Future Wireless Networks: Towards Learning-driven Sixth-generation Wireless Communications

Rabbia Saleem
4 January 2023

A thesis submitted for the degree of
Doctor of Philosophy of
The Australian National University

Declaration

Except where otherwise indicated, this thesis is my own original work.

Rabbia Saleem
January 4, 2023

Abstract

The evolution of wireless communication networks, from present to the emerging fifth-generation (5G) new radio (NR), and sixth-generation (6G) is inevitable, yet propitious. The thesis evolves around application of machine learning and optimization techniques to problems in spectrum management, internet-of-things (IoT), physical layer security, and intelligent reflecting surface (IRS).

The first problem explores License Assisted Access (LAA), which leverages unlicensed resource sharing with the Wi-Fi network as a promising technique to address the spectrum scarcity issue in wireless networks. An optimal communication policy is devised which maximizes the throughput performance of LAA network while guaranteeing a proportionally fair performance among LAA stations and a fair share for Wi-Fi stations. The numerical results demonstrate more than 75 % improvement in the LAA throughput and a notable gain of $8-9$ % in the fairness index.

Next, we investigate the unlicensed spectrum sharing for bandwidth hungry diverse IoT networks in 5G NR. An efficient coexistence mechanism based on the idea of adaptive initial sensing duration (ISD) is proposed to enhance the diverse IoT-NR network performance while keeping the primary Wi-Fi network's performance to a bearable threshold. A Q-learning (QL) based algorithm is devised to maximize the normalized sum throughput of the coexistence Wi-Fi/IoT-NR network. The results confirm a maximum throughput gain of 51 % and ensure that the Wi-Fi network's performance remains intact.

Finally, advanced levels of network security are critical to maintain due to severe signal attenuation at higher frequencies of 6G wireless communication. Thus, an IRS-based model is proposed to address the issue of network security under trusted-untrusted device diversity, where the untrusted devices may potentially eavesdrop on the trusted devices. A deep deterministic policy gradient (DDPG) algorithm is devised to jointly optimize the active and passive beamforming matrices. The results confirm a maximum gain of $2-2.5$ times in the sum secrecy rate of trusted devices and ensure Quality-of-Service (QoS) for all the devices.

In conclusion, the thesis has led towards efficient, secure, and smart communication and build foundation to address similar complex wireless networks.

Acknowledgments

This thesis is not just a piece of work, these are four precious years of my life's manuscript.

I would like to begin by thanking God for blessing me the opportunity, courage and patience to make it happen.

I can never thank enough my entire supervisory panel for their support. Dr. Lachlan Blackhall, Dr. Wei Ni, and Dr. Muhammad Ikram, thank you for believing in me, supporting me in the middle of crisis and giving me the confidence to be what I wanted to be. This would have not been possible without you all and I am filled with gratitude to have mentors like you.

My lovely parents, I hope I have made you proud. Over this period every single time when I thought of quitting you both gave me hope. To my amazing Dad, you were, and you will always be my greatest strength. I would also like to shower love on my siblings and my beautiful nieces for their encouragement and prayers throughout this journey.

Last but not the least, I would like to thank a bunch of beautiful people in my life for always being there. Ali, you are my family, and I wouldn't have survived last four years without your support. Dani, my always go to person, thank you so much for your love and help. My lovely friends and colleagues, I will always cherish your support throughout this journey.

Cheers to all of you.

Contents

List of Figures

List of Tables

List of Abbreviations

AI	Artificial Intelligence
ACK	Acknowledgement
ANN	Artificial Neural Network
AP	Access Point
AWGN	Additive White Gaussian Noise
BEB	Binary Exponential Backoff
BS	Base Station
CCA	Clear Channel Assessment
CDF	Cumulative Distribution Function
COT	Channel Occupancy Time
CSI	Channel State Information
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance
CW	Contention Window
DCF	Distributed Coordination Function
DDPG	Deep Deterministic Policy Gradient
DIFS	Distributed Interframe Space
DL	Deep Learning
DNN	Deep Neural Network
DRL	Deep Reinforcement Learning
eNB	eNodeB
IoE	Internet-of-Everything
IoT	Internet-of-Things
IRS	Intelligent Reflecting Surface
ISD	Initial Sensing Duration
LAA	Licensed Assisted Access

LBT	Listen Before Talk
LoS	Line of Sight
LTE	Long Term Evolution
LTE-U	Long Term Evolution Unlicensed
mmWave	Millimeter Wave
MDP	Markov decision process
MIMO	Multiple-Input Multiple-Output
MISO	Multiple-Input Single-Output
ML	Machine Learning
NR	New Radio
NR-U	New Radio Unlicensed
OFDMA	Orthogonal Frequency Division Multiple Access
QL	Q-Learning
QoS	Quality of Service
RIS	Reconfigurable Intelligent Surface
RL	Reinforcement Learning
SIFS	Short Interframe Space
SNR	Signal to Noise Ratio
SINR	Signal to Interference Noise Ratio
THz	Terahertz
UAV	Unmanned Air Vehicle
URLLC	Ultra-Reliable Low Latency Communications
WLAN	Wireless Local Area Networks
3GGP	3rd Generation Partnership Project
5G	Fifth-Generation
6G	Sixth-Generation

List of Publications

[1] **R. Saleem**, S. Alvi and S. Durrani, "Performance-Fairness Trade-off for Wi-Fi and LTE-LAA Coexistence," IEEE Access, vol. 9, pp. 62446-62459, Apr. 2021.

[2] **R. Saleem**, W. Ni, M. Ikram and A. Jamalipour, "Deep Reinforcement Learning-Driven Secrecy Design for Intelligent Reflecting Surface-Based 6G-IoT Networks," IEEE Internet of Things Journal, pp. 1–1, 2022.

[3] **R. Saleem**, W. Ni and M. Ikram, "Reinforcement Learning-based Unlicensed Spectrum sharing for IoT-devices of 5G New Radio," in 2022 IEEE International Mediterranean Conference on Communications and Networking (MeditCom), Athens, Greece, Sep. 2022, pp. 191-196.

"The thesis is dedicated to the memory of my late father, Saleem."

With love,
Rabbia.

Chapter 1

Introduction

1.1 Introduction

Wireless communication industry has experienced a never stopping evolution over the decades. On one hand, the evolution has rendered tremendous development of networks such as cellular, wireless local area networks (WLAN), satellite, terrestrial, wireless sensor networks, optical wireless networks, etc. Whereas, on the other hand, with advancement comes challenges such as increased levels of network complexity, resource utilization, ultra-low latency, inter-network interference, signal quality, energy and power control, network security, reliability, and so on [4]. Hence, there is a constant need to overcome the originated issues for a seamless growth of future wireless communication technology.

Currently, we are living in the era of fifth-generation (5G) communication which is entitled to bring all the previous and existing wireless networks on a single platform leading towards the future road of sixth-generation (6G) communication [5]. Previously mobile communication was simply focused on the betterment of network performance, but 5G pioneered to facilitate the use cases of wireless services for people, devices and industry [6]. The initiation of smart cities, virtual reality, massive machine-type communication, ultra-reliable low latency communications (URLLC), unlicensed spectrum sharing, vehicle-to-vehicle communication and industrial internet-of-things (IoT) are some of the use cases introduced by 5G new radio (NR).

Moreover, all the existing and emerging technologies are shaping a landscape of massive access for wireless networks. These huge number of devices and machines are realizing a new design of future 6G network for efficient, low power, reliable, and secure access [7]. The new infrastructure is also aimed to deal with large of amount data from massive IoT devices, leading to the enhanced cloud and edge computing methods [8]. The growth of devices, machines, and their communication data invites a forthcoming shortage of available radio resources. In that regard, bringing

a necessity of improvised radio resource management for the large-scale heterogeneous networks in licensed and unlicensed spectrum [9]. Also, acquiring wider unexplored bandwidths for forecasted high connection volume, data rates, efficiency, and capacity [10].

Established on the idea of revolutionized wireless systems from 5G NR to beyond 5G and 6G, this thesis has focused on some of the core challenges of spectrum access to the unlicensed band for cellular and massive devices (which defies conventional centralised resource management) and learning-based radio resource management, security and integrity. Also, the problem of massive devices' access to the network, from different and potentially complementary angles, including access frequency, adaptive access using learning, and secure access using learning. In conclusion, all the studies collectively contribute to a safe and massive access scheme in support of future 6G IoT networks.

1.2 Current and Emerging Wireless Communication Networks

The present 5G networks have exploited technologies such as massive multiple-input multiple-output (MIMO) and millimetre-wave (mmWave) to be able to reach higher data rates and better efficiency [11]. The multi-array antenna structures have manifested increased spectral efficiency by combating signal interference and encouraging spatial multiplexing [12]. Similarly, mmWaves are a hub of available bandwidth to cope with huge traffic volumes of 5G NR. Edge computing is also trending in 5G networks to support low latency by combining the services of telecommunications and information technology [13]. All the three technologies have their own advantages as well as challenges. Recent studies are looking into cell-free massive MIMO to control interference, severe signal attenuation in mmWaves, and data privacy with edge computing.

Recently, unmanned air vehicles (UAVs) and satellite communication are the talk of the town in the wireless community. In near future, UAVs are expected to become a fundamental entity in a wireless network because of their mobility, strong line-of-sight link and flexibility [14]. But there are challenges related to 3D deployment, wireless connectivity, trajectory planning, and energy efficiency. Likewise, satellite-based communication is aimed exercise global coverage for heterogeneous networks and reducing the burden from terrestrial stations [15]. Satellite and UAVs are predicted to handle the large-scale connectivity of emerging IoT networks. Aiming to resolve the issues of delays, link budget, Doppler shift and power consumption, the low orbit satellites will succeed in making their way to future wireless networks [16]. However, the coexistence with ground-based networks cannot be overlooked.

In addition, the fierce expansion in wireless data traffic can't be unseen at this stage. Hence, both academia and industry are looking into emerging technologies to shape the future of wireless communication networks. Research towards the terahertz (THz) band is prevalent now to foster wider bandwidth capacity and extremely high data rates. The attractively high bitrate cannot be achieved without monitoring the challenges of channel modelling, signal fading, molecular absorption, and THz beamforming [17]. On the same note, intelligent reflective surface (IRS) is a new trend that has been proposed to boost the signal quality in an uncontrollable fading environment by continuously configuring its passive reflecting elements [18]. IRS offers programmable control over shifting the phase and amplitude of reflected signals, hence, contributing to smart radio transmission [19]. Blockchain is another popular technology with decentralized control over large amounts of data that is predicted to manage secure and private communication for wireless networks [20]. However, there are issues like energy consumption and scalability for highly complex and large-scale systems.

The future 6G networks are expected to extend the digital communication and physical realms of the current systems to the human world. A digital twin technology is anticipated which will allow digital clones of the physical world processes to help the in real-life interactions and decision-making [21]. 6G is forecasted to offer diverse and inclusive services such as, smart society, sustainable ecosystem, telemedicine, e-education, mixed reality, holographic telepresence, quantum computing, optical wireless communication, hybrid access, and tactile internet [22]. This whole new technology will foresee issues of massive twining, reliable connectivity, sustainability, diverse level of security and privacy, and more.

Having said that, the increased complexity of the present and emerging wireless communication problems has shifted towards a paradigm of more sophisticated and intelligent solutions due to the high computational consumption of traditional mathematical methods [21]. Additionally, the complex and dynamic nature of problems has identified artificial intelligence (AI) and machine learning (ML) as key tools to implement autonomous adaptability and take advantage of experience while making decisions. Therefore, the wireless communication industry has also started leveraging these techniques and is seeking AI as one of the building blocks of future 6G networks [23].

1.3 Evolution of ML and AI Tools in Wireless Communication

In recent years, ML emerged as an application of AI for learning, decision making and problem-solving. It enables the devices to perform a

task relying on self-observations and inferences instead of being explicitly programmed [24]. ML and AI are starting to influence wireless communication and we are entering into an era of everything being smart, e.g., smart devices, smart homes, smart cities, etc. Similarly, due to the growing complexity of wireless networks, there is a prominent interest in developing intelligent and adaptive systems. This has given a boost to adopting techniques of ML in wireless systems which can efficiently organize and optimize the networking devices [25].

1.3.1 Types of Machine Learning Techniques

There are three different types of ML techniques: *a) Supervised Learning b) Unsupervised Learning* and *c) Reinforcement learning* (RL) [26]. Supervised learning is used in applications that are based on the classification, estimation, or prediction of a parameter. It takes labelled input data sets to establish a predictive model. The basic principle is to train the model with a set of training data and test data to predict the output value [27]. On the other hand, unsupervised learning works on unlabelled data to extract a feature, detect an anomaly, identify a pattern, or reduce the dimensionality. It uses the concept of clustering and grouping [28]. Lastly, RL is applied to the problems which need online adaptation of the system parameters [29]. In these real-time running scenarios, the decision-making entities (*agents*) interact with the *environment* in the form of some *actions* and receive *rewards* in return. The agents then try to maximize the reward in order to achieve the optimal value [30]. All the three types of ML previously mentioned have gained momentum in solving open research problems [31].

1.3.2 Classification of Machine Learning Literature

Recently, the research trend in wireless communication has shifted towards ML and its especially important to group the present study based on its nature. The state-of-the-art literature on ML in wireless networks can be classified as illustrated in Figure 1.1. The research work has been categorized on the use of ML in *channel modelling, transceiver design, application areas* and *techniques*.

ML is being used in the physical layer design of wireless systems. ML-based feature extraction for channel estimation and channel prediction is a new paradigm in channel modelling. These channel estimation and equalization methods have proven to provide a better bit error rate, therefore, improving the performance of network [32]. Lately, some research work has been directed to constructing radio environment maps based on ML and its applications to wireless communications. The models offer more accuracy as compared to empirical channel models [33]. Though,

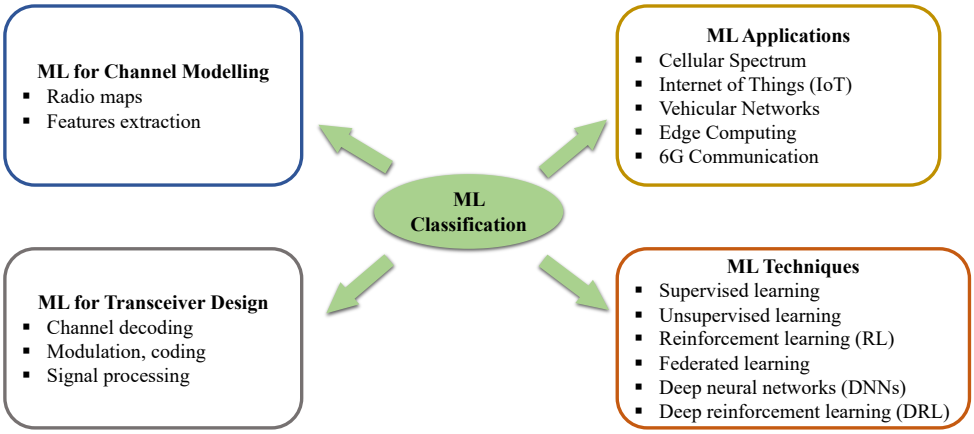


Figure 1.1: Classification of Machine learning literature in wireless communication.

ML-based channel modelling has its advantages, but it comes with some limitations which need to be explored.

Apart from the conventional ML techniques, an interrelated class of learning with AI, deep learning has demonstrated physical layer application as well. The latest study on learning-based physical layer processing includes transceiver design, optimization of modulation, channel coding and decoding schemes [34]. Thus, introducing an auto decoder and encoder in communication systems [35]. These learning-driven information coding and signal processing techniques also have their own limitations, thereby opening gates to future research.

Over the past years, ML has successfully marked its ground in multiple wireless communication applications. Next-generation wireless networks demand massive connectivity, intelligent adaptation and decision-making. These features invite ML-based solutions to counter the challenges in current and future self-organizing cellular systems [36]. ML has also found its application in massive MIMO and cognitive radios [37]. Moreover, widespread use of UAVs is seen in the recent aerial wireless communication for better coverage and capacity. ML techniques are being applied to UAV networks as well [38,39]. Furthermore, complex networks such as IoT and vehicular networks are also being catered by ML to resolve the issues [40–42]. Other wireless communication areas which are leveraging ML-based techniques are edge computing [43, 44], cyber security [45], backscatter communications [46] and satellite communication [47]. These applications offer vast room for future research.

The last part of the literature is grouped based on advancement and research in different ML techniques. The conventional techniques fall under three different types of machine learning mentioned in Section 1. Some of the modelling techniques for supervised learning are support vector ma-

chines, artificial neural networks (ANN), deep neural networks (DNN), Bayesian networks, decision trees and genetic algorithms. Clustering, dimensionality reduction and anomaly detection fall under unsupervised learning. Whereas, dynamic programming, Monte Carlo, Q-learning (QL) and Sarsa are the few techniques for RL [48]. Recently another technique, federated learning, has emerged which provides distributed learning reducing the complexity of dealing with large data sets [43]. Further, a combination of RL and deep learning, deep reinforcement learning (DRL) has been proposed over the last few years [40]. To summarize, AI and ML techniques are also evolving.

This thesis aims to explore optimization and ML techniques as a tool to effectively tackle some of the feasible challenges of wireless systems. ML is suitable for complex scenarios where traditional methods are unable to provide significant improvements.

1.4 Thesis Structure and Contributions

The following subsections provide detail about the thesis outline and the key contributions.

1.4.1 Thesis Outline

The remaining part of the thesis is structured into five main chapters. The chapter 2 elaborates on the targeted wireless communication areas and the problem-solving techniques adopted for their solution. The next three chapters 3, 4 and 5 represent the research works investigated in this thesis. Moreover, each of these chapters is aimed to provide the literature background, system model, problem formulation, proposed solution, results, and key highlights of their respective research problem. The last chapter 6 extracts the thesis findings and proposes a prospective future research direction.

1.4.2 Main Contributions of the Thesis

The main contributions of this thesis are summarized in the following chapters as.

- **Chapter 2** serves as background knowledge for establishing the main research body. Initially, it throws light on the major motivation behind the research and later presents a big picture by connecting the distinct aspects of the study.

  The emphasis is laid on the targeted areas and prospective techniques covered by the thesis. Furthermore, it familiarizes the reader with imminent research problems in the said areas. Finally, it builds

up the foundation of the chosen problems and the importance of addressing them.

- **Chapter 3** provides an efficient performance-fairness trade-off for licensed assisted access (LAA) and Wi-Fi coexisting networks in the unlicensed spectrum. The model proposes a novel idea of unique transmission probability for each LAA station dependent upon the wireless channel conditions, which promotes proportional fairness among LAA stations. Previously, an equal and independent transmission probability for Wi-Fi and LAA stations has been considered in the literature.

  An optimal access strategy is devised which uses non-overlapping transmission phases for Wi-Fi and LAA network to utilize the unlicensed channel. The policy rules out inter-network collision and more importantly considers the near-far effect of the different stations and the physical channel conditions.

  A mathematical model for the average throughput per LAA station is constructed and a joint optimization problem is formulated to maximize the total throughput of the LAA network with a fair performance guarantee for the Wi-Fi network. Also, an optimal solution is obtained analytically.

  The proposed scheme provides more than 75 % throughput gain for the LAA network as compared to the conventional schemes of uniform transmission probabilities. Moreover, an improved $8 - 9$ % of the fairness index confirms the intra-network fairness of the proposed model, compared to the benchmark.

- **Chapter 4** introduces a diversity inclusive unlicensed spectrum sharing for IoT devices in 5G NR. It proposes a contention-based coexistence mechanism with dynamically adaptive initial sensing duration (ISD) for diverse IoT-NR devices rather than the conventional uniform ISD.

  A mathematical model is designed to maximize the sum throughput of the coexistence network with a bearable performance guarantee of the Wi-Fi network.

  It proposes a QL-based algorithm to learn from environment dynamics and eventually converge to an optimal ISD for each IoT-NR device.

  The proposed model yields a maximum gain of 51% in the normalized sum throughput when compared to the benchmark cases.

- **Chapter 5** offers a learning-driven IRS-assisted secure communication for the diverse 6G-IoT networks in the presence of potential untrusted eavesdroppers.

It presents the concept of trusted-untrusted classification of legitimate IoT devices and provides an IRS-supported approach to overcome the signal attenuation and blockages of networks operating at the mmWave/THz band.

An optimization problem is designed that jointly optimizes the transmit beamforming matrix and IRS phase shift matrix to maximize the sum secrecy rate of trusted devices with a minimum throughput guarantee for all the legitimate devices.

A DRL-based solution is provided using a deep deterministic policy gradient (DDPG) algorithm suitable for the complex, dynamic, and continuous formulated maximization problem. The simulation model is trained, and optimal results are furnished.

The proposed DRL-based model provides a maximum gain of $2 - 2.5$ times the sum secrecy rate of trusted devices and ensures the throughput performance of all trusted and untrusted devices. Moreover, the performance of the proposed DRL model is also tested on various values of hyper-parameters to select the best suited.

- **Chapter 6** highlights the key findings of this thesis and discusses future research directions. In the essence, the thesis explores the evolution of wireless communication technology from 5G to 6G while addressing the issues of bandwidth hunger, signal attenuation and diverse physical layer security. It brings a flavour of sophisticated learning-based real-time solutions suitable for large-scale complex dynamic problems when the traditional mathematical models struggle with high computational complexity.

  The research problems discussed in chapters 3, 4 and 5 offer a solid baseline for similar wireless communication problems in the areas of IoT, UAVs, and satellite and non-terrestrial communication.

Chapter 2

Targeted Research Areas and Problem Solving Techniques

2.1 Motivation

A smooth transition from 5G to future 6G communication entails vigilantly addressing the prominent issues among the networks such as access, connectivity, resource management, integration, coexistence, and privacy [49]. Therefore, the thesis is aimed to build up a strong motivation toward understanding and solving a few of the pivotal problems in the current and emerging wireless networks. Based on the pillars of efficient, fair, secure, and smart solutions, it progresses to highlight some of the challenges. The proposed solutions are aimed to be flexible and adaptive to explore similar research problems, thus, furnishing a wider scope of wireless application areas.

Following the trend toward 6G networks, this research work as a whole is planned to play a significant role in emphasizing the problems of access, diversity and security for the massive devices. The core highlights of the thesis are given as:

- **A fair access to massive devices:** Massive device connectivity of cellular, IoT, and Wi-Fi networks calls for fair access to the available radio resources. This is ensured in terms of a fairly-managed coexistence in the licensed as well as unlicensed spectrum [50–52].
- **An inclusive wireless environment:** Diversity is a critical aspect of wireless networks in terms of service requirements [5, 53]. The future wireless technology is expected to be inclusive of these diverse features and assures quality communication based on different demands.
- **A secure communication:** To maintain the integrity and privacy of the massive heterogeneous devices through secure and reliable communication links [54–56].

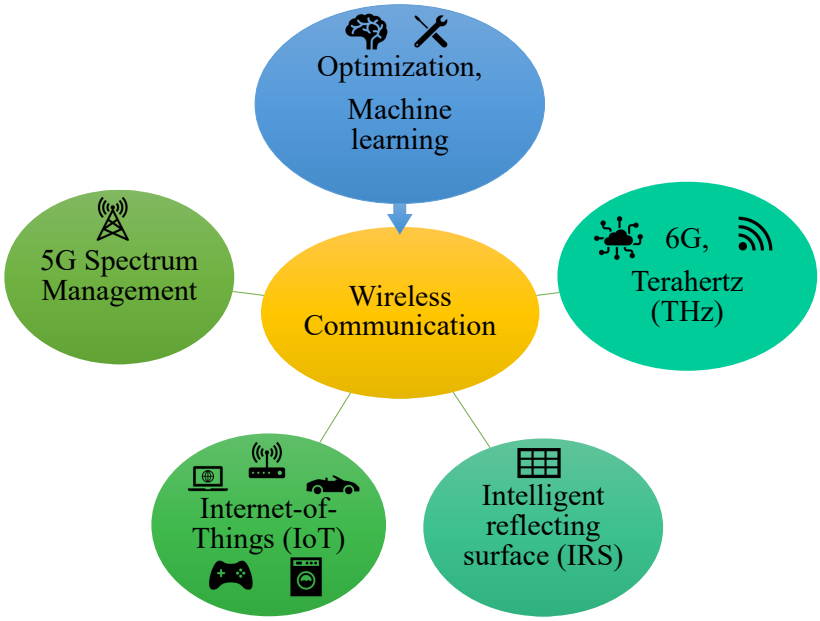


Figure 2.1: Targeted application areas and problem-solving techniques.

- **A smart and self-controlled network:** To provide intelligent and learning-based control to the complex dynamic networks [23, 57]. It develops scalable, adaptable and flexible characteristics for the system to behave autonomously in the changing environment.

2.2 Targeted Application Areas

The vast wireless communication technology encompasses boundless research gaps. This thesis targets a few of these application areas as illustrated in Figure 2.1. Motivated by 5G NR use cases and its emerging technologies, the research problems in spectrum management, IoT network, IRS-assisted communication, and 6G and THz networks are explored.

2.2.1 5G NR and Beyond Spectrum Management

The advanced wireless communication technology of 5G NR is burdened with catering a huge number of devices on various applications simultaneously. Therefore, to manage the systems efficiently, the concept of self-organised network management has been introduced rather than centralized control [58]. This indulges the current and future wireless systems in addressing the spectrum management issues such as:

- Spectrum access [59]

- Spectrum sharing [60–62]
- Resource allocation [63, 64]

One of the key 5G NR use cases is *unlicensed spectrum sharing* to counter the bandwidth capacity shortage. In this regard, chapter 3 and chapter 4 are based on two different research works of unlicensed spectrum sharing for cellular and IoT networks, respectively.

2.2.2 IoT Network

The idea of connecting machines, devices and networks together have given birth to large-scale IoT systems for 5G NR and beyond communication. The deployment of IoT, such as smart homes devices, gadgets, vehicles, mobiles, computing devices, etc has given rise to the following issues:

- Massive connectivity [65, 66],
- Resource management and energy efficiency [67, 68],
- Heterogeneous coexistence [69],
- Security and privacy [55].

One of the major concerns for the IoT networks is the inclusion of diverse demands of the connected things. Keeping this in mind, a diverse IoT coexistence in the unlicensed spectrum is visited in chapter 4. Also, the problem of diverse IoT network security is studied in chapter 5.

2.2.3 IRS-Assisted Communication

The IRS technology is being investigated as a favourable option for beyond MIMO communication for future wireless networks [70]. The passive reflecting surfaces, IRS, are mirror-like surfaces that can smartly control the signal propagation in the radio environment [71]. This cost-effective and low power consumption technology has taken over wireless communication to improve the signal quality in worse conditions [72]. Nevertheless, IRS-assisted communication faces issues such as:

- Phase angle configuration [73–75]
- Channel estimation [76]
- User detection [77, 78]

The research work in chapter 5 is based on phase angle configuration for IRS-assisted IoT networks.

2.2.4 6G and THz Network

The THz band is a gateway to wide available bandwidth resource, which is need of the hour for data capacity predicted for 6G networks [79]. However, extremely high data rates will not be achieved without experiencing the following problems:

- Channel modelling [80]
- Signal blockages and high propagation attenuation [81]
- Low coverage range [82]
- Security and privacy [56]

Hence, the last research study in chapter 5 examines the severe signal attenuation and security of a 6G-IoT network.

2.3 Problem Structures and Applied Techniques

The thesis contributes toward the application of machine learning algorithms and optimization techniques to suitable wireless communication problems. The research problems discussed in Chapters 3, 4 and 5 are formulated as optimization problems that are solved based on the following techniques:

2.3.1 Linear Optimization Technique

The first problem discussed in chapter 3 is proven to be a convex optimization problem that supports a closed-form optimal solution with low computational complexity. Therefore, a classical convex optimization technique is employed to yield optimal results.

2.3.2 ML Techniques

When the formulated optimization problems are non-convex and NP-hard in nature, the traditional mathematical solutions become impractical and analytical solution is nearly impossible to obtain. Thus, ML is an effective tool to resolve these problems because of its robustness, low computational complexity and autonomous adaptability [83]. The research problems formulated in this thesis have also leveraged from the following ML algorithms:

- RL has become a popular real-time self-learning process when no prior information from the surrounding is available. QL is one of the important RL techniques which incorporates agent and environment interface as shown in Figure 2.2. Agents are the decision-making

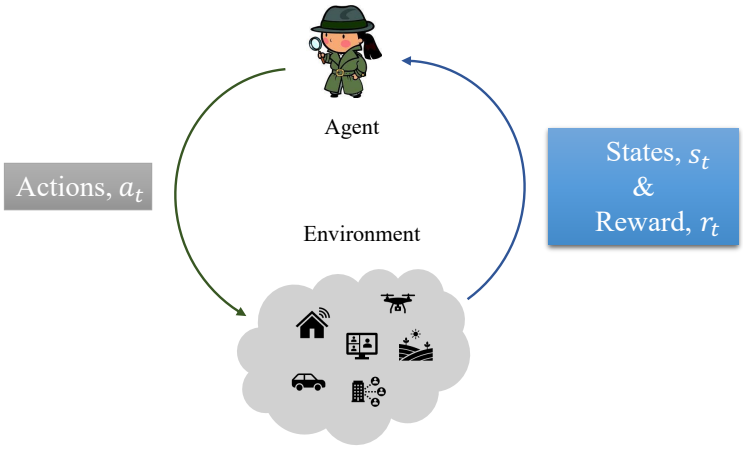


Figure 2.2: QL agent and environment framework involving actions, states, and rewards.

entities that respond to the changes in the environment through their actions. Based on their actions they transition between states and earn rewards [84]. The cumulative value of the reward converges the algorithm to optimal results. Resource management problems have found RL algorithms as a favourable solution for their complex and dynamic design [85]. Hence, in regard to the nature of the problem in Chapter 4, a QL-based solution is presented.

- DRL, a hybrid of RL with DNN, has emerged as a new technique that improves the learning rate for RL and adds to the neural network quality of handling high-scale complex problems. The effectiveness of these learning algorithms for mission-oriented challenges involving flexibility and decision-making under high uncertainty for IRS, IoT networks and future 6G networks has been extraordinary [57,86,87]. Therefore, it is a desirable technique for large-dimensional complex dynamic problems 6G-IoT networks such as the one investigated in chapter 5.

The PhD thesis is aimed to target the potential research works involving the application of optimization and ML techniques to the shortlisted application areas. The implementation of these techniques in wireless networks is a rich contribution to state-of-the-art research. The goal is to provide smarter and more efficient solutions to relevant wireless network challenges. The problems and their proposed solutions are elaborated in the following chapters.

Chapter 3

Performance-Fairness Trade-off for Wi-Fi and LTE-LAA Coexistence

3.1 Introduction

Recently wireless communication industry has turned its attention towards utilizing the unlicensed spectrum as an efficient means to address the spectrum scarcity and rapidly growing demand for data traffic by users [88, 89]. In this regard, different variants of the fourth-generation Long Term Evolution (LTE) have been proposed to leverage the unlicensed 5 GHz band which is mainly used by the Wi-Fi network, e.g., LTE unlicensed (LTE-U), LTE licensed assisted access (LAA) and MulteFire [90]. LTE-U emerged as the first standard for unlicensed sharing presented by the LTE-U forum on the basis of LTE specifications in release 12 [91]. In release 13, the 3rd generation partnership project (3GPP) issued LTE-LAA as a global standard to coexist with the unlicensed band [92]. Multefire was proposed in 2017 as a radio technology to self-deploy LTE in the unlicensed band without the need of an anchor in a licensed band [93]. LTE-U and LAA are based on carrier aggregation while MulteFire is based on the standalone operation.

The overall aim of the above-mentioned variants is the fair coexistence among different radio technologies, i.e., an LTE network should not impact the Wi-Fi network more than an additional Wi-Fi network in terms of throughput [94]. The idea of unlicensed spectrum sharing is extended to become the part of 5G new radio unlicensed (NR-U) standard and emerging 6G unlicensed spectrum sharing [5, 95]. LAA has been proposed as the basis for the channel access mechanism for 5G and 6G NR-U by industry and academia [96–99]. In this work, we focus on LTE-LAA/ Wi-Fi coexistence.

The standardization for the LTE-LAA and Wi-Fi coexistence in the

unlicensed 5 GHz band is still in under development. In this regard, the latest specifications for tuning the transmission opportunity, which is considered in this work, were discussed in the 3GPP working group meeting in Jan 2020 [100]. Moreover, the recent performance analysis studies such as the one done in [101] have identified many issues with the current coexistence specification. Multiple problems with the current existing releases were reported in [102]. The solution to these problems is not only crucial for 5 GHz unlicensed band, but the lessons learned from the 5 GHz coexistence are vital for fair and efficient coexistence deployments in 6 GHz and all future bands.

LTE can adversely affect the Wi-Fi throughput performance without a fair coexistence scenario. This is because of the difference in access mechanisms for Wi-Fi and LTE networks. A Wi-Fi network follows IEEE 802.11a Distributed Coordination Function (DCF) based on the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) mechanism to coexists with one another [103]. In order to maintain fair spectrum sharing with Wi-Fi, the conventional LTE-LAA adopts a contention-based channel accessing protocol. It is equipped with Listen before talk (LBT) mechanism which uses the clear channel assessment (CCA) to access the unlicensed channel and considers equal channel access probability for each LAA station [94]. The standard LTE-LAA mechanism has been shown to provide better data rates and higher airtime efficiency than the standalone Wi-Fi network operating in the unlicensed spectrum [104]. However, equal access probability does not take into the near-far effect of the LAA stations, and their individual throughput is affected. *Therefore, it is important to design a coexistence scheme that addresses both inter-network fairness as well as intra-network fairness in an LTE-LAA/Wi-Fi network.*

3.1.1 Related Work

The goal of Wi-Fi and LTE-LAA coexistence is to increase the throughput of the LAA network but not to degrade the performance of the Wi-Fi network. Some papers have looked at the performance analysis of standard or modified LBT mechanisms for LTE-LAA/Wi-Fi networks [105–110], without necessarily focusing on the fairness issue. Other works have addressed spatial reuse [111, 112], energy consumption [113], traffic offloading [114], power allocation [115] and resource allocation [116] issues. In this work, we are interested in the fairness issue. Two main approaches have been proposed in the literature to handle the fairness issue: (i) varying the transmission or channel occupancy time (COT) and (ii) varying the idle time. The transmission time of the LAA station is the time for which it keeps the channel occupied and idle time is defined as the time it keeps the channel vacant. An LAA station switches from the transmitting

mode and rests in the idle mode when it has no data to transmit, or it is undergoing the initial contention phase.

The first approach deals with transmission time modification. The percentage of time an LAA or Wi-Fi station occupies a channel account for overall channel efficiency and fairness between the two networks. Varying the COT under different load conditions is a coexistence solution. The work in [117] proposed adjusting COT for LAA stations from a range of values depending upon Wi-Fi load. Another COT modification of appending the Clear-To-Send frame for reserving the channel to make LAA less intrusive was studied in [118]. Recently, some papers have used machine learning algorithms to optimize Wi-Fi and LTE-LAA coexistence. A QL technique for estimating the channel occupancy time of LAA under different Wi-Fi traffic conditions was proposed in [119]. The learning outcome was then used to adapt COT and interference power constraints. A channel sharing scheme was proposed in [120] where LAA stations monitor the Wi-Fi activity for adaptive duty cycling. A QL-based approach was presented to intelligently select an optimal combination of transmission and mute time for LTE-U and Wi-Fi networks in [121]. A DRL-based approach was adopted to optimize the transmission time of LAA in [122].

The second approach deals with idle time modification. Under this approach, a key strategy is dynamically adapting the wait time to inculcate fairness among coexisting nodes. An LAA station undergoes a backoff phase where it waits for its turn to transmit relying upon the CW size. The contention window CW size adaptation has been presented to get the most out of available resources for both the LAA and Wi-Fi network in [123]. A QL-based solution to cope with the challenge of fair coexistence for LAA and Wi-Fi was proposed in [124, 125]. The focus was on controlling the CW size for every LAA eNB in accordance with channel state and traffic load information. Another supervised machine learning-based scheme was proposed in [126] to learn from past collisions and predict the CW size on the basis of negative acknowledgements of packets. All the stations contending on a channel had to sense it idle for a continuous duration. Making this contention time adaptive reduces the intra-network collisions and increases the system fairness. Initial sensing time optimization schemes were introduced in [127, 128]. Initial CW size and sensing time-based adaptation was proposed in [109] to achieve proportional fairness for LTE-LAA and Wi-Fi network.

The work in [129] considers LTE supplemental downlink methodology which is used only in the downlink to support higher downlink support for LTE stations. The work in [130], on the other hand, employs an alternating slot assignment model which assigns different time slots to LTE-U BS or Wi-Fi stations. Whereas the system model we consider in this work is based on uplink communication and employs a time division multiplexing-based solution. Hence, the perspective of resource sharing is

different in comparison to [129] and [130]. Note that we do not advocate that one model is better than the other and believe that different access schemes are designed to serve different purposes.

Some works have considered both idle time and transmission time modification. A fairness evaluation for Wi-Fi and LTE in the unlicensed spectrum for three different coexistence procedures: continuous transmission, discontinuous transmission and LBT was addressed in [131]. The work in [132] presented a dynamic transmission time and a fixed waiting time configuration for LTE-LAA and Wi-Fi coexistence.

With current 5G NR-U and advancement towards 6G spectrum sharing, enhanced LAA models are required to provide efficient unlicensed coexistence [133]. The future 6G technology is also looking into LAA for a capacity boost in the unlicensed spectrum [50]. Prior works on LTE-LAA are mainly focused on inter-network fairness, i.e., fairness between Wi-Fi and LTE-LAA networks. Whereas little attention is given to resource allocation based on intra-network fairness among LTE-U and Wi-Fi networks. The work in [134] highlighted the orthogonal channel allocation scheme based on wireless conditions of the unlicensed channel being shared by LTE-U and Wi-Fi. Proportional fairness for LTE-U/ Wi-Fi coexistence network was proposed for a resource allocation problem in [135]. However, the throughput-based intra-network fairness for an LAA network was ignored in the literature. This critical improvement is addressed in our work to provide a throughput trade-off to an enhanced level of fairness.

3.1.2 Contributions

In this study we consider a fair Wi-Fi and LAA coexistence model for a single unlicensed channel, maximize the LAA network throughput and retain Wi-Fi's performance. The main contributions of this work are given as follows:

- We model different transmission probabilities for each LAA station dependent upon the wireless channel conditions and fair resource sharing. This ensures a proportionally fair resource utilization among the LAA stations. To the best of our knowledge, all the aforementioned literature for Wi-Fi and LAA coexistence has assumed equal and independent packet transmission probability for Wi-Fi and LAA stations.

- We propose an optimal access scheme for LAA, rather than the conventional LAA-LBT policy, to coexist with the Wi-Fi network over the unlicensed spectrum. We divide the channel utilization into two non-overlapping phases for Wi-Fi and LAA. Although prior work in [122] has presented the idea of inter-network collision avoidance

through a similar strategy, they do not take into account the near-far effect of the different stations and the physical channel conditions.

- We formulate the mathematical model for the average throughput per LAA station. In addition, we formulate the optimization problem in order to simultaneously maximize the total throughput of the LAA network with a performance guarantee for the Wi-Fi network in terms of fair throughput share. The formulated optimization problem is analytically solved, and optimal design parameters are obtained.

- The proposed scheme provides more than 75 % throughput gain as compared to the benchmark scheme in which transmission probability is uniform for all LAA stations. The performance gain is profound when LAA stations are far from the eNB. In addition, a notable gain of $8-9$ % in the fairness index is observed. This reflects the improved intra-network fairness of the proposed LAA network over the conventional LAA network.

A list of the important variables and parameters is given in Table 3.1.

The following notation is used in this chapter. $p(\cdot)$ and $F(\cdot)$ represent the probability and the cumulative distribution function (CDF), respectively. $\mathcal{L}(\cdot)$ denotes the Lagrangian function. $\nabla(\cdot)$ is used for the gradient and $[\cdot]^\top$ is used for the transpose operator. The rest of this chapter is organized as follows. The system model is described in Section 3.2. The optimization problem and its derived solution are presented in Section 3.3. The numerical results are discussed in Section 3.4. Lastly, Section 3.5 provides the concluding remarks of the chapter.

3.2 System Model

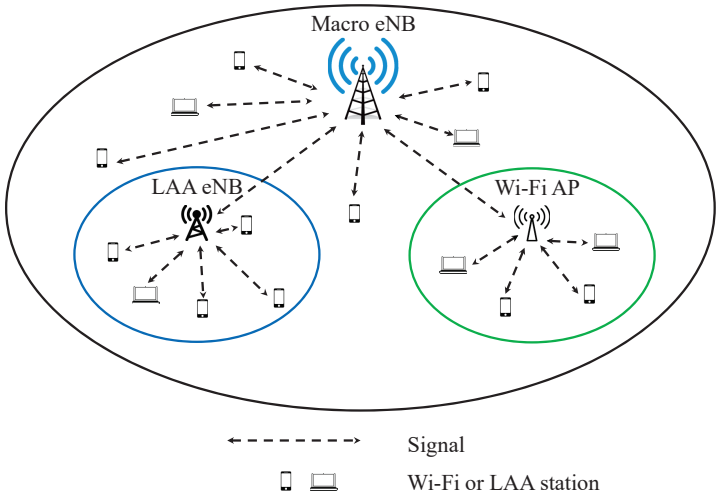
This section describes the Wi-Fi and LAA coexistence model in detail. Furthermore, it gives the analytical expressions for calculating the average throughput of Wi-Fi stations and LAA stations.

3.2.1 Wi-Fi and LAA Coexistence Model

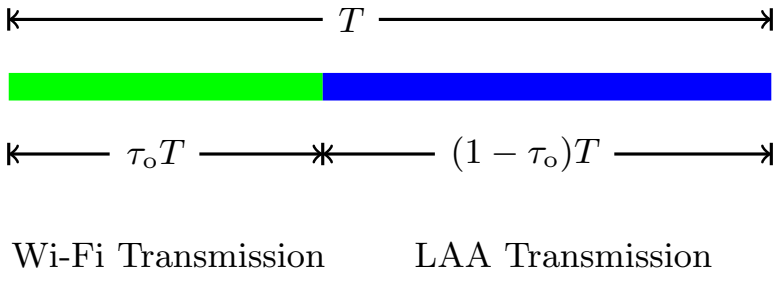
We consider a Wi-Fi and LTE-LAA network to coexist on a single channel of 20 MHz in the 5 GHz unlicensed band as illustrated in Fig. 3.1. In this work, our goal is to advocate for a cooperative LTE-LAA and Wi-Fi communication policy for fair coexistence. Therefore, we limit our system model to the fundamental single 20 MHz (non-aggregate) channel scenario which serves the purpose. Considering a multi-channel scenario with 40 MHz, 80 MHz, or 160 MHz aggregated channels gives rise to interference issues and other challenges which are outside the scope of this work [101].

Table 3.1: List of important parameters and variables.

Parameter Symbol	**Description**
W	Number of Wi-Fi stations.
w	Index of Wi-Fi station, where $w = 1, 2, ..., W$.
L	Number of LAA stations.
ℓ	Index of the LAA station, where $\ell = 1, 2, ..., L$.
τ_0	Fraction of time sharing between Wi-Fi and LAA.
θ	Idle slot time duration for Wi-Fi and LAA.
m	Backoff stage of Wi-Fi DCF.
$\mathrm{CW}_{\min}$	Initial contention window size of Wi-Fi DCF.
τ_{w}	Transmission probability of Wi-Fi station in a slot.
$p_{\mathrm{c,w}}$	Collision probability of Wi-Fi station in a slot.
$R_w^{\min}$	Minimum average throughput per Wi-Fi station.
$R_w^{\max}$	Maximum average throughput per Wi-Fi station.
$\tilde{R}_w$	Achievable average throughput per Wi-Fi station.
τ_ℓ	The channel access probability of the LAA station ℓ in a slot.
$p_{\mathrm{s},\ell}$	The probability of successfully accessing the channel by LAA station ℓ.
$R_{\mathrm{s},\ell}$	The average rate of ℓth LAA station during the successful channel access.
R_ℓ	The fixed rate of ℓth LAA station during the transmission phase.
$\tilde{R}_\ell$	The average throughput per LAA station.
$\mathcal{U}$	A set of all LAA stations excluding the station ℓ.
P_ℓ	The transmitted power for the LAA station ℓ.
$P_{\max}$	The available average power allocated per LAA station.
d_ℓ	The distance between the LAA eNB and the ℓth station.
h_ℓ	The fading channel gain of ℓth station.
γ_ℓ	The signal to noise ratio (SNR) of ℓth station.
α	The path loss exponent.
κ	The path loss factor.
σ^2	The additive white Gaussian noise (AWGN) power.



(a) Wi-Fi and LAA network.



(b) Wi-Fi and LAA transmission time model.

Figure 3.1: Illustration of system model: Wi-Fi and LAA coexisting on a single unlicensed channel.

Note that typically the LAA deployments are for the outdoors and Wi-Fi deployments are for the indoors [136–138]. The signal strength of both LAA and Wi-Fi devices operating outdoors is comparable. This leads to increased coexistence and hidden node problems. However, in the last few years, the proprietary Wi-Fi networks are increasingly being deployed in urban areas to support Wi-Fi in the outdoors [95]. This new rapidly spreading scenario is one of the most critical deployments in regard to the coexistence issues of LAA and Wi-Fi. Therefore, we consider the outdoors deployment setting. Nevertheless, the mathematical model in this work is valid for any channel size as long as both LAA and Wi-Fi operate on the same channel.

The network in Fig. 3.1a consists of an LTE-LAA eNB with L LAA stations and a Wi-Fi Access Point (AP) with W Wi-Fi stations. The macro eNB controls the unlicensed channel assignment to Wi-Fi and the

LAA eNB in its transmission range. Our focus is the coexistence of Wi-Fi and LAA after the channel has been assigned. We assume both Wi-Fi and LAA stations are in the saturated mode, i.e., the nodes always have a packet to transmit after successful transmission. Moreover, we model the total transmission time over the channel as N slots each of duration T. Each time slot T is further divided into t equal mini slots of length θ. In order to support the coexistence between Wi-Fi and LAA networks, each time slot T is divided into two phases: the Wi-Fi transmission phase with duration $\tau_0 T$ and the LAA transmission phase with duration $(1-\tau_0)T$ as shown in the Fig. 3.1b. Here, $0 < \tau_0 < 1$ is used to control the fraction of time assigned to Wi-Fi and LAA. During the Wi-Fi transmission phase, the Wi-Fi stations follow the standard DCF protocol [135] to access the channel with uniform access probabilities.

Moreover, during the LAA transmission phase, the LAA stations contend with each other for the remaining $(1-\tau_0)T$ time based on their channel access probabilities. Generally, in a Wi-Fi and LAA coexistence network, a collision occurs when two or more nodes transmit on a given channel at the same time, while an intended receiver is in their transmission range. This introduces two types of collisions: inter-network collisions and intra-network collisions. In our system model, Wi-Fi and LAA stations transmit on their non-overlapping transmission slots, thus, there will be no collisions between Wi-Fi and LAA stations. However, the probability of intra-network collisions is non-zero.

We analyze the performance of the proposed network setting in terms of the channel utilization for the successful transmission of data. This is referred to as the average throughput per station for the Wi-Fi and LAA. Firstly, for the proposed coexistence scenario, we obtain an achievable average throughput for a Wi-Fi station. Secondly, we calculate the average throughput of each LAA station based on the wireless channel conditions.

For a multi-channel channel scenario, there are two operations for LBT [139], i.e., Type-A and Type-B. For Type-A, our proposed problem setting can straightforwardly be extended. This will require disjoint rate adaptation for each of the aggregate channels. On the other hand, Type-B will require joint rate adaptation as well as optimization of the short clear channel assessment parameter used in Type-B. Nevertheless, a multi-channel scenario specifically for LBT Type-B operation poses an important challenge which is a promising future work direction.

3.2.2 Average Throughput for Wi-Fi Stations

IEEE 802.11 standard specifies Wi-Fi stations to adopt CSMA/CA [103]. Each station employs the binary exponential backoff (BEB) mechanism to transmit packets. It senses the channel for a distributed interframe space (DIFS) period to check its availability. If the channel is found idle the

backoff counter is decremented. The station attempts once the backoff counter is zero and the channel if found idle. After the packet transmission, the node waits for a period of short interframe space (SIFS) to receive an acknowledgement (ACK) from the AP. After a successful packet delivery, the station switches to the initial contention window (CW) size, $\mathrm{CW}_{\min}$. Under collision, the Wi-Fi station doubles the CW until it reaches the maximum contention window size, $\mathrm{CW}_{\max}$.

A homogeneous Wi-Fi network when it fully utilizes the channel (the case when $\tau_0 = 1$), provides a maximum throughput/rate to the stations. For a total number of W Wi-Fi stations, let $R_w^{\max}$ be the maximum average rate per station, where $w = 1, 2, \ldots, W$. Whereas, for a coexistence network, when the channel is being shared by both Wi-Fi and LAA (the case when $0 < \tau_0 < 1$), then performance is affected. According to the 3GPP standard [94], ideally, a single LAA station coexisting with Wi-Fi on the unlicensed channel must not degrade the Wi-Fi performance more than an additional Wi-Fi station. Therefore, for maximum allowable performance degradation, we consider that each LAA station acts just like a Wi-Fi station, and then the minimum average rate for each station is calculated as [140].

$$R_w^{\min} = \frac{p_{\mathrm{t,w}} p_{\mathrm{s,w}} D}{(1 - p_{\mathrm{t,w}})\theta + p_{\mathrm{t,w}} p_{\mathrm{s,w}} T_{\mathrm{s}} + p_{\mathrm{t,w}}(1 - p_{\mathrm{s,w}}) T_{\mathrm{c}}}, \tag{3.1}$$

where D is the average packet size and

$$p_{\mathrm{t,w}} = 1 - (1 - \tau_{\mathrm{w}})^n \tag{3.2}$$

is the transmission probability of at least one station in a mini-slot t and $n = W + L$ denotes the total number of stations (actual Wi-Fi stations plus LAA stations behaving like Wi-Fi stations). In (3.2), τ_{w} is the stationary transmission probability of each station which is calculated by solving [140]

$$\tau_{\mathrm{w}} = \frac{2(1 - 2p_{\mathrm{c,w}})}{(1 - 2p_{\mathrm{c,w}})(\mathrm{CW}_{\min} + 1) + p_{\mathrm{c,w}}(1 - (2p_{\mathrm{c,w}})^m)}, \tag{3.3}$$

where

$$p_{\mathrm{c,w}} = 1 - (1 - \tau_{\mathrm{w}})^{n-1} \tag{3.4}$$

is the stationary probability of collision in a single slot, $\mathrm{CW}_{\min}$, m are the initial window size and maximum backoff stage, respectively, based on the DCF mechanism. We assume that all the Wi-Fi stations have equal payload size and τ_{w}. In (3.1),

$$p_{\mathrm{s,w}} = \frac{n\tau_{\mathrm{w}}(1 - \tau_{\mathrm{w}})^{n-1}}{p_{\mathrm{t,w}}} \tag{3.5}$$

is the probability of any station to successfully transmit in the slot, given the fact that there is at least one transmission. Let T_{s} be the average period for the successful transmission and T_{c} be the average period for a collision. From [140]

$$T_{\mathrm{s}} = \frac{(H + D + \mathrm{ACK})}{R_{\mathrm{b}}} + \mathrm{SIFS} + 2\delta + \mathrm{DIFS} \tag{3.6}$$

and

$$T_{\mathrm{c}} = \frac{(H + D)}{R_{\mathrm{b}}} + \delta + \mathrm{DIFS}, \tag{3.7}$$

where H is the header size, ACK is the acknowledge size, δ is the propagation delay, R_{b} is the channel bitrate, SIFS is the short interframe space and DIFS is the distributed interframe space for Wi-Fi.

Since, we have assumed Wi-Fi and LAA networks to have non-overlapping transmission phases, therefore the conventional Wi-Fi activity is not being altered during its own transmission phase.

3.2.3 Average Throughput for LAA Stations

During the LAA transmission phase, all the L stations contend to access the channel with their fixed probabilities. We model the channel between LAA eNB and ℓth station over the mini slot t, as a quasi-static block fading channel that follows a Rayleigh distribution. The fading power gain of the channel from the eNB to the station ℓ or vice versa is denoted as h_ℓ.

In a mini-slot t, the channel access probability for the ℓth station is denoted as τ_ℓ where, $\ell = 1, 2, \ldots, L$. The probability $p_{\mathrm{s},\ell}$ is defined as the probability of the ℓth station to successfully access the channel when only the ℓth station accesses the channel, given the condition that at least one station has accessed the channel. This is written as

$$p_{\mathrm{s},\ell} = \frac{L\tau_\ell \prod\limits_{i \in \mathcal{U}} (1 - \tau_i)}{\left(1 - \prod\limits_{k=1}^{L} (1 - \tau_k)\right)}. \tag{3.8}$$

where, $\mathcal{U}$ is a set of all LAA stations excluding the station ℓ. The average throughput for the ℓth station for a given slot is calculated as

$$\tilde{R}_\ell = R_{\mathrm{s},\ell}\ p_{\mathrm{s},\ell}, \tag{3.9}$$

where, $R_{\mathrm{s},\ell}$ is the average rate of the station ℓ during the successful channel access and $p_{\mathrm{s},\ell}$ is the probability of successfully accessing the channel from (3.8). We assume that the station ℓ transmits at a fixed transmission

rate $(1-\tau_0)R_\ell$ during the whole transmission phase. This yields the average rate $R_{\mathrm{s},\ell}$ as a product of the fixed rate and the probability of successful transmission/non-outage, p(non-outage) and is given by

$$\begin{aligned} R_{\mathrm{s},\ell} &= R_\ell(1-\tau_0)\ p(\text{non-outage}) \\ &= R_\ell(1-\tau_0)\ p\bigg(\log_2\big(1+\gamma_\ell\big) \geq R_\ell\bigg), \end{aligned} \tag{3.10}$$

where γ_ℓ is the Signal-to-Noise-Ratio (SNR). It is obtained as

$$\gamma_\ell = \frac{\kappa P_\ell |h_\ell|^2}{d_\ell^\alpha \sigma^2}, \tag{3.11}$$

where $\kappa = (c/4\pi f)^2$ is the pathloss factor, c is the speed of light and f is the carrier frequency, P_ℓ is the transmit power for station ℓ, d_ℓ is the distance between the eNB and station ℓ, α is the path loss exponent, $|h_\ell|^2$ is the fading channel gain and σ^2 is the additive white Gaussian noise (AWGN) power. Substituting (3.11) into (3.10) and solving, we get

$$R_{\mathrm{s},\ell} = R_\ell(1-\tau_0)\bigg(1 - F_{\mathrm{h}_\ell}\bigg(\frac{d_\ell^\alpha \sigma^2(2^{R_\ell}-1)}{\kappa P_\ell}\bigg)\bigg), \tag{3.12}$$

where, $F_{\mathrm{h}_\ell}(\cdot)$ represents the CDF for the fading channel gain $|h_\ell|^2$. Transforming (3.12) in terms of Rayleigh fading channel it becomes an exponential which is expressed as

$$R_{\mathrm{s},\ell} = R_\ell(1-\tau_0)\exp\bigg(-\frac{\lambda d_\ell^\alpha \sigma^2(2^{R_\ell}-1)}{\kappa P_\ell}\bigg), \tag{3.13}$$

where, λ is the fading parameter.

3.3 Problem Formulation and Solution

In order to develop a fair utilization of the unlicensed channel for Wi-Fi and LAA networks, we consider two levels of fairness: *(a) the fairness between Wi-Fi and LAA networks*, and *(b) the fairness among LAA stations.*

3.3.1 Fairness Between Wi-Fi and LAA

Fairness between both the networks is maintained by sharing the channel access opportunity between them. During Wi-Fi and LAA coexistence scenario ($0 < \tau_0 < 1$), we observe that Wi-Fi stations will have a very low data rate if LAA stations aggressively use the unlicensed channel. This requires appropriate resource-sharing opportunities such that Wi-Fi

stations' performance is not compromised more than a minimum threshold and LAA stations get the chance to make the most out of the allotted time. We capture this by finding the favourable value of the sharing parameter, τ_0, which ensures that Wi-Fi stations can maintain an achievable average throughput per station, $\tilde{R}_w$, calculated as [134]

$$\tilde{R}_w = R_w(\tau_0) = \tau_0 \cdot R_w^{\max}, \tag{3.14}$$

and varies between

$$R_w^{\min} \leqslant R_w(\tau_0) \leqslant R_w^{\max} \tag{3.15}$$

during the time $\tau_0 T$. Here, $R_w^{\max}$ is calculated from (3.1) when $n = W$, i.e., the maximum achievable rate for a Wi-Fi-only system when the channel is not being shared by LAA. When Wi-Fi shares the channel, it gives a fraction of the opportunity to LAA stations, thus, it is expected to maintain a minimum throughput of $R_w^{\min}$.

3.3.2 Proportional Fairness Among LAA Stations

The time-division access network setting for Wi-Fi and LAA allows disjoint design for both kinds of stations. This opportunity is exploited in our work. In particular, an optimal transmission strategy is designed for the LAA stations. Once the Wi-Fi performance thresholds are met, the objective is to enhance the performance of LAA by devising a transmission control policy for the LAA transmission period $(1 - \tau_0)T$. It is easy to notice that considering equal access probability, τ_ℓ, for all the LAA stations, the stations at a greater distance from the LAA eNB will have more outage probability. Consequently, their average throughput will be low as compared to those at a smaller distance from eNB. Therefore, an equal access probability design results in unfairness among LAA stations.

We consider distinct access probability for each LAA station in order to incorporate the effect of wireless channel conditions. We also consider that, a station ℓ can transmit at an average power $P_\ell(1-\tau_0)$ which cannot be greater than the maximum transmit power allocated per station, $P_{\max}$. This can be given by the following inequality

$$P_\ell \tau_\ell (1 - \tau_0) \leqslant P_{\max}. \tag{3.16}$$

The overall goal is to improve the LAA network throughput while encouraging fairness among its stations. Unfair distribution of the resources will result in degraded performance of individual stations. The parameters which affect an LAA station performance are the fraction of channel utilization time, channel state, channel access probability, data rate, and transmitted power. Hence, an optimal transmission strategy based on the aforementioned parameters is required to enhance individual LAA station

performance (throughput) with fair consideration, which eventually will contribute to the overall network performance. In particular, the system objective is to be modelled in a particular fashion by considering the trade-off between the overall system performance and the level of fairness in terms of individual station performance.

In the literature, there are three popular system objectives for throughput/rate maximization, which differ in terms of the overall system performance and fairness among the stations [141]. These system objectives are (i) sum throughput maximization, (ii) min-max throughput maximization, and (iii) proportionally fairness throughput maximization.

The sum throughput maximization objective prioritizes the stations with better signal strength, thereby allocating more system resources to boost their throughput. As a result, the overall system throughput performance increases at the cost of throughput-unfairness among the stations. On the other hand, the max-min throughput maximization objective targets strict throughput fairness at the cost of reduced overall system throughput performance. The motivation behind the proportionally fair throughput maximization objective is to strike a balance between the system throughput and fairness among stations. This objective achieves some level of fairness among stations by providing each station with a performance that is proportional to its signal strength. This is achieved by reducing the opportunity of the stations with strong signal strength, getting a larger share of system resources compared to the weak stations. More system resources are allocated to the stations when their instantaneous channel condition is better relative to their channel statistics. Thereby, proportional fairness is achieved without compromising much throughput efficiency performance. Since the signal strength fluctuates independently for different stations, this strategy effectively exploits multi-user diversity. This is achieved by maximizing the sum of logarithmic throughput cost function of the individual stations [141–143], i.e., $\sum_{\ell=1}^{L} \log \tilde{R}_\ell$, where $\tilde{R}_\ell$ is defined in (3.9).

To address the above problem, we propose a transmission policy for LAA stations that maximizes their average throughput in a proportionally fair manner, when LAA and Wi-Fi stations coexist. This proposed transmission policy is given by the solution of the following optimization

problem

$$\begin{aligned} \underset{\tau_0,\tau_\ell,P_\ell,R_\ell}{\text{maximize}} \quad & \sum_{\ell=1}^{L} \log \tilde{R}_\ell \\ \text{subject to} \quad & C_1 : 0 < \tau_0 < 1, \\ & C_2 : 0 < \tau_\ell < 1, \quad \forall \ell \\ & C_3 : P_\ell \tau_\ell (1-\tau_0) \leqslant P_{\max}, \quad \forall \ell \\ & C_4 : R_w^{\min} \leqslant R_w(\tau_0) \leqslant R_w^{\max}, \quad \forall w. \end{aligned} \tag{3.17}$$

where $\tau_0, \tau_\ell, P_\ell$ and R_ℓ are the design variables. Here, the constraints C_1 and C_2 provide the range for the channel sharing parameter τ_0 (between LAA and Wi-Fi) and the probability of channel access for the LAA stations, respectively. The power allocation for each LAA station based on its channel access probability is ensured by condition C_3. Lastly, constraint C_4 accounts for the condition that LAA's coexistence with Wi-Fi does not compromise Wi-Fi's performance more than a Wi-Fi only system. This establishes inter-network fairness as well. The $\tilde{R}_\ell$ in the objective function is defined in (3.9). Using the basic calculus and algebraic calculations it can be shown that the problem in (3.17) is a non-convex optimization problem.

In order to solve the maximization problem in (3.17), we first present Lemma 3.3.1.

Lemma 3.3.1. *The optimal P_ℓ to maximize the objective function in problem* (3.17) *while satisfying the constraints is given as*

$$P_\ell = \frac{P_{\max}}{\tau_\ell(1-\tau_0)}. \tag{3.18}$$

Proof. Every station can transmit less than or equal to the maximum available power, $P_{\max}$. In order to maximize the objective function in problem (3.17), the constraint C_3 must attain the maximum available value of the average power, i.e., $P_\ell \tau_\ell (1-\tau_0) = P_{\max}$ which yields (3.18). ■

Based on Lemma 3.3.1, we remove P_ℓ as a design variable from (3.17) and plug in its value from (3.18) in the objective function of (3.17). The equivalent problem is then given by

$$\begin{aligned}\underset{\tau_0,\tau_\ell,R_\ell}{\text{maximize}} \quad & \sum_{\ell=1}^{L}\log\Bigg\{\frac{L\tau_\ell\prod\limits_{i\in\mathcal{U}}(1-\tau_i)}{\left(1-\prod\limits_{k=1}^{L}(1-\tau_k)\right)}R_\ell(1-\tau_0)\times \\ & \exp\left(-\frac{\lambda d_\ell^{\alpha}\sigma^2\tau_\ell(2^{R_\ell}-1)(1-\tau_0)}{\kappa P_{\text{av}}}\right)\Bigg\} \\ \text{subject to} \quad & C_1,\ C_2 \text{ and } C_4.\end{aligned} \tag{3.19}$$

Now we only have τ_0, τ_ℓ and R_ℓ as the design variables. Using the basic calculus and algebraic calculations it can be shown that the problem in (3.19) is a non-convex optimization problem. The solution to the optimization problem in (3.19) is given by the following theorem.

Theorem 3.3.2. *The optimal channel sharing parameter between Wi-Fi and LAA is given as*

$$\tau_0^* = \frac{R_w^{\min}}{R_w^{\max}}, \tag{3.20}$$

and the optimal rate of LAA station ℓ is given as

$$R_\ell^* = \hat{R}_\ell, \tag{3.21}$$

and the optimal channel access probability of the LAA station ℓ is given as

$$\tau_\ell^* = \hat{\tau}_\ell, \tag{3.22}$$

where $\hat{R}_\ell$ and $\hat{\tau}_\ell$ are obtained by simultaneously solving the following set of equations for all the LAA stations

$$\hat{R}_\ell = \frac{W_0\left(\frac{P_{\max}}{\lambda D_\ell\hat{\tau}_\ell(1-\frac{R_w^{\min}}{R_w^{\max}})}\right)}{\log(2)}, \quad \forall\ell, \tag{3.23}$$

$$\begin{aligned}&\frac{\hat{\tau}_\ell(L-1)\left(1-(1-\hat{\tau}_\ell)\prod\limits_{k\in\mathcal{U}}(1-\tau_k)\right)-(1-\hat{\tau}_\ell)\left(1-\prod\limits_{k\in\mathcal{U}}(1-\tau_k)\right)}{\hat{\tau}_\ell(1-\hat{\tau}_\ell)\left(1-(1-\hat{\tau}_\ell)\prod\limits_{k\in\mathcal{U}}(1-\tau_k)\right)} \\ &\qquad = -\frac{\lambda D_\ell(2^{\hat{R}_\ell}-1)(1-\frac{R_w^{\min}}{R_w^{\max}})}{P_{\max}}, \quad \forall\ell,\end{aligned} \tag{3.24}$$

where $W_0(\cdot)$ is the principal branch of the Lambert-W function and

$$D_\ell = \frac{d_\ell^\alpha \sigma^2}{\kappa}.$$

Proof. The optimization problem in (3.19) can be transformed into an equivalent problem written in the standard form as follows

$$\begin{aligned}
\underset{\tau_0,\tau_\ell,R_\ell}{\text{minimize}} \quad & -\sum_{\ell=1}^{L} \log \left\{ \frac{L\tau_\ell \prod_{i\in\mathcal{U}}(1-\tau_i)}{\left(1-\prod_{k=1}^{L}(1-\tau_k)\right)} R_\ell \times \right. \\
& \left. (1-\tau_0)\exp\left(-\frac{\lambda D_\ell \tau_\ell (2^{R_\ell}-1)(1-\tau_0)}{P_{\max}}\right)\right\} \qquad (3.25) \\
\text{subject to} \quad & 0<\tau_0, \quad \tau_0<1, \\
& 0<\tau_\ell, \quad \tau_\ell<1, \quad \forall \ell \\
& R_w^{\min} \leqslant \tau_0 \cdot R_w^{\max}, \quad \tau_0 \cdot R_w^{\max} \leqslant R_w^{\max}, \quad \forall w.
\end{aligned}$$

where,

$$D_\ell = \frac{d_\ell^\alpha \sigma^2}{\kappa}. \qquad (3.26)$$

Using the basic calculus and algebraic calculations it can be shown that the problem in (3.25) is a convex optimization problem.

The Lagrangian function for (3.25) can be given as

$$\begin{aligned}
\mathcal{L}(\tau_0,\tau_\ell,R_\ell,\boldsymbol{\mu}) = & -\sum_{\ell=1}^{L} \log \left\{ \frac{L\tau_\ell \prod_{i\in\mathcal{U}}(1-\tau_i)}{\left(1-\prod_{k=1}^{L}(1-\tau_k)\right)} R_\ell \times \right. \\
& \left. (1-\tau_0)\exp\left(-\frac{\lambda D_\ell \tau_\ell (2^{R_\ell}-1)(1-\tau_0)}{P_{\max}}\right)\right\} \\
& -\mu_1\tau_0 + \mu_2(\tau_0-1) - \mu_3\tau_\ell + \mu_4(\tau_\ell-1) + \\
& \mu_5(R_w^{\min} - R_w^{\max}\tau_0) + \mu_6 R_w^{\min}(\tau_0-1), \qquad (3.27)
\end{aligned}$$

where $\mu_i \in \boldsymbol{\mu} = \{\mu_1,\mu_2,\mu_3,\mu_4,\mu_5,\mu_6\}$ is the Lagrange multiplier corresponding to the ith constraint.

The Karush-Kuhn-Tucker (KKT) conditions for (3.25) are

$$
\begin{aligned}
&-\tau_0 < 0,\ \tau_0 - 1 < 0,\ -\tau_\ell < 0,\ \tau_\ell - 1 < 0, \\
&R_w^{\min} - R_w^{\max}\tau_0 < 0,\ R_w^{\max}(\tau_0 - 1) < 0, && \text{(3.28a)} \\
&\mu_1 \geqslant 0,\ \mu_2 \geqslant 0,\ \mu_3 \geqslant 0,\ \mu_4 \geqslant 0,\ \mu_5 \geqslant 0,\ \mu_6 \geqslant 0, && \text{(3.28b)} \\
&-\mu_1\tau_0 = 0,\ \mu_2(\tau_0 - 1) = 0,\ -\mu_3\tau_\ell = 0, \\
&\mu_4(\tau_\ell - 1) = 0,\ \mu_5(R_w^{\min} - R_w^{\max}\tau_0) = 0, \\
&\mu_6 R_w^{\max}(\tau_0 - 1) = 0, && \text{(3.28c)} \\
&\nabla_{\tau_0,\tau_\ell,R_\ell}\mathcal{L}(\tau_0,\tau_\ell,R_\ell,\boldsymbol{\mu}) = \left[\frac{\partial\mathcal{L}}{\partial\tau_0}\ \frac{\partial\mathcal{L}}{\partial\tau_\ell}\ \frac{\partial\mathcal{L}}{\partial R_\ell}\right]^\top \\
&\qquad\qquad = [\,0\ 0\ 0\,]^\top. && \text{(3.28d)}
\end{aligned}
$$

Here, ∇ is the gradient operator and $[\cdot]^\top$ represents the transpose of the matrix.

It can be shown that the problem in (3.25) is convex in τ_0 (there exists a global minima). From (3.28d) we have $\frac{\partial\mathcal{L}}{\partial\tau_0} = 0$. Taking the first derivative of (3.27) with respect to τ_0 and setting it equal to zero yields

$$
\sum_{\ell=1}^{L}\frac{1}{1-\tau_0} - \sum_{\ell=1}^{L}\frac{\lambda D_\ell\tau_\ell(2^{R_\ell}-1)}{P_{\max}} - \mu_1 + \mu_2 - R_w^{\max}\mu_5 + R_w^{\max}\mu_6 = 0. \tag{3.29}
$$

It can be shown that the problem in (3.25) is convex in τ_ℓ and R_ℓ, $\forall\ell$. It means there lies a global minima for every τ_ℓ and R_ℓ, respectively, $\forall\ell$. Taking the first derivative of (3.27) with respect to τ_ℓ and setting it equal to zero ($\frac{\partial\mathcal{L}}{\partial\tau_\ell} = 0$) yields

$$
\frac{L-1}{1-\tau_\ell} - \frac{1-\prod_{k\in\mathcal{U}}(1-\tau_k)}{\tau_\ell\left(1-(1-\tau_\ell)\prod_{k\in\mathcal{U}}(1-\tau_k)\right)} + \frac{\lambda D_\ell(2^{R_\ell}-1)(1-\tau_0)}{P_{\max}} - \mu_3 + \mu_4 = 0, \quad \forall\ell \tag{3.30}
$$

where $\mathcal{U}$ is a set of all LAA stations excluding the station ℓ.

Similarly, taking the first derivative of (3.27) with respect to R_ℓ and setting it equal to zero ($\frac{\partial\mathcal{L}}{\partial R_\ell} = 0$) yields

$$
-\frac{1}{R_\ell} + \frac{\lambda D_\ell\tau_\ell 2^{R_\ell}\log(2)(1-\tau_0)}{P_{\max}} = 0, \quad \forall\ell. \tag{3.31}
$$

For the complementary slackness conditions (3.28c) to be satisfied, either the constraints or the corresponding Lagrange multiplier should be

zero. Considering the case when the Lagrange multipliers $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_6 = 0$, $\mu_5 \neq 0$, i.e., when $\mu_1, \mu_2, \mu_3, \mu_4, \mu_6$ do not exist and μ_5 exists. It implies that the constraint $R_w^{\min} - R_w^{\max}\tau_0$ must follow the equality and be set to zero. These yields

$$\hat{\tau_0} = \frac{R_w^{\min}}{R_w^{\max}}. \tag{3.32}$$

Substituting the values of the Lagrange multipliers $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_6 = 0$, $\mu_5 \neq 0$, the expressions (3.29), (3.30) and (3.31) can be rewritten as

$$\mu_5 = \frac{L}{R_w^{\max}(1-\tau_0)} - \frac{1}{R_w^{\max}} \sum_{\ell=1}^{L} \frac{\lambda D_\ell \tau_\ell (2^{R_\ell} - 1)}{P_{\max}}, \quad \forall \ell, \tag{3.33}$$

$$\frac{\tau_\ell(L-1)\left(1-(1-\tau_\ell)\prod_{k\in\mathcal{U}}(1-\tau_k)\right) - (1-\tau_\ell)\left(1-\prod_{k\in\mathcal{U}}(1-\tau_k)\right)}{\tau_\ell(1-\tau_\ell)\left(1-(1-\tau_\ell)\prod_{k\in\mathcal{U}}(1-\tau_k)\right)} = -\frac{\lambda D_\ell (2^{R_\ell}-1)(1-\tau_0)}{P_{\max}}, \quad \forall \ell, \tag{3.34}$$

and

$$R_\ell = \frac{\mathrm{W}_0\left(\frac{P_{\max}}{\lambda D_\ell \tau_\ell (1-\tau_0)}\right)}{\log(2)}, \quad \forall \ell, \tag{3.35}$$

respectively. Here, $\mathrm{W}_0(\cdot)$ is the principal branch of the Lambert-W function. By setting $\tau_0 = \hat{\tau_0}$ and numerically solving (3.34) and (3.35) for R_ℓ and τ_ℓ, we get their values $\hat{R}_\ell$ and $\hat{\tau}_\ell$ as in (3.23) and (3.24), respectively. Plugging $\hat{R}_\ell$, $\hat{\tau}_\ell$ and $\hat{\tau_0}$ in (3.33) gives a positive solution of μ_5. It can be shown that $\hat{R}_\ell$, $\hat{\tau}_\ell$ and $\hat{\tau_0}$ satisfy all the KKT conditions when the Lagrange multiplier μ_5 is positive and $\mu_1, \mu_2, \mu_3, \mu_4, \mu_6$ are zero. Therefore, it is the optimal solution for the problem in (3.25).

For all other cases, similar steps can be followed, and it can be shown that those cases violate one or more KKT conditions. Hence, for all other cases, the corresponding solution becomes invalid. ■

The insights from Theorem 3.3.2 are discussed in the following remarks.

Remark 3.3.3. *In order to account for the Wi-Fi performance posed by the constraint C4 in an optimization problem in* (3.19)*, it is observed*

Algorithm 1 *Proposed System Operation*

1: **Inputs:** System parameters given in Table 2.
2: **Initialize:** τ_0 following (20), P_l following (18), and compute LAA station parameters for the given channel realization following Theorem 1.
3: **Repeat** for each transmission slot T.
4: Each Wi-Fi station follows the standard CSMA/CA with BEB within Wi-Fi transmission duration, $\tau_0 T$.
5: After $\tau_0 T$, each LAA station l attempts channel access with probability τ_l and transmits its packet at a transmission rate R_l with transmit power level P_l within the LAA transmission duration $(1-\tau_0)T$.
6: Update LAA station parameters for the given channel realization following Theorem 1.

that the optimum value of the time-sharing variable between Wi-Fi and LAA in (3.20) *is keeping the Wi-Fi throughput to the minimum threshold. Accordingly,* (3.20) *ensures that each Wi-Fi station achieves at least the minimum average throughput, and this will then allow the system to maximize the throughput of the LAA stations. As the ratio* $\frac{R_w^{\min}}{R_w^{\max}}$ *is always less than 1, therefore a fair chance of channel utilization is provided to LAA as well. Note that we are not prioritizing Wi-Fi's performance improvement as our goal lies in the overall LAA throughput improvement without significantly impacting Wi-Fi's performance.*

Remark 3.3.4. *The solution of* $\hat{R}_\ell$ *and* $\hat{\tau}_\ell$ *(by solving* (3.23) *and* (3.24)*) exists, when the constraint for* τ_ℓ *in the optimization problem* (3.25) *is a strict inequality. It is interpreted as* $0 < \hat{\tau}_\ell < 1$ *for each* ℓ *station of LAA. This is a practical condition for a coexistence scenario.*

The proposed scheme for WiFi and LAA coexistence is summarized in the Algorithm 1.

3.4 Results and Discussion

In this section, we evaluate the solution proposed in Section 3.3 using numerical simulations. Without loss of generality, we consider $L = 2, 6, 10, 14, 18$ LAA stations coexisting with $W = 5, 10$ Wi-Fi stations. According to the 3GPP standard modelling [139], an LAA eNB model allows the distribution of the stations within a radius of 40 m. Therefore, among the L stations of LAA, we populate one half at the radius of $r_1 = 5$ m and the other half at the radius of $r_2 = 30$ m from the eNB in order to include the near-far effect. The other parameter values are summarized in Table 3.2.

The optimal solution for the above simulation values is obtained after solving (3.20), (3.23) and (3.24) in MATLAB. The optimum numerical

Table 3.2: Simulation parameters and their values.

Parameter	**Value**	**Parameter**	**Value**
θ	9 μs	f	5 GHz
m	6	λ	1
$\mathrm{CW}_{\min}$	16	α	4
R_{b}	1 Mbps	σ^2	-90 dBm
SIFS	16 μs	$P_{\max}$	30 dBm
DIFS	34 μs	D	8184 bits
δ	14 μs	ACK + H	240 + 400 bits

values of τ_0^*, R_ℓ^* and τ_ℓ^* are then used to calculate the individual average throughput and sum average throughput of the LAA network.

3.4.1 Settings for the Schemes Compared

For comparative performance analysis with the proposed scheme we consider multiple existing schemes. Note that the proposed scheme is novel and its comparison with the existing schemes in their original form would not be fair. Hence, we adapt some closely related existing schemes in the following to compare them with the proposed scheme.

LAA-LBT mechanism: We adapt the analytical model of the LAA-LBT mechanism for the heterogeneous LAA/Wi-Fi coexistence network scenario presented in [105] to our scenario. Similar to the proposed scheme, this scheme considers Wi-Fi and LAA stations to coexist over a single channel under saturated traffic. Different from the proposed scheme, this scheme does not consider non-overlapping contention times for Wi-Fi and LAA, thus it may experience both intra-network and inter-network collisions. Moreover, unlike the proposed scheme, the individual sum throughput for the LAA and Wi-Fi networks are calculated under ideal channel conditions.

Fairness-constrained coexistence scheme: In [110], a fairness constrained LAA/Wi-Fi coexistence scheme based on optimal tuning of the initial window sizes and LAA transmission opportunity was proposed. Similar to the proposed scheme, this scheme considers Wi-Fi and LAA stations to coexist over a single channel under saturated traffic and maintains the 3GPP notion of fairness for the Wi-Fi stations to not be affected more than an additional Wi-Fi. In contrast to our proposed scheme which considers a more realistic scenario based on channel fading and outage, the scheme in [110] does not consider any packet loss and assumes ideal channel conditions. Another differentiating factor from the proposed scheme is that this scheme does not consider non-overlapping contention times for Wi-Fi and LAA, thus it may experience both intra-network and inter-network collisions. We implement the fairness scheme considering case

2 of Theorem 1 in [110] to calculate the maximum sum rate for LAA stations.

Benchmark scheme: We also consider a more comparable scheme and refer to it as the benchmark scheme. This benchmark scheme is inspired by the idea presented in [94, 139] for Wi-Fi and LAA coexistence in which all the LAA stations have an equal probability of accessing the channel in a given time slot [94, 139], i.e., for each station ℓ, $\tau_\ell = 1/L$. We also set a fixed transmission rate for all the LAA stations over the whole transmission period, i.e., $R_\ell = R_{\mathrm{m}}$, $\forall \ell$, where the transmission rate R_{m} is the one that maximizes the average throughput per station at the average distance of two different radii from the eNB, i.e., $\frac{r_1+r_2}{2}$. In addition, the benchmark scheme uses the same transmission time distribution between Wi-Fi and LAA in terms of the time-sharing fraction τ_0. Thus, it maintains the fair share between Wi-Fi and LAA but does not consider the fairness among LAA stations. This is the differentiating factor between the benchmark and the proposed scheme.

Moreover, for the brevity of analysis we calculate the benchmark optimal values of τ_0, τ_ℓ and R_{m} similarly as Theorem 3.3.2. Later we plug in these values in (3.9) to calculate the benchmark value for average throughput per LAA station.

3.4.2 Throughput Comparison of the Proposed Scheme with Other Existing Schemes

We compare the LAA sum throughput of the network for the proposed scheme with the other state-of-the-art schemes, i.e., the benchmark scheme, the fairness constrained coexistence scheme in [110] and the conventional LAA-LBT coexistence mechanism in [105]. The LAA sum throughput is calculated as $\sum_{\ell=1}^{L} \tilde{R}_\ell$.

Fig. 3.2 plots the LAA sum throughput in bps/Hz versus the number of LAA stations, L, for $W = 5$ and 10 Wi-Fi stations for the four considered schemes. We can see that for all the schemes, the sum throughput increases as the number of LAA stations increases, which is to be expected.

First, we compare the proposed scheme with the benchmark scheme as both follow no inter-network collision mechanism. From Fig. 3.2, we can see that the proposed scheme provides a higher sum throughput compared to the benchmark scheme. More importantly, the results show that for a given number of Wi-Fi stations, as the number of LAA stations decreases, the relative gap between proposed and benchmark schemes increases. We quantify this gap in terms of the throughput gain of the proposed scheme over the benchmark scheme, defined as a ratio of (LAA sum throughput of the proposed scheme - LAA sum throughput of bench-

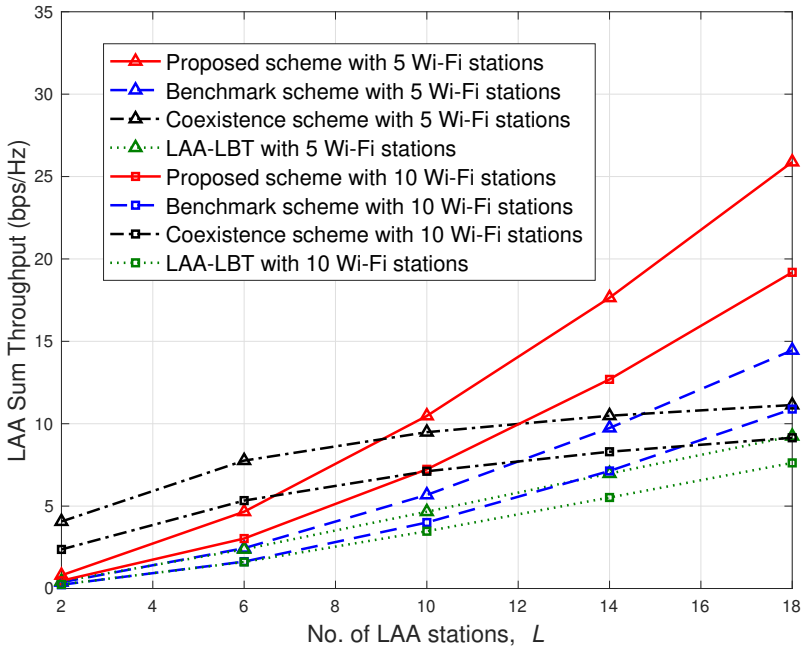
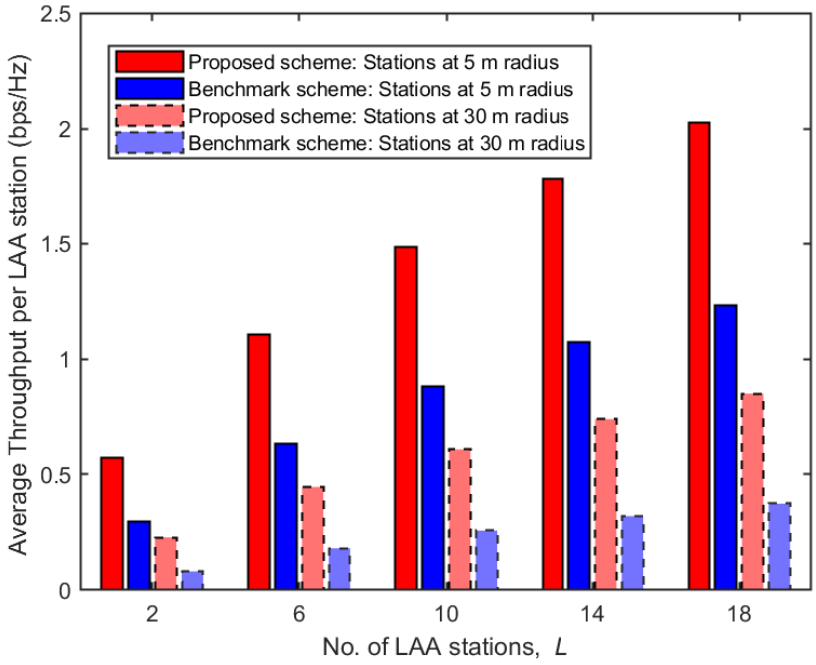


Figure 3.2: Sum throughput for LAA stations under the proposed and benchmark schemes.

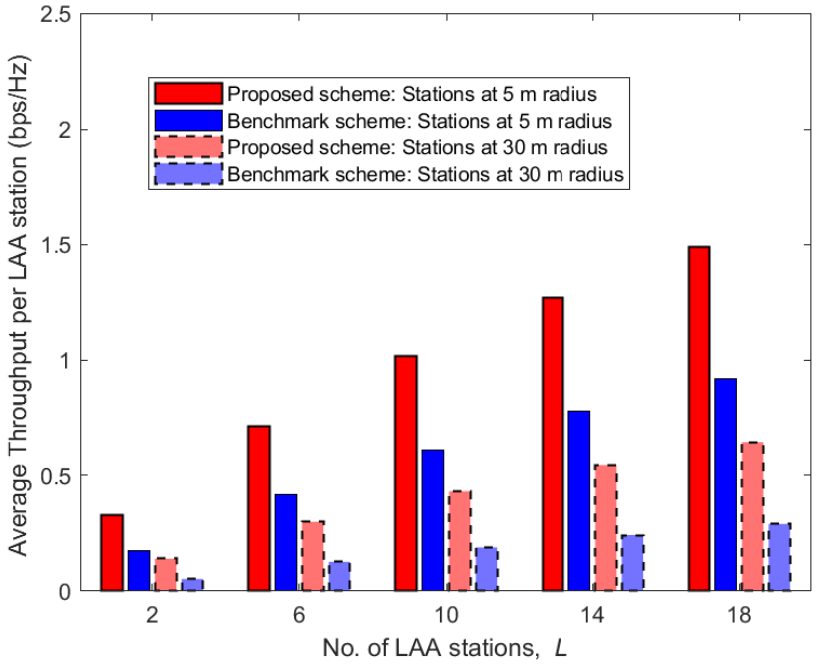
mark scheme)/(LAA sum throughput of benchmark scheme), expressed as a percentage. For example, when $W = 5$ and $L = 14$, the proposed scheme provides a throughput gain of approximately 81% over the benchmark scheme. As the number of Wi-Fi stations increases this gain does not decrease significantly. For instance, for $L = 14$, the throughput gain only decreases slightly from 81% for $W = 5$ to 79% for $W = 10$.

Secondly, we compare the proposed scheme to the LAA-LBT scheme [105] and the fairness constrained coexistence scheme [110]. These schemes suggest operating LAA and Wi-Fi simultaneously which may cause inter-network and intra-network collisions. In Fig. 3.2 we see that the proposed scheme outperforms the LAA-LBT coexistence mechanism. Fig. 3.2 shows that the sum throughput for the fairness constrained scheme is greater than our proposed scheme when LAA stations are lesser in number. However, when the number of LAA stations increases, our proposed scheme provides a better throughput performance. This is because when the number of LAA stations coexisting with the Wi-Fi increases, the inter-network and intra-network collisions increases for the model in [110] and the sum throughput saturates. Thus, our proposed scheme with zero inter-network collisions provides a scalable solution under practical conditions of higher number of coexisting users.

Next, we look at the average throughput per LAA station for the pro-



(a) Coexistence with 5 Wi-Fi stations.



(b) Coexistence with 10 Wi-Fi stations.

Figure 3.3: Average throughput per LAA station located at 5 m and 30 m radius from the eNB.

posed and the benchmark scheme. Fig. 3.3 plots the average throughput per LAA station versus number of LAA stations, L, coexisting with (a) $W = 5$ Wi-Fi stations and (b) $W = 5$ Wi-Fi stations. For each value of the number of LAA stations, we assume that half are located randomly at a radius of 5 m and the other half are randomly located at a radius of 30 m. We can see that the average throughput per LAA station is higher for the proposed scheme, compared to the benchmark scheme. For the number of LAA stations in the range 2-18, the increase in the average throughput of an LAA station for the proposed scheme compared to the benchmark scheme is in the range of 88% to 180% for $W = 5$ (Fig. 3.3a) and 87% to 167% for $W = 10$ (Fig. 3.3b). In addition, the LAA stations at the greater distance (30 m) show more throughput improvement compared to the ones at the closer distance (5 m). This indicates that proportional fairness is achieved among the LAA stations. The fairness is examined in more detail in the next section.

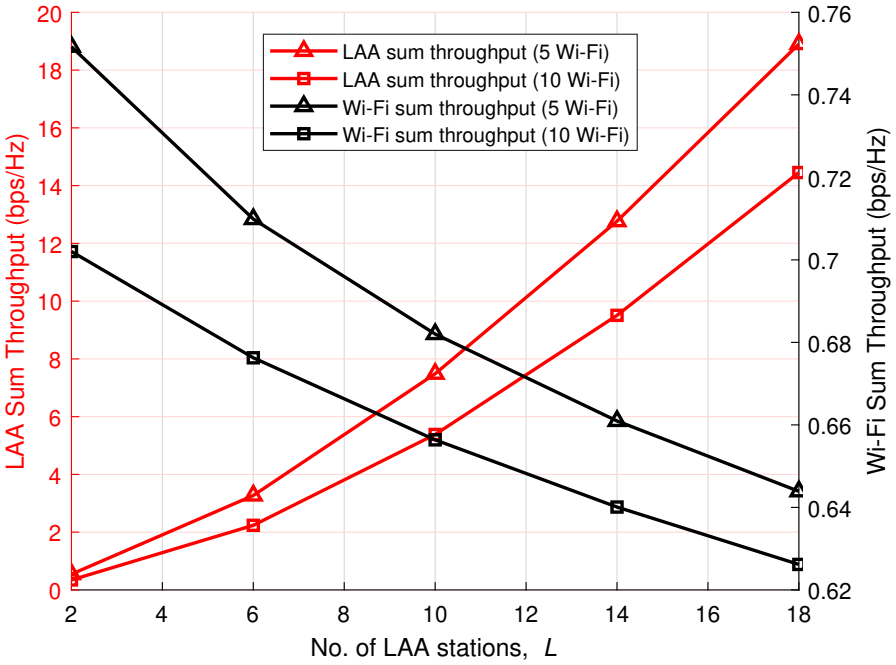
3.4.3 Inter-network and Intra-network Fairness

First, we look at the fairness between Wi-Fi and LAA networks. Fig. 3.4 plots the sum throughput of LAA stations and the sum throughput of the Wi-Fi stations versus the number of coexisting LAA stations, L, for (a) the proposed scheme and (b) the benchmark scheme. We can see that as the number of LAA stations increases, the sum throughput of the LAA stations increases and the sum throughput of Wi-Fi stations decreases for both the schemes. This is because the transmission time is being shared among the Wi-Fi and LAA stations. Although both the benchmark and proposed schemes maintain a fair coexistence with Wi-Fi by not sabotaging Wi-Fi throughput more than the minimum threshold, the proposed scheme provides enhanced throughput performance for the LAA, in Fig. 3.4a as compared to the benchmark scheme in Fig. 3.4b.

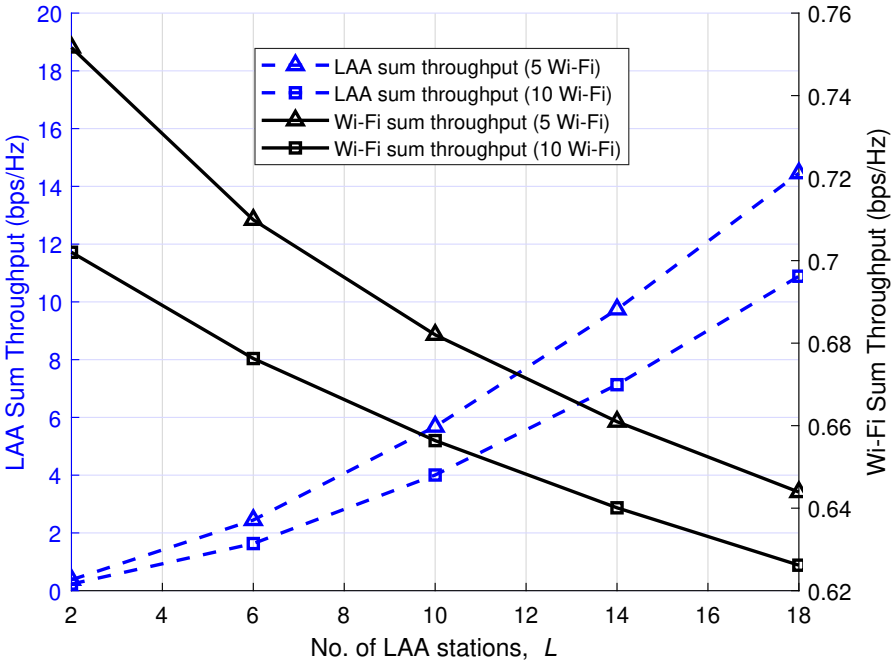
Next, we examine the intra-network fairness among the LAA stations. The fairness of the LAA system is determined in terms of Jain's fairness index (J_{I}) [144] as

$$J_{\mathrm{I}} = \frac{\left(\sum_{\ell=1}^{L} \tilde{R}_\ell\right)^2}{L \sum_{\ell=1}^{L} \tilde{R}_\ell^2}. \tag{3.36}$$

Fig. 3.5 plots the Jain's fairness index in (3.36) versus the number of LAA stations, L, for the proposed and benchmark schemes. The four different curves represent the two cases of coexistence scenarios for the proposed scheme and the benchmark scheme with $W = 5$ and $W = 10$ Wi-Fi stations. The results show that for both cases, there is a notable gain of $8 - 9\%$ in the fairness index for the proposed scheme is compared



(a) Performance for the proposed scheme.



(b) Performance for the benchmark scheme.

Figure 3.4: Sum throughput for Wi-Fi and LAA coexisting on a single unlicensed channel.

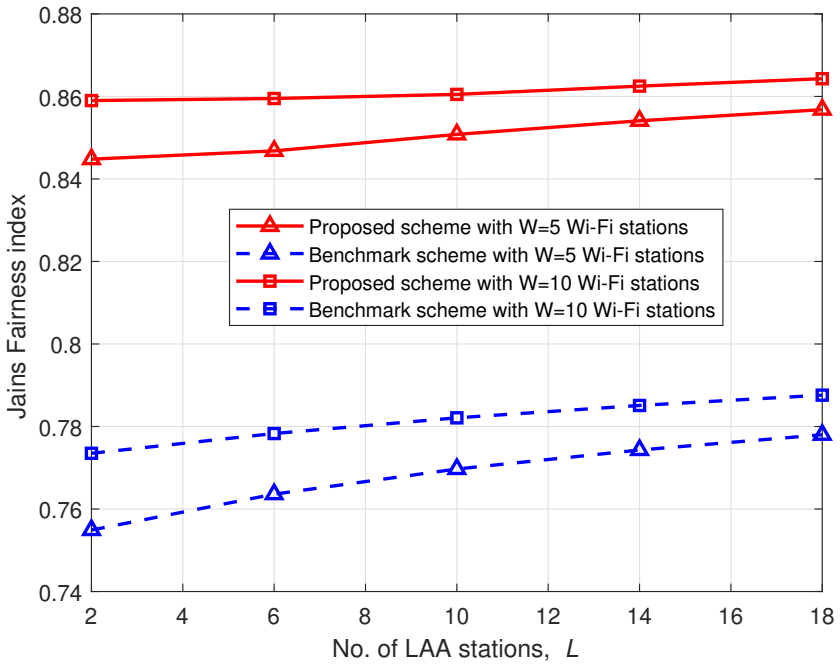


Figure 3.5: Jain's Fairness index for the LAA under the proposed and benchmark scheme.

to the benchmark scheme. In addition, the proposed system is fairer when the number of Wi-Fi stations increases from 5 to 10. This can be explained as follows. When the number of Wi-Fi stations is increased, the average throughput per LAA station reduces resulting in a decreased value of $L\sum_{\ell=1}^{L}\tilde{R}_{\ell}^{2}$, hence an increased value of J_{I}. In other words, the greater number of stations coexisting, the more fairly resources are distributed among them. This demonstrates the advantage of the proposed scheme.

3.4.4 Optimal Channel Sharing Parameter: Impact of LAA on Wi-Fi.

We can see the impact of a number of LAA stations on the Wi-Fi network performance in terms of the channel sharing parameter. Fig. 3.6 plots the optimal channel sharing parameter, τ_0 in (3.20) versus the number of LAA stations, L, coexisting with $W = 5$ and $W = 10$ Wi-Fi stations. The results show that for a fixed number of LAA stations when the number of Wi-Fi stations increases, τ_0 increases. A larger fraction of time is provided for the Wi-Fi network to accommodate an increased number of Wi-Fi stations. In addition, it can also be seen that the percentage drop in the value of τ_0 for an increasing number of LAA stations is more

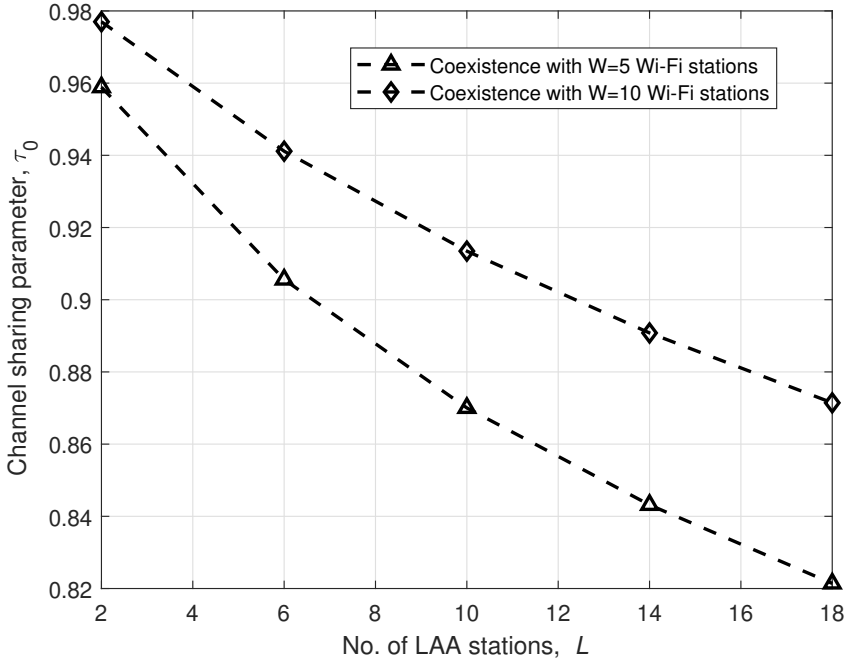


Figure 3.6: Channel sharing parameter, τ_0 for LAA and Wi-Fi coexistence network.

when LAA is coexisting with $W = 5$ Wi-Fi stations as compared to the case when it is coexisting with $W = 10$ Wi-Fi stations. The lesser congested Wi-Fi network has a larger capacity to accommodate LAA stations while maintaining its own network throughput to a minimum bearable threshold.

3.5 Conclusion

In this chapter, a fair coexistence scheme is proposed for the Wi-Fi and LAA networks. The proposed mechanism incorporates the inter-network fairness between the Wi-Fi and LAA networks, and the intra-network fairness among the LAA system. The core new idea of the proposed scheme is to devise different transmission probabilities for each LAA station dependent upon the wireless channel conditions and fair resource utilization, rather than the conventional coexistence approach based on uniform transmission probabilities.

We design a non-overlapping transmission policy in which Wi-Fi throughput tolerance is tuned through the optimum time-sharing fraction between Wi-Fi and LAA. In addition, we formulate a joint optimization problem in order to maximize the sum throughput of the LAA network with a performance guarantee for the Wi-Fi network in terms of fair throughput

share. The analytical solution provides the optimal design parameters which are validated through a comparison of proposed scheme with the state-of-the-art coexistence scheme. Our results show that, the proposed scheme ensures a significantly high LAA sum throughput while maintaining notable proportional fairness among its stations. The throughput performance gain is profound for the far LAA stations. On the other hand, Wi-Fi network throughput is also maintained to the minimum threshold.

This work can easily be extended for alternative approaches, for example when Wi-Fi is prioritized over LAA. Moreover, access probabilities of both Wi-Fi and LAA can also simultaneously be optimized to increase overall throughput of the systems.

Chapter 4

Reinforcement Learning-based Unlicensed Spectrum Sharing for IoT Devices of 5G New Radio.

4.1 Introduction

In the current era of Wireless technology, IoT devices have exhibited a tremendous increase. The 5G NR and beyond architectures are focusing on the large deployment of IoT networks [145]. This overwhelming growth of IoT devices once again raises the grave issue of bandwidth deficiency for radio access. Recently, the idea of sharing the unlicensed bands for 5G NR is gaining attention to alleviate the problem [5]. There is also a popular trend towards operating IoT devices for 5G-NR in unlicensed spectrum [52]. It resolves the licensed radio scarcity and also maintains QoS for the IoT networks [146]. However, unlicensed spectrum sharing gives birth to its own challenges of fair and efficient coexistence among diverse devices [97].

Typically, the unlicensed spectrum is occupied by Wi-Fi devices that utilize random access for radio services [147]. In order to coexist with the Wi-Fi devices, the secondary unlicensed devices follow a coexistence mechanism. Without a coexistence protocol among these heterogeneous networks, the secondary devices can heavily damage the performance of the primary Wi-Fi network [148]. In recent years, various coexistence technologies have evolved and been studied e.g., LTE-U, LAA, and MulteFire [90]. Lately, the 5G-NR and beyond has proposed to use LAA as a coexistence scheme for leveraging the unlicensed spectrum [4]. LAA is fundamentally a Wi-Fi-friendly protocol that follows an LBT mechanism to mimic Wi-Fi's CSMA/CA or DCF mechanism and creates a harmonious coexistence [94].

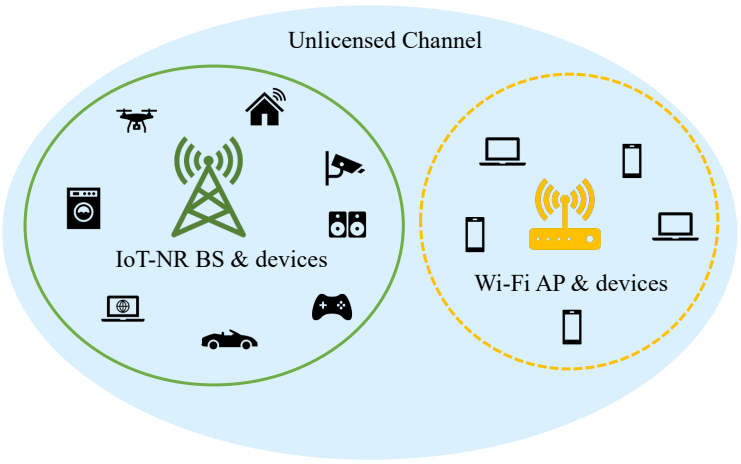


Figure 4.1: The coexistence of heterogeneous networks i.e., IoT-NR and Wi-Fi, over the unlicensed channel.

The coexistence of an IoT-NR network with a Wi-Fi network in an unlicensed channel is shown in Fig. 4.1. The IoT-NR network comprises an IoT base station (BS) and diverse IoT-NR devices. Meanwhile, the Wi-Fi network consists of a Wi-Fi AP and Wi-Fi devices. Both the networks follow their respective contention-based processes to randomly access and utilize the channel. During channel contention, the devices are prone to collide with each other and affect the capacity performance of their own network as well as the other network. Therefore, it is necessary to develop a policy for fair channel sharing [149].

Although, the major concern is a fair share of resources between the two heterogeneous networks. On the other hand, it is also important to take the diverse requirements of individual networks into consideration. Wi-Fi is the primary network is expected to not get affected while sharing the spectrum with other diverse devices. The IoT-NR devices such as displayed in Fig. 4.1 may possess diverse features in regard to their variable transmission times, data rates, security, or delays [150]. Hence, the diversity among IoT-NR devices must be paid attention to ensure efficient and fair inter-network and intra-network coexistence [151].

Furthermore, as 5G NR is a road map toward the emerging 6G communication, hence, the enhanced protocols of existing unlicensed radio access technologies such as LAA are expected to operate for the future coexistence frameworks [133]. Based on the advancement beyond 5G, more compatible coexistence protocols are required for future unlicensed spectrum sharing networks [50]. In this work, we have focused on an enhanced LAA scheme considering the capacity boost and QoS of the Wi-Fi and diverse IoT-NR networks.

4.1.1 Related Work

Related works have focused on the analysis of different coexistence mechanisms for heterogeneous networks such as IoT, cellular, and Wi-Fi. The work in [152] proposed a cellular user aggregated coexistence in the licensed and unlicensed spectrum for IoT users. A random access mechanism is presented for load estimation and dynamic contention of cellular IoT in the unlicensed [153]. Another work in [154] adopted a dynamic spectrum access technique to provide autonomous interference mapping for IoT networks in unlicensed bands. The study in [155], compared Wi-Fi and 5G NR performance in the unlicensed band industrial IoT.

In addition, some works have identified the LBT mechanism for unlicensed band sharing. A load-based and frame-based channel access mechanisms for a 5G IoT network are presented in [156]. Authors of [157] have proposed an advanced LBT mechanism for IoT, cellular and Wi-Fi coexistence in the 5G-NR unlicensed band. Another work in [105] has evaluated the performance of LBT for cellular and Wi-Fi coexistence. In [158] a three-dimensional Markov chain model is designed to obtain optimal CW for fair coexistence of IoT with Wi-Fi network. The work in [159] discussed a two-step LTE-advanced random access procedure to support URLLC IoT communication in the unlicensed spectrum.

Moreover, recent studies have also considered machine learning as a promising technique to solve the coexistence among different networks. A federated learning model is proposed in [160] for industrial IoT for 5G unlicensed resource utilization. The work in [119] presented an RL-based estimation for adapting the occupancy time and power for Wi-Fi and LAA coexistence. Another work in [121] discussed a QL-based approach for an LTE-U and Wi-Fi network. The authors of [122] presented a DRL-based solution for an LAA network to learn the Wi-Fi traffic pattern and adapt its transmission time accordingly. The study in [161] proposed a neural network-based algorithm for a friendly coexistence between IoT and Wi-Fi. Similarly, an RL-based algorithm is used to provide centralized scheduling for an LTE-U-enabled IoT network in [162].

Some research works have discussed diversity within a secondary network of a coexistence model with a primary network. A coexistence model considering different transmission probability for each LAA station dependent upon the wireless channel conditions was presented in [1]. The authors of [51] presented a study on the coexistence of intra-network LAA-LBT devices for a dense health care scenario. A bargaining game-based unlicensed spectrum sharing algorithm was presented in [163] for different classes of IoT devices. A massive MIMO-based connectivity was analyzed for IoT networks with different requirements of a number of devices, power consumption, and reliability [164]. However, all these works have proposed a different approach than ours.

Additionally, a diverse IoT-NR network coexistence has not been ex-

plored much. Therefore, in this work, we cater to the diverse transmission needs of IoT-NR devices in terms of dynamically adapting their initial sensing duration (ISD) for contention with Wi-Fi devices. The principal idea of adaptive ISD is to reduce the collisions among the devices in order to maximize the overall resource utilization and keep Wi-Fi performance intact.

4.1.2 Contributions

This chapter focused on the problem of diverse IoT-NR devices sharing an unlicensed spectrum with Wi-Fi devices. A learning-based solution is given to improve the coexistence performance of both networks. The main contributions of this work are summarized as follows:

- A coexistence mechanism is proposed for diverse IoT-NR devices in an unlicensed spectrum based on the idea of dynamically adaptive instead of similar values of ISD while guaranteeing the Wi-Fi network performance.
- A mathematical framework is constructed and an optimization problem is formulated to address the throughput performance of the proposed model.
- A QL framework is modelled to learn from environment dynamics and to find the optimum ISD for each IoT-NR device in order to maximize the sum throughput.
- The proposed model confirms a 51% maximum gain in the normalized sum throughput as compared to the benchmark scheme of fixed ISD.

The rest of the paper is structured as follows. The detailed coexistence mechanism, system model, and problem are explained in Section 4.2. Section 4.3 provides the proposed QL-based solution and algorithm. The simulation results and conclusion are presented in Section 4.4 and Section 4.5 respectively. Table 4.1 gives the summary of important variables and parameters.

4.2 System Model

In this section we elaborate on the system model for the proposed Wi-Fi and IoT-NR coexistence scheme. Later the mathematical expressions are designed to calculate the throughput of the network. In the end, the optimization problem is discussed.

Table 4.1: List of important parameters and variables.

Symbol	Description
W	Number of Wi-Fi devices.
N	Number of IoT-NR devices.
n	Index of the IoT-NR device, where $n = 1, 2, ..., N$.
σ	Idle slot time duration for Wi-Fi and IoT-NR device
m	Maximum backoff stage.
$CW_{\min}$	Initial contention window size
$T_{\text{ISD},n}$	The adapted ISD for the nth IoT-NR device.
$T_{\text{ISD}}^{\min}$	The minimum value of $T_{\text{ISD},n}$.
$T_{\text{ISD}}^{\max}$	The maximum value of $T_{\text{ISD},n}$.
$R_{\text{b,W}}$	The bitrate of the channel for a Wi-Fi device.
$R_{\text{b,NR}}$	The bitrate of the channel for an IoT-NR device.
T_{c}^{W}	The duration of collision for a Wi-Fi device.
T_{s}^{W}	The duration of successful transmission for a Wi-Fi device.
$T_{\text{c},n}^{\text{NR}}$	The duration of collision for the nth IoT-NR device.
$T_{\text{s},n}^{\text{NR}}$	The duration of successful transmission for the nth IoT-NR device.
R_{Wo}	The normalized system throughput for Wi-Fi only network.
R_{W}	Average normalized throughput for each Wi-Fi device in the heterogeneous network.
$R_{\text{NR},n}$	Average normalized throughput for the nth IoT-NR device in heterogeneous network.
S_{T}	The normalized sum throughput of all Wi-Fi and IoT-NR devices.

4.2.1 Coexistence Model for Wi-Fi and IoT-NR Network

We consider a coexistence scenario among W Wi-Fi devices and N IoT-NR devices on a single unlicensed channel. Both the Wi-Fi and IoT devices are assumed to be in saturation and always contending to access the channel for data transmission. The Wi-Fi network follows the DCF mechanism to coexist with each other. Whereas the IoT devices follow the proposed NR LBT protocol to coexist with the Wi-Fi network. The DCF and LBT access behaviours are framed as two-dimensional discrete Markov chains elaborated in [140] and in [105] respectively. Following the DCF protocol, the Wi-Fi device senses the channel for a period of DIFS and waits for a period of SIFS to receive an ACK for successful transmission. In case of a collision, the device observes a BEB for a random counter value from the window size, CW. The counter is decremented after an idle slot, σ. The size of CW doubles every time the device faces retransmission until it reaches its maximum limit when it is reset to its

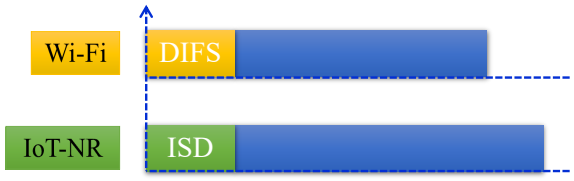


Figure 4.2: Initial contention for transmission between Wi-Fi and IoT-NR device over the unlicensed channel.

initial values. On the other hand, the LBT mechanism uses CCA to ensure a just sharing of unlicensed spectrum. It observes initial CCA or ISD and extended CCA to sense the channel before transmission and during a collision, respectively. Like DCF, it also experiences BEB upon collision and after K failed retransmission attempts the counters and CW are reset.

The initial contention for transmission between a Wi-Fi and an IoT-NR device over the unlicensed channel is illustrated in Fig. 4.2. We assume that different from the conventional LBT with fixed ISD our proposed NR-LBT has dynamically adaptive ISD, given as

$$T_{\text{ISD}}^{\min} \leq T_{\text{ISD},n} \leq T_{\text{ISD}}^{\max}, \quad \forall n. \tag{4.1}$$

where, $n = 1, 2, ..., N$ is the index of IoT-NR device, $T_{\text{ISD}}^{\min}$ and $T_{\text{ISD}}^{\max}$ are the minimum and maximum limits, respectively for $T_{\text{ISD},n}$ which is then adapted ISD for the nth IoT-NR device forming an analogy to Wi-Fi's DIFS. The proposed NR-LBT protocol is aimed to coexist with the Wi-Fi network by not impacting the performance latter network more than an additional Wi-Fi.

4.2.2 Analytical Sum Throughput of Wi-Fi and IoT-NR Devices

Performance of the system is expressed in terms of the amount of time the channel has served for the successful transmission of data, which is denoted as the normalized throughput R. We first describe the mathematical model for an independent Wi-Fi network and then for Wi-Fi and IoT-NR coexistence network.

Wi-Fi Only: The homogeneous Wi-Fi model in [1] is considered assuming that all the Wi-Fi devices have an equal probability of collision P_{wo}. For W Wi-Fi devices, the collision probability, P_{Wo} and the probability of a device transmitting in any slot, τ_{Wo} are given as

$$P_{\text{Wo}} = 1 - (1 - \tau_{\text{Wo}})^{W-1}, \tag{4.2}$$

and

$$\tau_{\mathrm{Wo}} = \frac{2(1-2P_{\mathrm{Wo}})}{(1-2P_{\mathrm{Wo}})(CW_{\min}+1) + P_{\mathrm{Wo}}CW_{\min}(1-(2P_{\mathrm{Wo}})^{m})}, \tag{4.3}$$

where, m is the backoff stage and $CW_{\min}$ is the initial contention window size. The solution of the nonlinear equations (4.2) and (4.3) is then used to find the normalized system throughput, R_{Wo} for Wi-Fi only situation given as

$$R_{\mathrm{Wo}} = \frac{P_{\mathrm{s}}P_{\mathrm{tr}}T_{\mathrm{tr}}}{(1-P_{\mathrm{tr}})\sigma + P_{\mathrm{s}}P_{\mathrm{tr}}T_{\mathrm{s}}^{\mathrm{W}} + P_{\mathrm{tr}}(1-P_{\mathrm{s}})T_{\mathrm{c}}^{\mathrm{W}}} \tag{4.4}$$

where, σ is the duration of an idle slot and T_{tr} is the channel occupancy time for a packet transmission. Besides, the average period when channel is found busy because of collision and successful transmission are $T_{\mathrm{c}}^{\mathrm{W}}$ and $T_{\mathrm{s}}^{\mathrm{W}}$ respectively. $P_{\mathrm{tr}} = 1-(1-\tau_{\mathrm{Wo}})^{W}$ is the probability accounting for at least a single transmission in the examined time slot and $P_{\mathrm{s}} = (W\tau_{\mathrm{Wo}}(1-\tau_{\mathrm{Wo}})^{W-1})(P_{\mathrm{tr}})^{-1}$ is the successful transmission probability conditioned on the event that at least one device transmits. According to the basic DCF mechanism, $T_{\mathrm{c}}^{\mathrm{W}}$ and $T_{\mathrm{s}}^{\mathrm{W}}$ are given as

$$T_{\mathrm{c}}^{\mathrm{W}} = \frac{H}{R_{\mathrm{b,W}}} + T_{\mathrm{tr}} + DIFS + \delta, \tag{4.5}$$

and

$$T_{\mathrm{s}}^{\mathrm{W}} = \frac{(H+ACK)}{R_{\mathrm{b,W}}} + SIFS + 2\delta + T_{\mathrm{tr}} + DIFS, \tag{4.6}$$

where, $R_{\mathrm{b,W}}$ is the bitrate of the channel for a Wi-Fi device and δ is the propagation delay. Header, H and ACK bits are configured from MAC layer description of Wi-Fi frame.

Wi-Fi and IoT-NR: To evaluate the performance of coexisting Wi-Fi and IoT-NR devices the analytical model of a heterogeneous network in [105] is followed. As described in [105], the collision probability of a Wi-Fi device is given as

$$P_{\mathrm{W}} = 1-(1-\tau_{\mathrm{W}})^{W-1}(1-\tau_{\mathrm{NR}})^{N} \tag{4.7}$$

and that of an NR device is given as

$$P_{\mathrm{NR}} = 1-(1-\tau_{\mathrm{NR}})^{N-1}(1-\tau_{\mathrm{W}})^{W}. \tag{4.8}$$

Moreover, the probability that a Wi-Fi device transmits in a heterogeneous network is given as

$$\tau_{\mathrm{W}} = \frac{2(1-2P_{\mathrm{W}})}{(1-2P_{\mathrm{W}})(CW_{\min}+1) + P_{\mathrm{W}}CW_{\min}(1-(2P_{\mathrm{W}})^{m})} \tag{4.9}$$

and the probability of an IoT device transmitting is given as

$$\tau_{\mathrm{NR}} = \frac{V}{X+Y+Z}, \tag{4.10}$$

where,

$$V = 2(1-2P_{\mathrm{NR}})(1-P_{\mathrm{NR}}+{P_{\mathrm{NR}}}^{K}-{P_{\mathrm{NR}}}^{m+K}), \tag{4.11a}$$
$$X = (1-P_{\mathrm{NR}})(1-2P_{\mathrm{NR}})(1+CW_{\min}(2P_{\mathrm{NR}})^{m}), \tag{4.11b}$$
$$Y = P_{\mathrm{NR}}^{K}(1-{P_{\mathrm{NR}}}^{m})(1-2P_{\mathrm{NR}}), \tag{4.11c}$$
$$Z = CW_{\min}(1-P_{\mathrm{NR}})(1-(2P_{\mathrm{NR}})^{m})(1-P_{\mathrm{NR}}+{P_{\mathrm{NR}}}^{K}), \tag{4.11d}$$

where, the parameter K denotes the re-transmission attempts allowed to NR device at the last backoff stage.

For the proposed coexistence network, the transmission probability that at least one of the W Wi-Fi devices and one of the N IoT-NR devices is transmitting in a given slot duration is given as $P_{\mathrm{tr}}^{\mathrm{W}} = 1-(1-\tau_{\mathrm{W}})^{W}$ and $P_{\mathrm{tr}}^{\mathrm{NR}} = 1-(1-\tau_{\mathrm{NR}})^{N}$, respectively. Also, the probability that only one of W Wi-Fi devices transmits and no NR device tries to contend is given as $P_{\mathrm{s}}^{\mathrm{W}} = (W\tau_{\mathrm{W}}(1-\tau_{\mathrm{w}})^{W-1}(1-\tau_{\mathrm{NR}})^{N})(P_{\mathrm{tr}}^{\mathrm{W}})^{-1}$. Similarly, the probability that only one of the N IoT-NR devices transmits and no Wi-Fi device tries to contend is given as $P_{\mathrm{s}}^{\mathrm{NR}} = (N\tau_{\mathrm{NR}}(1-\tau_{\mathrm{NR}})^{N-1}(1-\tau_{\mathrm{W}})^{W})(P_{\mathrm{tr}}^{\mathrm{NR}})^{-1}$. The average normalized throughput for each Wi-Fi and nth IoT-NR device (where $n = 1,\ldots,N$) are given as

$$R_{\mathrm{W}} = \frac{P_{\mathrm{s}}^{\mathrm{W}}P_{\mathrm{tr}}^{\mathrm{W}}T_{\mathrm{tr}}^{\mathrm{W}}}{T_{\mathrm{ch}}} \tag{4.12}$$

and

$$R_{\mathrm{NR},n} = \frac{P_{\mathrm{s}}^{\mathrm{NR}}P_{\mathrm{tr}}^{\mathrm{NR}}T_{\mathrm{tr},n}^{\mathrm{NR}}}{T_{\mathrm{ch},n}} \tag{4.13}$$

respectively. Here, $T_{\mathrm{tr}}^{\mathrm{W}}$ is the channel occupancy time for each Wi-Fi device transmission, $T_{\mathrm{tr},n}^{\mathrm{NR}}$ is the channel occupancy time for nth IoT-NR device transmission, $T_{\mathrm{ch},n}$ is the total channel time which integrates all the feasible scenarios among a Wi-Fi and nth IoT device and $T_{\mathrm{ch}} = \mathrm{mean}(T_{\mathrm{ch},n})$. $T_{\mathrm{ch},n}$ is given as

$$\begin{aligned} T_{\mathrm{ch},n} = {} & (1-P_{\mathrm{tr}}^{\mathrm{W}})(1-P_{\mathrm{tr}}^{\mathrm{NR}})\sigma + P_{\mathrm{tr}}^{\mathrm{W}}P_{\mathrm{s}}^{\mathrm{W}}(1-P_{\mathrm{tr}}^{\mathrm{NR}})T_{\mathrm{s}}^{\mathrm{W}} \\ & + P_{\mathrm{tr}}^{\mathrm{NR}}P_{\mathrm{s}}^{\mathrm{NR}}(1-P_{\mathrm{tr}}^{\mathrm{W}})T_{\mathrm{s},n}^{\mathrm{NR}} + P_{\mathrm{tr}}^{\mathrm{W}}(1-P_{\mathrm{s}}^{\mathrm{W}})(1-P_{\mathrm{tr}}^{\mathrm{NR}})T_{\mathrm{c}}^{\mathrm{W}} \\ & + P_{\mathrm{tr}}^{\mathrm{NR}}(1-P_{\mathrm{s}}^{\mathrm{NR}})(1-P_{\mathrm{tr}}^{\mathrm{W}})T_{\mathrm{c},n}^{\mathrm{NR}} + P_{\mathrm{tr}}^{\mathrm{W}}P_{\mathrm{s}}^{\mathrm{W}}P_{\mathrm{tr}}^{\mathrm{NR}}P_{\mathrm{s}}^{\mathrm{NR}}T_{\mathrm{i}} \\ & + P_{\mathrm{tr}}^{\mathrm{W}}P_{\mathrm{s}}^{\mathrm{W}}P_{\mathrm{tr}}^{\mathrm{NR}}(1-P_{\mathrm{s}}^{\mathrm{NR}})T_{\mathrm{i}} + P_{\mathrm{tr}}^{\mathrm{W}}(1-P_{\mathrm{s}}^{\mathrm{W}})P_{\mathrm{tr}}^{\mathrm{NR}}P_{\mathrm{s}}^{\mathrm{NR}}T_{\mathrm{i}} \\ & + P_{\mathrm{tr}}^{\mathrm{W}}(1-P_{\mathrm{s}}^{\mathrm{W}})P_{\mathrm{tr}}^{\mathrm{NR}}(1-P_{\mathrm{s}}^{\mathrm{NR}})T_{\mathrm{i}}. \end{aligned} \tag{4.14}$$

Eq. (4.14) takes into account the cross collision between Wi-Fi and IoT-NR devices where, $T_{\mathrm{i}} = \max\{T_{\mathrm{c}}^{\mathrm{W}}, T_{\mathrm{c},n}^{\mathrm{NR}}\}$ as explained in [106]. $T_{\mathrm{c}}^{\mathrm{W}}$ and $T_{\mathrm{s}}^{\mathrm{W}}$ can be obtained from (4.5) and (4.6) respectively. $T_{\mathrm{c},n}^{\mathrm{NR}}$ and $T_{\mathrm{s},n}^{\mathrm{NR}}$ represent the time period of the channel found busy when nth IoT-NR device undergoes collision and transmission. These are calculated as

$$T_{\mathrm{c},n}^{\mathrm{NR}} = \frac{H}{R_{\mathrm{b,\ NR}}} + T_{\mathrm{tr},n}^{\mathrm{NR}} + T_{\mathrm{ISD},n} + \delta, \tag{4.15}$$

and

$$T_{\mathrm{s},n}^{\mathrm{NR}} = \frac{(H + ACK)}{R_{\mathrm{b,NR}}} + T_{\mathrm{tr},n}^{\mathrm{NR}} + T_{\mathrm{ISD},n} + \delta, \tag{4.16}$$

where $R_{\mathrm{b,\ NR}}$ is the bitrate of the channel for an IoT-NR device. Furthermore, the sum normalized throughput of all Wi-Fi and IoT-NR devices is given as

$$S_{\mathrm{T}} = WR_{\mathrm{W}} + \sum_{n=1}^{N} R_{\mathrm{NR},n}. \tag{4.17}$$

4.2.3 Problem Formulation

In a multiple device coexistence network, collisions are often synchronized when the devices have a similar value of ISD. It causes all the collided devices to observe binary exponential backoffs before sensing the channel again. Thus, ISD plays a vital role in channel contention of IoT-NR devices coexisting with Wi-Fi devices. Adapting this period for each IoT-NR device will help collision reduction and avoid prolonged backoffs. This is because, with variable ISD rather than fixed, IoT-NR devices can choose to wait for more or less depending on the overall network performance and efficiently achieve their diverse transmission needs. It will not only benefit the IoT-NR devices but will also provide a mitigated intrusion for the Wi-Fi devices. Moreover, a Wi-Fi performance guarantee is needed to ensure it is not degraded beyond an additional Wi-Fi which is a Wi-Fi only network with total devices of $(W1 = W + N)$ which is calculated from (4.4) as R_{W1o} and the condition is given as

$$R_{\mathrm{W}} \geq R_{\mathrm{W1o}} \tag{4.18}$$

The objective is to find the optimum value of T_{ISD}, i.e., $T_{\mathrm{ISD}}{}^{*}$ for each IoT-NR device such that the total normalized sum throughput is maximized and Wi-Fi performance is ensured to a bearable threshold. These definitions constitute the maximization problem of network sum throughput which is formulated as

$$\begin{aligned} \underset{T_{\mathrm{ISD},n}}{\text{maximize}} \quad & S_{\mathrm{T}} \\ \text{subject to} \quad & C_1 : 0 \leq P_{\mathrm{W}} \leq 1, \\ & C_2 : 0 \leq P_{\mathrm{NR}} \leq 1, \\ & C_3 : T_{\mathrm{ISD}}^{\min} \leq T_{\mathrm{ISD},n} \leq T_{\mathrm{ISD}}^{\max}, \quad \forall n \\ & C_4 : R_{\mathrm{W}} \geq R_{\mathrm{W1o}}. \end{aligned} \tag{4.19}$$

The argument in (4.19) stresses finding the optimal $T_{\mathrm{ISD},n}$ for individual IoT-NR device to maximize the network sum throughput. The first two constraints C_1 and C_2 refer to the probability of collision in a slot duration for each Wi-Fi and IoT-NR device. The third condition limits the value of $T_{\mathrm{ISD},n}$ for n^{th}, IoT-NR device to minimum $T_{\mathrm{ISD}}^{\min}$ and maximum $T_{\mathrm{ISD}}^{\max}$ feasible range. The last constraint C_4 ensures that the proposed solution doesn't degrade the Wi-Fi performance more than that of a standalone Wi-Fi network.

The formulated problem in (4.19) demands online adjustment of contention parameters based on real-time environment feedback i.e., diverse transmission requirements and overall network performance. Since there is no communication between Wi-Fi and IoT-NR devices it is extraordinarily complex to calculate R_{W} without them knowing the actual transmission times and sensing times of the other devices. RL proves to be a suitable tool for this nature of the problem as the objects actively learn and react to changing environments with the ultimate goal of reaching an optimum result [29]. Additionally, it is less computationally complex, more robust, and provides autonomous management [26].

4.3 The Proposed QL Scheme

QL has emerged as one of the major classes of RL algorithms [84]. It is a model-free and off-policy approach that helps machines learn without policy binding. Instead, it follows a self-optimization and decision-making process which falls in the Markovian domain. We propose a QL-based solution for the formulated problem in section 4.2.3. In the subsections below, initially, the fundamentals of a Markov decision process (MDP) for QL are illustrated to form the base of the proposed multi-agent MDP. It is then followed by a proposed QL algorithm to efficiently utilize the unlicensed spectrum.

4.3.1 MDP Formation

Our MDP is formulated as a tuple ($\mathcal{S}$, $\mathcal{A}$, $\mathcal{P}$, $\mathcal{R}$) where:

- $\mathcal{S}$ is a set of finite states such that a single state, $s \in \mathcal{S}$.

- $\mathcal{A}$ is a set of actions by the agents such that, $a \in \mathcal{A}$ represents the action of the agent.
- $\mathcal{P}\ (s,\ s') = P\ (s_{t+1} = s' \mid s_t = s,\ a_t = a\ )$ is the transition probability representing that the action a taken at state s at time t will result in the state s' at time $t+1$.
- $\mathcal{R}\ (s,\ s')$ is the received reward function to calculate the reward value by the agents when they transition from state s to s' after acting on action a.

The relationship among the entities of the above MDP is explained in terms of an agent-environment interface which is an ongoing process of observation and interaction. The learning agents in the Markov chain transit between different states with respect to their transition probabilities and consequently interact with the environment by observing actions. These actions earn rewards for them. Based on the reward attained in each step t, a long-term estimation of rewards leads to the convergence of the problem towards the optimum value. The value termed, Q-value, is determined from a function following a maximum or greedy policy for each iteration [30]. The Q-value function is represented as

$$Q(s_t, a_t) \leftarrow (1-\alpha)Q(s_t, a_t) + \alpha\left[r_{t+1} + \gamma \underset{a}{\max} Q(s_{t+1}, a)\right] \tag{4.20}$$

where $0 \leq \gamma \leq 1$ is the discount factor and $0 \leq \alpha \leq 1$ is the learning factor in the iterative process of value estimation. The Q-value at time instant t, $Q(s_t, a_t)$ is updated as a discounted value of the action a (which provides the maximum Q-value at time instant t+1, $Q(s_{t+1}, a)$) and the weighted experience (captured by α). Besides, r_{t+1} is the resulting reward at state s_t for undertaking action a_t to reach state s_{t+1}. The optimum Q-value then becomes

$$\lim_{t\to\infty} Q(s, a) = Q^*(s, a) \tag{4.21}$$

as the learning saturates.

4.3.2 Wi-Fi and IoT-NR Coexisting Network as Multi-Agent MDP

The proposed solution is structured into a multi-agent MDP in the following manner:

1. **Agent** The one actively learning and making decisions is the agent. Hence, each n^{th} IoT-NR device is the agent, $\forall n$. The agent senses the channel of Wi-Fi activity and approximates its transmission times based on the feedback of the environment to regulate their actions.

2. **Action** The action of the agent 'n' is to select the initial sensing duration T_{ISD}, from the range of [$T_{\mathrm{ISD}}^{\min}$, $T_{\mathrm{ISD}}^{\max}$].

$$T_{\mathrm{ISD}}^{\min} = \mathrm{SIFS} + \sigma, \tag{4.22}$$

and

$$T_{\mathrm{ISD}}^{\max} = \mathrm{SIFS} + j\sigma \tag{4.23}$$

where, $j = 1, 2, \ldots, J$ is the current state and J is the total number of finite states.

3. **State** The state of the agent n at a particular instant t will depend upon the value of T_{ISD}. Each state, s, represents the value for the range $[T_{\mathrm{ISD}}^{\min}, T_{\mathrm{ISD}}^{\max}]$ which an IoT-NR device can observe based on its transmission time, $T_{\mathrm{tr},n}^{\mathrm{NR}}$.

4. **Reward** The reward is the cost predicted at each stage of the self-learning procedure, which is the total sum normalized throughput. A positive response indicates the learning is heading in the right direction and vice versa. The value of reward at time instant t for a state transition is calculated as

$$r_t = \begin{cases} \Omega S_{\mathrm{T}}, & \text{if} \quad R_{\mathrm{W}} \geq R_{\mathrm{W1o}} \\ \eta S_{\mathrm{T}}, & \text{otherwise} \end{cases} \tag{4.24}$$

where Ω and η represent the positive and negative values of the reward factor at a given state, respectively.

The dynamic interaction between the agents and environment is recorded by an experience $\{s_t,\ a_t,\ s_{t+1},\ r_t\}$. The IoT-NR devices transition between the finite number of J states with a uniform probability. At the time instant t, when the IoT-NR device is in a state s_t, it can choose one of the feasible T_{ISD} values as the action a_t to transition into a next state of s_{t+1}. This action results in a reward r_t. The associated reward influences the learning process of the IoT-NR device to update its next action and state transitions.

4.3.3 Proposed QL Algorithm

The proposed QL algorithm to find out the optimal value of $T_{\mathrm{ISD},n}$ has been summarized in algorithm 2. The agents need initial values of the $Q(s, a)$, J, $T_{\mathrm{ISD}}^{\min}$ and $T_{\mathrm{ISD}}^{\max}$ $\forall j$, ε, γ, α, η and Ω. The values of the learning parameters γ and α are chosen very carefully, too small, or too large values can cause inefficient learning. Once initialized, the learning process begins for every agent by choosing an initial state and selecting an initial action. After implementing the action, its reward is calculated. The next state is assigned to the agent and the Q-value table is updated based on(4.20). The process continues until the Q-value converges to the optimal result.

Algorithm 2 Proposed QL Algorithm for Wi-Fi and IoT-NR Coexistence

Input: diverse transmission time $T_{\mathrm{tr},n}^{\mathrm{NR}}$, $\forall n$ and Wi-Fi standalone rate, R_{W1o}
print random number, $k \in (0,1)$, total number of states, J, $T_{\mathrm{ISD}}^{\min}$ and $T_{\mathrm{ISD}}^{\max}$ $\forall j$, ε, γ, α, η and Ω

for each agent $n = 1, 2, ..., N$
  let iteration index, $t = 0$ set initial Q-values $Q(s, a)$ for all $s_j \in \mathcal{S}$ and $a \in \mathcal{A}$ randomly choose an initial state to start with.
end for
Learning:
for all n
  if $k \leq \varepsilon$ **then**
    select action $a_{n,t}$ randomly
  else
    select $a_{n,t} = \max_{a^*} Q(s_t, a^*)$
  end if
  implement action $a_{n,t}$
  evaluate reward r_{t+1}
  go to state s_{t+1}
  update $Q_n(s_t, a_{n,t})$ according to (4.20)
end for
Keep iterating until Q-value converges

Output: optimal action $a = \{T_{\mathrm{ISD},n}\}$ $\forall n$, Q-value

Table 4.2: Summary of simulation parameters with their values.

Parameter	**Value**	**Parameter**	**Value**
DIFS	56 μs	SIFS	16 μs
m	4	$CW_{\min}$	16
σ	9 μs	$T_{\mathrm{tr}}^{\mathrm{W}}$	8 ms
H	400 bits	ACK	240 bits
$R_{\mathrm{b,W}}$	1 Mbps	$R_{\mathrm{b,NR}}$	100 Mbps
γ	0.8	α	0.01
Ω	50	η	-50

4.4 Numerical and Simulation Results

A real-time simulator is built to imitate the channel access mechanisms of Wi-Fi devices and NR devices. All the simulations have been performed in MATLAB. The analytical model presented in section 4.2.2 is simulated for an equal number of Wi-Fi and NR devices, i.e., $W = N = D/2$, where D is the total number of devices contending. Two different values of transmission times i.e., 4 ms and 8 ms are randomly assigned to each half of the IoT-NR devices to include diversity. The simulations are averaged over 100 realizations. Table 4.2 summarizes the simulation parameters

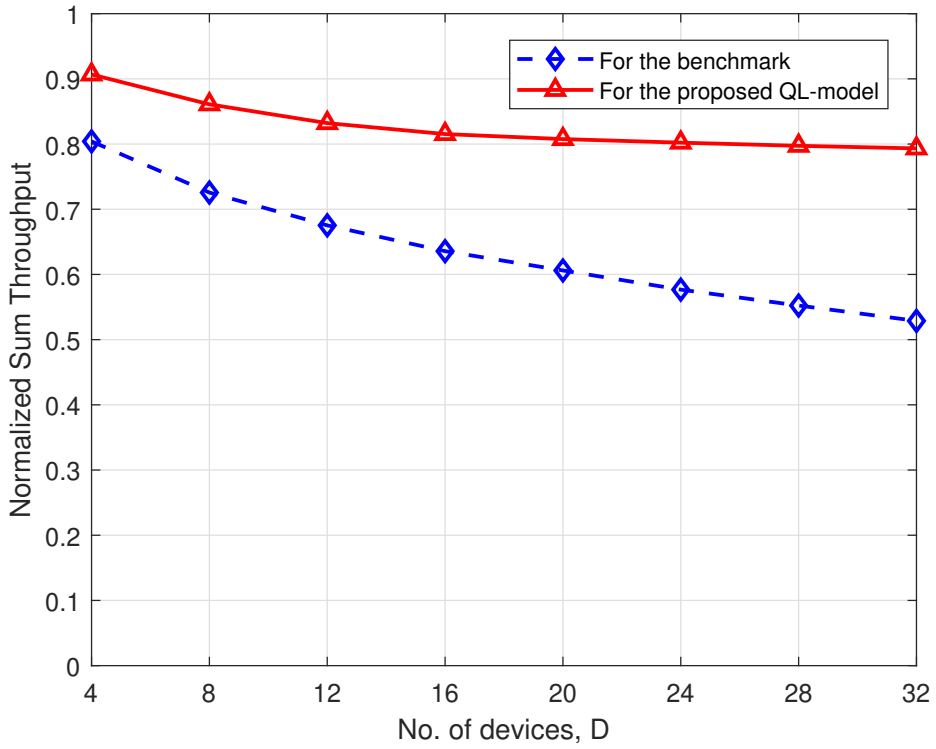


Figure 4.3: Normalized sum throughput of the Wi-Fi/IoT-NR coexistence network for the benchmark and the proposed model.

and their corresponding values.

For the benchmark scenario, it is assumed that the T_{ISD} of the IoT-NR devices is fixed and equivalent to Wi-Fi's DIFS similar to the work in [105]. The normalized sum throughput of the coexistence model is calculated for the benchmark and then used to compare with the proposed QL model.

For the proposed QL model we set $\varepsilon = 0.9$ initially to explore the new states and then the parameter value is reduced by a fraction of 0.99999 at every step till $\varepsilon \geq 0.5$, after that it is reduced by a fraction of 0.9999 until it reaches a minimum value of 0.01 or until the QL algorithm converges. We consider $J = 16$ states for the agents. Once the learning process is completed the optimal values are used to calculate the normalized sum throughput. The performance of the proposed QL model is elaborated in the following results.

4.4.1 Comparison of the Proposed Model with the Benchmark

Fig. 4.3 plots the normalized sum throughput of the Wi-Fi/IoT-NR coexistence network versus the number of devices, D for the benchmark and the proposed QL model. It can be observed that the sum throughput for

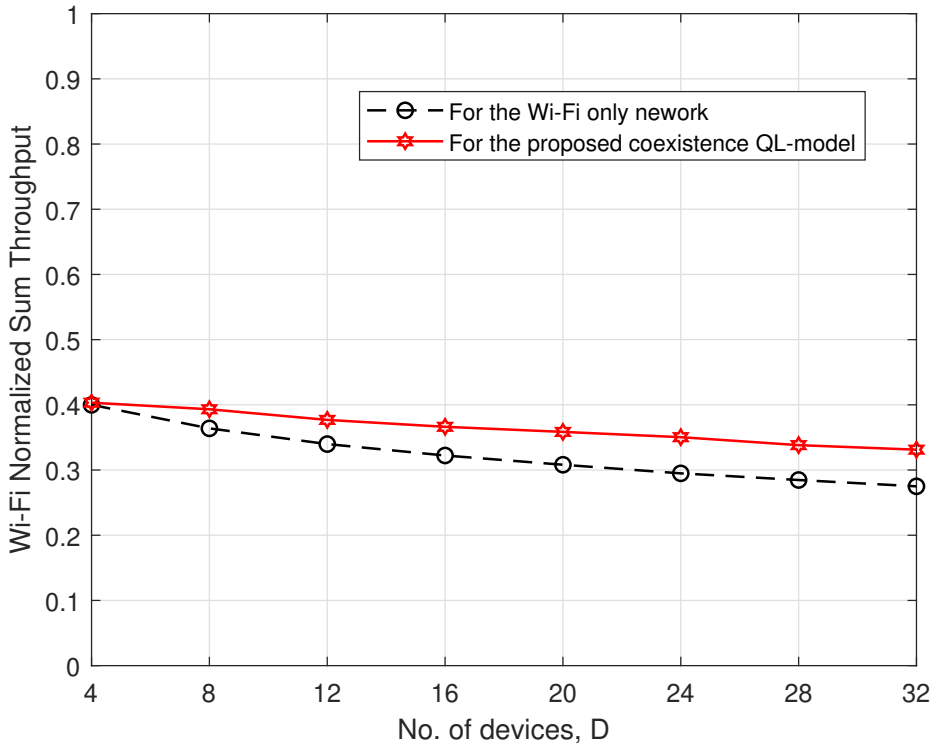


Figure 4.4: Normalized sum throughput for Wi-Fi standalone and Wi-Fi/IoT-NR Coexistence network.

both cases decreases as the number of coexisting devices increases. This is because a larger number of devices induces more collisions among each other hence, introducing longer backoffs. On the other hand, our proposed model obtains a maximum of 51% throughput gain as compared to the benchmark model. Also, the % gap increases as the number of devices increases indicating a better performance for densely populated networks.

4.4.2 Performance Guarantee for the Wi-Fi Network

Fig. 4.4 plots the normalized sum throughput for Wi-Fi standalone and Wi-Fi/IoT-NR coexistence network versus the number of devices, D. The plot verifies the constraint $C_4 : R_{\mathrm{W}} \geq R_{\mathrm{W1o}}$ of (4.19) for the proposed QL-based solution. It can be observed that the performance of the Wi-Fi network coexisting with the IoT-NR network is always greater than the standalone Wi-Fi network. Hence, the Wi-Fi network's performance remained intact and is never degraded beyond the threshold.

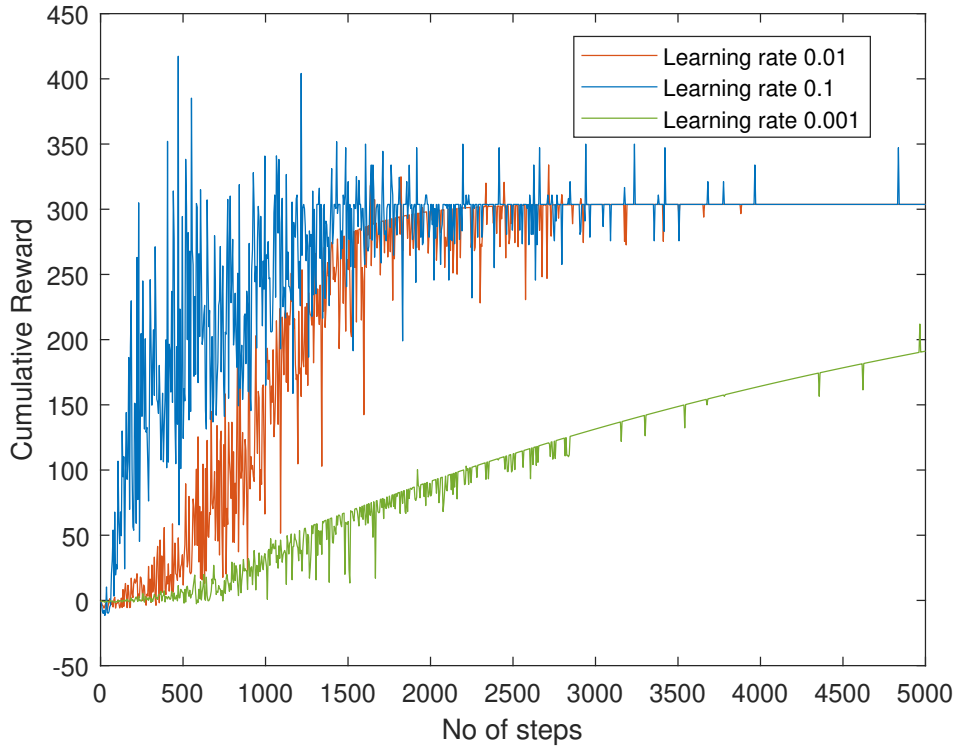


Figure 4.5: The accumulative reward for different learning rates for the proposed QL-model when D=12 devices.

4.4.3 Performance Evaluation of the Proposed QL Model

The sensitivity of the QL algorithm to its parameters is shown in Fig. 4.5, which plots the accumulative reward of the proposed QL-model versus the number of steps for three different values of learning rates i.e., $\alpha = 0.1$, 0.01 and 0.001. These results are obtained for the $D = 12$ number of devices. Fig. 4.5 shows that a low value of 0.001 takes too long for the algorithm to converge but with high precision. On the other hand, a higher learning rate of 0.1 makes the algorithm to converge too quickly but with less precision. Therefore, a suitable learning rate of 0.01, neither too large nor too small, is selected for the algorithm to converge.

4.5 Conclusion

In this work, an IoT-NR network was investigated to coexist with a primary Wi-Fi network in the unlicensed spectrum. An NR-LBT coexistence scheme was proposed which presented the idea of adaptive ISD over fixed ISD for IoT-NR devices based on their diverse transmission times. Established on the system model an optimization problem was formulated to

maximize the network's normalized sum throughput. To converge to an optimal solution for the maximization problem, a QL algorithm was provided. The results displayed a 51 % maximum gain in the normalized sum throughput of the proposed Wi-Fi/IoT-NR network as compared to the benchmark. Also, a performance guarantee was presented to the Wi-Fi network.

Chapter 5

Deep Reinforcement Learning-Driven Secrecy Design for Intelligent Reflecting Surface-Based 6G-IoT Networks

5.1 Introduction

Due to the rapidly growing number of IoT devices, there has been a never-quenching need for higher data rates and ultra-wide bandwidth capacity than there is in the present time. Both mmWave and THz bands are being explored as a major component of beyond 5G and 6G wireless communication networks [165, 166]. But higher frequencies confront their own challenges, such as signal attenuation, low coverage range, and signal blockages. Moreover, the increased connectivity of devices demands different levels of network security to address the diversity among IoT devices [54, 167].

Recently, IRS, also referred to as reconfigurable intelligent surface (RIS) [168], has emerged as a low-cost and power-efficient enhancement of wireless communication systems [169, 170]. These surfaces come with the property of reconfigurable features, such as amplitude, phase, and polarization of electromagnetic waves with frequencies ranging from microwaves to visible light [171]. With these features, IRS is expected to be a feasible option for a sustainable, energy-efficient architecture of 5G and 6G networks [172]. It has the potential to alleviate the problem of blockages in mmWave cellular networks [173]. IRS-aided IoT networks are promising to resolve severe signal attenuation at higher frequencies, e.g., THz [174].

Now, attention is being drawn towards IRS-assisted high frequencies

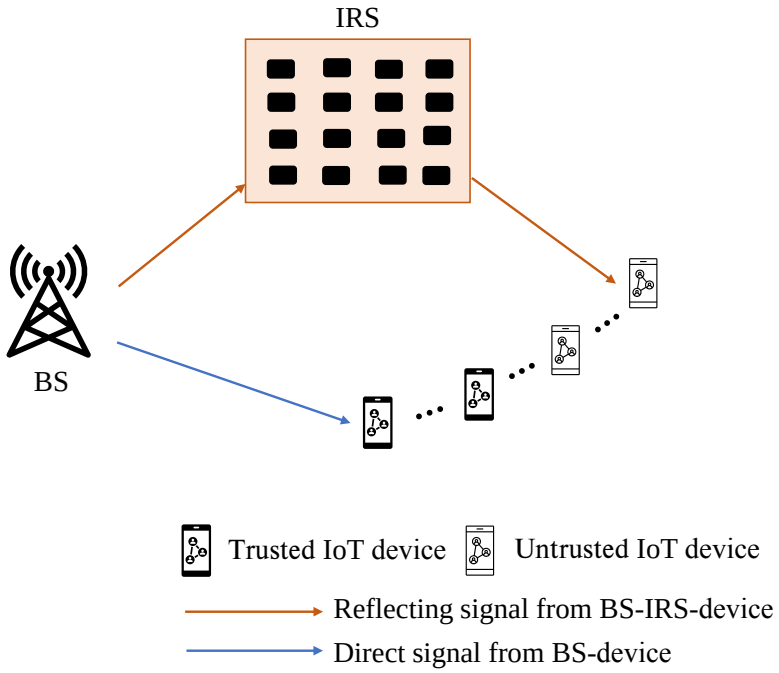


Figure 5.1: An IRS-aided communication for IoT trusted and untrusted devices with different security requirements in the presence of a direct BS-device link and a reflected BS-IRS-device link.

communication networks to join the benefits of both for the future secure wireless communications [175, 176]. One of the important security issues for these networks is handling diverse levels of security for trusted and untrusted IoT devices with potential eavesdropping threats [53, 177]. As depicted in Fig. 5.1, two different kinds of IoT devices, i.e., trusted devices and untrusted devices, require different levels of securities because the untrusted devices, though legitimate, may potentially eavesdrop on the trusted devices. The IoT BS communicates through all the legitimate (both trusted and untrusted) devices via a direct link and the reflected link from BS to IRS and then devices. For this reason, this work addresses the crucial problem of ensuring confidentiality for the trusted devices while also maintaining an essential throughput requirement of all legitimate devices.

The classification of IoT devices based on their diverse requirement of security embeds a few critical challenges. The dynamically changing radio environment for a massive IoT network adds an extra degree of complications. Also, when these devices are expected to operate at high-frequency channels of GHz, the appalling effects on signal quality cannot be disregarded. Consequently, forming an overall large-scale complex mathematical model. Thus, the chapter builds up on a sophisticated solution suitable to the nature of the problem.

5.1.1 Related Works

With recent interests in secure wireless communications, some research works have focused on IRS-aided network security. The work in [178] proposed an alternating optimization-based beamforming algorithm of IRS-aided secure network for legitimate devices in the presence of multiple pure eavesdroppers. Other alternating optimization-based solutions are presented in [179, 180] for maximizing the secrecy rate of a single device and an eavesdropper. An IRS-assisted multi-input and multi-output (MIMO) secure network with continuous and discrete phase shift coefficients is proposed in [181]. The authors in [182, 183] present a secure IRS-aided network under imperfectly known channel state information (CSI).

IRS has also found immense applications to support ultra-high-speed communication networks such as mmWave and THz. The work in [184] proposed a cross-entropy and local search algorithm to enhance the sum rate of the IRS-aided THz communication system. The work in [185] provided a joint optimization of IRS phase shift, IRS location, sub-band allocation and device power control-based solution for sum-rate maximization of an IRS-aided THz network. The authors in [186], jointly optimized the active and passive beamforming of a multi-device multi-input single-output (MISO) single-cell IRS-assisted network. Other joint optimization designs of IRS-aided MISO systems were presented for perfect and imperfect CSI in [187, 188]. The work in [189], proposed an IRS-aided decentralized framework for cooperative beamforming in a cell-free network.

Moreover, there are a few works that considered uplink communication models to focus on the issues of channel estimation and user detection. In [190], two optimum schemes based on pilot signals and IRS reflection patterns were proposed to minimize the channel estimation error for an uplink IRS-assisted orthogonal frequency division multiple access (OFDMA) system. A channel estimation algorithm for IRS-assisted THz MIMO was presented in [191]. The authors in [192], proposed a joint power control, user detection, and passive beamforming design to overcome the effect of blockage in the mmWave systems. An IRS element-grouping strategy was adopted for channel estimation via a pilot signal reduction in [193]. Joint channel estimation and rate maximization is presented for the IRS-based Terahertz MIMO network in [194].

While most of the works have adopted a classical or alternating optimization approach, some of the works have considered machine learning techniques to resolve IRS-based networks. The work in [195] presented a DRL-based algorithm for IRS-assisted secure communication in the presence of external pure eavesdroppers. The authors in [196] provided a DRL-based design for a multi-IRS framework to improve the convergence range of a downlink network operating at the THz-band. The work in [197],

proposed a DRL framework for a stand-alone IRS that self-configured its reflection coefficients to achieve an optimal rate under perfect CSI without any control from the BS. In [198], a joint design for transmit beamforming and IRS phase shifts is proposed based on a DRL algorithm.

All the prior works for IRS-aided communication have mainly addressed the network security in the presence of either a single or multiple external pure eavesdroppers. Whereas, to the best of our knowledge the issue of legitimate but potential eavesdroppers in an IRS-aided communication system has not been explored more. In addition, the previous works have not taken into the account the trusted-untrusted classification of the legitimate devices. Also, secure communication for a THz IRS-aided network has been ignored in the literature. Therefore, we address the IRS-aided network security in THz communication and we consider trusted and untrusted but legitimate devices with untrusted devices having the ability to potentially eavesdrop. In addition, a DDPG-based solution is proposed to leverage the complex, continuous and dynamic nature of the RF environment.

5.1.2 Contributions

In this study we investigated the diverse level of security for IoT devices leveraging from IRS-assisted communication for a 6G network. The problem involves a large set of complex and dynamic variables, therefore a sophisticated learning-based approach is adopted. The contributions of this chapter are listed as follows:

- A DRL-based joint active and passive beamforming design is proposed to optimize the secrecy rate of an IRS-aided network operating at the mmWave/THz band. We present the concept of trusted and untrusted legitimate devices where the untrusted devices can potentially eavesdrop on the trusted devices.

- A mathematical model is designed to calculate the secrecy rate of trusted devices for the proposed system. In addition, we formulate the optimization problem in order to simultaneously maximize the sum secrecy rate of trusted devices with a performance guarantee for both trusted and untrusted devices.

- A DDPG algorithm is developed to solve the optimization problem due to its complex and continuous nature. A MATLAB-based training network is established for the system model and the optimal solution is obtained.

The proposed DRL scheme is compared with the three different benchmark cases and the results confirm a maximum gain of $2 - 2.5$ times the sum secrecy rate of trusted devices. The throughput performance of

all trusted and untrusted devices is also ensured. Moreover, the performance of the DRL model is analyzed under meticulously selected best hyper-parameters.

5.1.3 Notation and Chapter Organization

The rest of this chapter is organized as follows. In Section 5.2, the system model is defined and then the problem is discussed. A description of the proposed DRL-based approach to solving the formulated problem is given in Section 5.3. The simulation model and results are explained in Section 5.4. The highlights of the chapter are summarized in the last Section 5.5.

Notations used in this chapter: $p(\cdot)$ and $E[\cdot]$ represent the probability and the expectation, respectively. $tr\{\cdot\}$ denotes the trace of the function enclosed. $\nabla(\cdot)$ is the gradient operator. x_t or x_{t+1} denotes the value of x at time instant t or $t+1$, respectively. The list of important variables and parameters is given in Table 5.1.

5.2 System Model and Problem Formulation

In this section, we elaborate on the system model and design the analytical representation for the achievable rates and secrecy rates of the devices. Also, the channel model is presented. Lastly, we explain the optimization problem to address the issue of security for trusted devices in the presence of potential eavesdropping devices.

5.2.1 System Model

We consider a multi-device MISO downlink system, where the base station (BS) uses M transmit antennas to communicate with K single-antenna devices. The K legitimate devices are classified into two categories of either a trusted device, k_{TD} or an untrusted device, k_{UD}. The legitimate but untrusted device may eavesdrop internally on the trusted devices. Let,

$$\mathcal{K} = \mathcal{K}_{\mathrm{TD}} \cup \mathcal{K}_{\mathrm{UD}}, \tag{5.1}$$

be the union of the sets of the trusted devices $\mathcal{K}_{\mathrm{TD}}$ and the untrusted devices $\mathcal{K}_{\mathrm{UD}}$. Also, $k = 1, 2, ..., K$ and $k \in \{k_{\mathrm{TD}}, k_{\mathrm{UD}}\}$ be the index to a device belonging either to $\mathcal{K}_{\mathrm{TD}}$ or $\mathcal{K}_{\mathrm{UD}}$. The communication is assisted by an IRS employed between the BS and devices to reflect the signals and enhance the received signal strength for the secure transmission of data. The K data signals are transmitted from the BS aimed at each one of the K devices. These transmitted signals from the BS are the first incident on the IRS and then reflected from the IRS towards the devices.

Table 5.1: List of important parameters and variables.

Symbol	Description
M	Number of base station antennas.
K	Number of all the devices.
$\mathcal{K}$	Union of sets of trusted devices, $\mathcal{K}_{\text{TD}}$ and untrusted devices, $\mathcal{K}_{\text{UD}}$, i.e., $\mathcal{K}_{\text{TD}} \cup \mathcal{K}_{\text{UD}}$.
k	Index of the device, where $k = 1, 2, ..., K$ and $k \in \{k_{\text{TD}}, k_{\text{UD}}\}$
N	Total number of reflecting elements of IRS.
n	Index of the reflecting elements, where $n = 1, 2, ..., N$.
$\lvert a_n \rvert^2$	Magnitude of the nth reflecting element.
ψ_n	Phase of the nth reflecting element.
$\mathbf{\Psi}$	The reflection beamforming matrix of the IRS.
$\mathbf{h}_{\text{b}k}$	The direct channel matrix between BS to kth device.
$\mathbf{h}_{\text{r}k}$	The channel matrix between IRS and the kth device.
$\mathbf{G}_{\text{br}}$	The channel matrix between the BS to the IRS.
$\mathbf{W}$	The transmit beamforming matrix.
$P_{\max}$	The maximum value of BS's transmit power.
y_k	The received signal at the kth device.
γ_k	The signal to interference plus noise ratio (SINR) of kth device.
$\gamma_{\min}$	The minimum threshold SINR for all devices.
σ^2	The additive white Gaussian noise (AWGN) power.
R_k	The achievable effective rate of kth device.
$R_{k_{\text{TD}}}$	The achievable rate of a trusted device, k_{TD}.
$R_{k_{\text{UD}}}$	The achievable rate of an untrusted device, k_{UD}.
$R^{\text{e}}_{k_{\text{UD}},k_{\text{TD}}}$	The eavesdropping rate of untrusted device, k_{UD} on trusted device.
$R^{\text{s}}_{k_{\text{TD}}}$	The secrecy rate of trusted device, k_{TD}.
R_{S}	The sum secrecy rate of trusted devices.

Furthermore, the IRS is composed of a micro-controller and N reflecting elements placed along the x-y dimension on the mirror surface. These passive elements can be intelligently programmed to change their phase shifts according to the dynamics of the wireless environment. Exploiting the reflection property of an IRS for a multi-device MISO system [170], we further assume that the IRS can serve all the devices at a particular time.

The passive beamforming matrix of IRS is given as $\mathbf{\Psi} = \text{diag}([a_1, ..., a_N]^\top)$ where, $a_n = e^{j\psi_n}$ and ψ_n is the element phase shift, $\forall\ n = 1, 2, ..., N$. We consider the magnitude of each element of the IRS as unity, i.e., $|a_n|^2 = 1$.

Under a particular setting of IRS and CSI for all the individual chan-

nels, the received signal at the k-th device is

$$y_k = \left(\mathbf{h}_{\mathrm{b}k}^{\mathrm{H}} + \mathbf{h}_{\mathrm{r}k}^{\mathrm{H}} \mathbf{\Psi} \mathbf{G}_{\mathrm{br}}\right) \mathbf{x} + w_k, \quad \forall k \in \mathcal{K}, \tag{5.2}$$

where, $\mathbf{h}_{\mathrm{b}k} \in \mathbb{C}^{M\times 1}$, $\mathbf{h}_{\mathrm{r}k} \in \mathbb{C}^{N\times 1}$ and $\mathbf{G}_{\mathrm{br}} \in \mathbb{C}^{N\times M}$ represents the channels between BS to kth device, IRS to the kth device and BS to the IRS respectively. w_k is the additive white Gaussian noise (AWGN) with zero mean and σ_k^2 variance. In (3.9), $\mathbf{x} \in \mathbb{C}^{M\times 1}$ is the transmit signal given as

$$\mathbf{x} = \mathbf{W}\mathbf{s} = \sum_{k=1}^{K} \mathbf{w}_k s_k, \quad k = 1, \cdots, K, \tag{5.3}$$

where, $\mathbf{W} = [\mathbf{w}_1, \cdots, \mathbf{w}_K] \in \mathbb{C}^{M\times K}$ is the beamforming matrix with $\mathbf{w}_k$ as the beamforming vector and s_k is the transmitted data for the kth device. We consider, the maximum value of transmit power $P_{\max}$ of the base station as

$$tr\left\{\mathbf{W}^{\mathbf{H}}\mathbf{W}\right\} \leq P_{\max}, \tag{5.4}$$

where, $tr\{\mathbf{W}^{\mathbf{H}}\mathbf{W}\}$ is the trace of the beamforming matrix. We assume $s_k \sim \mathcal{N}(0,1)$ to be Gaussian distributed independent random variable. The received signal in (5.2) for the kth device takes into account the co-channel interference from the other devices and is written as

$$\begin{aligned} y_k = & \underbrace{\left(\mathbf{h}_{\mathrm{b}k}^{\mathrm{H}} + \mathbf{h}_{\mathrm{r}k}^{\mathrm{H}} \mathbf{\Psi} \mathbf{G}_{\mathrm{br}}\right) \mathbf{w}_k s_k}_{\text{Desired signal}} \\ & + \underbrace{\sum_{j\in\mathcal{K}, j\neq k} \left(\mathbf{h}_{\mathrm{b}k}^{\mathrm{H}} + \mathbf{h}_{\mathrm{r}k}^{\mathrm{H}} \mathbf{\Psi} \mathbf{G}_{\mathrm{br}}\right) \mathbf{w}_j s_j}_{\text{Co-channel interference}} + w_k, \quad \forall k \in \mathcal{K}. \end{aligned} \tag{5.5}$$

Also, for the device k, the received signal-to-interference-plus-noise ratio (SINR) is given as

$$\gamma_k = \frac{\left|\left(\mathbf{h}_{\mathrm{b}k}^{\mathrm{H}} + \mathbf{h}_{\mathrm{r}k}^{\mathrm{H}} \mathbf{\Psi} \mathbf{G}_{\mathrm{br}}\right) \mathbf{w}_k\right|^2}{\sum_{j\in\mathcal{K}, j\neq k} \left|\left(\mathbf{h}_{\mathrm{b}k}^{\mathrm{H}} + \mathbf{h}_{\mathrm{r}k}^{\mathrm{H}} \mathbf{\Psi} \mathbf{G}_{\mathrm{br}}\right) \mathbf{w}_j\right|^2 + \sigma_k^2}, \quad \forall k \in \mathcal{K}, \tag{5.6}$$

and the achievable rate is given as

$$R_k = \log_2\left(1 + \gamma_k\right), \quad \forall k \in \mathcal{K}, \tag{5.7}$$

where R_k is measured in bps/Hz and k is the index of either a trusted device, k_{TD} or an untrusted device, k_{UD}.

In case if any of the untrusted device $k_{\mathrm{UD}} \in \mathcal{K}_{\mathrm{UD}}$ eavesdrop on the trusted device $k_{\mathrm{TD}} \in \mathcal{K}_{\mathrm{TD}}$, then the eavesdropping rate of k_{UD} at the trusted device is given as [177]

$$R^{\mathrm{e}}_{k_{\mathrm{UD}},k_{\mathrm{TD}}} = \log_2\left(1+\frac{\left|\left(\mathbf{h}^{\mathrm{H}}_{\mathrm{b}k_{\mathrm{UD}}}+\mathbf{h}^{\mathrm{H}}_{\mathrm{r}k_{\mathrm{UD}}}\boldsymbol{\Psi}\mathbf{G}_{\mathrm{br}}\right)\mathbf{w}_{k_{\mathrm{TD}}}\right|^2}{\sum_{j\in\mathcal{K}_{\mathrm{TD}},j\neq k_{\mathrm{TD}}}\left|\left(\mathbf{h}^{\mathrm{H}}_{\mathrm{b}k_{\mathrm{UD}}}+\mathbf{h}^{\mathrm{H}}_{\mathrm{r}k_{\mathrm{UD}}}\boldsymbol{\Psi}\mathbf{G}_{\mathrm{br}}\right)\mathbf{w}_j\right|^2+\sigma^2_{k_{\mathrm{UD}}}}\right), \quad \forall k_{\mathrm{UD}}\in\mathcal{K}_{\mathrm{UD}}, \tag{5.8}$$

where $\mathbf{h}_{\mathrm{b}k_{\mathrm{UD}}}$ and $\mathbf{h}_{\mathrm{r}k_{\mathrm{UD}}}$ represent the channels between the BS and the untrusted device k_{UD} and between the IRS to the untrusted device k_{UD}, respectively. $\sigma^2_{k_{\mathrm{UD}}}$ is the noise variance of the untrusted device, k_{UD}.

Taking into consideration that any of the untrusted devices can try to eavesdrop then the secrecy rate of a trusted device, k_{TD}, is given as [195]

$$R^{\mathrm{s}}_{k_{\mathrm{TD}}} = \left[R_{k_{\mathrm{TD}}} - \max_{\forall k_{\mathrm{UD}}} R^{\mathrm{e}}_{k_{\mathrm{UD}},k_{\mathrm{TD}}}\right]^+, \quad \forall k_{\mathrm{TD}}\in\mathcal{K}_{\mathrm{TD}}, \tag{5.9}$$

where $[a]^+ = \max(a,0)$, and $R_{k_{\mathrm{TD}}}$ is the achievable rate of the trusted device and can be calculated from (3.12), (3.13) and (3.1) when $k = k_{\mathrm{TD}}$, as given by

$$R_{k_{\mathrm{TD}}} = \log_2\left(1+\frac{\left|\left(\mathbf{h}^{\mathrm{H}}_{\mathrm{b}k_{\mathrm{TD}}}+\mathbf{h}^{\mathrm{H}}_{\mathrm{r}k_{\mathrm{TD}}}\boldsymbol{\Psi}\mathbf{G}_{\mathrm{br}}\right)\mathbf{w}_{k_{\mathrm{TD}}}\right|^2}{\sum_{j\in\mathcal{K},j\neq k_{\mathrm{TD}}}\left|\left(\mathbf{h}^{\mathrm{H}}_{\mathrm{b}k_{\mathrm{TD}}}+\mathbf{h}^{\mathrm{H}}_{\mathrm{r}k_{\mathrm{TD}}}\boldsymbol{\Psi}\mathbf{G}_{\mathrm{br}}\right)\mathbf{w}_j\right|^2+\sigma^2_{k_{\mathrm{TD}}}}\right), \quad \forall k_{\mathrm{TD}}\in\mathcal{K}_{\mathrm{TD}}, \tag{5.10}$$

where $\mathbf{h}_{\mathrm{b}k_{\mathrm{TD}}}$ and $\mathbf{h}_{\mathrm{r}k_{\mathrm{TD}}}$ represents the channels between the BS to the trusted device k_{UD} and the IRS to the trusted device k_{TD}, respectively. $\sigma^2_{k_{\mathrm{TD}}}$ is the noise variance of the trusted device, k_{TD}. In (5.10), the co-channel interference from all the devices is taken into account in order to confirm the legitimacy of devices irrespective of being trusted or untrusted.

5.2.2 Channel Model of the Proposed System

At THz frequency, the communication channel is affected by molecular absorption, spreading, and other unfavourable conditions which result in path loss. It is composed of the line of sight (LoS) and a few non-LoS scattering components. The power of non-LoS scattering components is very negligible as compared to the LoS components [54, 199]. Therefore, we ignore the non-LoS components and only consider the LoS part as it majorly dominates the THz channel. The overall path loss of the proposed THz channel for the direct path from BS to devices is given as [54, 196]

$$L(f,d) = \frac{c}{4\pi f d}e^{-\frac{1}{2}\tau(f)d}, \tag{5.11}$$

and for the reflected path from BS to IRS to devices, the combined pathloss is given as

$$L(f, d_1, d_2) = \frac{c}{8\sqrt{\pi^3} f d_1 d_2} e^{-\frac{1}{2}\tau(f)(d_1+d_2)}, \tag{5.12}$$

where f is the operating frequency, d is the distance between the BS and device, d_1 is the distance between the BS and IRS, d_2 is the distance between the IRS and device, $\tau(f)$ is the medium absorption factor, and c is the speed of light.

Moreover, we consider that $\mathbf{h}_{\mathrm{b}k}$, $\mathbf{h}_{\mathrm{r}k}$ and $\mathbf{G}_{\mathrm{br}}$ are the channels between the BS to the kth device, the IRS to the kth device and the BS to the IRS, respectively, and they are complex in nature. These complex channels are modeled individually as the product of zero-mean, unit-variance Gaussian variables and their respective path loss.

5.2.3 Problem Formulation

Existing and emerging mmWave/THz communications suffer from high signal attenuation, as the coverage range of the signals is low and prone to blockages. This results in unfair throughput distribution for different devices trusted or untrusted. The quality of service (QoS) of individual devices in terms of their achievable rate and the network security in terms of the secrecy rate is affected as interference and chances of eavesdropping from other devices also increase at these frequencies.

To address this problem for multi-device diversity at high frequencies an IRS-supported network is implemented. The objective is to reduce the propagation attenuation and provide improved beamforming for the devices with the goal of maximizing the overall secrecy of the network. It involves providing each device with performance relative to its instantaneous channel conditions. Also, the proposed IRS-aided network tends to resolve the throughput fairness among devices by maintaining the minimum throughput threshold for all devices. Therefore, we make a trade-off between continuously configuring reflecting elements of IRS and QoS of individual devices by finding the optimum IRS configuration and digital beamforming matrix at the BS. This maintains fairness among the devices and achieves network secrecy.

This paper aims to maximize the sum secrecy rate of the trusted devices, denoted by $R_{\mathrm{S}}(\mathbf{h}_{\mathrm{b}k}, \mathbf{h}_{\mathrm{r}k}, \mathbf{G}_{\mathrm{br}}, \mathbf{\Psi}, \mathbf{W})$:

$$R_{\mathrm{S}}(\mathbf{h}_{\mathrm{b}k}, \mathbf{h}_{\mathrm{r}k}, \mathbf{G}_{\mathrm{br}}, \mathbf{\Psi}, \mathbf{W}) = \sum_{k_{\mathrm{TD}} \in \mathcal{K}_{\mathrm{TD}}} R^{\mathrm{s}}_{k_{\mathrm{TD}}}, \tag{5.13}$$

while maintaining QoS for all the legitimate devices (both trusted and untrusted). It is mathematically written as the following maximization

problem

$$\begin{aligned} \max_{\mathbf{\Psi},\mathbf{W}} \quad & R_{\mathrm{S}}(\mathbf{h}_{\mathrm{b}k}, \mathbf{h}_{\mathrm{r}k}, \mathbf{G}_{\mathrm{br}}, \mathbf{\Psi}, \mathbf{W}) \\ \text{s.t.} \quad & C_1 : tr\left\{\mathbf{W}^{\mathbf{H}}\mathbf{W}\right\} \leq P_{\max}, \\ & C_2 : \gamma_k \geq \gamma_{\min}, \quad \forall k \in \mathcal{K} \\ & C_3 : |a_n|^2 = 1, \quad \forall n = 1, 2, \ldots, N \\ & C_4 : \psi_n \in (0, 2\pi], \forall n = 1, 2, \ldots, N. \end{aligned} \tag{5.14}$$

where constraint C_1 represents the active beamformer transmit power which is limited to a maximum level of $P_{\max}$. The QoS requirement of all the devices (both trusted and untrusted) is defined by constraint C_2 with $\gamma_{\min}$ as the threshold or minimum SINR. Constraint C_3 and C_4 define the amplitude and phase shifts of all the N elements of IRS for passive beamforming.

Using the basic calculus and algebraic calculations it can be shown that the optimization problem in (5.14) is non-convex and NP-hard in nature. The computational complexity of solving such a problem with traditional mathematical methods is too high and the analytical solution is highly not possible to obtain. In recent years, deep learning and DRL have emerged as eminent techniques to solve IRS-based communication networks, especially for complex system models like ours as in [193, 196]. Our proposed network demands online adjustment of transmit beamforming matrix and IRS phase-shifting matrix based on a real-time device environment. Furthermore, operating in the THz spectrum adds difficult channel conditions for the overall model. Due to the continuous dynamic nature, DRL is highly desirable to solve such goal-directed high-dimensional problems [200]. Therefore, we aim to leverage the DRL-based approach for converging to an optimal solution with low computational complexity and a higher level of flexibility for a continuously varying environment. Once a DRL model is trained, it can optimize the solution for any given CSI.

5.3 DRL-Based Joint Active and Passive Beamforming Design

In this section, we initially describe the fundamentals of DRL based on the MDP, RL, and DL. We provide an overview of DDPG as a type of DRL technique. Secondly, we present our proposed DDPG solution for the joint active and passive beamforming problem in section 5.2.3. Lastly, we explain the DDPG algorithm for the designed solution.

5.3.1 DRL and DDPG Overview

Artificial intelligence and machine learning techniques are being widely adopted to reduce the computational complexity involved in large-scale wireless communication systems. One of the major types of machine learning applications is reinforcement learning. RL has enabled us to make real-time decisions based on the challenging dynamics of the network's environment [29]. Typically, an RL model is an MDP that involves decision-making entities known as agents to interact with the environment by transitioning between different states and taking actions. Every action has a return associated with it which is also known as a reward. This iterative interaction eventually enables the agent in learning the environment through the quality of instantaneous reward calculated at each stage. Therefore, agents lead towards the best actions from their past experiences.

Another subclass of machine learning is deep learning which exploits multiple neural network layers to intelligently extract useful information from high-dimensional data [201]. Initially, a training model is run to calculate the weighted values of neural nodes of each layer. Once the training process is completed, then the trained network can perform decisions with high accuracy on any set of input data.

The following subsections provide insights about DRL, a hybrid of DL and RL, and secondly one of the important DRL techniques, DDPG.

5.3.1.1 DRL Fundamentals

Currently, DRL has evolved as a combination of deep neural networks (DNNs) and RL techniques to cater to the performance advantages of both in one [200]. DL aids in overcoming the complexity in training the high dimensional problem for the learning process of RL algorithms, consequently improving the learning rate and efficiency. The fundamental elements in an RL network are summarized as

- ***State space***: $\mathcal{S}$ is a set of states such that, $s \in \mathcal{S}$, which are observed from the environment.
- ***Action space***: $\mathcal{A}$ is a set of actions available for agent to chose, such that $a \in \mathcal{A}$.
- ***Probability function***: $p(t) = p(s_{t+1} = s' \mid s_t = s, a_t = a)$ is the transition probability representing that the action a taken at state s at time t will result in the state s' at time $t+1$. The probability of selecting an action a at any state s is termed as policy, $\pi(s,a) = p\left(a_t = a \mid s_t = s\right)$ and the sum of these probabilities for all actions at time t from state s must be unity i.e., $\sum_{a \in \mathcal{A}} \pi(s,a) = 1$.

- ***Reward function***: $r(t) = r_t(s_t, a_t)$ is the received reward function to calculate the instantaneous reward by the agents when they take action a at state s.
- ***Value function***: Based on the immediate reward obtained in a given state a value function estimates a long-term future reward of taking the action a in that given state.

The whole set of transitions (a_t, s_t, r_t, s_{t+1}) forms an *experience*. Since it is an episodic process, assuming that the time instants in an episode follows a sequence of $(t, t+1, t+2, ..., T)$ as $T \to \infty$ and T is the total number of steps in an episode. The cumulative reward is given by [200]

$$Z_t = \sum_{\tau=0}^{\infty} \delta^{\tau} r_{t+\tau+1} \tag{5.15}$$

where $\delta \in (0, 1]$ is the discount factor. The most common state-action value function is know as Q-value function. When the policy π is adopted at time t to take action a_t for a state s_t in order to obtain an instant reward r_t, then the Q-value function to calculate the future long term reward is given as

$$Q_{\pi}(s_t, a_t) = E_{\pi}\left[Z_t \mid s_t = s, a_t = a\right], \tag{5.16}$$

where $E[\cdot]$ represents the expectation. Using the Bellman Equation the Q-value function in (5.16) is expanded as [30]

$$Q_{\pi}(s_t, a_t) = E_{\pi}\left[r_{t+1} \mid s_t = s, a_t = a\right] + \delta \sum_{s' \in \mathcal{S}} p(t) \left(\sum_{a' \in \mathcal{A}} \pi(s', a') Q_{\pi}(s', a') \right), \tag{5.17}$$

where $p(t) = p(s_{t+1} = s' \mid s_t = s, a_t = a)$ is the transition probability representing that the action a taken at state s at time t will result in the state s' at time $t+1$ and the action a' represents the next iteration action at $t+1$ from state s' to the next state. For the state-action value functions there is an optimal function that provides the maximum cumulative reward for all states. This value is achieved through adopting an optimal policy which is given as

$$\pi^*(s, a) = \arg\max_{\pi} Q^*(s, a), \forall s \in \mathcal{S} \tag{5.18}$$

From (5.17) and (5.18), the optimal Q-value function adopted by the learning agent for an optimal policy is given as

$$Q^*(s_t, a_t) = r_{t+1}(s_t, a_t, \pi = \pi^*) + \delta \sum_{s' \in \mathcal{S}} p(t) \max_{a' \in \mathcal{A}} Q^*(s', a'). \tag{5.19}$$

The recursive iterations of (5.19) provide the optimum solution by adopting a greedy policy or maximum value policy for the agents. It involves constructing and updating the Q-value table for the action and state spaces. The updating is done as

$$Q^*\left(s_t, a_t\right) \leftarrow \left(1-\alpha\right) Q^*\left(s_t, a_t\right) + \alpha\left(r_{t+1} + \delta \max_{a' \in \mathcal{A}} Q_\pi(s', a')\right), \quad (5.20)$$

where $\alpha \in (0, 1]$ is the learning rate of the model. However this process becomes too complex for high dimensional action and state spaces because the transition probabilities and rewards are unknown and it is hard to approximate the optimal value. Therefore, deep neural networks (DNNs) are used as learning agents to approximate the Q-value and policy functions. The combination is termed as a DRL model.

DRL algorithms can be categorized into three different types depending on estimation techniques: (a) policy-based (b) value-based and (c) a hybrid of both. When it comes to large continuous action and state spaces, the application of only policy-based or value-based algorithms becomes tedious to provide better convergence. Therefore, a hybrid of both is preferred [57]. According to the formulated problem in Section 5.2.3, the required action space i.e., selection of passive and active beamforming matrix elements, is continuous in nature. Therefore, we employ a hybrid algorithm, DDPG.

5.3.1.2 Functioning Blocks of DDPG

A DRL technique with an *actor-critic*-based architecture that supports a large continuous action space is known as DDPG [57]. The actor and critic are both neural network-based structures with multiple layers. The actor and critic frameworks combine the features of value functions as well as policy parameters to evaluate the learning experience. The purpose of actor-network is to parameterize the policy by interacting with the environment's states and to output an action. The parameter θ is used to distinguish the weighted policy for DNN agents. The policy function is parameterized as $\pi\left(\theta \mid s, a\right)$ by the actor-network. On the other hand, a critic network approximates the value function based on the actions and parameterized policy. The resulting parameterized Q-value function is denoted as $Q\left(\theta \mid s_t, a_t\right)$ and the weighted parameter is updated as

$$\theta_{t+1} = \theta_t - \alpha \nabla_\theta F(\theta), \quad (5.21)$$

where α is learning rate of DDPG agent, $F(\theta)$ is the loss function associated with the weight parameter, θ of DNN and ∇_θ represents the gradient of the loss function. In a neural network a loss function defines the values difference between the predicted value and the target value. While DRL

involves no prior knowledge of the target value therefore we use two similar neural networks termed as a training network and the target network to estimate the actual target value. Let, the Q-value functions of these two networks are $Q\left(\theta_{\text{train}} \mid s_t, a_t\right)$ and $Q\left(\theta_{\text{target}} \mid s_t, a_t\right)$ then the actual target Q-value is given as

$$v = r_{t+1} + \delta \max_{a' \in \mathcal{A}} Q\left(\theta_{\text{target}} \mid s', a'\right), \tag{5.22}$$

and the loss function is given as

$$F(\theta) = \Big(v - Q\left(\theta_{\text{train}} \mid s_t, a_t\right)\Big)^2. \tag{5.23}$$

Moreover, the training and the target DDPG models require updates for both their actor and critic networks. This is elaborated as follows:

Training critic network: The critic network in the training DDPG is updated through

$$\theta_{t+1,\text{c1}} = \theta_{t,\text{c1}} - \alpha_{\text{c1}} \nabla_{\theta_{\text{c1}}} F(\theta_{\text{c1}}), \tag{5.24}$$

$$F(\theta_{\text{c1}}) = \Big(r_t + \delta Q\left(\theta_{\text{c2}} \mid s', a'\right) - Q\left(\theta_{\text{c1}} \mid s_t, a_t\right)\Big)^2, \tag{5.25}$$

where, θ_{c1} and θ_{c2} represent the parameters of training and target critic network respectively. Here, α_{c1} is the learning rate of training critic network. The loss function of the training critic network, $F(\theta_{\text{c1}})$, in (5.25) is updated based on both the target network critic network and the training critic network.

Training actor-network: While the training critic network updates based on the value functions, the training actor-network updates as

$$\theta_{t+1,\text{a1}} = \theta_{t,\text{a1}} - \alpha_{\text{a1}} \nabla_{a_t} Q\left(\theta_{\text{c2}} \mid s_t, a_t\right) \nabla_{\theta_{\text{a1}}} \pi\left(\theta_{\text{a1}} \mid s_t, a_t\right), \tag{5.26}$$

where α_{a1} and θ_{a1} are the learning rate and parameter of the training actor network, respectively; $\nabla_{a_t} Q\left(\theta_{\text{c2}} \mid s_t, a_t\right)$ is the gradient of the target critic network with respect to the target action a_t; and $\nabla_{\theta_{\text{a1}}} \pi\left(\theta_{\text{a1}} \mid s_t, a_t\right)$ is the gradient of the training actor network with respect to the parameter θ_{a1}. The gradient of the target critic network plays a key role in the next action selection of the target which favours the optimal Q-value.

Target critic network: This network is updated as

$$\theta_{\text{c2}} \leftarrow \alpha_{\text{c2}} \theta_{\text{c1}} + \left(1 - \alpha_{\text{c2}}\right) \theta_{\text{c2}}, \tag{5.27}$$

where α_{c2} is the learning rate of the target critic network.

Target actor-network: This network is updated as

$$\theta_{\text{a2}} \leftarrow \alpha_{\text{a2}} \theta_{\text{a1}} + \left(1 - \alpha_{\text{a2}}\right) \theta_{\text{a2}}, \tag{5.28}$$

where α_{c2} and θ_{a2} is the learning rate and the parameter of the target actor network, respectively.

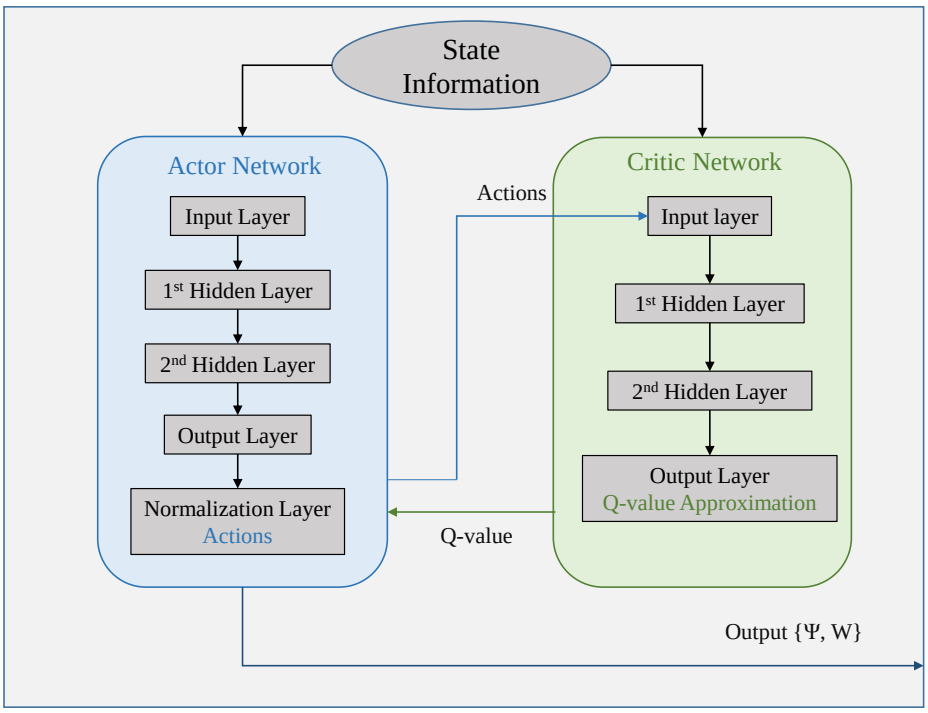


Figure 5.2: The proposed DDPG framework of Actor and Critic as multi-layer neural networks with a single input and output layer, and two hidden layers each.

5.3.2 Proposed DDPG Model

Our proposed DDPG framework comprises an actor and a critic DNN network as shown in Fig. 5.2. These learning agents continuously interact with the environment to update states, actions, and rewards. States information from the environment is input to both the networks as observations of instantaneous channels, transmit and receive powers, phase angles of IRS elements, and beamforming matrix of BS. Both the actor and critic are multi-layer neural networks with a single input and output layer and two hidden layers each.

Actor Neural Network: The actor DNN consists of fully connected layers. The input layer which is taking the state information is then followed by two hidden layers. These hidden layers are using *Tangent hyperbolic*, $\tanh(\cdot)$ as the activation function for the training. The advantage of using the $\tanh(\cdot)$ function is to cater to both the positive and negative values. The outcome from the second hidden layer is then passed to the last layer which is the output layer. The output layer provides the actions set to the critic based on the Q-value approximated from the critic network.

The regression function is used as the activation function for the layer. The normalization layer after output layer performs the normalization to satisfy $tr\left\{\mathbf{W^H W}\right\} \leq P_{\max}$ for the power constraint.

Critic Neural Network: It is composed of all fully connected layers. The input layer is taking state information from the environment and actions from the actor as input. This layer is then connected to two hidden layers both using the $\tanh(\cdot)$ function for activation. The outcome of the second hidden layer is then passed as an input to the output layer. The output layer of the critic network provides the Q-value estimated based on the DDPG policy. A regression activation function is used in the output layer to find the mean square error from the target Q-value.

5.3.3 Proposed DDPG Algorithm

The proposed DDPG algorithm will obtain the optimal values for the design variables of the optimization problem defined in Section 5.2.3. It will jointly optimize the active, $\mathbf{W}$ and passive, $\mathbf{\Psi}$ beamforming matrices. The proposed algorithm can adapt itself to the given channel information i.e., between BS-device, BS-IRS and IRS-device and there is no need for estimating the channels. This is because the estimation of those channels is very challenging and out of the scope of this work. The given channel information will be input to the agent along with the maximum transmit power and SINR constraints, i.e, $P_{\max}$ and $\gamma_{\min}$. We also observe $|a_n|^2 = 1$ and $\psi_n \in (0, 2\pi]$ for all the IRS elements.

To begin with the DDPG algorithm initializes a Replay memory, B, which is composed of the experience (a_t, s_t, r_t, s_{t+1}) transitions. It also provides the initial values for the DNN parameters of both the training and target networks such that, $\theta_{\text{a1}} = \theta_{\text{a2}}$ and $\theta_{\text{c1}} = \theta_{\text{c2}}$. The active and passive beamforming matrices are both initialized as identity matrices. Later, we allocate a mini-batch of experience samples to train the DDPG agents. Once training is completed and the Q-value is obtained, it will be used to store new experiences in the memory B. In this way, the Agents will be trained more effectively based on past and new experiences. Moreover, a target network helps stabilize the Q-value estimation of the training network and reduce any correlations while training.

The proposed DDPG algorithm has been summarized in algorithm 3. The algorithm runs for D number of episodes and perform T iterations in each episode. At each iteration the state of the DDPG agent s_t is defined as the the channel matrices $\mathbf{h}_{\text{b}k}, \mathbf{h}_{\text{r}k}$ and $\mathbf{G}_{\text{br}} \forall k$, the action taken in previous iteration and the transmit and receive powers of all devices. The action a_t taken is the selection of transmit beamforming matrix and the IRS phase shift matrix. Whereas the reward is estimated as the sum secrecy rate $R_{\text{S}}(\mathbf{h}_{\text{b}k}, \mathbf{h}_{\text{r}k}, \mathbf{G}_{\text{br}}, \mathbf{\Psi}, \mathbf{W})$ of the trusted devices in the network.

At a time step t, the action $a_t = \{\mathbf{\Psi}_t, \mathbf{W}_t\}$ results in a next state

Algorithm 3 Proposed DDPG Algorithm for joint active and passive beamforming

Input: channel matrices $\mathbf{h}_{\mathrm{b}k}, \mathbf{h}_{\mathrm{r}k}$ and $\mathbf{G}_{\mathrm{br}} \forall k$, transmit power, $P_{\max}$, and minimum SINR threshold, $\gamma_{\min}$

print replay memory size, B, the DNN parameters for training actor, θ_{a1}, training critic, θ_{c1}, target actor, θ_{a2}, target critic, θ_{c2}, the IRS phase shift matrix, $\mathbf{\Psi}$, the transmit beamforming matrix, $\mathbf{W}$

for each episode $= 1, 2, ..., D$
  set the first state s from the information of all channel matrices
  for iteration $t = 0, 1, 2, ..., T-1$
    select the action $a_t = \{\mathbf{\Psi}_t, \mathbf{W}_t\}$ for the given state s_t updated from the actor network and go to state s_{t+1}
    evaluate reward r_{t+1}
    keep the experience in replay memory
    update the Q-value function of the training critic network as $Q(\theta_{\mathrm{c1}} \mid s_t, a_t)$
    select a mini-batch experience samples of size X from the replay memory
    Set up the loss function $F(\theta_{\mathrm{c1}})$ in (5.25) for the training critic network
    calculate the gradient $\nabla_{\theta_{\mathrm{c1}}} F(\theta_{\mathrm{c1}})$ from the training critic network
    calculate the gradient $\nabla_{\theta_{\mathrm{a1}}} \pi(\theta_{\mathrm{a1}} \mid s_t, a_t)$ from the training actor network
    calculate the gradient $\nabla_{a_t} Q(\theta_{\mathrm{c2}} \mid s_t, a_t)$ from the target actor network
    using (5.24) update the parameter θ_{c1}
    using (5.26) update the parameter θ_{a1}
    using (5.27) update the parameter θ_{c2} after every U iteration
    using (5.28) update the parameter θ_{a2} after every U iteration
    return the updates as next state s_{t+1} to the learning agents
  end for
end for

Output: optimal action $a = \{\mathbf{\Psi}, \mathbf{W}\}$, Q-value function

s_{t+1}. The reward earned for this action, r_{t+1}, is used by training the critic DNN to estimate and update the Q-value function, $Q(\theta_{\mathrm{c1}} \mid s_t, a_t)$. The experience is continuously stored in the memory, and another set of minibatches is selected to activate the learning. The loss function $F(\theta_{\mathrm{c1}})$ and the gradient $\nabla_{\theta_{\mathrm{c1}}} F(\theta_{\mathrm{c1}})$ are calculated for the training critic network. Also, the gradients $\nabla_{\theta_{\mathrm{a1}}} \pi(\theta_{\mathrm{a1}} \mid s_t, a_t)$ and $\nabla_{a_t} Q(\theta_{\mathrm{c2}} \mid s_t, a_t)$ are computed for the training actor network and the target actor network, respectively. The parameters of the training and target agents, i.e., θ_{c1}, θ_{a1}, θ_{c2}, and θ_{a2}, are updated accordingly. The algorithm iterates until the last iteration or the convergence to the optimal values of $\mathbf{W}$ and $\mathbf{\Psi}$.

Also, in our proposed model the complex input entries of the action and state space are dealt as two separate absolute values contributing from the real and imaginary part respectively. The action space is $\mathbf{W} = \text{Re}\{\mathbf{W}\} + \text{Im}\{\mathbf{W}\}$ and $\mathbf{\Psi} = \text{Re}\{\mathbf{\Psi}\} + \text{Im}\{\mathbf{\Psi}\}$. Hence, the dimension of action space is $2(MK + N)$. Similarly, the state space is also doubled to compensate for its complex nature. The channel matrices, $\mathbf{h}_{\text{b}k}$, $\mathbf{h}_{\text{r}k}$ and $\mathbf{G}_{\text{br}}$ contribute $2(MK + NK + NM)$ observations. The transmit and the received power of all the K devices adds $2(K + K^2)$ entries. Therefore, combining all the state space entries of channel observations, transmit power, receive power, and previous action, the overall dimension becomes $2(MK + NK + NM + K + K^2 + MK + N)$.

Moreover, the time-complexity of our DNN-based model is linear to its size. The training and target agents in total execute DT iterations. Given our training and target actor and critic networks have similar structures, we assume ℓ_a and ℓ_c as the number of neurons in the actor and critic hidden layers, respectively. Hence, the time-complexity of the DDPG agent with an actor and critic is $\mathcal{O}(DT[(2MK+2N)\ell_a+(2MK+2NK+2NM+2K+2K^2+2MK+2N)\ell_c])$.

5.4 Simulations and Results

In this section, we evaluate the solution proposed in Section 5.3 using simulations. The proposed system model and DDPG algorithm are simulated in MATLAB. We consider $K = 4$ devices, populated randomly to capture the trusted-untrusted classification, in two different radii of 20 m centered at a distance of 120 m and 220 m from the BS. The trusted devices are closer to the BS than the untrusted devices, as suggested in [53,177]. We also consider the first half of the K devices as trusted devices and the remaining half as untrusted devices. The IRS with $N = 128$ elements is located at a distance of 100 m from the BS. The channel matrices $\mathbf{h}_{\text{b}k}$, $\mathbf{h}_{\text{r}k}$ and $\mathbf{G}_{\text{br}}$ are randomly generated as complex normal distributed. The operating frequency, f, of the proposed network is 1 THz [185]. The respective path loss is added to the channel matrix entities from (5.11) and (5.12). The flow of the simulation network is illustrated in Fig. 5.3. First of all, the training dataset is generated and fed to the DDPG training module. The training module creates the DDPG learning agents, i.e., actor and critic, the environment and training options. Once training is completed, it is saved and loaded for the test dataset to achieve the optimum solution. The optimum result of active and passive beamforming matrices is utilized to calculate the sum secrecy rate and effective sum throughput. We run the simulations for different values of the BS transmit power, $P_{\max}$. Other simulation parameter values for the proposed system model are summarized in Table 5.2.

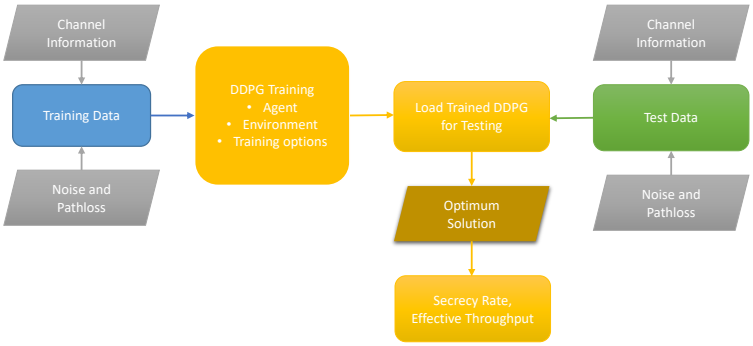


Figure 5.3: Flow diagram of the simulation network.

Table 5.2: Simulation parameters of the system model and their values.

Parameter	**Value**	**Parameter**	**Value**
K	4	M	4
N	128	f	1 THz
σ^2	-174 dBm/Hz	$\tau(f)$	0.033

5.4.1 Settings of the Benchmark Scenarios

We consider three existing benchmark scenarios in order to compare the proposed DRL-based jointly optimized transmit power beamforming and the IRS phase shift matrices.

Case 1: The first scenario takes into account the classical case of no IRS-supported communication. In this case, we only consider direct communication channel link, $\mathbf{h}_{\mathrm{b}k}$ between devices and the BS. For this scenario, the received signal y_k in (5.2) changes to

$$y_k = \left(\mathbf{h}_{\mathrm{b}k}^{\mathrm{H}}\right)\mathbf{x} + w_k, \quad \forall k \in \mathcal{K}, \tag{5.29}$$

where w_k is the noise and $\mathbf{x} = \mathbf{W}\mathbf{s}$, as defined in Section 5.2.1.

Case 2: The second scenario uses IRS-based communication similar to the proposed scenario but adopts a simple selection of passive beamforming matrix. The phase shift elements are randomly assigned a value i.e., $\psi_n \in (0, 2\pi]\,, \forall n = 1, 2, \ldots, N$. For both the first and second scenarios we consider the optimal BS transmit beamforming matrix under the given channel matrices.

Case 3: For the third scenario we adapt the closely related work [195] to our proposed scenario since its comparison with our proposed scenario in its original form would not be fair. This scenario is similar to the proposed scenario as it also optimizes the active and passive beamforming

Table 5.3: DDPG agent parameters for training and simulation.

Parameter and description	**Value**
α_{a1} , learning rate of training actor network	$n = 1e-3$
α_{c1}, learning rate of training critic network	$n = 1e-3$
α_{a2}, learning rate of target actor network	$n = 1e-3$
α_{c2}, learning rate of target critic network	$n = 1e-3$
λ_{a1}, decaying rate of training actor network	$n = 1e-4$
λ_{c1}, decaying rate of training critic network	$n = 1e-4$
δ, discount factor for reward function	0.99
B, length of replay experience memory window	$n = 2e5$
D, number of episodes	$n = 5e2$
T, number of steps per episode	$n = 2e3$
U, number of steps synchronizing target with the training network	1
X, number of experiences in the mini-batch	16

matrices of the IRS-based downlink MISO network to maintain secrecy in presence of eavesdropping devices. In contrast to our proposed scenario which considers legitimate but potential eavesdroppers and the scenario in [195] considers pure external eavesdroppers. As opposed to our proposed model that accounts for the minimum QoS of all the legitimate devices (trusted and untrusted devices), the scenario in [195] maintains QoS for its legitimate devices which are the trusted devices only. We adapt the equations from Section II-A of [195] to calculate the secrecy rate of trusted devices.

5.4.2 Settings of the Proposed DDPG

We use *Adam* as the optimizer for the actor and critic network. For both the actor and critic networks, the dimensions of input and output layers correspond to their respective input and output dimensions. According to the general DNN practice, the number of neurons in all the hidden layers is kept larger than the input and output layer dimensions [198]. For our model, hidden layer 2 is four times larger than hidden layer 1 for both DDNs. The training DDPG network trains the actor and critic through adaptive learning rates, $\alpha_{t+1,\text{a1}} = \lambda_{\text{a1}}\alpha_{t,\text{a1}}$ and $\alpha_{t+1,\text{c1}} = \lambda_{\text{c1}}\alpha_{t,\text{c1}}$ at each iteration of the episode. λ_{a1} is the decaying rate of the training actor-network. λ_{c1} is the decaying rate of the training critic network. The DDPG agent hyper-parameters are listed in Table 5.3.

The MATLAB simulator of the DDPG actor-critic network is illustrated in Fig. 5.4. The actor and critic are fully connected DNNs. The nodes in the actor-network represent the different layers. It begins with the input layer, *Observation*, and then is followed by two fully connected

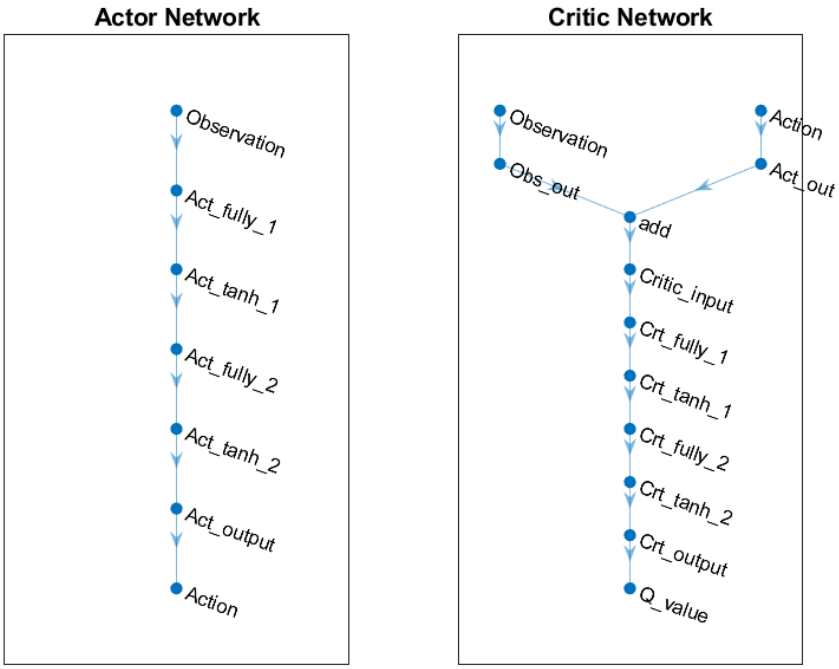


Figure 5.4: The proposed DDPG actor-critic fully connected DNNs in MATLAB with each node representing the different neural layer.

layers *Act_fully_1* and *Act_fully_2* using the tanh(·) activation function. The second last node represents the output layer, *Act_output* using the regression activation function. This is followed by the final node, *Action*, for the normalized output. The two initial node branches of the critic network represent an input *Observation* layer and input *Action* layer (from the actor) which are added as a combined input to the Critic network. These are followed by two fully connected layers *Crt_fully_1* and *Crt_fully_2* using tanh(·) activation function layers. The last two nodes represent the output layers formation for the Q-value approximation i.e., *Q_value*, using the regression layer.

5.4.3 Comparison of Proposed with Benchmark

The simulation results of the proposed solution are compared with the benchmark results. Initially, two different comparisons of the sum secrecy rate of trusted devices for the proposed and benchmark cases are illustrated with respect to the change in $P_{\max}$ and N, respectively. Later, the effective sum throughput of all the legitimate devices for the proposed and benchmark cases at different values of $P_{\max}$ is provided.

5.4.3.1 Sum Secrecy Rate of Trusted Devices

First, we compare the sum secrecy rate of the trusted devices for the proposed and the benchmark scenarios by varying $P_{\max}$ from -10 dB to 30 dB for IRS $N = 128$ elements. Fig. 5.5 plots the sum secrecy rate of the trusted devices in bps/Hz versus the maximum transmit power $P_{\max}$, for the proposed DRL and the benchmark cases. We can see that for all the cases, the sum secrecy increases as $P_{\max}$ increases, which is to be expected. It can be seen from Fig. 5.5 that the proposed DRL solution outperforms the first two benchmarks of no IRS and randomly phased IRS. We can also note that the gap between the proposed and first two benchmarks increases as the value of transmit power increases. The proposed model provides a best-case gain of 15 times when compared to benchmark case 1. This validates the idea of performance improvement for an IRS-assisted network as compared to the network without the assistance of the IRS. When compared to benchmark case 3, our proposed model provides a higher sum secrecy rate till $P_{\max} = 10$ dB but for higher values of $P_{\max}$ the benchmark case 3 is greater than our proposed model. This later increase in the values of case 3 compared to ours is because our models take into account the SINR from trusted as well as from untrusted devices as their legitimacy. But in case 3 untrusted devices (pure eavesdroppers) are not legitimate devices so the achievable rate of trusted devices increases as $P_{\max}$ increases. Overall, our proposed model provides a maximum of $2 - 2.5$ times gain in the secrecy rate of trusted users as compared to all the benchmark cases.

Secondly, we compare the sum secrecy rate of the trusted devices for the proposed and the benchmark scenarios for $N = 32, 64$ and 128 elements, when the BS transmit power $P_{\max} = 10$ dB. Fig. 5.6 plots the sum secrecy rate of the trusted devices (in bps/Hz) versus the maximum transmit power $P_{\max}$, for the proposed DRL and the benchmark cases. It can be observed that the secrecy rate of our proposed DRL model is greater than all benchmarks for three different values of N. From the result, we can see that the secrecy rate is not affected by the number of IRS elements for the first two cases with no IRS and when IRS elements are randomly phased. Whereas, for our proposed solution and benchmark case 3, the sum secrecy rate increases as the number of IRS elements increases. This is because more elements are optimally phased to enhance the received signal power at the devices. The percentage increase for benchmark case 3 is more as compared to our proposed case, as the untrusted devices are not legitimate in case 3 but our model also takes into account the legitimacy of untrusted devices while evaluating the secrecy rate at trusted devices.

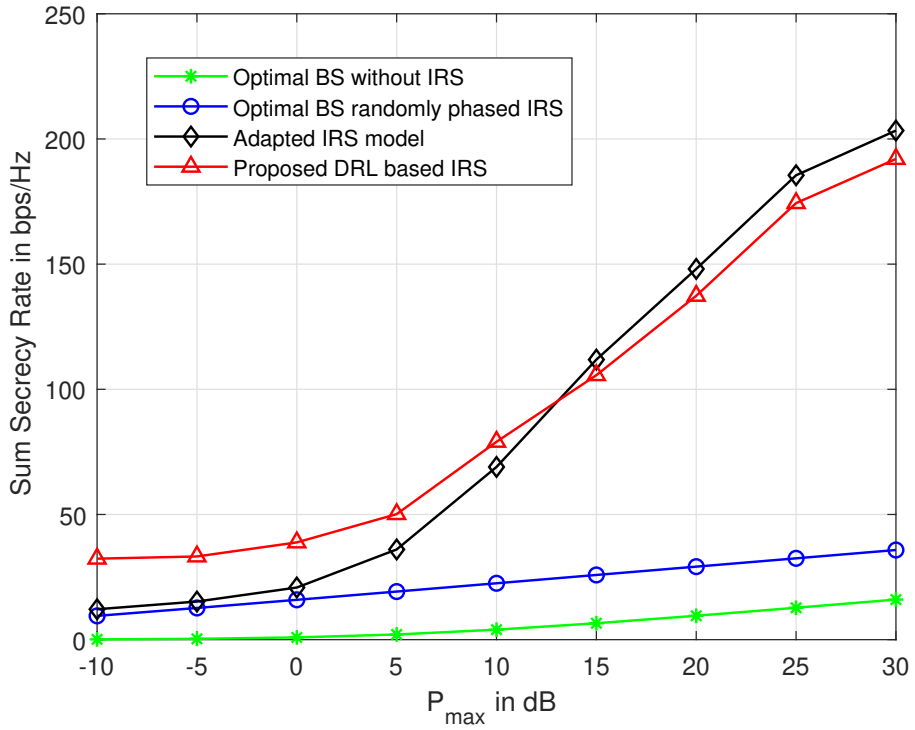


Figure 5.5: The sum secrecy rate of trusted devices for the proposed and benchmark cases at different values of $P_{\max}$.

5.4.3.2 Sum Effective Throughput of all Legitimate Devices

Now we compare the sum effective throughput of all the legitimate devices for the proposed and the benchmark scenarios for $N = 128$ elements. For our model and benchmark case, 1 and 3 the legitimate devices are all trusted and untrusted devices whereas for benchmark cases 3 only trusted devices are legitimate devices. Fig. 5.7 plots sum throughput in bps/Hz versus the maximum transmits power $P_{\max}$, for the proposed DRL and the benchmark cases. We can see that for all the cases, the sum throughput increases as $P_{\max}$ increases, which is to be expected. More importantly, it can be seen that our proposed DRL model always provides higher sum throughput than all the benchmark cases. The gap between our proposed model and benchmark case 3 remains nearly the same for all values of $P_{\max}$. On the other hand, the gap between our model and benchmark case 1 and 2 increases, as $P_{\max}$ increases. This is because of increased received signal strength for our optimally phased IRS model at a higher BS transmit power.

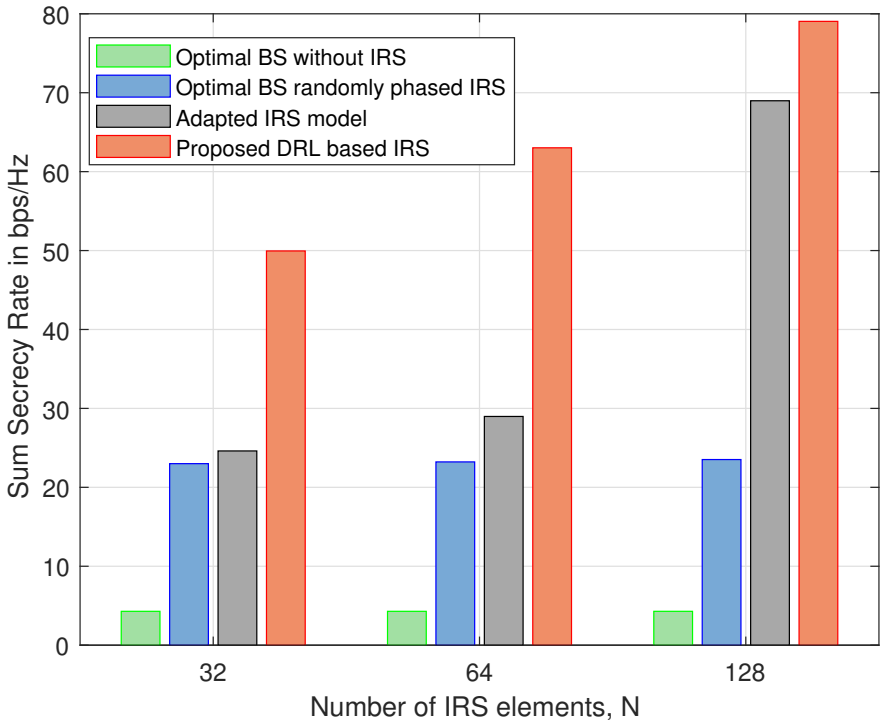


Figure 5.6: The sum secrecy rate of trusted devices for the proposed and benchmark cases at different values of N.

5.4.4 Confirmation of the Proposed Solution to the Problem Constraints

The constraints of the problem in (5.14) are verified through the following results.

5.4.4.1 Verification of SINR Constraint

As we know that our proposed DRL solution guarantees the throughput performance for all the legitimate devices either trusted or untrusted via the constraint $C_2 : \gamma_k \geq \gamma_{\min}, \forall k \in \mathcal{K}$ in (5.14), where γ_k is the achieved SINR of device k $\gamma_{\min}$ is the minimum SINR threshold. We calculate the average SINR of all devices as

$$\gamma_{\text{avg}} = \frac{\sum_{k \in \mathcal{K}} \gamma_k}{K}, \tag{5.30}$$

Fig. 5.8 plots the SINR for all devices versus the simulation index. The result shows three plots of the threshold SINR, the minimum achieved SINR, and the average of achieved SINR for all devices. It can be observed that the minimum value of γ_k is always greater than the threshold, $\gamma_{\min} =$

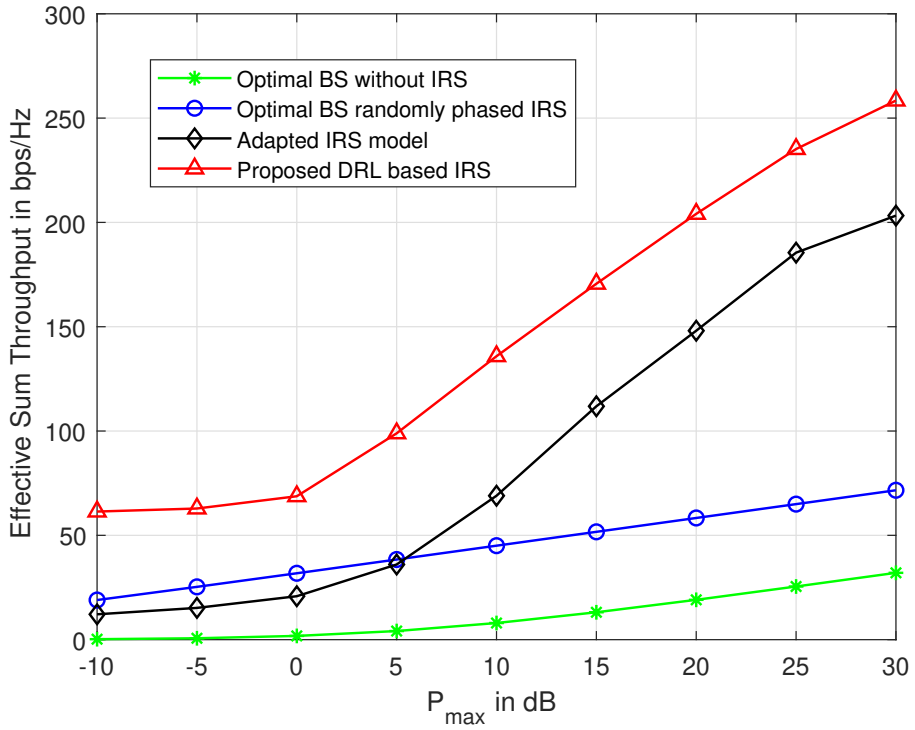


Figure 5.7: The effective sum throughput of all the legitimate devices for the proposed and benchmark cases at different values of $P_{\max}$.

-40 dB. This result confirms that the throughput condition is maintained. In addition, the higher values of the average SINR are a valid indication of the DRL performance under our configuration of the hyper-parameters.

5.4.4.2 Verification of Transmit Power Constraint

The proposed DRL solution requires the condition $C_1 : tr\left\{\mathbf{W^H W}\right\} \leq P_{\max}$ in (5.14) to be met. We take the condition into account while training the DDPG agents.

Fig. 5.9 plots the transmit power of BS versus the simulation index for the target maximum value and the one achieved through the DRL model i.e. the trace of the beamforming matrix, $tr\left\{\mathbf{W^H W}\right\}$. This result is plotted for $P_{\max} = 30$ dB. It can be observed that the trace value is always less than $P_{\max}$. Therefore, our proposed DRL solution satisfies the transmit power constraint.

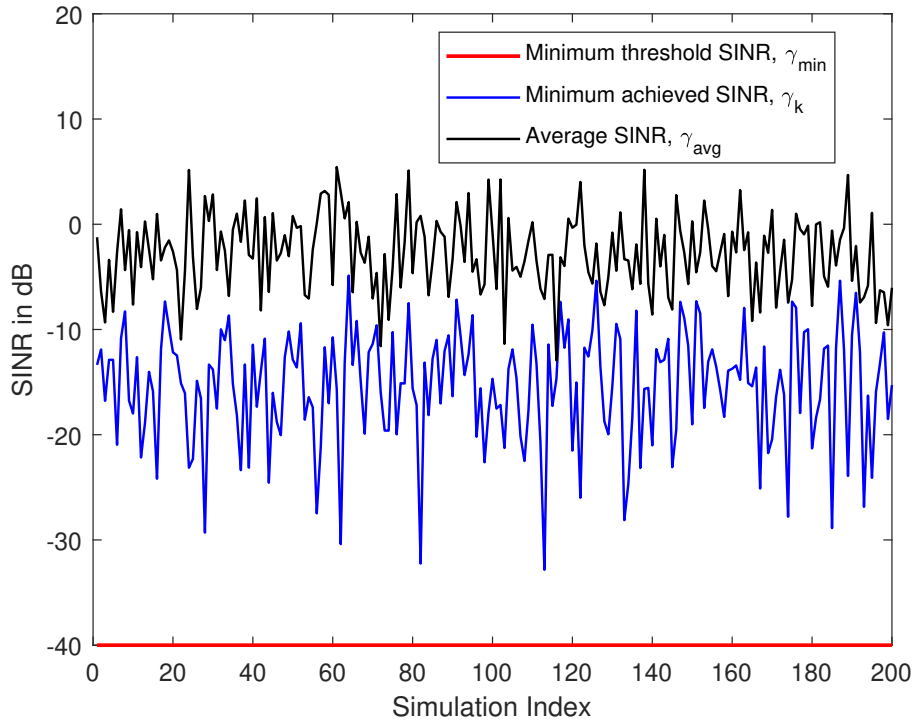


Figure 5.8: The minimum QoS constraint of all the trusted and untrusted devices for the proposed DRL solution.

5.4.5 Effect of System Parameters and DNN Hyper-parameters on the Proposed DRL Model

The sensitivity of the DDPG model is also analyzed over the different values of the system parameters and the hyper-parameters.

5.4.5.1 Effect of Maximum Transmit Power on Reward

DRL performance is affected by different values of system parameters. We examine our DRL training model in terms of the convergence of the reward calculated per episode. Fig. 5.10 plots the average reward and instantaneous reward of the training model versus the number of time steps for two values of maximum transmit power, i.e, 20 dB and 10 dB. The instantaneous reward value is calculated at each step and then it is averaged over each episode. It can be seen that the instantaneous and reward rewards follow the same trend. As our reward is the sum secrecy rate of trusted devices, we can see that the reward values increase when $P_{\max}$ increases from 10 dB to 20 dB. Moreover, the reward value converges faster for a lower value of $P_{\max}$ because the learning agents have a smaller range of data to select from.

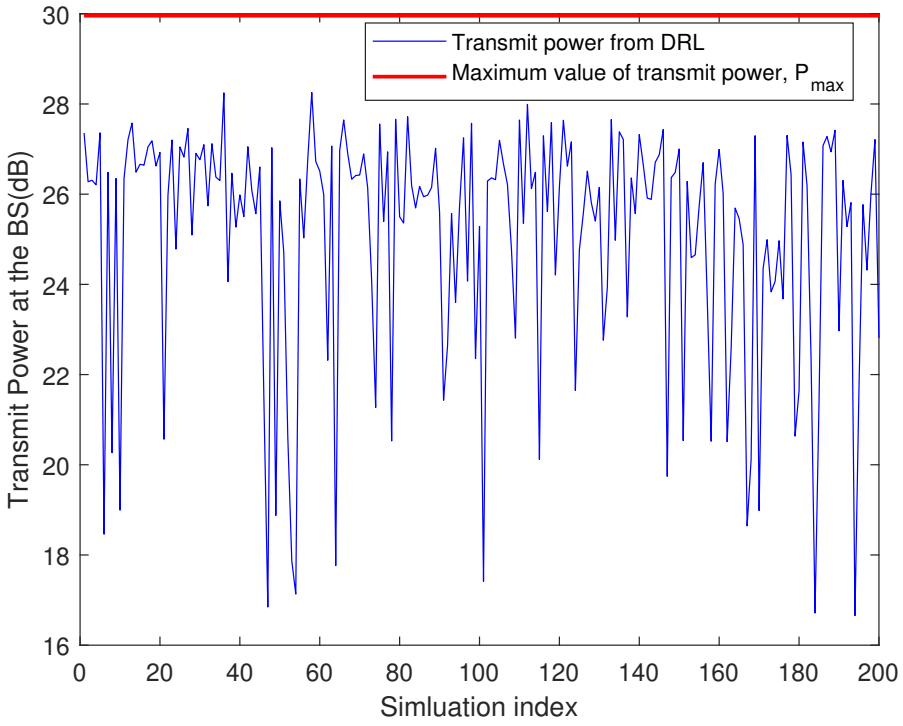


Figure 5.9: The maximum transmit power constraint for active beamforming matrix at $P_{\max} = 30$ dB.

5.4.5.2 Effect of Learning Rates of Agents on Reward

The DRL model is very sensitive to the values of the hyper-parameters. Therefore, the selection of one fixed value of each hyper-parameter is very important to ensure effective learning. This is interpreted in terms of average reward which shows if the learning process is headed in the right direction. Fig. 5.11 plots the average reward versus the number of time steps for five different values i.e., $(0.1, 0.01, 0.001, 0.0001, 0.00001)$ of learning rates ($\alpha_{\text{a1}} = \alpha_{\text{c1}} = \alpha_{\text{a2}} = \alpha_{\text{c2}}$) of agents. It can be seen that when the learn rate value is too low e.g., 0.00001 the average reward value is lower and it converges very soon which indicates that it is not a suitable value for the DRL. Similarly, when the learning rate is too high e.g., 0.1 the reward value is still low. Fig. 5.11 shows that the DRL performs to its best potential at a moderate value of learning rate i.e., 0.001. At the rate of 0.001, it takes longer to converge but the average reward is considerably greater, compared to other rate values.

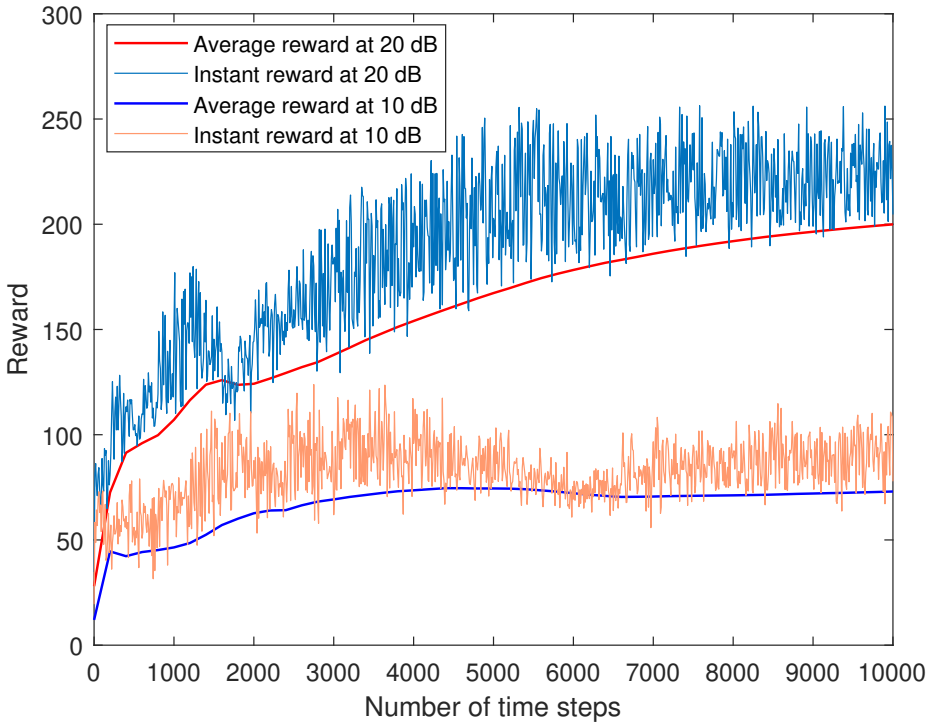


Figure 5.10: DRL instantaneous and average reward for $P_{\max} = 20$ dB and $P_{\max} = 10$ dB.

5.5 Conclusion

In this study, a critical problem of secure communication in diverse multi-device IoT environments was discussed. The trusted-untrusted classification of legitimate devices was taken into account to address different levels of network security in the presence of potential eavesdroppers. A joint active and passive beamforming optimization problem was formulated to maximize the sum secrecy rate of trusted devices while ensuring a performance guarantee to all trusted and untrusted devices. A DDPG-based algorithm was proposed to obtain the optimal IRS phases and transmit the beamforming matrix.

Simulation results were compared with existing cases of optimal BS with no IRS-assisted communication, optimal BS with randomly-phased IRS-assisted communication and an adaptation of jointly optimized active and passive beamforming for IRS-assisted network in the presence of pure eavesdroppers. The results showed a maximum gain of $2 - 2.5$ times in the secrecy rate of trusted devices for the proposed scheme. In addition, the throughput threshold for all trusted and untrusted devices was maintained. The performance sensitivity of the proposed DRL model

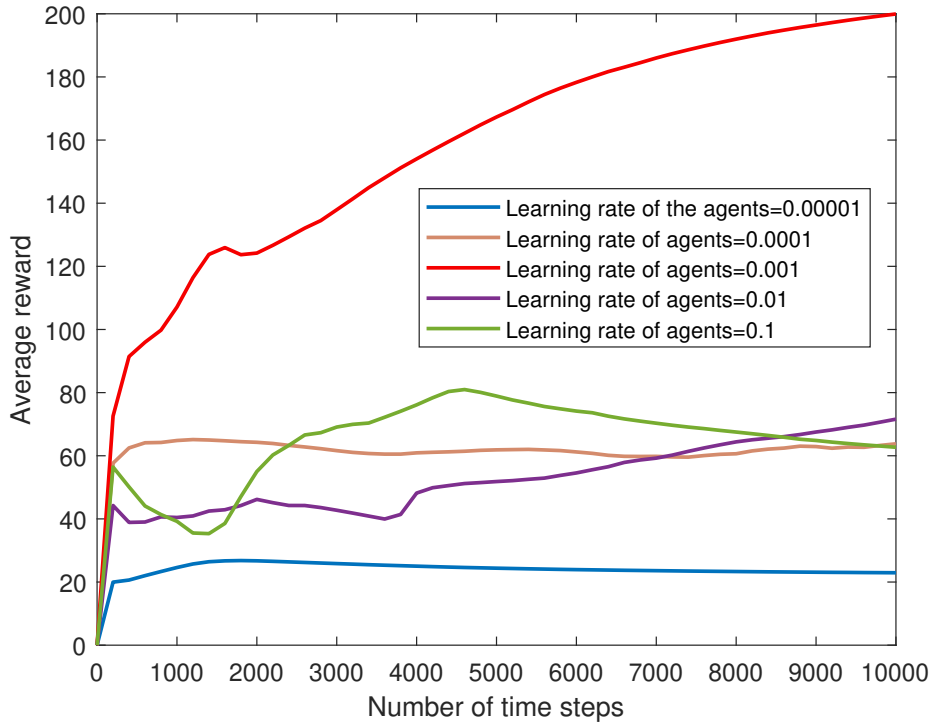


Figure 5.11: The effect of different learning rates of agents on DRL performance.

was analyzed to validate the best values of hyper-parameters.

Chapter 6

Conclusion

The thesis explored research problems on the road towards 6G with the aim to provide efficient, secure, and smart wireless communication. Mainly addressing issues of bandwidth hunger in 5G NR, IoT diversity, severe signal attenuation, and system secrecy for 6G networks. The Figure 6.1 summarizes the key challenges and their offered solutions during the journey. A fair unlicensed spectrum sharing mechanism was proposed to assist the bandwidth-hungry cellular network. A friendly coexistence between Wi-Fi and the diverse IoT-NR was enabled in the unlicensed band. An IRS-assisted communication was introduced to cope with blockages and severe signal attenuation at high frequency. A secure 6G-IoT network with a diverse levels of security was ensured.

The following subsections conclude the key findings, future research direction and overall summary of the thesis.

6.1 Key Findings of Thesis

The three major research works in chapters 3, 4 and 5 have a vital contribution to the thesis. Similarly, the solution presented in each of these chapters adds its own significance to solve different nature of problems.

Firstly, chapter 3 was focused on developing an optimal coexistence mechanism for Wi-Fi and LAA networks offering both inter-network and intra-network fairness. A non-overlapping transmission policy was designed between the networks to exploit the unlicensed channel. As opposed to the conventional idea of constant probability, a different transmission probability for each LAA station was configured which yielded proportionally fair resource utilization among the LAA stations. A joint optimization problem was formulated to maximize the total throughput of the LAA network while guaranteeing a fair throughput share for the Wi-Fi network. The analytical optimal solution was obtained, and the results confirmed more than 75 % throughput gain for the proposed scheme is compared to the benchmark scheme. Also, the proportional intra-network

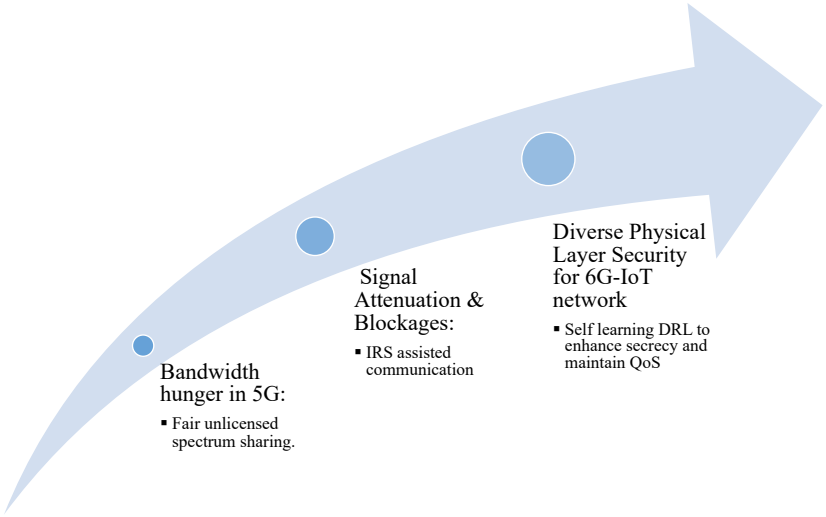


Figure 6.1: Towards efficient, secure, and smart 6G communication.

fairness index depicted an improved gain of $8-9$ % for the proposed LAA network over the conventional LAA network.

Secondly, chapter 4 introduced an inclusive 5G NR network with diverse IoT devices to coexist in the unlicensed spectrum with Wi-Fi devices. The idea of dynamically adaptive ISD was presented instead of uniform ISD to serve the transmission time diversity of IoT-NR devices. An optimization problem was formulated to maximize the sum throughput of the network with a bearable performance warranty of the Wi-Fi network. The problem was dynamic in nature and without the exact information-analytical solution was computationally exhaustive. Thus, a QL framework was proposed to learn from the dynamically changing environment and converge to an optimal solution. The simulation results provided a 51% maximum gain in the normalized sum throughput as compared to the benchmark scheme of fixed ISD.

Finally, chapter 5 addressed a diverse level of secure communication for a 6G-IoT network where the untrusted devices could potentially eavesdrop on the trusted devices. An IRS-aided communication was developed to cope with signal attenuation and blockages at the mmWave/THz band. A mathematical model was obtained for the proposed system and an optimization problem was formulated to maximize the sum secrecy rate of trusted devices with a performance guarantee for both trusted and untrusted devices. Suitable to the complex and continuous nature of the problem a DRL-based joint active and passive beamforming solution was proposed. The proposed DDPG algorithm yielded a maximum gain of $2-2.5$ times for the sum secrecy rate of trusted devices and a throughput assurance of all trusted and untrusted devices. The performance of the

proposed DDPG model was evaluated under meticulously selected hyper-parameters.

6.2 Future Potential Research Direction

The research problems visited in this thesis are complex and dynamic in nature and the solutions are highly flexible to adapt the different network conditions. Also, the proposed system models and solutions can influence a broader horizon of similar problems for other wireless networks, not only limited to 5G NR, IoT, IRS and THz communications. As already mentioned in Chapter 1, there are open research challenges to mark the safety, security and capacity of next-generation wireless communications [11]. Some of these potential research directions are mentioned as follows:

6.2.1 Non-terrestrial Communication

The journey of wireless communication networks from 5G to 6G envisions a new era of bringing non-terrestrial communication, autonomous IoT connectivity, AI and UAVs on a single platform [202]. The future use cases forecast a UAV-based cellular communication to empower the 6G technology for achieving higher reliability and exceptional performance [5]. Although the marriage of aerial communications networks with terrestrial networks is being acknowledged to be the most promising trend of future wireless networks then this may arise multiple challenges which need to be investigated. Some of the challenges are channel modelling, privacy and security, mobility and trajectory management, spectrum coexistence, 3D deployment, addressing UAVs diversity in terms of dynamic load traffic, wireless capacity, and delay between the satellite and terrestrial radio links [203–205].

6.2.2 Internet of Everything (IoE)

The next age of IoT communication is inclining towards a progressive paradigm of IoE. The concept is to connect everyone and everything, combine their data, process it into useful information, and make beneficial decisions for society [206]. To facilitate the features of IoE, there is a need for a compatible design of future 6G networks which can accommodate machines and people [207]. As an example of such a project, coupling the human, physical, and digital world for future 6G infrastructure is Hexa-X i.e., aimed to be deployed by 2030 [208]. However, with the advent of IoT, there are issues such as massive connectivity, autonomous control, complex computing, ultra-high reliability, security, and sustainability that need to be investigated [209].

These emerging technologies are imagined to have large-scale frameworks and radically complex computing requirements. Hence, AI and wireless communication are bonded together to influence self-sustaining, sophisticated, and intelligent architecture for the future. The learning-driven solutions proposed in the thesis can contribute toward a smart realization of similar research problems.

6.3 Summary

To summarise, this thesis addressed the critical problems of spectrum sharing, bandwidth scarcity, signal attenuation, and diverse levels of security in the current and emerging wireless communication networks. Also, it provided optimization and learning based solutions to these problems. Moreover, it highlighted the directions and challenges for the future research problems widely for IoE, and non-terrestrial 6G communications.

Bibliography

1. R. Saleem, S. A. Alvi, and S. Durrani, "Performance-fairness trade-off for Wi-Fi and LTE-LAA coexistence," *IEEE Access*, vol. 9, pp. 62 446–62 459, Apr. 2021.

2. R. Saleem, W. Ni, M. Ikram, and A. Jamalipour, "Deep reinforcement learning-driven secrecy design for intelligent reflecting surface-based 6G-IoT networks," *IEEE Internet of Things Journal*, pp. 1–1, 2022.

3. R. Saleem, W. Ni, and M. Ikram, "Reinforcement learning-based unlicensed spectrum sharing for IoT Devices of 5G new radio," in *2022 IEEE International Mediterranean Conference on Communications and Networking (MeditCom)*, Athens, Greece, Sep. 2022, pp. 191–196.

4. Q. Cui, W. Ni, S. Li, B. Zhao, R. P. Liu, and P. Zhang, "Learning-assisted clustered access of 5G/B5G networks to unlicensed spectrum," *IEEE Wireless Communications*, vol. 27, no. 1, pp. 31–37, Feb. 2020.

5. M. Vaezi, A. Azari, S. R. Khosravirad, M. Shirvanimoghaddam, M. M. Azari, D. Chasaki, and P. Popovski, "Cellular, wide-area, and non-terrestrial IoT: A survey on 5G advances and the road toward 6G," *IEEE Communications Surveys Tutorials*, vol. 24, no. 2, pp. 1117–1174, Feb. 2022.

6. W. Jiang, B. Han, M. A. Habibi, and H. D. Schotten, "The road towards 6G: a comprehensive survey," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 334–366, Feb. 2021.

7. X. Chen, D. W. K. Ng, W. Yu, E. G. Larsson, N. Al-Dhahir, and R. Schober, "Massive access for 5G and beyond," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 3, pp. 615–637, Mar. 2021.

8. C. L. Stergiou, K. E. Psannis, and B. B. Gupta, "IoT-based big data secure management in the fog over a 6G wireless network,"

IEEE Internet of Things Journal, vol. 8, no. 7, pp. 5164–5171, Apr. 2021.

9. Y. Xu, G. Gui, H. Gacanin, and F. Adachi, "A survey on resource allocation for 5G heterogeneous networks: Current research, future trends, and challenges," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 668–695, Feb. 2021.

10. C. D. Alwis, A. Kalla, Q.-V. Pham, P. Kumar, K. Dev, W.-J. Hwang, and M. Liyanage, "Survey on 6G frontiers: Trends, applications, requirements, technologies and future research," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 836–886, Apr. 2021.

11. B. Ji, Y. Han, S. Liu, F. Tao, G. Zhang, Z. Fu, and C. Li, "Several key technologies for 6G: challenges and opportunities," *IEEE Communications Standards Magazine*, vol. 5, no. 2, pp. 44–51, Jun. 2021.

12. J. Zhang, E. Björnson, M. Matthaiou, D. W. K. Ng, H. Yang, and D. J. Love, "Prospective multiple antenna technologies for beyond 5G," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1637–1660, Jun. 2020.

13. Y. Liu, M. Peng, G. Shou, Y. Chen, and S. Chen, "Toward edge intelligence: Multiaccess edge computing for 5G and internet of things," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 6722–6747, Jun. 2020.

14. B. Li, Z. Fei, and Y. Zhang, "UAV communications for 5G and beyond: Recent advances and future trends," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2241–2263, Dec. 2019.

15. A. Guidotti, A. Vanelli-Coralli, M. Conti, S. Andrenacci, S. Chatzinotas, N. Maturo, B. Evans, A. Awoseyila, A. Ugolini, T. Foggi, L. Gaudio, N. Alagha, and S. Cioni, "Architectures and key technical challenges for 5G systems incorporating satellites," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 2624–2639, Jan. 2019.

16. O. Kodheli, E. Lagunas, N. Maturo, S. K. Sharma, B. Shankar, J. F. M. Montoya, J. C. M. Duncan, D. Spano, S. Chatzinotas, S. Kisseleff, J. Querol, L. Lei, T. X. Vu, and G. Goussetis, "Satellite communications in the new space era: A survey and future challenges," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 70–109, Jan. 2021.

17. H.-J. Song and N. Lee, "Terahertz communications: Challenges in the next decade," *IEEE Transactions on Terahertz Science and Technology*, vol. 12, no. 2, pp. 105–117, Nov. 2022.

18. S. Gong, X. Lu, D. T. Hoang, D. Niyato, L. Shu, D. I. Kim, and Y.-C. Liang, "Toward smart wireless communications via intelligent reflecting surfaces: A contemporary survey," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 4, pp. 2283–2314, Jun. 2020.

19. C. Sun, W. Ni, Z. Bu, and X. Wang, "Energy minimization for intelligent reflecting surface-assisted mobile edge computing," *IEEE Transactions on Wireless Communications*, pp. 1–1, Feb. 2022.

20. M. Z. Chowdhury, M. Shahjalal, S. Ahmed, and Y. M. Jang, "6G wireless communication systems: Applications, requirements, technologies, challenges, and research directions," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 957–975, Jul. 2020.

21. M. A. Uusitalo, P. Rugeland, M. R. Boldi, E. C. Strinati, P. Demestichas, M. Ericson, G. P. Fettweis, M. C. Filippou, A. Gati, M.-H. Hamon, M. Hoffmann, M. Latva-Aho, A. Pärssinen, B. Richerzhagen, H. Schotten, T. Svensson, G. Wikström, H. Wymeersch, V. Ziegler, and Y. Zou, "6G vision, value, use cases and technologies from european 6G flagship project Hexa-X," *IEEE Access*, vol. 9, pp. 160 004–160 020, Nov. 2021.

22. A. Dogra, R. K. Jha, and S. Jain, "A survey on beyond 5G network with the advent of 6G: architecture and emerging technologies," *IEEE Access*, vol. 9, pp. 67 512–67 547, Oct. 2021.

23. K. B. Letaief, W. Chen, Y. Shi, J. Zhang, and Y.-J. A. Zhang, "The roadmap to 6G: AI empowered wireless networks," *IEEE Communications Magazine*, vol. 57, no. 8, pp. 84–90, Aug. 2019.

24. M. Wang, Y. Cui, X. Wang, S. Xiao, and J. Jiang, "Machine learning for networking: Workflow, advances and opportunities," *IEEE Network*, vol. 32, no. 2, pp. 92–99, Mar. 2018.

25. C. Jiang, H. Zhang, Y. Ren, Z. Han, K. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," *IEEE Transactions on Wireless Communications*, vol. 24, no. 2, pp. 98–105, Apr. 2017.

26. J. Moysen and L. Giupponi, "From 4G to 5G: Self-organized network management meets machine learning," *Computer Communications*, vol. 129, pp. 248–268, Jul. 2018.

27. E. Alpaydin, *Introduction to Machine Learning*. MIT Press, 2014.

28. M. Usama, J. Qadir, A. Raza, H. Arif, K. A. Yau, Y. Elkhatib, A. Hussain, and A. Al-Fuqaha, "Unsupervised machine learning for networking: Techniques, applications and research challenges," *IEEE Access*, vol. 7, pp. 65 579–65 615, May 2019.

29. R. S. Sutton and A. G. Barto, *Introduction to Reinforcement Learning*, 1st ed. Cambridge, MA, USA: MIT Press, 1998.

30. M. Martin. (2011, Autumn) *Reinforcement Learning Searching for optimal policies I:Bellman equations and optimal policies*. [Online]. Available: http://www.lsi.upc.es/~mmartin/Ag4.pdf

31. F. Pacheco, E. Exposito, M. Gineste, C. Baudoin, and J. Aguilar, "Towards the deployment of machine learning solutions in network traffic classification: A systematic survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1988–2014, Nov. 2019.

32. J. Liu, K. Mei, X. Zhang, D. Ma, and J. Wei, "Online extreme learning machine-based channel estimation and equalization for OFDM systems," *IEEE Communications Letters*, Early access 2019.

33. C. A. Oroza, Z. Zhang, T. Watteyne, and S. D. Glaser, "A machine-learning-based connectivity model for complex terrain large-scale low-power wireless deployments," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 4, pp. 576–584, Dec. 2017.

34. T. Wang, C. Wen, H. Wang, F. Gao, T. Jiang, and S. Jin, "Deep learning for wireless physical layer: Opportunities and challenges," *China Communication*, vol. 14, no. 11, pp. 92–111, Nov. 2017.

35. E. Nachmani, E. Marciano, L. Lugosch, W. J. Gross, D. Burshtein, and Y. Be'ery, "Deep learning methods for improved decoding of linear codes," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 119–131, Feb. 2018.

36. P. V. Klaine, M. A. Imran, O. Onireti, and R. D. Souza, "A survey of machine learning techniques applied to self-organizing cellular networks," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2392–2431, Jul. 2017.

37. M. Bkassiny, Y. Li, and S. K. Jayaweera, "A survey on machine-learning techniques in cognitive radios," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1136–1159, Oct. 2013.

38. M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, and M. Debbah, "A tutorial on uavs for wireless networks: Applications, challenges, and open problems," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2334–2360, Mar. 2019.

39. M. Chen, W. Saad, and C. Yin, "Liquid state machine learning for resource and cache management in LTE-U unmanned aerial vehicle (UAV) networks," *IEEE Transactions on Wireless Communications*, vol. 18, no. 3, pp. 1504–1517, Mar. 2019.

40. N. C. Luong, D. T. Hoang, S. Gong, D. Niyato, P. Wang, Y.-C. Liang, and D. I. Kim, "Applications of deep reinforcement learning in communications and networking: A survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3133–3174, May. 2019.

41. L. Liang, H. Ye, and G. Y. Li, "Toward intelligent vehicular networks: A machine learning framework," *IEEE Internet of Things Journal*, vol. 6, no. 1, pp. 124–135, Feb. 2019.

42. X. Wang, Y. Wang, Q. Cui, K.-C. Chen, and W. Ni, "Machine learning enables radio resource allocation in the downlink of ultra-low latency vehicular networks," *IEEE Access*, vol. 10, pp. 44 710–44 723, Apr. 2022.

43. J. Ren, H. Wang, T. Hou, S. Zheng, and C. Tang, "Federated learning-based computation offloading optimization in edge computing-supported internet of things," *IEEE Access*, vol. 7, pp. 69 194–69 201, Jun. 2019.

44. Q. Cui, Z. Zhu, W. Ni, X. Tao, and P. Zhang, "Edge-intelligence-empowered, unified authentication and trust evaluation for heterogeneous beyond 5G systems," *IEEE Wireless Communications*, vol. 28, no. 2, pp. 78–85, Apr. 2021.

45. Y. Xin, L. Kong, Z. Liu, Y. Chen, Y. Li, H. Zhu, M. Gao, H. Hou, and C. Wang, "Machine learning and deep learning methods for cybersecurity," *IEEE Access*, vol. 6, pp. 35 365–35 381, May 2018.

46. X. Wen, S. Bi, X. Lin, L. Yuan, and J. Wang, "Throughput maximization for ambient backscatter communication: A reinforcement learning approach," in *Proc. IEEE Inf. Technol. Netw. Electron. Autom. Control Conf. (ITNEC)*, Mar. 2019, pp. 997–1003.

47. P. V. R. Ferreira, R. Paffenroth, A. M. Wyglinski, T. M. Hackett, S. G. Bilén, R. C. Reinhart, and D. J. Mortensen, "Multiobjective reinforcement learning for cognitive satellite communications using deep neural network ensembles," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 5, pp. 1030–1041, May 2018.

48. J. Xie, F. R. Yu, T. Huang, R. Xie, J. Liu, C. Wang, and Y. Liu, "A survey of machine learning techniques applied to software defined networking (SDN): Research issues and challenges," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 393–430, Aug. 2019.

49. F. Guo, F. R. Yu, H. Zhang, X. Li, H. Ji, and V. C. M. Leung, "Enabling massive IoT toward 6G: A comprehensive survey," *IEEE Internet of Things Journal*, vol. 8, no. 15, pp. 11 891–11 915, Mar. 2021.

50. G. Gür, "Expansive networks: Exploiting spectrum sharing for capacity boost and 6G vision," *Journal of Communications and Networks*, vol. 22, no. 6, pp. 444–454, Dec. 2020.

51. S. Muhammad, H. H. Refai, and M. O. Al Kalaa, "5G NR-U: homogeneous coexistence analysis," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, Dec. 2020, pp. 1–6.

52. X. Lu, V. Petrov, D. Moltchanov, S. Andreev, T. Mahmoodi, and M. Dohler, "5G-U: conceptualizing integrated utilization of licensed and unlicensed spectrum for future IoT," *IEEE Communications Magazine*, vol. 57, no. 7, pp. 92–98, May 2019.

53. C. Yuan, X. Tao, W. Ni, N. Li, A. Jamalipour, and R. P. Liu, "Joint power allocation and beamforming for overlaid secrecy transmissions in MIMO-OFDM channels," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 9, pp. 10 019–10 032, Jun. 2020.

54. Kamaldeep, M. Dutta, and J. Granjal, "Towards a secure internet of things: A comprehensive study of second line defense mechanisms," *IEEE Access*, vol. 8, pp. 127 272–127 312, Jun. 2020.

55. J. Zheng, K. Li, N. Mhaisen, W. Ni, E. Tovar, and M. Guizani, "Exploring deep reinforcement learning-assisted federated learning for online resource allocation in privacy-preserving EdgeIoT," *IEEE Internet of Things Journal*, pp. 1–1, May 2022.

56. W. Gao, C. Han, and Z. Chen, "DNN-powered SIC-free receiver artificial noise aided terahertz secure communications with randomly distributed eavesdroppers," *IEEE Transactions on Wireless Communications*, vol. 21, no. 1, pp. 563–576, Jan. 2022.

57. W. Chen, X. Qiu, T. Cai, H.-N. Dai, Z. Zheng, and Y. Zhang, "Deep reinforcement learning for internet of things: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1659–1692, Apr. 2021.

58. O. O. Erunkulu, A. M. Zungeru, C. K. Lebekwe, M. Mosalaosi, and J. M. Chuma, "5G mobile communication applications: A survey and comparison of use cases," *IEEE Access*, vol. 9, pp. 97 251–97 295, Jun. 2021.

59. Y. Shi, Q. Cui, W. Ni, and Z. Fei, "Proactive dynamic channel selection based on multi-armed bandit learning for 5G NR-U," *IEEE Access*, vol. 8, pp. 196 363–196 374, Oct. 2020.

60. W. S. H. M. W. Ahmad, N. A. M. Radzi, F. S. Samidi, A. Ismail, F. Abdullah, M. Z. Jamaludin, and M. N. Zakaria, "5G technology:

Towards dynamic spectrum sharing using cognitive radio networks," *IEEE Access*, vol. 8, pp. 14 460–14 488, Jan. 2020.

61. S. Jacob, V. G. Menon, S. Joseph, P. Vinoj, A. Jolfaei, J. Lukose, and G. Raja, "A novel spectrum sharing scheme using dynamic long short-term memory with CP-OFDMA in 5G networks," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 3, pp. 926–934, Jan. 2020.

62. H. Bao, Y. Huo, X. Dong, and C. Huang, "Joint time and power allocation for 5G NR unlicensed systems," *IEEE Transactions on Wireless Communications*, vol. 20, no. 9, pp. 6195–6209, 2021.

63. D. Candal-Ventureira, F. J. González-Castaño, F. Gil-Castiñeira, and P. Fondo-Ferreiro, "Coordinated allocation of radio resources to Wi-Fi and cellular technologies in shared unlicensed frequencies," *IEEE Access*, vol. 9, pp. 134 435–134 456, Sep. 2021.

64. M. Le, Q.-V. Pham, H.-C. Kim, and W.-J. Hwang, "Enhanced resource allocation in D2D communications with NOMA and unlicensed spectrum," *IEEE Systems Journal*, vol. 16, no. 2, pp. 2856–2866, Jan. 2022.

65. D. Thomas, M. Orgun, M. Hitchens, R. Shankaran, S. C. Mukhopadhyay, and W. Ni, "A graph-based fault-tolerant approach to modeling QoS for IoT-based surveillance applications," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3587–3604, Sep. 2021.

66. X. Lyu, C. Ren, W. Ni, H. Tian, Q. Cui, and R. P. Liu, "Online learning of optimal proactive schedule based on outdated knowledge for energy harvesting powered internet-of-things," *IEEE Transactions on Wireless Communications*, vol. 20, no. 2, pp. 1248–1262, Feb. 2021.

67. S. Wang, T. Lv, W. Ni, N. C. Beaulieu, and Y. J. Guo, "Joint resource management for MC-NOMA: a deep reinforcement learning approach," *IEEE Transactions on Wireless Communications*, vol. 20, no. 9, pp. 5672–5688, Apr. 2021.

68. Z. Wang, Z. Lin, T. Lv, and W. Ni, "Energy-efficient resource allocation in massive MIMO-NOMA networks with wireless power transfer: A distributed ADMM approach," *IEEE Internet of Things Journal*, vol. 8, no. 18, pp. 14 232–14 247, Mar. 2021.

69. Q. Cui, Z. Zhang, Y. Shi, W. Ni, M. Zeng, and M. Zhou, "Dynamic multichannel access based on deep reinforcement learning in distributed wireless networks," *IEEE Systems Journal*, pp. 1–4, Dec. 2021.

70. E. Basar, "Reconfigurable intelligent surface-based index modulation: A new beyond MIMO paradigm for 6G," *IEEE Transactions on Communications*, vol. 68, no. 5, pp. 3187–3196, May 2020.

71. Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1546–1577, May. 2021.

72. Y. Zhu, B. Mao, and N. Kato, "Intelligent reflecting surface in 6G vehicular communications: A survey," *IEEE Open Journal of Vehicular Technology*, pp. 1–1, May 2022.

73. Y. Cao, T. Lv, and W. Ni, "Two-timescale optimization for intelligent reflecting surface-assisted MIMO transmission in fast-changing channels," *IEEE Transactions on Wireless Communications*, pp. 1–1, Jun. 2022.

74. Y. Cao, T. Lv, W. Ni, and Z. Lin, "Sum-rate maximization for multi-reconfigurable intelligent surface-assisted device-to-device communications," *IEEE Transactions on Communications*, vol. 69, no. 11, pp. 7283–7296, Aug. 2021.

75. L. Lv, Q. Wu, Z. Li, Z. Ding, N. Al-Dhahir, and J. Chen, "Covert communication in intelligent reflecting surface-assisted NOMA systems: Design, analysis, and optimization," *IEEE Transactions on Wireless Communications*, vol. 21, no. 3, pp. 1735–1750, Mar. 2022.

76. C. Pan, H. Ren, K. Wang, J. F. Kolb, M. Elkashlan, M. Chen, M. Di Renzo, Y. Hao, J. Wang, A. L. Swindlehurst, X. You, and L. Hanzo, "Reconfigurable intelligent surfaces for 6G systems: Principles, applications, and research directions," *IEEE Communications Magazine*, vol. 59, no. 6, pp. 14–20, Jun. 2021.

77. S. Zhang and R. Zhang, "Intelligent reflecting surface aided multi-user communication: Capacity region and deployment strategy," *IEEE Transactions on Communications*, vol. 69, no. 9, pp. 5790–5806, Sep. 2021.

78. L. Feng, X. Que, P. Yu, W. Li, and X. Qiu, "IRS assisted multiple user detection for uplink URLLC non-orthogonal multiple access," in *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, Toronto, ON, Canada, Jul. 2020, pp. 1314–1315.

79. H. Chen, H. Sarieddeen, T. Ballal, H. Wymeersch, M.-S. Alouini, and T. Y. Al-Naffouri, "A tutorial on terahertz-band localization

for 6G communication systems," *IEEE Communications Surveys & Tutorials*, pp. 1–1, May 2022.

80. K. Guan, H. Yi, D. He, B. Ai, and Z. Zhong, "Towards 6G: Paradigm of realistic terahertz channel modeling," *China Communications*, vol. 18, no. 5, pp. 1–18, May 2021.

81. Z. Chen, X. Ma, B. Zhang, Y. Zhang, Z. Niu, N. Kuang, W. Chen, L. Li, and S. Li, "A survey on terahertz communications," *China Communications*, vol. 16, no. 2, pp. 1–35, Feb. 2019.

82. I. F. Akyildiz, C. Han, Z. Hu, S. Nie, and J. M. Jornet, "Terahertz band communication: An old problem revisited and research directions for the next decade," *IEEE Transactions on Communications*, vol. 70, no. 6, pp. 4250–4285, 2022.

83. S. Hu, X. Chen, W. Ni, E. Hossain, and X. Wang, "Distributed machine learning for wireless communication networks: Techniques, architectures, and applications," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1458–1493, Jun. 2021.

84. C. J. C. H. Watkins and P. Dayan, "Q-learning," *Machine Learning*, vol. 8, no. 3, pp. 279–292, May 1992.

85. A. Alwarafy, M. Abdallah, B. S. Çiftler, A. Al-Fuqaha, and M. Hamdi, "The frontiers of deep reinforcement learning for resource management in future wireless HetNets: Techniques, challenges, and research directions," *IEEE Open Journal of the Communications Society*, vol. 3, pp. 322–365, Feb. 2022.

86. W. Wang and W. Zhang, "Intelligent reflecting surface configurations for smart radio using deep reinforcement learning," *IEEE Journal on Selected Areas in Communications*, pp. 1–1, Jun. 2022.

87. C. She, C. Sun, Z. Gu, Y. Li, C. Yang, H. V. Poor, and B. Vucetic, "A tutorial on ultra-reliable and low-latency communications in 6G: Integrating domain knowledge into deep learning," *Proceedings of the IEEE*, vol. 109, no. 3, pp. 204–246, Mar. 2021.

88. J. Zhang, M. Wang, M. Hua, T. Xia, W. Yang, and X. You, "LTE on license-exempt spectrum," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 1, pp. 647–673, Nov. 2017.

89. Q. Cui, Y. Gu, W. Ni, X. Zhang, X. Tao, P. Zhang, and R. P. Liu, "Preserving reliability of heterogeneous ultra-dense distributed networks in unlicensed spectrum," *IEEE Communications Magazine*, vol. 56, no. 6, pp. 72–78, Jun. 2018.

90. M. Labib, V. Marojevic, J. H. Reed, and A. I. Zaghloul, "Extending LTE into the unlicensed spectrum: Technical analysis of the proposed variants," *IEEE Communications Magazine*, vol. 1, pp. 31–39, Dec. 2017.

91. H. Cui, V. C. M. Leung, S. Li, and X. Wang, "LTE in the unlicensed band: Overview, challenges, and opportunities," *IEEE Transactions on Wireless Communications*, vol. 24, no. 4, pp. 99–105, Feb. 2017.

92. Qualcomm Incorporated. *Progress on LAA and its relationship to LTE-U and MulteFire*. [Online]. Available: https://www.qualcomm.com/media/documents/files/laa-webinar-feb-2016.pdf

93. C. Rosa, M. Kuusela, F. Frederiksen, and K. I. Pedersen, "Standalone LTE in unlicensed spectrum: Radio challenges, solutions, and performance of MulteFire," *IEEE Communications Magazine*, vol. 56, no. 10, pp. 170–177, Oct. 2018.

94. *Study on Licensed-Assisted Access to Unlicensed Spectrum*, 3GPP Std. 36.889 V13.0.0, Jun. 2015.

95. S. Lagen, L. Giupponi, S. Goyal, N. Patriciello, B. Bojović, A. Demir, and M. Beluri, "New radio beam-based access to unlicensed spectrum: Design challenges and solutions," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 1, pp. 8–37, Oct. 2020.

96. N. Patriciello, S. Lagén, B. Bojović, and L. Giupponi, "NR-U and IEEE 802.11 technologies coexistence in unlicensed mmWave spectrum: Models and evaluation," *IEEE Access*, vol. 8, pp. 71 254–71 271, Apr. 2020.

97. M. Hirzallah, M. Krunz, B. Kecicioglu, and B. Hamzeh, "5G new radio unlicensed: Challenges and evaluation," *IEEE Transactions on Cognitive Communications and Networking*, pp. 1–1, Dec. 2020.

98. Qualcomm Incorporated. How does unlicensed spectrum with NR-U transform what 5G can do for you? [Online]. Available: https://www.qualcomm.com/media/documents/files/presentation-how-nru-can-transform-what-5g-can-do-for-you.pdf

99. P. Yang, L. Kong, and G. Chen, "Spectrum sharing for 5G/6G URLLC: research frontiers and standards," *IEEE Communications Standards Magazine*, vol. 5, no. 2, pp. 120–125, Apr. 2021.

100. 3GPP. *3GPP RAN1 meeting on the standardization of NR-Unlicensed. Date: 2020-01-16. Doc.: IEEE 802.11-20/0195r0.*

101. V. Sathya, M. I. Rochman, and M. Ghosh, "Measurement-based coexistence studies of LAA Wi-Fi deployments in Chicago," *IEEE Transactions on Wireless Communications*, vol. 28, no. 1, pp. 1–8, Nov. 2020.

102. I. Latif. (Jul. 2019.) Efficient and fair medium sharing enabled by a common preamble. [Online]. Available: http://grouper.ieee.org/groups/802/11/Workshops/2019-July-Coex/Quantenna{_}Contribution.pdf

103. *IEEE Standard for Information Technology-Telecommunications and Information Exchange Between Systems Local and Metropolitan Area Networks-Specific Requirements Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Std. 802.11-2016 (Revision of IEEE Std. 802.11-2012), Dec. 2016.

104. B. Chen, J. Chen, Y. Gao, and J. Zhang, "Coexistence of LTE-LAA and Wi-Fi on 5 GHz with corresponding deployment scenarios: A survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 7–32, Jul. 2016.

105. N. Bitar, M. O. Al Kalaa, S. J. Seidman, and H. H. Refai, "On the coexistence of LTE-LAA in the unlicensed band: Modeling and performance analysis," *IEEE Access*, vol. 6, pp. 52 668–52 681, Sep. 2018.

106. E. Pei and J. Jiang, "Performance analysis of licensed-assisted access to unlicensed spectrum in LTE Release 13," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1446–1458, Feb. 2019.

107. Y. Gao, X. Chu, and J. Zhang, "Performance analysis of LAA and WiFi coexistence in unlicensed spectrum based on Markov chain," in *2016 IEEE Global Communications Conference (GLOBECOM)*, Washington, DC, USA, Dec. 2016, pp. 1–6.

108. J. Xiao, J. Zheng, L. Chu, and Q. Ren, "Performance modeling and analysis of the LAA Category-4 LBT procedure," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 10 045–10 055, Jul. 2019.

109. Y. Gao and S. Roy, "Achieving proportional fairness for LTE-LAA and Wi-Fi coexistence in unlicensed spectrum," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3390–3404, Feb. 2020.

110. Y. Gao, "LTE-LAA and WiFi in 5G NR unlicensed: Fairness, optimization and win-win solution," in *IEEE SmartWorld,*

Ubiquitous Intelligence Computing, Advanced Trusted Computing, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), Leicester, UK, Aug. 2019, pp. 1638–1643.

111. R. Yin, G. Y. Li, and A. Maaref, "Spatial reuse for coexisting LTE and Wi-Fi systems in unlicensed spectrum," *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 1187–1198, Feb. 2018.

112. C. S. Yang, C. K. Kim, J. Moon, S. Park, and C. G. Kang, "Channel access scheme with alignment reference interval adaptation (aria) for frequency reuse in unlicensed band LTE: Fuzzy Q-learning approach," *IEEE Access*, vol. 6, pp. 26 438–26 451, Jun. 2018.

113. M. K. Maheshwari, A. Roy, and N. Saxena, "DRX over LAA-LTE-A new design and analysis based on Semi-Markov model," *IEEE Transactions on Mobile Computing*, vol. 18, no. 2, pp. 276–289, Feb. 2019.

114. Q. Chen, G. Yu, H. Shan, A. Maaref, G. Y. Li, and A. Huang, "Cellular meets WiFi: Traffic offloading or resource sharing?" *IEEE Transactions on Wireless Communications*, vol. 15, no. 5, pp. 3354–3367, May 2016.

115. H. Lee and H. J. Yang, "Downlink mu-mimo lte-laa for coexistence with asymmetric hidden wi-fi aps," *IEEE Transactions on Mobile Computing*, vol. 21, no. 1, pp. 93–109, Jun. 2022.

116. U. Challita, L. Dong, and W. Saad, "Proactive resource management for LTE in unlicensed spectrum: A deep learning perspective," *IEEE Transactions on Wireless Communications*, vol. 17, no. 7, pp. 4674–4689, Jul. 2018.

117. C. K. Kim, C. S. Yang, and C. G. Kang, "Adaptive listen-before-talk (LBT) scheme for LTE and Wi-Fi systems coexisting in unlicensed band," in *2016 13th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, Las Vegas, NV, USA, Jan. 2016, pp. 589–594.

118. Q. Chen, G. Yu, and Z. Ding, "Enhanced LAA for unlicensed LTE deployment based on TXOP contention," *IEEE Transactions on Communications*, vol. 67, no. 1, pp. 417–429, Jan. 2019.

119. A. Galanopoulos, F. Foukalas, and T. A. Tsiftsis, "Efficient coexistence of LTE with WiFi in the licensed and unlicensed spectrum aggregation," *IEEE Transactions on Cognitive Communications and Networking*, vol. 2, no. 2, pp. 129–140, Jun. 2016.

120. M. Maule, D. Moltchanov, P. Kustarev, M. Komarov, S. Andreev, and Y. Koucheryavy, "Delivering fairness and QoS guarantees for LTE/Wi-Fi coexistence under LAA operation," *IEEE Access*, vol. 6, pp. 7359–7373, Jan. 2018.

121. V. Maglogiannis, D. Naudts, A. Shahid, and I. Moerman, "A Q-learning scheme for fair coexistence between LTE and Wi-Fi in unlicensed spectrum," *IEEE Access*, vol. 6, pp. 27 278–27 293, Jun. 2018.

122. J. Tan, L. Zhang, Y.-C. Liang, and D. Niyato, "Deep reinforcement learning for the coexistence of LAA-LTE and WiFi systems," in *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, May 2019, pp. 1–6.

123. F. Hao, C. Yongyu, H. Li, J. Zhang, and W. Quan, "Contention window size adaptation algorithm for LAA-LTE in unlicensed band," in *2016 International Symposium on Wireless Communication Systems (ISWCS)*, Poznan, Poland, Sep. 2016, pp. 476–480.

124. E. Pei, J. Jiang, L. Liu, Y. Li, and Z. Zhang, "A chaotic Q-learning-based licensed assisted access scheme over the unlicensed spectrum," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 9951–9962, Oct. 2019.

125. Y. Li, T. Zhou, Y. Yang, H. Hu, and M. Hamalainen, "Fair downlink traffic management for hybrid LAA-LTE/Wi-Fi networks," *IEEE Access*, vol. 5, pp. 7031–7041, Jun. 2017.

126. Z. Ali, L. Giupponi, J. Mangues-Bafalluy, and B. Bojovic, "Machine learning based scheme for contention window size adaptation in lte-laa," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Montreal, QC, Canada, Oct. 2017, pp. 1–7.

127. A. Garcia-Saavedra, P. Patras, V. Valls, X. Costa-Perez, and D. J. Leith, "ORLA/OLAA: Orthogonal coexistence of LAA and WiFi in unlicensed spectrum," *IEEE Transactions on Networking*, vol. 26, no. 6, pp. 2665–2678, Dec. 2018.

128. M. Mehrnoush, S. Roy, V. Sathya, and M. Ghosh, "On the fairness of Wi-Fi and LTE-LAA coexistence," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 4, pp. 735–748, Dec. 2018.

129. A. K. Bairagi, S. F. Abedin, N. H. Tran, D. Niyato, and C. S. Hong, "QoE-Enabled unlicensed spectrum sharing in 5G: A Game-Theoretic approach," *IEEE Access*, vol. 6, pp. 50 538–50 554, Sep. 2018.

130. M. Alsenwi, I. Yaqoob, S. R. Pandey, Y. K. Tun, A. K. Bairagi, L. Kim, and C. S. Hong, "Towards coexistence of cellular and WiFi networks in unlicensed spectrum: A neural networks based approach," *IEEE Access*, vol. 7, pp. 110 023–110 034, Aug. 2019.

131. Y. Li, F. Baccelli, J. G. Andrews, T. D. Novlan, and J. C. Zhang, "Modeling and analyzing the coexistence of Wi-Fi and LTE in unlicensed spectrum," *IEEE Transactions on Wireless Communications*, vol. 15, no. 9, pp. 6310–6326, Jun. 2016.

132. M. Alhulayil and M. López-Benítez, "Novel LAA waiting and transmission time configuration methods for improved LTE-LAA/Wi-Fi coexistence over unlicensed bands," *IEEE Access*, vol. 8, pp. 162 373–162 393, Sep. 2020.

133. S. Lagen, N. Patriciello, and L. Giupponi, "Cellular and Wi-Fi in unlicensed spectrum: Competition leading to convergence," in *2020 2nd 6G Wireless Summit (6G SUMMIT)*, Levi, Finland, Mar. 2020, pp. 1–5.

134. A. K. Bairagi, N. H. Tran, W. Saad, Z. Han, and C. S. Hong, "A game-theoretic approach for fair coexistence between LTE-U and Wi-Fi systems," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 1, pp. 442–455, Jan. 2019.

135. H. He, H. Shan, A. Huang, L. X. Cai, and T. Q. S. Quek, "Proportional fairness-based resource allocation for LTE-U Coexisting With Wi-Fi," *IEEE Access*, vol. 5, pp. 4720–4731, Sep. 2016.

136. V. Sathya, S. M. Kala, M. I. Rochman, M. Ghosh, and S. Roy, "Standardization advances for cellular and Wi-Fi coexistence in the unlicensed 5 and 6 GHz bands," *GetMobile: Mobile Comp. and Comm.*, vol. 24, no. 1, p. 5–15, Mar. 2020.

137. M. Mehrnoush, V. Sathya, S. Roy, and M. Ghosh, "Analytical modeling of Wi-Fi and LTE-LAA coexistence: Throughput and impact of energy detection threshold," *IEEE/ACM Transactions on Networking*, vol. 26, no. 4, pp. 1990–2003, Aug. 2018.

138. S. M. Kala, V. Sathya, E. Yamatsuta, H. Yamaguchi, and T. Higashino, "Operator data driven cell-selection in LTE-LAA coexistence networks," in *International Conference on Distributed Computing and Networking 2021*. New York, NY, USA: Association for Computing Machinery, Jan. 2021, p. 206–214.

139. *Evolved Universal Terrestrial Radio Access (E-UTRA) Physical Layer Procedures*, 3GPP Std. 36.213 V13.6.0, Jun. 2017.

140. G. Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 3, pp. 535–547, Mar. 2000.

141. S. Alvi, X. Zhou, and S. Durrani, "Optimal compression and transmission rate control for node-lifetime maximization," *IEEE Transactions on Wireless Communications*, vol. 17, no. 11, pp. 7774–7788, Nov. 2018.

142. P. D. Diamantoulakis and G. K. Karagiannidis, "Maximizing proportional fairness in wireless powered communications," *IEEE Wireless Communications Letters*, vol. 6, no. 2, pp. 202–205, Apr. 2017.

143. S. C. Liew and Y. J. Zhang, "Proportional fairness in multi-channel multi-rate wireless networks-part I: The case of deterministic channels with application to AP association problem in large-scale WLAN," *IEEE Transactions on Wireless Communications*, vol. 7, no. 9, pp. 3446–3456, Sep. 2008.

144. H. Shi, R. V. Prasad, E. Onur, and I. G. M. M. Niemegeers, "Fairness in wireless networks: Issues, measures and challenges," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 1, pp. 5–24, May 2013.

145. A. Mahmood, L. Beltramelli, S. Fakhrul Abedin, S. Zeb, N. I. Mowla, S. A. Hassan, E. Sisinni, and M. Gidlund, "Industrial IoT in 5G-and-beyond networks: Vision, architecture, and design trends," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 6, pp. 4122–4137, Jun. 2022.

146. H. He, H. Shan, A. Huang, Q. Ye, and W. Zhuang, "Edge-aided computing and transmission scheduling for LTE-U-Enabled IoT," *IEEE Transactions on Wireless Communications*, vol. 19, no. 12, pp. 7881–7896, Aug. 2020.

147. N. Patriciello, S. Lagén, B. Bojović, and L. Giupponi, "NR-U and IEEE 802.11 technologies coexistence in unlicensed mmWave spectrum: Models and evaluation," *IEEE Access*, vol. 8, pp. 71 254–71 271, Apr. 2020.

148. Q. Cui, Y. Gu, W. Ni, and R. P. Liu, "Effective capacity of licensed-assisted access in unlicensed spectrum for 5G: From theory to application," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 8, pp. 1754–1767, May 2017.

149. H. Song, J. Bai, Y. Yi, J. Wu, and L. Liu, "Artificial intelligence enabled internet of things: Network architecture and spectrum access," *IEEE Computational Intelligence Magazine*, vol. 15, no. 1, pp. 44–51, Jan. 2020.

150. Y. Chen, Y. A. Sambo, O. Onireti, and M. A. Imran, "A survey on LPWAN-5G integration: Main challenges and potential solutions," *IEEE Access*, vol. 10, pp. 32 132–32 149, Mar. 2022.

151. M. Stoyanova, Y. Nikoloudakis, S. Panagiotakis, E. Pallis, and E. K. Markakis, "A survey on the internet of things (IoT) forensics: Challenges, approaches, and open issues," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 1191–1221, Jan. 2020.

152. H. Zhang, B. Di, K. Bian, and L. Song, "IoT-U: cellular internet-of-things networks over unlicensed spectrum," *IEEE Transactions on Wireless Communications*, vol. 18, no. 5, pp. 2477–2492, May 2019.

153. R. F. Sari, R. Harwahyu, and R.-G. Cheng, "Load estimation and connection request barring for random access in massive C-IoT," *IEEE Internet of Things Journal*, vol. 7, no. 7, pp. 6539–6549, Jan. 2020.

154. S. Grimaldi, A. Mahmood, S. A. Hassan, M. Gidlund, and G. P. Hancke, "Autonomous interference mapping for industrial internet of things networks over unlicensed bands: Identifying cross-technology interference," *IEEE Industrial Electronics Magazine*, vol. 15, no. 1, pp. 67–78, Mar. 2021.

155. R. Maldonado, A. Karstensen, G. Pocovi, A. A. Esswie, C. Rosa, O. Alanen, M. Kasslin, and T. Kolding, "Comparing Wi-Fi 6 and 5G downlink performance for industrial IoT," *IEEE Access*, vol. 9, pp. 86 928–86 937, Jun. 2021.

156. S. Park, H. Ryu, Y. Kim, and J.-K. Han, "Performance of channel access mechanisms for 5G Industrial-IoT over unlicensed bands," in *2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)*, Sep. 2021, pp. 1–5.

157. L. Wang, Z. Fei, M. Zeng, B. Li, Y. Huo, X. Dong, and Q. Cui, "Spatial-reuse-based efficient coexistence for cellular and WiFi systems in the unlicensed band," *IEEE Internet of Things Journal*, vol. 9, no. 3, pp. 1885–1898, Jun. 2022.

158. Q. Zhang, Y. Xue, Z. Han, and Y. Lin, "Design and performance analysis of 3D Markov chain model based fair spectrum sharing access for IoT services," *IEEE Internet of Things Journal*, pp. 1–1, Feb. 2022.

159. J.-B. Seo, W. T. Toor, and H. Jin, "Analysis of two-step random access procedure for cellular ultra-reliable low latency communications," *IEEE Access*, vol. 9, pp. 5972–5985, Jan. 2021.

160. Q. Chen, X. Xu, Z. You, H. Jiang, J. Zhang, and F.-Y. Wang, "Communication-efficient federated edge learning for NR-U based IIoT networks," *IEEE Internet of Things Journal*, pp. 1–1, 2021.

161. H. Gu, Y. Wang, S. Hong, and G. Gui, "Deep learning aided friendly coexistence of WiFi and LTE in unlicensed bands," in *2019 11th International Conference on Wireless Communications and Signal Processing (WCSP)*, Xian, China, Oct. 2019, pp. 1–5.

162. H. He, H. Shan, A. Huang, Q. Ye, and W. Zhuang, "Reinforcement learning-based computing and transmission scheduling for LTE-U-enabled IoT," in *2018 IEEE Global Communications Conference (GLOBECOM)*, Abu Dhabi, United Arab Emirates, 2018, pp. 1–6.

163. S. Kim, "A new bargaining game-based unlicensed spectrum sharing scheme for TVWS platform," *IEEE Access*, vol. 9, pp. 95 528–95 537, 2021.

164. G. Callebaut, S. Gunnarsson, A. P. Guevara, A. J. Johansson, L. Van Der Perre, and F. Tufvesson, "Experimental exploration of unlicensed Sub-GHz massive MIMO for massive internet-of-things," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 2195–2204, Sep. 2021.

165. J. Chen, S. Li, J. Xing, J. Wang, and S. Fu, "Multiple nodes access of wireless beam modulation for 6G-enabled internet of things," *IEEE Internet of Things Journal*, vol. 8, no. 20, pp. 15 191–15 204, Dec. 2021.

166. K. Guan, H. Yi, D. He, B. Ai, and Z. Zhong, "Towards 6G: paradigm of realistic terahertz channel modeling," *China Communications*, vol. 18, no. 5, pp. 1–18, May 2021.

167. M. T. Mamaghani and Y. Hong, "Aerial intelligent reflecting surface enabled terahertz covert communications in beyond-5G internet of things," *IEEE Internet of Things Journal*, pp. 1–1, 2022.

168. C. Pan, H. Ren, K. Wang, J. F. Kolb, M. Elkashlan, M. Chen, M. Di Renzo, Y. Hao, J. Wang, A. L. Swindlehurst, X. You, and L. Hanzo, "Reconfigurable intelligent surfaces for 6G systems: Principles, applications, and research directions," *IEEE Communications Magazine*, vol. 59, no. 6, pp. 14–20, Jul. 2021.

169. M. Di Renzo, M. Debbah, D.-T. Phan-Huy, A. Zappone, M.-S. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J. Hoydis, H. Gacanin *et al.*, "Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come,"

EURASIP Journal on Wireless Communications and Networking, vol. 2019, no. 1, pp. 1–20, May 2019.

170. Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface aided wireless communications: A tutorial," *IEEE Transactions on Communications*, vol. 69, no. 5, pp. 3313–3351, Jan. 2021.

171. W. Tang, M. Z. Chen, X. Chen, J. Y. Dai, Y. Han, M. Di Renzo, Y. Zeng, S. Jin, Q. Cheng, and T. J. Cui, "Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, p. 421–439, Jan. 2021.

172. Q. Liu, S. Sun, B. Rong, and M. Kadoch, "Intelligent reflective surface based 6G communications for sustainable energy infrastructure," *IEEE Wireless Communications*, vol. 28, no. 6, pp. 49–55, Dec. 2021.

173. L. Jiao, P. Wang, A. Alipour-Fanid, H. Zeng, and K. Zeng, "Enabling efficient blockage-aware handover in RIS-Assisted mmWave cellular networks," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2021.

174. W. Hao, G. Sun, M. Zeng, Z. Chu, Z. Zhu, O. A. Dobre, and P. Xiao, "Robust design for intelligent reflecting surface-assisted MIMO-OFDMA terahertz IoT networks," *IEEE Internet of Things Journal*, vol. 8, no. 16, pp. 13 052–13 064, Mar. 2021.

175. S. Ma, Y. Zhang, H. Li, J. Sun, J. Shi, H. Zhang, C. Shen, and S. Li, "Covert beamforming design for intelligent-reflecting-surface-assisted IoT networks," *IEEE Internet of Things Journal*, vol. 9, no. 7, pp. 5489–5501, Sep. 2022.

176. S. Xu, J. Liu, and Y. Cao, "Intelligent reflecting surface empowered physical-layer security: Signal cancellation or jamming?" *IEEE Internet of Things Journal*, vol. 9, no. 2, pp. 1265–1275, May 2022.

177. C. Yuan, X. Tao, W. Ni, N. Li, A. Jamalipour, and R. P. Liu, "Optimal power allocation for superposed secrecy transmission in multicarrier systems," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 2, pp. 1332–1346, Feb. 2021.

178. J. Chen, Y.-C. Liang, Y. Pei, and H. Guo, "Intelligent reflecting surface: A programmable wireless environment for physical layer security," *IEEE Access*, vol. 7, pp. 82 599–82 612, Jun. 2019.

179. H. Shen, W. Xu, S. Gong, Z. He, and C. Zhao, "Secrecy rate maximization for intelligent reflecting surface assisted multi-antenna communications," *IEEE Communications Letters*, vol. 23, no. 9, pp. 1488–1492, Jun. 2019.

180. M. Cui, G. Zhang, and R. Zhang, "Secure wireless communication via intelligent reflecting surface," *IEEE Wireless Communications Letters*, vol. 8, no. 5, pp. 1410–1414, May 2019.

181. W. Jiang, Y. Zhang, J. Wu, W. Feng, and Y. Jin, "Intelligent reflecting surface assisted secure wireless communications with multiple-transmit and multiple-receive antennas," *IEEE Access*, vol. 8, pp. 86 659–86 673, May 2020.

182. X. Yu, D. Xu, Y. Sun, D. W. K. Ng, and R. Schober, "Robust and secure wireless communications via intelligent reflecting surfaces," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2637–2652, Jul. 2020.

183. Y. Ge and J. Fan, "Robust secure beamforming for intelligent reflecting surface assisted full-duplex MISO systems," *IEEE Transactions on Information Forensics and Security*, vol. 17, pp. 253–264, Jan. 2022.

184. W. Chen, X. Ma, Z. Li, and N. Kuang, "Sum-rate maximization for intelligent reflecting surface based Terahertz communication systems," in *IEEE/CIC International Conference on Communications Workshops in China (ICCC Workshops)*, Changchun, China, Aug. 2019, pp. 153–157.

185. Y. Pan, K. Wang, C. Pan, H. Zhu, and J. Wang, "Sum rate maximization for intelligent reflecting surface assisted Terahertz communications," *arXiv preprint arXiv:2008.12246*, 2020.

186. Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019.

187. G. Zhou, C. Pan, H. Ren, K. Wang, M. D. Renzo, and A. Nallanathan, "Robust beamforming design for intelligent reflecting surface aided MISO communication systems," *IEEE Wireless Communications Letters*, vol. 9, no. 10, pp. 1658–1662, Oct. 2020.

188. H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3064–3076, Feb. 2020.

189. S. Huang, Y. Ye, M. Xiao, H. V. Poor, and M. Skoglund, "Decentralized beamforming design for intelligent reflecting surface-enhanced cell-free networks," *IEEE Wireless Communications Letters*, vol. 10, no. 3, pp. 673–677, Mar. 2021.

190. B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface assisted multi-user OFDMA: Channel estimation and training design," *IEEE Transactions on Wireless Communications*, vol. 19, no. 12, pp. 8315–8329, Dec. 2020.

191. X. Ma, Z. Chen, Y. Chi, W. Chen, L. Du, and Z. Li, "Channel estimation for intelligent reflecting surface enabled Terahertz MIMO systems," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, Dublin, Ireland, Jun. 2020, pp. 1–6.

192. Y. Cao, T. Lv, Z. Lin, and W. Ni, "Delay-constrained joint power control, user detection and passive beamforming in intelligent reflecting surface-assisted uplink mmWave system," *IEEE Transactions on Cognitive Communications and Networking*, pp. 1–1, Mar. 2021.

193. S. Zhang, S. Zhang, F. Gao, J. Ma, and O. A. Dobre, "Deep learning based RIS channel extrapolation with element-grouping," *arXiv preprint arXiv:2105.06850*, 2021.

194. X. Ma, Z. Chen, W. Chen, Z. Li, Y. Chi, C. Han, and S. Li, "Joint channel estimation and data rate maximization for intelligent reflecting surface assisted terahertz MIMO communication systems," *IEEE Access*, vol. 8, pp. 99 565–99 581, May 2020.

195. H. Yang, Z. Xiong, J. Zhao, D. Niyato, L. Xiao, and Q. Wu, "Deep reinforcement learning-based intelligent reflecting surface for secure wireless communications," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 375–388, Jan. 2021.

196. C. Huang, Z. Yang, G. C. Alexandropoulos, K. Xiong, L. Wei, C. Yuen, Z. Zhang, and M. Debbah, "Multi-hop RIS-empowered terahertz communications: A DRL-based hybrid beamforming design," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 6, pp. 1663–1677, Apr. 2021.

197. A. Taha, Y. Zhang, F. B. Mismar, and A. Alkhateeb, "Deep reinforcement learning for intelligent reflecting surfaces: Towards standalone operation," in *IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Atlanta, GA, USA, May 2020, pp. 1–5.

198. C. Huang, R. Mo, and C. Yuen, "Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1839–1850, Aug. 2020.

199. B. Ning, Z. Chen, W. Chen, Y. Du, and J. Fang, "Terahertz multiuser massive MIMO with intelligent reflecting surface: Beam training and hybrid beamforming," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 2, pp. 1376–1393, Jan. 2021.

200. N. C. Luong, D. T. Hoang, S. Gong, D. Niyato, P. Wang, Y.-C. Liang, and D. I. Kim, "Applications of deep reinforcement learning in communications and networking: A survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3133–3174, May 2019.

201. Q. Mao, F. Hu, and Q. Hao, "Deep learning for intelligent wireless networks: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2595–2621, Jun. 2018.

202. G. Geraci, A. Garcia-Rodriguez, M. M. Azari, A. Lozano, M. Mezzavilla, S. Chatzinotas, Y. Chen, S. Rangan, and M. Di Renzo, "What will the future of UAV cellular communications be? a flight from 5G to 6G," *IEEE Communications Surveys & Tutorials*, pp. 1–1, May 2022.

203. M. Giordani and M. Zorzi, "Non-terrestrial networks in the 6Gera: Challenges and opportunities," *IEEE Network*, vol. 35, no. 2, pp. 244–251, Jun. 2021.

204. Z. Wang, T. Lv, J. Zeng, and W. Ni, "Placement and resource allocation of wireless-powered multiantenna UAV for energy-efficient multiuser NOMA," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2022.

205. K. Li, W. Ni, A. Noor, and M. Guizani, "Employing intelligent aerial data aggregators for the internet of things: Challenges and solutions," *IEEE Internet of Things Magazine*, vol. 5, no. 1, pp. 136–141, Mar. 2022.

206. A. Asheralieva and D. Niyato, "Optimizing age of information and security of the next-generation internet of everything systems," *IEEE Internet of Things Journal*, pp. 1–1, May 2022.

207. W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, vol. 34, no. 3, pp. 134–142, May 2020.

208. M. A. Uusitalo, M. Ericson, B. Richerzhagen, E. U. Soykan, P. Rugeland, G. Fettweis, D. Sabella, G. Wikström, M. Boldi, M.-H. Hamon, H. D. Schotten, V. Ziegler, E. C. Strinati, M. Latva-aho, P. Serrano, Y. Zou, G. Carrozzo, J. Martrat, G. Stea, P. Demestichas, A. Pärssinen, and T. Svensson, "Hexa-X the European 6G flagship project," in *2021 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, Porto, Portugal, Jun. 2021, pp. 580–585.

209. M. A. Raza, M. Abolhasan, J. Lipman, N. Shariati, W. Ni, and A. Jamalipour, "Statistical learning-based grant-free access for delay-sensitive internet of things applications," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 5, pp. 5492–5506, Mar. 2022.