rayleigh fading distribution, and the channel parameters are changing randomly at each transmission time slot. The transmitted signals are modulated using Quadrature phase shift keying (QPSK) and the path loss is assigned to 4. Bandwidth of the system is $B = 1$ MHz, the noise spectral density is $N_0 = -174$ dBm/Hz and R_{min} is specified to be 1 bps and SIMO-TDMA is used as a traditional OMA scheme for the sake of comparison.

In Figure 3.5, simulation results for SIMO-NOMA based fixed power allocation and power optimized schemes are compared in terms of the possible sum rate versus transmitted power. Simulation results for SIMO-OMA system is also created in the figure for the purpose of comparison. It can be noticed that the sum rate for the SIMO-NOMA based optimized power scheme is showing a dominance over the SIMO-NOMA scenario based on fixed power allocation. In addition, it is clearly observed that existence of multiple antennas at receiver side in NOMA system has meaningful impact in boosting the sum rate compared to OMA case with more than 3 b/s/Hz.
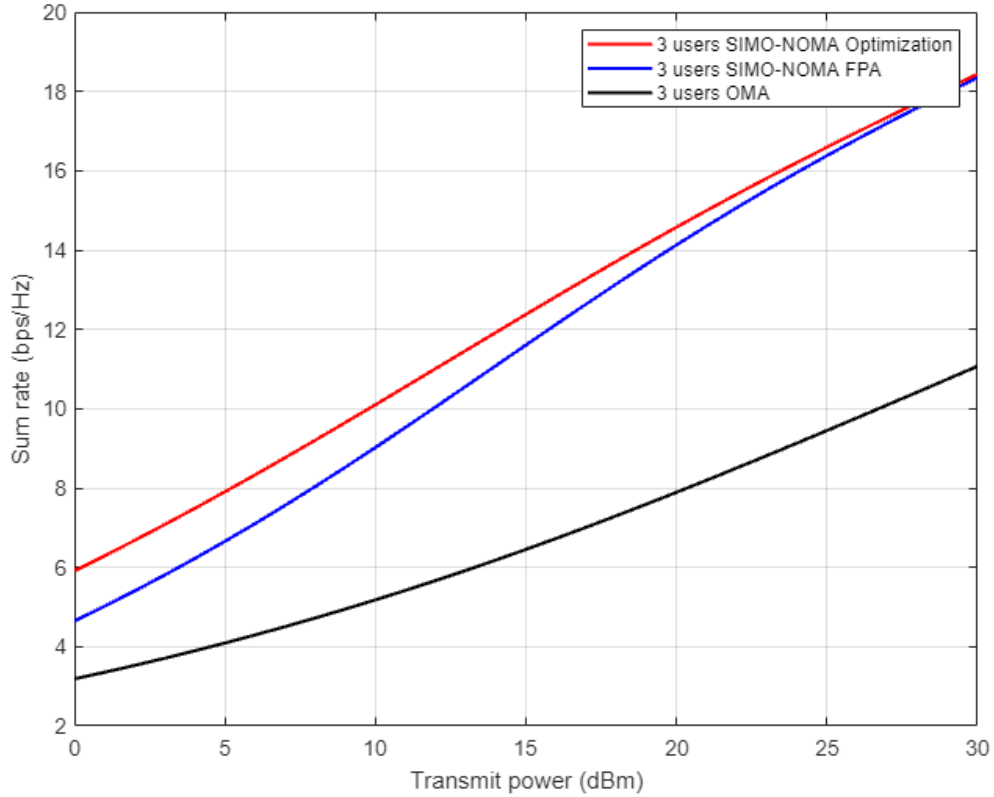
Figure 3.5 Sum rate vs Power for SIMO-NOMA and SIMO-OMA.

Figure 3.6 shows the outage probability versus transmitted power for far, middle, and near users in SIMO-NOMA system when both the optimized power scheme and fixed power scheme are considered. Simulation results for far user and middle user based optimized power scheme are clearly indicating an enhancement in power saving with approximately 2-3 dBm compared to the fixed power setting. On the other hand, near user with fixed power scheme shows some sort of improvement in outage compared to the power optimization case. This can be justified that both the good quality channel with fixed power factors can be more effective for near user than the optimized power case.
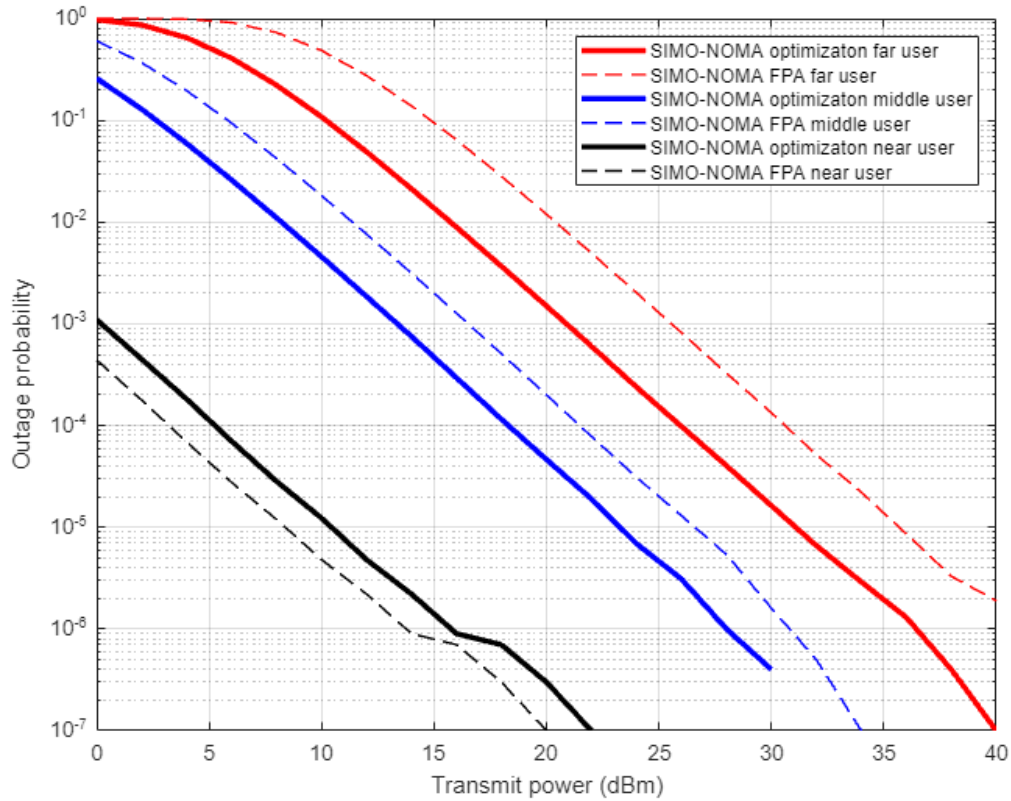
Figure 3.6 Outage probability vs Power for SIMO-NOMA (optimized - FPA).

In Figure 3.7, the achieved capacity is simulated versus transmitted power for every user in the examined SIMO-NOMA system, when both the optimized and fixed power scenarios are implemented. The simulation outcomes for the near user show dominance in the achievable rate compared to other users, and results based on the optimized power scheme show improvement by 1 b/s approximately over the fixed power scheme. In middle user case, a substantial reduction in the rate is noticed due to interference, but still the optimized power scheme can show some sort of enhancement than the fixed power scenario. On the other hand, the simulation results indicate that the optimized power scheme is not sufficient for far user, and it is noticed that fixed power scheme can provide a little enhancement over the optimized power scenario, this can reveal that the optimized power scheme may not be an appropriate selection for far user in SIMO-NOMA system.
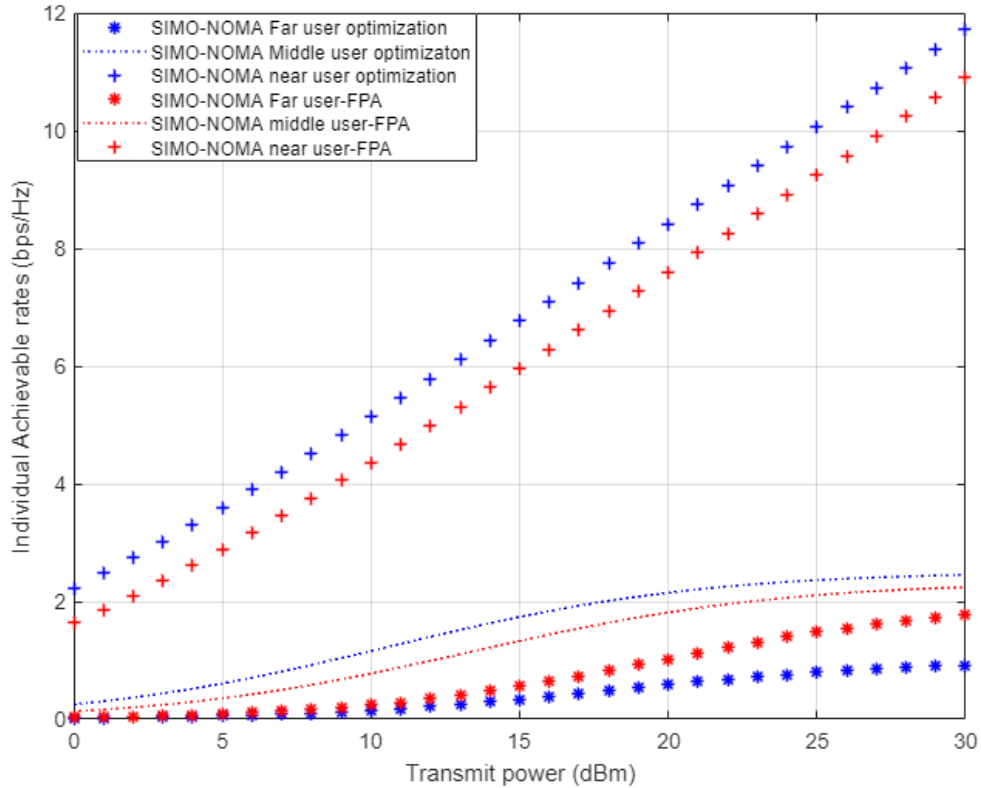
Figure 3.7 Individual rates vs Power for SIMO-NOMA (optimized - FPA).

3.17 Summary

A downlink NOMA system is discussed and evaluated on the basis of sum rate maximization problem subject to the constraints of total power budget and minimum transmission rate. In this chapter, a structured mathematical analysis is presented to derive a closed form expression for the optimum power factor for each user device in the examined SISO-NOMA, MISO-NOMA, and SIMO-NOMA systems. The achievable rate for the three users in the inspected NOMA systems is formulated, then the objective function is verified to be concave. Lagrange function and KKT optimality conditions are applied to find the optimum power factors. Several performance metrics such as BER, sum rate, outage probability and achievable capacity are utilized to inspect the enhancement accomplished in the system performance based on the optimized power scheme. Overall, simulation outcomes showed that NOMA system with optimized power policy can deliver higher sum rate and capacity compared to fixed power scheme. In terms of outage probability, simulation results for near user based fixed power allocation showed better results compared to optimized scenario, which may clarify that the channel condition is more valuable for near user than the assigned power.

Chapter 4

Deep Neural Networks based Long Short-Term Memory Architecture and Framework

4.1 Introduction

In a non-orthogonal multiple access (NOMA) system, the successive interference cancellation (SIC) procedure is classically employed at the receiver side, where several user's signals can be decoded in a successive manner. Fading channels may scatter the transmitted signal and this may initiate dependencies among its samples, which may affect the channel estimation process and consequently disturb the SIC process which correspondingly may degrade the signal detection accuracy [77]. In this chapter, the influence of deep neural network (DNN) in explicitly estimating the channel parameters for each user equipment in the considered NOMA cell is investigated in different fading channels such as Rayleigh and Rician channels. The proposed approach mainly integrates the Long Short-Term Memory (LSTM) network into the NOMA system and the LSTM network will be accountable to estimate the channel coefficients before the signal detection process. Initially, the DNN is trained using different channel statistics and the DNN model is trained to update the weights, afterward the trained network will be utilized to predict the desired channel parameters, and these channel coefficients will be exploited by the receiver to retrieve the original data [78].

In addition, in this chapter we will investigate how the channel estimation procedure based on LSTM network and the optimized power scheme that derived in chapter 3 can both jointly utilized for multiuser (MU) detection in the considered downlink Power domain NOMA system. Simulation results for different performance metrics, such as bit error rate (BER), sum

rate, outage probability and individual user capacity, have proved the dominance of the proposed DNN based LSTM network to estimate the channel parameters for each user device in the examined NOMA system. Furthermore, the performance of the optimized power scheme and the fixed power scheme are evaluated and compared when DNN based LSTM is implemented as a channel estimator.

4.2 Related Works

In [79], authors have proposed a recognition system based on deep learning (DL) scheme in downlink OFDM-NOMA system. Authors mostly count on the pilot signals for predicting the channel information and according to the pilot's feedback, symbol detection can be realized using DL scheme without the need for a separate channel estimation phase. Simulation's results revealed that the suggested DL scheme outperforms the standard SIC procedure. On the other hand, the proposed model needs to be initially trained offline for different channel conditions and the simulation outcomes were presented in terms of BER.

In [80], authors suggested a deep learning framework to retrieve the desired signal for each user in MIMO-NOMA system when Rayleigh fading channel is applied. The presented deep learning method can accomplish the channel prediction and signal detection concurrently. In terms of symbol error rate (SER) and throughput, simulations were conducted, and results were compared with the conventional SIC procedure. According to the simulation outcomes, the recommended deep learning scheme can address channel impairments, but the examined NOMA cell was restricted for two users. Also, an offline training stage is mandatory where the DNN depend on the received signal with the labels, and these labels are considered as supervised data to help the neural network to optimize the network parameters.

Authors in [81], introduced a data driven deep learning channel estimator for frequency selective channels. The proposed framework is developed such that a pre-training scheme and pilot symbols are employed as inputs for the DNN to realize the desired initialization for the network to further augment the performance of the estimator. The presented DNN is trained offline in both the pre-training (initialize the network) and training stages (update weights). In the testing stage, the features or changes in the channel can be dynamically tracked by the DNN when only pilot symbols are used, and then the transmitted information can be recovered. Different number of layers has been investigated in that DNN estimator and numerical results

demonstrate that the proposed estimator can outperform the traditional channel estimator scheme in terms of efficiency and robustness (dependability).

Deep learning approach is also introduced in [82], to approximate the downlink channel parameters and to reduce the training overhead in a fog radio access network. Gated Recurrent Unit (GRU) is employed to learn the correlations between fading channels related to various users in the system, and to further enhance the estimation process. Simulation outcomes are provided to show the performance gains, but the examined performance metrics were limited to mean square error (MSE) and loss function.

On the basis of deep Learning algorithm, authors in [83] introduced a sliding window estimator that implement Gated Recurrent Unit (GRU) to learn the characteristics associated with a time varying Rayleigh fading channel. Channel coding and Interleaving schemes are also adopted with the proposed channel estimator to boost the system performance. Simulation results showed the ability of the indicated technique to follow or track the changes in the channel in a reliable way and achieve better mean square error (MSE). Besides, the presented GRU estimator is examined with various numbers of pilot symbols, and the robustness of that estimator against the variations in the channel parameters is inspected.

In [84], authors went to conclude that deep learning algorithm can also be exploited in signal detection in uplink NOMA transmission. Authors proposed a DNN approach to differentiate the complex channel parameters, and restricted Boltzmann machines (RBM) is implemented as a pre-training phase for the input sequence transmitted in the network. In the suggested deep learning scheme, offline learning is applied, and the proposed LSTM layer is designed to track the environment statistics automatically and then an iterative detection procedure is implemented to identify the transmitted symbols. Performance analysis for the suggested DNN scheme is evaluated simply in terms of sum data rate and block error rate.

Based on DNN, authors proposed a semi-blind detection method in [85] to distinguish users' symbols in co-operative NOMA system. The presented DNN model has proven the capability to detect the user signal based on the assistance from pilot responses without the need for additional channel estimation process. The DNN model is trained offline over Rayleigh fading channel and then the trained network is adopted in the online detection phase. Additionally, the trained DNN model is further examined using Nakagami and Rician fading channels and

simulation outcomes have demonstrated the ability of the suggested DNN scheme in outperforming the classical detectors.

In an OFDM system, authors in [86] have introduced a DNN structure for joint channel estimation and signal detection. This joint structure considers OFDM system and the fading channel as black box and the presented DNN network is trained offline using simulated data. Simulation results declared that the joint approach has the potential to learn and investigate the complicated attributes of the wireless channels. Besides, the joint approach has proven its dominance over conventional methods (MMSE) when limited pilots' symbols are employed.

4.3 Recurrent Neural Networks and long short-term memory Networks

Recurrent Neural Networks (RNNs) are observed as a class of supervised learning procedures, where they can exploit consecutive sequences for prediction and detection [87]. As shown in Figure 4.1, RNN involves hidden layer that composed of artificial neurons with feedback loop. Therefore, RNNs have dual inputs, i.e., the current input and the recent previous response. Typically, in RNNs, hidden layers are qualified to play a role as the storage for the network at any given time, hence this structure enables RNN to handle the preceding data for an extended period of time. Typically, RNN can be used to solve problems with sequential input data such as time series. On the other hand, traditional RNN encounters slow learning and a vanishing gradient problem. Accordingly, RNNs may not be the best candidate for signals that may be sent through fading channels, where these fading channels may scatter the signal and initiate long-term dependencies among its samples [10][87]. Long short-term memory (LSTM) network, which is a one category of RNNs, is commonly used with time series data and sequences for classification, where LSTM layer can take a benefit or manage the long-term dependency for the time series data [10][57][58].

Based on LSTM underlying design, the LSTM network includes LSTM cells, and each cell contains a set of gates that are capable of storing and gaining access to data over extended periods of time. In addition, LSTM cell can manage a vector of complex data, hence encompassing the magnitude and phase parts of the received sequence instantaneously. Therefore, LSTM layer can be considered as an appropriate selection to realize channel estimation process when time series data or simulated data for different channel environments are available [88].

(a) Single neuron element. (b) Recurrent neural network (RNN).
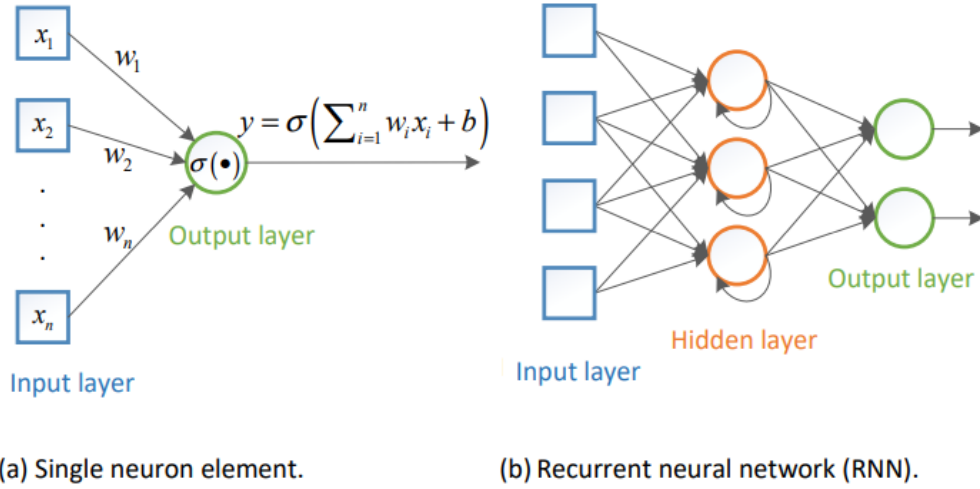
Figure 4.1 RNN network architecture.

4.3.1 Proposed Deep Neural Networks Architecture and Framework

A framework that combines the LSTM network with NOMA system is introduced in this section, where the LSTM model is mainly trained to be employed to estimate the channel coefficients for each user device in the cell. Data driven communication scheme is generally depending on empirical observations to determine the amount of LSTM cells in each layer and the numbers of LSTM layers that are essential for the estimation process. Furthermore, it is important to take in consideration that adding more LSTM layers may not offer an evident gain in learning phase, or it may not positively affect the network convergence [10][12][88].

Figure 4.2 illustrates the architecture of the proposed DNN network for channel estimation that consists of four layers [10][12], and each layer contains several neurons, and the weighted sum of each neuron will be the input to an activation function. The length of each training sequence is specified as L, which is the size of the input layer. We decide that the size of the input layer is $L = 128$, and the construction for the 128 inputs for the input layer can be clarified as follows:

- Initially generate random channel coefficients with vector size 120 based on the
    a. channel model parameters
    b. User distance
    c. Path loss
- Estimate the channel coefficients based on the pilot symbols with vector size 8.

- Combine the random generated channel coefficients and the estimated channel coefficients to form a vector of channel parameters with size 128.
- The constructed channel parameters frame will be used as input for the first layer in the proposed DNN Architecture.

Accordingly, the first input layer involves 128 neurons, and the inputs to the input layer will be shifted to the subsequent layer (LSTM layer). In the second layer, one LSTM layer is implemented which includes 300 hidden cells. The learnable weights of LSTM layer are the input weights W, the recurrent weights R, and the bias b. LSTM cell structure and mechanism will be discussed in the next section. The third layer is a fully connected layer that is accountable for processing the outputs of the LSTM layer. The fully connected layer will multiply the input by a weight matrix W and then adds a bias vector b. All neurons in a fully connected layer are connected to all the neurons in the preceding layer, which indicate that the fully connected layer can bring together all of the features and internal information gathered by the prior layers.
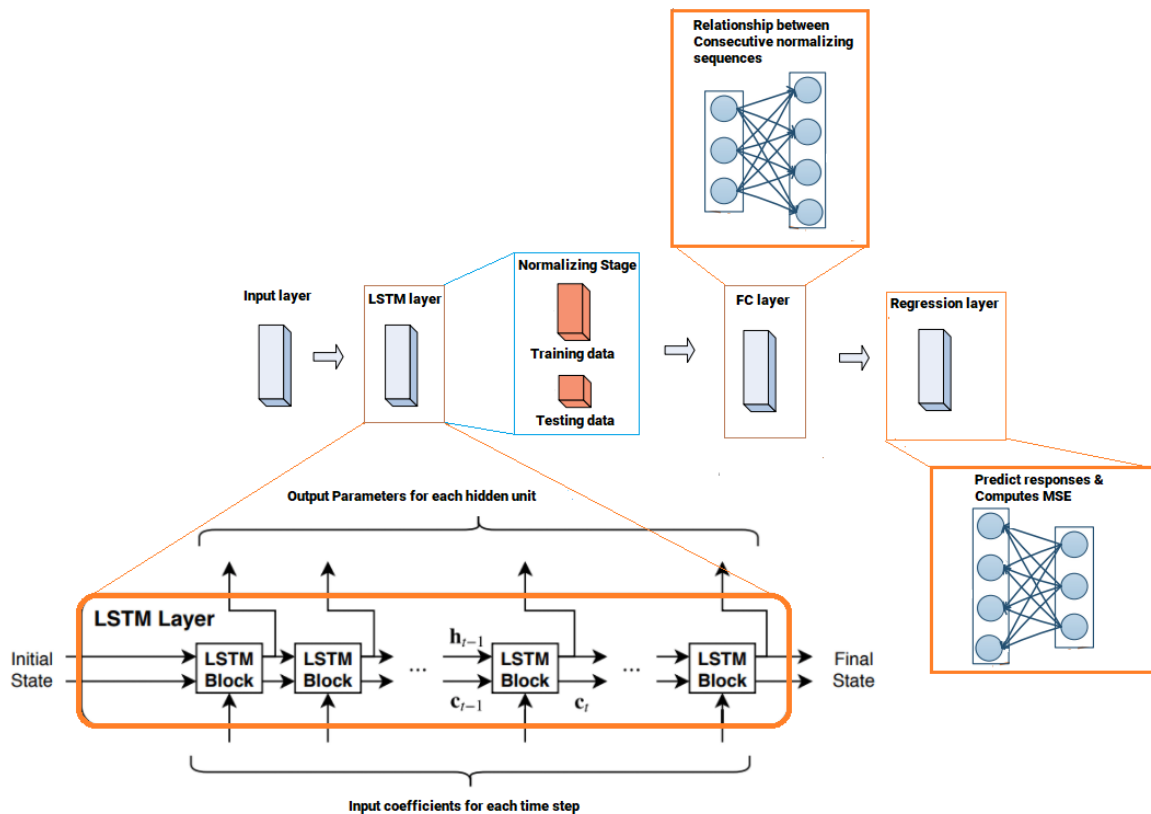
Figure 4.2 Architecture of the proposed DNN network**.**

73

The last adopted layer in the proposed DNN framework is the regression layer, which is liable of updating the network weights, cell status, and biases. A regression layer can also compute the mean square error (MSE) and predicts the responses of a trained network [89]. Normalizing the training data can facilitate stabilizing and accelerating the training process of the neural networks. For a single observation, the mean square error can be calculated as [10][88][89].

$$MSE = \sum_{j=1}^{r} \frac{(y_{Tj} - y_{Pj})^2}{r} \qquad (4.1)$$

where r is the number of responses, y_{Tj} is the target output, and y_{Pj} is the predicted output at response j.

4.3.2 LSTM Cell Structure and Mechanism

In LSTM cell, the output is created based on the cell state, and the previous hidden state, and the current input. In order to remember the preceding cell state and decide if the prior state will be used or not, the LSTM cell consists of different types of gates, these gates are the forget gate, the input gate, and the output gate. In LSTM, there are two states, the hidden state h_{t-1}, which is used for computing the output and the cell state C_{t-1}, which is called internal memory where all information is accumulated [10][88][89].

Figure 4.3 illustrates the internal configuration of LSTM cell [10][12][90], where x_t is the current input, and h_{ti} represents next hidden state and also it can represent the output channel coefficients for user i at time t. In addition, C_{t-1} represents the previous cell state, and at same time it will be one of the inputs to the current LSTM cell at time instant t. At every time step, the LSTM cell can add up information or remove information from the cell state and LSTM cell can control these updates using numerous gates that can be simply described as follows:

1. The forget gate is responsible for controlling the level of cell state that need to be reset:
    $f_t = \sigma(W_f x_t + R_f h_{t-1} + b_f)$.

2. The input gate is responsible for controlling the level of cell state that need to be updated:
    $i_t = \sigma(W_i x_t + R_i h_{t-1} + b_i)$.

3. The candidate state is responsible for adding information to the cell state:
    $g_t = tanh(W_g x_t + R_g h_{t-1} + b_g)$.

4. Updated cell state: $C_t = (C_{t-1} \odot f_t) + (i_t \odot g_t)$, where $\odot$ is element-wise multiplication.

5. The output gate is responsible for controlling the level of cell state added to hidden state: $O_t = \sigma(W_o x_t + R_o h_{t-1} + b_o)$.

6. Estimated output coefficients: $h_t = O_t \odot tanh(C_t)$.
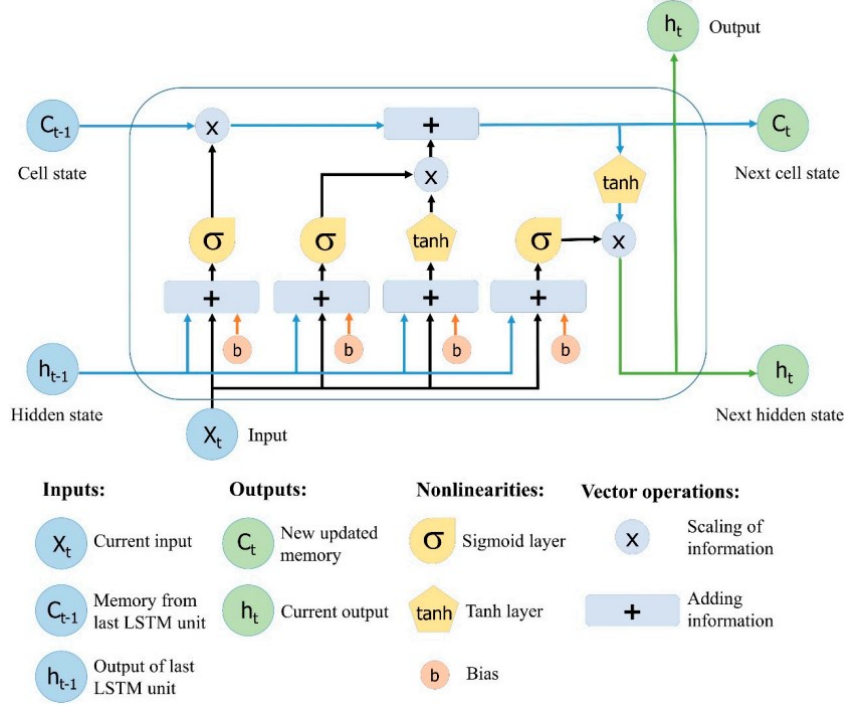
Figure 4.3 Internal structure of LSTM cell [10].

4.3.3 Channel Estimation Based DNN Model

The transmitted frame involves data and pilot symbols. The applied channel model is assumed to be static during transmission of one frame of pilot and data symbols and the channel coefficients are changing from one frame to another. In the proposed DNN scheme, two major phases are implemented to achieve an effective DNN model for channel parameters estimation. The first phase involves two stages, the training stage, and the testing stage and both stages include the usage of the combination between the random generation of the channel coefficients and estimated channel parameters based on the pilot symbols. During the model iterations, we use 70 % of the iterations for training, and 30 % of the iterations for testing [91]. In the second phase, the trained DNN model will be utilized to explicitly predict the practical channel parameters for each user based on the assistance from pilot symbols only. and these

estimated channel parameters will be employed to recover the desired transmitted data symbols for each user.

4.3.4 Dataset Generation

At the beginning of every training stage, the weights and bias values of LSTM layer are initialized, and during the training phase, weights and biases are updated according to a gradient descent procedure [10][46][92]. The distance of each user from the BS and the path loss needs to be assigned in the dataset, so that the channel parameters for each user can be randomly initialized to model the Rayleigh fading channel between the user and BS. Pilot symbols are also generated at random and recognized at the BS and at the receiver side of each user in order to assist in the channel estimation process.

Based on the initial channel parameters generated and the pilot symbols, the size of the training and testing frames can be identified. The training model are carefully designed based on the selected layers, the quantity of the hidden cells assigned for LSTM layer, and the training parameters [93]. In order to further accelerate and stabilize the training process for the training network, we choose to normalize the training data.

Throughout the training phase, the performance of the proposed DNN scheme based on LSTM model is assessed using RMSE function. In the testing period, a new fading coefficient will be randomly generated, such that these coefficients are not the same as those generated for training. Once the training and testing stages are inspected by the training network, the trained model will be employed as online channel estimator with the help of pilot symbols only. The proposed channel estimation algorithm based on LSTM can be outlined as shown in Algorithm 4.1.

Algorithm 4.1 Proposed Channel Estimation scheme based on LSTM.

---

1. Initialize the learnable weights of an LSTM layer (W, R, b), W is the input weights, R is the recurrent weights, and b is the bias.

Inputs
- Number of Iterations.
- The distance of each user from the BS.
- Path loss exponent.
- Generate Initial random Rayleigh channel coefficients for each user based on channel model.
- Specify the pilot symbols.
- Assign the initial power factors for each user.

- Identify the size of training frame (L_T) and size of testing frame (L_S).

Procedure

2. Assign the training sequence (Z_T), testing sequence (Z_S) and the desired coefficients (Z_D).
3. Calculate the **mean** and **variance** (μ_T, σ_T^2) for channel coefficients in training sequence at each iteration.
4. Normalizing the training data $Z_T \rightarrow Z_{NT}$ based on (μ_T, σ_T^2).
5. Characterize the relationship between Consecutive normalized training sequences as $Z_{NT} \approx (X_{NT}, Y_{NT})$.
6. Initialize the **training network** (T_{net}), and assign the following:
- Number of layers.
- Number of hidden units.
- Training parameters.
7. Use $Z_{NT} \approx (X_{NT}, Y_{NT})$ as inputs for the training model (T_{net})
8. **Predict** the output normalized coefficients (Y_{NP}) .
9. **Update** the state of training model (T_{net}).

   For i =1 : L_T

   $[T_{net}, Y_{NP}]$ = **Predict and Update state** (T_{net}, Z_{NT})

   End
10. Denormalize $Y_{NP} \rightarrow Y_P$
11. Calculate **RMSE ($Z_D - Y_P$)** & **Loss** function
12. **Update** the network and the state of (T_{net}) and **reset** the values for Y_P
13. Normalize **testing** data $Z_S \rightarrow Z_{NS}$, using (μ_T, σ_T^2),
14. Use Normalized testing data (Z_{NS}) as inputs for updated trained network (T_{net})

    For i =1:L_S

    $[T_{net}, Y_{NP}]$ = **Predict and Update state** (T_{net}, Z_{NS})

    End

Outputs

15. Predicted normalized channel coefficients Y_{NP}
16. Denormalize $Y_{NP} \rightarrow Y_P$.
17. Calculate **RMSE ($Z_D - Y_P$)** & **Loss** function. (update weights for trained model)
18. Estimate channel parameters based on the trained DNN model and assisted pilot symbols.

Detailed LSTM Procedure and Workflow:

Based on algorithm 4.1, the detailed LSTM procedure and workflow for estimating the channel parameters for each user in NOMA cell can be listed as follows:

1. In the first phase, we initially generate random channel parameters with vector size 120 based on the applied channel model parameters such as the user distance, and path loss.
2. Estimate the channel coefficients based on the pilot symbols with vector size 8.

3. Combine the random generated channel coefficients in step 1, and the estimated channel coefficients in step 2, to compose a vector of channel parameters with size 128.

4. Based on (3), The LSTM will generate one vector of channel coefficients for each frame of data symbols

5. In each iteration (repeat steps 1-4) for 10^6

6. 70% of the iterations is used for training, and 30% of the iterations is used for testing.

7. In the second phase, we start estimating the channel coefficients based on the learned LSTM model + assistance from pilot symbols).

8. Activation functions used in hidden layers are sigmoid and tanh, while activation functions at output layer can be a linear, sigmoid, or Relu.

9. We generate 10 frames of data bits.

4.4 Simulation Environment

In this section, a clarification for the simulation settings and parameters is presented. Our examined downlink NOMA cell contains one base station (BS) and three different user devices in which the BS and each user in the cell is equipped with one antenna. For the downlink NOMA scenario, the modulated signals for each user are superimposed and transmitted by BS to all users via uncorrelated Rayleigh or Rician fading channels. At the receiver side, additive white gaussian noise (AWGN) is considered and the noise spectral density is $N_0 = -174$ dBm/Hz and path loss is 4.

In this chapter, simulations mainly are conducted using MATLAB to simulate and to highlight the following: first, to evaluate the efficiency of adopting the proposed DL based LSTM network to accurately estimate the channel parameters for each device in downlink NOMA cell. Second, to combine the proposed LSTM algorithm as a channel estimator with the derived optimized power allocation scheme and compare this integrated model with NOMA system when fixed power factors are considered along with the proposed LSTM network. Monte-Carlo simulations are conducted with 10^6 iterations, and at the start of every set of iterations, pilot symbols are generated at random and identified at the BS and at the receiver side of each device. The main simulation parameters are summarized in Table 4.1. In our simulation environment, we assume that the channel state information (CSI) is not available at the receiver

side. Therefore, for the sake of comparison and in order to verify the efficiency of the proposed DL algorithm, two methods for channel parameters estimation are implemented at the receiver side for each device. The first method, which is our proposed approach that employs DNN-based LSTM layer to estimate the desired channel parameters, and the gradient descent algorithm is applied in conjunction with LSTM layer to update the network weights. The second implemented channel estimation scheme at each receiver side is initiated based on the minimum mean square error (MMSE) procedure [62][94]. The MMSE procedure will be applied as a conventional channel estimation technique for each user in the NOMA cell, and in the simulations results, we refer to the MMSE procedure as conventional NOMA, to clarify that users are using the MMSE procedure for estimating the channel parameters before recovering the desired signal.

Channel taps that are employed to model the Rayleigh fading wireless channel are generated on the basis of ITU channel models. Throughout the simulations, NOMA system parameters are created [95]. Training and testing phases are conducted online throughout the simulations, and the fading parameters in the testing stage are generated such that these parameters are not the same as in the training stage. At the end of the training stage, which involves adopting training and testing data, the trained network will be prepared to be utilized as online channel estimator.

Initially, different power factors are allocated for each user device according to their distance from the BS and the current channel status. Power allocation coefficients α_n, α_m, and α_f are defined for near, middle, and far users, respectively. In the fixed power allocation (FPA) scenario, we can assign $\alpha_f = 0.7$, $\alpha_m = 0.2$, and $\alpha_n = 0.1$. Alternatively, in the optimized power scheme, power factors are assigned for users according to the analytical forms derived earlier in chapter 3 for each user . The propagation distance for each user with respect to BS is initially assigned in the simulation files as follows: $d_f = 1000$ m, $d_m = 500$ m, and $d_n = 200$ m. Quadrature phase shift keying QPSK is employed as a modulation scheme for the data symbols and pilot sequences. The applied transmitted power mainly varies from 0 to 30 dBm.

Table 4.1 Summary of the simulation parameters.

Parameter	Value
Simulation Tool	MATLAB
Modulation type	QPSK
Number of Users	3, [2–10]
System Bandwidth B	1 MHz
Fading channel	(Rayleigh, Rician)
Path loss	4
Number of Iterations	10^6
Optimizer	ADAM
Learning Rate α	0.001
Number of LSTM cells	300
Pilot symbols	8
Frame size	128

4.5 Simulation Results and Discussion

In Figure 4.4, the simulation results show the comparison between the proposed DL based LSTM scheme for channel estimation and the conventional NOMA scheme that applies MMSE technique for estimating the channel coefficients. The estimated channel parameters using both procedures will be used in recovering the desired signals for far, middle, and near users and the simulation results are shown in terms of bit error rate (BER) versus transmitted power. All users in the NOMA cell-based LSTM for channel estimation show noticeable improvement in lowering the bit errors compared to the standard NOMA scenario, especially when the assigned power is keep increasing. It can be noticed that for certain BER values, such as 10^{-2}, the power saving achieved by DL based LSTM network is approximately between 5–8 dBm for far and middle users, while for the near user, the power saving is up to 5 dBm.
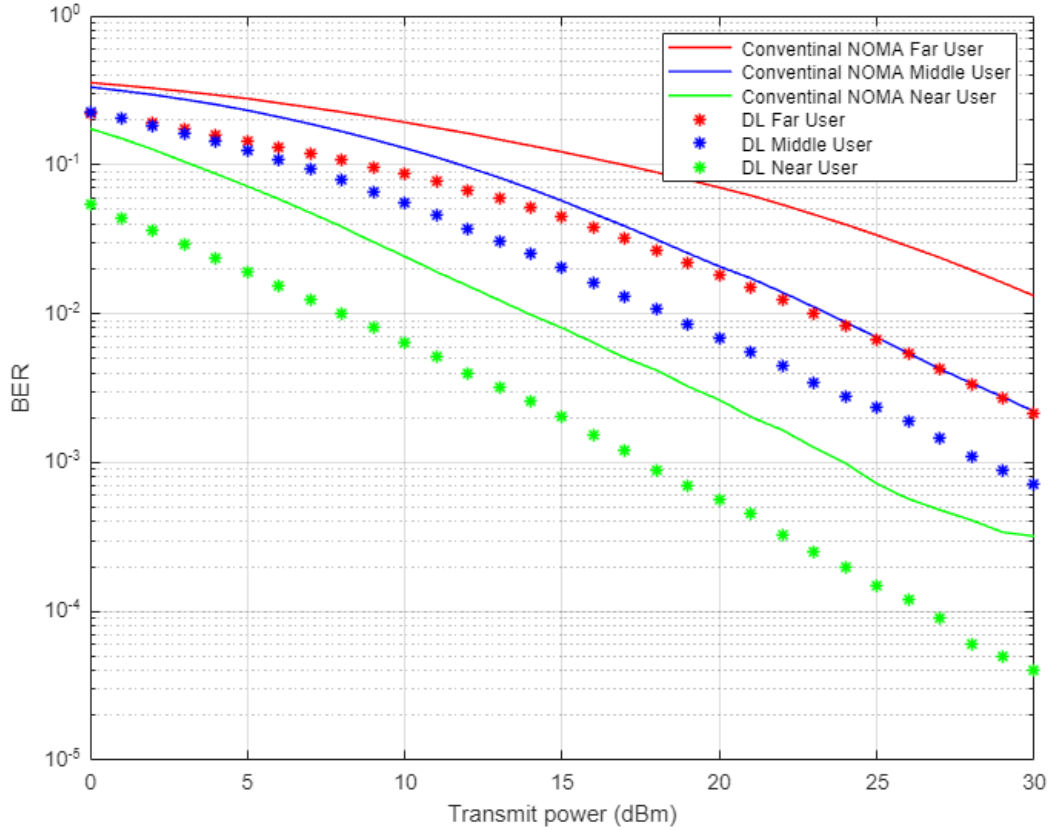
Figure 4.4 BER vs. power-for proposed DL-NOMA and conventional NOMA (Rayleigh).

Figure 4.5 shows the results for outage probability metric versus transmitted power for the three examined users in NOMA system when DL based LSTM and conventional NOMA procedures are adopted for channel estimation. Far and near users' simulation results indicate an approximately 5 dBm improvement in power saving to achieve a certain outage probability (10^{-2}) when DL based LSTM method is implemented compared to the conventional estimation scheme. Also, the middle user with the DL estimation scheme shows more enhancement in power saving compared to the MMSE procedure, by 5–7 dBm approximately, which verifies the superiority of the proposed DL based LSTM estimation scheme.
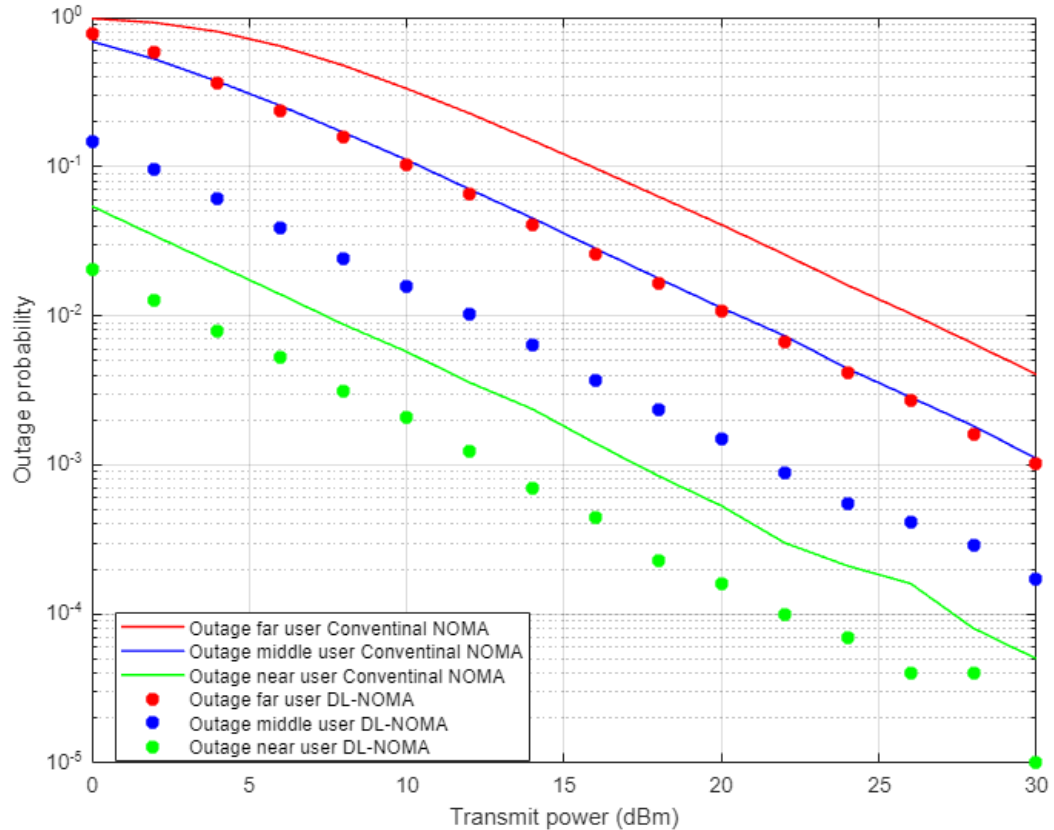
Figure 4.5 Outage probability vs. power for DL and conventional NOMA (Rayleigh).

Figure 4.6 shows the simulation outcomes for the sum rate versus the transmitted power for the three examined users in the NOMA cell. In this Figure, and as a benchmark comparison, three different channel estimation schemes are inspected, i.e., the proposed DL based LSTM approach, conventional NOMA based on the MMSE scheme, and the DL algorithm for joint channel estimation and signal detection that was applied in [86]. Based on the simulation outcomes, it can be evidently noticed that the developed LSTM channel estimation scheme shows dominance over the conventional NOMA scenario, nearly with 6 b/s/Hz. The results also indicate an improvement over the DL algorithm implemented in [86] by 2 b/s/Hz. These results support the efficiency of the developed DL based LSTM scheme in estimating the channel parameters before being employed in the decoding process.
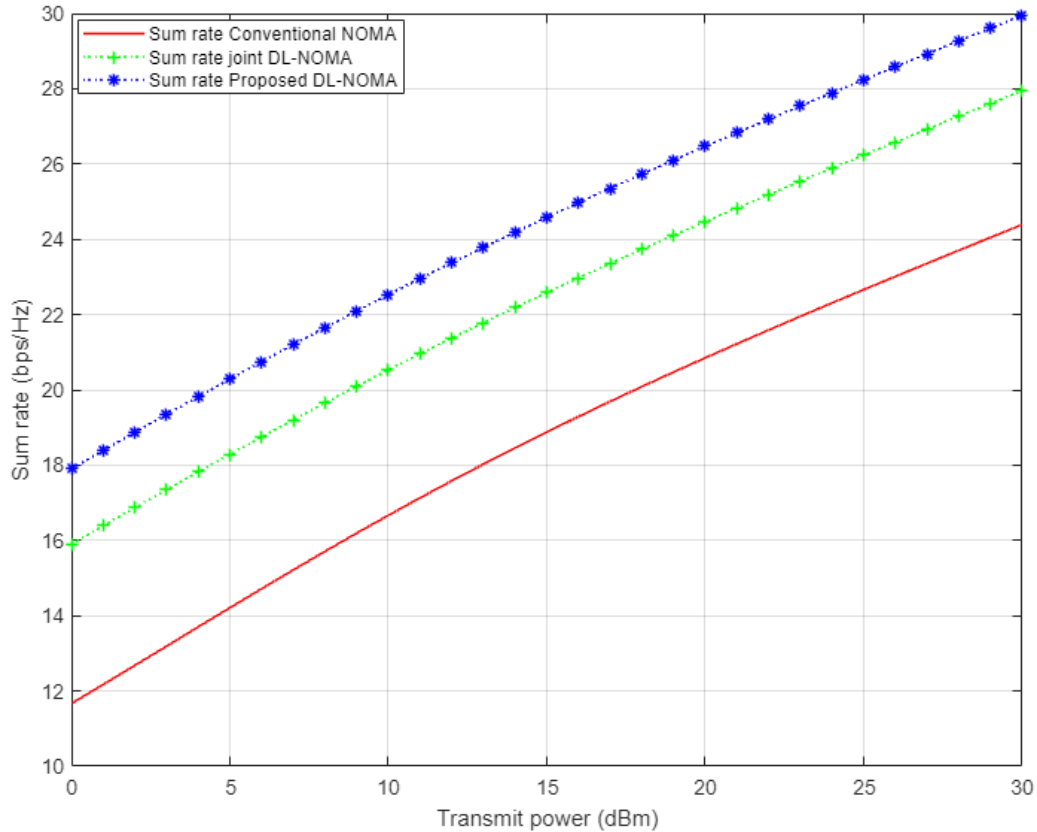
Figure 4.6 Sum rate vs. power for conventional NOMA, joint DL-NOMA, and proposed DL-NOMA (Rayleigh).

Figure 4.7 shows the simulation outcomes for the individual capacity metric for each device in the NOMA cell when the developed DL based LSTM and the traditional NOMA based MMSE are both utilized for channel estimation process. As expected, when power level starts to increase, the achieved capacity for the near user provides a substantial difference by at least 5 b/s/Hz over far and middle users' rates. This may be justified by the good channel state for the near user compared to other users in the cell. Furthermore, for far and middle users the developed LSTM approach still can deliver a visible improvement compared to MMSE procedure, but with little impact especially for the far user, due to interference and weak channel environment.
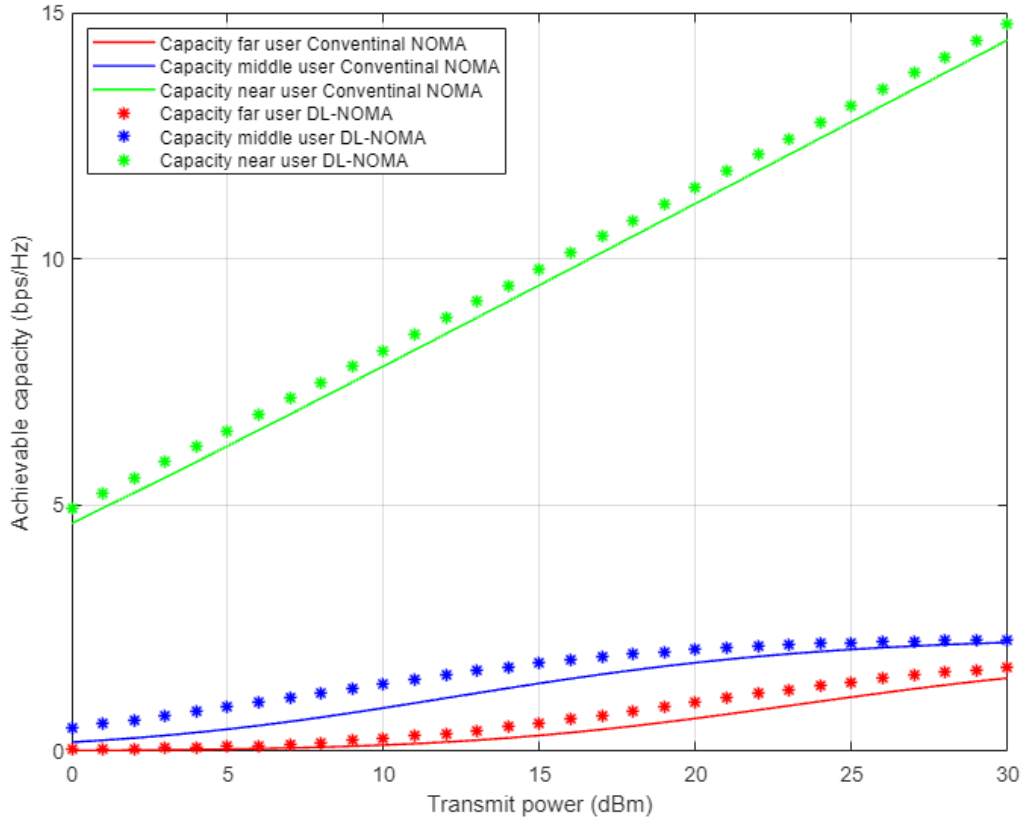
Figure 4.7 Individual capacity vs. power for conventional NOMA and DL-NOMA (Rayleigh).

Figure 4.8 shows the simulation results for BER versus power when the Rician channel is applied. The developed LSTM channel estimation scheme and traditional MMSE procedure will be further inspected using the Rician distribution. Rician fading is a stochastic model for radio propagation, where the signal arrives at the receiver by different paths, and hence, exhibits multipath fading. Typically, Rician fading happens when one of the paths is a line of sight (LOS) signal or some strong reflection signals, are much stronger than the other received components [96][97]. A Rician fading channel can be described by two parameters The first one is the Rician factor K defined as the ratio of the signal power in the line-of-sight component to the scattered power in other components. The other primary parameter is Ω, which signifies the total power from both paths and can be considered as a scaling factor to the distribution. In our simulation file for the Rician channel, we assign $K = 10$, sample rate = 9600 Hz, and maximum doppler shift = 100 Hz, which is consider as small compared to transmitted signal frequency 10^4 Hz.

In Figure 4.8, simulation results for middle and near users show a visible improvement in lowering the BER when the developed DL based LSTM method is applied compared to the MMSE scenario. The near user exhibits a substantial improvement in terms of power saving due to the elimination of interference by the SIC procedure and the relaxed channel condition. In the far user situation, the impact of DL in tracking the channel features is limited due to the weak channel conditions and noise effect from other users.
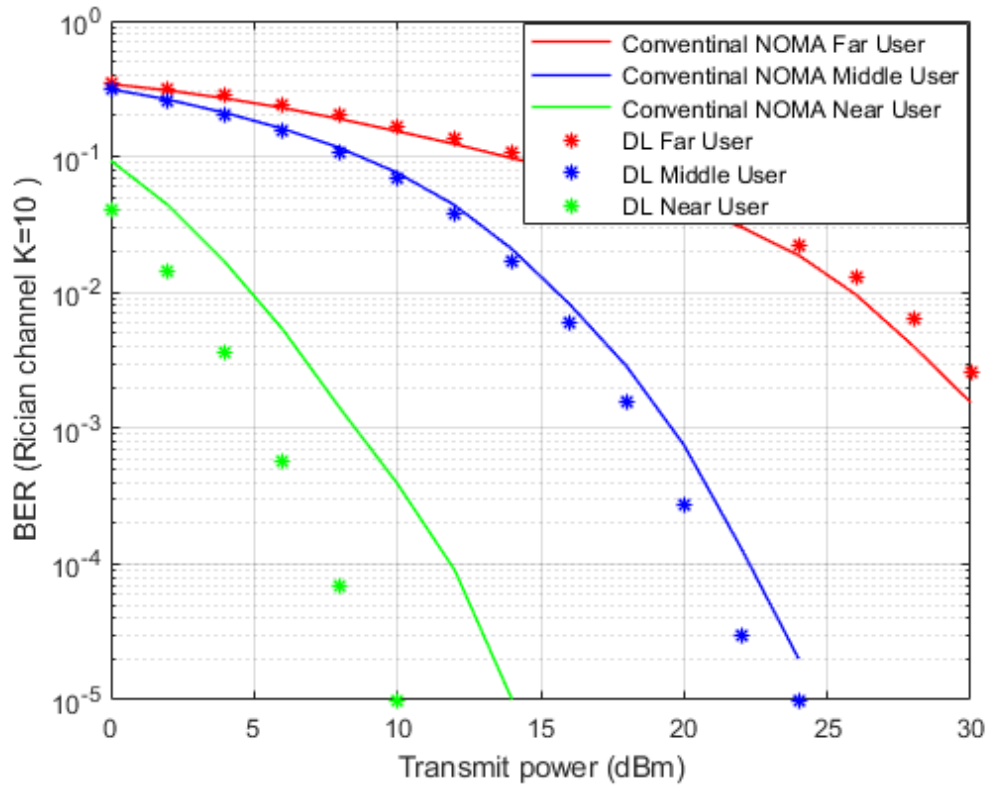
Figure 4.8 BER vs. power for conventional NOMA and DL-NOMA (Rician).

Figure 4.9 shows the simulation outcomes for outage probability metric versus power when Rician channel is applied, and both LSTM and MMSE channel estimation schemes are adopted. Overall, when DL based LSTM channel estimation scenario is conducted, far and middle users' simulation outcomes show an improvement within approximately 4 dBm in terms of power saving compared to the MMSE procedure. In terms of the near user, the simulation results reveal that the developed DL channel estimation scheme starts showing clear enhancement regarding the outage probability when the power assigned to the near user is more than 5 dBm.
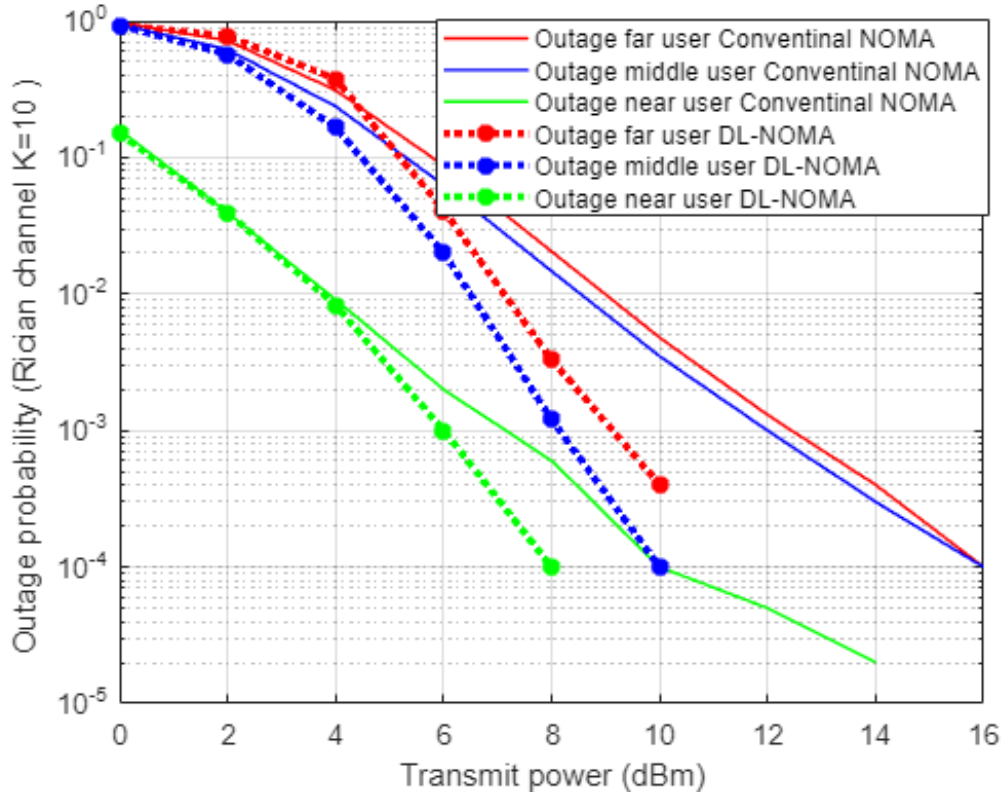
Figure 4.9 Outage probability vs. power for conventional NOMA and DL-NOMA (Rician).

In Figure 4.10, the simulation outcomes regarding the individual capacity for each device are shown where the Rician channel model is applied, and both the DL based LSTM and conventional NOMA based on MMSE procedures are considered as channel estimator. It is clearly noticed that for both far and middle users, DL based LSTM shows comparable capacity compared to the MMSE scheme. This can be clarified that the current LSTM parameters are not sufficient enough to mitigate the interference coming from other users in the cell when Rician distribution is considered. In contrast, and comparable to the Rayleigh fading results in Figure 4.7, the achieved capacity for the near user shows a substantial difference by at least 6 b/s/Hz compared to far and middle users for same applied power level. This improvement in capacity can be defended by the line of site component between transmitter and receiver in the Rician channel and the relaxed fading channel between BS and the near user device.
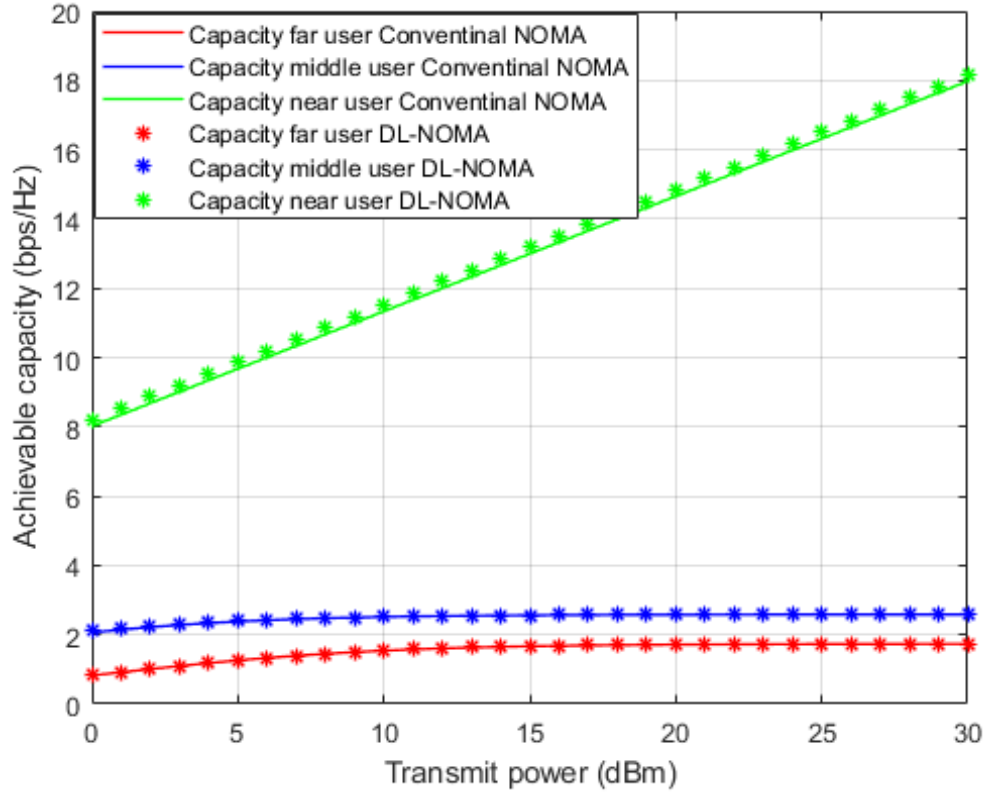
Figure 4.10 Individual capacity vs. power for conventional NOMA and DL-NOMA (Rician).

Figure 4.11 shows the simulation outcomes for the sum rate versus number of users considered in the NOMA cell when the Rayleigh fading model is applied. In addition to our developed LSTM scheme and conventional MMSE procedure that both used as a channel estimator, in this figure, and as an additional benchmark comparison, we have created the simulation parameters related to the work in [86], that implements the DL structure for the purpose of joint channel estimation and signal detection as a one-shot process. As revealed from the results, our proposed DL based LSTM scheme can achieve a significant higher sum rate compared to both the MMSE procedure, and the DL scheme for joint channel estimation and signal detection presented in [86]. It can be clearly observed that while increasing the number of users in the cell, our developed DL channel estimation scheme remains superior in showing higher rates compared to other approaches. These outcomes imply that reliability can be ensured by the proposed DL scheme even when the cell capacity is increased. On the other hand, it is worth mentioning that as the total number of users keeps increasing in the NOMA cell, the interference will also increase, and accordingly, the performance will be degraded.
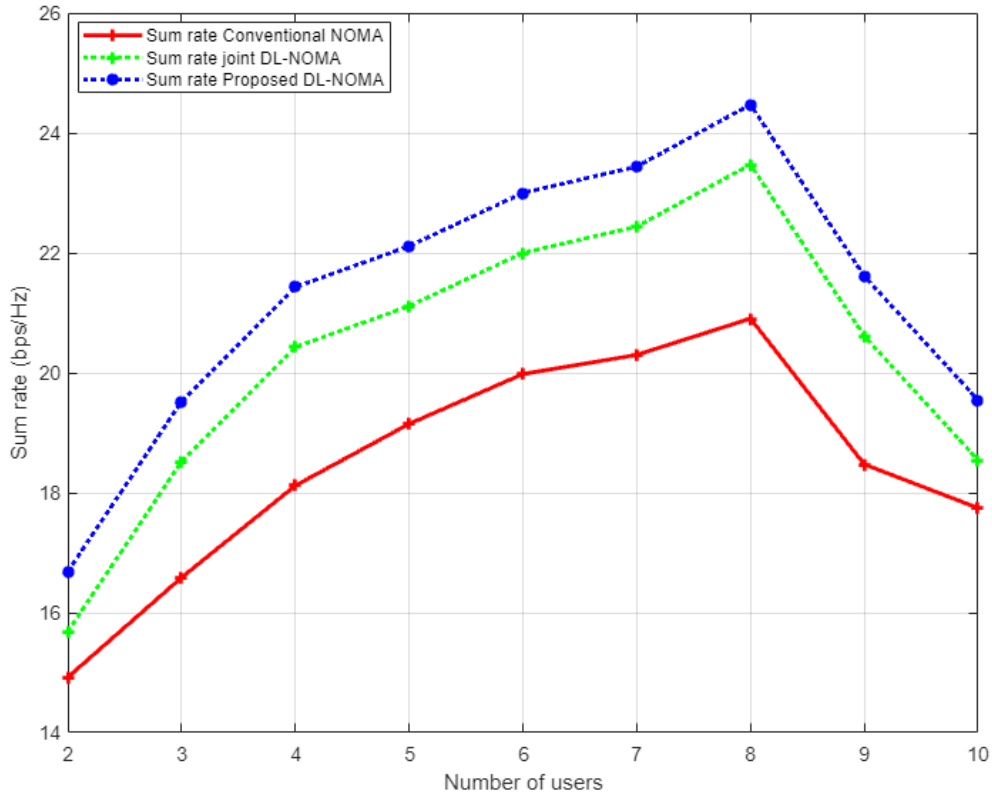
Figure 4.11 Sum rate vs. number of users for conventional NOMA, joint DL-NOMA, and DL-NOMA (Rayleigh).

In Figure 4.12, two different simulation scenarios are performed here to produce this figure. The first scenario mainly depends on applying a fixed power allocation (FPA) scheme for each user in the system. The second scenario depends on applying the optimized power scheme that is derived earlier in chapter 3. Both scenarios are conducted when the proposed DNN based LSTM approach is implemented as a channel estimator. Simulation results for far and middle users verify the dominance of the power optimized structure over the FPA structure in terms of BER. For the near user results, the developed DL based LSTM for channel estimation jointly with FPA scheme provides little enhancement in terms of the received bits error over the optimized power scheme, and this could be justified that for the near user scenario, the good channel condition is more valuable than the assigned power.
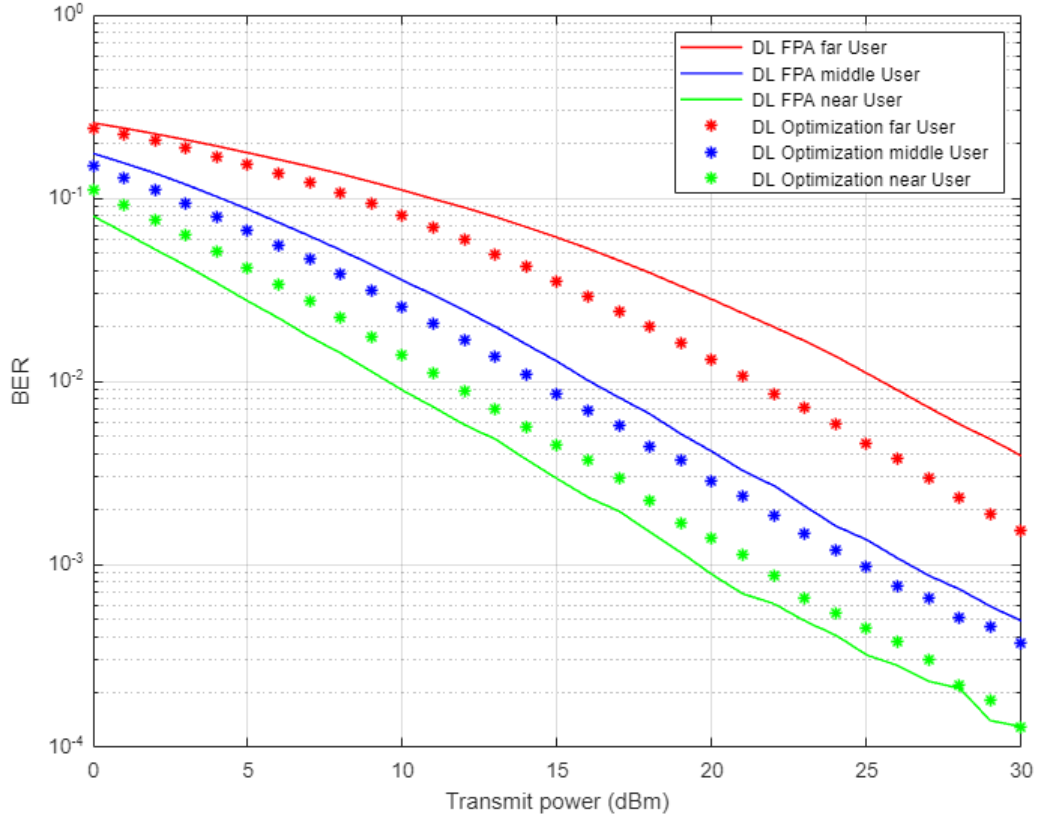
Figure 4.12 BER vs. power for DL-based optimized and FPA schemes (Rayleigh).

Figure 4.13 displays the outage probability results versus the power for far, middle, and near users when the optimized power and FPA scenarios are applied, and both scenarios are conducted in combination with the proposed DL based LSTM for channel estimation task. Far user results show an improvement in outage probability and the power saving is recorded to approximately 5–6 dBm when the optimized power scheme is applied compared to the FPA scheme. Likewise, for the middle user case, both the DL and optimized scheme provide an evident enrichment in the outage probability, with power saving nearly 2–3 dBm. Otherwise, the near user with DL for channel estimation and FPA scenario, shows a considerable outage amelioration compared to the optimized power case. These results also confirm the results obtained for the BER metric, which indicates that FPA jointly with high channel gain, are more sufficient for the near user than the power optimization scheme.
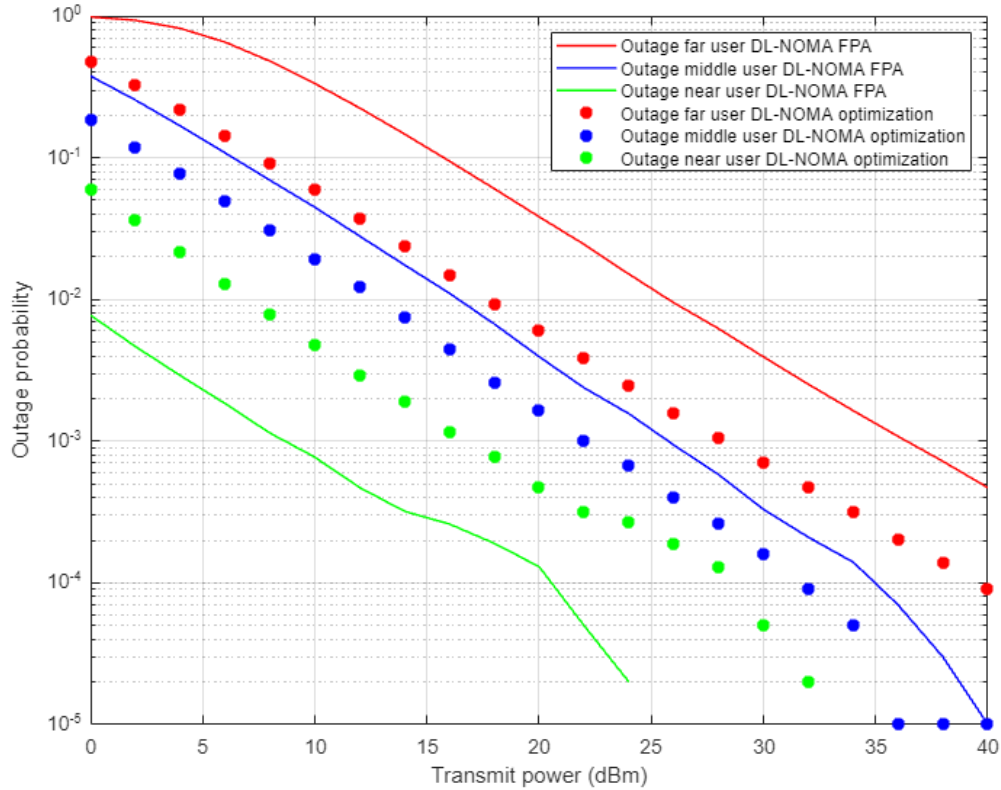
Figure 4.13 Outage probability vs. power for DL-based optimized and FPA schemes (Rayleigh).

In Figure 4.14, the simulation results for the sum rate for the three examined users in NOMA cell are shown. Each of the optimized power scheme and FPA scheme is incorporated with the proposed DL algorithm utilized for estimating the channel coefficients prior to calculating the rate for each user. On the basis of the simulation outcomes, it can be clearly noticed that the channel estimation based on DL combined with the optimized power scheme show little improvement in the sum rate compared to the FPA scenario when the applied power level is low. Starting from 15 dBm, both the optimized power and FPA schemes provide a comparable sum rate when our proposed DL channel estimation scheme is implemented.
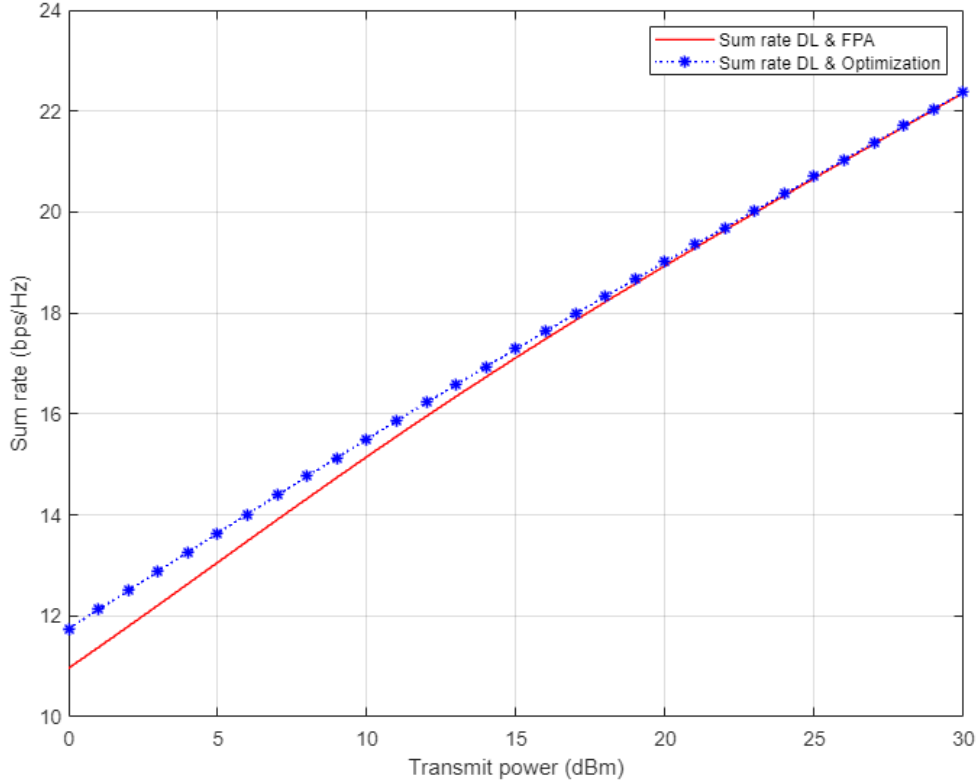
Figure 4.14 Sum rate vs. power for DL-based optimized and FPA schemes (Rayleigh).

4.6 Results Summary

In this section, we decide to provide a sort of summary for the simulation results in terms of a sample statistics to highlight the average percentage improvement achieved by the inspected performance metrics when DL based LSTM is applied as a channel estimator. Table 4.2 presents the average numerical values that reflects the estimated amount of improvement attained by DL based LSTM compared to conventional NOMA scheme when different performance metrics such as BER, outage probability, sum rate, and the individual capacity are considered. Sample power values 5 dBm, and 20 dBm are selected to record the percentage improvements for each metric in Rayleigh fading channel. Based on the sample measurements shown in the table, it is clearly noticed DL based LSTM is providing a sufficient improvement for different performance metrics for all users. In same scenario, Table 4.3 is indicating to the average improvement realized by DL based LSTM over the conventional NOMA scheme when both are implemented as channel estimator for the examined users in Rician channel. It can be noticed that the average improvement in Rician channel is not as that shown in the Rayleigh channel, and this was justified by the line-of-sight component in Rician model that clearly

enhance the conventional NOMA scheme based on MMSE procedure. The numerical values that reflect the comparison between the Optimized power scheme and the fixed power allocation (FPA) scheme is shown in Table 4.4 when the DL based LSTM is employed for the channel estimation process. As noticed from recorded values, the optimized power scheme is showing a noticeable improvement over the FPA scheme when far and middle users are considered. On the other hand, the metrics values related to the FPA scheme are indicating the dominance of the FPA scheme over the optimized power scheme when the near user is considered. Also, it is worth mentioning that both optimized power and FPA schemes are providing a comparable performance at high power level when sum rate metric is examined and the DL based LSTM is employed for the channel estimation.

Table 4.2 Sample statistics for average percentage improvement (DL- LSTM vs Conventional NOMA - Rayleigh)

Performance metric	5 dBm	20 dBm	Criteria Improvement	Users	Channel
BER	≈ 70.6%	≈ 68%	DL-LSTM vs Conventional NOMA	All	Rayleigh
Outage Prob.	≈ 64.5%	≈ 71.6%	DL-LSTM vs Conventional NOMA	All	Rayleigh
Sum Rate	≈ 42%	≈ 26%	DL-LSTM vs Conventional NOMA	All	Rayleigh
Individual Capacity	≈ 11.5%	≈ 14.7%	DL-LSTM vs Conventional NOMA	All	Rayleigh

Table 4.3 Sample statistics for average percentage improvement (DL- LSTM vs Conventional NOMA- Rician)

Performance metric	5 dBm	20 dBm	Criteria Improvement	Users	Channel
BER	≈ 31.2%	≈ 27.7%	DL-LSTM vs Conventional NOMA	All	Rician
Outage Prob.	≈ 44%	≈ 91.3%	DL-LSTM vs Conventional NOMA	All	Rician
Individual Capacity	≈ 4%	≈ 4%	DL-LSTM vs Conventional NOMA	All	Rician

Table 4.4 Sample statistics for average percentage improvement (Optimization vs FPA)

Performance metric	5 dBm	20 dBm	Criteria Improvement	Users	Channel
BER	≈ 16.5%	≈ 24 %	Optimization vs FPA (DL-LSTM)	Far & Middle	Rayleigh
BER	≈ 22%	≈ 37 %	FPA vs Optimization (DL-LSTM)	Near	Rayleigh
Outage Prob.	≈ 72.5%	≈ 67.5 %	Optimization vs FPA (DL-LSTM)	Far & Middle	Rayleigh
Outage Prob.	≈ 83%	≈ 62.5 %	FPA vs Optimization (DL-LSTM)	Near	Rayleigh
Sum Rate	≈ 5%	≈ 1 %	Optimization vs FPA (DL-LSTM)	All	Rayleigh

4.7 Summary

In this work, the impact of the Deep Neural Network (DNN) in explicitly estimating the channel coefficients for each user in the NOMA cell is investigated, where the LSTM network is developed for complex data processing. In the proposed DL algorithm, the DNN model is trained online based on both the normalized channel statistics and the relationship between successive training sequences. The validity and efficiency of the proposed DL channel estimation scheme is emphasized by inspecting the proposed DNN model using the Rayleigh fading channel and Rician fading channel. Furthermore, we introduce a framework that investigates how the proposed channel estimation based on the DL and the power optimization scheme are jointly utilized for multiuser detection in the PD-NOMA system. A systematic mathematical analysis for the optimization problem is introduced and the Lagrange function and KKT conditions are employed to deduce the optimal power factors. The simulation results in terms of the BER, outage probability, sum rate, and individual capacity have verified that the proposed DL model-assisted NOMA can realize reliable performance compared to the conventional NOMA scheme, even when cell capacity is increased.

Chapter 5

Reinforcement Learning based Q-Learning for Channel Estimation in MISO-NOMA System

5.1 Introduction

In downlink NOMA system, each user equipment can receive a multiplexing of signals and these signals are related to several user devices in the NOMA cell and this multiplexed signal is transmitted by the base station (BS). Therefore, excluding the interference generated by other user devices come to be important for the sake of coordinated detection. Frequently in power domain NOMA (PD-NOMA), multiuser detection can be handled via successive interference cancellation (SIC) [98]. In the SIC procedure, signals related to different users are decoded successively on the basis of the power percentage assigned for each user and the channel state information (CSI). A broad investigation of CSI for various users in NOMA cell is demanding because pilot data that traditionally exploited in channel prediction, might interfere with symbols from other user terminals, which may lead to disturb the performance of a conventional prediction scheme that based on the minimum mean square error (MMSE) estimator [94]. Besides, power allocation policy is also considered as a critical process for user devices when PD-NOMA is applied in different communication networks.

Deep Learning (DL) and Reinforcement Learning (RL) techniques, have the possibility to track the changes in the channels between user devices and BS, thus, they are recently considered a

powerful tool for the forthcoming radio systems [99]. Hence, estimating the channel parameters or allocating the power factors for user terminals based on the machine learning (ML) algorithms, prompted the authors for more deep investigations in this field in order to boost the system performance and enhance the quality of the signal reconstruction process.

5.2 Related Works

Authors in [100], introduced two approaches to approximate the optimal MMSE channel estimator when the reconfigurable intelligent surfaces (RIS) based MISO system is considered. Basically, the reconfigurable intelligent surface is a programmable surface structure that can be used to control the reflection of electromagnetic (EM) waves by changing the electric and magnetic properties of the surface. In the first presented approach in [100], authors suggest an analytical linear channel estimator to control the phase shift matrix of the RIS scheme during the training phase, and the estimator based on that technique exhibited a sensible accuracy when the statistical properties of the applied channel and noise are both considered. In the second approach, authors analysed the channel estimation problem as an image denoising problem, then they adopt a convolutional neural network (CNN) to achieve the denoising and predict the channel parameters. Numerical outcomes have illustrated that the channel estimator based on CNN approach can offer a better performance compared to the linear estimation method and low computational complexity is also attained.

A neural network model for wireless channel estimator is suggested in [101] to be used with uncoded multi-input multi-output (MIMO) system, to enhance the channel estimation process reliability. Based on the presented neural network structure, a channel estimator is proposed, and also a mathematical scheme is introduced to derive an optimum power scheme to assist in the lowering the channel prediction bandwidth utilization. Simulation outcomes proved that the channel estimator based on the presented neural network structure, can deliver an improvement in mean square error (MSE) and bit error rate (BER) compared to the traditional MMSE channel estimation technique.

Based on deep autoencoder scheme, authors in [102] performed experimental verifications in a massive MIMO system on two tasks, the first task is related to investigate a channel model for wireless links, and the second task is associated with managing the power allocation policy. The proposed deep learning autoencoder scheme is also adopted to handle the issue raised from

inadequate training datasets that may produce critical overfitting problems and therefore affect the model's reliability. Results based on the proposed autoencoder model, have clarified that the suggested structure could successfully enhance the performance particularly when the extent of the training dataset is within a specified threshold selection.

The work in [103] introduced a deep learning framework to handle the limitations raised up when standard iterative techniques for power control are utilized, such as unnecessary latency and high complexity. In the presented DNN framework, the outdated and partial CSI is exploited, to construct an optimization problem that can boost the spectral efficiency in device-to-device communication system. User fairness and energy efficiency constraints were considered, and simulation outcomes showed that the suggested deep learning framework can attain better spectral and energy efficiency compared to the MMSE estimator when several channel correlation factors are considered.

Based on the position of the user device with respect to BS, the path loss, and the CSI, a deep learning model labelled PowerNet is introduced in [104]. The authors attempt to prove that PowerNet scheme has the capability to avoid the time consuming with complex channel estimation procedures, and at the same time, power control can be also achieved. Different from conventional DNNs that depend on a fully connected structure, the proposed PowerNet model is implementing a CNN layer to differentiate the interference characteristics through several channels in wireless networks. Simulation outcomes verified that the introduced PowerNet scheme can attain a stable performance with no explicit channel estimation procedure.

Recently, predicting the channel parameters or assigning the power factors with the assistance of Reinforcement learning (RL), has been explored by many researchers. The authors in [105] provided an end-to-end channel estimation framework based on actor-critic procedure for a downlink multiuser multiple antenna system. Authors principally relies on the agent to utilize the pilot symbols to estimate the channel parameters that will be utilized to create a downlink beamforming matrix. To achieve the purpose of maximizing the sum rate reward, the deep network parameters are updated based on the deep policy gradient method. Simulation outcomes proved that the suggested RL based actor critic procedure for channel estimation can provide a stable performance under diverse channel statistics and can show noticeable

convergence compared to the typical MMSE procedure when the sum rate metric is inspected.

In [106], authors have developed a deep reinforcement learning (DRL) framework for device to device pairing to distinguish the correlation (relationship) patterns in wireless networks. The presented DRL algorithm is implemented to jointly explore the channel selection and the power control problem for device-to-device pairing in order to enhance the weighted sum rate in the system. The proposed DRL learning procedure, can make use of the outdated and local channel information to update the network parameters and perform decisions independently for each device-to-device pair. Simulation results indicated that even without a global channel state information, the suggested DRL scheme is capable to attain a steady performance close to that attained using ordinary analytical approaches.

5.3 Reinforcement Learning Theory and Framework

Reinforcement learning (RL) is typically developed on the basis of a Markov Decision Process (MDP) design, that includes a group of elements [107] as follows: a state space 'S', which is the set of observations or states in the environment and these states can be recognized or observed by the agent. An action space 'A', which is the set of all actions that can be chosen by the agent at each state. An instant reward 'R', which is the direct reward that is sent to the agent after selecting a certain action a $\in A$ to transfer to another state $s \in S$. The policy 'P' that characterizes how the agent can move from the current observed state to the new state based on the action decided by an agent. Another important element in the RL process is the State-action value function $Q(s,a)$, which is formally labelled as the expectation of the cumulative discounted rewards when a specific action a $\in A$ is chosen by an agent and a particular policy is considered. Typically, RL can be observed as a approach for understanding the agent's interaction in a stochastic environment where the agent is progressively selecting actions during a sequence of time slots. Based on the aforementioned discussion, we can conclude that the main objective of reinforcement learning algorithm is to train an agent to carry out a certain task within non static environment [108].

The interaction between the RL agent and the environment can be described by this scenario: at each time step, the agent can observe or recognize the states in the environment, and based on the current state, the agent can decide and perform a specific action. Then, an instant reward will be sent to the agent from the environment. The reward reflects how effective the action is,

when the agent decides a certain action to achieve a specific goal. Principally, at each learning iteration, the RL agent can interact with the environment by following a predefined policy that can regulate the transition between states [109].

Based on the above-mentioned discussion and as shown in Figure 5.1, the RL agent can be represented by two components: the policy and the learning algorithm. The policy is the mapping criterion that decides actions based on the states observed in the environment. In DRL, the policy can be represented as a function approximator with tuneable parameters, such as DNN. The learning algorithm constantly update the weights, or the parameters of the policy based on observations, actions, and rewards. Overall, the objective of the learning algorithm is to realize the optimum or the best possible policy that can maximize the received cumulative reward.
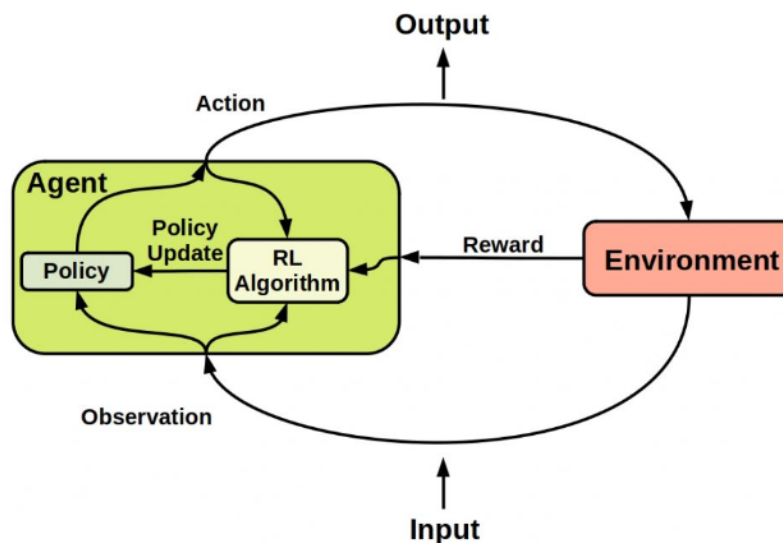
Figure 5.1 Reinforcement Learning Framework.

5.4 Channel Estimation Based Q-Learning Algorithm

In this chapter and in the channel estimation process, we assume that the action space is discrete, therefore, we manage to use an RL based Q-learning algorithm as one of the RL candidate's procedures to decide the channel parameters in the examined NOMA cell [110]. The Q-learning algorithm is categorized as a model free, and off policy reinforcement learning procedure. Also, a Q-learning agent is described as a value-based RL agent that has the role of updating a specific Q-value function to enhance the future rewards. At a certain state, the agent can examine and decide the action for which the expected reward can be maximized. In this

section, RL based Q-learning is employed for channel estimation task in the MISO-NOMA cell and pilot symbols are also utilized to assist in the channel approximation process. Therefore, it is assumed that there is coordination between user devices and BS, such that the pilot symbols can be identified at the BS and the user terminals.

In this chapter, we have considered the BS as the Q-learning agent, and we assume that the BS will start estimating the channel parameters for each user once each user equipment finishes sending the pilot signals in uplink transmission [111]. Therefore, the proposed Q-learning algorithm can be utilized to estimate the channel parameters based on the assistance from the received pilot symbols.

Based on our developed Q learning algorithm, the scenario for the channel estimation process can be summarized in this way. Firstly, at the start of each transmission time slot, user devices can send pilot symbols to BS through the uplink channel. Secondly, based on our developed Q-learning algorithm and the availability of network information such as the distance from the BS, power allocation factor for each user device and the path loss, the BS (agent) can approximate the channel coefficients for user devices. Thirdly, BS will create the superposition coding signal and performs downlink data transmission. Finally, the receiver of each device will receive the superimposed downlink transmitted signal beside the channel parameters that have been estimated based on Q learning algorithm. The predicted channel coefficients will be utilized to decode the desired signal and each device can feedback the SINR or the achieved rate to the BS to boost the detection scenario [11][112].

In this chapter, the main aim of our developed RL based Q-learning algorithm is to maximize the downlink sum rate and minimize the estimation loss. Rather than estimating the received signal, we principally focus on incorporating the developed Q-learning algorithm in the NOMA system for the purpose of channel estimation. The Q agent (BS) is planned to estimate the channel parameters by interacting with the environment, and during the learning iteration, the Q learning agent can decide on the action that may both enhance the estimated state-action value function $Q(s, a)$ and the long-term reward. It is worth mentioning that increasing the number of learning iterations, can sufficiently update the Q-values which may results in improving the channel approximation process and the sum rate reward [112].

In the proposed Q-learning procedure, the instantaneous sum rate R_t at time instant t can be

expressed as follows [11][105][113]:

$$R_t = \sum_{i=1}^{M} \log(1 + SINR_{it})$$
(5.1)

where $SINR_{it}$ is the signal to interference plus noise ratio of user i at time instant t, and M is the number of users in the examined MISO-NOMA cell. In this chapter, the optimum goal of the developed Q-learning algorithm is to maximize the total discounted reward R^γ starting from time instant t, which can be denoted as

$$R_t^\gamma = \sum_{k=t}^{\infty} \gamma^{k-t} R_{k+1}$$
(5.2)

where R_t^γ is the discounted reward at time slot t, and γ is the discount factor. Substituting the sum rate from Equation (5.1) into Equation (5.2), then the discounted sum rate reward, can be expressed as:

$$R_t^\gamma = \sum_{l=t}^{\infty} \gamma^{l-t} \sum_{i=1}^{M} \log(1 + SINR_{i(l+1)})$$
(5.3)

As stated earlier, the Q-learning agent is the BS, whose target is to enhance the accumulative sum rate. Accordingly, two value functions can be inspected while considering the RL maximization problem [11][105], the first function is called the state value function $V(s)$

$$V(s) = E[R^\gamma / (S_t = s)]$$
(5.4)

and the other function is the state-action value function $Q(s, a)$

$$Q(s, a) = E[R^\gamma / (S_t = s, A_t = a)]$$
(5.5)

where E denotes the expectation parameter given that the agent can follow a certain policy within the applied procedure. Owing to the unspecified transition probabilities and limited observed states, an optimal policy is not easy to achieve. Therefore, in this chapter the Q-learning procedure is developed to approximately achieve the best possible policy to select the

best action. In our developed Q-learning algorithm, the state-action value function $Q(s, a)$ values can be learned via trial and error and are updated according to the following formula [11][105][113]:

$$Q(s, a) \leftarrow (1 - \alpha)Q(s, a) + \alpha \left[R(s, a) + \gamma \max_{a' \in A} Q(s', a') \right] \quad (5.6)$$

where α is the learning rate, s' and a' denote the new state, and the new action that will be considered and decided by the agent to maximize the new state-action value function $Q(s', a')$.

5.5 Q-Learning Network Architecture

In data transmission, the transmitted frame involves data and pilot symbols, and in our analysis, we assumed that the applied channel model is stationary throughout one frame transmission and these channel parameters can be changed from one frame to another. The basic framework for the channel estimation process using the developed Q algorithm that employed in our examined MISO-NOMA system is illustrated in Figure 5.2. This structure can be principally described as a sequence of stages [11[113]. The steps for the first stage can be listed as follows:

1- Initialize the Q table (rows = states, columns = Actions)
2- States (initial distance, initial power factor, path loss)
3- Actions (generate random channel coefficients based on the channel model parameters with size 12 + estimate channel coefficients based on the pilot symbols with size 8)
4- We create 10 different states (Increase or decrease the initial distance by a step size 5)
5- We create 10 different actions (10 different channel realizations)
6- Each channel realization is a combination between the random channel coefficients generated from the channel model parameters and channel coefficients estimated based on pilot symbols.
7- The Q table size is 10×10.
8- Each cell in the Q table is now representing a state-action value function $Q(s, a)$ which is in our model is represented by the channel parameters vector with size 20
9- Assign the target rate "$\boldsymbol{R}_T$"
10- Based on
    a. Average generated channel coefficients h_k
    b. Power factor for each user α_k

  c. Total assigned power P_t

  d. Noise power σ^2

We can calculate the rate for each user based on the following expression

$$R_k = \log_2\left(1 + \frac{|h_k|^2 P_t \alpha_k}{|h_k|^2 \sum_{j=1}^{k-1} P_t \alpha_j + \sigma^2}\right)$$

and compare with target rate to update the Q values.

In the second stage, the reward matrix R is initialized with zero values. In our scenario, the state-action value function $Q(s,a)$ can be adjusted based on the difference between the assigned target rate R_T and the calculated user rate R_k for every device k.
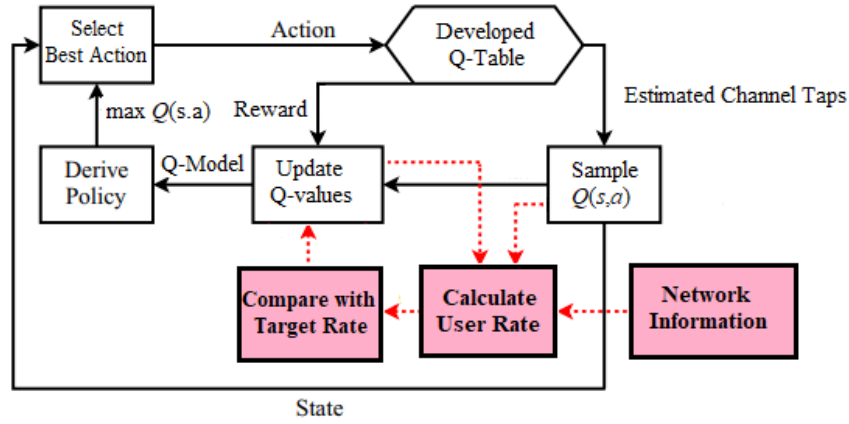
Figure 5.2 Architecture of the proposed Channel prediction scheme based on the developed Q- algorithm.

As mentioned above, the state space S contains the distance between the BS and each user device, the power allocation factor for each user in the NOMA cell, and the path loss. Similarly, the actions that can be chosen by the agent can be selected from the action space A. Based on the aforementioned analysis we created 10 different actions (10 different channel realizations. Hence, in the third stage, based on the trial-and-error scheme, the best policy can be explored, where the best action can be decided and implemented by the Q agent (BS), and then updating the values for the Q table. Additionally, the reward matrix R values can be dynamically assigned according to the actions decided by the Q agent [11][113]. In the fourth stage, the values for the Q-table can be further adjusted according to a Q algorithm procedure with the

aid of the following parameters, the discount factor γ, the learning rate α , and the assigned immediate reward matrix R. It is worth mentioning, that throughout the learning phase, the generated state action values $Q(s,a)$ will be sampled to calculate the new user rate and at the same time update the Q table values till the optimum rate or the terminal state is attained [11][113][114].

5.5.1 Dataset Preparation

Principally, the distance between each user device and the BS and the path loss parameters need to be identified in the dataset to enable random initialization of the channel weights for every user device in the considered MISO-NOMA system. Pilot symbols are generated, transmitted, and identified at the BS and at the receiver side of every user to assist in channel estimation process. As mentioned earlier in section 5.5, based on the state space S and the action space A, the Q-table values can be set up and during the algorithm iterations, and the Q-values can be adjusted according to a Q-learning procedure [11][113]. Throughout the learning process and for the sake of adjusting the values of the Q-table, the discount factor γ, learning rate α, the target rate R_T, current state, and the terminal state should be identified.

In our developed Q-learning algorithm, the Q-agent will decide the next state at random and set it as the next state, then the Q learning agent will inspect all possible actions available in the next state. Next, the Q agent will carefully identify the best action a, that satisfies the maximum value for $Q(s,a)$. After deciding the most suitable action among all available actions, a reward value will be given to the agent as a measure of how successful this action was in order to move to the new state [11][112][113].

In the update phase, we can compute ΔQ, which denotes the difference between the new resultant value function and the preceding value function of $Q(s,a)$. Then, the state action value function in the Q-table will be updated according to the following formula.

$$Q(s,a) = Q(s,a) + \alpha \cdot \Delta Q \qquad (5.7)$$

Based on the updated Q-values in the Q-table, the channel coefficients and the channel gain for each user can also updated. Concurrently, a new attained rate can be calculated and compared to the target assigned rate for each user device. In our developed Q-learning algorithm, once the terminal state or the optimum rate is attained, the developed Q-matrix will

be employed to compose the channel taps for each user device. The developed Q-learning algorithm for channel parameters estimation can be summarized as presented in algorithm 5.1.

Algorithm 5.1 Developed Q-learning Channel Prediction Structure

---

1.　　Initialize the $\boldsymbol{Q}$ table values and initialize the reward matrix $\boldsymbol{R}$ with zeroes.

Inputs

2.　　Number of Iterations and the size for the channel parameters for every user device.

3.　　Initial distance "$\boldsymbol{d_i}$" between each user device and the BS.

4.　　Path loss parameter "$\boldsymbol{\vartheta}$".

5.　　Assign the pilot symbols to assist in the channel estimation process.

6.　　Initialize the random channel parameters for each user "$\boldsymbol{h_{ij}}$" based on the fading Channel model parameters, $j \in [1, 2, ..., N]$ and $i \in [1, 2, ..., M]$. N is the number of antennas at BS and M is the number of users in the cell.

7.　　Assign the power factor for each user. Determine the total transmit power "$\boldsymbol{P_T}$", and noise power spectral density "$\boldsymbol{N_o}$"

8.　　Assign the target rate "$\boldsymbol{R_T}$"

Procedure

9.　　Based on the following parameters: channel gain $\left|\boldsymbol{h_{ij}}\right|^2$, total transmit power "$\boldsymbol{P_T}$", and initial power factor for each user, the signal to interference noise ratio "$\boldsymbol{SINR_i}$" and minimum required rate "$\boldsymbol{R_i}$" can be calculated for each device.

10.　　At each iteration, **compare** the initial generated rate "$\boldsymbol{R_i}$" with the target rate "$\boldsymbol{R_T}$", to update the initial values for the Q-table.

Q-algorithm

13.　Identify the discount factor "$\boldsymbol{\gamma}$", learning rate "$\boldsymbol{\alpha}$", the current state, and the terminal state.

13.　　Choose the next state at random and set it as the next new state.

14.　　**Inspect** all possible actions "$\boldsymbol{a_i}$" in the next state.

15.　　Select the **best action $\boldsymbol{a_i} \in \boldsymbol{A}$**, which satisfies the maximum value for the Q-value function argmax $\boldsymbol{Q(s, a)}$.

16.　　Identify the immediate Reward "$\boldsymbol{R}$"

17.　　Based on the following: maximum Q-value $\boldsymbol{Q(s, a)}$ obtained in step 15, the corresponding reward "$\boldsymbol{R}$", and the discount factor "$\boldsymbol{\gamma}$", then $\boldsymbol{Q(s, a)}$ can be updated based on bellman's equation as follows:
$$\boldsymbol{Q(s, a)} \leftarrow \boldsymbol{R} + \boldsymbol{\gamma} \text{ argmax } \boldsymbol{Q(s, a)}$$

Outputs

18.　　Move to a new state with different distance and different power factor.

19.　　Use the new state with the updated Q values (updated channel parameters) to calculate the new rate until reach target rate or terminal state.

20.　　Compute the difference "$\Delta\boldsymbol{Q}$" between the updated Q value function $\boldsymbol{Q_{new}(s, a)}$ and the old $\boldsymbol{Q(s, a)}$.

21.	Based on (20), $Q(s, a)$ value in the Q-table can be further updated according to $Q(s, a) \leftarrow Q(s, a) + \alpha \cdot \Delta Q$
22.	Check whether the terminal state or **optimum rate** has been reached or the episode completed.
23.	Compose the predicted channel parameters $\widehat{h}_i$

Detailed Q learning Procedure and Workflow:

Based on algorithm 5.1, we can list the detailed workflow for the proposed Q learning algorithm that is developed for estimating the channel parameters for each user in the examined MISO-NOMA system as follows:

1- Initialize the Q table (rows = states, columns = Actions)

2- States are represented by (initial distance, initial power factor, path loss)

3- Actions are represented by (generate a random channel coefficient based on the channel model parameters with size 12 + estimate the channel coefficients based on the pilot symbols with size 8)

4- We create 10 different states (Increase or decrease the initial distance by a step size 5)

5- We create 10 different actions (10 different channel realizations)

6- Each channel realization is a combination between the random channel coefficients generated from the channel model parameters and the channel coefficients estimated based on pilot symbols.

7- The Q table size is 10×10.

8- Each cell in the Q table is now representing a state-action value function $Q(s, a)$ and each $Q(s, a)$ will initially represent the channel parameters h_k vector with vector size 20.

9- Assign a certain target rate $"R_T"$

10- Based on the generated channel coefficients h_k, Power factor for each user α_k, total assigned power P_T, noise power σ^2, we can calculate the rate $"R_k$ for each user.

11- Compare the calculated rate $"R_k$ with target rate $"R_T$ to update the initial Q values.

12- Choose the next state by random and inspect all available actions (all available generated channel parameters in step 6) and choose one that can provide the highest rate.

13- Assign the Reward R

14- Use bellman equation $Q(s, a) \leftarrow R + \gamma \operatorname{argmax} Q(s, a)$ to update the Q values.

15- Move to a new state with different distance and different power factor.

16- Use the new state with the updated Q values (updated channel parameters) to calculate the new rate until reach target rate or terminal state.

17- For each user, repeat steps (1-16) for 10^5

18- We create 10 frames of data bits.

19- For each data frame, we estimate the channel coefficients by averaging the updated channel parameters with vector size 20, from the Q table at the end of each iteration.

5.6 Simulation Parameters

Characterization of the simulation parameters and settings is discussed in this section. The examined downlink MISO-NOMA system contains three distinct user devices and one BS in which the BS is supplied with two antennas and every user device in the cell is provided with a single antenna. In the examined NOMA structure, the modulated signals in downlink transmission are superimposed and transferred by BS to user devices via independent Rayleigh or Rician fading channels that are influenced by AWGN with noise power density assigned as $N_0 = -174$ dBm/Hz and the path loss is set to 3.5.

MATLAB platform has been utilized as a simulation tool to satisfy the following aims, (1) inspect, characterize, and evaluate the performance of the developed RL based Q-learning algorithm when implemented as a channel estimator in the considered MISO-NOMA system, (2) investigate the reliability of incorporating the developed Q-algorithm as channel estimator scheme with the derived optimized power policy derived in chapter 3, for the examined MISO-NOMA network, and different performance metrics are considered to explore the impact of this integration process. (3) The optimized power allocation scheme and fixed power allocation scheme are both compared when the developed RL based Q algorithm is utilized as a channel estimator in the MISO-NOMA cell. Monte-Carlo simulations are performed with $N = 10^5$ iterations, and at the beginning of each iteration, pilot symbols are randomly generated and recognized at the BS and each device. The main simulation parameters are summarized in Table 5.1.

Table 5.1: Simulation environment parameters.

Parameter	Value
Simulation Tool	MATLAB
Modulation type	QPSK
Number of Users	3, [2–10]
System Bandwidth B	1000 kHz
Fading channel	(Rayleigh, Rician)
Path loss exponent ϑ	3.5
Number of Iterations	10^5
Noise PSD N_0	-174 dBm/Hz
Learning Rate α	0.1
Discount factor γ	0.99
Epoch	1

5.6.1 Simulation Setup

The simulation figures are generated based on the assumption that the explicit channel coefficients are not available at each user device in advance. Thus, in order to examine the effectiveness of the developed RL based Q-learning procedure, and for the sake of comparison, we established four additional simulation environments for the channel estimation process. (1) Standard minimum mean square error (MMSE) for channel prediction scheme [94]; (2) DNN algorithm based on LSTM network for channel estimation applied in [11]; (3) RL based actor-critic procedure for channel prediction applied in [105]; (4) the fourth simulation environment is created depending on RL based State-Action-Reward-State-Action (SARSA) procedure [115]. Throughout the simulations, we point out to MMSE technique as conventional NOMA, to denote that user devices are applying the MMSE technique for predicting the channel parameters prior to reconstructing the desired signal.

In the simulation environment, NOMA parameters are generated, and the channel parameters are created to initially model the Rayleigh fading channels. In our developed Q-learning algorithm, at the end of the training episode, or if the terminal state is reached, the updated

$Q(s,a)$ values in the Q-table will be employed to generate a practical channel coefficient for the user devices. Different power percentages are initially assigned for every user device according to channel gain and based on the current distance from the BS. Power factors η_n, η_m, and η_f are specified for near, middle, and far users respectively. In a fixed power allocation setup, we designate $\eta_f = 0.65$, $\eta_m = 0.25$, and $\eta_n = 0.1$. In the optimized power structure (OPS), power factors are allocated for the devices in proportion to the analytical formula concluded for every device in chapter 3. In the simulation files, the transmission distance for each user device with respect to BS is assigned as follows: $d_f = 900$ m, $d_m = 400$ m, and $d_n = 100$ m. Data and pilot symbols are modulated using Quadrature phase shift keying (QPSK) as the modulation format and the applied transferred power is mostly varying from 0 to 30 dBm.

5.6.2 Results Discussion and Analysis

Simulation results that illuminate the comparison between the proposed RL based Q-learning algorithm and the conventional NOMA scheme that adopts MMSE procedure to estimate the channel parameters for each user device are shown in Figure 5.3 in terms of BER versus power. The estimated channel parameters based on Q algorithm and MMSE procedure are employed for the signal detection for each device and the simulated results are shown where fixed power allocation (FPA) scheme is applied. When the proposed Q-algorithm is applied for channel estimation, every user equipment in the examined MISO-NOMA cell offers a visible enhancement in lowering the BER compared to the MMSE technique. At particular BER value such as 10^{-2}, the achieved power saving by the Q-learning algorithm is within 2 dBm for far and middle user devices, while 1 dBm power reduction is observed for the near user.
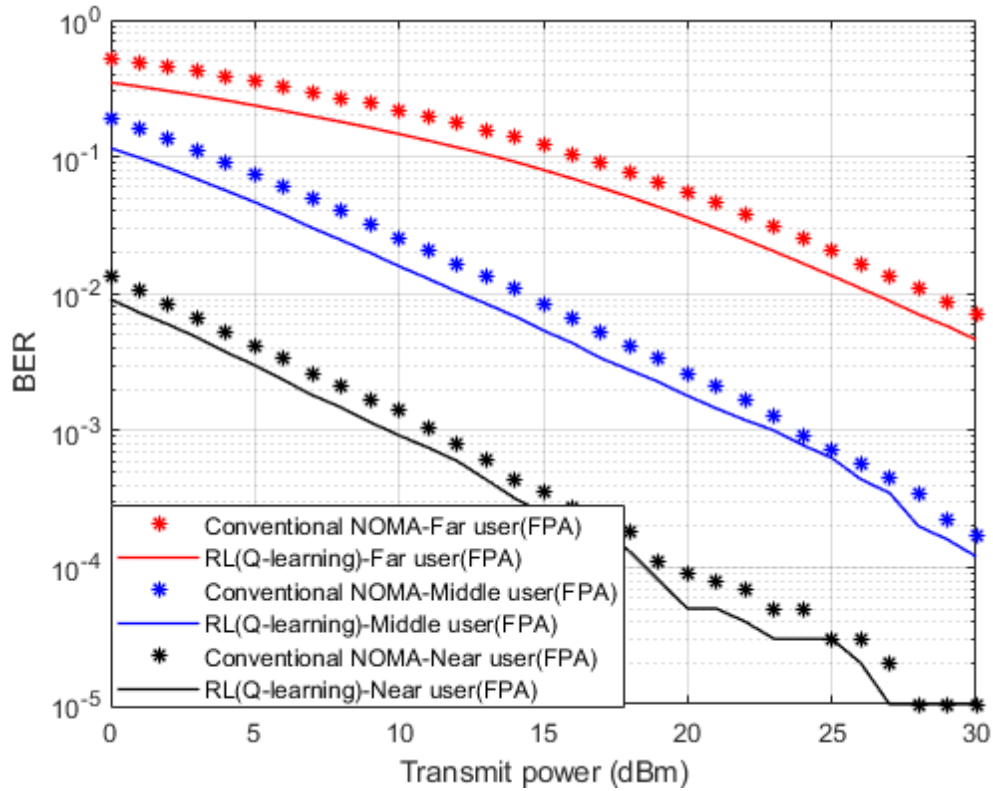
Figure 5.3 BER vs. power (Q-learning & Conventional NOMA (MMSE)).

In terms of the outage probability metric versus power, Figure 5.4 shows the results for the examined user devices in the MISO-NOMA cell when the proposed Q-learning algorithm and standard MMSE are utilized as a channel estimator. Far, and middle user devices simulation results point toward 2 dBm improvement in saving power to realize 10^{-2} outage probability when the proposed Q-learning algorithm scenario is applied compared to the MMSE procedure. Likewise, a near user with the developed Q-learning procedure exhibits a 1 dBm enhancement in power saving with respect to the MMSE scheme. This clear improvement in power saving can verify the advantage of the developed Q-model as a channel estimator compared to the MMSE procedure.
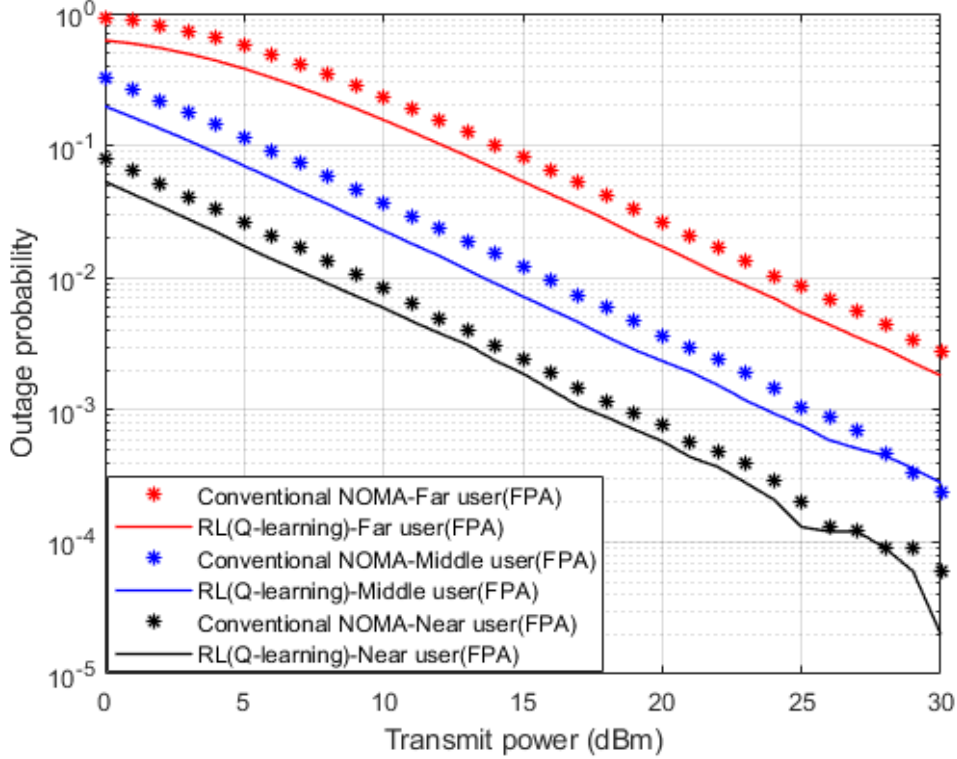
Figure 5.4 Outage Prob. vs. power (*Q*-learning & Conventional NOMA (MMSE)).

In Figure 5.5, we simulate three benchmarks' environments for comparison: (1) standard minimum mean square error (MMSE) procedure for channel estimation [94]; (2) DNN algorithm based on LSTM for channel estimation applied in [10]; and RL based actor-critic procedure for channel prediction applied in [105].

Based on Equation (5.3), for the discounted sum rate reward, and the detailed workflow for the proposed Q learning algorithm, Figure 5.5 displays the simulation results for the sum rate for all users in the MISO-NOMA cell versus power. Based on the simulation results, it is apparently shown that the developed RL based Q algorithm shows superiority over the classical MMSE procedure by 12 b/s/Hz. Moreover, the proposed Q learning scheme performs an improvement over the DL based LSTM scheme presented in [10] by 2 b/s/Hz. For the third benchmark in [105], we create the simulation environment according to the following parameters: the actor and critic networks are both consist of two hidden layers with 400 and 300 nodes, respectively. The learning rate for actor and critic networks are 10^{-4} and 10^{-3} respectively. The discount factor γ is set to be 0.9 and has a buffer size of 10^5. Our proposed RL based Q-learning procedure, shows superiority over the RL based actor-critic procedure at

low power levels while starting from 23 dBm the actor-critic procedure starts showing clear improvement in terms of sum rates compared to the Q-learning algorithm. These results can simply verify that the developed Q-learning algorithm can be a competitive procedure compared to other algorithms that primarily depend on hidden layers to estimate the channel parameters.
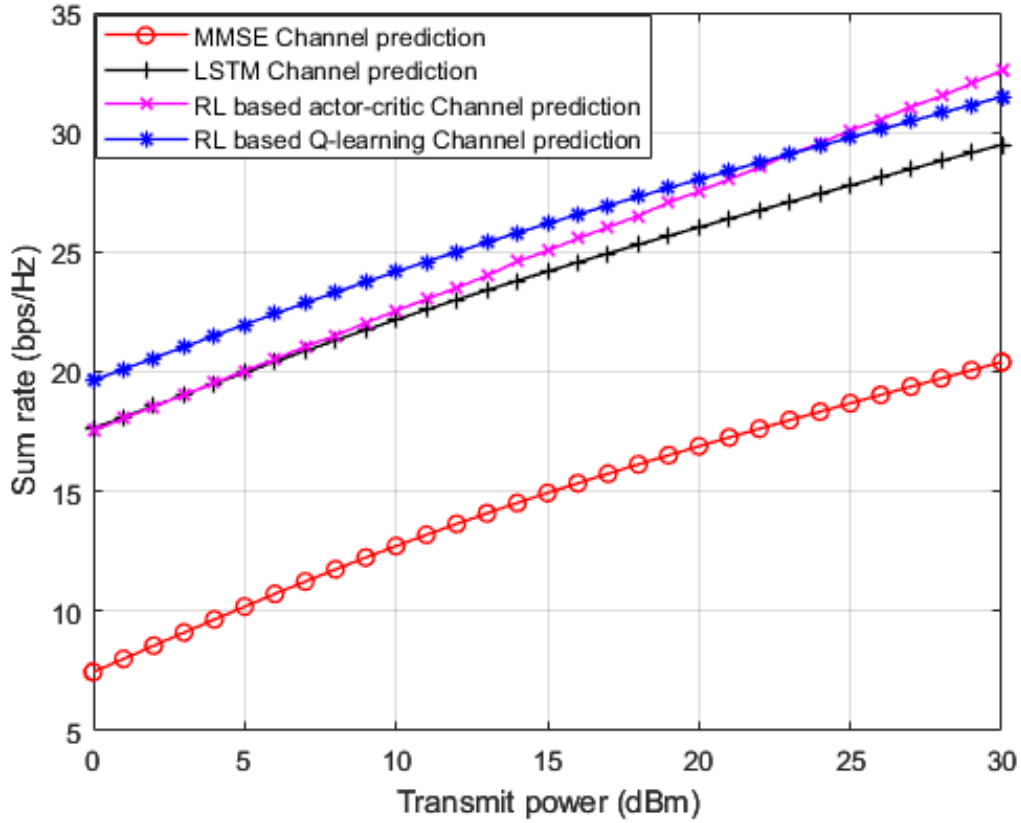
Figure 5.5 Sum rate vs. power (MMSE, LSTM, RL Actor-Critic, RL *Q*-learning).

Simulation results for sum rate maximization against different number of users in MISO-NOMA cell are shown in Figure 5.6, where the reference power is assumed to be 1 dBm. In addition to our developed Q-learning algorithm, three different channel estimation procedures are explored as a benchmark comparison: (1) standard minimum mean square error (MMSE) procedure for channel estimation [94]; (2) DL algorithm based on LSTM for channel estimation applied in [10]; and RL based actor-critic procedure for channel prediction applied in [105]. As shown in the results, our proposed RL based Q-learning algorithm can realize a noticeable greater sum rate compared to MMSE procedure, by at least 2 b/s/Hz. It can be noticed that as the number of the users in the cell is increasing, the performance of RL based

Q-learning algorithm still shows dominance in achieving higher rates compared to MMSE and LSTM channel estimation schemes. Similar to Figure 5.5, the RL actor-critic procedure applied in [105] is created in our MISO-NOMA environment with the following parameters: the actor and critic networks are both consist of two hidden layers with 400 and 300 nodes, respectively. The learning rate for actor and critic networks are 10^{-4} and 10^{-3} respectively. The discount factor γ is set to be 0.9 and has a buffer size of 10^5. As shown in the results, the proposed Q-learning algorithm is showing an advantage over the actor-critic scheme with up to 6 users in the cell. Then, the two hidden layers feature in the actor-critic procedure starts showing enhancement in the sum rates compared to the Q-learning algorithm while the number of users in the cell is increasing. Overall, these outcomes confirm that dependability can be guaranteed by the developed Q-learning algorithm even when the devices in the cell are increased. Furthermore, it is worth saying that while raising the user devices in the system, the interference will also grow up, hence the sum rate could be affected.
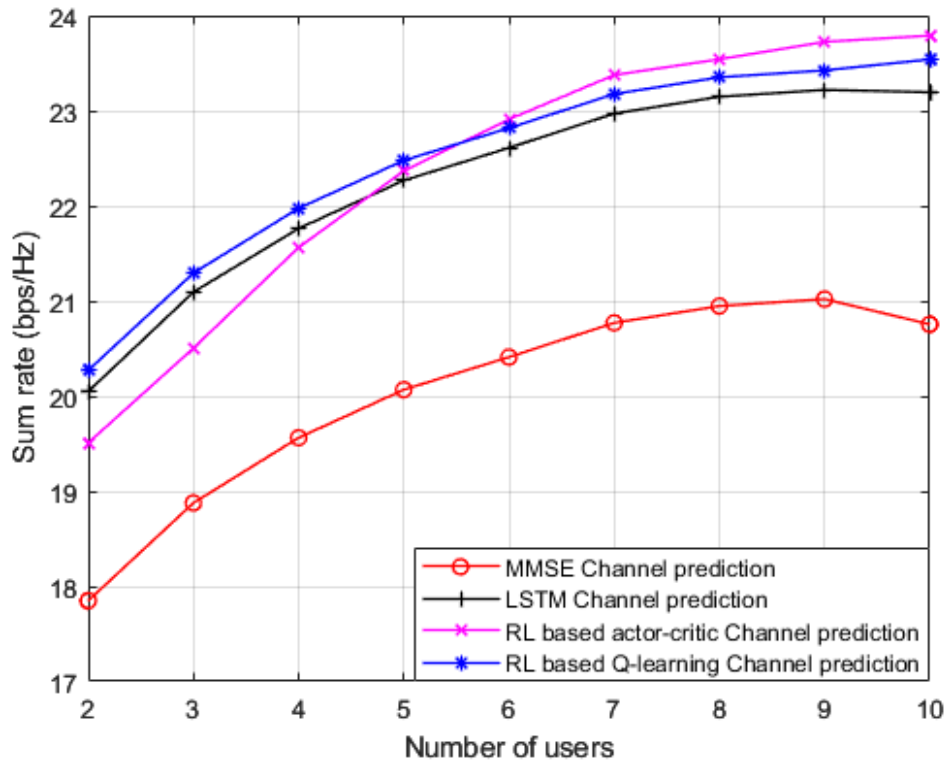
Figure 5.6 Sum rate vs. number of users (MMSE, LSTM, RL actor-critic, RL *Q*-learning).

The proposed Q-learning algorithm and conventional MMSE procedure will be further judged when the Rician channel is applied for the path among BS and each device. Rician channel is

a stochastic model for wireless transmission where the signal gets at the receiver device via various scattered paths. Figure 5.7 illustrate simulation results for BER against power transmitted when the Rician fading channel is applied. In the Rician simulation environment, we assign parameter $K = 10$, where K is described as the fraction of the signal power of the line-of-sight path to the signal power of the remaining scattered components. In addition, maximum doppler shift = 100 and sample rate = 9600 Hz are assumed. Results for the Rician channel reveal that the Q-learning scheme still can present some sort of improvement in decreasing the BER compared to the MMSE technique. This slight improvement can be justified by the presence of a line of site component among BS and user device which can also enhance the work of the MMSE technique.
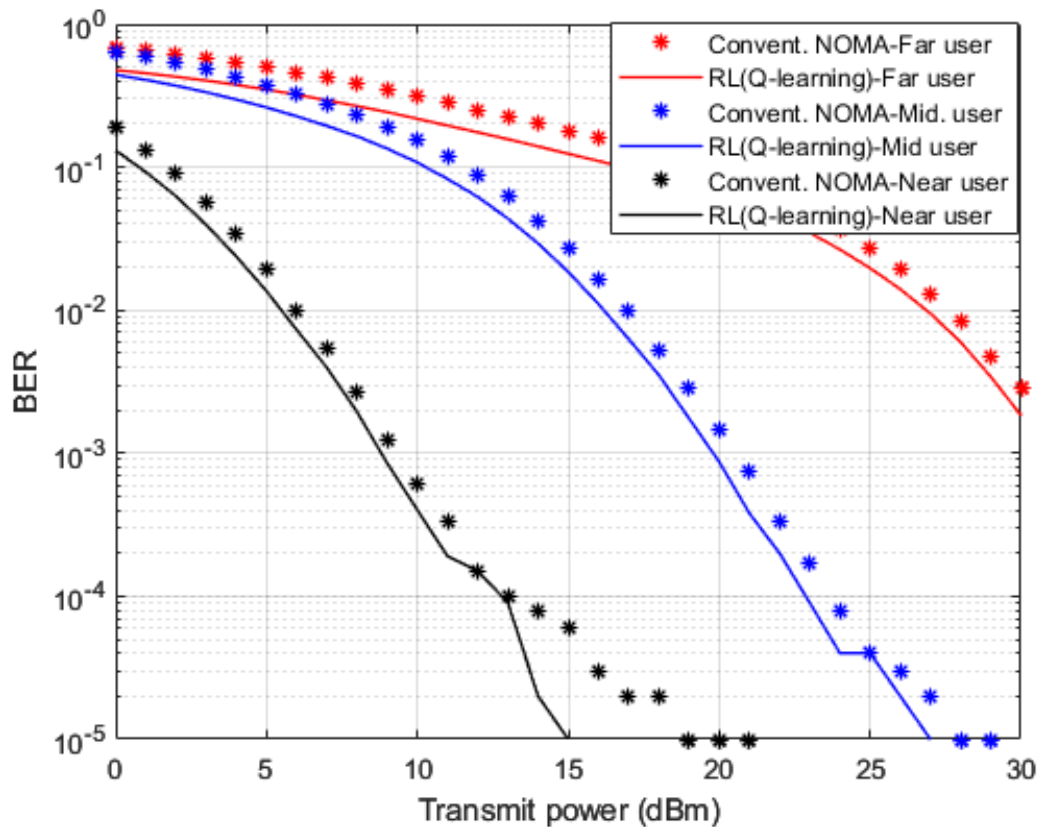
Figure 5.7 BER vs. Power (Q-learning, Conventional NOMA (MMSE) – Rician channel)

In addition to the three benchmarks comparisons realized in Figures 5.5 and 5.6, we also create and implement RL based State-Action-Reward-State-Action (SARSA) algorithm [115] in Figures (5.8–5.10) for the sake of more explorations and comparisons. The parameters and features of the SARSA algorithm are designed in order that the SARSA technique can be used

as a channel estimator and compare the results of SARSA algorithm with the outcomes attained based on our proposed Q-learning algorithm.

The Q-learning algorithm and SARSA algorithm are two efficient RL algorithms, they are both table-based techniques with a Q-table to record the equivalent Q-values of each state-action pair. However, when the dimension of state space increases, it will require a considerable amount of memory. Similar to Q-learning procedure, the SARSA algorithm also working based on the exploration and exploitation procedures, and it also uses a Q-table to record $Q(s_t, a_t)$ value corresponding to state s_t and action a_t. Differently, the running steps of the SARSA algorithm [115] are as follows. First, according to the action selection policy, the gent at the current state s_t, can select the action a_t randomly not based on maximum value. Then, the agent will get an immediate reward R based on the corresponding $Q(s_t, a_t)$ value. Then, s_t will transfer to s_{t+1} and the agent will choose the next action a_{t+1} based on same action selection policy happened in last selection. Hence, the SARSA algorithm is a bit different from the Q-learning procedure, where the Q-value in SARSA algorithm is updated based on the action a_t selected by the agent at the state s_t. While in the Q-learning procedure, the action that has the greatest Q-value in the next state s_{t+1} will be selected to update Q-table.

In Figure 5.8, the BER performance metric is simulated against the transmitted power, when the proposed Q-learning algorithm, SARSA algorithm, and Actor-critic algorithm are all implemented in the simulation environment. It can be noticed that the both the proposed Q-learning algorithm and SARSA algorithm show a comparable performance. However, at high power levels, our developed Q-learning algorithm shows some improvement compared to the SARSA algorithm, which is expected since that the Q agent can decide or select the greedy action, which is the action that offers the maximum Q value for the current state. As a benchmark comparison, the RL based actor-critic algorithm applied in [105] is created in simulation environment with the following parameters: the actor and critic networks are both consist of two hidden layers with 400 and 300 nodes, respectively. The learning rate for actor and critic networks are 10^{-4} and 10^{-3} respectively. The discount factor γ is set to be 0.9 and has a buffer size of 10^5. As shown in the results, with respect to far user, the proposed Q- algorithm, SARSA algorithm, and Actor-Critic algorithm all show a comparable performance up to 20 dBm level. Afterward, Actor-Critic algorithm start showing a noticeable improvement in terms

of lowering the error rate. In terms of the middle user results, the simulation outcomes indicate a clear performance improvement starts from 15 dBm when Actor-critic algorithm is applied. For the near user scenario, it can be noticed that a apparent enhancement in the BER is achieved when actor-critic algorithm is applied compared to the Q algorithm and SARSA algorithm. This apparent enhancement can be justified by the two hidden layers feature in the actor-critic procedure, and the good channel condition for the near user after applying the SIC procedure.
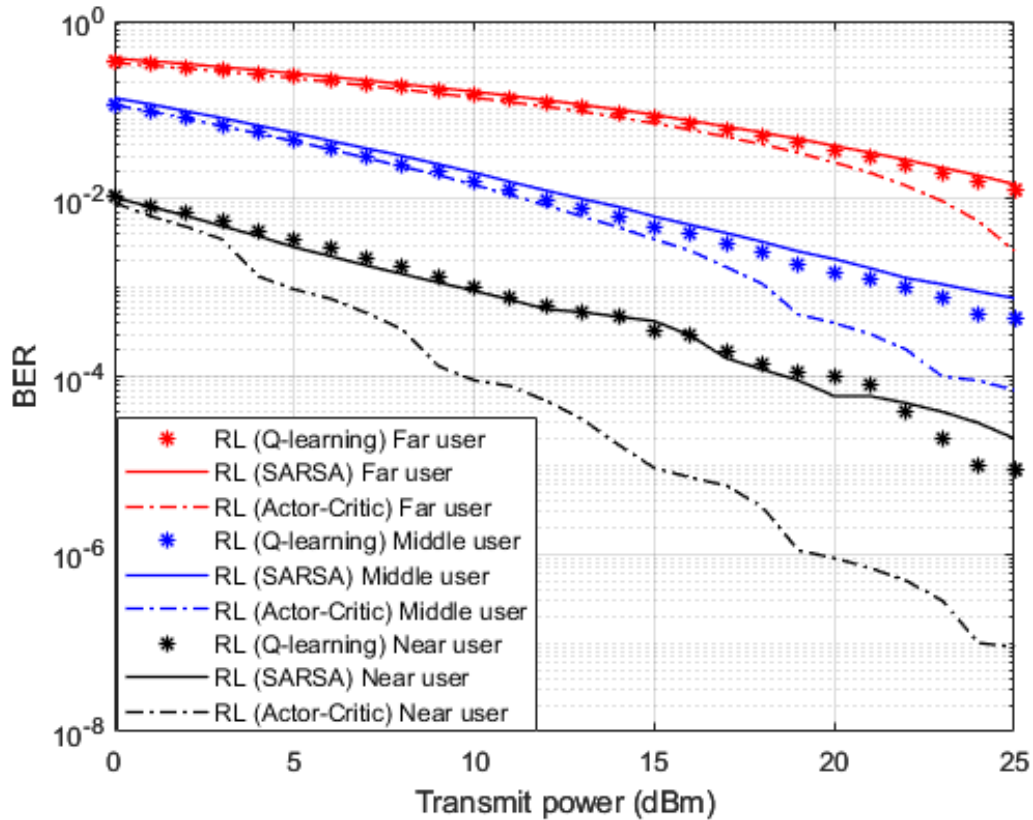
Figure 5.8 BER vs. power (*Q*-learning, SARSA, Actor-Critic).

In Figure 5.9, The outage probability metric is simulated against transmitted power where both the proposed Q-learning algorithm and the SARSA algorithm exhibit a comparable performance. Similar to Figure 5.8, at high power levels, our developed Q-learning algorithm shows some improvement compared to the SARSA algorithm, which can be justified that the Q agent apply the greedy action, which is the action that offers the maximum Q value in the current state.
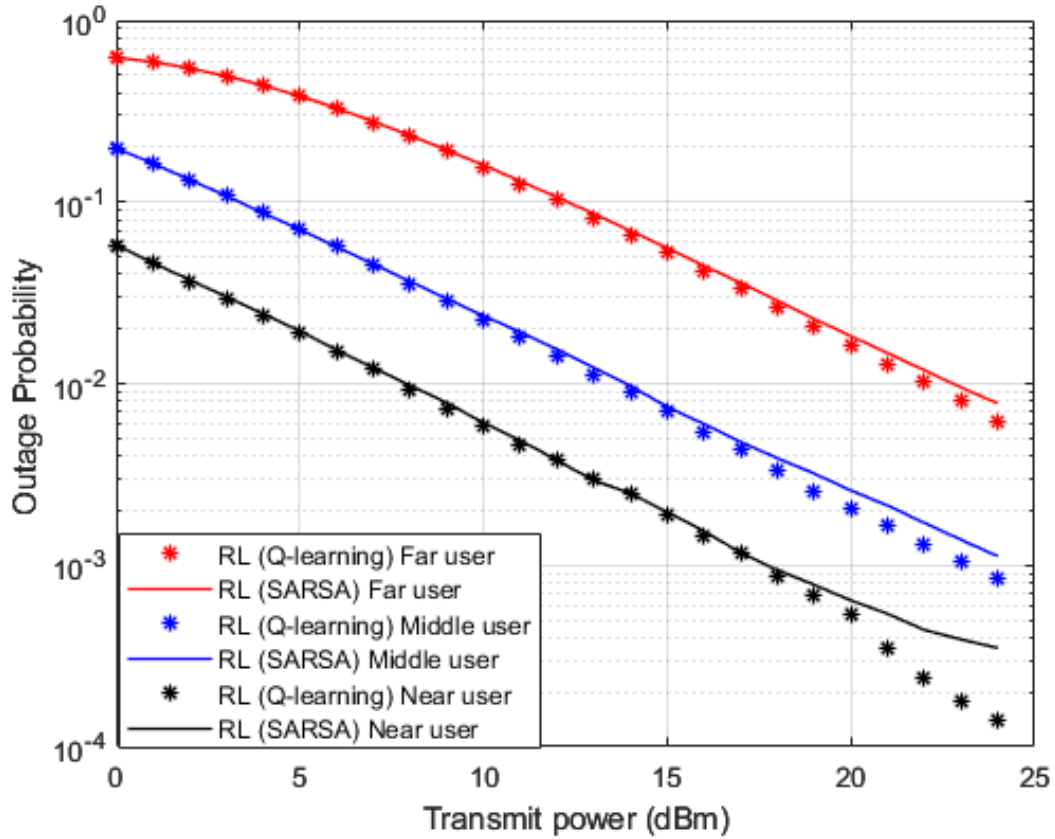
Figure 5.9 Outage probability vs. power (Q-learning, SARSA).

More simulation outcomes for the comparison between our developed Q-learning algorithm and SARSA algorithm are shown in Figure 5.10, where the user's sum rates versus transmitted power are simulated and it is apparent that our developed Q-learning algorithm can deliver more rates over the SARSA algorithm, and at same time a power saving is recorded by 1–2 dBm approximately at any power level applied.
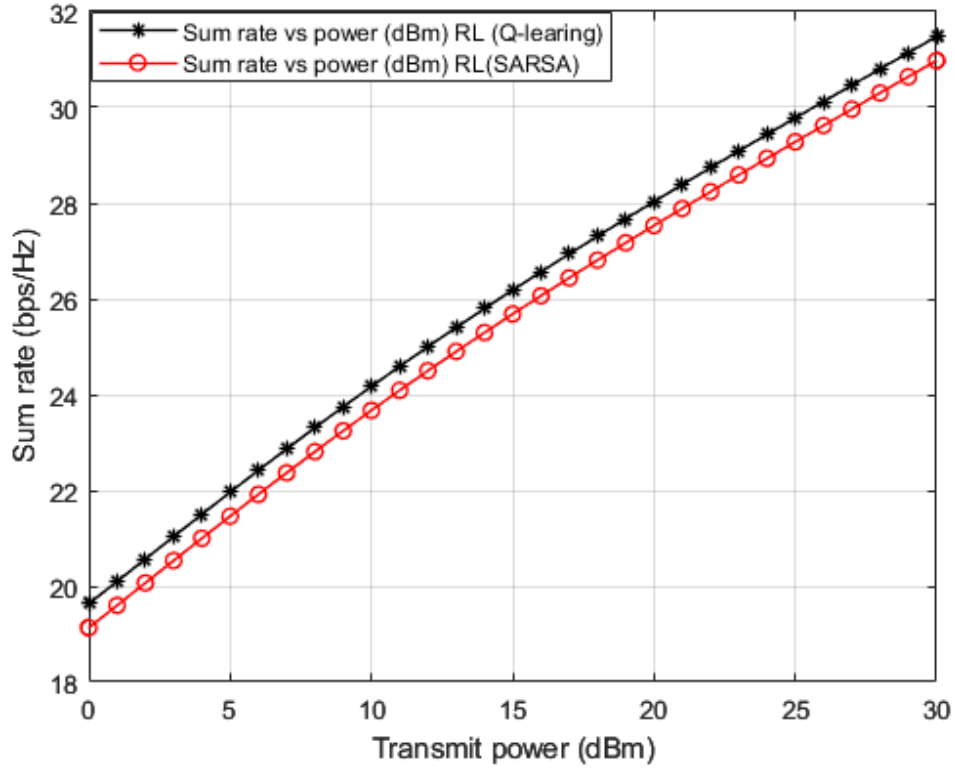
Figure 5.10 Sum rate vs. power (Q-learning, SARSA).

In Figure 5.11, two independent simulation environments are accomplished here to produce these results. In the first simulation environment, the Fixed Power Allocation (FPA) scheme is applied for every user device in the MISO-NOMA cell. The second simulation environment depends on the optimized power scheme (OPS) that previously derived in chapter 3. Both FPA and OPS will be applied in conjunction with the developed Q-learning algorithm that will be adopted as a channel estimator. Simulation results in terms of BER indicate that far and middle devices confirm the dominance of the OPS over the FPA. It can be noted that at specific BER values such as 10^{-2}, the attained power saving by OPS is about 5 dBm for the far device, and 1–2 dBm approximately for the middle user device. On the other hand, for near user results, the proposed Q-learning algorithm jointly with the FPA scheme provide obvious improvement in terms of BER than the OPS, this could be explained that for near device scenario, the stable channel condition plays a major impact in enhancing the system performance than the assigned power scheme.
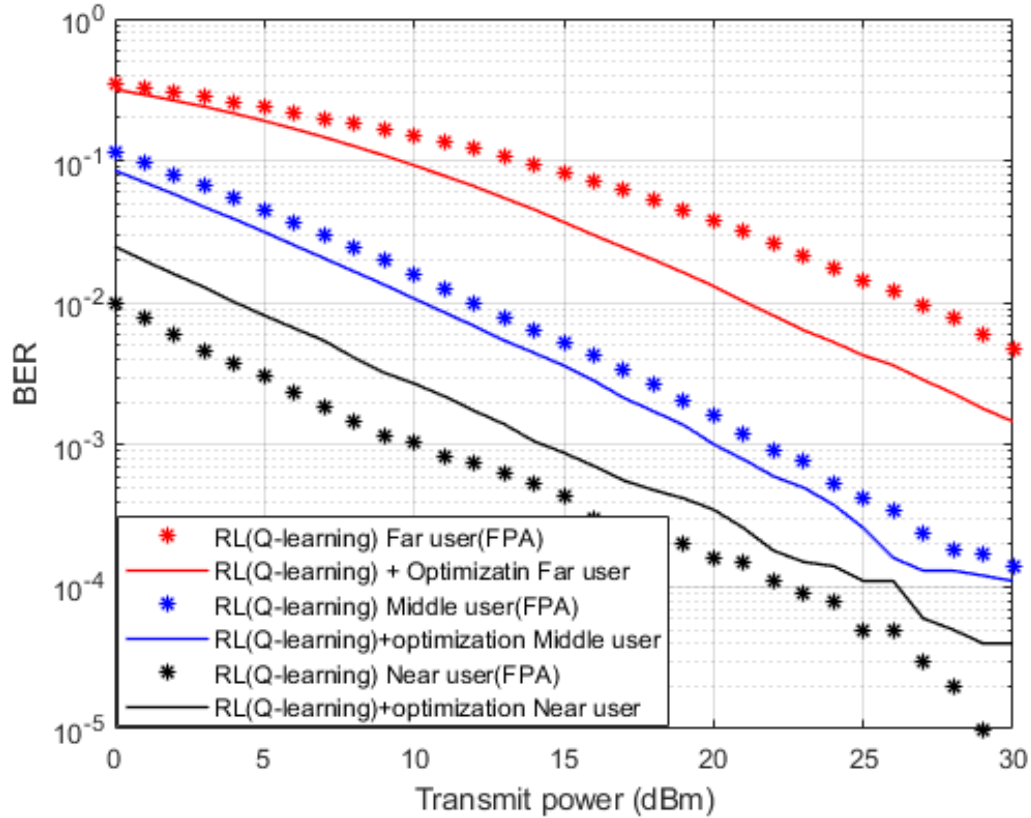
Figure 5.11 BER vs. Power (*Q*-learning, Optimization, FPA).

Based on the OPS scheme and the FPA scheme, Outage probability results versus power are displayed in Figure 5.12 and both schemes of OPS and FPA are applied in conjunction with the developed Q-learning algorithm as a channel estimator in our examined MISO-NOMA cell. Both far user and middle user results still show an amelioration in outage probability where a power saving can be detected within 1–2 dBm when OPS scheme is applied compared to the FPA scheme. Alternatively, near user with a Q-learning algorithm and FPA scenario exhibits a noticeable outage amelioration compared to the OPS case. Approximately, a power reduction within 4 dBm is realized when the FPA scheme is applied. These outcomes validate the results obtained for BER in Figure 5.11, which indicate that the FPA scheme can provide a more satisfactory performance for user devices with good channel condition.
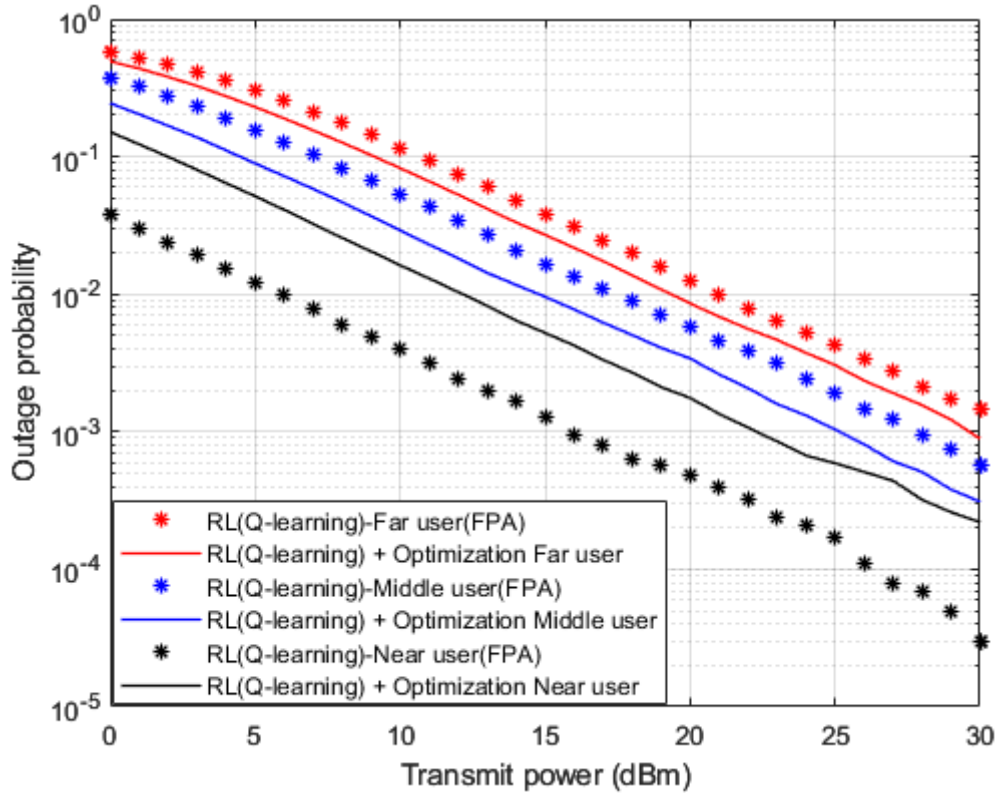
Figure 5.12 Outage Prob. vs. Power (*Q*-learning, Optimization, FPA).

In Figure 5.13, the expected bit rates for the examined user devices in the considered MISO-NOMA cell are simulated against power transmitted when the OPS scheme and the FPA scheme are applied in conjunction with the developed RL based Q-learning algorithm that is utilized as a channel estimator. Results for far and middle users indicate that OPS provides 1 b/s/Hz improvement compared to the FPA scheme. This limited enhancement might be justified that the management of the power allocation for the users is not necessarily adequate enough to reduce the effect of interference particularly for far and middle users that mainly experience unstable links environments. As expected, results for near user device reveal dominance in the attained rate with respect to middle and far devices with at least 10 b/s/Hz. Likewise, the results for the near user with FPA point out a evident improvement compared to OPS, and this outcomes also validate the results obtained in Figure 5.11, and Figure 5.12.
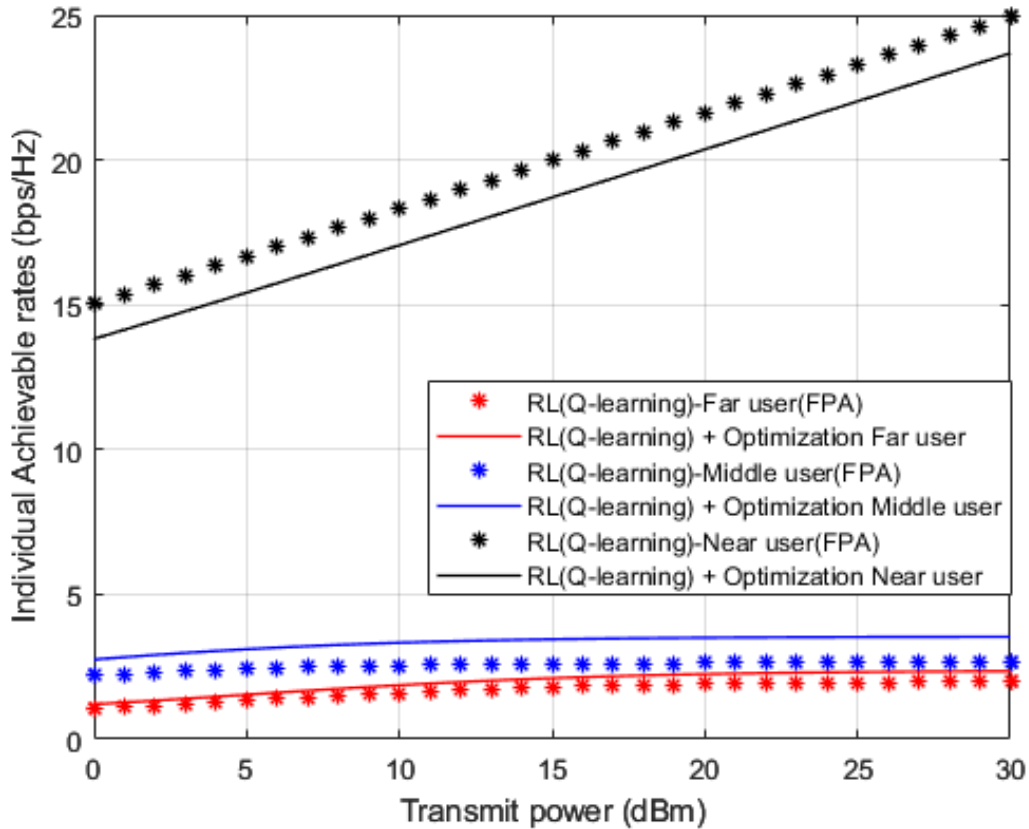
Figure 5.13 Individual rate vs. Power (*Q*-learning, Optimization, FPA).

5.7 Results Summary

In this section, we can present a sort of summary for the simulation results in terms of a sample statistics to emphasize the average percentage improvement attained in different performance metrics when RL based Q algorithm is applied as a channel estimator. Table 5.2 contains a sample numerical values and these values represent the estimated amount of improvement achieved by RL based Q algorithm compared to conventional NOMA scheme when different performance metrics such as BER, outage probability, sum rate, and the individual capacity are examined. Sample power values 5dBm, and 20dBm are chosen to record the percentage improvements for each metric in Rayleigh and Rician fading channel. Table 5.2 clearly indicates that RL based Q algorithm is providing an adequate improvement for different performance metrics in Rayleigh and Rician channel, and the most substantial enhancement is shown for sum rate metric. Table 5.3 is providing the comparison between RL based Q algorithm and RL based SARSA algorithm when both methods are implemented as channel estimator for the users in Rayleigh channel. When BER metric is inspected, the recorded values

reflect the improvement in Q algorithm when far middle users are inspected, while for near user case the SARSA algorithm verifies some dominance of Q algorithm. Overall, the Q algorithm recodes some enhancement over SARSA procedure when outage probability metric and sum rate metric are considered. This little enhancement was justified that the framework functionality of each algorithm is nearly similar.

The statistical values that facilitate the comparison between the Optimized power scheme and the fixed power allocation (FPA) scheme is shown in Table 5.4 when the RL based Q algorithm is utilized for the channel estimation scenario. As perceived from recorded values, the optimized power scenario is indicating a visible improvement over the FPA scheme when far and middle users are examined. In contrast, performance metrics values related to the FPA scheme are pointing to the dominance of the FPA scheme over the optimized power scheme when the near user is measured. Also, it is worth mentioning that both the optimized power and FPA schemes are recording a slight enhancement when the individual capacity metric is examined compared to other performance metrics.

Table 5.2 Sample statistics for average percentage improvement (RL- Q learning vs Conventional NOMA).

Performance metric	5 dBm	20 dBm	Criteria Improvement	Users	Channel
BER	≈ 39 %	≈ 37 %	RL- Q learning vs Conventional NOMA	All	Rayleigh
Outage Prob.	≈ 32 %	≈ 30 %	RL- Q learning vs Conventional NOMA	All	Rayleigh
Sum Rate	≈ 115 %	≈ 65 %	RL- Q learning vs Conventional NOMA	All	Rayleigh
BER	≈ 24 %	≈ 28.3 %	RL- Q learning vs Conventional NOMA	All	Rician

Table 5.3 Sample statistics for average percentage improvement (RL- Q learning vs RL- SARSA).

Performance metric	5 dBm	20 dBm	Criteria Improvement	Users	Channel
BER	≈ 26.5 %	≈ 24.5 %	RL- Q learning vs RL- SARSA	Far & Middle	Rayleigh
BER	≈ 14 %	≈ 22 %	RL- SARSA vs RL- Q learning	Near	Rayleigh
Outage Prob.	≈ 1 %	≈ 16 %	RL- Q learning vs RL- SARSA	All	Rayleigh
Sum Rate	≈ 2 %	≈ 2 %	RL- Q learning vs RL- SARSA	All	Rayleigh

Table 5.4 Sample statistics for average percentage improvement (Optimization vs FPA).

Performance metric	5 dBm	20 dBm	Criteria Improvement	Users	Channel
BER	≈ 32 %	≈ 50 %	Optimization vs FPA (RL- Q learning)	Far & Middle	Rayleigh
BER	≈ 40 %	≈ 33.3 %	FPA vs Optimization (RL- Q learning)	Near	Rayleigh
Outage Prob.	≈ 33.5 %	≈ 38.3 %	Optimization vs FPA (RL- Q learning)	Far & Middle	Rayleigh
Outage Prob.	≈ 65 %	≈ 69 %	FPA vs Optimization (RL- Q learning)	Near	Rayleigh
Individual Capacity	≈ 7.5 %	≈ 12.5 %	Optimization vs FPA (RL- Q learning)	Far & Middle	Rayleigh
Individual Capacity	≈ 11 %	≈ 11 %	FPA vs Optimization (RL- Q learning)	Near	Rayleigh

5.8 Summary

In this study, the influence of adopting a developed RL based Q-learning algorithm to distinctly predict the channel parameters for every user device in the MISO-NOMA system is analysed. In the developed Q-learning algorithm, the Q-model is created on the basis of the initialized channel statistics then updated based on the interaction between the Q-agent and the environment to maximize the received downlink sum rate and minimize the estimation loss. The efficacy of the developed Q-learning procedure is investigated by inspecting the performance of the proposed algorithm against different benchmark channel estimation schemes. The first benchmark scheme is based on standard MMSE procedure, the second approach is applying DL based LSTM network, the third scheme is implementing RL based actor-critic algorithm, and the fourth benchmark scheme is using RL based SARSA algorithm. In addition, the reliability of the proposed Q-learning procedure is validated by analysing the behaviour of the developed Q-learning algorithm in different fading channels.

Chapter 6

Deep Reinforcement Learning Framework in NOMA System

6.1 Introduction

It can be noticed that the high energy consumption by connected terminals in wireless networks can create an essential challenge in designing the upcoming 6G wireless networks, that are planned to enable an extensive range of important applications but with additional energy consuming [116]. Hence, it is important to consider the energy consumption issue in the future wireless communication networks while preserving the required quality of service (QoS). In this context, it is valuable to indicate that the utilization of superposition coding in NOMA can contribute to the energy efficient wireless transmission scheme. Thus, with the aim to ensure a desired quality of service (QoS) levels for all superimposed user devices, numerous research efforts have been dedicated to find out an efficient design of power allocation mechanisms. To that end, a resource allocation problem is analysed in NOMA system, and the channel estimation task is also investigated [117].

Furthermore, many authors recently have investigated the machine learning and artificial intelligence tools to optimize the resource allocation problems in NOMA system [118]. Recently, reinforcement learning (RL) based Q-learning algorithm and deep reinforcement learning based Q network (DQN) have achieved a remarkable interest by authors `in various fields. Q-learning algorithm is a subclass of reinforcement learning that depends on Q tables to store the optimal Q-values for state and action pairs, in order to maximize the future reward.

6.2 Related Works

In the context of optimizing communication networks, several works have employed Q-learning algorithm to enhance the performance of wireless systems based on different perspectives. The work in [119] used the Q-learning algorithm to introduce a framework for

enabling mobile edge computing in NOMA system. In [120], authors suggested dynamic reinforcement learning scheme-based power allocation to maximize the sum rate and the spectral efficacy in MIMO-NOMA system when smart jamming is considered. Authors have employed the Q-learning algorithm to allocate a power to user terminals to mitigate the jamming effects.

By incorporating deep learning into RL, deep reinforcement learning (DRL) can address the challenges associated with Q-learning in terms of Q-table storage. Based on that, the work in [121] used a deep Q-network (DQN) to model a multiuser NOMA offloading problem, while the work in [122], proposed a power allocation procedure in systems using deep reinforcement learning. In addition, authors in [123] have introduced the actor-critic based RL algorithm to handle a dynamic power allocation mechanism. Likewise, RL based actor-critic procedure was also applied in [124] to attain the optimal policy for user scheduling and resource allocation in HetNets. Authors designed the actor network to decide the parameters of the policy network to decide a stochastic action based on gaussian distribution, while the critic network evaluates the value function and assists the actor network to discover or learn the optimal policy.

Deep reinforcement learning was also introduced in [125] to arrive at sub-optimal power allocation for an uplink multicarrier NOMA cell. The work in [126], has discussed a joint channel assignment and power distribution procedure for NOMA system using deep reinforcement learning. Authors have derived a near-optimal power allocation scheme by considering two users per channel and the channel assignment has been performed using DRL algorithm to boost the overall sum rate while minimum capacity for user terminals is considered.

6.3 System Model

In NOMA cell, numerous user devices can be served via the same resource block (RB) based on employing the power domain in both uplink and downlink transmissions. In this chapter, we are considering a downlink NOMA cell, where the BS can serve distinct types of users or devices at same time and via different fading channels. The BS can assign one channel to every set of user devices, and the signals of the that user devices can be multiplexed using unique power levels. Then, user devices will receive their desired signals beside the signals of the other devices in same channel as interference or noise. The undesirable received signals will

be considered as noise if the power level of the desired signal is high, otherwise, the additional signals will be regarded as interference. To decode the required signal, each user device will use the successive interference cancelation (SIC) procedure. SIC technique will decode the signal with the highest power and then subtracts that signal from the principal signal until the desired signal is decoded.

Typically, SIC technique depends on the channel state information (CSI) [127] for each user device such as signal to interference plus noise ratio (SINR). In fact, the SINR itself may include the channel gain $|h_i|^2$ and power allocation α_i, where h_i is the fading channel between BS and user device i, and α_i is the power level allocated to user i . In NOMA scenario, the data rate R_i for user device i can be expressed as follows:

$$R_i = \log_2\left(1 + \frac{P_T \alpha_i \eta_i}{\sum_{j=1}^{i-1} P_T \alpha_j \eta_i + 1}\right) \tag{6.1}$$

where η_i is the channel to noise ratio (CNR) for user i and P_T is the total power transferred by the BS. The channel to noise ratio η_i for user i, can be expressed as

$$\eta_i = \frac{|h_i|^2}{\sigma_n^2} \tag{6.2}$$

As mentioned above, h_i is the fading channel between BS and user device i, while σ_n^2 is the noise power. In this chapter, we are considering a downlink of NOMA system, and the total number of devices in the cell is N, and for a practical wireless environment, we can assume that the channel parameters are not perfectly known. In the NOMA cell, the signals related to the N user devices are combined, and the BS will transmit this composed signal. The composed signal X can be represented as follows [128]:

$$X = \sum_{i=1}^{N} \sqrt{P_T \alpha_i}\, x_i \qquad i = 1,2, \dots, N \tag{6.3}$$

where x_i is the modulated desired signal for user device i. The resultant transmitted signal X will be received at the receiver side of each user terminal, with path loss and additive white Gaussian noise (AWGN), hence the received signal Y can be represented as

$$Y = \sum_{i=1}^{N} \sqrt{P_T \alpha_i}\, h_i x_i + n \qquad i = 1,2, \dots, N \tag{6.4}$$

where h_i is the fading channel between BS and user device i and n denotes the AWGN component. After receiving the composed signal, the receiver at each user device will activate the SIC procedure to decode its signal and perfect SIC depends on the availability of SINR of that device. The SINR itself varies according to the channel gain and power allocation for each user device. In PD-NOMA, a distinct power levels will be given to user terminals in the cell, and the highest power will be given to the user device with the lowest CNR and contrariwise. Therefore, for user devices having CNR's as follows:

$$\eta_1 > \eta_2 > \ldots > \eta_N \tag{6.5}$$

These user devices will be given power levels as follows:

$$P_1 < P_2 < \ldots < P_N \tag{6.6}$$

The SINR for user device i can be represented as shown [127][128]

$$SINR_i = \frac{P_T \alpha_i \eta_i}{\sum_{j=1}^{i-1} P_T \alpha_j \eta_i + 1} \qquad i = 1,2, \ldots, N \tag{6.7}$$

Basically, to implement the SIC, the BS can allocate power P_i to any user terminal as shown in the following expression:

$$P_i = \left(P_T - \left(\sum_{j=1}^{i-1} P_T \alpha_j \right) \right) \geq P_{th} \tag{6.8}$$

The expression in (6.8), can be interpreted as follows: for proper achievement for the SIC technique, the user device with low CNR must have a higher power than the sum of power levels for other devices that have high CNR.

6.4 Channel Estimation in Multiuser Environment

In this section, the channel estimation task will be analysed based on a distributed user devices and single base station (BS) in a NOMA cell. The BS is supplied with single antenna and each user terminal is also equipped with a one antenna. The network is assumed to work with a specific length time slots and each time slot allows for a one frame transmission, either uplink or downlink transmission. The pilot assisted channel prediction is considered in this chapter,

where pilot symbols can be recognized by BS and user devices [10][11][105]. Initially, each user terminal can transmit its pilot symbols to BS via the uplink channel. Then, prior to downlink data transmission, BS can examine or inspect both the pilot symbols and the available network information to facilitate estimating the channel features and parameters.

The principal aim of this chapter is to use deep Q-network (DQN) to realize the channel estimation task. The matrix H, that includes the downlink channel coefficients vectors that describe the channel properties from BS to each user device in the NOMA cell can simply be shown as follows:

$$H = [h_1; \ h_2; \dots; h_N] \tag{6.9}$$

where h_i is the vector that represents the channel coefficients from BS to user device i, where $i \in [1, 2, \dots, N]$, and N is the number of user terminals in the examined NOMA cell. Furthermore, we can denote the data signal transmitted to each user i as

$$s_i = [s_{i1}, s_{i2}, \dots, s_{iK}] \tag{6.10}$$

where K is the length of the data signal, and the matrix that contains the desired information transmitted for all the user devices can be expressed as follows:

$$S = [s_1; s_2; \dots; s_N] \tag{6.11}$$

Many of the current research depend on pilot symbols to estimate the uplink channel parameters and then utilize the channel reciprocity to realize the prediction of downlink channel parameters [11][105]. Using channel reciprocity for channel prediction may not be a reliable scheme, especially in cases of imperfect channel reciprocity owing to hardware impairments [105]. Hardware impairments are mainly introduced by phase noise of the radio frequency (RF) components that negatively affects the system performance. Hardware impairments have been studied and can be modelled as an additive distortion noise or nonlinear multiplicative factor [128]. Furthermore, using channel reciprocity for channel prediction may introduce estimation errors when the uplink and downlink channel parameters are not stationary within a particular transmission slot.

In this chapter, the deep reinforcement algorithm based deep Q-network (DQN) [129] will be

developed to explicitly predict the downlink channel parameters for every user device in the NOMA cell based on the support given from the network information and the pilot symbols. The matrix $\hat{H}$, that represents the estimated channel parameters vectors between BS and N users, can be shown as follows:

$$\hat{H} = [\hat{h}_1; \hat{h}_2; \dots; \hat{h}_N] \tag{6.12}$$

where $\hat{h}_i$ represents the vector of generated channel parameters between BS and user device i.

6.5 Deep Reinforcement Learning Basic Concept

In this section, we will introduce the concept of deep reinforcement learning (DRL), which is a special case of reinforcement learning procedure [130]. Reinforcement learning is a fork of machine learning, where an agent interacts with the environment to carry out the best sequences of actions that can maximize the expected future reward in an interactive environment. Generally, reinforcement learning (RL) can be classified as single-agent or multi-agent based on the quantity of agents in the environment. In the scenario of a single agent RL, the 1) agent can recognize the entire states in the environment and the 2) decision-making task can be modelled using Markov decision process (MDP) framework. In this chapter, our proposed structure assumes a single agent, and the best sequence of actions that can be chosen by an agent will be determined based on the deep neural network (DNN), therefore, the combination between DNN and reinforcement learning (RL) can exemplify the universal function approximator.

The fundamental elements for deep reinforcement learning (DRL) can be listed as follows [11][129][130]:

- Observations: the continuous measurements of the properties of the environment, and all of the observed properties in the environment can be included in the state space S.

- States: the discretized observations at time step t can be denoted as state $s_t \in S$.

- Actions: an action a_t is one of the valid decisions that the agent can select at time step t from the action space A.

- Policy: a policy denoted by $\pi(.)$, is the criteria that control how to select a certain action in the environment at any given state.

- Rewards: the immediate reward r_t, is obtained after an agent carries out a specific action a_t in a given state s_t, which leads to move to new state s_{t+1} .

- State-action value function: denoted by $Q_\pi(s, a)$, and represents the expected discounted reward when the agent starts at state s_t and selects a certain action a_t based on the policy π.

When an agent selects an action a_t at a given time step t, then the agent's state will change from the current state s_t to the subsequent state s_{t+1} and as a result of this transition, the agent will receive an immediate reward r_t from the environment. Based on that scenario, the network can generate an experience tuple $e = (s_t, a_t, r_t, s_{t+1})$ that can be stored in experience replay buffer $\mathcal{D}$. Hence, the primary target of the agent is to maximize the long-term cumulative discounted reward R_t^γ, which is defined as follows [11][105][130]:

$$R_t^\gamma = \sum_{i=0}^{\infty} \gamma^i \, r_{t+i} \tag{6.13}$$

where discount factor $\gamma \in [0:1]$. To satisfy R_t^γ, an optimal policy π^* is essential to map the best actions to states. In other words, the optimal policy π^* can assist the agent to decide which action should be taken at any given state, to enhance the long-term cumulative reward. Typically, the state action Q-value function is defined as the expectation of the cumulative reward R_t^γ.

Based on the following: starting state s_t, the considered policy π, and the selected action a_t, the Q value function can be expressed as follows [11][105]:

$$Q_\pi(s_t, a_t) = \mathbb{E}[R_t^\gamma | s_t, a_t] = \mathbb{E}\left[\sum_{i=0}^{\infty} \gamma^i \, r_{t+i} | s_t, a_t\right]$$

$$= [r_t + \gamma Q_\pi(s_{t+1}, a_{t+1}) | s_t, a_t] \tag{6.14}$$

where $\mathbb{E}[\,.\,]$ denotes the expectation parameter. When the optimal policy π^* is adopted for maximizing for all states and actions pairs, then the optimal Q value function Q_{π^*} that follows the optimal policy π^* can be expressed as follows:

$$Q_{\pi^*}(s_t, a_t) = [r_t + \gamma Q_{\pi^*}(s_{t+1}, a_{t+1}) | s_t, a_t] \tag{6.15}$$

The expression in (6.15) is known as the Bellman equation. The benefit of Bellman equation is to represent the state-action Q value function into two components: the instantaneous reward r_t and the long-term discounted reward. However, the Bellman equation is nonlinear, and hence, there is no closed form solution to it. As a result, an iterative procedure such as the Q-learning algorithm has been emerged to converge the Bellman equation to obtain the optimal Q value function [11][105]. On the other hand, the computation of the Q-learning algorithm will become complex in multi-user environments that have a huge state and action spaces, and as a result the size of the Q-table will be extremely large. Hence, the regular solution to this problem is to estimate the Q-values using a function approximation, by adopting deep neural networks (DNN), which is the core component in our proposed deep Q-network (DQN) [11][105][130]. Basically, the basic DQN structure, shown in Figure 6.1, consists of three main phases: the first phase represents the input layer that shows the states of the environment. The second phase is the hidden layers that act as a function approximator. Mainly in the hidden layer, the Rectified Linear Unit (ReLU) activation function can be applied to compute the hidden layer values. The primary gain of utilizing ReLU as an activation function is the computational efficiency [131], that may lead to faster convergence. At the end stage, the output layer is responsible to predict optimal state-action value function, $Q_{\pi^*}(s, a, W_t)$, where W_t is the weights at time instant t.
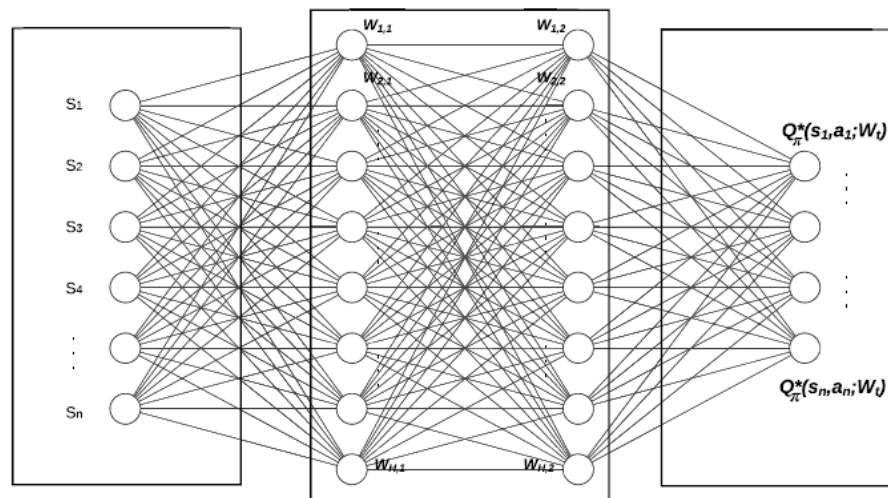
Figure 6.1 DQN basic structure with two hidden layers.

6.6 DQN Training phase

In this section, we will discuss how the proposed DRL based DQN network is trained, and Figure 6.2 illustrates the structure of the developed DQN algorithm that mainly relies on the LSTM network to achieve the most appropriate performance. The DQN network will be trained, and the weights of the hidden layers will be updated to approximate the state-action value function $Q_\pi(s, a)$. As indicated in the aforementioned discussion, each experience tuple is described as $e_t = (s_t, a_t, r_t, s_{t+1})$, and the agent will store all experience tuples in an experience replay buffer $\mathcal{D} = \{e_1 \ e_2 \ e_3 \ ... \ e_t\}$, and these learned experience tuples can be utilized to train the DQN [46][132]. It is optimal for the DQN algorithm to exploit all experiences tuples in each training iteration, but this will be costly when the training set is huge. A more effective procedure is to update the DQN network weights in each iteration using an arbitrary subset from the replay buffer $\mathcal{D}$, and this subset is described as mini batch. Hence, the loss function can be defined as follows [11][129][130]:

$$\mathcal{L}(W) = \sum_{e \in \mathcal{D}} \left(r_t + \gamma \ maxQ_{\pi^*}(s_{t+1}, a_{t+1}, \widehat{W}) - Q_{\pi^*}(s_t, a_t, W) \right)^2 \quad (6.16)$$

where $\mathcal{L}(W)$ denotes the DQN loss function for a random mini batch sampled from the replay buffer $\mathcal{D}$ at time slot t and $\widehat{W}$ represents the nearly static target weights that are updated every T time steps. With the purpose of minimizing the loss function $\mathcal{L}(W)$, the weights W of the policy DQN will be updated at every time step t using a stochastic gradient descent (SGD) algorithm on a mini batch sampled from the replay buffer $\mathcal{D}$. Also, the SGD algorithm can update the DNN weights W in an iterative process with a learning rate of $\mu > 0$ as follows:

$$W_{t+1} = W_t - \mu \ \nabla \mathcal{L}_t(W_t) \quad (6.17)$$

6.7 DQN Based LSTM Network

Long short-term memory (LSTM) network is a developed design from recurrent neural network (RNN), which can 1) inspect long term dependencies and has 2) the ability to remember previous information for future usage. The 3) LSTM network has a chain structure consists of multiple LSTM cells [89][90] and the proposed DQN structure shown in Figure 6.2 is clearly adopting the LSTM network as the DNN hidden layers. The DNN based LSTM is mainly consists of four layers, and each layer contains several neurons, where the weighted sum of each neuron will be the input to an activation function. In our proposed DQN approach,

the length of each training sequence is specified as L, which is the dimension of the input layer. In our scenario, we choose the input layer to include 128 neurons (same number of input states), In the second layer, we have applied one LSTM layer that includes 300 hidden cells.
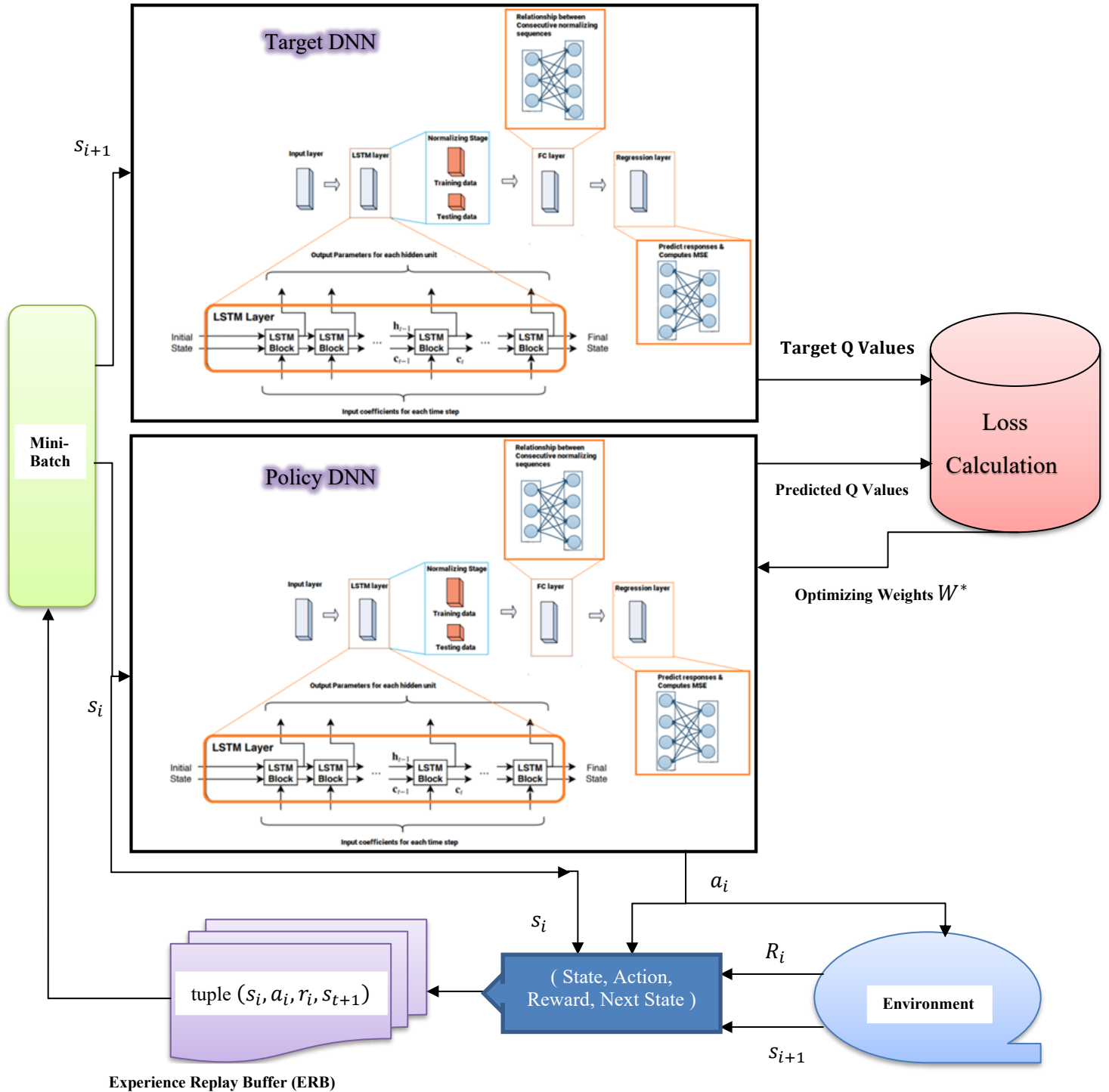
Figure 6.2 Proposed DQN Architecture.

For each hidden cell the learnable weights are specified as follows [10][90]: the input weights *W*, the recurrent weights *R*, and the bias *b*. The third layer is a fully connected layer that processes the outputs of the LSTM layer, and it assembles all of the characteristics and internal information that gathered by the prior layers. Also, the fully connected layer multiplies the input by a weight matrix *W* and then adds a bias vector *b*. In addition, the fully connected layer behaves separately at each time step, and all neurons in a fully connected layer are connected to all the neurons in the previous layer.

The last adopted layer in the DNN in the DQN structure, is the regression layer, which is accountable to compute the mean square error (MSE), improve the cell status, and update the cell weights. A regression layer can also predict response of the trained network. It is worth mentioning that normalizing the training data in LSTM network enable stabilizing and accelerating training process for neural networks. In the proposed DQN structure, it is shown in Figure 6.2, that the input states are established according to the size of the input layer, then these states will be passed through the DNN policy network to generate the state action value functions at the output of the policy DNN.

The proposed DQN Architecture in Figure 6.2, can be simplified as a sequence of blocks or stages as shown in Figure 6.3, and these blocks can be briefly explained as follows:

Stage 1: Initialization and generation of training data

1. Perform a few random actions with the environment to initialize the experience replay data.

2. Initialize the weights for the policy DNN and copy these weights to the target DNN.

3. Starting with the first-time step, do the following:

    - Based on the initial interaction with the environment, random states can be generated to be used as input for the policy DNN.

    - The policy DNN will predict the Q-values for all actions that can be decided in the current state, and then those Q-values will be inspected to select or identify a certain Q-value based on the most suitable action.

- Based on the selected and executed action, the experience replay will receive the reward and move to the next state.

- The experience replay will store these results in the replay buffer.

- Each of these results will be considered as sample training data, that can be later used as training data.

Stage 2: Select a random batch for training.

- Select a batch of random samples from the replay buffer and use these selected samples as an input for both the policy DNN and the target DNN.

- From the random sample, use the current state as input to the policy DNN.

- The policy DNN can predict the Q-values for all actions that can be performed in the current state.

- Based on the decided or selected action, the policy DNN will identify the predicted Q-value.

- From the selected random sample, the next state will be used as input to the target DNN.

- The target network will predict the Q-values for all actions that can be performed in the next state, then the target DNN will select the maximum of those Q-values.

Stage 3: Get the Target Q-value.

1- The Target Q-value can be decided based on two components.

- The immediate reward from the environment

- The max Q value that has been predicted by the target DNN in the next state.

Stage 4: Compute the Loss function.

1. Compute the loss function between the Target Q value and the predicted Policy Q value in terms of mean squared error (MSE) as shown in equation (6.16).

Stage 5: Back-propagate the Loss function.

    1- Back-propagate the loss in order to update the weights of the policy DNN using SGD.

    2- At this stage, the weights of the Target DNN are not updated and remain fixed, and this completes the processing for this time step.

Stage 6: Repeat for next time step

    1. The process will be repeated for the next time step.

       • The policy DNN weights will be updated but not the Target DNN.

       • This allows the policy DNN to learn to predict more accurate Q-values, while the weights for the target DNN remain fixed for a while.

Stage 7: After T time steps, copy the policy DNN weights to the target DNN. This step will enable the target DNN to get the updated weights so that it can also predict more accurate target Q-values.
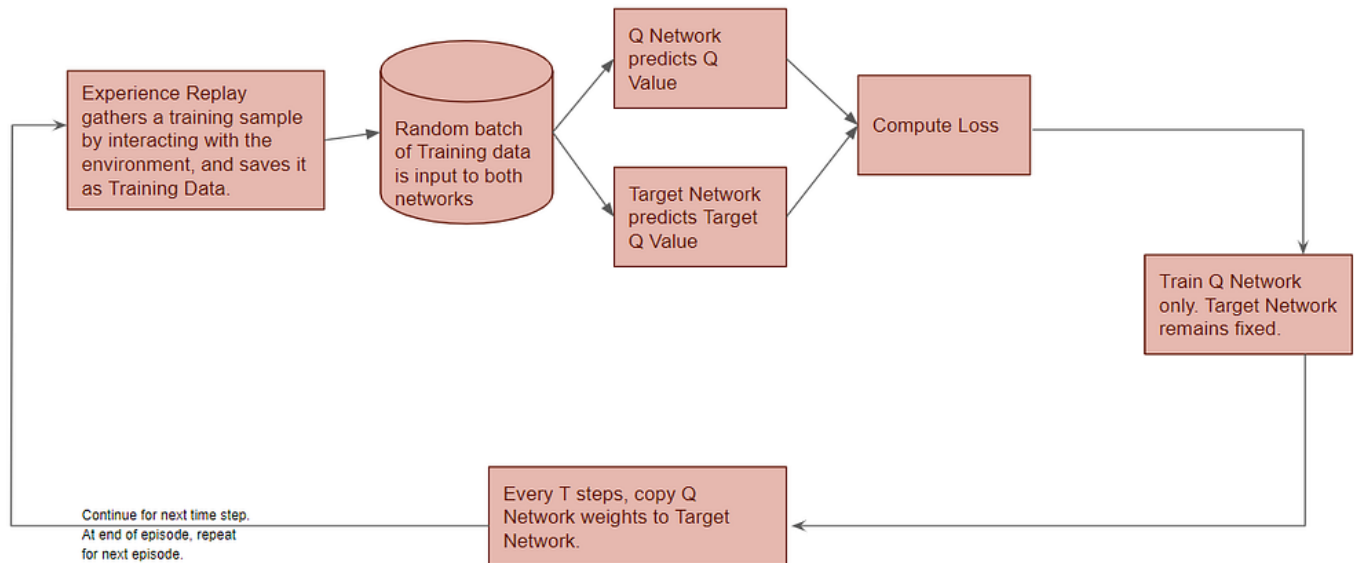
Figure 6.3 Basic DQN Architecture.

The design of a single LSTM cell is shown in Figure 6.4 [10][90]. Each LSTM cell has three inputs and two output parameters. The hidden state h_{t-1} and the cell state c_{t-1} are the shared

parameters between inputs and outputs and the other parameter is the current input. The LSTM cell also includes three sigmoid functions and two tanh functions to control the flow of information. In the initialization stage, a random hidden states will be generated along with the input for the first LSTM cell. Then the current outputs that include the hidden state h_t and cell state c_t and the new input x_t will comprise the three inputs to the next cell.
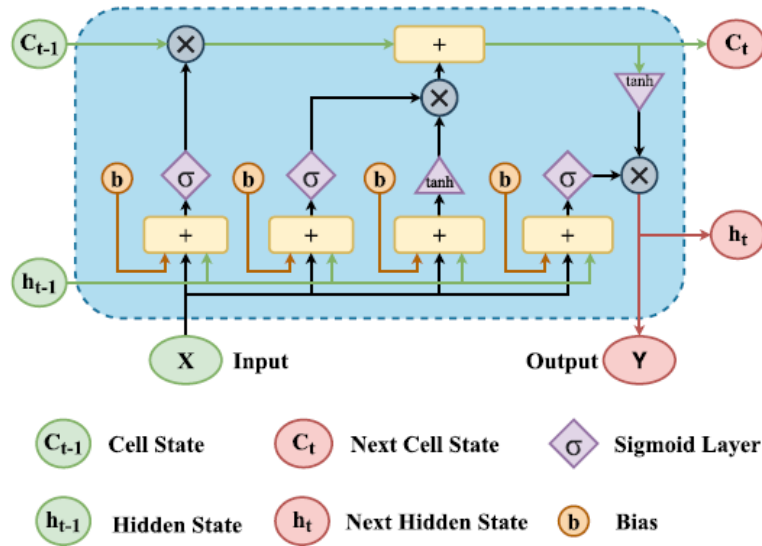
Figure 6.4 LSTM Cell Structure [10].

6.8 DQN Dataset Generation

Typically, the DQN scheme involves an agent, deep neural network (DNN), and an environment. The agent works together with the environment and decides which action to take. In our proposed scheme, the BS will be considered as an agent and interacts with the environment that involve the user devices and fading channels. At the start, the agent (BS) will start exploring the environment to collect the sample information for each user device such as power distribution, user distance, channel model, and path loss [11][130][133].

At each time step t and based on the current state s_t for each user device, the agent can take a decision on a certain action a_t using the DNN to maximize the sum rate for all users in NOMA network. Accordingly, the agent will receive an instant reward r_t and move to the next state s_{t+1} in the environment. By taking decisions on actions, the agent (BS) can learn more about the environment to achieve an optimal channel prediction policy π_c. This optimal policy π_c for predicting or estimating the channel parameters for each user device can be learned and

enhanced at each time step t via the target DNN as shown in Figure 6.2. Furthermore, the agent can adjust and enhance the policy π_c by repeating the channel estimation process for multiple episodes. Based on the aforementioned discussion, it is clearly noticed that in this chapter, the DNN replaces the Q-table to estimate the Q-values for each state-action pair in the environment, and this designed DNN can be considered as the policy controller for channel parameters estimation.

6.8.1 Near-Greedy Action Policy

The period of time slot in which the agent interacts with the environment is termed as episode, and every episode has a total duration time of T time steps. At each episode, the main aim is to maximize the sum rates for all users in the NOMA cell while minimizing the total loss. In our proposed DQN architecture, the dimension of the input layer for either the policy DNN or the target DNN is set to equal to the available states in state space S and correspondingly, the dimension of the output layer is equal to the number of possible actions in the action space A [10][11][90][134].

As stated before, we decide to choose four layers as a depth of the implemented DNN hidden layers, this may provide a reasonable balance between the network performance and computational complexity. Typically, the Q learning procedure is considered as an off-policy algorithm, which means that without applying any greedy policy, the Q algorithm can iteratively estimate the best action for future reward. In this chapter, and in our proposed DQN procedure, we decide to apply a near-greedy action selection policy, that has two approaches [113]:

1) Exploration: the agent discovers and carries out random actions at every time step t.

2) Exploitation: Based on the previous experience and the current network weights, the agent can decide an action to maximize the state-action value function $Q_\pi(s_t, a_t, W_t)$.

In our proposed near-greedy action policy, the agent has an exploration rate of ϵ and an exploitation rate of $(1 - \epsilon)$ where $0 < \epsilon < 1$, and ϵ is considered as a hyper-parameter that can control the trade-off between exploitation and exploration during the training process. Hence and based on that designated action selection policy, the agent can select a specific action a_t at a given state s_t at every time step t and correspondingly, the agent can receive a

positive or negative reward and moves to a new state s_{t+1}.

6.8.2 DQN Algorithm

In this subsection, we will discuss the channel estimation task based on DQN framework. In the channel estimation process, we need to predict the channel parameters of each user device in NOMA cell, and also, we need to maximize the sum rate for all users in the considered NOMA system at each time step t. The state space S contains the user distance d_i that represents the distance between base station (BS) and the user device i, the power allocation coefficient α_i for each user in the NOMA cell, and the path loss ϑ. Accordingly, the resultant state space can be represented as [10][113][134]

$$S = \{d_1 \ d_2 \ d_3 \dots d_N \ \alpha_1 \ \alpha_2 \ \alpha_3 \ \dots \alpha_N \dots \vartheta\} \tag{6.18}$$

All the actions that can be chosen by the agent are selected from the action space A. The possible actions in the action space A can be described as follows: change the distance of the user device, change the path loss exponent, and increase or decrease the power distribution factor α_i by a certain step size.

The reward function plays an essential role in the DQN algorithm, and there are many ways to assign the rewards based on the selected action. In our proposed scenario, we decide to calculate the rate for each user in NOMA system using (6.1), to represent the instant reward r_i returned from environment to the agent after choosing a certain action a_t at state s_t. Hence, based on the selected action, the higher data rate will reflect a good reward for the agent, while lower data rate will reflect a bad reward. Based on the proposed DQN framework, we can describe in algorithm 6.1, the algorithm steps for approximating the channel parameters for each user device in NOMA cell.

Algorithm 6.1 Proposed DQN Algorithm for channel estimation

---

1. Initialize policy DNN and target DNN networks with random weights $(W, \widehat{W})$.
2. Initialize experience replay memory (ERM).
3. Initialize exploration rate ϵ.
4. **for** each **episode do**
5.   **for** each instance (**step**) **do**
6.   **for** each **user** device **do**

7.　　　　based on ϵ, and based on the current state s_i , Select the channel parameters and add to action space a_i

8.　　　　**end for**

9.　Observe the immediate rewards r_i (based on data rate) and move to the next state s_{t+1}.

10. Insert (Store) (s_i, a_i, r_i, s_{t+1}) in ERM.

11. Create a mini batch with random sample of tuple (s_i, a_i, r_i, s_{t+1}) from ERM.

12. **for** each tuple in mini batch **do**

13.　　Obtain or generate Q-values (length 128 taps) using policy DNN (Q network).

14.　　 Approximate Q^* values (Target value) using target DNN. (Target Network)

15.　　 Calculate the loss between Q values from Policy DNN and Q^* values from Target DNN.

16.　　Optimize the parameters W of the policy DNN using SGD.

17. **end for**

18. **end for**

19. $\widehat{W} \leftarrow W$ after a certain number of steps.

20. **end for**

Detailed DQN Procedure and Workflow

Based on algorithm 6.1, we can list the detailed workflow for the developed DQN algorithm that is responsible for estimating the channel parameters for each user in the examined NOMA system:

- Initialize the weights for both the policy DNN and the target DNN.

- Initialize the ERM with a typical size of 10,000 (it can be 10^6).

- Initialize the ϵ parameter for near-greedy action selection policy with a large value of $\epsilon = 0.999$ (start by exploration then decay).

- Initialize data records (tuples).

  a) Generate a random channel coefficient based on the fading model parameters with size = 120).

  b) Based on the pilot symbols, approximate the channel coefficients with size = 8).

  c) For each user, both the randomly generated channel parameters and the channel coefficients estimated based on the pilot symbols will be combined and used as initial channel parameters with size 120.

139

- For each user, assign the initial distance, initial power factor, and path loss, and prepare the state space S.

- Select a random state s_t from the sate space and use it as an input for policy DNN.

- The policy DNN will select a random action and correspondingly select a random Q value, and the selected Q value will actually represent the channel coefficients for current state.

- Calculate the rate, and based on the calculated rate the reward can be assigned.

- Go to the next state s_{t+1}

- Compose a tuple $e_1 = (s_t, a_t, r, s_{t+1})$

- Store a tuple e_1 in ERM.

- Generate experience tuples = 1000 and store these tuples in ERM.

- Select a random batch of the tuples from ERM with batch size 32 tuples.

- Number of episodes = 20, and number of steps $T = 10^4$

- For each tuple in the random batch do the following:

   a) From the policy DNN, select the Q-values (channel coefficients) randomly.

   b) From the Target DNN select the Q-values based on the greedy policy

   c) Assign the Reward.

   d) Calculate the Loss function as follows:

      Loss= (Target Q value (Reward + Q_{max} value) - Policy Q value).

   e) Based on the Loss function, update the weights of the policy DNN using the gradient descent (GD) procedure.

- Every $T = 10^2$ steps, copy the weights of the policy DNN to the Target DNN.

- Activation functions used in LSTM layers are (sigmoid and tanh), while the activation functions used at the output layer are (linear or Relu).

- SGD optimizer is utilized for weight updates.

6.9 Simulation Parameters

Discussion for the simulation parameters and settings is described in this section. The simulated downlink NOMA system includes three distinct user devices and one BS in which the BS is equipped with single antenna and each user device in the cell is also equipped with a single antenna. In the examined NOMA environment, the user's modulated signals in the downlink transmission are superimposed and transmitted by the BS to each user device via an independent Rayleigh channel that are affected by AWGN at the receiver side, and the noise power density is allocated as $N_0 = -174$ dBm/Hz and the path loss is set to 3.

MATLAB simulation tool is employed to realize the following objectives, inspect, characterize, and measure the performance of the proposed deep reinforcement learning (DRL) based deep Q network (DQN) algorithm that will be utilized as a channel estimator in the examined NOMA system, explore the reliability of integrating the proposed DQN algorithm as channel estimator scheme with the optimized power scheme in the examined NOMA cell. Standard performance metrics will be utilized to evaluate the impact of adopting DQN algorithm as a channel estimator. Monte-Carlo simulations are accomplished with $N = 10^4$ iterations, and at the start of the iteration, pilot symbols are generated and recognized at the BS and each user. The main simulation parameters can be summarized in Table 6.1.

Table 6.1: Simulation parameters for DQN Framework.

Parameter	Value
Simulation Tool	MATLAB
Modulation type	QPSK
Number of Users	3, [2–10]
System Bandwidth B	1000 kHz
Fading channel	Rayleigh
Path loss exponent	3
Number of steps	10^4
Noise PSD N_0	-174 dBm/Hz
Learning Rate α	0.01
Discount factor γ	0.99

Episode	20
Optimizer	SGD
ERM size	10,000

6.10 Simulation Environment

The simulation figures are created based on the assumption that the channel parameters are not perfectly available at each user device. Therefore, in order to examine the impact of the proposed DRL based DQN procedure, the standard minimum mean square error (MMSE) based channel estimation technique [94] is also simulated for the sake of comparison. In the simulation environment, NOMA parameters are assigned, and channel parameters are initially generated to model the Rayleigh fading channel. In our developed DQN algorithm, at the end of each training episode, the predicted $Q(s, a)$ values generated from the policy DNN will be employed as a practical channel parameters for each user device to recover the desired signal for each user. Distinct power factors are initially assigned for every user device according to the current distance from the BS and the channel condition. Power factors η_n, η_m, and η_f are assigned for near, middle, and far users respectively. In a fixed power allocation setup, we assign $\eta_f = 0.65$, $\eta_m = 0.3$, and $\eta_n = 0.05$. In the simulation files, the transmission distance for every user device with respect to BS is initially defined as follows: $d_f = 1000$ m, $d_m = 500$ m, and $d_n = 100$ m. user's data and pilot symbols are modulated using Quadrature phase shift keying (QPSK) modulation format and the applied transmitted power is mostly varying from 0 to 30 dBm.

6.11 Results Discussion and Analysis

Simulation results that describe the comparison between the proposed DRL based DQN algorithm and the MMSE procedure when being utilized to estimate the channel parameters for each device are shown in Figure 6.5 in terms of BER versus power. The estimated channel parameters using both procedures will be employed for the signal recovery for each user and the simulated results are generated where fixed power allocation (FPA) scheme is considered. It is clearly noticed that when the developed DQN algorithm is applied for channel estimation, task, each user device in the examined NOMA cell shown the ability to provide a visible

enhancement in lowering the BER compared to the MMSE technique. As an example, at particular transmitted power 20 dBm, the realised BER value for far user using MMSE procedure is 10^{-1}, while the achieved BER in case of DQN is 10^{-2}. Similarly, the amelioration in the BER for middle and near users are obviously observed when the developed DQN algorithm is applied compared to MMSE scenario.
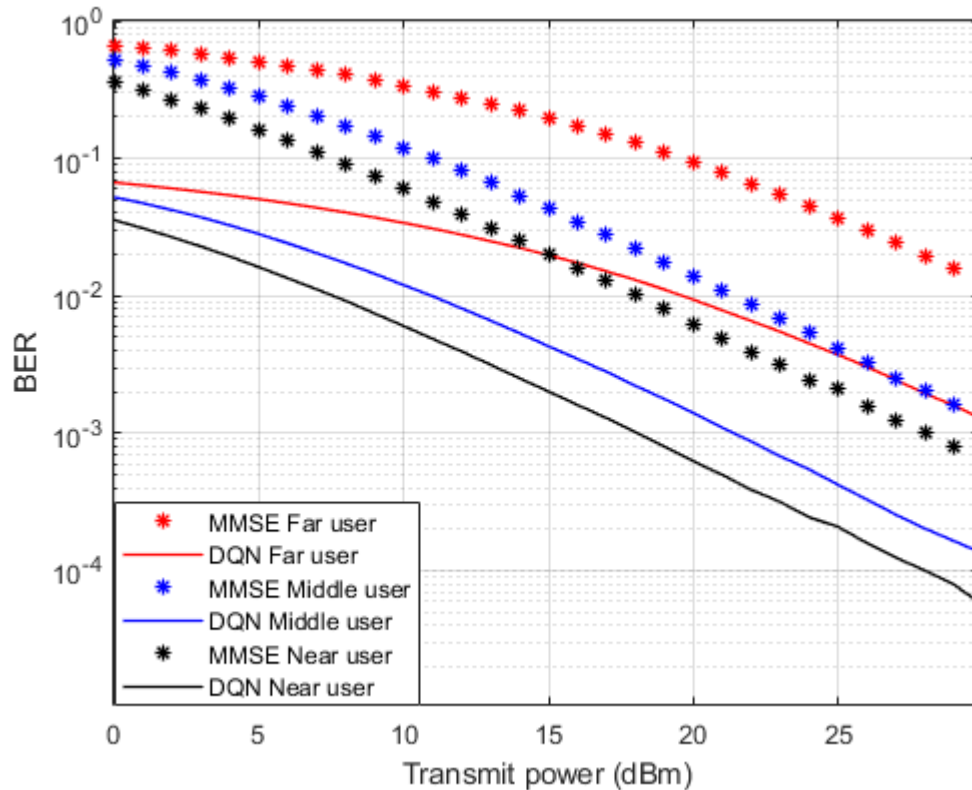
Figure 6.5 BER vs. power (DQN - MMSE).

In terms of the outage probability against applied power, Figure 6.6 illustrates the simulation results for the inspected user devices in the NOMA cell when both the developed DRL based DQN algorithm and the standard MMSE technique are implemented as a channel estimator. Similar to BER results, all user devices simulation outcomes indicate about 10dBm enhancement in the power saving when the developed DRL based DQN algorithm scenario is applied compared to the MMSE technique. The reduction in the power transmitted also supports the improvement noticed in minimizing the outage probability when DQN algorithm

is adopted. These results verify the advantage of the developed DQN as a channel estimator compared to the traditional MMSE technique.
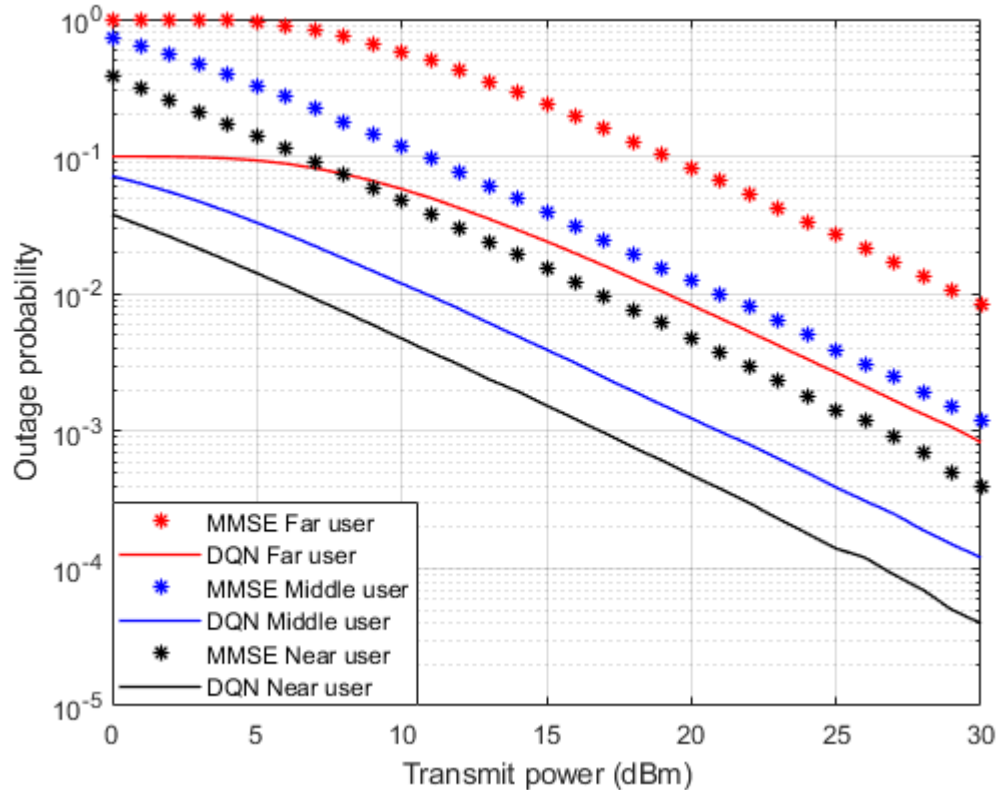
Figure 6.6 Outage Probability vs. power (DQN - MMSE).

Figure 6.7 shows simulation results for the attainable capacity for each device in the examined NOMA system when both the proposed DRL based DQN algorithm and the standard MMSE channel estimation procedures are applied. The achieved rate for the near device shows significant enhancement by more than 20 b/s/Hz over far and middle users' rates. The dominance of the near user in terms of the possible rate can be imagined, due to the stable channel condition for the near user plus SIC effect compared to other users in the system. Moreover, the proposed DQN algorithm still can deliver a stable capacity compared to the MMSE technique for far and middle users' scenarios, and this slight amelioration can be justified by the existing interference and inadequate link conditions for far and middle user devices.
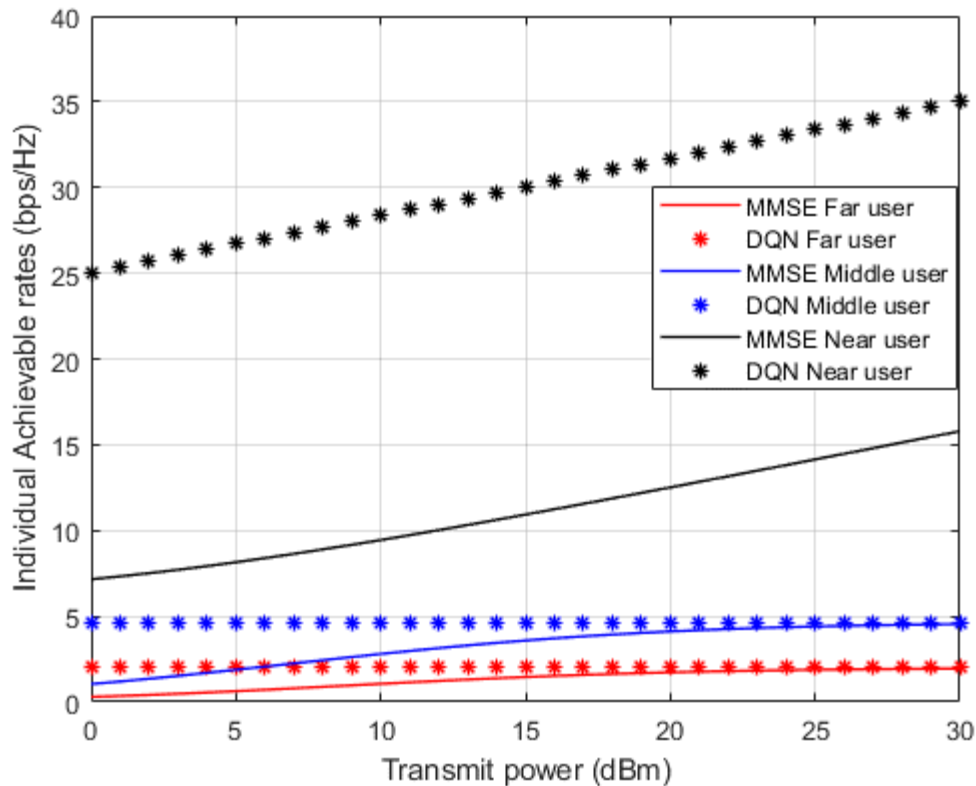
Figure 6.7 Capacity vs. power (DQN - MMSE).

In Figure 6.8, we establish three benchmarks for comparisons: (1) standard minimum mean square error (MMSE) procedure for channel estimation [94]; (2) DL based LSTM structure for channel prediction applied in [10]; and RL based Q algorithm for channel estimation applied in [11]. Figure 6.8 displays the simulation outcomes for the sum rate for all user devices in the NOMA cell versus applied power. It is evidently apparent that the developed DRL based DQN algorithm shows superiority over the standard MMSE procedure by more than 20 b/s/Hz approximately. Furthermore, the developed DQN algorithm performs an improvement over the DL based LSTM procedure presented in [10] by nearly 10 b/s/Hz. For the third benchmark in [11], The proposed DQN procedure, performs an amelioration over the RL based Q algorithm by 8 b/s/Hz approximately. These findings can support that the developed DQN algorithm can be a strong candidate technique compared to other procedures when it is being utilized as a channel estimator.
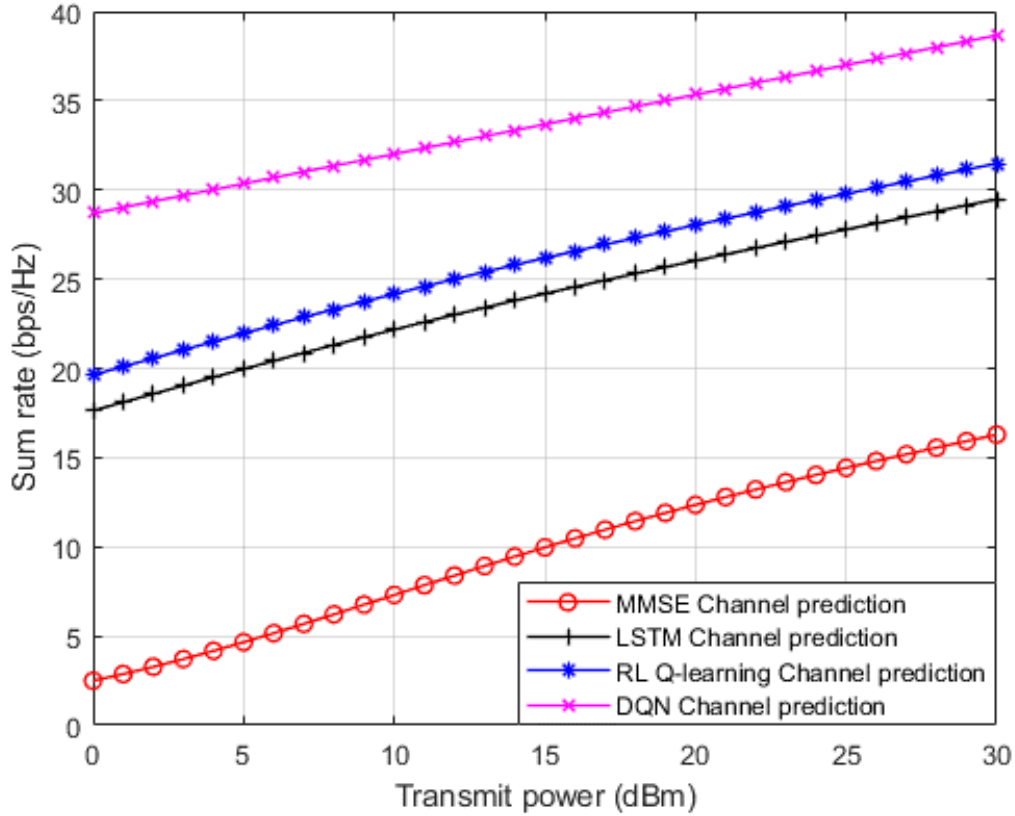
Figure 6.8 Sum rate vs. power (MMSE, LSTM, RL Q-learning, DQN).

Simulation results for the sum rate performance metric versus different number of users in the examined NOMA cell are shown in Figure 6.9, where the reference power is assigned to be 1 dBm. Similar to the simulation environment in Figure 6.8, three distinct channel prediction schemes are also investigated here as a benchmark comparison: (1) standard minimum mean square error (MMSE) procedure for channel estimation [94]; (2) DNN based LSTM structure for channel prediction applied in [10]; and RL based Q algorithm for channel estimation applied in [11]. As revealed from the results in Figure 6.9, it is clearly noticed that our developed DQN algorithm can realize a substantial greater sum rate with respect to the MMSE procedure, by at least 4 b/s/Hz when the cell capacity is initialized with 2 users. It can also be noticed that as the number of user devices in the cell is keep increasing, the developed DQN algorithm still shows dominance in accomplishing higher sum rates compared to DNN based LSTM procedure by 2 b/s/Hz approximately. likewise, the hidden layers feature in the DQN procedure is showing a noticeable enhancement in the sum rates compared to the Q-learning

procedure when the number of user devices in NOMA cell is increasing. Overall, these findings reveal that reliability can be ensured by our suggested DQN algorithm even when the user devices in the cell increased. Furthermore, it is worth saying that while increasing the user devices in the cell, the interference will also grow up, thus the performance and the sum rate could be affected.
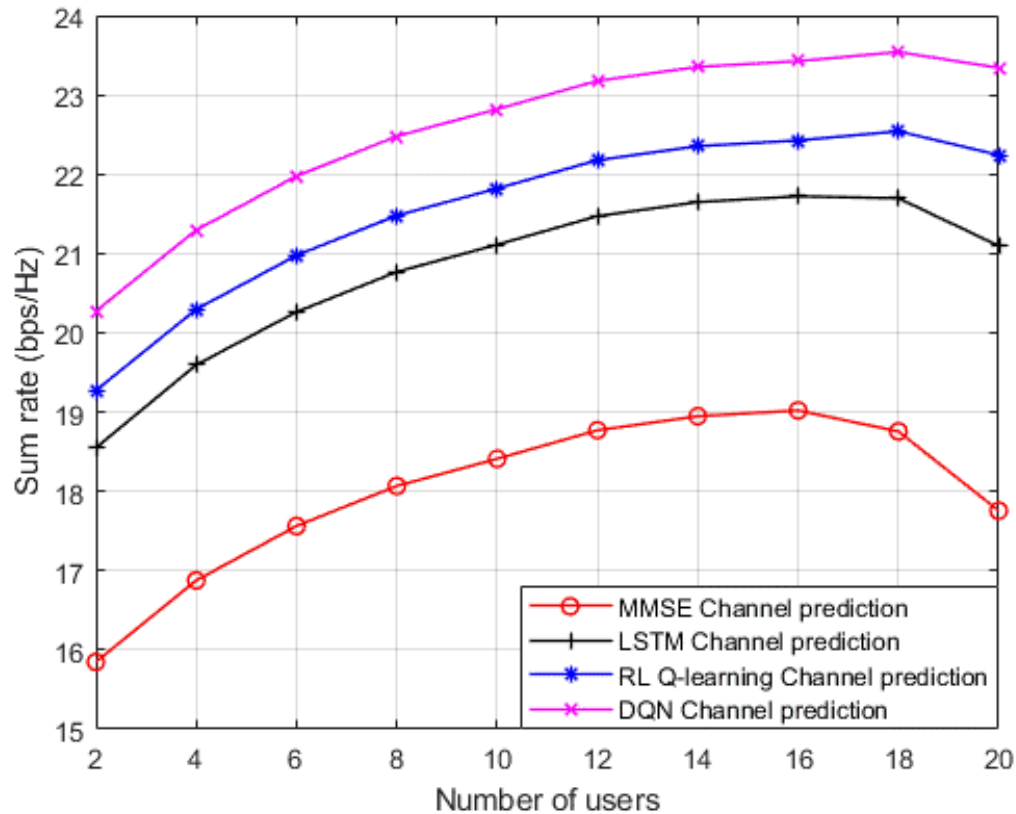
Figure 6.9 Sum rate vs. number of users (MMSE, LSTM, RL Q-learning, DQN).

In Figure 6.10, simulation results in terms of BER versus power are shown where both the proposed DRL based DQN technique, and the RL based Q-learning algorithm [11] are utilized as a different approaches for channel estimation. In addition, the optimized power coefficients that previously derived in chapter 3 for the examined NOMA cell are also applied in this simulation environment. Simulation results, indicate that all user devices in the cell can provide a perceivable enhancement in the performance when the developed DQN algorithm is applied as a channel estimator scheme compared to the case when the Q learning algorithm is adopted. Furthermore, and based on the simulation outcomes, it can be clearly noticed that when both

the optimized power scheme and DQN algorithm for channel estimation are combined, the achieved power saving is recorded at least to 5 dBm compared to Q algorithm procedure scenario.
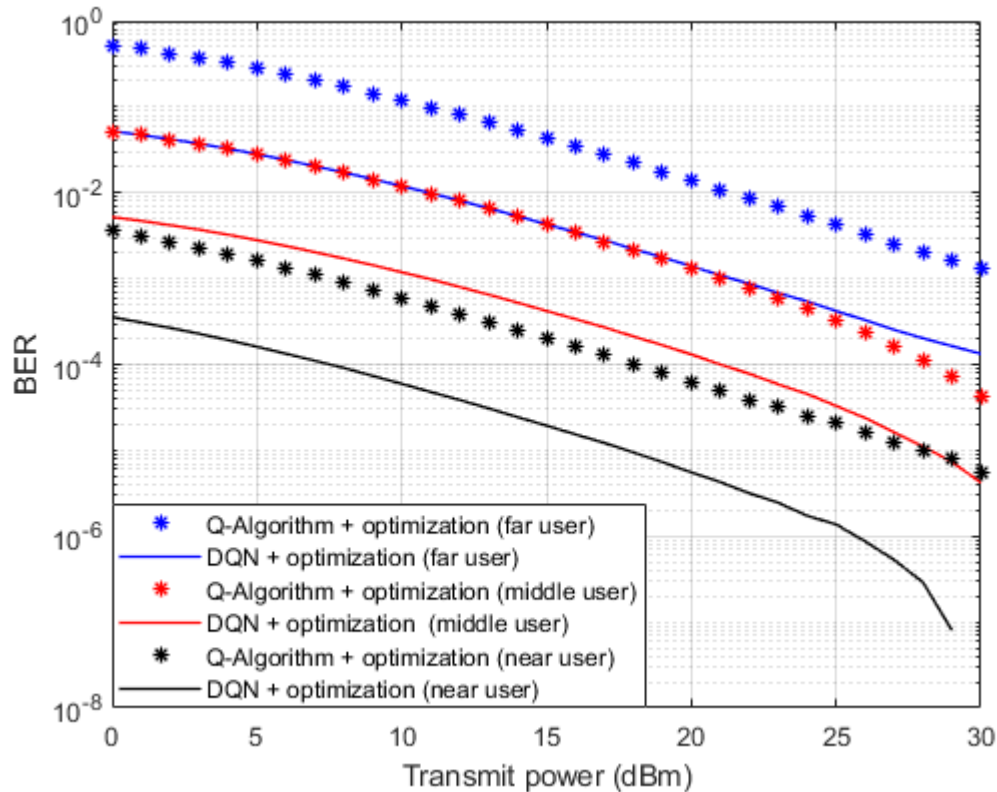
Figure 6.10 BER vs. power (DQN – Q learning - Optimization).

6.12 Results Summary

In this section, a kind of summary is presented for the above-generated simulation results in terms of a sample statistics to highlight the average percentage enhancement achieved in the examined performance metrics when RL based DQN algorithm is implemented as a channel estimator. Table 6.2 includes a sample numerical values and these values characterize the estimated amount of enhancement achieved by RL based DQN algorithm compared to conventional MMSE scheme when various performance metrics such as BER, outage probability, and the individual capacity are inspected. Sample power values 5dBm, and 20dBm are selected to record the percentage improvements for each metric in Rayleigh fading channel. Table 6.2 clearly indicates that the proposed RL based DQN procedure is providing a significant improvement for different performance metrics in Rayleigh channel for all users.

Table 6.3 is providing the comparison between the proposed RL based DQN algorithm, and the following benchmarks schemes: MMSE scheme, DL based LSTM, and RL based Q algorithm when all of these approaches are realized as channel estimator for all users in Rayleigh fading channel. Sum rate metric is inspected at two different power levels, and the recorded values reflect a sufficient improvement in the sum rate when DQN algorithm is applied compared to other benchmark schemes. It is clearly noticed that the DQN algorithm recodes a major enhancement over MMSE procedure when sum rate metric is considered. This huge improvement was clarified based on the policy and target networks in DQN approach that play a sufficient role maximizing the long-term reward.

Similarly, Table 6.4 also shows the comparison between the proposed RL based DQN algorithm, and the following benchmarks schemes: MMSE scheme, DL based LSTM, and RL based Q algorithm when all of these approaches are realized as channel estimator. Sum rate metric is inspected when the number of users in NOMA cell is 5 or 20. As expected the highest gain is mainly achieved when DQN algorithm is compared to MMSE procedure, while for LSTM scheme and Q learning scheme the rate enhancement is noticeable because of the resemblance in the technique of estimation for DQN algorithm, LSTM network, and Q algorithm.

Table 6.2 Sample statistics for average percentage improvement (DQN vs MMSE)

Performance metric	5dBm	20dBm	Criteria Improvement	Users	Channel
BER	≈ 88.7%	≈ 89.2%	DQN vs MMSE	All	Rayleigh
Outage Prob.	≈ 88.4%	≈ 89.3%	DQN vs MMSE	All	Rayleigh
Individual Capacity	≈ 94.8%	≈ 61.6%	DQN vs MMSE	All	Rayleigh

Table 6.3 Sample statistics for average percentage improvement (Sum rate vs Power)

Performance metric	5dBm	20dBm	Criteria Improvement	Users	Channel
Sum Rate	≈ 500%	≈ 160%	DQN vs MMSE	All	Rayleigh
Sum Rate	≈ 50%	≈ 34%	DQN vs LSTM	All	Rayleigh
Sum Rate	≈ 36%	≈ 25%	DQN vs Q-algorithm	All	Rayleigh

Table 6.4 Sample statistics for average percentage improvement (Sum rate vs Number of users)

Performance metric	5 users	20 users	Criteria Improvement	Users	Channel
Sum Rate	≈ 23%	≈ 30%	DQN vs MMSE	All	Rayleigh
Sum Rate	≈ 7%	≈ 10.7%	DQN vs LSTM	All	Rayleigh
Sum Rate	≈ 4.8%	≈ 4%	DQN vs Q-algorithm	All	Rayleigh

6.13 Summary

In this chapter, the impact of utilizing a developed DRL based DQN algorithm to specifically estimate the channel coefficients for each user device in NOMA system is discussed. In the developed DRL based DQN algorithm, the DQN model is initialized based on random channel statistics then DQN model is updated based on the interaction between the Q-agent and the environment to minimize the estimation loss and at same time maximize the received downlink sum rate. The reliability of the proposed DQN procedure to estimate the channel parameters is explored by inspecting the performance of the proposed algorithm against selective benchmark channel estimation schemes, such as MMSE, LSTM, RL based actor-critic procedure, and RL based SARSA algorithm. Also, the reliability of the proposed DQN procedure is validated by inspecting the behaviour of the proposed DQN algorithm when the cell capacity is increased.

Chapter 7

Conclusions and Future Work

7.1 Introduction

Next generation of communication networks are not just an outspread form of the previous generation because of the growing request of network data traffic, but the upcoming generation must support the new emerged technologies such as the internet of things (IoT) devices, and the web-based artificial intelligence (AI) applications. Also, the forthcoming generation should have the capability to assist the massive connectivity and enhance the spectral efficiency. Typically, orthogonal multiple access (OMA) scheme can be considered as the standard multiple access scheme but in limited communication environment, and OMA scheme may not have the capability to serve huge networks that may require a diverse quality of services (QoS).

In addition, OMA may suffer from the issue of restricted degrees of freedom (DoF), where user with a good channel condition is served first with respect to user with weak channel condition who need to postpone for channel access. In order to achieve the requirements of QoS and DoF that needed for high scale communication networks, non-orthogonal multiple access (NOMA) was introduced as a new multiple access scheme to enhance the system capacity. To achieve high spectral efficiency and huge connectivity, power domain NOMA (PD-NOMA) is considered as one of the multiple access candidate's schemes, since it can handle signals that have significant difference in power levels. Also, the integration between PD-NOMA and multiple antenna system is considered as a suitable approach to improve the system capacity.

To enhance the system capacity, we work on maximizing the sum rates for user devices in the PD-NOMA system based on optimizing the power factors for each user in accordance with the channel conditions and based on the overall power transmitted and the QoS constraints. A systematic mathematical analysis for the optimization problem is introduced and the Lagrange function and optimality conditions are employed to derive the optimal power factors. The simulation results in terms of the BER, outage probability, and sum rate have verified that the

PD-NOMA can provide a reliable performance compared to OMA counterpart.

7.2 Conclusions

In this thesis, the impact of the utilizing deep neural network (DNN) in explicitly estimating the channel parameters for each user in NOMA system is investigated, where the LSTM layer is developed and implemented for complex data processing. In the proposed DNN based LSTM, the LSTM model is trained online based on both the generated normalized channel statistics and the relationship between the successive training sequences. The validity and efficiency of the proposed channel estimation scheme based on LSTM is emphasized by inspecting the proposed model in Rayleigh and Rician fading channels. Compared to the classical channel estimation schemes, the simulation results based on the proposed LSTM layer have recorded performance improvement in terms of different performance metrics such as BER, outage probability and sum rate by about 70%, 68%, and 35% respectively when Rayleigh fading channel is considered. Furthermore, performance enhancement was also achieved in terms of metrics such as BER, outage probability and capacity by about 29%, 77%, and 4% respectively when Rician fading channel model is applied. Moreover, the proposed DNN based LSTM for channel estimation and the optimized power scheme are both jointly utilized to examine how this joint model can be used for multiuser detection in the PD-NOMA system. As expected, an enhancement is achieved by about 31%, 65%, and 3% for BER, outage probability and sum rate performance metrics respectively compared to the detection scheme achieved when fixed power allocation is applied.

In addition, in this thesis, the impact of implementing a developed model for reinforcement learning based Q algorithm to distinctly estimate the channel parameters for user devices in the MISO-NOMA system is also investigated. In the developed Q learning algorithm, the Q model is created and initialized based on a random channel statistics, then the Q model is updated based on the interaction between the Q agent and the environment to maximize the received downlink sum rates for all users and at same time minimize the estimation loss. The efficiency of the developed Q learning procedure is explored by inspecting the performance of the proposed algorithm against different benchmark channel estimation techniques. The first considered benchmark scheme is based on the classical MMSE procedure, the second benchmark is DNN based LSTM structure, the third adopted scheme is RL based actor-critic algorithm, and the fourth applied benchmark technique is based on RL based SARSA

algorithm.

The dependability of the developed Q learning algorithm in estimating the channel parameters is validated by analysing the behaviour of the proposed Q algorithm against conventional NOMA channel estimation scheme in terms of diverse performance metrics in different fading channels. The simulation results for the proposed Q model emphasized a performance improvement in terms of several performance metrics such as BER, outage probability and sum rate by about 38%, 31%, and 94% respectively when Rayleigh fading channel is considered. Furthermore, Q learning performance enhancement against conventional NOMA channel estimation scheme was also noticed in terms of BER by about 26%, when Rician fading channel is considered.

In addition, we have also simulated a new scenario that explores how the channel estimation process based on the developed Q algorithm and the optimized power scheme can both cooperatively exploited for multiuser detection in the MISO-NOMA system. The simulation results in terms of BER, Outage probability, and individual capacity have demonstrated that the developed Q learning procedure for channel estimation jointly with the optimized power scheme can both realize a consistent performance. In the joint scenario, an improvement is recorded by about 50%, 38%, and 12% for BER, outage probability and individual capacity performance metrics respectively when being compared with the fixed power allocation scenario.

Moreover, in this thesis, the deep reinforcement learning (DRL) approach based deep Q-network (DQN) is developed and proposed to clearly estimate the downlink channel parameters for users in the NOMA cell based on the support given from the network information and the pilot symbols. As a benchmark comparison, a distinct channel prediction schemes are also implemented such as MMSE procedure, DNN based LSTM structure, and RL based Q algorithm. Simulation results have verified that developed DQN algorithm can realize a substantial performance improvement compared to the simulated benchmark channel prediction schemes. DQN simulation outcomes have clarified a performance enhancement in terms of sum rate metric by about 500%, 50%, 36% respectively when compared with MMSE procedure, LSTM network, and Q algorithm respectively in Rayleigh fading channel and at 5 dBm power level. Similarly, when 20 users are considered in NOMA cell, DQN simulation

results also have proved a performance improvement in terms of sum rate metric, by about 30%, 10%, 4% respectively when compared with MMSE procedure, LSTM network, and Q algorithm respectively in Rayleigh fading channel. Based on these outcomes, we can claim that the reliability can be ensured by our developed DQN algorithm even when the users in the NOMA cell is increased.

Finally, the simulation outcomes have also verified that when both the optimized power scheme and the proposed DQN algorithm for channel estimation are both applied in NOMA system, an enhancement in the BER is also achieved and a reduction in the power needed can be recorded approximately by 10 dBm.

7.3 Future Work

In this section we will introduce and outline some future research directions that can clarify how ML algorithms can make NOMA system more flexible, and easier to implement. Most of the current research in NOMA systems is based on stationary environments. Power allocation scheme, user pairing, and SIC technique are mainly discussed based on the static behaviour of the user devices. As the user device is moving from one location to another location, then the channel gain will vary with respect to the user's location with respect to the BS. Therefore, proposing a dynamic ML algorithms that can globally adapt or react with the continuous change of the channel condition for the user devices can be considered as one of the future research areas.

In NOMA cell, the user devices are multiplexed at the transmitter side using SC procedure and decoded at the receiver side using the SIC technique. As the number of devices in the NOMA cell increases, the allocation of powers to the devices may need additional time and the complexity can increase. At receiver side, decoding the high-power user and continue this process until reaching the low power user is an exhausting task, and this process can increase the latency when the number of multiplexed devices is huge. Therefore, a powerful ML algorithm is needed to manage the complexity and latency either at the transmitter side or at the receiver side especially for huge NOMA networks, and this can be a one of the challenging research areas.

Deep reinforcement learning (DRL) is an alternative method for solving resource allocation issues. For the next generation of wireless communication technologies, there exist a numerous

resource allocation tasks to be considered, the resource allocation can vary between power allocation, channel assignment, transmission scheduling, user selection, beamforming, and joint schemes of these techniques. The current level of deep learning knowledge is still not sufficient to handle this large amounts of tasks. Therefore, DRL can be a suggesting candidate algorithm that needs to be investigated carefully for future communication networks for the purpose of jointly handling more than one resource allocation task.

Furthermore, there are some major challenges that need to be considered in 6G such as global coverage considering both terrestrial and non-terrestrial networks, big data sets generated using extremely heterogeneous networks, and security need to be strengthened. Also, the wide range of channel encoding techniques, can also be exploited in 6G digital communication systems to improve the reliability and efficiency of data transmission over noisy channels. There are different types of channel coding schemes such as Low-Density Parity-Check (LDPC) coding scheme. LDPC is characterized by their excellent error correction capabilities and low encoding and decoding complexity. Hence, as a future research areas, all the above-mentioned fields can be investigated and integrated with ML algorithms and then exploited by NOMA system to further enhance the system performance.

References

[1] Dai, L.; Wang, B.; Ding, Z.; Wang, Z.; Chen, S.; Hanzo, L. "A survey of non-orthogonal multiple access for 5G", IEEE Commun. Surveys Tuts. pp. 2294–2323, 2018.

[2] L. Dai, B. Wang, Y. Yuan, S. Han, C. I, and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends," IEEE Commun. Mag., vol. 53, no. 9, pp. 74-81, 2015

[3] Aldababsa, M.; Toka, M.; Gökçeli, S.; Kurt, G.K.; Kucur, O. "A Tutorial on Nonorthogonal Multiple Access for 5G and beyond." Wireless Commun. Mob. Computing, 2018.

[4] Gaballa, M.; Abbod, M.; Jameel, A. "Power Optimization Analysis using Throughput Maximization in MISO Non-Orthogonal Multiple Access System" In Proceedings of the 2021 IEEE Globecom Workshops, Madrid, Spain, 7–11 December 2021.

[5] AbdelMoniem, M.; Gasser, S.M.; El-Mahallawy, M.S.; Fakhri, M.W.; Soliman, A. "Enhanced NOMA system using adaptive coding and modulation based on LSTM neural network channel estimation." Appl. Sci. 2019, 9, 3022.

[6] O'Shea, T.; Hoydis, J. An introduction to deep learning for the physical layer. IEEE Trans. Cogn. Commun. Netw. 2017, 3, 563–575.

[7] C. Luo, J. Ji, Q. Wang, X. Chen, and P. Li, "Channel State Information Prediction for 5G Wireless Communications: A Deep Learning Approach," in IEEE Transactions on Network Science and Engineering, vol. 7, no. 1, pp. 227-236, 1 Jan.-March 2020, doi: 10.1109/TNSE.2018.2848960.

[8] Rothmann, M.; Porrmann, M. "A Survey of Domain-Specific Architectures for Reinforcement Learning." IEEE Access 2022, 10, 13753–13767.

[9] Naeem, M.; De Pietro, G.; Coronato, A. "Application of Reinforcement Learning and Deep Learning in Multiple-Input and Multiple-Output (MIMO) Systems." Sensors 2022, 22, 309.

[10] Gaballa, M.; Abbod, M.; Aldallal, A. "Investigating the Combination of Deep Learning for Channel Estimation and Power Optimization in a Non-Orthogonal Multiple Access

System." Sensors 2022, 22, 3666. https://doi.org/10.3390/s22103666.

[11] Gaballa, M.; Abbod, M.; Aldallal, A. A Study on the Impact of Integrating Reinforcement Learning for Channel Prediction and Power Allocation Scheme in MISO-NOMA System. Sensors 2023, 23, 1383. https://doi.org/10.3390/s23031383.

[12] M. Gaballa, M. Abbod and A. Aldallal, "Deep Learning and Power Allocation Analysis in NOMA System," 2022 Thirteenth International Conference on Ubiquitous and Future Networks (ICUFN), Barcelona, Spain, 2022, pp. 196-201, doi: 10.1109/ICUFN55119.2022.9829643.

[13] Neumann, D.; Wiese, T.; Utschick, W. "Learning the MMSE Channel Estimator. IEEE Trans. Signal Process. 2018, 66, 2905–2917.

[14] Chu, M.; Liu, A.; Jiang, C.; Lau, V.K.N.; Yang, T. Wireless Channel Prediction for Multi-user Physical Layer with Deep Reinforcement Learning. In Proceedings of the 2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring), Helsinki, Finland, 19–22 June 2022.

[15] Ahsan, W.; Yi, W.; Qin, Z.; Liu, Y.; Nallanathan, A. Resource Allocation in Uplink NOMA-IoT Networks: A Reinforcement-Learning Approach. IEEE Trans. Wireless Communication, 2021, 20, 5083–5098.

[16] E. Hossain and M. Hasan, ''5G cellular: Key enabling technologies and research challenges,' IEEE Instrum. Meas. Mag., vol. 18, no. 3, pp. 11–21, Jun. 2015.

[17] M. Baghani, S. Parsaeefard, M. Derakhshani and W. Saad, "Dynamic Non-Orthogonal Multiple Access and Orthogonal Multiple Access in 5G Wireless Networks," in IEEE Transactions on Communications, vol. 67, no. 9, pp. 6360-6373, Sept. 2019, doi: 10.1109/TCOMM.2019.2919547.

[18] P. N. Thakre and S. B. Pokle, "A survey on Power Allocation in PD-NOMA for 5G Wireless Communication Systems," 2022 10th International Conference on Emerging Trends in Engineering and Technology - Signal and Information Processing (ICETET-SIP-22), Nagpur, India, 2022, pp. 1-5, doi: 10.1109/ICETET-SIP-2254415.2022.9791576.

[19] A. Mahmood, S. Khan, S. Hussain, and M. Zeeshan, "Performance Analysis of Multi-

User Downlink PD-NOMA Under SUI Fading Channel Models," in IEEE Access, vol. 9, pp. 52851-52859, 2021, doi: 10.1109/ACCESS.2021.3070147.

[20] C. -X. Wang et al., "On the Road to 6G: Visions, Requirements, Key Technologies, and Testbeds," in IEEE Communications Surveys & Tutorials, vol. 25, no. 2, pp. 905-974, Second quarter 2023, doi: 10.1109/COMST.2023.3249835.

[21] H. Zuo and X. Tao, "Power allocation optimization for uplink non-orthogonal multiple access systems," 2017 International Conference on Wireless Communications and Signal Processing (WCSP), 2017, pp. 1-5, 2017, IEEE.

[22] M. S. Ali, H. Tabassum and E. Hossain, "Dynamic User Clustering and Power Allocation for Uplink and Downlink Non-Orthogonal Multiple Access (NOMA) Systems," in IEEE Access, vol. 4, pp. 6325-6343, 2016, IEEE.

[23] T. Park, G. Lee, W. Saad, and M. Bennis, "Sum Rate and Reliability Analysis for Power-Domain Nonorthogonal Multiple Access (PD-NOMA)," in IEEE Internet of Things Journal, vol. 8, no. 12, pp. 10160-10169, 2021, IEEE.

[24] B. Narottama and S. Y. Shin, "Dynamic Power Allocation for Non-Orthogonal Multiple Access with User Mobility," 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2019, pp. 0442-0446, IEEE.

[25] T. Assaf, A. Al-Dweik, M. E. Moursi and H. Zeineldin, "Exact BER Performance Analysis for Downlink NOMA Systems Over Nakagami-m Fading Channels," in IEEE Access, vol. 7, pp. 134539-134555, 2019.

[26] P. Gupta and D. Ghosh, "Channel Assignment with Power Allocation for Sum Rate Maximization in NOMA Cellular Networks," 2020 5th International Conference on Computing, Communication and Security (ICCCS), 2020, pp. 1-5, IEEE.

[27] S. H. Amin, A. H. Mehana, S. S. Soliman and Y. A. Fahmy, "User Capacity in Downlink MISO-NOMA Systems," IEEE Global Communications Conference (GLOBECOM), pp. 1-7, 2018.

[28] Z. Sun and Y. Jing, "Average Power Analysis and User Clustering Design for MISO-NOMA Systems," IEEE International Workshop on Signal Processing Advances in

Wireless Communications, pp. 1-5, 2020.

[29] Y. Fu, L. Salaün, C. W. Sung and C. S. Chen, "Distributed Power Allocation for the Downlink of a Two-Cell MISO-NOMA System," IEEE 87th Vehicular Technology Conference, pp. 1-6, 2018.

[30] T. -V. Nguyen, V. -D. Nguyen, T. -N. Do, D. B. da Costa and B. An, "Spectral Efficiency Maximization for Multiuser MISO-NOMA Downlink Systems with SWIPT," IEEE Global Communications Conference, pp. 1-6, 2019.

[31] S. Li, M. Derakhshani, C. S. Chen and S. Lambotharan, "Outage Probability Analysis for Two-Antennas MISO-NOMA Downlink with Statistical CSI," IEEE Global Communications Conference, pp. 1-6, 2019.

[32] F. Alavi, K. Cumanan, M. Fozooni, Z. Ding, S. Lambotharan and O.A. Dobre, "Robust Energy-Efficient Design for MISO Non-Orthogonal Multiple Access Systems," in IEEE Transactions on Communications, vol. 67, no. 11, pp. 7937-7949, 2019.

[33] Xu, F., & Zhang, H. "Min-SINR maximization for mmWave massive MISO-NOMA system with randomly directional beamforming", IEEE Access, 8, 2020.

[34] Fu, Y., Zhang, M., Salaün, L., Sung, C. W., & Chen, C. S. "Zero-forcing oriented power minimization for multi-cell MISO-NOMA systems: A joint user grouping, beamforming, and power control perspective". IEEE Journal on Selected Areas in Communications, 38(8), 2020.

[35] Jeong, Y., Lee, C., & Kim, Y. H. "Power minimizing beamforming and power allocation for MISO-NOMA systems". IEEE Transactions on Vehicular Technology, 68(6), 2019.

[36] Karthik, M., Reddy, N. S., Kasyapa, G., Sriram, K., & Kirthiga, S. "Cooperative NOMA and Energy Harvesting using MISO in 5G Networks". 5th International Conference on Computer, Communication and Signal Processing, pp. 216-221, 2021.

[37] N. Zhang, K. Cheng, and G. Kang, "A Machine-Learning-Based Blind Detection on Interference Modulation Order in NOMA Systems," in IEEE Communications Letters, vol. 22, no. 12, pp. 2463-2466, Dec. 2018, doi: 10.1109/LCOMM.2018.2874218.

[38] A. K. Jain, Jianchang Mao and K. M. Mohiuddin, "Artificial neural networks: a tutorial,"

in Computer, vol. 29, no. 3, pp. 31-44, March 1996, doi: 10.1109/2.485891.

[39]   Y. Lu, P. Cheng, Z. Chen, W. H. Mow, Y. Li, and B. Vucetic, "Deep Multi-Task Learning for Cooperative NOMA: System Design and Principles," in IEEE Journal on Selected Areas in Communications, vol. 39, no. 1, pp. 61-78, Jan. 2021, doi: 10.1109/JSAC.2020.3036943.

[40]   A. Ly and Y. -D. Yao, "A Review of Deep Learning in 5G Research: Channel Coding, Massive MIMO, Multiple Access, Resource Allocation, and Network Security," in IEEE Open Journal of the Communications Society, vol. 2, pp. 396-408, 2021, doi: 10.1109/OJCOMS.2021.3058353.

[41]   M. Kim, N. -I. Kim, W. Lee, and D. -H. Cho, "Deep Learning-Aided SCMA," in IEEE Communications Letters, vol. 22, no. 4, pp. 720-723, April 2018, doi: 10.1109/LCOMM.2018.2792019.

[42]   X. Yi and C. Zhong, "Deep Learning for Joint Channel Estimation and Signal Detection in OFDM Systems," in IEEE Communications Letters, vol. 24, no. 12, pp. 2780-2784, Dec. 2020, doi: 10.1109/LCOMM.2020.3014382.

[43]   Zhao, Zheng, Weihai Chen, Xingming Wu, Peter CY Chen, and Jingmeng Liu. "LSTM network: a deep learning approach for short-term traffic forecast." IET Intelligent Transport Systems 11, no. 2 (2017): 68-75.

[44]   K. N. Doan, M. Vaezi, W. Shin, H. V. Poor, H. Shin, and T. Q. S. Quek, "Power Allocation in Cache-Aided NOMA Systems: Optimization and Deep Reinforcement Learning Approaches," in IEEE Transactions on Communications, vol. 68, no. 1, pp. 630-644, Jan. 2020, doi: 10.1109/TCOMM.2019.2947418.

[45]   Y. Lu, P. Cheng, Z. Chen, W. H. Mow, Y. Li, and B. Vucetic, "Deep Multi-Task Learning for Cooperative NOMA: System Design and Principles," in IEEE Journal on Selected Areas in Communications, vol. 39, no. 1, pp. 61-78, Jan. 2021, doi: 10.1109/JSAC.2020.3036943.

[46]   X. Wang, L. Yan, and Q. Zhang, "Research on the Application of Gradient Descent Algorithm in Machine Learning," 2021 International Conference on Computer Network, Electronic and Automation (ICCNEA), Xi'an, China, 2021, pp. 11-15, doi:

10.1109/ICCNEA53019.2021.00014.

[47] A. Shrestha and A. Mahmood, "Review of Deep Learning Algorithms and Architectures," in IEEE Access, vol. 7, pp. 53040-53065, 2019, doi: 10.1109/ACCESS.2019.2912200.

[48] R. Saravanan and P. Sujatha, "A State of Art Techniques on Machine Learning Algorithms: A Perspective of Supervised Learning Approaches in Data Classification," 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2018, pp. 945-949, doi: 10.1109/ICCONS.2018.8663155.

[49] H. U. Dike, Y. Zhou, K. K. Deveerasetty and Q. Wu, "Unsupervised Learning Based on Artificial Neural Network: A Review," 2018 IEEE International Conference on Cyborg and Bionic Systems (CBS), Shenzhen, China, 2018, pp. 322-327, doi: 10.1109/CBS.2018.8612259.

[50] W. Qiang and Z. Zhongli, "Reinforcement learning model, algorithms and its application," International Conference on Mechatronic Science, Electric Engineering and Computer (MEC), Jilin, China, 2011, pp. 1143-1146, doi: 10.1109/MEC.2011.6025669.

[51] A. Mohammadian, C. Tellambura and G. Y. Li, "Deep Learning-Based Phase Noise Compensation in Multicarrier Systems," in IEEE Wireless Communications Letters, vol. 10, no. 10, pp. 2110-2114, Oct. 2021, doi: 10.1109/LWC.2021.3093574.

[52] N. Ye, J. Pan, X. Wang, P. Wang, and X. Li, "Online Reconfigurable Deep Learning-Aided Multi-User Detection for IoT," 2021 International Wireless Communications and Mobile Computing (IWCMC), Harbin City, China, 2021, pp. 133-137, doi: 10.1109/IWCMC51323.2021.9498949.

[53] H. Yu, Z. Fei, Z. Zheng, N. Ye and Z. Han, "Deep Learning-Based User Activity Detection and Channel Estimation in Grant-Free NOMA," in IEEE Transactions on Wireless Communications, vol. 22, no. 4, pp. 2202-2214, April 2023, doi: 10.1109/TWC.2022.3209667.

[54] Y. Ahn, W. Kim, and B. Shim, "Active User Detection and Channel Estimation for Massive Machine-Type Communication: Deep Learning Approach," in IEEE Internet of Things Journal, vol. 9, no. 14, pp. 11904-11917, 15 July15, 2022, doi:

10.1109/JIOT.2021.3132329.

[55] J. Connor and L. Atlas, "Recurrent neural networks and time series prediction," IJCNN-91-Seattle International Joint Conference on Neural Networks, Seattle, WA, USA, 1991, pp. 301-306 vol.1, doi: 10.1109/IJCNN.1991.155194.

[56] T. Faghani, A. Shojaeifard, K. -K. Wong and A. H. Aghvami, "Recurrent Neural Network Channel Estimation Using Measured Massive MIMO Data," 2020 IEEE 31st Annual International Symposium on Personal, Indoor and Mobile Radio Communications, London, UK, 2020, pp. 1-5, doi: 10.1109/PIMRC48278.2020.9217192.

[57] S. Siami-Namini, N. Tavakoli and A. S. Namin, "The Performance of LSTM and BiLSTM in Forecasting Time Series," 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 2019, pp. 3285-3292, doi: 10.1109/BigData47090.2019.9005997.

[58] M. Gaballa, M. Abbod and S. Alnasur, "Hybrid Deep Learning for Channel Estimation and Power Allocation for MISO-NOMA System," 2022 IEEE Future Networks World Forum (FNWF), Montreal, QC, Canada, 2022, pp. 361-366, doi: 10.1109/FNWF55208.2022.00070.

[59] X. Miao, D. Guo, and X. Li, "Grant-Free NOMA With Device Activity Learning Using Long Short-Term Memory," in IEEE Wireless Communications Letters, vol. 9, no. 7, pp. 981-984, July 2020, doi: 10.1109/LWC.2020.2976992.

[60] Long Short-Term Memory (Sepp Hochreiter and Jürgen Schmidhuber), In Neural Computation, volume 9, 1997.

[61] I. Budhiraja et al., "A Systematic Review on NOMA Variants for 5G and Beyond," in IEEE Access, vol. 9, pp. 85573-85644, 2021, doi: 10.1109/ACCESS.2021.3081601.

[62] Y. Deng and T. Ohtsuki, "Low-Complexity Subspace MMSE Channel Estimation in Massive MU-MIMO System," in IEEE Access, vol. 8, pp. 124371-124381, 2020, doi: 10.1109/ACCESS.2020.3006242.

[63] I. -H. Lee and H. Jung, "User Selection and Power Allocation for Downlink NOMA Systems with Quality-Based Feedback in Rayleigh Fading Channels," in IEEE Wireless

Communications Letters, vol. 9, no. 11, pp. 1924-1927, Nov. 2020, doi: 10.1109/LWC.2020.3008174.

[64]	Tang, Z.; Wang, J.; Wang, J.; Song, J. On the achievable rate region of NOMA under outage probability constraints. IEEE communication Lett. 2019, 23, 370–373.

[65]	Zhu, J.; Wang, J.; Huang, Y.; He, S.; You, X.; Yang, L. On Optimal Power Allocation for Downlink Non-Orthogonal Multiple Access Systems. IEEE J. Sel. Areas Commun. 2017, 35, 2744–2757.

[66]	Yang, Z.; Xu, W.; Pan, C.; Pan, Y.; Chen, M. On the Optimality of Power Allocation for NOMA Downlinks with Individual QoS Constraints. IEEE Commun. Lett. 2017, 21, 1649–1652.

[67]	Ding, Z.; Yang, Z.; Fan, P.; Poor, H.V. On the Performance of Non-Orthogonal Multiple Access in 5G Systems with Randomly Deployed Users. IEEE Signal Process. Lett. 2014, 21, 1501–1505.

[68]	Gaballa, M.; Abbod, M.; Jameel, A.; Khaled, N. Throughput Maximization & Power Optimization Analysis in Non-Orthogonal Multiple Access System. In Proceedings of the 2021 IEEE 4th 5G World Forum, Montreal, QC, Canada, 13–15 October 2021.

[69]	Boyd, S.; Vandenberghe, L. Convex Optimization; Cambridge University Press: Cambridge, UK, 2004.

[70]	Ben-Tal, A.; Nemirovski, A. Lecture on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications; MPS-SIAM: Philadelphia, PA, USA, 2018.

[71]	Y. Fu, L. Salaün, C. W. Sung and C. S. Chen, "Distributed Power Allocation for the Downlink of a Two-Cell MISO-NOMA System," 2018 IEEE 87th Vehicular Technology Conference (VTC Spring), Porto, Portugal, 2018, pp. 1-6, doi: 10.1109/VTCSpring.2018.8417837.

[72]	Ghous, M.; Hassan, A.K.; Abbas, Z.H.; Abbas, G.; Hussien, A.; Baker, T. Cooperative Power-Domain NOMA Systems: An Overview. Sensors 2022, 22, 9652.

[73]	Gaballa, M.; Abbod, M.; Albasman, M. Power Allocation & MRC Analysis for Single Input Multi Output Non-Orthogonal Multiple Access System. In Proceedings of the 2021

IEEE International Conferences on Internet of Things (iThings) and IEEE Green Computing & Communications (GreenCom) and IEEE Cyber, Physical & Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics), Melbourne, Australia, 6–8 December 2021.

[74] Wang, X.; Zhu, P.; Li, D.; Xu, Y.; You, X. Pilot-Assisted SIMO-NOMA Signal Detection with Learnable Successive Interference Cancellation. IEEE communication Lett. 2021, 25, 2385–2389.

[75] D. -T. Do, T. -L. Nguyen, S. Ekin, Z. Kaleem and M. Voznak, "Joint User Grouping and Decoding Order in Uplink/Downlink MISO/SIMO-NOMA," in IEEE Access, vol. 8, pp. 143632-143643, 2020, doi: 10.1109/ACCESS.2020.3013835.

[76] B. E. Youcef Belmekki, A. Hamza, and B. Escrig, "Outage Analysis of Cooperative NOMA Using Maximum Ratio Combining at Intersections," 2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), Barcelona, Spain, 2019, pp. 1-6, doi: 10.1109/WiMOB.2019.8923449.

[77] B. Makki, K. Chitti, A. Behravan and M. -S. Alouini, "A Survey of NOMA: Current Status and Open Research Challenges," in IEEE Open Journal of the Communications Society, vol. 1, pp. 179-189, 2020, doi: 10.1109/OJCOMS.2020.2969899.

[78] W. Jiang and H. D. Schotten, "Deep Learning for Fading Channel Prediction," in IEEE Open Journal of the Communications Society, vol. 1, pp. 320-332, 2020, doi: 10.1109/OJCOMS.2020.2982513.

[79] Emir, A.; Kara, F.; Kaya, H.; Li, X. Deep learning-based flexible joint channel estimation and signal detection of multi-user OFDM-NOMA. Phys. Communication. 2021, 48, 101443.

[80] Chuan, L.; Chang, Q.; Li, X. A deep learning approach for MIMO-NOMA downlink signal detection. Sensors 2019, 19, 2526.

[81] Yang, Y.; Gao, F.; Ma, X.; Zhang, S. Deep learning-based channel estimation for doubly selective fading channels. IEEE Access 2019, 7, 36579–36589.

[82] Mao, Z.; Shi, Y. Deep learning-based channel estimation in fog radio access networks. China Communication. 2019, 16, 16–28.

[83] Bai, Q.; Wang, J.; Zhang, Y.; Song, J. Deep Learning-Based Channel Estimation Algorithm over Time Selective Fading Channels. IEEE Trans. Cognitive Communication Networks. 2019, 6, 125–134.

[84] Gui, G.; Huang, H.; Song, Y.; Sari, H. Deep learning for an effective nonorthogonal multiple access scheme. IEEE Trans. Veh. Technol. 2018, 67, 8440–8450.

[85] Emir, A.; Kara, F.; Kaya, H.; Yanikomeroglu, H. Deep Learning Empowered Semi-Blind Joint Detection in Cooperative NOMA. IEEE Access 2021, 9, 61832–61852.

[86] Ye, H.; Li, G.Y.; Juang, B.-H.F. Power of deep learning for channel estimation and signal detection in OFDM systems. IEEE Wireless Communication Lett. 2018, 7, 114–117.

[87] A. C. S. Kumar, S. M. Bhandarkar and M. Prasad, "DepthNet: A Recurrent Neural Network Architecture for Monocular Depth Prediction," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, USA, 2018, pp. 396-3968, doi: 10.1109/CVPRW.2018.00066.

[88] A. K. Gizzini, M. Chafii, S. Ehsanfar and R. M. Shubair, "Temporal Averaging LSTM-based Channel Estimation Scheme for IEEE 802.11p Standard," 2021 IEEE Global Communications Conference (GLOBECOM), Madrid, Spain, 2021, pp. 01-07, doi: 10.1109/GLOBECOM46510.2021.9685409.

[89] V. Pattana-Anake and F. J. J. Joseph, "Hyper Parameter Optimization of Stack LSTM Based Regression for PM 2.5 Data in Bangkok," 2022 7th International Conference on Business and Industrial Research (ICBIR), Bangkok, Thailand, 2022, pp. 13-17, doi: 10.1109/ICBIR54589.2022.9786465.

[90] T. He and J. Droppo, "Exploiting LSTM structure in deep neural networks for speech recognition," 2016 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Shanghai, China, 2016, pp. 5445-5449, doi: 10.1109/ICASSP.2016.7472718.

[91] X. Fu, B. F. d. Silva and D. Le Ruyet, "Multidimensional Codebook Design Using Deep Learning Techniques for Rayleigh Fading Channels," in IEEE Wireless Communications Letters, vol. 10, no. 9, pp. 1974-1978, Sept. 2021, doi: 10.1109/LWC.2021.3089024.

[92] H. He, S. Jin, C. -K. Wen, F. Gao, G. Y. Li, and Z. Xu, "Model-Driven Deep Learning

for Physical Layer Communications," in IEEE Wireless Communications, vol. 26, no. 5, pp. 77-83, October 2019, doi: 10.1109/MWC.2019.1800447.

[93] S. Dörner, S. Cammerer, J. Hoydis and S. t. Brink, "Deep Learning Based Communication Over the Air," in IEEE Journal of Selected Topics in Signal Processing, vol. 12, no. 1, pp. 132-143, Feb. 2018, doi: 10.1109/JSTSP.2017.2784180.

[94] Van de Beek, J.-J.; Edfors, O.; Sandell, M.; Wilson, S.K.; Borjesson, P.O. On channel estimation in OFDM systems. In Proceedings of the IEEE 45th VTC, Chicago, IL, USA, 25–28 July 1995; Volume 2, pp. 815–819.

[95] Recommendation ITU-R M.1225; Guidelines for Evaluation of Radio Transmission Technologies for IMT-2000. International Telecommunication Union: Geneva, Switzerland, 1997

[96] L. Wu, Z. Zhang, J. Dang, Y. Wu, H. Liu, and J. Wang, "Joint User Identification and Channel Estimation Over Rician Fading Channels," in IEEE Transactions on Vehicular Technology, vol. 69, no. 6, pp. 6803-6807, June 2020, doi: 10.1109/TVT.2020.2986466.

[97] P. Liu, D. Kong, J. Ding, Y. Zhang, K. Wang, and J. Choi, "Channel Estimation Aware Performance Analysis for Massive MIMO With Rician Fading," in IEEE Transactions on Communications, vol. 69, no. 7, pp. 4373-4386, July 2021, doi: 10.1109/TCOMM.2021.3073111.

[98] W. Saetan and S. Thipchaksurat, "Application of Deep Learning to Energy-Efficient Power Allocation Scheme for 5G SC-NOMA System with Imperfect SIC," 2019 16th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications, and Information Technology (ECTI-CON), Pattaya, Thailand, 2019, pp. 661-664, doi: 10.1109/ECTI-CON47248.2019.8955410.

[99] Jeon, Y.-S.; Li, J.; Tavangaran, N.; Poor, H.V. Data-aided channel estimator for MIMO systems via reinforcement learning. In Proceedings of the ICC 2020-IEEE International Conference on Communications (ICC), Dublin, Ireland, 7–11 June 2020.

[100] Kundu, N.K.; McKay, M.R. Channel Estimation for Reconfigurable Intelligent Surface Aided MISO Communications: From LMMSE to Deep Learning Solutions. IEEE Open J. Commun. Soc. 2021, 2, 471–487.

[101] Mthethwa, B.; Xu, H. Deep Learning-Based Wireless Channel Estimation for MIMO Uncoded Space-Time Labeling Diversity. IEEE Access 2020, 8, 224608–224620.

[102] Li, L.; Zhang, Z.; Yang, L. Influence of Autoencoder-Based Data Augmentation on Deep Learning-Based Wireless Communication. IEEE Wireless Communication Lett. 2021, 10, 2090–2093.

[103] Kim, D.; Jung, H.; Lee, I.H. Deep Learning-Based Power Control Scheme with Partial Channel Information in Overlay Device-to-Device Communication Systems. IEEE Access 2021, 9, 122125–122137.

[104] Zhang, T.; Mao, S. Energy-Efficient Power Control in Wireless Networks with Spatial Deep Neural Networks. IEEE Trans. Cognitive Communication. Networks, 2020, 6, 111–124.

[105] Chu, M.; Liu, A.; Jiang, C.; Lau, V.K.N.; Yang, T. Wireless Channel Prediction for Multi-user Physical Layer with Deep Reinforcement Learning. In Proceedings of the 2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring), Helsinki, Finland, 19–22 June 2022.

[106] Tan, J.; Liang, Y.C.; Zhang, L.; Feng, G. Deep Reinforcement Learning for Joint Channel Selection and Power Control in D2D Networks. IEEE Trans. Wireless Communication. 2021, 20, 1363–1378.

[107] Sutton, R.S.; Barto, A.G. Reinforcement Learning: An Introduction; MIT Press: Cambridge, MA, USA, 2018.

[108] Rothmann, M.; Porrmann, M. A Survey of Domain-Specific Architectures for Reinforcement Learning. IEEE Access 2022, 10, 13753–13767.

[109] Naeem, M.; De Pietro, G.; Coronato, A. Application of Reinforcement Learning and Deep Learning in Multiple-Input and Multiple-Output (MIMO) Systems. Sensors 2022, 22, 309.

[110] Nie, J.; Haykin, S. A Q-learning-based dynamic channel assignment technique for mobile communication systems. IEEE Trans. Veh. Technol. 1999, 48, 1676–1687.

[111] Azar, Mohammad Gheshlaghi, Remi Munos, Mohammad Ghavamzadeh, and Hilbert

Kappen. "Speedy Q-learning." In Advances in neural information processing systems. 2011.

[112] Mete, E.; Girici, T. Q-learning based scheduling with successive interference cancellation. IEEE Access 2020, 8, 172034–172042.

[113] Zhai, Q.; Bolić, M.; Li, Y.; Cheng, W.; Liu, C. A Q-Learning-Based Resource Allocation for Downlink Non-Orthogonal Multiple Access Systems Considering QoS. IEEE Access 2021, 9, 72702–72711.

[114] P. Aermsa-Ard, C. Wangsamad and K. Mamat, "NOMA Power Allocation Based on Q-Learning," 2022 37th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), Phuket, Thailand, 2022, pp. 892-895, doi: 10.1109/ITC-CSCC55581.2022.9894960.

[115] W. Ahsan, W. Yi, Z. Qin, Y. Liu, and A. Nallanathan, "Resource Allocation in Uplink NOMA-IoT Networks: A Reinforcement-Learning Approach," in IEEE Transactions on Wireless Communications, vol. 20, no. 8, pp. 5083-5098, Aug. 2021, doi: 10.1109/TWC.2021.3065523.

[116] Alsabah, Muntadher, Marwah Abdulrazzaq Naser, Basheera M. Mahmmod, Sadiq H. Abdulhussain, Mohammad R. Eissa, Ahmed Al-Baidhani, Nor K. Noordin, Sadiq M. Sait, Khaled A. Al-Utaibi, and Fazirul Hashim. "6G wireless communications networks: A comprehensive survey." IEEE Access 9 (2021): 148191-148243.

[117] M. Almekhlafi, M. A. Arfaoui, C. Assi and A. Ghrayeb, "Joint Resource and Power Allocation for URLLC-eMBB Traffics Multiplexing in 6G Wireless Networks," ICC 2021 - IEEE International Conference on Communications, Montreal, QC, Canada, 2021, pp. 1-6, doi: 10.1109/ICC42927.2021.9500443.

[118] J. Du, C. Jiang, J. Wang, Y. Ren, and M. Debbah, "Machine Learning for 6G Wireless Networks: Carrying Forward Enhanced Bandwidth, Massive Access, and Ultrareliable/Low-Latency Service," in IEEE Vehicular Technology Magazine, vol. 15, no. 4, pp. 122-134, Dec. 2020, doi: 10.1109/MVT.2020.3019650.

[119] Z. Yang, Y. Liu, Y. Chen, and N. Al-Dhahir, "Cache-aided NOMA mobile edge computing: A reinforcement learning approach," IEEE Trans. Wireless Communication,

vol. 19, no. 10, pp. 6899–6915, Oct. 2020, doi: 10.1109/TWC.2020.3006922.

[120] L. Xiao, Y. Li, C. Dai, H. Dai, and H. V. Poor, ``Reinforcement learning-based NOMA power allocation in the presence of smart jamming,'' IEEE Trans. Veh. Technol., vol. 67, no. 4, pp. 3377_3389, Apr. 2018.

[121] P. Yang, L. Li, W. Liang, H. Zhang, and Z. Ding, "Latency optimization for multi-user NOMA-MEC offloading using reinforcement learning," in Proc. 28th Wireless Opt. Communication Conf. (WOCC), May 2019, pp. 1–5, doi: 10.1109/WOCC.2019.8770605.

[122] K. N. Doan, M. Vaezi, W. Shin, H. V. Poor, H. Shin, and T. Q. S. Quek, "Power allocation in cache-aided NOMA systems: Optimization and deep reinforcement learning approaches," IEEE Trans. Communication, vol. 68, no. 1, pp. 630–644, Jan. 2020.

[123] S. Zhang et al., "A dynamic power allocation scheme in power-domain NOMA using actor-critic reinforcement learning," in Proc. IEEE/CIC Int. Conf. Communication China (ICCC), Aug. 2018, pp. 719–723.

[124] Y. Wei, F. R. Yu, M. Song, and Z. Han, ``User scheduling and resource allocation in HetNets with hybrid energy supply: An actor-critic reinforcement learning approach,'' IEEE Trans. Wireless Communication, vol. 17, no. 1, pp. 680_692, Jan. 2018.

[125] H. T. H. Giang, T. N. K. Hoan, P. D. Thanh, and I. Koo, "Hybrid NOMA/OMA-based dynamic power allocation scheme using deep reinforcement learning in 5G networks," Appl. Sci., vol. 10, no. 12, p. 4236, Jun. 2020.

[126] C. He, Y. Hu, Y. Chen, and B. Zeng, ``Joint power allocation and channel assignment for NOMA with deep reinforcement learning,'' IEEE J. Sel. Areas Communication., vol. 37, no. 10, pp. 2200_2210, Oct. 2019.

[127] Z. Ding, R. Schober and H. V. Poor, "Unveiling the Importance of SIC in NOMA Systems—Part 1: State of the Art and Recent Findings," in IEEE Communications Letters, vol. 24, no. 11, pp. 2373-2377, Nov. 2020, doi: 10.1109/LCOMM.2020.3012604.

[128] S. Rezvani, E. A. Jorswieck, R. Joda, and H. Yanikomeroglu, "Optimal Power Allocation in Downlink Multicarrier NOMA Systems: Theory and Fast Algorithms," in IEEE

Journal on Selected Areas in Communications, vol. 40, no. 4, pp. 1162-1189, April 2022, doi: 10.1109/JSAC.2022.3143237.

[129] Y. Cao, G. Zhang, G. Li, and J. Zhang, "A Deep Q-Network Based-Resource Allocation Scheme for Massive MIMO-NOMA," in IEEE Communications Letters, vol. 25, no. 5, pp. 1544-1548, May 2021, doi: 10.1109/LCOMM.2021.3055348.

[130] N. C. Luong et al., "Applications of Deep Reinforcement Learning in Communications and Networking: A Survey," in IEEE Communications Surveys & Tutorials, vol. 21, no. 4, pp. 3133-3174, Fourthquarter 2019, doi: 10.1109/COMST.2019.2916583.

[131] R. Parhi and R. D. Nowak, "The Role of Neural Network Activation Functions," in IEEE Signal Processing Letters, vol. 27, pp. 1779-1783, 2020, doi: 10.1109/LSP.2020.3027517.

[132] Tian, Yingjie, Yuqi Zhang, and Haibin Zhang. 2023. "Recent Advances in Stochastic Gradient Descent in Deep Learning" Mathematics 11, no. 3: 682. https://doi.org/10.3390/math11030682

[133] Ling, Jing, Junjuan Xia, Fusheng Zhu, Chongzhi Gao, Shiwei Lai, and Venki Balasubramanian. "DQN-based resource allocation for NOMA-MEC-aided multi-source data stream." EURASIP Journal on Advances in Signal Processing 2023, no. 1 (2023): 44.

[134] Y. Liu, Y. Deng, H. Zhou, M. Elkashlan and A. Nallanathan, "Deep Reinforcement Learning-Based Grant-Free NOMA Optimization for mURLLC," in IEEE Transactions on Communications, vol. 71, no. 3, pp. 1475-1490, March 2023, doi: 10.1109/TCOMM.2023.3238061.