On Enhancing Connectivity and Efficiency in 5G and Emerging 6G Mobile Networks

Muhammad Farhan Khan
MASTER OF SCIENCE IN TELECOMMUNICATION ENGINEERING
BACHELOR OF SCIENCE IN COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF IRELAND, CORK

COLLEGE OF SCIENCE, ENGINEERING AND FOOD SCIENCE

SCHOOL OF COMPUTER SCIENCE & INFORMATION TECHNOLOGY

Thesis submitted for the degree of
Master of Philosophy

2025

Head of School: Professor Dirk H J Pesch

Supervisors: Professor Dirk H J Pesch
Professor Cormac J. Sreenan

Research supported by Science Foundation Ireland

Contents

List of Figures

List of Tables

List of Abbreviations

4G	Fourth Generation
5G	Fifth Generation
6G	Sixth Generation
AI	Artificial Intelligence
AWGN	Additive White Gaussian Noise
BPSK	Binary Phase Shift Keying
CBS	Composite Base Station
CCDF	Complementary Cumulative Distribution Function
CNN	Convolutional Neural Network
CoMP	Coordinated Multipoint
CQI	Channel Quality Indicator
CR	Compression Ratio
D2D	Device-to-Device Communication
DSS	Dynamic Spectrum Scheduling
ECDF	Empirical Cumulative Distribution Function
eMBB	Enhanced Mobile Broadband
HetNets	Heterogeneous Networks
IIoT	Industrial Internet of Things
iUE	Industrial User Equipment
IRS	Intelligent Reflecting Surfaces
IoT	Internet of Things
LOS/NLOS	Line-of-Sight/Non-Line-of-Sight
MIMO	Multiple-Input Multiple-Output
ML/DL	Machine Learning/Deep Learning
mMTC	Massive Machine-Type Communication
mmWave	Millimeter-Wave
MSE	Mean Squared Error
MRP	Maximum Received Power
MRT	Maximum Ratio Transmission
NOMA	Non-Orthogonal Multiple Access
NMSE	Normalized Mean Square Error
OFDM	Orthogonal Frequency Division Multiplexing
PAPR	Peak-to-Average Power Ratio
QoS	Quality of Service
RBs	Resource Blocks
RIS	Reconfigurable Intelligent Surfaces
RNN	Recurrent Neural Network
RoI	Region of Interest
SINR	Signal-to-Interference-plus-Noise Ratio
STBC	Space-Time Block Coding
STBC-ST	Space-Time Block Coding with Superposition Transmission
URLLC	Ultra-Reliable and Low Latency Communication
V5SLS	Vienna 5G System Level Simulator
ZF	Zero Forcing

This is to certify that the work I am submitting is my own and has not been submitted for another degree, either at University College Cork or elsewhere. All external references and sources are clearly acknowledged and identified within the contents. I have read and understood the regulations of University College Cork concerning plagiarism and intellectual property.

Muhammad Farhan Khan

To my parents, whose steadfast encouragement and enduring belief in my abilities laid the foundation for my endeavors; to Professor Dirk H. J. Pesch, whose profound dedication to Wireless Communication inspired in me a lifelong commitment to the pursuit of knowledge; and to Professor Cormac J. Sreenan, whose thoughtful guidance and steady support enriched every stage of this journey. I am deeply grateful to them for their belief in me and for the pivotal roles they played in shaping this transformative experience.

List of Publications

Aspects of the work described in this dissertation have been published as follows:

- **[Pub 1] Muhammad Farhan Khan**, "An Approach for Optimal Base Station Selection in 5G HetNets for Smart Factories," 2020 IEEE 21st International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM), Cork, Ireland, 2020, pp. 64-65, doi: 10.1109/WoWMoM49955.2020.00025.

  Statement of Contributions: As the primary contributor, first author, and corresponding author, Muhammad Farhan Khan developed the base station selection framework for 5G HetNets, designed and implemented the Maximum SINR and Maximum Received Power cell association strategies, conducted extensive simulations, and analyzed throughput and load balancing performance. He also authored the manuscript. While the paper had no formal co-authors, research advisors Professor Dirk Pesch and Professor Cormac J. Sreenan contributed through technical discussions and critical review of the work.

- **[Pub 2] Muhammad Farhan Khan** and D. Pesch, "On Performance of Multi-User Massive MIMO for 5G and Beyond," 2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring), Helsinki, Finland, 2022, pp. 1-5, doi: 10.1109/VTC2022-Spring54318.2022.9860567.

  Statement of Contributions: As the primary contributor, first author, and corresponding author, Muhammad Farhan Khan along with Professor Dirk Pesch formulated the problem statement, developed the system model for multi-user massive MIMO, designed and implemented the Zero Forcing (ZF) and Maximum Ratio Transmission (MRT) precoding techniques, conducted performance evaluations, and wrote the manuscript. Other Co-authors contributed through discussions and manuscript refinement.

- **[Pub 3] Muhammad Farhan Khan**, D. Pesch, A. Iqbal, and J. M. Hamamreh, "A Novel Multi-User Space-Time Block Coding Based Superposition Transmission for Future Generation Wireless Networks," 2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring), Florence, Italy, 2023, pp. 1-7, doi: 10.1109/VTC2023-Spring57618.2023.10200935.

  Statement of Contributions: As the primary contributor, first author, and corresponding author, Muhammad Farhan Khan along with Professor Dirk Pesch proposed the novel MU-STBC-ST transmission scheme, developed the superpositioned transmission approach, formulated the mathematical model, conducted extensive simulations, and analyzed throughput, spectral efficiency, and error rate performance. They also wrote the manuscript. Co-authors provided technical insights and manuscript review.

- **[Pub 4] Muhammad Farhan Khan**, A. Rashid, A. Iqbal, and D. Pesch, "Enabling High-Speed Connectivity in Urban Environments Through Composite Base Stations and Dynamic Spectrum Scheduling," 2023 IEEE 34th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Toronto, ON, Canada, 2023, pp. 1-6, doi: 10.1109/PIMRC56721.2023.10294017.

Statement of Contributions: As the primary contributor, first author, and corresponding author, Muhammad Farhan Khan along with Professor Dirk Pesch conceptualized the Composite Base Station (CBS) and Dynamic Spectrum Scheduling (DSS) framework, developed system models integrating sub-6 GHz and mmWave bands, designed spectrum allocation algorithms, implemented performance evaluations, and wrote the manuscript. Other Co-authors assisted in discussions and refinement.

- **[Pub 5] Muhammad Farhan Khan**, A. Iqbal, A. Rashid, and D. Pesch, "Enhancing Industrial 4.0 Connectivity: A D2D-Based Algorithm for Blind Spot Mitigation in 5G Future Networks Enabled Smart Industry," 2023 IEEE Globecom Workshops (GC Wkshps), Kuala Lumpur, Malaysia, 2023, pp. 2012-2017, doi: 10.1109/GCWkshps58843.2023.10464805.

  Statement of Contributions: As the primary contributor, first author, and corresponding author, Muhammad Farhan Khan along with Professor Dirk Pesch developed the novel D2D-based blind spot mitigation algorithm, designed the system framework for Industrial IoT networks, conducted extensive simulations, analyzed performance metrics, and wrote the manuscript. Other Co-authors contributed through discussions and technical feedback.

- **[Pub 6] Muhammad Farhan Khan**, A. Iqbal, A. Rashid, M. A. Jamshed, and D. Pesch, "Performance Analysis of Intelligent Reflecting Surfaces for 5G/6G-Enabled Future Smart Industries with a Focus on Millimeter-Wave Band Communications," 2024 IEEE International Conference on Communications Workshops (ICC Workshops), Denver, CO, USA, 2024, pp. 2040-2045, doi: 10.1109/ICCWorkshops59551.2024.10615318.

  Statement of Contributions: As the primary contributor, first author, and corresponding author, Muhammad Farhan Khan along with Professor Dirk Pesch formulated the Intelligent Reflecting Surfaces (IRS) system model for mmWave communications, developed IRS-assisted beamforming techniques, conducted performance evaluations, analyzed energy and spectral efficiency, and authored the manuscript. Other Co-authors provided review and refinement.

- **[Pub 7] Muhammad Farhan Khan**, M. Ahmed Mohsin, Aimen Naeem, Adeel Iqbal, Ali Jamshed "Machine Learning-Driven Performance Analysis of Compressed Communication in Aerial-RIS Networks for Future 6G Networks," Submitted to Annals of Telecommunications.

  Statement of Contributions: As the primary contributor, first author, and corresponding author, Muhammad Farhan Khan conceptualized and developed the ML-driven framework for compressed communication in

Aerial-RIS networks, implemented AI-based resource allocation algorithms, designed and conducted simulations, performed outage probability and energy efficiency analysis, and authored the manuscript. Co-authors contributed through discussions and manuscript review.

Other Publications

The following publications and submitted manuscripts are derived from work carried out during the study, but are not presented in this dissertation.

- **[Pub 8]** Muhammad Ahmed Mohsin, Hassan Rizwan, Muhammad Jazib, Muhammad Iqbal, Muhammad Bilal, Tabinda Ashraf, Muhammad Farhan Khan, Jen-Yi Pan “Deep Reinforcement Learning Optimized Intelligent Resource Allocation in Active RIS-Integrated TN-NTN Networks”, IEEE Wireless Communications and Networking Conference, 24–27 March 2025, Mico Milano Congressi, Milan, Italy.
- **[Pub 9]** Muhammad Ahmed Mohsin, Muhammad Farhan Khan, Zeeshan Alam, Ali Jamshed, "Vision Transformer Based Semantic Communications for Next Generation Wireless Networks" IEEE International Conference on Communications, 8–12 June 2025, Montreal, Canada.
- **[Pub 10]** Ali Raza, Muhammad Farhan Khan, Muhammad Ali Jamshed, "TeraRIS NOMA-MIMO Communication System for 6G and Beyond" 22nd International Conference on Frontiers of Information Technology (FIT), December 15 to 16, 2025, Islamabad, Pakistan.
- **[Pub 11]** Adnan Rashid, Adeel Iqbal, Muhammad Furqan Khan, Muhammad Farhan Khan, Giancarlo Sciddurlo, Arcangela Rago “A Comprehensive Layered Framework for Telesurgery: Technologies, Challenges, and Future Directions, Submitted to Computer Networks.

Acknowledgements

I extend my deepest gratitude to Prof. Dirk H. J. Pesch, whose expert guidance and consistent support have been pivotal throughout my research. He has provided both encouragement and critical insight at every stage. As a mentor, his broad knowledge, thoughtful feedback, and patience played a key role in shaping and maturing this work.

I am equally thankful to my co-supervisor, Prof. Cormac J. Sreenan, for his valuable mentorship and steady guidance during my time at University College Cork. I also extend heartfelt thanks to Professor Ken Brown, whose assistance was instrumental in securing this opportunity, and to my colleagues and friends whose camaraderie enriched this journey.

I wish to express my profound appreciation to Prof. Dr. Aamir Saeed Malik, currently at Brno University of Technology, Czech Republic, who introduced me to the world of research during my time at Universiti Teknologi PETRONAS Malaysia. I am deeply grateful to Dr. Muhammad Imran Khan and Dr. Amna Imtiaz for their generous support and insightful guidance, which helped shape the direction of this work. I also thank my M.Sc. supervisors, Dr. Farrukh Aziz Bhatti, Dr. Aamir Habib and Dr Sobia Jangsher, for introducing me to the field of 5G. A special mention goes to Ehsan Yar Khan, Chaudhry Mudassar Naveed, Sikandar Abbas, Ikram Ul Haq Kayani, Faheem Iqbal, Viqar Rashid Khan, Miraj Gul, the late Muhammad Umer Farooq, whose contributions and support have been invaluable.

I would like to sincerely thank Dr. Adnan Rashid, Dr. Imran Khan, Dr. Adeel Iqbal, Mr. Ali Raza, and Ms. Aiman Naeem for their collegial support during the preparation of this thesis. Their involvement was limited to providing informal assistance with formatting, feedback, proofreading suggestions. These contributions, while appreciated, were solely supportive in nature and did not relate to the technical content, analysis, or intellectual framing of the research.

I would also like to thank the friends and colleagues I met during my time at UCC, especially Saeid Rezaei, Dr. Mohamed Seliem, Dr. Zohaib Ijaz, Hamzeh Asadi, Amirhossein Ebrahimi Moghaddam, Dr. Diana Gabriela Morillo Fueltala, Dr. Leila Sedghi and Dr. Fadhil Firyaguna, whose companionship made my time in Cork memorable. I also thank my internal and external examiners for taking the time to evaluate this dissertation.

My sincere thanks go to the School of Computer Science and Information Technology and its staff. Special thanks to the research support team — Caitríona Walsh, James Duggan, Mary Noonan, and Linda O'Sullivan — and to the technical and administrative staff, including Peter McHale, James Delaney, David O'Byrne, Youngsin Ó Sé, Julie Walsh, and Joanne Murphy, for their dedicated assistance throughout.

I also wish to acknowledge the Institute of Telecommunications at TU Wien for providing the Vienna Cellular Communications Simulators (VCCS) software suite under an academic license. I am grateful to all my co-authors and research collaborators during this journey.

I am profoundly grateful to my family — my father, mother, sister, and brother — for their unconditional love, encouragement, and wise counsel. Their belief in me has been my greatest source of strength.

I would also like to acknowledge the profound influence of George Boole FRS (1815–1864), the first Professor of Mathematics at Queen's College, Cork. Often regarded as the father of the Information Age, his pioneering work in Boolean algebra laid the foundation for modern digital systems — a legacy that inspired me every time I passed his bust at the university library during this journey.

Finally, I gratefully acknowledge the funding support from Science Foundation Ireland now Research Ireland under Grant No. 16/RC/3918, without which this research would not have been possible.

Muhammad Farhan Khan
2025

Abstract

The evolution of 5G and emerging 6G mobile networks introduces stringent requirements on connectivity, efficiency, reliability, and adaptability, particularly for smart factory and dense urban environments. This thesis investigates key challenges in next-generation wireless systems, including base station selection, interference management, spectral efficiency enhancement, latency reduction, and feedback overhead, in support of Industry 4.0, massive machine-type communication (mMTC), and ultra-reliable low-latency communication (URLLC).

The thesis first examines base station selection strategies in 5G heterogeneous networks, comparing Maximum Received Power (MRP) and Maximum SINR-based association. Results show clear trade-offs between interference awareness, load distribution, and user throughput, highlighting the importance of interference-aware association for fairness and capacity in dense deployments. The performance of massive MIMO precoding is then analyzed, demonstrating that Zero Forcing (ZF) significantly outperforms Maximum Ratio Transmission (MRT) in interference-limited scenarios, yielding notable gains in spectral efficiency and downlink and uplink sum throughput with increasing antenna counts.

To further improve reliability and efficiency, a multi-user STBC-OFDM transmission scheme is investigated, achieving reduced error rates, a twofold throughput gain, and lower peak-to-average power ratio compared to conventional OFDM. The thesis also proposes a Composite Base Station architecture with Dynamic Spectrum Scheduling, enabling coordinated use of sub-6 GHz and mmWave bands, and demonstrating substantial latency and throughput improvements in dense urban scenarios.

In addition, a device-to-device relay selection framework is developed for smart industrial IoT environments, significantly improving system efficiency and mitigating coverage blind spots. The role of Intelligent Reflecting Surfaces (IRS) is then explored for mmWave and sub-6 GHz communications, showing that IRS-assisted links enhance spectral efficiency while substantially reducing required transmit power.

Finally, the thesis investigates machine learning-driven optimization for compressed communication in aerial RIS-assisted CoMP-NOMA networks. Autoencoder-based feedback compression is employed to reduce RIS phase feedback overhead under fading and noise, while preserving reconstruction accuracy. Simulation results confirm that aerial RIS deployment improves network sum rate, spectral efficiency, energy efficiency, and outage performance, particularly for cell-edge users.

Overall, this thesis presents a unified and experimentally grounded framework for enhancing performance, scalability, and energy efficiency in 5G and future 6G wireless networks, providing practical insights for smart industrial and urban communication systems.

Chapter 1

Introduction and Background

> *"The day will come when the man at the telephone will be able to see the distant person to whom he is speaking."*
> *Alexander Graham Bell in 1906*

The world has witnessed four generations of mobile communications in the last 5 decades. 1G which was analog transmission came around 1980 and was limited to voice services, 2G emerged in early 1990s, and was based on digital transmission. 2G was mainly Voice with limited data services. 3G was introduced in early 2000 and allowed high-quality mobile broadband. 4G was introduced in 2010 providing higher efficiency and enhanced mobile broadband with high data rates. The discussion on 5G began in 2012, and it was used to refer to the new 5G radio access technology. Initially, much of the discussion around 5G centered on the physical layer and radio access network (RAN) components, as these are fundamental to achieving the promised speeds and capacity. While the focus began on radio access technology, the vision for 5G expanded to encompass much more than just faster speeds. It includes concepts like massive IoT, ultra-reliable low-latency communications (URLLC), and network slicing. The term "5G" has evolved over time to represent a broader ecosystem of technologies and applications, not just radio access.The 3rd Generation Partnership Project (3GPP) after standardizing 3G continued into the development of 4G and 5G. It is the only significant organization developing Technical specification for mobile communication [DPS18]. The emergence of 5G wireless networks represents a pivotal transformation in the realm of Cyber-Physical Manufacturing Systems, particularly within the smart factory landscape. The 5G standards, which encapsulate Ultra-Reliable and Low La-

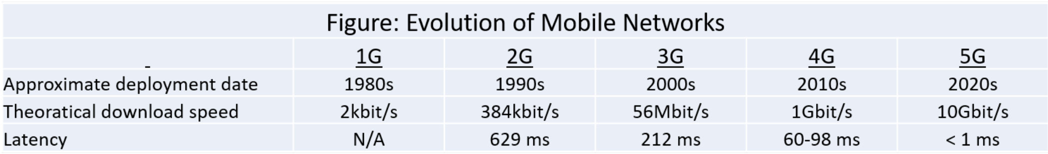

Figure: Evolution of Mobile Networks					
-	1G	2G	3G	4G	5G
Approximate deployment date	1980s	1990s	2000s	2010s	2020s
Theoratical download speed	2kbit/s	384kbit/s	56Mbit/s	1Gbit/s	10Gbit/s
Latency	N/A	629 ms	212 ms	60-98 ms	< 1 ms

Figure 1.1: Evolution of Wireless Technologies : Image Courtesy ITU [ITU]

tency Communications (URLLC), Enhanced Mobile Broadband (eMBB), Massive Machine-Type Communications (mMTC), are foundational in propelling Industry 4.0 use cases into a new era. While some may argue that 5G is merely an evolution of 4G with higher data rates, this overlooks the revolutionary advancements it brings. 5G standards, including URLLC and mMTC, are not just incremental improvements. These innovations enable high-speed, reliable communication that supports the diverse and critical needs of modern industrial environments, transforming the landscape of connectivity. This diversity spans a spectrum of communication requirements, from short-range wireless devices handling localized tasks to long-range devices orchestrating broader factory operations.

1.1 Evolution of Wireless Technologies

The evolution of wireless technologies, as shown in Figure 1.1, highlights the transition from analog to digital and now to high-speed broadband communication. From its initial 1G phase to the advanced 5G era, has been marked by significant strides in network efficiency. This evolution encompasses not only the increased spectral and energy efficiency but also the critical reduction in latency. Furthermore, the support for mobile broadband and Internet-of-Things (IoT) devices has been substantially enhanced. The impending proliferation of IoT devices, projected to reach an astonishing 25 billion by 2025, underscores the burgeoning demand for robust and seamless connectivity[JNJJ20]. This demand is not just quantitative in the sheer number of connections but also qualitative in terms of the reliability and speed of these connections[AHG$^+$20]. The evolution from 1G to 5G can be visualized as a journey from basic voice communication to high speed communication with an interconnected web of smart devices, each requiring and contributing to a more sophisticated network.

1.2 5G: The Next Generation

5G-NR is not just an upgrade; it's a paradigm shift. It is not just a radio-access technology but visualized as new services envisioned to be enabled by future mobile communication. As per [DPS18] there are many 5G use cases, but the three significant classes are: Ultra-Reliable and Low Latency Communications, Enhanced Mobile Broadband, Massive Machine-Type Communications.

Massive Machine-Type Communications (mMTC) refers to services involving a vast number of connected devices, such as remote sensors, actuators, and equipment monitoring systems. These services prioritize extremely low device cost and minimal energy consumption, allowing devices to operate on battery power for several years. Each device typically generates and consumes only small amounts of data, making high data rate support less essential. In this context, 5G supports applications that require cost-effective, energy-efficient solutions for managing a large number of devices that produce small data volumes, such as smart agriculture, logistics, fleet management, smart meters, and tracking systems.

Enhanced Mobile Broadband (eMBB) constitutes an evolutionary advancement in mobile broadband technologies, characterized by its capability to handle significantly higher data volumes and deliver markedly enhanced end-user data rates. This advancement facilitates a superior user experience across a broad spectrum of applications. eMBB is designed to support high-speed, high-capacity data services in various domains, including enterprise networks, residential broadband, mobile and wireless/fixed communications, large-scale venues, virtual and augmented reality (VR/AR), broadcasting, and high-definition content streaming, such as 4K/8K Ultra High Definition (UHD).

Ultra-Reliable Low-Latency Communications (URLLC) are expected to require minimal latency and exceptionally high reliability, essential for applications such as traffic safety systems, autonomous control mechanisms, and industrial automation processes. Critical Machine-Type Communications (MTC), characterized by the need for ultra-reliable, low-latency, and highly available connections, encompasses applications including traffic safety and control, industrial automation, remote manufacturing, remote training, and remote surgery.

5G is crucial for applications central to smart manufacturing, the Internet of Things (IoT), and autonomous robotics. These applications are driven by various use cases of 5G technology, including Enhanced Mobile Broadband

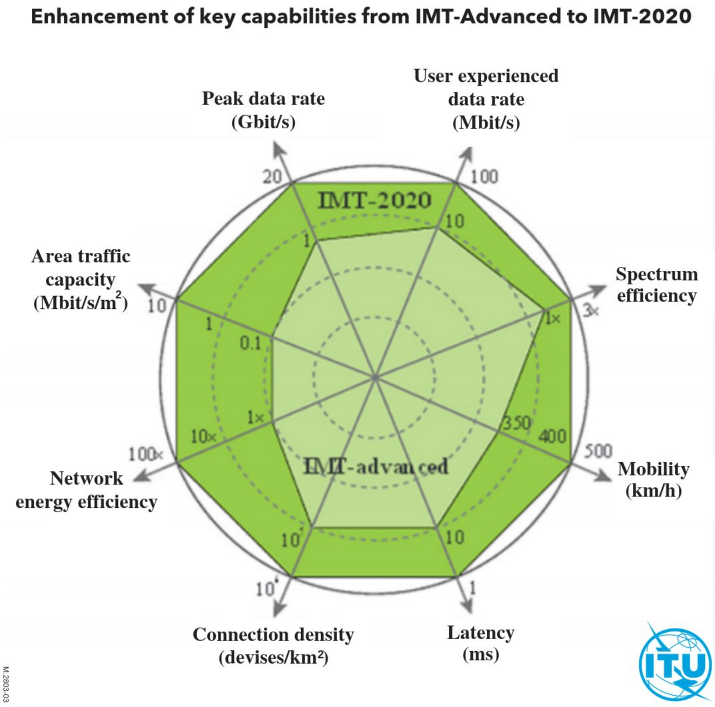

Figure 1.2: IMT-Advanced to IMT-2020 : Image courtesy ITU

(eMBB), Ultra-Reliable Low-Latency Communications (URLLC), and Massive Machine-Type Communications (mMTC). Each use case addresses distinct and demanding requirements of modern industrial systems, ranging from high data throughput to immediate and reliable communication in critical operations [Stu17].

5G represents the current iteration of mobile standards as defined by the ITU. Referred to as IMT-2020 (5G), it encompasses systems, components, and associated elements designed to provide advanced capabilities surpassing those of IMT-2000 (3G) and IMT-Advanced (4G) systems. As illustrated in Figure 1.2, the transition from IMT-Advanced (4G) to IMT-2020 (5G) demonstrates significant advancements in peak data rates and coverage capabilities. The anticipated peak data rate for enhanced mobile broadband in IMT-2020 is projected to achieve 10 Gbit/s. However, specific conditions and scenarios may enable IMT-2020 to support a peak data rate of up to 20 Gbit/s. IMT-2020 is designed to accommodate diverse user-experienced data rates, catering to various environments for enhanced mobile broadband. In instances requiring wide area coverage, such as urban and suburban areas, an anticipated user-experienced data rate of 100 Mbit/s is expected. Conversely, in hotspot scenarios, the user-experienced data rate is anticipated to reach higher values, potentially reaching 1 Gbit/s indoors.

Figure 1.3 illustrates the progression of mobile subscriptions by technology from 2021 to 2031, highlighting the rapid adoption of 5G technology and the forecasted decline of older technologies [Eri].

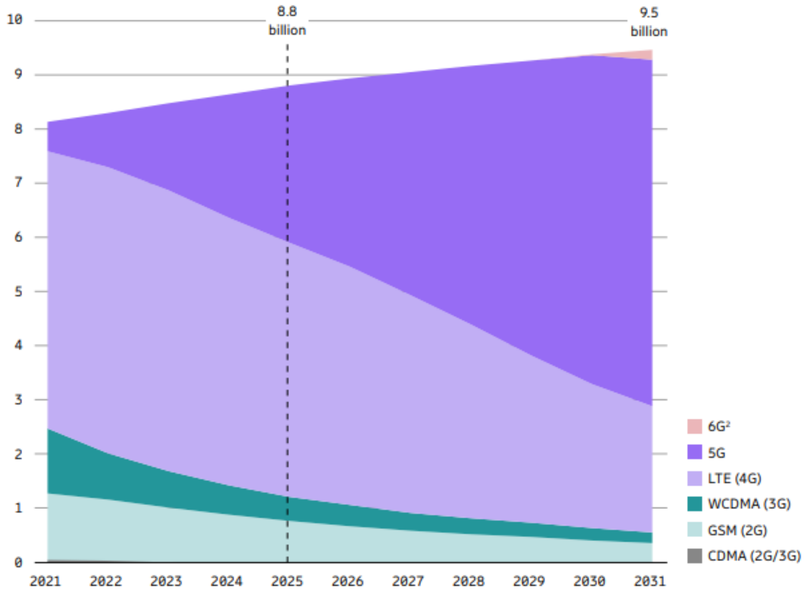

Figure 1.3: Mobile subscriptions by technology (billion) in November 2025 : Image courtesy Ericsson [Eri].

1.3 The Role of 5G in Smart Industries and Urban Environments-Enhancing Connectivity

Within the domains of smart manufacturing and Industrial IoT (IIOT) the advent of 5G marks a transformative era. It heralds a significant enhancement in network capacity, offering high data rates coupled with low latency, and marks a substantial improvement in energy efficiency over preceding technologies. This enhancement is not merely incremental but is a cornerstone in the ongoing journey toward advanced industrial automation. In this context, 5G serves as a key enabler that unlocks new possibilities - from real-time data processing and analytics to the seamless integration of various manufacturing processes. These capabilities are integral to the development of smarter, more efficient, and highly responsive manufacturing environments / smart factories[Kha20].

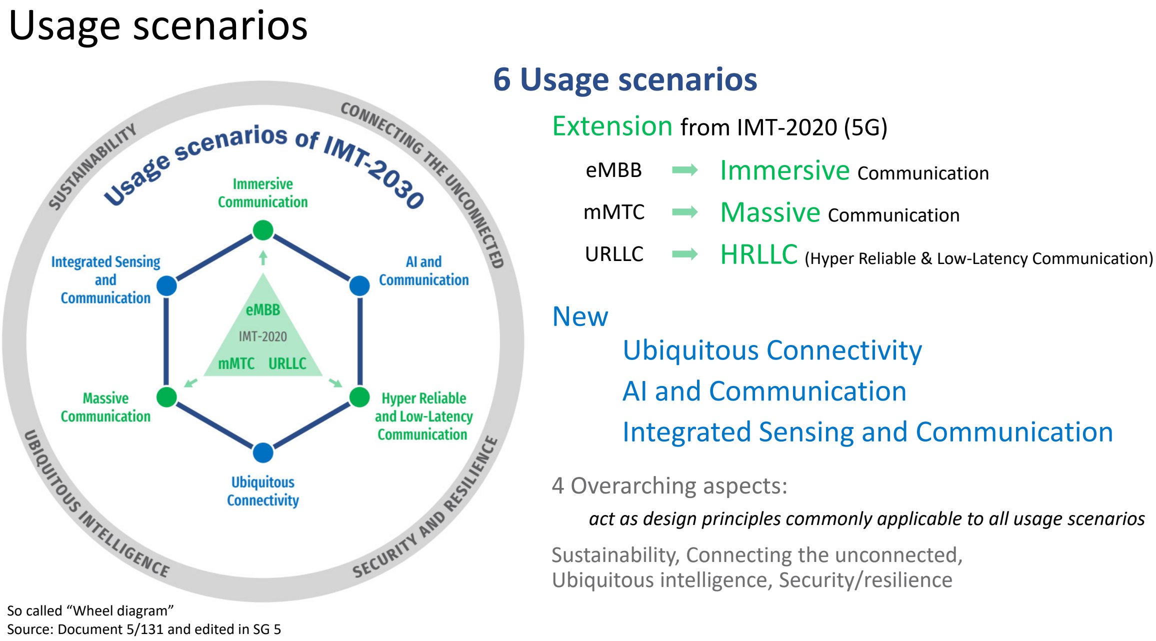

Figure 1.4: IMT-2030, 6G Vision : Image courtesy ITU

1.3.1 Motivation

The exponential growth of connected devices and bandwidth-intensive applications has placed unprecedented demands on wireless infrastructure. Fifth-generation (5G) networks, designed around enhanced Mobile Broadband (eMBB), ultra-Reliable Low-Latency Communications (uRLLC), and massive Machine-Type Communications (mMTC), have achieved remarkable progress but still face inherent challenges in scalability, spectral efficiency, and energy optimization. The advent of 5G technology marks a revolutionary shift in the landscape of factories of the future, playing a pivotal role in driving the IIoT. This section, dig into how 5G acts as a key element for the transformation and advancement of industrial processes and Urban environments. The integration of 5G Heterogeneous Networks (HetNets) in smart factories is not just an enhancement but a fundamental enabler of this evolution. Figure1.4 illustrates the overarching vision of IMT-2030 (6G), highlighting key advancements such as Hyper Reliable Low Latency Communication (HRLLC), AI-enabled networks, and integrated sensing and communication. The optimization of base station selection within these networks is critical for maintaining robust connectivity across a diverse array of IoT devices. By examining cell association strategies, particularly the efficacy of the Maximum Received Power method, the importance of 5G in ensuring enhanced network efficiency and robustness in smart manufacturing environments is highlighted [CSRKP+21].

Furthermore, the role of 5G is exemplified through the implementation of multi-user massive (MIMO) technology. This technology heralds significant improvements in connectivity, speed, and reliability, which are vital for the complexities of industrial automation and smart manufacturing. The application of advanced techniques like Zero Forcing precoding in massive MIMO systems, as evidenced by simulation results, illustrates the critical role of 5G in optimizing technological selections to maximize the potential of IIOT. These advancements enable real-time data analysis, predictive maintenance, and seamless machine-to-machine communications, thereby fostering a more efficient, productive, and innovative industrial landscape [SONA22].

The exploration of 5G's role is particularly pertinent in the context of the advancements in Space-Time Block Coding (STBC) and Orthogonal Frequency Division Multiplexing (OFDM), as discussed in the works [CRH20], and [SCW+11]. These technologies are central to the development of multi-user transmission schemes in 5G networks, enhancing the efficiency and reliability. This highlights not just the technological advancements inherent in 5G, but also its potential to catalyze significant improvements in security and efficiency.

The integration of advanced technologies such as CBS and DSS within 5G and LTE networks is crucial for enhancing connectivity and data transmission speeds essential for Urban Environments. The comparison considers sub-6GHz and mmWave frequencies (28GHz) and a Dynamic Spectrum Scheduling (DSS) to improve network performance and increase overall system capacity. The deployment of these technologies, particularly in urban and settings, optimizes spectrum usage and improves network capacity, ensuring robust and reliable communications. The innovative use of D2D communication and IoT technologies, including the development of algorithms that utilize IoT nodes as relays to overcome coverage gaps and improve overall system efficiency, thus driving forward the capabilities and potential of Industrial IoT in the era of 5G [KRIP23].

Lastly, 5G deployment with the Intelligent Reflecting Surfaces (IRS) in Future Smart Industries with a Focus on Millimeter-Wave Band Communications. These technologies are central for enhancing the efficiency and reliability of Industrial applications. The integration of IRS in 5G aids in enhancing coverage and network reliability for , addressing the critical challenges in industrial environments such as energy efficiency and robust communication. The progression towards 6G networks introduces novel paradigms such as Aerial Reconfigurable Intelligent Surfaces (Aerial-RIS), which promise to revolutionize connectivity in both urban and industrial environments by enhancing spectral and energy efficiency in aerial architectures. The integration of AI-driven optimization techniques within Aerial-RIS assisted networks facilitates adaptive resource allocation, improving outage probability, power consumption, and network reliability. This emerging technology is pivotal for realizing the full potential of future 6G networks, addressing key challenges in flexible, dynamic, and sustainable connectivity.

1.4 Research Objectives and Contributions

1.4.1 Problem Statement

This thesis is centered on the advancement of 5G technologies and the advent of Industry 4.0 which necessitates robust and efficient 5G wireless communication systems to support a wide array of applications, including autonomous robotics, IoT integration, and real-time data processing. While 5G and Industry 4.0 promise unprecedented levels of automation and connectivity, their realization is hindered by persistent technical and operational challenges. Despite continuous evolution in radio-access technologies, modern cellular networks still face key bottlenecks:

- Sub-optimal base-station association and load imbalance in heterogeneous networks lead to inefficient spectrum utilization and degraded user experience.
- Limited spectral and energy efficiency under imperfect channel estimation and interference conditions restrict the practical gains of Massive MIMO and cooperative schemes.
- Static spectrum allocation fails to adapt to real-time traffic dynamics, particularly in ultra-dense urban and industrial environments.
- Coverage-enhancement via D2D and IRS techniques and Machine-learning-based optimization for Aerial-RIS-enabled 6G networks.

- latency issues affecting real-time industrial applications, Connectivity gaps, including blind spots in smart factories and rural settings.

Addressing these issues is critical to meeting the performance requirements of next-generation networks. Although 5G technologies have made significant advancements, certain challenges such as limited coverage, high latency, and interference in heterogeneous networks still prevail. These issues impede smart factory environments to achieve the highest performance in terms of reliability and high-speed connectivity which is crucial for operational efficiency. Existing solutions such as IRS, CBS, and D2D communication, have proven effective in mitigating these challenges [CYYL20] [DMP+23] [NARFJ+22] [MPC22] . Additionally, the integration of mmWave communication offers potential for high data rates but faces limitations such as atmospheric absorption and coverage. Although technologies such as massive MIMO and DSS can enhance capacity and optimize resource allocation, their effective integration is still under investigation.

Therefore, the central research problem addressed in this thesis is formulated as: “How can connectivity, efficiency, and reliability be jointly optimized across multi-tier 5G/6G architectures—spanning base-station selection, multi-antenna transmission, spectrum scheduling, and intelligent surface-aided communication—through integrated modeling and data-driven optimization?”

1.4.2 Research Objectives

This research aims to:

- Develop optimized base station selection algorithms to improve user association and load balancing.
- Enhance massive MIMO transmission for improved spectral efficiency and reliability.
- Introduce multi-user STBC-OFDM schemes to reduce BER and improve network robustness.
- Develop Composite Base Station (CBS) and Dynamic Spectrum Scheduling (DSS) techniques for efficient spectrum utilization.
- Develop a D2D-based algorithm for blind spot mitigation in IoT smart industries.

- Design and performance evaluation of IRS-assisted millimeter-wave communication frameworks to improve signal coverage, reliability, and energy efficiency in smart industrial environmentss.
- Design ML-driven compressed communication schemes for quantized phase shift feedback in Aerial RIS-assisted CoMP-NOMA networks to enhance efficiency and reliability.

1.5 Research Questions

To address the challenges associated with enhancing connectivity, efficiency, and reliability in 5G and emerging 6G mobile networks, this research is guided by the following key research questions:

1. **Optimization of Base Station Selection in 5G HetNets**
    - How can optimal base station selection strategies, including *Maximum Received Power* and *Signal-to-Interference-plus-Noise Ratio (SINR)-based association*, improve connectivity and load balancing in heterogeneous 5G networks in a smart factory setting?
2. **Enhancing Throughput and Reliability in Multi-User Massive MIMO**
    - What is the impact of Zero Forcing (ZF) and Maximum Ratio Transmission (MRT) precoding techniques on the performance of multi-user massive MIMO systems in terms of throughput, spectral efficiency, and reliability in dense industrial and urban environments?
3. **Optimization of Multi-User Transmission in 5G Networks**
    - How can the integration of *Space-Time Block Coding (STBC) with Orthogonal Frequency Division Multiplexing (OFDM)* be optimized to *reduce bit error rates (BER), enhance link robustness, and improve spectral and energy efficiency* in next-generation wireless communication systems?
4. **Efficient Spectrum Utilization and Network Performance in Urban 5G Networks**
    - How can *Composite Base Stations (CBS)* and *Dynamic Spectrum Scheduling (DSS)* dynamically allocate *sub-6 GHz and millimeter-wave (mmWave) spectrum* to optimize *spectrum utilization, network*

efficiency, and service quality in urban 5G networks?

5. **Improving IoT Connectivity in Smart Industry Environments**
   - How can a *Device-to-Device (D2D) communication framework* be designed and optimized to *mitigate blind spots, enhance resource allocation, and improve network resilience* in *IoT-driven smart industry environments*?
6. **Intelligent Reflecting Surfaces (IRS) for mmWave Communication in Industrial Deployments**
   - How can the integration of *Intelligent Reflecting Surfaces (IRS)* improve *signal propagation, spectral efficiency, and energy conservation* in *millimeter-wave (mmWave) communication* within high-density industrial networks?
7. **Machine Learning-Driven Optimization in Compressed Aerial-RIS Networks**
   - How can *machine learning-based autoencoder architectures* be utilized to compress *resource allocation, spectral efficiency, and outage probability reduction* in quantized phase shift feedback in Aerial RIS-assisted CoMP-NOMA networks to reduce feedback overhead while maintaining spectral efficiency and minimizing outage probability in dense 6G urban environments?

It is hypothesized that integrating optimized base station selection, massive MIMO, multi-user STBC-OFDM transmission, D2D communication, composite base stations with dynamic spectrum scheduling, IRS-assisted mmWave frameworks, and machine learning-driven compressed feedback schemes will significantly enhance the connectivity, spectral efficiency, reliability, and energy efficiency of 5G and emerging 6G networks, particularly in urban and smart industrial environments.

1.5.1 Key Contributions

This thesis makes a set of coherent and experimentally grounded contributions toward enhancing connectivity, efficiency, and reliability in 5G and emerging 6G wireless networks, with a particular focus on smart industrial and dense urban environments. The contributions span radio access optimization, multi-user transmission techniques, spectrum utilization, device cooperation, intelligent surfaces, and machine-learning-assisted communication, providing a unified framework for next-generation wireless system design.

- **(*Pub 1*) Base Station Selection and Load Balancing in 5G Heterogeneous Networks:** This work presents a systematic evaluation of base station selection strategies based on Maximum Received Power (MRP) and Maximum SINR in dense 5G heterogeneous networks representative of smart factory deployments. The study characterizes the trade-offs between signal quality, load distribution, and user throughput under interference-limited conditions, demonstrating that interference-aware association improves fairness and high-percentile throughput, while proximity-based association offers more stable received power. These results provide practical insights into adaptive cell association mechanisms for industrial 5G networks.
- **(*Pub 2*) Performance Analysis of Precoding Techniques in Multi-User Massive MIMO:** An in-depth link-level evaluation of multi-user massive MIMO precoding is conducted using the Vienna 5G simulator, comparing Zero Forcing (ZF) and Maximum Ratio Transmission (MRT). Under controlled simulation assumptions, ZF is shown to consistently outperform MRT in terms of per-user and sum throughput in both downlink and uplink as the number of base-station antennas increases. The results highlight the capacity advantages of interference-aware precoding in dense multi-user scenarios.
- **(*Pub 3*) Multi-User Space–Time Block Coded OFDM Transmission:** This thesis investigates a multi-user auxiliary-signal-assisted STBC-OFDM transmission scheme that enables simultaneous transmission to multiple users over shared time–frequency resources. Simulation results demonstrate an effective throughput gain of approximately two compared to conventional STBC-OFDM, while preserving diversity benefits and reducing peak-to-average power ratio. The scheme improves throughput efficiency without increasing receiver complexity, making it suitable for link-level multi-user communication scenarios.
- **(*Pub 4*) Composite Base Stations and Dynamic Spectrum Scheduling:** A Composite Base Station (CBS) architecture supporting LTE and 5G NR across sub-6 GHz and 28 GHz bands is evaluated under a traffic-aware Dynamic Spectrum Scheduling framework using system-level simulations. Results demonstrate how dynamic, demand-driven spectrum al-

location can balance the coverage robustness of sub-6 GHz bands with the high-capacity potential of mmWave links in dense urban and campus environments.

- **(*Pub 5*) Device-to-Device Relay Selection for Industrial Blind-Spot Mitigation:** A lightweight, event-driven D2D relay selection algorithm is proposed to improve connectivity availability in industrial IoT environments affected by coverage blind spots. The algorithm prioritizes candidate relays based on link quality and resource constraints, achieving significant improvements in connectivity availability and reducing outage time compared to random relay selection. This work demonstrates the effectiveness of D2D cooperation for resilient industrial connectivity.
- **(*Pub 6*) Intelligent Reflecting Surfaces for mmWave Smart Industry Communications:** The impact of Intelligent Reflecting Surfaces (IRS) is analyzed for smart-factory scenarios operating at mmWave and sub-6 GHz frequencies. Results show that IRS assistance significantly enhances spectral efficiency and enables substantial reductions in required transmit power, highlighting IRS as an energy-efficient solution for improving reliability and coverage in industrial wireless networks.
- **(*Pub 7*) Machine Learning–Driven Compressed Feedback in Aerial RIS-Assisted Networks:** This thesis investigates aerial RIS-assisted CoMP-NOMA downlink transmission and addresses the practical challenge of RIS phase feedback overhead. A machine-learning-based autoencoder framework is developed to compress quantized RIS phase shifts under noisy and fading channels. Simulation results show that compressed feedback preserves performance while reducing overhead, and that aerial RIS deployment improves sum rate, spectral efficiency, energy efficiency, and outage performance, particularly for cell-edge users in interference-limited environments.

In summary, the contributions of this thesis are marked by a thoughtful exploration of strategic approaches to enhancing connectivity, efficiency, and resource management within industrial and urban contexts. While the work introduces innovative methodologies, it remains grounded in practical applications, providing insights relevant to ongoing advancements in wireless communication technologies. These contributions are particularly applicable to the evolution of Industry 4.0 and the development of smart industries, where reliable and efficient communication systems are essential. The thesis offers valuable perspectives that address specific challenges within the field, contributing to the ongoing academic and practical discourse.

1.6 Structure of the Thesis

The remainder of this thesis is organized as follows:

Chapter 2: Optimal Base Station Selection in 5G Networks

Chapter 3: Advancements in Massive MIMO for 5G

Chapter 4: Multi-User Transmission in Future Wireless Networks

Chapter 5: Composite Base Stations and Dynamic Spectrum Scheduling

Chapter 6: D2D Communication for Industry 4.0

Chapter 7: Intelligent Reflecting Surfaces for 5G/6G-Enabled Future Smart Industries with a Focus on Millimeter-Wave Band Communications

Chapter 8: Machine Learning-Driven Optimization of Aerial RIS-Assisted CoMP- NOMA Networks for Enhanced 6G Wireless Performance

Chapter 9: Conclusion and Future Work

References Detailed references supporting the research are provided at the end of the thesis.

1.7 Research Methodology

1.7.1 Overview

This section presents the structured approach adopted in this research to ensure consistency in evaluating the various technologies explored throughout the thesis. It defines key performance metrics, describes the simulation framework, and details the evaluation methodology. The research primarily relies on numerical simulations, analytical modeling, and empirical evaluations to derive meaningful insights into 5G and emerging 6G networks. The methodology is designed to ensure scientific rigor and reproducibility in results.

1.7.2 Key Metrics and Definitions

To maintain clarity and coherence, the following metrics are consistently used across all chapters:

- **Quality of Service (QoS)**: Represents overall network performance, including throughput, latency, reliability, and packet loss.
- **Signal-to-Interference-plus-Noise Ratio (SINR)**: Defined as:

$$\text{SINR} = \frac{P_s}{P_i + P_n} \tag{1.1}$$

  where P_s is the received signal power, P_i is interference power, and P_n is noise power. SINR is crucial in determining link quality and network performance.
- **Wideband SINR**: An extension of SINR that considers the entire allocated spectrum bandwidth, making it essential for evaluating frequency-selective fading effects in wideband transmissions.
- **Throughput**: The effective data rate achieved at the receiver, calculated as:

$$T = B \times \log_2(1 + \text{SINR}) \tag{1.2}$$

  where B is the bandwidth in Hz. This metric is widely used to assess spectral efficiency and network performance.
- **Energy Efficiency**: The ratio of achievable throughput to power consumption, measured in bits per Joule, which is critical for optimizing next-generation wireless systems.
- **Block Error Rate (BLER)**: The probability of a block of data being incorrectly received, providing a measure of transmission reliability and error performance.
- **Latency (ms)**: End-to-end transmission delay.
- **Spectral Efficiency (bits/s/Hz)**: Data rate per unit bandwidth.
- **Fundamental Capacity Considerations**: The ultimate limit on information transfer is defined by Shannon's channel-capacity theorem [Sha48]

$$C = B \log_2(1 + \text{SNR})$$

where C is channel capacity (bit/s), B the bandwidth (Hz), and SNR the signal-to-noise ratio. This relation establishes the theoretical upper bound on spectral efficiency and underpins the design of modern modulation and coding schemes.

1.7.3 Simulation Framework

This research primarily employs computational simulations to evaluate network performance. The models incorporate real-world parameters, including path loss models, fading effects, and network congestion factors. The simulations were conducted using different tools depending on the nature of the study in each chapter:

- **Chapter 2 (System-Level Simulation)**: Vienna 5G System-Level Simulator was used to evaluate optimal base station selection strategies. Key parameters modified included base station power levels, user distributions, and scheduling strategies.
- **Chapter 3 (Link-Level Simulation)**: Vienna 5G Link-Level Simulator was used to analyze advancements in massive MIMO, including ZF and MRT precoding techniques. Parameters adjusted include antenna configurations, modulation schemes, and CSI feedback mechanisms.
- **Chapter 4 (MATLAB-Based Monte Carlo Simulations)**: MATLAB was employed for evaluating multi-user transmission techniques such as OFDM-STBC-based transmission. Monte Carlo simulations varied user channel conditions, noise levels, and coding schemes.
- **Chapter 5 (System-Level Simulation)**: Vienna 5G System-Level Simulator was used to analyze CBS and DSS strategies. The study adjusted spectral resource allocation, inter-cell interference management, and dynamic scheduling algorithms.
- **Chapter 6 (Python-Based Simulations)**: Google Colab (Jupyter Notebook) was used for ML-driven resource allocation in D2D communication. Key modifications included training dataset structures, neural network hyperparameters, and traffic prediction models.
- **Chapter 7 (Hybrid Approach - Link-Level and MATLAB Simulations)**: Vienna 5G Link-Level Simulator was used for IRS-based performance evaluation, while MATLAB was used for complementary analysis. Simulations

included optimization of RIS phase shifts and beamforming algorithms.

- **Chapter 8 (TensorFlow/PyTorch)**: Python-based machine learning simulations (Autoencoders for RIS feedback optimization) were implemented using Google Colab, supplemented with MATLAB for additional validations. Key parameters adjusted included feature extraction techniques, neural network layers, and channel estimation models.

The combination of these tools ensures a comprehensive evaluation of 5G/6G networks across various domains.

1.7.3.1 Simulation Environment

The simulations are conducted in a heterogeneous network (HetNet) topology comprising multiple base stations and user equipment (UEs) in urban and industrial environments. The key simulation parameters include:

- **Network Type**: 5G HetNet with macro, pico, and femto base stations.
- **Carrier Frequency**: 2.5 GH, 3.5 GHz (sub-6 GHz) and 28 GHz (mmWave).
- **Bandwidth**: 100 MHz (sub-6 GHz) and 400 MHz (mmWave).
- **Base Station Transmission Power**: 40 W (macro), 2 W (pico), 0.1 W (femto).
- **Path Loss Model**: 3GPP TR 38.901 Urban Micro (UMi). Free-Space and Urban Macro (UMa) models.
- **Modulation Schemes**: QPSK, 16-QAM, 64-QAM, and 256-QAM.
- **Antenna Configurations**: Single-user and multi-user massive MIMO with Zero-Forcing (ZF) and Maximum Ratio Transmission (MRT) precoding.
- **Scheduling Algorithms**: Round Robin, Proportional Fair, and Opportunistic.
- **Interference Management Techniques**: Dynamic Spectrum Scheduling (DSS) and Composite Base Station (CBS) strategies.
- **Machine Learning-based Optimization**: Applied in Chapters 8 to analyze traffic prediction, resource allocation, and intelligent spectrum management.

1.7.3.2 Evaluation Methodology

The evaluation process follows a structured sequence to ensure consistency and accuracy:

1. **Network Setup**: Define the simulation topology, including base station placement, user distribution, and propagation models.
2. **Parameter Initialization**: Configure transmission power, modulation schemes, channel models, and noise levels based on standardized 3GPP models.
3. **Performance Measurement**: Compute SINR, throughput, BLER, energy efficiency, and other performance indicators under different conditions.
4. **Algorithm Comparison**: Evaluate and compare various scheduling algorithms, cell association strategies, and interference mitigation techniques.
5. **Machine Learning Implementation**: Train and evaluate ML-based resource allocation and traffic prediction models in Python.
6. **Statistical Analysis**: Use Empirical Cumulative Distribution Function (ECDF) plots, mean performance values, and confidence intervals to ensure robust evaluation.

1.7.4 Validation and Benchmarking

To ensure the accuracy and credibility of the results, simulation outcomes are cross-validated against:

- Standardized models from 3GPP TR 38.901, ETSI, and ITU-R recommendations.
- Empirical data from real-world measurements available in the literature.
- Multiple independent simulation runs to confirm statistical reliability.
- Comparative analysis with existing algorithms to demonstrate performance improvements over state-of-the-art techniques.

Chapter 2

Optimal Base Station Selection in 5G Networks

"Wireless is freedom. It's about being unleashed from the telephone cord and having the ability to be virtually anywhere when you want to be."
Martin Cooper (Pioneer in the wireless communications industry)

This chapter offers a comprehensive analysis of base station selection strategies within 5G heterogeneous networks (HetNets), focusing on their application in smart factory environments. By examining the interplay between advanced cell association methods, such as Maximum Received Power and Maximum SINR, this study sheds light on their impact on network performance metrics, including user throughput and signal quality. Leveraging state-of-the-art simulation tools, the research demonstrates how strategic base station deployment enhances connectivity and operational efficiency in low and high-mobility scenarios. The findings underscore the critical role of precise base station selection in optimizing resource allocation, thereby advancing the capabilities of smart manufacturing in the era of Industry 4.0.

2.1 Key Metrics for Base Station Performance

2.1.1 Signal Quality Metrics

In the exploration of signal quality metrics for 5G networks in industrial IoT, considerable emphasis is placed on the pivotal roles of SNR and SINR

[MAP+17]. Metrics like SNR and SINR are essential for assessing the reliability and performance of wireless communication systems, particularly in complex smart factory environments. The accurate assessment of SNR and SINR is instrumental in ensuring the seamless operation of IoT devices, as they dictate the quality of the signal amidst various forms of interference and background noise prevalent in industrial settings [IUIL23].

Additionally, the relationship between signal quality metrics and the overall performance and efficiency of 5G networks in industrial applications is thoroughly analyzed, with an emphasis on their critical role in effective network resource management. Maintaining high signal quality necessitates a delicate balance between signal strength, capacity, and coverage to ensure robust and reliable communication. The comprehensive analysis in [NQA+20] underscores the importance of evaluating these metrics as essential indicators of base station efficacy, facilitating high-performance communication that is fundamental to the success of IoT implementations in modern industrial sectors. Building on this foundation, our study systematically investigates how these metrics contribute not only to overall communication quality but also to the resilience and adaptability of networks across diverse industrial scenarios, ensuring they meet the dynamic demands of these environments.

2.2 Network Throughput, Capacity, and Latency

A key consideration in real-time applications—such as robotic control, industrial automation, and mission-critical IoT—is transmission delay (latency). Low latency ensures timely data delivery, prevents disruptions in process automation, and enhances the reliability of industrial operations. Metrics such as end-to-end delay and jitter play a critical role in enabling real-time responsiveness, particularly in applications where even milliseconds of delay can impact decision-making and operational efficiency.

In high-density industrial environments, optimizing both network throughput and latency is essential. While high throughput facilitates bulk data handling and large-scale communication, minimizing transmission delay ensures that real-time systems operate seamlessly. Emerging technologies—such as edge computing and Ultra-Reliable Low-Latency Communication (URLLC) in 5G networks—enable latency-sensitive applications to meet stringent delay requirements while preserving high throughput levels.

This section, therefore, underscores the importance of a balanced approach that optimizes both throughput and latency, rather than focusing solely on data capacity. The ability to dynamically manage these parameters is crucial for achieving high-performance, low-latency industrial communication in next-generation 5G and 6G smart factory deployments.

2.2.1 Related State-of-the-Art

Efficient base station selection in heterogeneous 5G networks has been extensively studied to optimize user association, load balancing, and throughput. Dhillon *et al.* [DGA13] introduced a stochastic-geometry-based, load-aware model incorporating conditional thinning of the interference field, showing that associating users to the strongest received-power base station improves fairness and coverage probability in multi-tier deployments. Ye *et al.* [YRC+13] further demonstrated that traditional Max-SINR and Max-Received-Power (RSRP) rules, while simple, often lead to macrocell overloading, motivating bias-based and load-aware association techniques for better spectral efficiency. Related work by Singh and Andrews *et al.* [LWC+16] explored the impact of range-expansion biasing in heterogeneous networks, showing that optimized biasing significantly improves user throughput and network fairness compared to pure Max-SINR association.

Building on these foundations, Obi *et al.* [OUST19] proposed a hybrid association scheme integrating Max-SINR and biased-power selection to enhance load balancing in dense small-cell environments, outperforming conventional Max-RSRP and Max-SINR approaches in both throughput and energy metrics. Similarly, Khan *et al.* [QKJR+25] introduced a deep-Q-learning-based, energy-efficient association method that dynamically adapts to user battery levels, balancing macro- and small-cell loads while maintaining fairness and minimizing power consumption. Gharaei *et al.* [HF20] further extended this work by proposing interference-aware and context-adaptive association strategies, combining RSRP and SINR metrics under multi-cell interference and shadowing conditions to enhance reliability and throughput in industrial 5G scenarios.

To ensure consistency and comparability across research efforts, the 3GPP technical reports—TR 38.901 [ETS22] and TR 38.913 [ETS17]—provide standardized channel models, deployment scenarios, and performance metrics for 5G New Radio. These technical reports define baseline evaluation methodologies for hybrid power-based and SINR-based association strategies across frequencies from 0.5 GHz to 100 GHz, forming the foundation for simulation and benchmarking of next-generation HetNet association schemes.

2.3 Simulation and Modeling Techniques

2.3.1 Simulation Environments

We utilize the Vienna 5G System Level Simulator [MAD+18] developed by the Institute of Telecommunications (ITC) at TU Wien. This simulator is a member of the Vienna Cellular Communications Simulators (VCCS) software suite. This simulation is particularly focused on smart factory scenarios, where the integration of various types of base stations—including Macro, Femto, and Pico – is crucial for achieving optimal network performance. The simulated environment replicates a typical smart factory layout, incorporating both low-mobility Internet of Things (IoT) devices and high-mobility entities such as vehicles and robots. Figure 2.1 provides an illustration of mobile networks and also focuses on integrating various base stations - Macro, Femto, and Pico - essential for optimal network functionality in such environments. The setup emulates a typical smart factory layout, including both stationary Internet of Things (IoT) devices and dynamic entities like vehicles and robots, as illustrated in Figure 2.1.

Table 2.1 presents the simulation parameters for testing the effectiveness of a 5G heterogeneous network (HetNet) in the context of a smart factory. These include base station types and parameters, and cell association strategies for use in the simulation.

In this work, an investigation into cell association strategies is carried out, which is pivotal in determining the efficiency and reliability of the network in a smart factory context. Two primary strategies are examined: Maximum Received Power (MRP) and Maximum SINR. In both scenarios, the cell with the highest signal quality metric is selected as the serving (or preferred) cell for each user. Since cell association relies solely on macroscopic fading parameters, the cell association table is calculated for each segment, and handovers between base stations occur only at the beginning of a new segment.

2.3.1.1 Temporal Segmentation in Cell Association

In this work, cell association relies on macroscopic fading parameters, and segmentation is defined as a temporal division of network operation. A segment

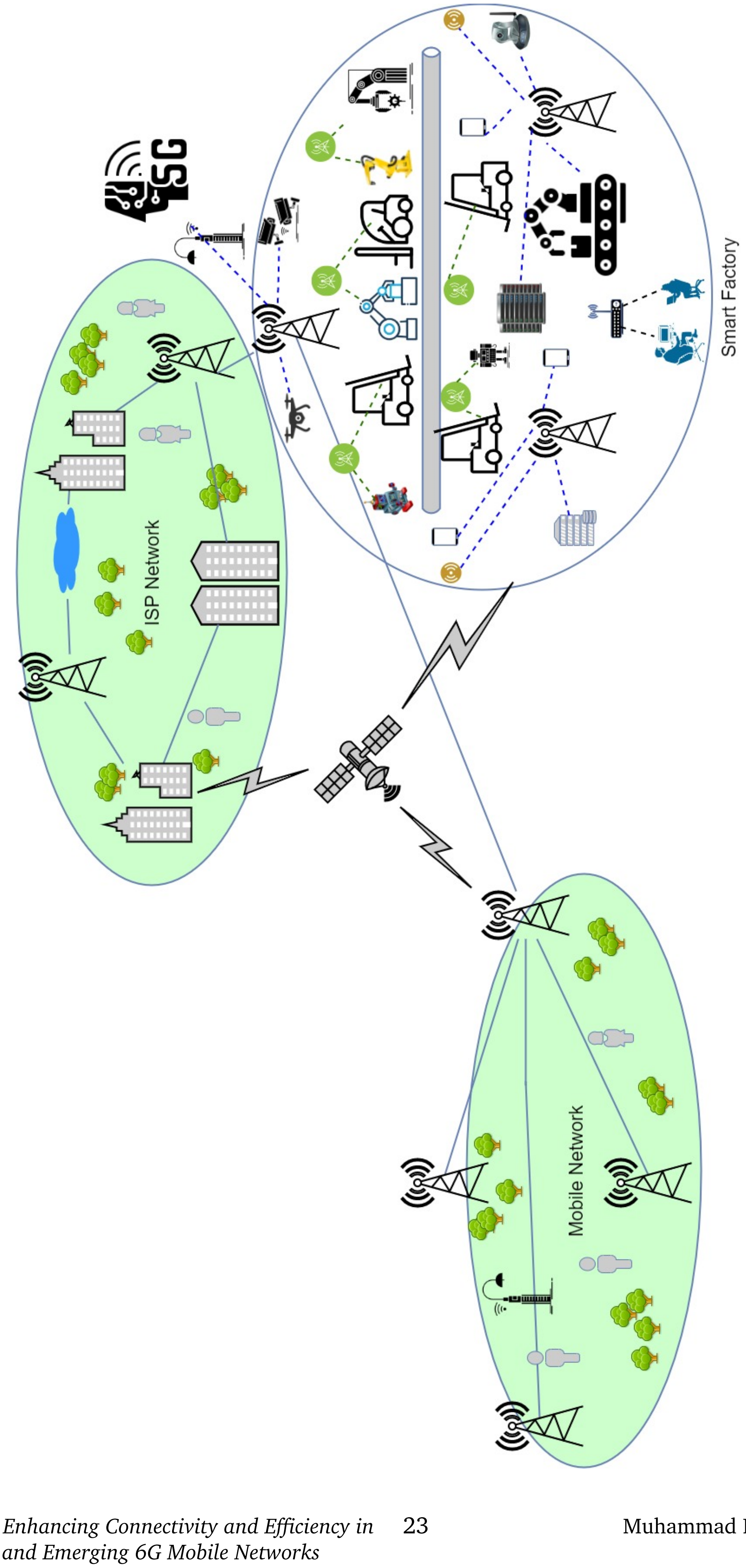

Figure 2.1: A Typical 5G Cellular Network Supporting a Smart Factory

Table 2.1: Simulation Parameters

Parameter	**Value**
BS Types	M, P, F
M Antenna Height	25 m
P Antenna Height	5 m
F Antenna Height	1.5 m
M Transmit Power	40 W
P Transmit Power	2 W
F Transmit Power	0.1 W
Path Loss Model (M)	UMa-3D
Path Loss Model (P)	Free-Space
Path Loss Model (F)	Indoor
UE Types	Low/High Mobility (IoT, Robots)
UE Distribution	Poisson2D/Predefined
Cell Assoc. Strategies	Max SINR, Max Received Power
Simulator	Vienna 5G System

represents a predefined time interval during which macroscopic fading parameters, such as path loss and large-scale interference, remain constant. This temporal segmentation enables efficient handover management and ensures that association decisions are made at the beginning of each segment rather than on a per-slot basis. At the start of a new segment, the cell association table is recalculated, incorporating updated macroscopic conditions to optimize network performance. Handover events between base stations are restricted to segment boundaries, reducing frequent handovers and minimizing signaling overhead. This approach aligns with time-based segmentation models used in advanced system-level simulations, where fading parameters are updated at discrete time intervals rather than continuously. By employing a temporal segmentation strategy, the system achieves a balance between network stability and adaptive association decisions, ensuring optimal performance in dynamic 5G heterogeneous environments.

2.4 Signal Quality Metrics in 5G HetNets

2.4.1 Signal-to-Noise Ratio (SNR)

Signal-to-Noise Ratio (SNR) is a fundamental metric in evaluating the quality of a received signal. It is defined as the ratio of the received signal power to the noise power, given by:

$$\text{SNR} = \frac{P_{\text{signal}}}{P_{\text{noise}}} \tag{2.1}$$

where P_{signal} is the power of the received signal from the serving base station, and P_{noise} is the noise power at the receiver. SNR is calculated using large-scale parameters like path loss, antenna gain, and wall loss. A higher SNR indicates better signal quality and improved communication reliability.

2.4.2 Signal-to-Interference-plus-Noise Ratio (SINR)

SINR extends the concept of SNR by incorporating interference from neighboring base stations. It is computed as:

$$\text{SINR} = \frac{P_{\text{signal}}}{P_{\text{interference}} + P_{\text{noise}}} \tag{2.2}$$

where $P_{\text{interference}}$ represents the total interference power from surrounding/ neighboring base stations. SINR is derived from large-scale network parameters such as Antenna gain, Path loss, Transmission power and Interference from surrounding base stations. SINR plays a crucial role in optimizing resource allocation and ensuring network efficiency.

2.4.3 Wideband SINR in 5G Networks

Wideband SINR provides a more comprehensive evaluation of signal quality across multiple frequency subcarriers. It accounts for small-scale fading effects, scheduling power allocation, and frequency-selective fading over the entire bandwidth. The calculation is given by:

$$\text{Wideband SINR} = \frac{\sum_{k=1}^{N} P_{\text{signal},k}}{\sum_{k=1}^{N}(P_{\text{interference},k} + P_{\text{noise},k})} \tag{2.3}$$

where N is the number of frequency subcarriers, and $P_{\text{signal},k}$ is the received signal power on subcarrier k, $P_{\text{interference},k}$ is the interference on subcarrier k, and $P_{\text{noise},k}$ is the noise on subcarrier k. Simple SINR is calculated based on large-scale parameters only (path loss, antenna gain, interference). Wideband SINR incorporates small-scale fading effects, scheduling power allocation, and frequency selectivity. Wideband SINR provides a more accurate representation of network performance compared to narrowband SINR.

2.4.4 Network Throughput and Capacity

Equation (2.4) [Sha48] expresses the achievable channel capacity (C) of a wireless link as defined by Shannon's theorem:

$$C = B \cdot \log_2(1 + \text{SINR}) \tag{2.4}$$

where B denotes the system bandwidth and SNR represents the signal-to-noise ratio. This equation provides the theoretical upper limit of data rate achievable under given channel conditions, forming the baseline for evaluating throughput and spectral-efficiency performance in the subsequent analysis.

2.4.5 Cell Association Strategies

Two major strategies for user association with base stations are:

- **Maximum Received Power (MRP)**: Users associate with the base station that provides the highest received power.
- **Maximum SINR (MaxSINR)**: Users associate with the base station that provides the highest SINR, reducing interference and optimizing resource utilization.

2.4.6 Path Loss Models

Path loss quantifies signal attenuation due to propagation. Common models include:

- **Urban Macrocell (UMa) Path Loss**: Used for macro base stations in urban settings and large-scale industrial sites.
- **Free-Space Path Loss**: Applied in scenarios with minimal obstructions, such as small cell deployments inside factories.

The general path loss model is given by:

$$PL(d) = PL_0 + 10 \cdot n \cdot \log_{10}(d/d_0) \tag{2.5}$$

[ETS22]

where PL_0 is the reference path loss at distance d_0, n is the path loss exponent, and d is the distance between the transmitter and receiver.

The Urban Macrocell (UMa) path loss model for urban environments, as specified in 3GPP TR 38.901, is given by:

$$PL_{\text{UMa}}(d) = 28 + 22\log_{10}(d) + 20\log_{10}(f_c) \tag{2.6}$$

where d is the distance in meters between the base station and the user equipment, and f_c is the carrier frequency in GHz.

2.4.7 Handover Rate

Handover rate measures how frequently users switch between base stations. It is particularly relevant for mobile users in HetNets and affects seamless connectivity in high-mobility scenarios such as industrial robotic arms or autonomous guided vehicles. The rate can be expressed as:

$$H = \frac{N_{\text{handover}}}{T} \tag{2.7}$$

where N_{handover} is the number of handovers occurring in a time window T.

2.4.8 Definition of Key Metrics with Equations

The selection of an optimal base station is highly dependent on SINR and Wideband SINR values. In this study, the following key metrics are utilized for cell association and performance evaluation. These equations are derived from the simulation code implementation and represent the core principles of network performance assessment.

- **Maximum SINR Selection:** The base station providing the highest SINR is selected, ensuring reduced interference and better user experience, particularly for fixed wireless devices.

  The SINR for a user equipment (UE) connected to a base station (BS) is defined as:

  $$\text{SINR}_i = \frac{P_i}{I + N_0} \tag{2.8}$$

  where:

  - P_i is the received power from the serving base station i.
  - I is the total interference from all other base stations.

  - N_0 is the noise power.

  The Maximum SINR Association selects the base station i^* that maximizes SINR:

$$i^* = \arg\max_i \text{SINR}_i \tag{2.9}$$

  where $\arg\max$ denotes the index of the base station that achieves the maximum SINR. In other words, the base station i^* is selected such that its SINR is higher than that of any other candidate base station. This ensures the user is connected to the base station offering the best signal quality and minimal interference.

- **Maximum Received Power Selection:** The base station with the highest received power is chosen, prioritizing strong signal strength but potentially leading to suboptimal interference conditions.

  The received power at a UE from a base station i is given by:

$$P_i = P_t G_i L_i^{-1} \tag{2.10}$$

  where:

  - P_t is the transmission power of the base station.
  - G_i is the channel gain.
  - L_i is the path loss between the base station and the UE.

  The Maximum Received Power Association assigns the UE to the base station that provides the highest received power:

$$i^* = \arg\max_i P_i \tag{2.11}$$

- **Empirical Cumulative Distribution Function (ECDF)** The Empirical Cumulative Distribution Function (ECDF) represents the probability distribution of a dataset by showing the proportion of values less than or equal to a given value. ECDF is used to compare the distribution of user throughput across different cell association strategies, providing insights into overall network performance and fairness.

  The ECDF of a random variable X is defined as: The ECDF of a random variable X is defined as:

$$F_X(x) = \frac{1}{N}\sum_{i=1}^{N}\mathbb{I}(X_i \leq x) \tag{2.12}$$

where:

- N is the total number of observations.
- $\mathbb{I}(X_i \leq x)$ is an indicator function that equals 1 if $X_i \leq x$, and 0 otherwise.

The ECDF is used to compare the distribution of user throughput across different cell association strategies, providing insights into overall network performance and fairness.

2.4.9 Simulation Framework and Cell Association Strategies

Macroscopic SINR-based association prioritizes signal quality, mitigating interference, while received power-based association optimizes coverage and load balancing. In this context, the Maximum Received Power (MRP) strategy ensures connectivity by associating devices with the strongest signal, while the Maximum SINR strategy enhances communication quality and throughput by selecting the cell with the best signal relative to interference and noise.

The simulator offers two distinct association strategies for the users. One is to attach the users to the BS from which they potentially receive the MRP. The other one takes the inter-cell interference into account by choosing the cell offering the highest wideband SINR (maxSINR) for the user. Wideband SINR is considered because it accounts for the overall signal quality across multiple frequency bands, providing a more comprehensive measure of communication performance. Both the calculation of the wideband SINR as well as the receive power is based on the transmit power of the BSs, the channel in terms of the macroscopic fading and the receivers' noise power.

The network architecture comprises Macro, Pico, and Femto Base Stations, each configured with a single transmit antenna for downlink (DL) cellular multi-tier Single-Input Single-Output (SISO) communication links. User Equipment (UEs) and Femto Base Stations are spatially distributed using a Poisson Point Distribution (Poisson2D), while Macro and Pico Base Stations are positioned at predefined locations. The cell association strategy determines the parameters influencing base station association decisions. Macro Base Stations are characterized by a 25-meter antenna height and a transmit power of 40W, whereas Pico Base Stations have a 5-meter antenna height and a transmit power of 2W, and Femto Base Stations feature a 1.5-meter antenna height with a transmit power of 0.1W. The path loss modeling adapts to the specific base station types: Urban Macro cell (UMa)-3D path loss model is employed for Macro Base Stations, while Pico and Femto Base Stations use the free-space and indoor path loss models, respectively. This configuration enables a versatile and optimized network framework suitable for diverse operational scenarios.

The simulation is designed to reflect real-world conditions in a smart factory. Parameters such as the number and placement of base stations, the density and type of IoT devices, are calibrated to mimic actual industrial environments. This includes varying levels of interference, signal propagation characteristics, and the impact of physical obstacles within the factory. The performance metrics evaluated include signal quality, data throughput, latency, and the handover rate between base stations, providing a comprehensive analysis of network performance under each cell association strategy.

The results from the simulation provide insightful revelations about the efficacy of cell association strategies in a smart factory setting. The findings underscore the superiority of the MRP strategy when paired with an optimized deployment scheme for base stations.

2.4.9.1 Modeling Criteria and Simulation Assumptions Including Mobility Model

The foundation of this simulation is built upon a set of modeling criteria designed to accurately represent a 5G heterogeneous network (HetNet) within a smart factory environment. The network topology consists of three types of base stations—Macro, Pico, and Femto—each serving distinct coverage and capacity roles. Macro base stations provide wide-area coverage, while Pico and Femto base stations enhance localized capacity and indoor connectivity within the factory layout.

Mobility and User Types: The simulation assumes the coexistence of *low-mobility IoT devices* and *high-mobility industrial robots or vehicles*, reflecting the realistic device diversity in modern smart factories. The mobility profiles are defined as follows:

- **IoT and Fixed Machines:** Stationary or quasi-static (0–1 m/s), representing sensors, controllers, and surveillance systems.
- **Pedestrian or Wearable Devices:** Random waypoint mobility with average speed of 1.5 m/s, simulating human operators with connected devices.
- **Vehicular and Robotic Platforms:** Manhattan grid mobility model with an average speed of 60 km/h, representing automated guided vehicles (AGVs) and mobile robotic systems navigating structured factory pathways.

Propagation and Channel Modeling: Signal propagation follows the 3GPP TR 38.901 Urban Macro (UMa) and Urban Micro (UMi) path loss models. Rayleigh fading is assumed for non-line-of-sight (NLOS) channels and Rician fading for line-of-sight (LOS) conditions. Inter-cell interference is incorporated based on user location, transmit power, and antenna configuration. The presence of physical obstacles such as metallic machinery and factory structures is modeled through additional attenuation factors to emulate realistic indoor propagation losses [SKYK11].

Simulation Parameters: Key parameters used in the simulation are summarized as follows:

- Carrier frequency: 3.5 GHz
- System bandwidth: 20 MHz
- Network area: 1 km^2
- Number of users: 200 (heterogeneous mix of IoT, pedestrian, and vehicular)
- Scheduler: Round-robin with adaptive MCS selection
- Antenna model: Sectorized with 17 dBi gain for Macro BS, 10 dBi for Pico/Femto

Performance Evaluation Criteria: The effectiveness of base station selection and association strategies is evaluated using three main performance metrics: signal quality (SINR), user throughput, and the frequency of handovers between base stations. The optimal configuration aims to maximize SINR and throughput while minimizing unnecessary handovers, ensuring both stability and efficiency in dynamic industrial environments.

This comprehensive modeling framework captures the unique characteristics of smart factories—combining diverse device mobility, dense deployment, and industrial interference—to provide a realistic and reproducible basis for evaluating 5G HetNet performance and optimization strategies.

2.5 Results and Discussion

2.5.1 Challenges in Smart Factories

Smart factories represent the pinnacle of industrial innovation, where the integration of advanced wireless systems is not just beneficial but essential. The need for high-performing wireless systems in these environments cannot be overstated. Such systems must offer high data rates and improved spectral efficiency to effectively support the array of mobile robots and autonomous vehicles integral to smart factory operations. This requirement forms the foundation of the discussion, emphasizing the critical role of 5G technology in meeting these demands [MAD+18].

The deployment of 5G networks in smart factories introduces a layer of complexity, particularly due to the multi-tier nature of these networks, comprising Femto, Pico, and Macrocell networks. Each tier plays a unique role in the network, catering to different areas and types of devices within the factory [GLL19].

To tackle these challenges, various wireless devices are embedded in an assembly line, simulated using the Vienna 5G System Level Simulator. This approach allows for a detailed examination of network dynamics and the effectiveness of different strategies in a controlled yet realistic smart factory environment.

We examine the empirical cumulative distribution functions (ecdf) of average user throughput and user SINR for each cell association strategy. This analysis is vital for understanding how different strategies affect the network's ability to serve various devices efficiently. Following this detailed analysis, Figure 2.2 below illustrates representation of the spatial and functional interaction between user devices (e.g., IoT sensors, mobile robots) and base stations (macro, pico, or femto cells) as well as cell associations in a smart factory environment. In this visual representation, the association between UEs and BS denoted by colored lines, where each color represents a different type of base station. High mobility users are well-supported by Macro Base Stations, ensuring consistent connectivity despite movement.Low mobility users benefit from Pico and Femto Base Stations, optimizing resource allocation and reducing load on Macro Base Stations.

2.5.1.1 Base Station Types

Macro Base Stations (Blue): Represented by large nodes emitting many connections. These provide wide coverage and are likely responsible for handling high mobility users like vehicles or robots. Pico Base Stations (Green): Smaller coverage areas, serving low-power and low-mobility devices such as sensors in localized zones. Femto Base Stations (Black): Extremely localized, likely used for high-density areas with specialized devices.

2.5.1.2 User Types

Low Mobility Users (Red): Devices such as stationary IoT nodes or machines that require stable, consistent connectivity. High Mobility Users (Pink): Likely to be moving robots, vehicles, or other mobile entities requiring robust and dynamic connectivity.

2.5.2 Association

In the context of Industrial IoT (IIoT), the association of devices with base stations is a critical factor in ensuring network efficiency, reliability, and scalability. While macro cells provide wide-area coverage, their role in IIoT deployments is often limited compared to small cells (pico/femto), which offer localized, high-capacity coverage with lower latency.

However, macro cells can still play a supporting role in hybrid network architectures, where they serve as a backbone layer to ensure seamless mobility and connectivity across large industrial zones. In such cases, small cells handle the primary IIoT traffic, but macro cells assist in maintaining network continuity, particularly for mobile industrial applications such as automated guided vehicles (AGVs) and remote monitoring systems.

Overall, small cells (pico/femto) are preferred for High-density deployments requiring localized coverage, Low-latency and interference-controlled environments. On the other hand, Macro Cells are useful for providing outdoor-to-indoor connectivity to complement small cells, ensuring mobility support for industrial robots and mobile IIoT devices, and acting as a fallback layer when localized networks experience congestion or failure.

2.5.3 Importance of Efficient Base Station Utilization

The varying demands of user equipment in a smart factory environment pose unique challenges for network management. Low-mobility IoT devices may require stable, continuous connectivity, while high-mobility vehicles demand robust and swift handovers between base stations. Efficient utilization of base stations becomes critical in meeting these diverse needs without compromising network performance as cited in [Pan20].

Choosing the right servicing base station for different types of devices is another critical aspect of efficient base station utilization. The network must dynamically allocate resources to cater to the varied requirements of the devices, ensuring that each base station is optimally utilized to provide the best possible service.

We present experimental results that compare the effectiveness of the Maximum Received Power and Maximum SINR strategies. These results are crucial in understanding how different cell association strategies influence base station utilization and, consequently, network performance. Figure **??** presents the empirical cumulative distribution function (ECDF) of the downlink user SINR for two different base-station association strategies: Maximum Received Power (MRP) and Maximum SINR. Each point on the ECDF curve represents the fraction of users whose SINR values fall below a given threshold, thereby illustrating the overall distribution of signal quality across the network.

Under the MRP association, users connect to the base station providing the strongest received power, regardless of interference levels. In contrast, the Max SINR association considers both desired-signal strength and interference, assigning each user to the base station that maximizes the instantaneous SINR.

The figure shows that the red dashed curve (Max SINR) lies to the left of the blue solid curve (MRP) over the SINR range of approximately –25 dB to –5 dB. This indicates that, within this interval, a larger fraction of users experience lower SINR values under the Max SINR strategy. The apparent degradation stems from the interference-limited nature of the deployment—where SINR-based association may steer users toward cells with transiently favorable instantaneous ratios but ultimately higher co-channel interference under dense

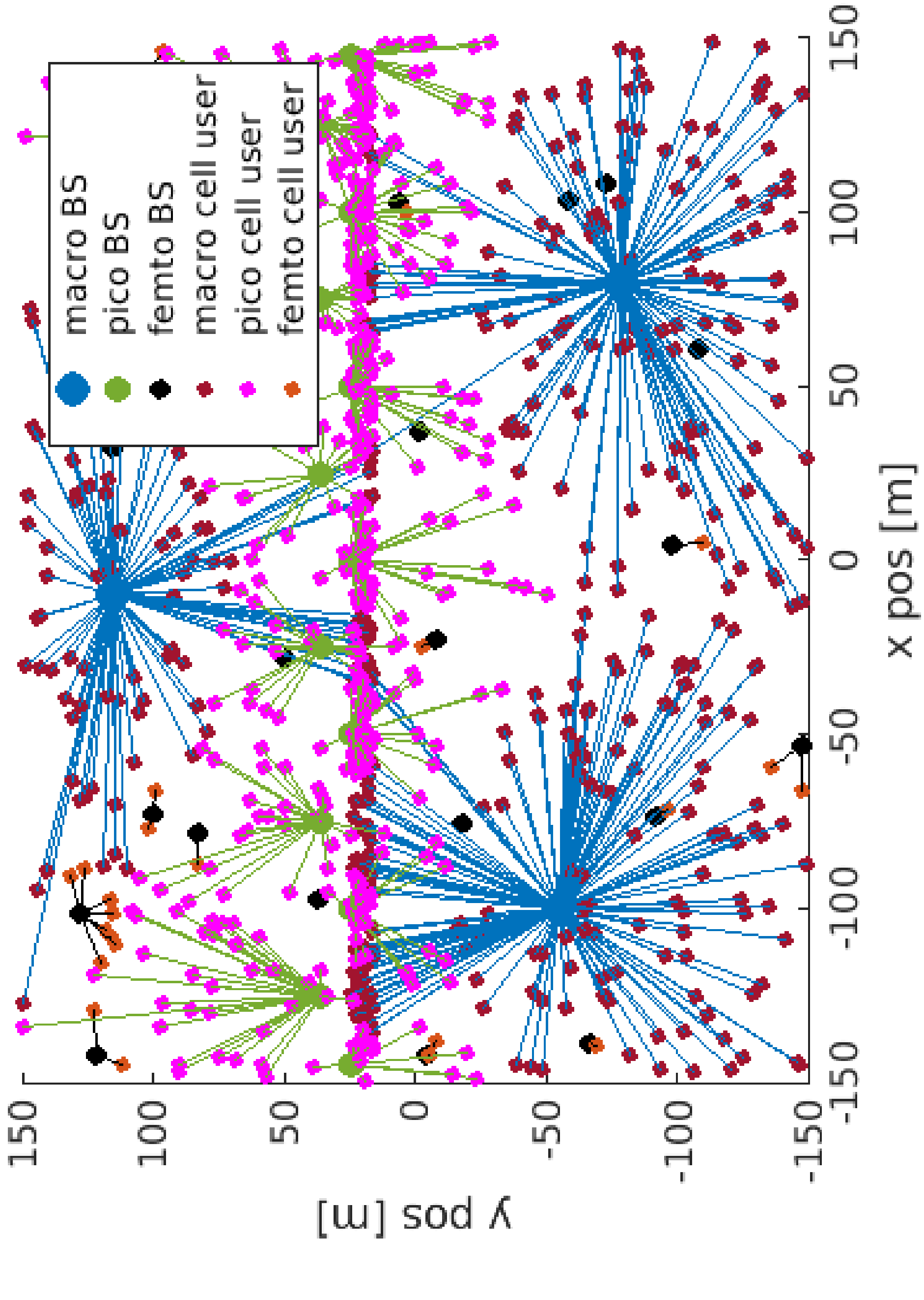

Figure 2.2: User and Base Station Association in a Smart Factory Environment: Users are spread across the factory, with clusters near Pico and Femto Base Stations indicating zones with high device density or specialized tasks. [Maximum SINR association strategy]

frequency-reuse and full-buffer conditions. In contrast, MRP tends to associate users with cells offering consistently stronger average received power, which, under near-uniform interference, results in slightly higher effective SINR values.

At higher SINR levels (above –5 dB), both curves converge, indicating that users located closer to their serving base stations experience nearly identical link quality regardless of the association rule. Overall, the observed left-shift of the Max SINR curve confirms that, under heavy interference coupling, instantaneous SINR-based association does not necessarily guarantee superior link quality compared with the simpler MRP criterion.

Finally, the generally low SINR values observed across both curves are consistent with the deliberately interference-limited baseline configuration employed in this study: frequency-reuse 1, full-buffer traffic, all co-channel cells active across all physical resource blocks, and an indoor smart-factory layout characterized by metallic reflections and penetration losses. Under such stress-test assumptions, aggregate co-channel interference dominates thermal noise, resulting in sub-0 dB SINR for a large user fraction—consistent with theoretical link-budget expectations.

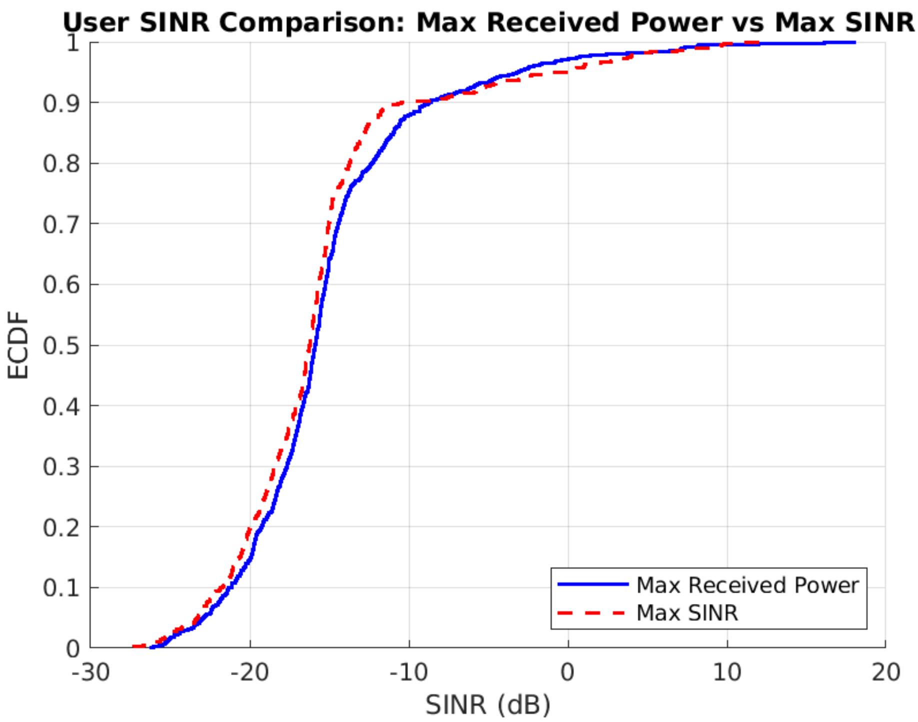

Figure 2.3: User SINR Distribution for Maximum SINR and Maximum Received Power Strategies

Figure 2.4, The plot presents the ECDF of user throughput for the Maximum SINR and Maximum Received Power cell association strategies. The x-axis represents user throughput in Mbps, while the y-axis shows the ECDF, indicating the cumulative probability of users achieving a given throughput or less.

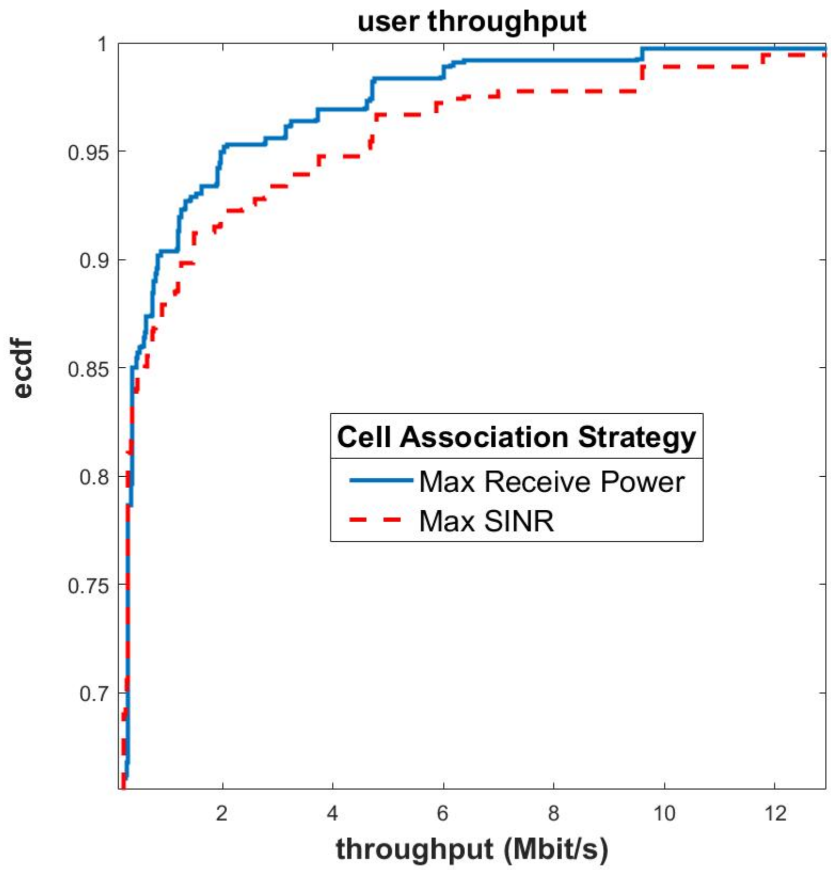

Figure 2.4: User Throughput Distribution for Maximum SINR and Maximum Received Power Strategies

The Maximum SINR (red dashed) strategy exhibits better overall throughput performance than the Maximum Received Power (MRP) strategy (blue line). As seen in the ECDF, for any given cumulative probability below 0.95, the Max SINR curve remains below and slightly to the right of the MRP curve, indicating that a larger fraction of users achieve higher data rates when association is based on SINR rather than purely on received power. For instance, roughly 80 % of users attain throughputs above 4 Mbit/s under the Max SINR policy, whereas the same fraction of MRP-associated users achieves lower rates. This improvement arises because SINR-based association accounts for interference during cell selection, steering users toward base stations that offer cleaner links rather than merely stronger signals.

At higher throughput values (approx. 10–12 Mbit/s), both curves converge, showing that peak users perform similarly under both association rules. However, the right-shifted red curve demonstrates that Max SINR improves link-quality fairness and raises the median user throughput, confirming its effectiveness in interference-limited heterogeneous networks. By mitigating cross-cell interference and distributing users more intelligently across base stations, Max SINR yields a more balanced and efficient use of radio resources without compromising top-end capacity.

In dense industrial or smart-factory deployments, this behavior is particularly valuable. SINR-based association ensures more consistent service quality across users, reduces outage probability for interference-sensitive links, and supports stable high-throughput connectivity required for real-time monitoring and control. While the MRP method remains simpler, it can overload certain cells and degrade spectral efficiency, whereas Max SINR offers a clear fairness and performance advantage in such interference-dominated environments.

Figure 2.5 The heatmap illustrates the cell association distribution for the MaxReceivedPower strategy, depicting the number of users connected to each base station (BS). The x-axis and y-axis range from -2000 to +2000, representing the spatial layout of the smart factory simulation.

The color-coded heatmap shows red for the highest number of associated users and blue for the lowest. Brighter colors indicate areas where users receive stronger signals (higher received power), while darker colors represent regions with weaker signals due to distance from the base station or obstacles like walls.

The color bar represents base station IDs (1 to 15), with each number corresponding to a specific BS, illustrating which areas are served by which stations. The visualization highlights how cell association is influenced by received power, ensuring users connect to the BS with the strongest signal in the given environment.

2.5.4 Evaluation of Base Station Placement Strategies in Smart Factory Networks

Below we have simulated and analyzed different cellular network scenarios, specifically focusing on cell association strategies and HetNets. The simulation produced a map showing which users are associated with which base stations. This map visually represents the distribution of users across the cells based on the maximum received power strategy. A visual representation of the deployment of macro, pico, and femto base stations. This layout helps in understanding the spatial arrangement and density of different types of base stations. A map indicating the LOS and NLOS conditions for users. This can be used to analyze the impact of these conditions on signal quality and network performance. A detailed map showing the received signal strength from different types of base stations at user locations. This can help in evaluating the coverage and capacity provided by the heterogeneous network.

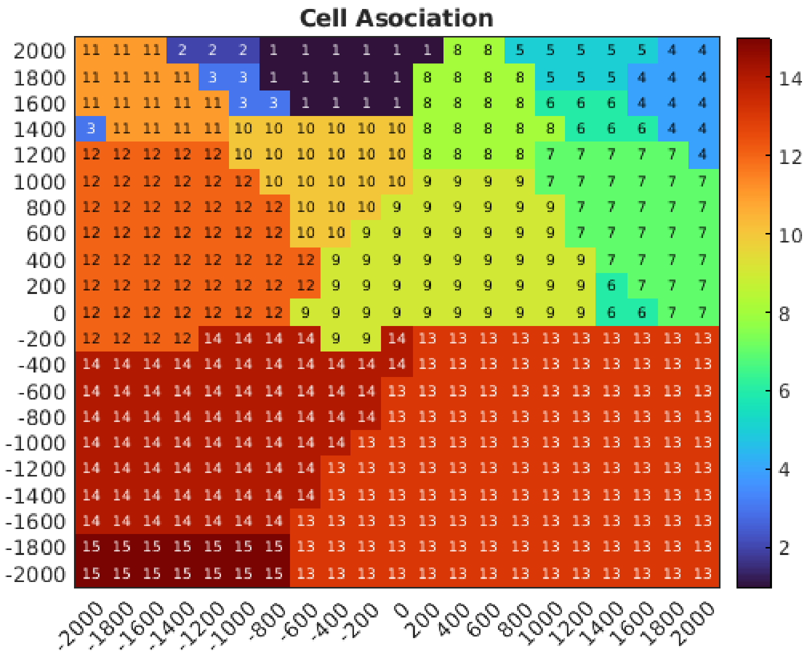

Figure 2.5: Heatmap of cell association for the MaxReceivedPower strategy, showing spatial grid assignment based on signal strength.

Figure 2.6 shows Base Station Position in Region of Interest. The plot represents the positions of the base stations (red dots) and spatial grid points (blue dots) used in the MaxReceivedPower cell association strategy within a smart factory environment. The x-axis and y-axis range from -2000 to +2000, representing the spatial layout of the simulation grid. The blue dots correspond to uniformly distributed grid points across the simulation area, which serve as reference locations to evaluate the coverage and received power from the base stations. The red dots represent the fixed positions of the base stations strategically placed to optimize coverage and signal quality. The black line in the plot represents a physical obstruction or barrier within the smart factory environment. Such lines are included in network models to simulate the impact of walls, machinery, or other obstacles that affect signal propagation and reception. These barriers can

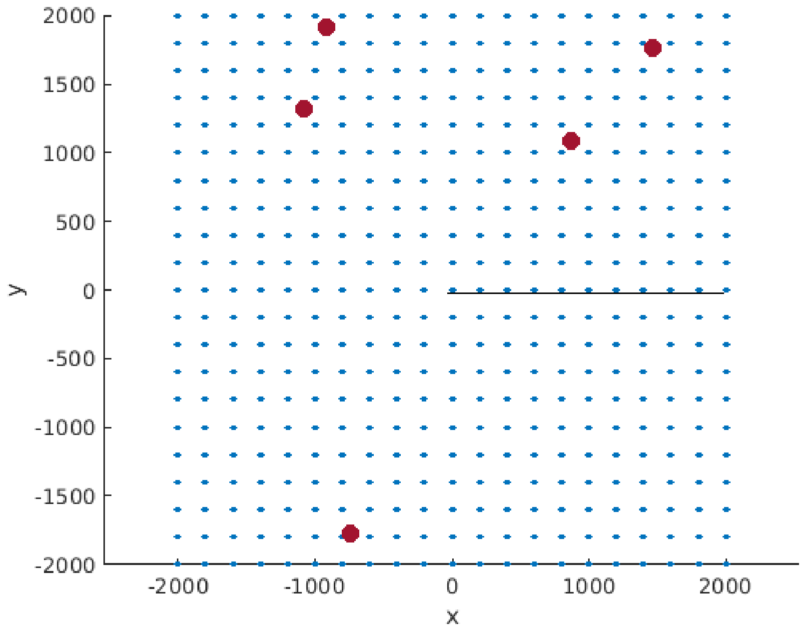

Figure 2.6: BS Positions for Cell Association Strategy: maxReceivePower

influence the received power and SINR at specific grid points (blue dots) by causing signal attenuation, reflection, or shadowing, ultimately affecting the association of those points with the base stations (red dots). This inclusion is essential to model real-world complexities in smart industrial environments.

Figure 2.7 is Distribution of user associations in the network using MRP.

Figure 2.8, 2.9 are for Max SINR.

Figure 2.10 shows Empirical Cumulative Distribution Function (ECDF) plot compares throughput performance across three user groups—IoT users, pedestrian users, and car users—under the MaxReceivedPower cell association strategy. ECDF of user throughput for IoT, pedestrian, and car user groups. The IoT users exhibit the highest potential throughput due to stationary links and favorable line-of-sight conditions, while car users show lower throughput owing to mobility-induced fading and frequent handovers. Pedestrian users maintain intermediate performance with stable but moderate data rates.

Figure 2.11 compares user throughput performance between DLFeedback and DLbestCQI under the MaxReceivedPower association strategy. Both schedulers allocate downlink resources based on Channel Quality Indicator (CQI) reports that reflect real-time channel conditions, including interference, path loss, and fading.

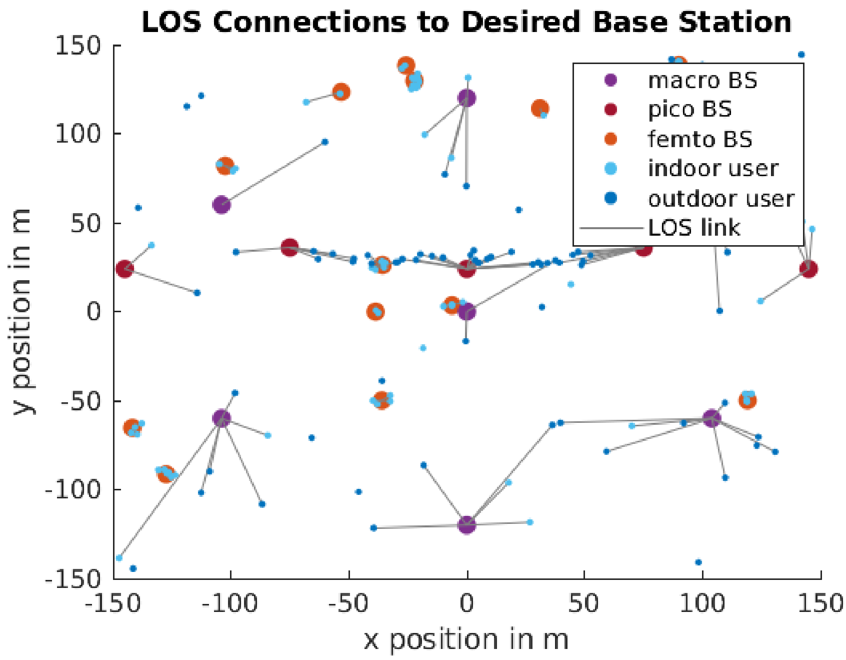

Figure 2.7: MaxReceivePower: Distribution of user associations in the network.

In the DLFeedback scheme, resource allocation dynamically adapts to reported channel quality to enhance signal utilization, manage interference, and promote fairness among users. Conversely, DLbestCQI selects users with the highest instantaneous CQI values, prioritizing peak channel conditions but often neglecting fairness and interference coupling.

The results show that DLbestCQI (orange line) achieves higher throughput, with most users exceeding 2 Mbps and peak values around 10 Mbps, as indicated by its right-shifted ECDF curve. The DLFeedback scheme (blue line) delivers lower throughput and a steeper, left-shifted curve, reflecting more conservative link adaptation. Overall, DLbestCQI demonstrates superior throughput performance due to its aggressive scheduling on subbands with the best instantaneous channel quality, consistent with findings in LTE/5G scheduling literature.

The figure 2.12 depicts the simulation scenario for a smart factory environment, highlighting the spatial distribution of various user types and base stations. The x-axis and y-axis represent spatial coordinates in meters, with nodes distributed across the grid. Macro Base Stations (BS), Pico BS, and Femto BS are shown as purple, red, and orange dots, respectively. Users are classified as vehicular users (blue dots), pedestrian users (cyan dots), and IoT users (green dots), with NOMA pairs represented by connecting lines.

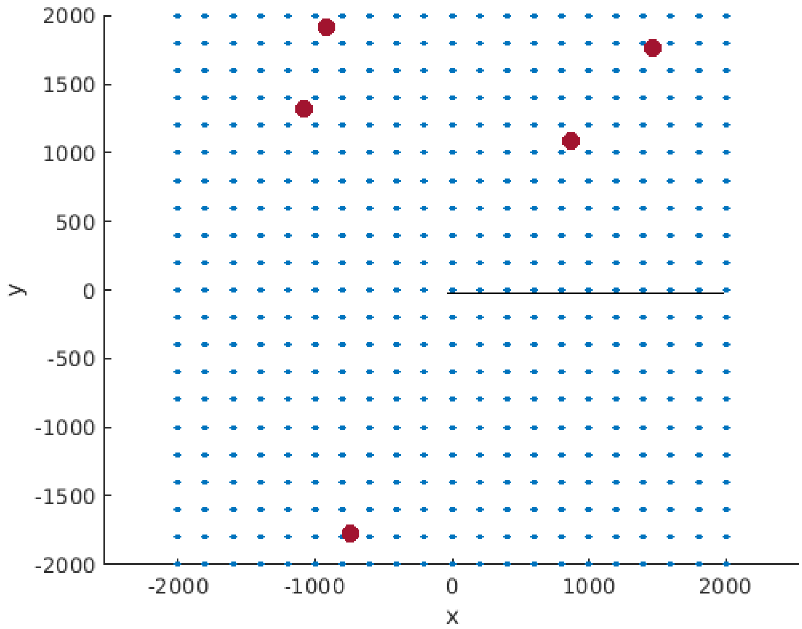

Figure 2.8: BS Positions for Cell Association Strategy: maxSINR

This visualization demonstrates the heterogeneous network (HetNet) setup, where macro BSs provide wide coverage, while pico and femto BSs serve localized areas for specific user groups. The simulation evaluates the efficiency of the MaxReceivedPower strategy, emphasizing optimized resource allocation and load balancing among base stations to support varying mobility and communication requirements in smart factory settings.

The figure 2.13 illustrates the user association patterns within the network using the MaxSINR strategy. The lines connecting users to base stations indicate Line-of-Sight (LOS) connections, emphasizing the association mechanism that prioritizes maximizing the SINR for each user. This visualization provides insights into the network topology, showing the distribution of base stations, users, and their respective associations for efficient connectivity.

The figure 2.14 shows how users are associated with different base stations based on the MaxSINR strategy, emphasizing the strategy's focus on signal quality and network efficiency. The graph shows the CDF of user throughput for two different feedback mechanisms: DLFeedback and DLbestCQI. The DLbestCQI scheme (orange line) achieves higher throughput, with most users exceeding 5 Mbit/s and peak values approaching 20 Mbit/s. Its right-shifted ECDF curve shows that a larger share of users experience higher data rates. In contrast, the DLFeedback scheme (blue line) yields lower throughput.

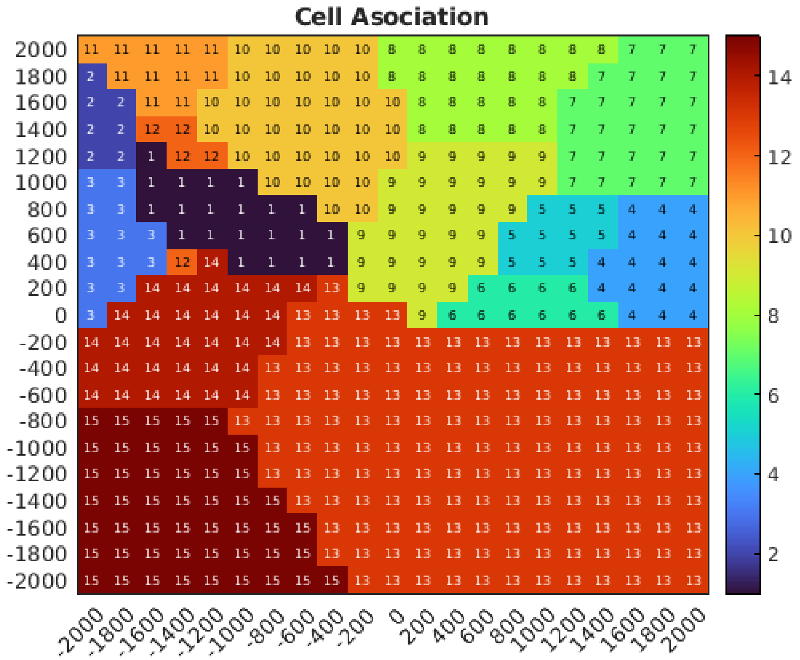

Figure 2.9: Heat Map of User Associations for MaxSINR Strategy

The figure 2.15 illustrates the distribution of SINR across the network under the MaxSINR strategy, showing areas of high and low signal quality and the effectiveness in managing interference. The visual representation of the network's layout and user distribution helps in understanding network efficiency and coverage.

Figure 2.16 presents an ECDF graph comparing throughput performance among IoT, pedestrian, and car users. This figure illustrates that stationary or low-mobility IoT devices benefit from stable channels and low interference, while vehicular users suffer from mobility-induced degradation despite short-term CQI gains.

2.6 Conclusion

This chapter presented a comprehensive performance evaluation of base-station association and scheduling strategies in 5G heterogeneous networks for industrial and urban scenarios. Through extensive ECDF-based analysis of SINR and user throughput, the study compared Maximum Received Power (MRP) and Maximum SINR association policies, as well as downlink scheduling methods (DLFeedback vs DLbestCQI) and user-type variations (IoT, pedestrian, vehicular).

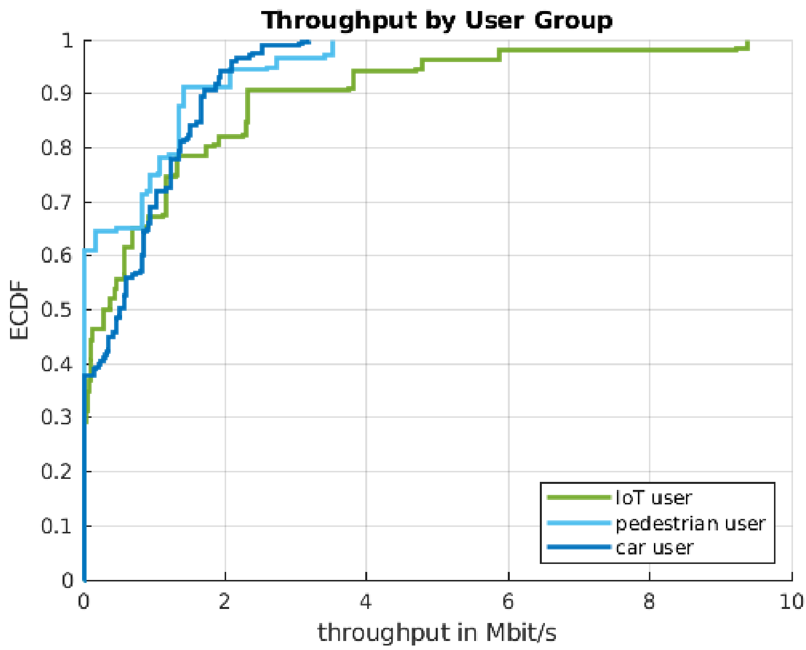

Figure 2.10: The ECDF plot shows throughput distribution for IoT, pedestrian, and car users.[MaxReceivedPower]

Results show that while the Maximum SINR association improves signal quality for interference-sensitive users, it can direct users toward cells with higher co-channel interference, resulting in generally lower SINR levels compared with MRP. However, throughput analysis reveals that Maximum SINR achieves better overall data-rate performance and improved link-quality fairness, since it accounts for interference when selecting serving cells. In contrast, MRP offers stronger average power connections and simpler implementation but may cause load imbalance and reduced spectral efficiency in dense deployments. In scheduling analysis, DLbestCQI outperformed DLFeedback, achieving higher average and peak throughput (up to 12 Mbit/s) owing to effective exploitation of instantaneous channel quality for link adaptation, while DLFeedback favored fairness through adaptive feedback-based allocation. Among user categories, IoT nodes exhibited the highest throughput, benefiting from static positioning and proximity to serving cells, followed by pedestrian users, while vehicular users experienced degraded performance due to mobility-induced fading and frequent handovers.

Overall, the chapter established that adaptive, interference-aware, and mobility-conscious cell-association and scheduling mechanisms are essential for sustaining high spectral efficiency and reliability in next-generation industrial 5G HetNets.

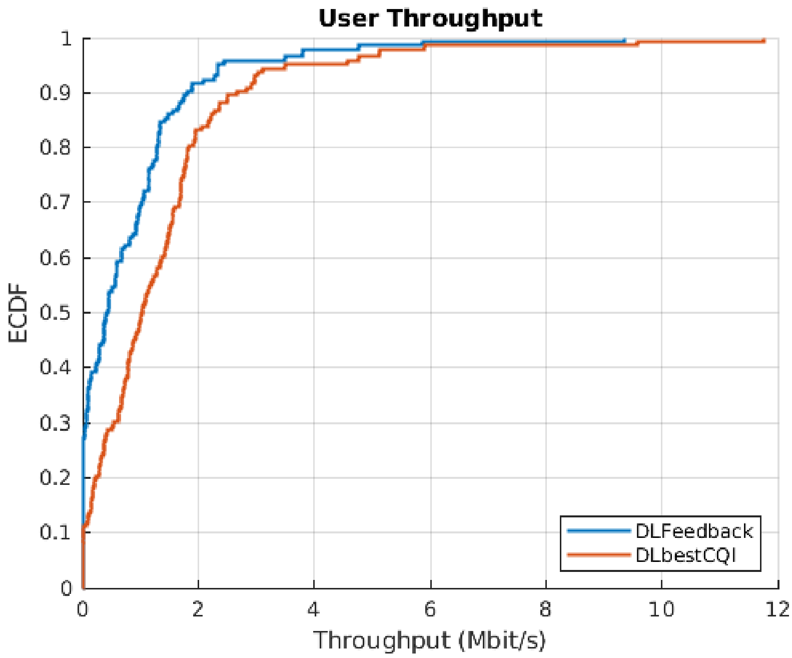

Figure 2.11: ECDF graph compares user throughput performance for DLFeedback and DLbestCQI methods. [MaxReceivedPower]

2.6.1 Key Findings

Key findings from this chapter highlight that:

- **Max SINR association** yields higher throughput and fairness under interference-limited conditions by accounting for co-channel interference during cell selection.
- **MRP association** provides stronger coverage and simpler implementation but can cause uneven load distribution.
- **DLbestCQI scheduling** achieves higher average and peak throughputs by prioritizing subbands with the best instantaneous CQI.
- **DLFeedback scheduling** favors fairness and stability, resulting in lower but more consistent throughput.
- **IoT, pedestrian, and car users** exhibit distinct throughput characteristics, where IoT $>$ pedestrian $>$ car, highlighting the influence of mobility and channel variability.
- Combining **SINR-aware association and adaptive scheduling** ensures better spectral efficiency, reliability, and service uniformity for industrial 5G HetNets.

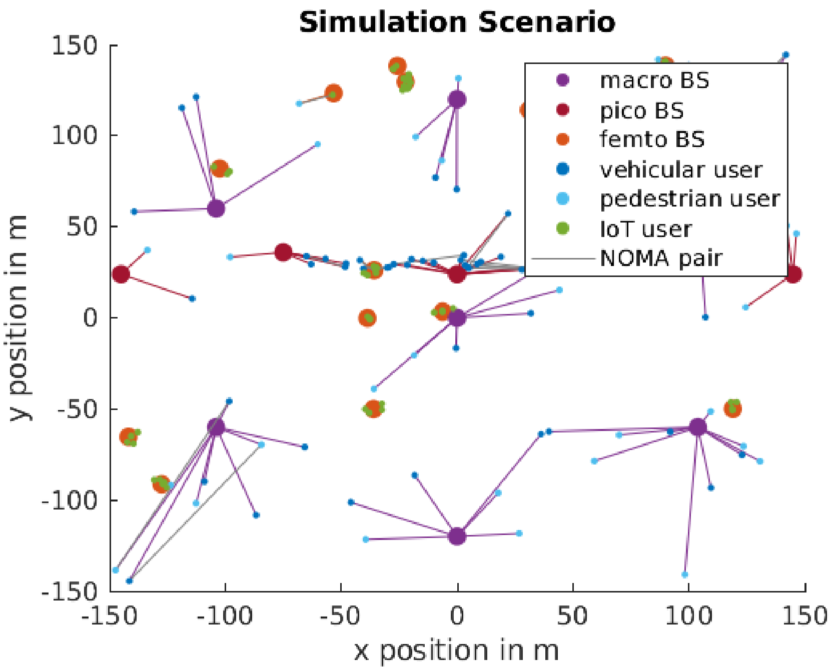

Figure 2.12: Figure illustrates the simulation scenario in a smart factory, showing macro (purple), pico (red), and femto (orange) base stations, along with vehicular (blue), pedestrian (cyan), and IoT (green) users. NOMA pairs are indicated by connecting lines, highlighting the heterogeneous network setup for evaluating resource allocation and load balancing. [MaxReceivedPower strateg]

2.6.2 Relevance to Research Questions and Novel Contributions

This chapter directly addresses the research questions posed in this study by:

- **Demonstrating how base-station association strategies influence network performance**, thereby addressing the problem of optimal cell selection in 5G heterogeneous networks for industrial and smart-factory environments.
- **Providing empirical validation through SINR and throughput ECDF analyses**, showing that Maximum SINR association improves throughput and fairness under interference-limited conditions, whereas Maximum Received Power (MRP) ensures stable coverage and balanced load distribution.
- **Evaluating scheduling strategies (DLbestCQI vs. DLFeedback)** to illustrate how CQI-based resource allocation enhances link adaptation and overall spectral efficiency, while feedback-driven scheduling promotes sta-

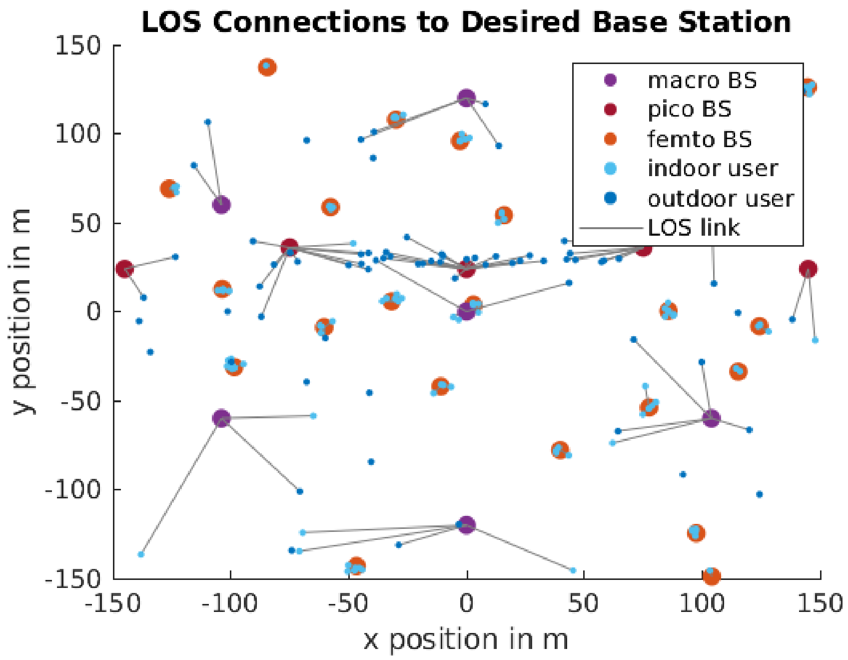

Figure 2.13: This figure visualizes the spatial distribution of users and their associated base stations under the MaxSINR strategy, incorporating Line-of-Sight (LOS) connections.

bility and fairness.

- **Analyzing user-type performance** (IoT, pedestrian, and vehicular) to reveal how mobility and channel dynamics affect throughput—critical for developing mobility-aware association and scheduling mechanisms.
- **Highlighting the importance of multi-tier integration** of macro, pico, and femto base stations to balance coverage, capacity, and reliability across diverse industrial traffic demands.
- **Establishing the analytical foundation** for subsequent chapters on multi-user MIMO, composite base-station architectures, and dynamic spectrum scheduling (DSS), which further extend these insights to enhance 5G/6G network connectivity and efficiency.

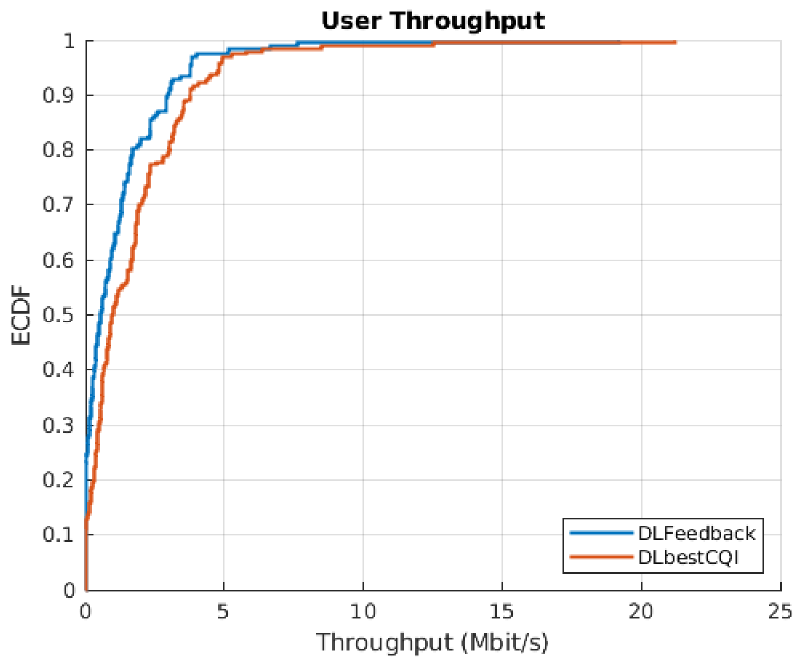

Figure 2.14: MaxSINR: Distribution of user associations based on signal quality.

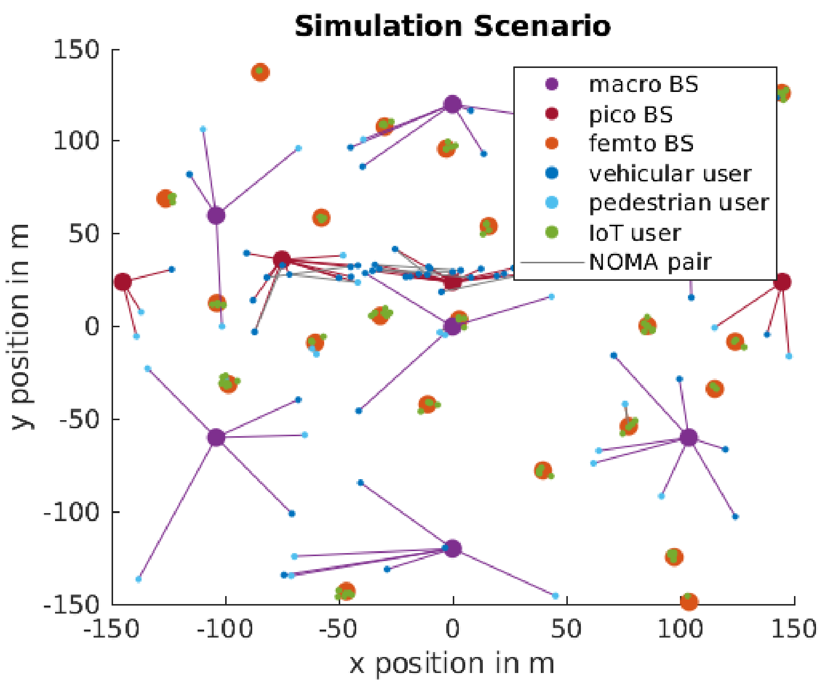

Figure 2.15: MaxSINR: Distribution of signal quality and interference across the network.

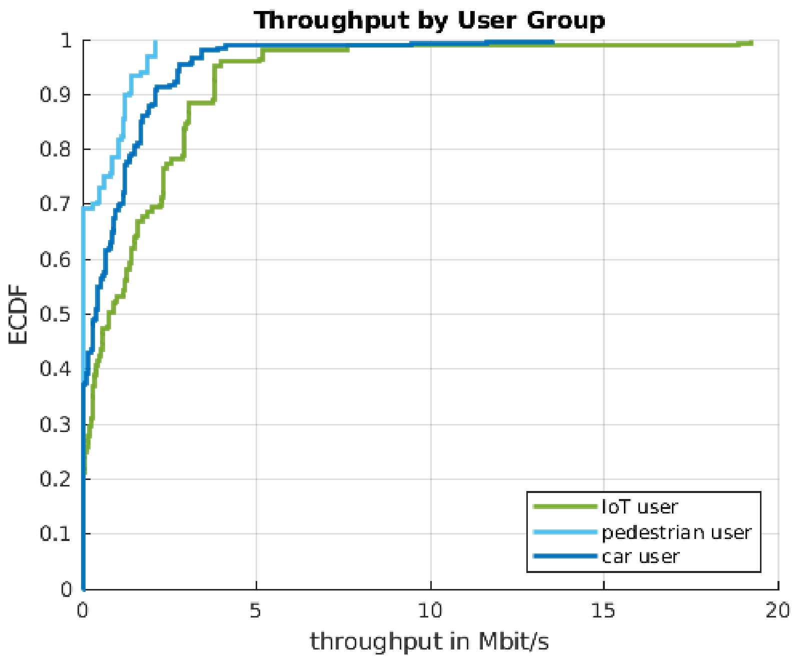

Figure 2.16: The ECDF graph compares throughput performance among IoT, pedestrian, and car users under different mobility scenarios.[MaxSINR]

Chapter 3

Advancements in Massive MIMO for 5G

"Scientific thought and its creation is the common and shared heritage of mankind."
Nobel laureate Professor Dr. Abdus Salam (1926-1996)

This chapter focuses on the advancements regarding massive MIMO in 5g wireless communications. MIMO stands at the promises to drastically enhance network capacity, efficiency, and reliability. Two precoding techniques, notably Zero Forcing (ZF) and Maximum Ratio Transmission (MRT), which have emerged as key players in optimizing the performance of massive MIMO systems are studied. The complexities of these techniques are analyzed to reveal their fundamental roles, particularly ZF's effectiveness in massive MIMO settings[IA17].

The introduction presents a comprehensive examination of the theoretical evolution of MIMO technologies. The chapter then transitions into an in-depth discussion on the practical aspects of implementing MIMO technologies, particularly emphasizing the challenges and considerations in system design and deployment. It covers a range of topics, from the intricate design considerations that dictate the architecture and functionality of MIMO systems to the practical challenges encountered during their implementation.

This comparative study goes beyond a simple juxtaposition of the techniques; it examines into the mechanics of how each technique functions, their respective advantages and limitations, and, most importantly, their performance and suitability.

3.1 Theoretical Underpinnings of MIMO Technologies

3.1.1 Fundamental Concepts

5G is looked at as a key enabler for mobile communication system based automation applications in smart manufacturing, Industrial Internet of Things (IIoT), smart cities and autonomous vehicles and robotics. With such heterogeneous use cases, the deployment of a 5G network has become a challenge. 5G networks will play an important role in future Cyber-Physical Manufacturing Systems (CPMS) [Kha20]. In order to support enhanced requirements, large-scale or massive MIMO systems are employed at BSs. They employ antenna arrays that have more elements (one hundred or more antenna elements) than traditional MIMO systems. Massive MIMO works on the principle of favorable propagation, which occurs when the channel vectors between the BS and each MS are nearly orthogonal as the number of BS antennas are increased [Stu17]. We examine the performance of massive MIMO systems in 5G for two MIMO pre-coding techniques, ZF and MRT. Related work includes [MOE15] where an evaluation of the performance of two linear precoding technologies, namely ZF and MRT for the downlink massive MIMO was carried out using two normalization methods namely vector and matrix normalization to eliminate or mitigate the Multi user Interference. The authors concluded that vector/matrix normalization for ZF performs better at high power, while MRT performs better at low power. Also, according to their findings, the ZF technique is superior than MRT for users near the cell centre (high power), whereas MRT is superior to ZF for users at the cell boundary (low power). Another comparable study is reported in [PKJ14], where a performance analysis and comparison of ZF and MRT based downlink massive MIMO systems was carried out focused on achievable data rate plus energy efficiency aspect and the required transmit power simultaneously, under same conditions / assumptions. Parameters which were studied are the achievable sum rate and the total downlink transmit powers, results showed that ZF achieves higher data rates and is also more power efficient than MRT. All these studies assumed a single-cell downlink massive MIMO system. The study in [Pan20] used spectrum efficiency and energy efficiency as metrics to examine the performance of large-scale multi-user MIMO systems. They used the Open Source Lena 5GNR and NS3-MmWave simulators and claimed that it is a first of its kind work reported. Compared with the MRT-MRC precoder, the ZF precoder has better performance. In [PT18], relative performance analysis of linear precoding in downlink multi-user MIMO was examined and it was shown that the ZF precoding gives a better reachable sum rate (bits/s/Hz). Not only does the low complexity implementation of MIMO systems, such as hybrid beamforming, become more significant as the number of antennas increases, but the spatial features of wireless communications channels also become more critical. The research presented in this paper aims to address performance of linear precoding / beamforming techniques under such conditions.

Table 3.1: Comparison of Zero Forcing (ZF) and Maximum Ratio Transmission (MRT) Beamforming Techniques

Feature	Zero Forcing (ZF)	Maximum Ratio Transmission (MRT)
Interference Management	Actively suppresses interference by forming nulls	Does not actively suppress interference
Signal Strength	Moderate (energy spent nulling interference)	High (focuses on maximizing SNR for intended user)
Computational Complexity	High	Moderate
Efficiency	Best in interference-heavy environments	Best for high-SNR, low-interference scenarios

Figure 3.1 visually depicts the operational mechanisms of both ZF and MRT in a multi-user network environment. This diagram compares Zero Forcing (ZF) and Maximum Ratio Transmission (MRT) beamforming techniques in MIMO systems. ZF focuses on suppressing interference by creating nulls in the beam pattern for unintended devices, ensuring cleaner communication for the intended user but at the cost of higher computational complexity and reduced power efficiency. In contrast, MRT maximizes the signal strength for the intended user by aligning phase and amplitude across antennas, prioritizing signal-to-noise ratio (SNR) over interference suppression. The green region in both cases represents the beam directed toward the intended device (desired user). For ZF, the red regions indicate nulls created to suppress interference for other devices, ensuring minimal signal overlap. For MRT, the green region highlights a strong, focused beam for the intended device, with less emphasis on interference suppression, leading to residual coverage near other devices. ZF is ideal for dense networks with high interference, while MRT performs better in low-interference environments where maximizing SNR is key. Both approaches address different network scenarios, balancing trade-offs between interference management and computational simplicity.

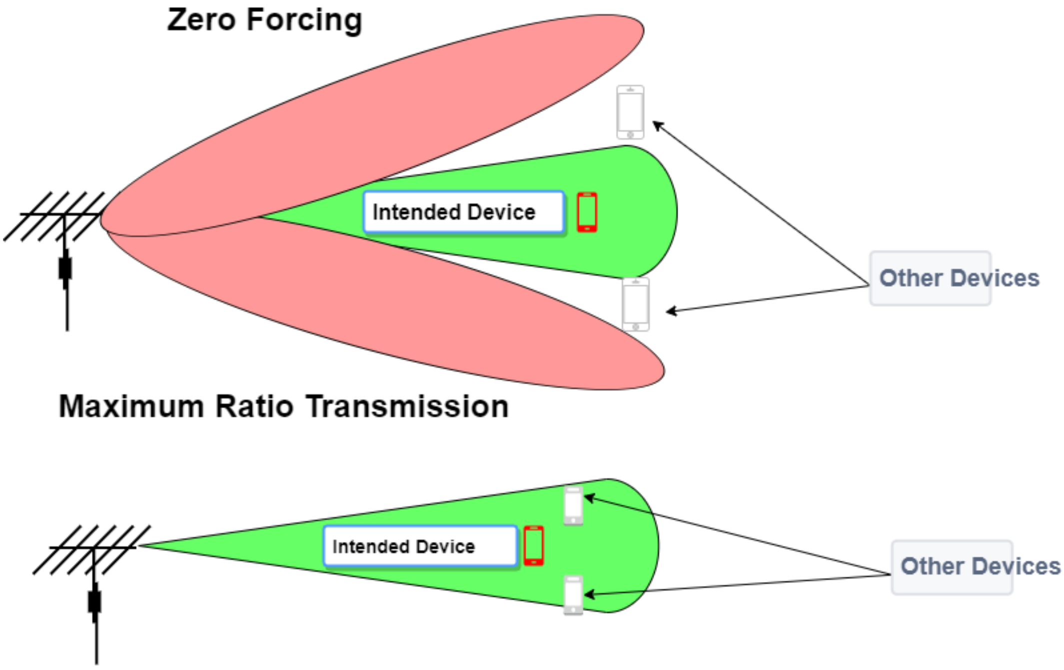

Figure 3.1: Comparison of Zero Forcing (ZF) and Maximum Ratio Transmission (MRT) Beamforming Techniques. Both approaches address different network scenarios, balancing trade-offs between interference management and computational simplicity.

3.1.2 Related State-of-the-Art in MIMO Systems

The concept of Multiple-Input Multiple-Output (MIMO) systems has evolved through several key milestones that underpin modern 5G and 6G wireless technologies. Early theoretical foundations were laid by Paulraj and Kailath (1993) [PK93], who introduced the spatial multiplexing concept that exploits multipath propagation to increase channel capacity without additional bandwidth or power.

Foschini (1996) [Fos96] expanded this idea with the Bell Labs Layered Space-Time (BLAST) architecture, which demonstrated that the achievable capacity of MIMO grows linearly with the number of transmit and receive antennas, provided rich scattering environments. This marked a paradigm shift in wireless communications, allowing multiple independent data streams to be transmitted simultaneously over the same frequency band.

Telatar (1999) [Tel99] provided the first mathematical capacity formulation for MIMO channels, establishing that capacity scales with $\min(M_t, M_r)$, where M_t and M_r are the numbers of transmit and receive antennas, respectively.

The introduction of Massive MIMO by Marzetta (2010) [Mar10] revolutionized cellular system design by scaling antenna arrays to hundreds of elements at the base station, thereby enabling high spatial multiplexing gains and energy-efficient communication. Building on this foundation, Larsson et al. (2014) [LETM14] demonstrated that the use of a large excess of service antennas over active terminals—operating in time-division duplex (TDD) mode—can increase capacity by more than an order of magnitude while improving radiated energy efficiency by up to two orders. Massive MIMO focuses transmitted energy into narrow spatial regions, enhancing robustness, reliability, and spectral efficiency even in interference-limited or fading environments. The approach also leverages low-cost, low-power components and exhibits resilience to hardware imperfections due to the law of large numbers. Subsequent empirical and analytical studies confirmed that favorable propagation and channel hardening properties observed in realistic deployments validate the scalability and feasibility of Massive MIMO for 5G and beyond

This study builds upon these foundational works, extending their principles toward optimized precoding (ZF and MRT) under multi-user and heterogeneous 5G scenarios to evaluate the trade-offs between spectral efficiency, power consumption, and computational complexity.

3.2 System Model

A MU-MIMO system is one in which the base station allows for multiple parallel communications to occur in the same time and frequency resource, which is known as Space Division Multiple Access (SDMA). In this work, we consider a MU-MIMO downlink system.

3.2.1 System Model for Massive MIMO

The received signal at a given user can be modeled as:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{3.1}$$

where:

- $\mathbf{H}$ is the channel matrix between the base station and users.
- $\mathbf{x}$ is the transmitted signal vector.

- $\mathbf{n}$ represents the additive white Gaussian noise (AWGN).

3.2.2 SINR and Spectral Efficiency Analysis

The Signal-to-Interference-plus-Noise Ratio (SINR) in a massive MIMO system with K users is given by:

$$\text{SINR}_k = \frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{\sum_{j \neq k} |\mathbf{h}_k^H \mathbf{w}_j|^2 + \sigma^2} \tag{3.2}$$

where:

- $\mathbf{h}_k$ is the channel vector for user k.
- $\mathbf{w}_k$ is the precoding vector for user k.
- σ^2 is the noise power.

The spectral efficiency for each user is defined as:

$$SE_k = \log_2(1 + \text{SINR}_k) \tag{3.3}$$

The sum spectral efficiency for all users is:

$$SE_{\text{sum}} = \sum_{k=1}^{K} SE_k \tag{3.4}$$

3.2.3 Simulator-Based Precoding Matrix Selection

The link-level simulator selects the precoding matrix using:

$$X_{\text{precoded}} = X_{\text{data}} \cdot W_{\text{precoding}} \tag{3.5}$$

where:

- X_{data} represents the transmitted data symbols.
- $W_{\text{precoding}}$ is the selected precoding matrix.

3.2.4 Path Loss Model for Massive MIMO

The simulator models the path loss attenuation for massive MIMO as:

$$PL(d)[\text{dB}] = PL(d_0) + 10n \log_{10}\left(\frac{d}{d_0}\right) + X_\sigma \tag{3.6}$$

where $PL(d_0)$ is the path loss at the reference distance $d_0 = 1\,\text{m}$, n is the environment-dependent path loss exponent, and X_σ is a log-normal shadowing component with standard deviation σ.

3.2.5 Throughput Analysis and Performance Metrics

For a given bandwidth B, the throughput per user is:

$$T_k = B \cdot \log_2(1 + \text{SINR}_k) \tag{3.7}$$

The total network throughput is:

$$T_{\text{total}} = \sum_{k=1}^{K} T_k \tag{3.8}$$

To further evaluate performance, the Bit Error Rate (BER), which measures the probability of bit errors, can be computed as:

$$BER = \frac{Q(\sqrt{2 \cdot \text{SNR}})}{2} \tag{3.9}$$

where $Q(.)$ is the Q-function, representing the tail probability of a standard normal distribution. Additionally, the Energy Efficiency (EE) of the system is given by:

$$EE = \frac{T_{\text{total}}}{P_{\text{total}}} \tag{3.10}$$

where P_{total} is the total power consumption of the base station.

This framework provides a comprehensive basis for evaluating the effectiveness of massive MIMO techniques, ensuring optimized performance in 5G HetNet deployments.

3.3 Precoding Techniques

3.3.1 Zero Forcing (ZF) Technique

The precoder takes the output of the layer mapper and maps it to the different antennas. The pre-coding can be configured in different ways, corresponding to different transmission schemes. We investigate the performance of linear precoding in order to analyse massive MU-MIMO downlink systems. Linear precoding is a simple process that involves multiplying the information data vector by a linear precoding matrix to obtain the transmission signal vector. We use two precoding techniques, ZF and MRT as they are common for massive MIMO. ZF minimizes inter-beam interference assuming that the base station has perfect channel state information for the downlink, which is the ideal case. In ZF the inter-user interference can be cancelled out at each user. We also assume that in our simulation each of the base stations receives pilot signals from user equipment and is able to determine perfect channel state information (CSI). We then calculate SINR and throughput for each of the users in the cells. ZF precoding implement a pseudo-inverse of the channel matrix [Pak13].

Zero Forcing (ZF) is a linear precoding technique that eliminates inter-user interference by inverting the channel matrix, assuming perfect channel state information (CSI) at the base station. The equations are borrowed from [SSCY17][ZCM+18][WLZ18] The ZF precoding matrix is given by:

$$A_{ZF} = \frac{1}{\beta} H^*(H^T H^*)^{-1} \tag{3.11}$$

where H is the channel matrix, which represents the wireless channel between the base station (BS) and the users (UEs). It captures the fading coefficients that model the signal propagation environment, and β is the normalization factor ensuring power constraints:

$$\beta = \sqrt{\frac{\mathrm{tr}(BB^H)}{P_{tr}}} \tag{3.12}$$

where:

$$B = H^*(H^T H^*)^{-1} \tag{3.13}$$

Here, A_{ZF} is the ZF precoding matrix, and P_{tr} represents the total transmit power. The SINR for ZF precoding is expressed as:

$$\text{SINR}_k^{\text{ZF}} = \frac{|\mathbf{h}_k^H A_{ZF,k}|^2}{\sum_{j\neq k} |\mathbf{h}_k^H A_{ZF,j}|^2 + \sigma^2} \tag{3.14}$$

where $A_{ZF,k}$ is the ZF precoding vector for user k, and σ^2 is the noise variance. ZF minimizes interference at the cost of noise enhancement due to the matrix inversion.

ZF sets beamforming weights to minimize interference to all other users in a cell, placing them within local nulls.

ZF attempts to minimize the amount of interference that the beam causes to other users while maximizing the signal strength to an intended user equipment.

3.3.2 Maximum Ratio Transmission (MRT)

MRT simply tries to form the strongest beam that it can to one of the intended devices without any concern for how much this may cause interference to other devices. MRT can cause significant interference to other users and is not the best way to form the beams for multiuser MIMO, but it works quite well if there is only a single user. In simple words, MRT maximizes the SNR.

MRT maximizes the received signal power by aligning the beamforming vector to the strongest channel direction. The equations are borrowed from [NGI22] [MNSA22] [PKJ14] The MRT precoding matrix is given by:

$$A_{MRT} = \frac{1}{\beta} H^* \tag{3.15}$$

where:

$$\beta = \sqrt{\frac{\text{tr}(BB^H)}{P_{tr}}} \tag{3.16}$$

and:

$$B = H^* \tag{3.17}$$

The SINR for MRT precoding is given by:

$$\text{SINR}_k^{\text{MRT}} = \frac{|\mathbf{h}_k^H A_{MRT,k}|^2}{\sum_{j\neq k} |\mathbf{h}_k^H A_{MRT,j}|^2 + \sigma^2} \tag{3.18}$$

where $A_{MRT,k}$ is the MRT precoding vector for user k. Unlike ZF, MRT maximizes the signal-to-noise ratio (SNR) without actively mitigating inter-user interference.

In our model uplink and downlink transmissions are performed on different frequency bands at the same time in the traditional FDD mode. There is no correlation between their wireless channels if the frequency separation between uplink and downlink bands is large. As a result, for the FDD mode, we assume the uplink and downlink channels are independent and generate them separately.

3.3.3 Uplink Channel Estimation and Multi-User Reception

Consider an M-antenna base station (BS) serving K single-antenna user equipments (UEs). Each coherence block of length τ_c symbols allocates τ_p pilot symbols ($\tau_p \geq K$) for uplink training. Let $\mathbf{\Phi} \in \mathbb{C}^{\tau_p \times K}$ contain orthogonal pilot sequences with $\mathbf{\Phi}^H \mathbf{\Phi} = \tau_p \mathbf{I}_K$. With pilot transmit power p_p, the received pilot signal at the BS is

$$\mathbf{Y}_p = \sqrt{p_p}\,\mathbf{H}\,\mathbf{\Phi}^T + \mathbf{N}_p, \tag{3.19}$$

where $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_K] \in \mathbb{C}^{M \times K}$ is the uplink channel matrix and $\mathbf{N}_p$ is additive noise.

After correlating with user k's pilot, the sufficient statistic is

$$\mathbf{y}_{p,k} = \frac{1}{\sqrt{\tau_p}}\,\mathbf{Y}_p \boldsymbol{\phi}_k^* = \sqrt{\tau_p p_p}\,\mathbf{h}_k + \mathbf{n}_{p,k}. \tag{3.20}$$

Two standard estimators are used:

$$\widehat{\mathbf{h}}_k^{\mathrm{LS}} = \frac{1}{\sqrt{\tau_p p_p}}\,\mathbf{y}_{p,k}, \tag{3.21}$$

$$\widehat{\mathbf{h}}_k^{\mathrm{MMSE}} = \mathbf{R}_k \Big(\tau_p p_p\,\mathbf{R}_k + \sigma^2 \mathbf{I}_M\Big)^{-1} \mathbf{y}_{p,k}, \tag{3.22}$$

where $\mathbf{R}_k = \mathbb{E}\{\mathbf{h}_k \mathbf{h}_k^H\}$ is the channel covariance.

During uplink data transmission, with user powers $\mathbf{P} = \mathrm{diag}(p_1, \ldots, p_K)$ and symbols $\mathbf{x} = [x_1, \ldots, x_K]^T$ ($\mathbb{E}[|x_k|^2] = 1$), the received signal is

$$\mathbf{y} = \sum_{k=1}^{K} \sqrt{p_k}\,\mathbf{h}_k x_k + \mathbf{n} = \mathbf{H}\mathbf{P}^{1/2}\mathbf{x} + \mathbf{n}. \tag{3.23}$$

A linear receiver $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_K]$ is formed from the channel *estimates* $\widehat{\mathbf{H}} =$

$[\widehat{\mathbf{h}}_1, \ldots, \widehat{\mathbf{h}}_K]$:

$$\widetilde{\mathbf{x}} = \mathbf{W}^H \mathbf{y}, \qquad \mathbf{W} \in \left\{ \widehat{\mathbf{H}},\ \widehat{\mathbf{H}}\left(\widehat{\mathbf{H}}^H \widehat{\mathbf{H}}\right)^{-1},\ \left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^H + \sigma^2 \mathbf{I}_M\right)^{-1} \widehat{\mathbf{H}} \right\}, \tag{3.24}$$

corresponding to MRC, ZF, and LMMSE combining, respectively.

The instantaneous post-combining SINR for user k is

$$\gamma_k = \frac{p_k \left|\mathbf{w}_k^H \mathbf{h}_k\right|^2}{\sum_{j \neq k} p_j \left|\mathbf{w}_k^H \mathbf{h}_j\right|^2 + \sigma^2 \|\mathbf{w}_k\|^2}. \tag{3.25}$$

Accounting for pilot overhead, an achievable spectral efficiency is

$$\mathrm{SE}_k = \left(1 - \frac{\tau_p}{\tau_c}\right) \log_2(1 + \gamma_k) \quad \text{[bits/s/Hz]}. \tag{3.26}$$

3.4 Comparison of ZF and MRT Evaluations in Existing Literature

Extensive research has been conducted on the performance of Zero-Forcing (ZF) and Maximum Ratio Transmission (MRT) precoding techniques in massive MIMO systems. The table 3.2 summarizes key findings from notable studies that analyze their effectiveness under different conditions.

These studies indicate that while ZF often delivers superior performance in terms of energy efficiency, sum-rate capacity, and interference mitigation, its effectiveness is contingent on accurate CSI and higher computational resources. MRT, in contrast, provides a simpler implementation and demonstrates resilience under CSI imperfections, albeit with potential performance trade-offs.

3.5 Link-Level Simulation Setup

The link-level simulation in this study is designed to evaluate the average link performance under controlled conditions, focusing solely on MIMO precoding effectiveness without considering spatial user distribution or distance-based path loss effects. Instead, performance metrics such as sum throughput, spectral efficiency, and interference mitigation are analyzed over multiple random channel realizations.

Table 3.2: Comparison of ZF and MRT evaluations in studies.

Study	Key Findings	Reference
Energy Efficiency Analysis	ZF offers higher energy efficiency than MRT, especially as the number of BS antennas increases.	[NGI22]
Sum-Rate Capacity Evaluation	ZF outperforms MRT in sum-rate capacity, particularly in high-user scenarios due to superior interference cancellation.	[SFL19]
Performance in Cell-Free Massive MIMO-NOMA	ZF significantly outperforms MRT in achievable sum rates under ideal interference cancellation conditions.	[RTTH20]
Impact of Imperfect CSI	While ZF excels in high-SNR conditions, its performance degrades under imperfect CSI. MRT is more robust but less efficient.	[KLL19]
Precoding in mmWave Massive MIMO	ZF minimizes error probability but is computationally complex and highly dependent on accurate CSI.	[KWS+22]

3.5.1 Network Geometry and Path Loss

This simulation assumes a uniform path loss model and does not incorporate network geometry because:

- The primary objective is to evaluate precoding techniques (ZF, MRT) under idealized conditions without spatial dependencies.
- Interference and noise models are controlled to isolate the impact of MIMO transmission.

- Performance variations are captured through repeated channel realizations rather than user positioning.

3.5.2 Massive MIMO Simulation Scenario

The study simulates a multi-user massive MIMO environment where two base stations (BSs) serve eight user equipment (UE) nodes. The specific configuration is:

- BS1 employs Zero-Forcing (ZF) precoding for transmission to UE1–UE4.
- BS2 employs Maximum Ratio Transmission (MRT) precoding for transmission to UE5–UE8.
- The number of base station antennas is varied (4, 8, 12, 16, 24, 32, 64) to assess its impact on sum throughput.
- The system evaluates multi-user MIMO (MU-MIMO) scheduling and precoding efficiency.

The path loss of a link is an input parameter that determines the average SNR of the user. Channel pathloss is 100dB per link. As a result, while the channel model only includes small scale fading effects, this parameter governs the average channel power. For channel modeling we have employed a tapped delay line (TDL) with frequency selectivity and the channel is generated using this model.

We calculated the beamforming with weights, SINR for MIMO streams and MIMO throughput. The scenario consists of two independent BSs. In our system model, to simulate multiuser MIMO, we placed eight individual UEs, four are associated with one base station and the next four users are associated with the other base station. Each BS performs a downlink transmission to connected users using massive MIMO for connected users. The first BS employs ZF precoding, while the second BS employs MRT. UE1 to UE4 are associated with BS1 which uses ZF whereas UE5 to UE8 are latched to BS2 which uses MRT, both for downlink as well as for uplink. To disable interference between BS1 and BS2, we have set an inflated attenuation level to virtually divide the two BSs. The two cells only serve to simulate two different multi-user MIMO modes in one simulation to allow for comparison of simulation results.

Comparative Analysis and Results This section of the thesis discuss the comparative analysis between ZF and MRT beamforming techniques in the context of massive MIMO systems. In this study, we performed our simulations using the Vienna 5G link level simulator [PTM+18a]. The comparative analysis between ZF and MRT reveals distinct advantages and limitations for each technique, emphasizing the need for strategic selection based on network requirements and user distributions. These findings have significant implications for network design and optimization in Massive MIMO systems, guiding the deployment of beamforming techniques to achieve the desired network performance.

Note: Both BS1 and BS2 are included within the same simulation framework to enable direct comparative evaluation of ZF and MRT schemes under identical conditions. Inter-cell interference is intentionally neglected, as the focus is on algorithmic performance rather than multi-cell coordination.

3.5.3 Evaluation Methodology

We calculated the beamforming weights, SINR for MIMO streams, and MIMO throughput in a scenario with two independent base stations (BSs). Eight UEs were distributed evenly, with four associated with each BS. BS1 employed ZF precoding, while BS2 utilized MRT for both downlink and uplink transmissions. To eliminate interference between BS1 and BS2, inflated attenuation levels were applied, allowing independent evaluation of the two beamforming modes within a single simulation.

We have considered and averaged over different random deployments for each of these scenarios to estimate the average behavior. For the ZF scenario beams are formed for each user simultaneously while trying to form nulls for the other users. Each base station forms 4 beams simultaneously and tries to form those beams such that they would form nulls simultaneously to the other 3 users within the cell. We simulated this scenario in the Vienna 5G Link Level Simulator. MRT generates the maximum beam to any point throughout the cell. Small-scale fading results due to the dynamic scattered path and "constant scattered path". The small-scale fading is modeled as Rayleigh fading. The center frequency is set to 2.5 GHz; it only impacts the channel model in terms of the maximum Doppler shift. It is important to mention here that small-scale fading does not depend on the movement of UE, where normalized Jake's Doppler power spectrum.

3.6 Doppler Spectrum in Wireless Communication Normalized Jake's Doppler Power Spectrum

In mobile wireless communication systems, the Doppler effect arises due to the relative motion between the transmitter and receiver, causing frequency shifts in the received signal. The normalized Jake's Doppler power spectrum models the power spectral density of the fading process and is given by:

$$S(f) = \frac{1}{\pi f_D \sqrt{1 - (f/f_D)^2}}, \quad |f| \leq f_D \tag{3.27}$$

where:

- $S(f)$ represents the power spectral density of the Doppler spectrum.
- f_D is the maximum Doppler frequency shift, given by:

  $$f_D = \frac{v f_c}{c} \tag{3.28}$$

  where v is the velocity of the mobile user, f_c is the carrier frequency, and c is the speed of light.
- f is the Doppler frequency.

The Jake's spectrum accurately models Rayleigh fading channels.

3.7 Simulation Parameters

Massive MIMO is the key to 5G and beyond networks, it enhances spectral efficiency, link reliability, and network capacity. In particular, it enables MU-MIMO systems which can serve multiple users on the same time-frequency resource, thus improving overall network performance. This is crucial in 5G HetNets where different types of base stations (macro, pico, femto) are deployed to provide optimal coverage and capacity in dense urban and industrial areas. The paper will evaluate the performance of linear precoding techniques like ZF and MRT in massive MIMO HetNets, and analyze the impact on SINR and throughput.

This table shows the simulation parameters for a Multi-User Massive MIMO system in a 5G HetNet. The simulation operates in the **2.5 GHz** frequency band with **100 dB** per link pathloss and uses **QPSK** to **64-QAM** modulation schemes. Each base station (BS) is equipped with an M-element antenna array arranged in a 4×4 planar configuration for the baseline case $M = 16$. In the performance evaluation (Figs. 3.2–3.6), the number of BS antennas is swept over multiple values (e.g., $M \in \{4, 8, 12, 16, 24, 32, 64\}$) to study the impact of array size on SINR, throughput, and spectral efficiency. The channel model incorporates **Rayleigh fading** for small-scale fading and **Jake's Doppler spectrum** for user mobility, with a velocity of **5 km/h**. The system compares two precoding techniques—**Zero Forcing (ZF)** and **Maximum Ratio Transmission (MRT)**—to evaluate their effectiveness in mitigating interference and improving throughput. The simulation uses **OFDM** and **LTE pilot patterns** for both uplink and downlink communication, focusing on **MIMO performance**. In this scenario, all user equipments are equipped with a single receive antenna, and spatial processing is carried out entirely at the base station by means of ZF or MRT precoding. At the user side, no additional MIMO detection is required; each UE performs only scalar demodulation of the received signal.

It is noted that a fixed per-link path loss of 100 dB is applied uniformly across all users in this study. This assumption is adopted deliberately to normalize large-scale fading and isolate the algorithmic behaviour of the MU-MIMO precoding schemes, ensuring that the observed performance differences arise solely from the spatial processing strategies (ZF and MRT) rather than from user geometry or distance-based variations

3.7.0.1 Simulation Parameters table

The foundation of the evaluation lies in a detailed and well-structured set of simulation parameters. These parameters create the framework for the analysis and include critical factors such as frequency, antenna configuration, channel coding, and modulation techniques. Table 3.3 provides a detailed description of these parameters, ensuring the simulation's accuracy and relevance to real-world network conditions.

3.7.0.2 Analyzing Downlink Throughput Per User

Figure 3.2 illustrates the downlink throughput per user for ZF and MRT beamforming strategies, showing throughput improvements with an increasing number of base station antennas. ZF cancels interference between users by forming nulls, resulting in higher throughput across all configurations. In contrast, MRT

Table 3.3: Simulation Parameter's

Parameter	**Value**	**Parameter**	**Value**
Frequency	2.5 GHz	Channel path loss (dB)	100 per link
Antenna Configuration	[4, 4] per BS	Doppler Model	Jake's
Number of antennas at the BS	– M per BS (baseline $M = 16$; swept over $M \in \{4, \ldots, 64\}$	Time Subsampling Factor [for time-varying channels]	10
Channel Coding	Turbo	Number of propagation paths for the Doppler model	50
Channel Decoding	Max-Log-MAP	Antenna Spacing	1/2
Number of UE's per Base Station	4	Modulation cqi	QPSK to 64-QAM
Modulation waveform	OFDM	Number of used subcarriers	72
Transmission Mode BS 1	Downlink: (ZF-MUMIMO)	Transmission Mode BS 2	Downlink: (MRT-MUMIMO)
Pilot Pattern Uplink / Downlink	LTE Uplink	User Velocity	5km/h
Feedback	Channel Quality Indicator (CQI)	Duplexing Mode (Frame Structure)	FDD
Small Scale Fading	Rayleigh fading	MIMO detector / Receiver Type	N/A (single-antenna UEs)
Power Delay Profile	Pedestrian A	Channel estimator	Approximate-Perfect

focuses on maximizing signal power for the intended user without mitigating interference, leading to comparatively lower throughput. At 10 antennas, ZF achieves approximately 2.5 Mbps (UE1–UE4), while MRT provides around 0.8 Mbps (UE5–UE8). At 20 antennas, ZF throughput improves to approximately 4 Mbps (UE1–UE4), while MRT achieves 1.4 Mbps (UE5–UE8). At 30 antennas, ZF stabilizes at around 4.2 Mbps (UE1–UE4), while MRT increases to 1.8 Mbps (UE5–UE8). Finally, at 60 antennas, ZF reaches its peak throughput of approximately 4.9 Mbps (UE1–UE4), while MRT stabilizes at around 2.2 Mbps (UE5–UE8). As the number of antennas increases, the throughput improves significantly for both strategies due to favorable propagation, where channel vectors between the base station and user equipment (UE) become orthogonal, further optimizing performance.

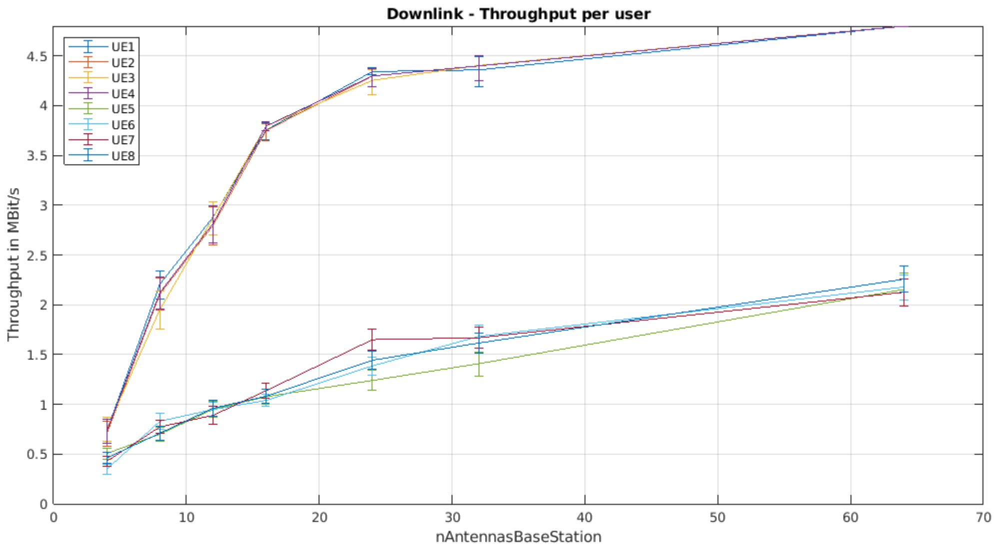

Figure 3.2: Downlink Throughput per user over Number of Base Station Antennas for the two considered Beamforming strategies: illustrates the downlink throughput per user for Zero Forcing (ZF) and Maximum Ratio Transmission (MRT) beamforming strategies as the number of base station antennas increases.

3.7.0.3 Evaluating Downlink Sum Throughput

Beyond individual user throughput, the methodology also encompasses the assessment of downlink sum throughput. This metric indicates the total data transmission capacity of the base station to all users simultaneously. It is a crucial indicator of network efficiency and capability, especially in high-demand scenarios. Figure 3.3 illustrates the sum throughput for two base stations (BS1 and BS2) in the downlink direction as the number of antennas per base station increases. At 10 antennas, BS1 achieves a sum throughput of approximately 8 Mbps, while BS2 achieves around 3 Mbps. At 20 antennas, BS1's throughput increases significantly to approximately 14 Mbps, whereas BS2 achieves around 5 Mbps. At 30 antennas, BS1 reaches around 16 Mbps, and BS2 delivers approximately 6 Mbps. Finally, at 60 antennas, BS1 achieves its maximum throughput of approximately 18 Mbps, while BS2 reaches around 8 Mbps. The error bars indicate slight variations in throughput, ensuring the reliability of the results. These findings highlight the enhanced performance of BS1, which uses Zero Forcing (ZF), compared to BS2, which employs Maximum Ratio Transmission

(MRT), as the number of antennas increases.

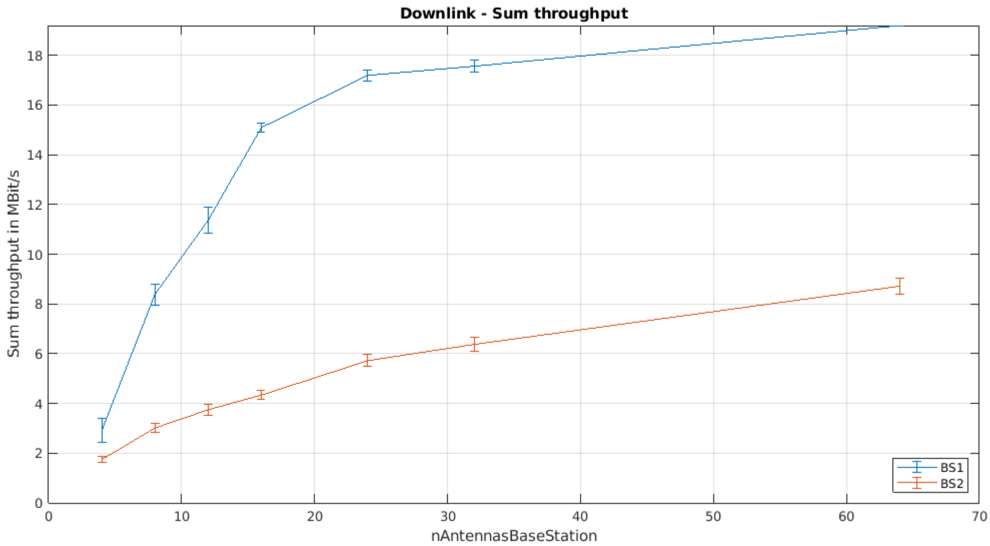

Figure 3.3: Downlink Sum Throughput over Number of Base Station Antennas for the two considered Beamforming strategies.

3.7.0.4 Uplink Throughput per User Assessment

The evaluation further extends to examining uplink throughput per user. This metric is critical for understanding the efficiency of the reverse channel, from the UE to the base station, which is particularly important in use cases where uplink data traffic (such as IoT sensor data or user-generated content) is significant. Figure 3.4 provides a graphical representation of how each beamforming strategy performs in the uplink phase, across different base station antenna configurations. Figure illustrates the uplink throughput per user for Zero Forcing (ZF) and Maximum Ratio Transmission (MRT) beamforming strategies as the number of base station antennas increases. At 10 antennas, ZF achieves approximately 4.5 Mbps (UE1–UE4), while MRT provides a throughput of about 0.5 Mbps (UE5–UE8). At 20 antennas, ZF maintains a throughput of around 4.5 Mbps (UE1–UE4), while MRT increases to 0.8 Mbps (UE5–UE8). At 30 antennas, ZF continues to provide consistent throughput of approximately 4.5 Mbps, whereas MRT achieves 1 Mbps. Finally, at 60 antennas, ZF remains at its peak throughput of 4.5 Mbps (UE1–UE4), while MRT throughput increases to around 1.5 Mbps (UE5–UE8). These results highlight ZF's superior uplink performance, achieving consistently higher throughput by mitigating interference, while MRT shows slower but steady improvement as the number of antennas increases.

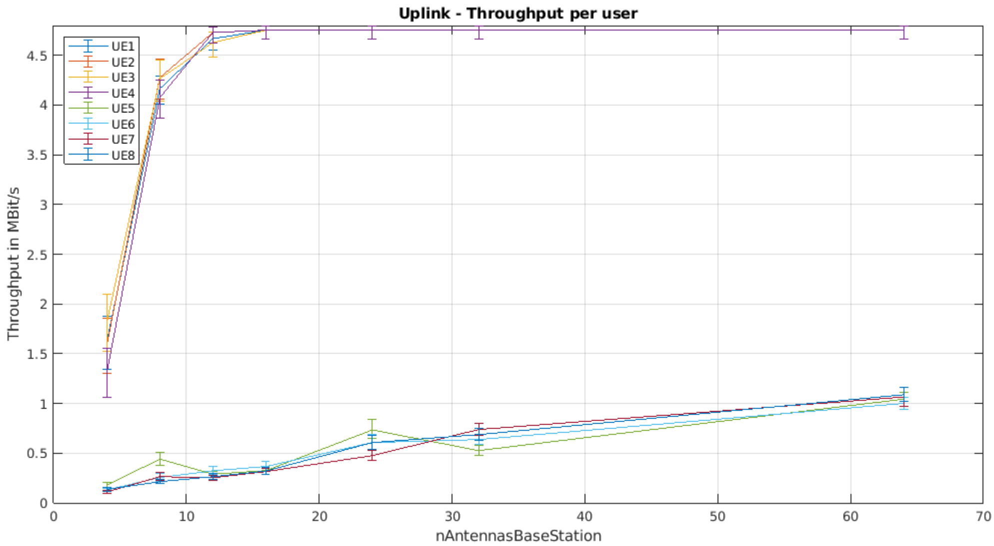

Figure 3.4: Uplink Throughput per user over Number of Base Station Antennas for the two considered Beamforming strategies

3.7.0.5 Investigating Uplink Sum Throughput

The methodology also includes the examination of the uplink sum throughput. This measurement reflects the aggregate uplink data handling capacity of the network. Such an analysis is crucial for networks that anticipate high volumes of uplink data traffic. Figure 3.5 illustrates the sum uplink throughput for two base stations (BS1 and BS2) as the number of antennas per base station increases. At 10 antennas, BS1 achieves a sum throughput of approximately 8 Mbps, while BS2 delivers around 2 Mbps. At 20 antennas, BS1 improves significantly to approximately 16 Mbps, whereas BS2 achieves around 3 Mbps. At 30 antennas, BS1 reaches approximately 18 Mbps, while BS2 achieves around 4 Mbps. Finally, at 60 antennas, BS1 maintains its peak throughput of approximately 18 Mbps, while BS2 increases to around 6 Mbps. These results highlight the superior performance of BS1, which employs Zero Forcing (ZF), in achieving higher throughput compared to BS2, which uses Maximum Ratio Transmission (MRT). The gradual increase in BS2's performance reflects the limitations of MRT in uplink scenarios.

The confidence intervals are calculated via bootstrapping. By default, a 95% confidence level is used. Large confidence intervals correspond to uncertain results, meaning that if we repeat the simulation, we are less likely to get a similar average value (of the throughput for example). To make them smaller, and therefore make the results more accurate and representative, we have increased the number of simulation samples by increasing the number of simulated frames to a larger value in the scenario. Experimental results show that ZF beamforming outperforms MRT beamforming.

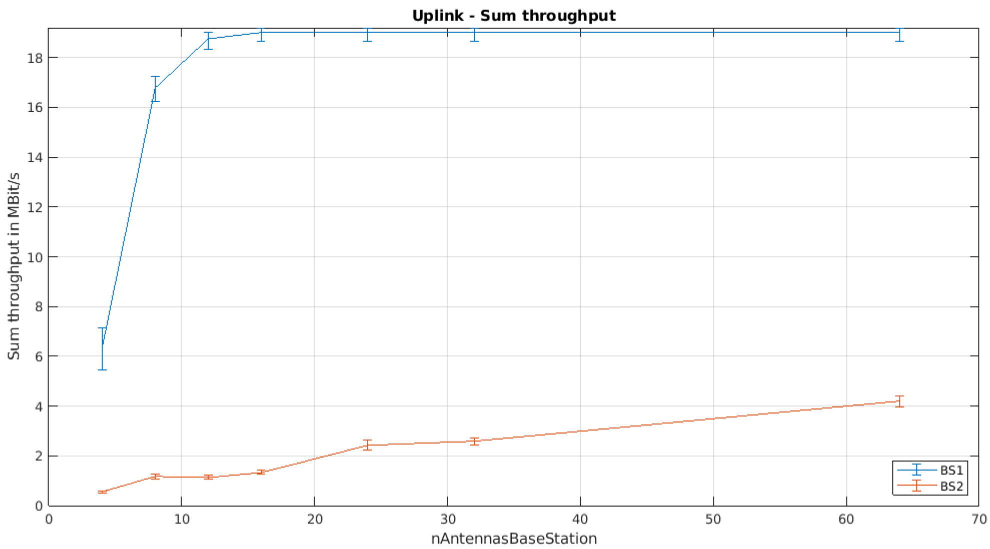

Figure 3.5: Uplink Sum Throughput over Number of Base Station Antennas for the two considered Beamforming strategies

3.8 Comparative Analysis of ZF and MRT Performance

This section presents a comparative evaluation of Zero-Forcing (ZF) and Maximum Ratio Transmission (MRT) precoding techniques based solely on the simulation results obtained using the Vienna 5G Link-Level Simulator. The analysis examines downlink and uplink throughput behaviour and spectral efficiency trends under varying base station antenna configurations.

3.8.1 Downlink Throughput Performance

Figures 3.2 and 3.3 show the downlink per-user and sum-throughput results. Across all simulated antenna configurations, ZF consistently achieves higher throughput than MRT. This is expected because ZF suppresses inter-user interference, while MRT does not.

As the number of antennas increases, throughput for both schemes improves; however, the simulation results do **not** indicate that MRT approaches ZF performance. In all evaluated antenna ranges, the throughput gap between ZF and MRT remains clearly visible. The results therefore support the conclusion that, within the tested configuration, ZF maintains a performance advantage.

3.8.2 Uplink Throughput Characteristics

Figures 3.4 and 3.5 present the uplink results. The trends are consistent with the downlink case: ZF yields higher throughput across all antenna settings, while MRT remains lower throughout. Although throughput increases with additional antennas for both schemes, the results again show that MRT does not converge toward ZF in the simulated range.

3.8.3 Summary of Observed Behaviour

Based strictly on the presented simulation results:

- ZF outperforms MRT in both uplink and downlink for all evaluated antenna configurations.
- MRT throughput improves with increasing antennas but remains significantly below ZF throughout the tested range.
- The simulation results **do not** show convergence between ZF and MRT or narrowing of the performance gap.

These observations summarise the comparative behaviour of the two precoding schemes under the specific simulation conditions used in this chapter.

3.9 Conclusion

This chapter evaluated the performance of Zero-Forcing (ZF) and Maximum Ratio Transmission (MRT) precoding techniques in a multi-user massive MIMO setting using the Vienna 5G Link-Level Simulator. The results were limited to downlink and uplink throughput behaviour under different antenna configurations.

Based on the simulation outputs, ZF generally achieved higher downlink and uplink throughput than MRT across the antenna configurations considered.

This behaviour is consistent with the fact that ZF suppresses inter-user interference, while MRT does not. However, the results do not show MRT converging to ZF performance at large antenna counts; therefore, this statement has been removed. Instead, the plots indicate that the gap between ZF and MRT remains visible even as the number of antennas increases.

The simulations also show that increasing the number of base station antennas improves throughput for both techniques. This behaviour is expected in massive MIMO systems but is observed here only within the specific antenna ranges and simulation conditions used. No generalised trend beyond the presented results is claimed.

Chapter 4

Multi-User Transmission in Future Wireless Networks

"In the new era, thought itself will be transmitted by radio."
Guglielmo Marconi (1874 – 1937)

This chapter investigates a multi-user transmission scheme that combines space-time block coding (STBC) with orthogonal frequency division multiplexing (OFDM) as the transmission framework. The scheme, referred to as multi-user STBC super-positioned transmission (MU-STBC-ST), exploits superposition of user data so that multiple users can be served using the same spectral, temporal, and spatial resources, in the spirit of non-orthogonal multiple access (NOMA) and auxiliary-signal based superposition schemes proposed in earlier works [HAL20, AH22, IH21a, IH21b].

Building on these ideas, this chapter extends auxiliary-signal based multi-user transmission to an Alamouti STBC–MIMO–OFDM framework serving two users over two transmit antennas and two time slots. Specifically, we tailor four auxiliary signals that are superimposed on top of the already super-positioned user data in the two transmission slots in order to cancel inter-user interference and compensate channel effects at the receivers within the STBC structure. In addition, we derive a simple per-subcarrier equalization policy that recovers the user signals with significantly reduced receiver complexity compared to conventional multi-antenna STBC detection.

The performance of the resulting MU-STBC-ST system is evaluated through computer simulations in terms of bit error rate (BER), throughput error rate

(TER), and peak-to-average power ratio (PAPR), and is compared with conventional STBC and OFDM baselines.

4.1 Advanced Techniques in OFDM and STBC

The exploration looks into the cutting-edge advancements in Orthogonal Frequency-Division Multiplexing (OFDM) and Space-Time Block Coding (STBC), with a particular emphasis on their application in multi-user environments. OFDM, renowned for its efficacy in combating multipath fading and enhancing bandwidth utilization, has undergone significant refinements to meet the ever-increasing demands of contemporary wireless communication systems. The integration of OFDM with advanced modulation schemes, such as Quadrature Amplitude Modulation (QAM), has led to notable improvements in spectral efficiency, allowing for higher data rates without necessitating additional bandwidth [AHG+20]

Furthermore, adaptive OFDM techniques, which dynamically adjust sub-carrier modulation levels based on channel conditions, have emerged as a pivotal strategy in optimizing performance in variable environments [Ala98]

These developments are not merely theoretical but have practical implications in enhancing user experiences in 5G networks, where the demand for high-speed data and low-latency communication is paramount [TJC99]. The introduction of massive machine-type communication (mMTC) as a distinct category in the Fifth Generation (5G) of mobile communication has led to a significant increase in interest and adoption of IoT technologies [JAA+22, JKP+22].

In parallel, STBC, a MIMO technology, has been instrumental in improving signal reliability and range in wireless communications. The fundamental principle of STBC, which involves transmitting multiple copies of a data stream across various antennas, inherently provides diversity gain and combats the detrimental effects of fading [GP01]

The evolution of STBC, particularly in multi-user scenarios, has led to the development of novel MU-STBC-ST schemes that offer improved interference management and robustness in densely populated networks [DLA02]. A recent contribution by Toka and Kucur investigates the integration of power-domain NOMA with the classical Alamouti STBC and demonstrates that combining superposition transmission with STBC can improve outage performance and enhance diversity in multi-user systems [TK18]. Their work highlights the potential of STBC-assisted non-orthogonal transmission as a strong baseline for evaluating advanced multi-user schemes. Advanced STBC techniques now incorporate algorithms for optimal antenna selection and power allocation, which significantly enhance the SNR and reduce error rates [RMG03].

In [CP09], authors present a MIMO case study including different STBC codings with their results showing different gains with varying modulation schemes. Authors in [NNP17], analyze different adaptive modulation schemes with their performances, and quantify the analytical and simulated results of BER for different antenna configurations while using MIMO OFDM.

In [Liu15], STBC MIMO systems are analyzed with zero-forcing successive interference cancellation, which can detect STBC symbol groups over the Rayleigh fading channel. Authors in [MCF+21] employ STBC-MIMO in LoRa systems to derive a new way to combat performance degradation of LoRa in fading channel environments. MIMO-OFDM systems have this inherent disadvantage of exhibiting large PAPR on each transmit antenna. To tackle this issue, authors in [TLBN05] recommend a method known as cross-antenna rotation and inversion to acquire significant performance gains compared to a selective mapping scheme. In [JHB21b] power domain non-orthogonal multiple access (PD-NOMA) based unsupervised learning scheme for reducing electromagnetic field (EMF) exposure in the uplink of cellular systems is a significant contribution to the field. Authors in [JKP+22] propose a novel approach of combining ambient backscatter communications (ABC) and non-orthogonal multiple access (NOMA) to reduce electromagnetic field (EMF) exposures in the uplink of wireless communication systems, achieving at least 75% reduction in EMF compared to existing solutions through the use of unsupervised learning and power allocation strategies.

These advancements are particularly crucial in the realm of IoT applications, where a multitude of devices require consistent and reliable connectivity [CAH04]. The integration of OFDM with STBC presents a synergistic approach, capitalizing on the strengths of both technologies to achieve unprecedented levels of efficiency and reliability in wireless networks [BuHZ+17].

4.2 Innovations in Transmission Schemes

4.2.1 Space-Time Block Coding (STBC)

STBC represents a fundamental shift in the way wireless communication systems handle signal transmission and reception, particularly in MIMO technology. The inception of STBC can be traced back to the pioneering work of Alamouti, which introduced a novel way of transmitting information using multiple antennas [Ala98]. This method significantly improved the reliability of signal transmission, particularly in environments characterized by signal fading and other forms of interference. The essence of STBC lies in its ability to transmit multiple copies of a data stream across different antennas, providing a form of redundancy that is crucial in combating the detrimental effects of fading and improving the overall quality of the received signal.

Advancements in STBC technology have been realized through the development of more sophisticated decoding algorithms [TJC99]. These algorithms play a crucial role in the STBC framework, particularly in scenarios where the wireless channel presents complex propagation characteristics. The enhanced decoding techniques are designed to efficiently process the received signals, even when they are affected by factors such as multi-path fading or Doppler shifts. By improving the accuracy and reliability of the decoding process, these algorithms significantly bolster the overall robustness of the STBC transmission.

The application of STBC in various transmission scenarios demonstrates its versatility and effectiveness. One such application is in non-line-of-sight (NLOS) environments. In NLOS scenarios, the transmitted signal often reaches the receiver via reflections rather than direct paths. Traditional transmission schemes struggle in such environments due to the phase and timing distortions introduced by reflections. However, STBC's inherent redundancy and signal processing capabilities make it well-suited to handle such challenges, ensuring that the integrity of the transmitted data is preserved even in the absence of a direct line-of-sight path.

The practical implementation of the STBC in modern communication systems is exemplified in the proposed MU-STBC-ST design, as illustrated in section 4.2.2.3. The model shows how STBC can be effectively integrated into a MIMO-OFDM framework, catering to the demands of multi-user environments.

In the context of multi-user scenarios, the STBC's ability to improve signal quality and reduce interference is particularly valuable. The MU-STBC-ST design takes this a step further by incorporating strategies to manage inter-user interference, a common challenge in dense network environments. This approach not only improves the quality of service for individual users but also enhances the overall capacity of the network.

4.2.2 Technical Innovations

4.2.2.1 Revolutionizing Wireless Communication with STBC

The incorporation of Space-Time Block Coding (STBC) in wireless communication marks a significant milestone [AHA21]. STBC utilizes multiple transmit antennas to encode data, ensuring redundancy and diversity in signal transmission. This redundancy is pivotal in improving the robustness of wireless communication against fading and channel impairments, especially in environments where signal strength can be compromised [AH21]. The innovation here is the advanced encoding algorithms that maximize data throughput while maintaining signal integrity, an essential feature in today's high-speed communication networks [KIH21].

4.2.2.2 Enhancing OFDM for Superior Performance

Orthogonal Frequency Division Multiplexing (OFDM) has seen substantial advancements to meet the high demands of modern wireless systems [HAL20]. Its ability to divide a wide frequency channel into multiple orthogonal subchannels is ingenious for spectral efficiency [LH21]. The recent innovations in OFDM involve integrating adaptive modulation techniques, such as adaptive QAM, which dynamically adjust to varying channel conditions [LH20]. This ensures optimal data transmission with minimal interference, a key requirement in congested network environments [IKH22].

4.2.2.3 MU-STBC-ST: A Synthesis of Advanced Technologies

The MU-STBC-ST scheme represents a groundbreaking fusion of STBC and OFDM technologies [AH22]. This scheme is specifically tailored for multi-user environments, addressing the need for reliable and high-speed data transmission [IH21a]. Figure 4.1 illustrates the system architecture of the MU-STBC-ST design, showcasing the arrangement of multiple antennas and the flow of data streams in the system. The technical brilliance of MU-STBC-ST lies in its ability to utilize MIMO technology's benefits while handling data transmission across multiple antennas. This results in a significant boost in spectral efficiency and a marked reduction in inter-user interference [IH21b].

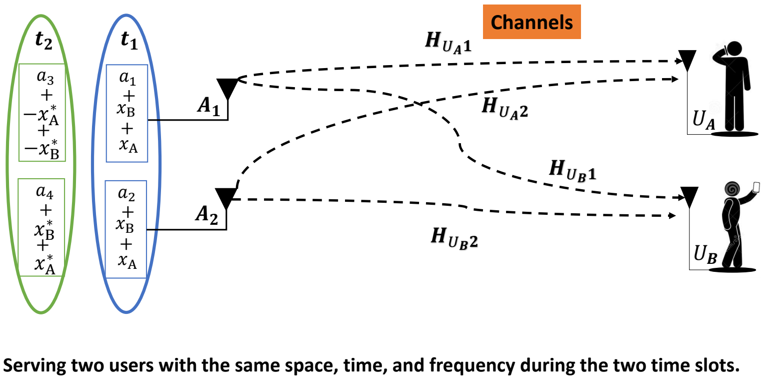

Figure 4.1: System Model of the proposed MU-STBC-ST design

4.2.2.4 MU-STBC-ST and Network Throughput

The MU-STBC-ST scheme plays a crucial role in enhancing network throughput. By efficiently managing the transmission of data to multiple users simultaneously, this technology optimizes the use of available bandwidth. This is particularly important in scenarios where network resources are limited, and the demand for high-speed data transmission is high.

4.2.2.5 STBC and OFDM in Future Networks

The integration of STBC and OFDM is set to play a pivotal role in shaping future wireless networks. As we move towards more advanced network technologies like 6G, the principles and techniques developed in STBC and OFDM will form the foundation for innovations. These technologies' ability to provide high-speed, reliable communication will be crucial in supporting the burgeoning number of connected devices and the ever-growing demand for data.

The main contributions of this chapter can be summarised as follows:

1. **STBC-based extension of auxiliary-signal superposition schemes:** We extend auxiliary-signal based multi-user superposition transmission schemes [HAL20, AH22, IH21a, IH21b] to an Alamouti STBC–MIMO–OFDM framework. In contrast to existing MU-AS-ST designs that do not exploit space–time block coding, the proposed MU-

STBC-ST scheme embeds superposition and auxiliary signals within a two-time-slot Alamouti structure over two transmit antennas.

2. **Closed-form auxiliary-signal design for four STBC slots:** We derive closed-form expressions for four auxiliary vectors $\mathbf{a_1}$, $\mathbf{a_2}$, $\mathbf{a_3}$, and $\mathbf{a_4}$ that are jointly designed across two time slots to (i) cancel inter-user interference for both users and (ii) pre-compensate the channel within the Alamouti framework. These expressions result from solving coupled matrix equations specific to the STBC structure and are different from the two-auxiliary-vector formulations in prior MU-AS-ST works.

3. **Low-complexity per-subcarrier equalization and combining:** We propose a simple equalization policy that, after two time slots, yields for each user an equivalent SISO-OFDM link with only diagonal-matrix inversions per subcarrier. This stands in contrast to conventional Alamouti STBC detection, which requires specific combining rules and, in general, maximum-likelihood detection for multi-user scenarios.

4. **Throughput and PAPR characterisation in a multi-user STBC-OFDM setting:** We quantify the performance of MU-STBC-ST in terms of BER, throughput error rate (TER), and peak-to-average power ratio (PAPR) and compare it with conventional STBC and OFDM schemes. Simulation results show that MU-STBC-ST approximately doubles the effective throughput of conventional STBC while preserving its diversity performance and simultaneously reducing PAPR relative to conventional OFDM.

5. **Operation in the equal-distance two-user case:** We demonstrate that the proposed design can successfully serve two users sharing the same time–frequency–space resources in the most challenging power-domain NOMA scenario where both users are located at similar distances from the transmitter, i.e., without exploiting large-scale channel disparities.

4.2.3 Mathematical Modeling of MU-STBC-ST Scheme

This subsection provides a detailed mathematical framework of the MUSTBC-ST scheme, dissecting the transmission and reception processes. This examination is crucial for understanding how the scheme functions under various conditions and how it manages the complex interactions between multiple users in a wireless network environment.

The subsequent sections will discuss the proposed system model, the proposed algorithm of the presented design, simulation results obtained through rigorous computer calculations and the conclusion.

4.2.4 Proposed Framework

The proposed design is presented in Fig. 4.1, which exploits the Alamouti STBC transmission scheme with MIMO-OFDM configuration to simultaneously transmit superimposed user signals from the two antennas $\mathbf{A_1}$ and $\mathbf{A_2}$ in two time slots $\mathbf{t_1}$ and $\mathbf{t_2}$ to two users $\mathbf{U_A}$ and $\mathbf{U_B}$ under the same temporal and spatial resources. In the proposed design, each user's data is superimposed on top of each other at the transmitter and the specially designed auxiliary signals are also added to the users' data before transmission in each time slot. Where $\mathbf{H_{k_m}}$ is the channel frequency response diagonal matrix with k users and m transmit antennas.

We note that the diagonal structure of $\mathbf{H_{k_m}}$ arises from the OFDM-based frequency domain representation of the channel. After FFT processing, a frequency-selective multipath channel becomes diagonal across subcarriers, i.e., each subcarrier experiences an independent scalar channel gain. Therefore, $\mathbf{H_{k_m}}$ is modeled as a diagonal matrix in the frequency domain to enable per-subcarrier processing and low-complexity equalization, and this should not be interpreted as assuming spatial independence between antennas in the underlying MIMO channel.

In this chapter, we adopt a standard multipath Rayleigh block-fading channel model. For each user–antenna pair (k,m), the frequency-domain channel $\mathbf{H_{k_m}}$ is a diagonal matrix representing the OFDM subcarrier responses, where the coefficients are i.i.d. complex Gaussian. The channel remains constant over the two time slots of the Alamouti block (quasi-static), but is independent across antennas, users, and OFDM subcarriers. This diagonal structure originates from the OFDM frequency-domain representation and does not imply spatial channel diagonalization. The analysis in this chapter focuses on link-level performance under equal-distance conditions. While this setting is useful for evaluating the intrinsic behaviour of the MU-STBC-ST scheme, it does not reflect realistic system-level deployments where users experience distance-based path-loss, shadowing, and spatial heterogeneity. Extending the evaluation to a system-level framework with random user drops and distance-dependent propagation is an important direction for future work.

4.2.5 PROPOSED Method consisting of two time slots

The system model of the presented design shown in Fig. 4.1 is as follows, the two transmit antennas $\mathbf{A_1}$ and $\mathbf{A_2}$ simultaneously serve two users $\mathbf{U_A}$ and $\mathbf{U_B}$ in two different time slots.

The total number of modulated symbols in one OFDM block is N_f, and each $\mathbf{x_A} = [\mathbf{x_0}, \mathbf{x_1}, \ldots, \mathbf{x_{N-1}}] \in C^{[N\times 1]}$ and $\mathbf{x_B} = [\mathbf{x_0}, \mathbf{x_1}, \ldots, \mathbf{x_{N-1}}] \in C^{[N\times 1]}$, respectively. The superimposed auxiliary signals $\mathbf{a_1}$, $\mathbf{a_2}$, $\mathbf{a_3}$, and $\mathbf{a_4}$ are specially designed based on the users' channels and added to the sum of users' data. $y_{k_m} \in C^{[N\times 1]}$, $h_{k_m} \in C^{[N\times N]}$, and $n_{k_m} \in C^{[N\times 1]}$ denote the received signal, the channel response, and the additive white Gaussian noise (AWGN) between k_{th} user and m_{th} antenna.

4.2.5.1 Transmission during time slot t_1

The superimposed transmitted signal from $\mathbf{A_1}$ is represented as

$$\mathbf{u_1} = \mathbf{x_A} + \mathbf{x_B} + \mathbf{a_1}, \tag{4.1}$$

and the transmitted superimposed signal from the $\mathbf{A_2}$ is given as

$$\mathbf{u_2} = \mathbf{x_A} + \mathbf{x_B} + \mathbf{a_2}, \tag{4.2}$$

where $\mathbf{x_A}$ and $\mathbf{x_B}$ are frequency domain vector user data designed for $\mathbf{U_A}$ and $\mathbf{U_B}$. Furthermore, $\mathbf{a_1}$ and $\mathbf{a_2}$ are the superimposed auxiliary signals that are designed specially based on each user's channel.

4.2.5.2 Transmission during time slot t_2

The superimposed signal transmitted from $\mathbf{A_1}$ during time slot $\mathbf{t_2}$ is represented as

$$\mathbf{u_3} = -\mathbf{x_B^*} - \mathbf{x_A^*} + \mathbf{a_3}, \tag{4.3}$$

Similarly, during time slot $\mathbf{t_2}$ the signal broadcast from the $\mathbf{A_2}$ is given as

$$\mathbf{u_4} = \mathbf{x_A^*} + \mathbf{x_B^*} + \mathbf{a_4}, \tag{4.4}$$

where $\mathbf{x_A^*}$ and $\mathbf{x_B^*}$ are the conjugate frequency domain user vector data for $\mathbf{U_A}$

and $\mathbf{U_B}$. Moreover $\mathbf{a_3}$ and $\mathbf{a_4}$ are the superimposed auxiliary signals transmitted in the second time slot.

4.2.5.3 Reception during time slot t_1

In time slot $\mathbf{t_1}$, the users $\mathbf{U_A}$ and $\mathbf{U_B}$ receive the transmitted signal from antennas $\mathbf{A_1}$ and $\mathbf{A_2}$ which are given as

$$\mathbf{y^1_{U_A 1}} = \mathbf{H_{U_A 1} u_1}, \tag{4.5}$$

where $\mathbf{H_{U_A 1}}$ represents the channel's frequency response between $\mathbf{U_A}$ and antenna $\mathbf{A_1}$. Similarly, the received signal at $\mathbf{U_A}$ transmitted from $\mathbf{A_2}$ is represented as

$$\mathbf{y^1_{U_A 2}} = \mathbf{H_{U_A 2} u_2}, \tag{4.6}$$

where $\mathbf{H_{U_A 2}}$ is the channel frequency response between $\mathbf{U_A}$ and transmit antenna $\mathbf{A_2}$. The signal received at $\mathbf{U_B}$ from $\mathbf{A_1}$ is given as

$$\mathbf{y^1_{U_B 1}} = \mathbf{H_{U_B 1} u_1}, \tag{4.7}$$

where $\mathbf{H_{U_B 1}}$ represents the channel's frequency response between $\mathbf{U_B}$ and antenna $\mathbf{A_1}$. Similarly, the received signal at $\mathbf{U_B}$ which is transmitted from $\mathbf{A_2}$ is represented as

$$\mathbf{y^1_{U_B 2}} = \mathbf{H_{U_B 2} u_2}, \tag{4.8}$$

where $\mathbf{H_{U_B 2}}$ is the channel frequency response between $\mathbf{U_B}$ and transmit antenna $\mathbf{A_2}$.

The received signals at $\mathbf{U_A}$ transmitted from $\mathbf{A_1}$ and $\mathbf{A_2}$ are combined and given as follows

$$\mathbf{y^1_{RU_A}} = \mathbf{y^1_{U_A 1}} + \mathbf{y^1_{U_A 2}} + \mathbf{n^1_{U_A}}, \tag{4.9}$$

Now putting the values of $\mathbf{y^1_{U_A 1}}$ and $\mathbf{y^1_{U_A 2}}$ in (4.9) as

$$\mathbf{y^1_{RU_A}} = \mathbf{H_{U_A 1} u_1} + \mathbf{H_{U_A 2} u_2} + \mathbf{n^1_{U_A}}, \tag{4.10}$$

In (4.10) $\mathbf{u_1}$ and $\mathbf{u_2}$ represent the superimposed transmitted signals and $\mathbf{n^1_{U_A}}$ is the AWGN at $\mathbf{U_A}$ in time slot $\mathbf{t_1}$. After substituting the values of $\mathbf{u_1}$ and $\mathbf{u_2}$ in (4.10) we get

$$\begin{aligned}\mathbf{y^1_{RU_A}} &= \mathbf{H_{U_A1}}(\mathbf{x_A} + \mathbf{x_B} + \mathbf{a_1}) \\ &\quad + \mathbf{H_{U_A2}}(\mathbf{x_A} + \mathbf{x_B} + \mathbf{a_2}) + \mathbf{n^1_{U_A}}.\end{aligned} \tag{4.11}$$

Combining like terms after rearranging (4.11) can give us

$$\begin{aligned}\mathbf{y^1_{RU_A}} &= (\mathbf{H_{U_A1}} + \mathbf{H_{U_A2}})\mathbf{x_A} + (\mathbf{H_{U_A1}} + \mathbf{H_{U_A2}})\mathbf{x_B} \\ &\quad + \mathbf{H_{U_A1}}\mathbf{a_1} + \mathbf{H_{U_A2}}\mathbf{a_2} + \mathbf{n^1_{U_A}}.\end{aligned} \tag{4.12}$$

The first term in (4.12) is the desired term concerning $\mathbf{U_A}$, whereas the remaining expressions are unwanted.

The received signals at $\mathbf{U_B}$ transmitted from $\mathbf{A_1}$ and $\mathbf{A_2}$ in time slot $\mathbf{t_1}$ are combined and given as follows

$$\mathbf{y^1_{RU_B}} = \mathbf{y^1_{U_B1}} + \mathbf{y^1_{U_B2}} + \mathbf{n^1_{U_B}}, \tag{4.13}$$

Now putting the values of $\mathbf{y^1_{U_B1}}$ and $\mathbf{y^1_{U_B2}}$ in (4.13) as

$$\mathbf{y^1_{RU_B}} = \mathbf{H_{U_B1}}\mathbf{u_1} + \mathbf{H_{U_B2}}\mathbf{u_2} + \mathbf{n^1_{U_B}}, \tag{4.14}$$

In (4.14) $\mathbf{u_1}$ and $\mathbf{u_2}$ represent the superimposed transmitted signals and $\mathbf{n^1_{U_B}}$ is the AWGN at $\mathbf{U_B}$ in time slot $\mathbf{t_1}$. Substituting the values of $\mathbf{u_1}$ and $\mathbf{u_2}$ in (4.14) we get

$$\begin{aligned}\mathbf{y^1_{RU_B}} &= \mathbf{H_{U_B1}}(\mathbf{x_A} + \mathbf{x_B} + \mathbf{a_1}) \\ &\quad + \mathbf{H_{U_B2}}(\mathbf{x_A} + \mathbf{x_B} + \mathbf{a_2}) + \mathbf{n^1_{U_B}}.\end{aligned} \tag{4.15}$$

Equation (4.15) can be rearranged and represented as

$$\begin{aligned}\mathbf{y^1_{RU_B}} &= (\mathbf{H_{U_B1}} + \mathbf{H_{U_B2}})\mathbf{x_A} + (\mathbf{H_{U_B1}} + \mathbf{H_{U_B2}})\mathbf{x_B} \\ &\quad + \mathbf{H_{U_B1}}\mathbf{a_1} + \mathbf{H_{U_B2}}\mathbf{a_2} + \mathbf{n^1_{U_B}}.\end{aligned} \tag{4.16}$$

The second term in (4.16) is the desired term concerning $\mathbf{U_B}$, whereas the remaining expressions are unwanted.

4.2.5.4 Reception during time slot t_2

In time slot $\mathbf{t_2}$, $\mathbf{U_A}$ and $\mathbf{U_B}$ receive the transmitted signals from antennas $\mathbf{A_1}$ and $\mathbf{A_2}$ that are given as follows

$$\mathbf{y^2_{U_A 1}} = \mathbf{H_{U_A 1} u_3}, \tag{4.17}$$

where $\mathbf{H_{U_A 1}}$ represents the channel's frequency response between $\mathbf{U_A}$ and antenna $\mathbf{A_1}$. Similarly, the received signal at $\mathbf{U_A}$ transmitted from $\mathbf{A_2}$ is represented as

$$\mathbf{y^2_{U_A 2}} = \mathbf{H_{U_A 2} u_4}, \tag{4.18}$$

where $\mathbf{H_{U_A 2}}$ is the channel frequency response between $\mathbf{U_A}$ and transmit antenna $\mathbf{A_2}$. The signals received from $\mathbf{A_1}$ at $\mathbf{U_B}$ is given as follows

$$\mathbf{y^2_{U_B 1}} = \mathbf{H_{U_B 1} u_3}, \tag{4.19}$$

where $\mathbf{H_{U_B 1}}$ represents the channel's frequency response between $\mathbf{U_B}$ and antenna $\mathbf{A_1}$. Similarly, the received signal at $\mathbf{U_B}$ which is transmitted from $\mathbf{A_2}$ is represented as

$$\mathbf{y^2_{U_B 2}} = \mathbf{H_{U_B 2} u_4}, \tag{4.20}$$

where $\mathbf{H_{U_B 2}}$ is the channel frequency response between $\mathbf{U_B}$ and transmit antenna $\mathbf{A_2}$.

The combined received signals at $\mathbf{U_A}$ from $\mathbf{A_1}$ and $\mathbf{A_2}$ are given below

$$\mathbf{y^2_{R U_A}} = \mathbf{y^2_{U_A 1}} + \mathbf{y^2_{U_A 2}} + \mathbf{n^2_{U_A}}, \tag{4.21}$$

Now putting the values of $\mathbf{y^2_{U_A 1}}$ and $\mathbf{y^2_{U_A 2}}$ in (4.21) as

$$\mathbf{y^2_{R U_A}} = \mathbf{H_{U_A 1} u_3} + \mathbf{H_{U_A 2} u_4} + \mathbf{n^2_{U_A}}, \tag{4.22}$$

In (4.22), $\mathbf{u_3}$ and $\mathbf{u_4}$ represent the superimposed transmitted signals and $\mathbf{n^2_{U_A}}$ is the AWGN at $\mathbf{U_A}$ during time slot $\mathbf{t_2}$. After substituting the values of $\mathbf{u_3}$ and

$\mathbf{u_4}$ in (4.22) we get

$$\mathbf{y^2_{RU_A}} = \mathbf{H_{U_A1}}(-\mathbf{x^*_B} - \mathbf{x^*_A} + \mathbf{a_3}) \\ + \mathbf{H_{U_A2}}(\mathbf{x^*_B} + \mathbf{x^*_A} + \mathbf{a_4}) + \mathbf{n^2_{U_A}}. \tag{4.23}$$

Combining like terms after rearranging (4.23) can give us

$$\mathbf{y^2_{RU_A}} = (-\mathbf{H_{U_A1}} + \mathbf{H_{U_A2}})\mathbf{x^*_B} + (-\mathbf{H_{U_A1}} + \mathbf{H_{U_A2}})\mathbf{x^*_A} \\ + \mathbf{H_{U_A1}}\mathbf{a_3} + \mathbf{H_{U_A2}}\mathbf{a_4} + \mathbf{n^2_{U_A}}. \tag{4.24}$$

The second term in (4.24) is the desired term concerning $\mathbf{U_A}$, whereas the remaining expressions are unwanted.

The received signals at $\mathbf{U_B}$ transmitted from $\mathbf{A_1}$ and $\mathbf{A_2}$ in time slot $\mathbf{t_2}$ are combined and given as follows

$$\mathbf{y^2_{RU_B}} = \mathbf{y^2_{U_B1}} + \mathbf{y^2_{U_B2}} + \mathbf{n^2_{U_B}}, \tag{4.25}$$

Now putting the values of $\mathbf{y^2_{U_B1}}$ and $\mathbf{y^2_{U_B2}}$ in (4.25) we get

$$\mathbf{y^2_{RU_B}} = \mathbf{H_{U_B1}}\mathbf{u_3} + \mathbf{H_{U_B2}}\mathbf{u_4} + \mathbf{n^2_{U_B}}, \tag{4.26}$$

In (4.26) $\mathbf{u_3}$ and $\mathbf{u_4}$ represent the superimposed transmitted signals and $\mathbf{n^2_{U_B}}$ is the AWGN at $\mathbf{U_B}$ during time slot $\mathbf{t_2}$. Substituting the values of $\mathbf{u_3}$ and $\mathbf{u_4}$ in (4.26) we get

$$\mathbf{y^2_{RU_B}} = \mathbf{H_{U_B1}}(-\mathbf{x^*_B} - \mathbf{x^*_A} + \mathbf{a_3}) \\ + \mathbf{H_{U_B2}}(\mathbf{x^*_A} + \mathbf{x^*_B} + \mathbf{a_4}) + \mathbf{n^2_{U_B}}. \tag{4.27}$$

Equation (4.27) can be rearranged and represented as

$$\mathbf{y^2_{RU_B}} = (-\mathbf{H_{U_B1}} + \mathbf{H_{U_B2}})\mathbf{x^*_B} + (-\mathbf{H_{U_B1}} + \mathbf{H_{U_B2}})\mathbf{x^*_A} \\ + \mathbf{H_{U_B1}}\mathbf{a_3} + \mathbf{H_{U_B2}}\mathbf{a_4} + \mathbf{n^2_{U_B}}. \tag{4.28}$$

The first term in (4.28) is the desired term concerning $\mathbf{U_B}$ and the rest of the expressions are unwanted. Equations (4.12) and (4.16) represent the received signals at $\mathbf{U_A}$ and $\mathbf{U_B}$ in time slot $\mathbf{t_1}$ respectively. Likewise, during time slot $\mathbf{t_2}$

the signals received by $\mathbf{U_A}$ and $\mathbf{U_B}$ are denoted by equations (4.24) and (4.28) respectively.

4.2.5.5 Design of Auxiliary Signals

In this subsection, the design structure of auxiliary signals $\mathbf{a_1}$, $\mathbf{a_2}$, $\mathbf{a_3}$, and $\mathbf{a_4}$ will be presented. These superimposed signals are the function of wireless channels and the unwanted user data. The motivation behind the design of these auxiliary signals stems from [AHA21, AH21, KIH21, HAL20, LH21, LH20, IKH22, AH22, IH21a, IH21b, IKH21]. Conceptually, this approach follows the auxiliary-signal superposition philosophy introduced in the MU-AS-ST family of schemes [HAL20, AH22, IH21a, IH21b], where additional signals are constructed as functions of the users' symbols and channels in order to suppress interference and, in some cases, enhance physical-layer security. The contribution of this work is therefore not the auxiliary-signal concept itself, but its adaptation and reformulation for an Alamouti STBC–MIMO–OFDM framework. In particular, we design four auxiliary vectors $(\mathbf{a_1}, \mathbf{a_2}, \mathbf{a_3}, \mathbf{a_4})$ jointly across two time slots, consistent with the conjugate and sign structure of the Alamouti code, and derive the closed-form expressions in (4.32) and (4.36) that simultaneously cancel inter-user interference and pre-compensate channel effects.

As explained earlier the desired terms for $\mathbf{U_A}$ during time slots $\mathbf{t_1}$ and $\mathbf{t_2}$ are $(\mathbf{H_{U_A1}} + \mathbf{H_{U_A2}})\mathbf{x_A}$ and $(-\mathbf{H_{U_A1}} + \mathbf{H_{U_A2}})\mathbf{x_A^*}$ respectively. For $\mathbf{U_B}$ the desired expressions are $(\mathbf{H_{U_B1}}+\mathbf{H_{U_B2}})\mathbf{x_B}$ and $(-\mathbf{H_{U_B1}}+\mathbf{H_{U_B2}})\mathbf{x_B^*}$ respectively. To calculate the values of the auxiliary signals, first, we have to design the conditions and based on these the values of signals $\mathbf{a_1}$, $\mathbf{a_2}$, $\mathbf{a_3}$, and $\mathbf{a_4}$ can be determined.

Condition for $\mathbf{a_1}$ and $\mathbf{a_2}$ during $\mathbf{t_1}$: To derive the condition for the signals $\mathbf{a_1}$ and $\mathbf{a_2}$ the second, third, and fourth terms in the equation (4.12) and the first, third and fourth expressions in equation (4.16) are equated to null matrix $\mathbf{O}$ as given below

$$(\mathbf{H_{U_A1}} + \mathbf{H_{U_A2}})\mathbf{x_B} + \mathbf{H_{U_A1}}\mathbf{a_1} + \mathbf{H_{U_A2}}\mathbf{a_2} = \mathbf{O}. \tag{4.29}$$

$$(\mathbf{H_{U_B1}} + \mathbf{H_{U_B2}})\mathbf{x_A} + \mathbf{H_{U_B1}}\mathbf{a_1} + \mathbf{H_{U_B2}}\mathbf{a_2} = \mathbf{O}. \tag{4.30}$$

From the two above equations, we can obtain a conditional system of equations such as

$$\begin{cases}(\mathbf{H}_{\mathbf{U_A}\mathbf{1}}+\mathbf{H}_{\mathbf{U_A}\mathbf{2}})\mathbf{x_B}+\mathbf{H}_{\mathbf{U_A}\mathbf{1}}\mathbf{a_1}+\mathbf{H}_{\mathbf{U_A}\mathbf{2}}\mathbf{a_2}=\mathbf{O},\\(\mathbf{H}_{\mathbf{U_B}\mathbf{1}}+\mathbf{H}_{\mathbf{U_B}\mathbf{2}})\mathbf{x_A}+\mathbf{H}_{\mathbf{U_B}\mathbf{1}}\mathbf{a_1}+\mathbf{H}_{\mathbf{U_B}\mathbf{2}}\mathbf{a_2}=\mathbf{O}.\end{cases} \tag{4.31}$$

The above system of equations can simultaneously be solved to obtain the values of $\mathbf{a_1}$ and $\mathbf{a_2}$ such as

$$\begin{cases}\mathbf{W}=\mathbf{H}_{\mathbf{U_A}\mathbf{2}}-\mathbf{H}_{\mathbf{U_A}\mathbf{1}}\mathbf{H}_{\mathbf{U_B}\mathbf{1}}{}^{-1}\mathbf{H}_{\mathbf{U_B}\mathbf{2}},\\\mathbf{a_2}=\mathbf{W}^{-1}((\mathbf{H}_{\mathbf{U_A}\mathbf{1}}+\mathbf{H}_{\mathbf{U_A}\mathbf{1}}\mathbf{H}_{\mathbf{U_B}\mathbf{1}}{}^{-1}\mathbf{H}_{\mathbf{U_B}\mathbf{2}})\mathbf{x_A}\\-(\mathbf{H}_{\mathbf{U_A}\mathbf{1}}+\mathbf{H}_{\mathbf{U_A}\mathbf{2}})\mathbf{x_B}),\\\mathbf{a_1}=-(\mathbf{I}+\mathbf{H}_{\mathbf{U_B}\mathbf{1}}{}^{-1}\mathbf{H}_{\mathbf{U_B}\mathbf{2}})\mathbf{x_A}-\mathbf{H}_{\mathbf{U_B}\mathbf{1}}{}^{-1}\mathbf{H}_{\mathbf{U_B}\mathbf{2}}\mathbf{a_2}.\end{cases} \tag{4.32}$$

Condition for $\mathbf{a_3}$ and $\mathbf{a_4}$ during $\mathbf{t_2}$: To derive the condition for the super positioned signals $\mathbf{a_3}$ and $\mathbf{a_4}$ the first, third, and fourth terms in the equation (4.24) and the second, third and fourth parts of the equation (4.28) are linked to null matrix $\mathbf{O}$ as given below

$$(-\mathbf{H}_{\mathbf{U_A}\mathbf{1}}+\mathbf{H}_{\mathbf{U_A}\mathbf{2}})\mathbf{x_B^*}+\mathbf{H}_{\mathbf{U_A}\mathbf{1}}\mathbf{a_3}+\mathbf{H}_{\mathbf{U_A}\mathbf{2}}\mathbf{a_4}=\mathbf{O}. \tag{4.33}$$

$$(-\mathbf{H}_{\mathbf{U_B}\mathbf{1}}+\mathbf{H}_{\mathbf{U_B}\mathbf{2}})\mathbf{x_A^*}+\mathbf{H}_{\mathbf{U_B}\mathbf{1}}\mathbf{a_3}+\mathbf{H}_{\mathbf{U_B}\mathbf{2}}\mathbf{a_4}=\mathbf{O}. \tag{4.34}$$

From the two above equations, we can obtain a conditional system of equations such as

$$\begin{cases}(-\mathbf{H}_{\mathbf{U_A}\mathbf{1}}+\mathbf{H}_{\mathbf{U_A}\mathbf{2}})\mathbf{x_B^*}+\mathbf{H}_{\mathbf{U_A}\mathbf{1}}\mathbf{a_3}+\mathbf{H}_{\mathbf{U_A}\mathbf{2}}\mathbf{a_4}=\mathbf{O},\\(-\mathbf{H}_{\mathbf{U_B}\mathbf{1}}+\mathbf{H}_{\mathbf{U_B}\mathbf{2}})\mathbf{x_A^*}+\mathbf{H}_{\mathbf{U_B}\mathbf{1}}\mathbf{a_3}+\mathbf{H}_{\mathbf{U_B}\mathbf{2}}\mathbf{a_4}=\mathbf{O}.\end{cases} \tag{4.35}$$

By repeating the same process utilized in (4.31), the above system of equations can be solved to obtain the values of $\mathbf{a_3}$ and $\mathbf{a_4}$ as shown below

$$\begin{cases}\mathbf{W}=\mathbf{H}_{\mathbf{U_A}\mathbf{2}}-\mathbf{H}_{\mathbf{U_A}\mathbf{1}}\mathbf{H}_{\mathbf{U_B}\mathbf{1}}{}^{-1}\mathbf{H}_{\mathbf{U_B}\mathbf{2}},\\\mathbf{a_4}=\mathbf{W}^{-1}((-\mathbf{H}_{\mathbf{U_A}\mathbf{1}}+\mathbf{H}_{\mathbf{U_A}\mathbf{1}}\mathbf{H}_{\mathbf{U_B}\mathbf{1}}{}^{-1}\mathbf{H}_{\mathbf{U_B}\mathbf{2}})\mathbf{x_A^*}\\-(-\mathbf{H}_{\mathbf{U_A}\mathbf{1}}+\mathbf{H}_{\mathbf{U_A}\mathbf{2}})\mathbf{x_B^*}),\\\mathbf{a_3}=-(-\mathbf{I}+\mathbf{H}_{\mathbf{U_B}\mathbf{1}}{}^{-1}\mathbf{H}_{\mathbf{U_B}\mathbf{2}})\mathbf{x_A^*}-\mathbf{H}_{\mathbf{U_B}\mathbf{1}}{}^{-1}\mathbf{H}_{\mathbf{U_B}\mathbf{2}}\mathbf{a_4}.\end{cases} \tag{4.36}$$

Using the system of equations such as (4.32) and (4.36), the unwanted terms and the channel's effects are eliminated during time slots $\mathbf{t_1}$ and $\mathbf{t_2}$. Therefore through (4.32), the equations (4.12) and (4.16) can be modified as

$$\mathbf{y^1_{RU_A}}=(\mathbf{H}_{\mathbf{U_A}\mathbf{1}}+\mathbf{H}_{\mathbf{U_A}\mathbf{2}})\mathbf{x_A}+\mathbf{n^1_{U_A}}. \tag{4.37}$$

$$\mathbf{y^1_{RU_B}} = (\mathbf{H_{U_B1}} + \mathbf{H_{U_B2}})\mathbf{x_B} + \mathbf{n^1_{U_B}}. \tag{4.38}$$

Form the operation of (4.36), the equations such as (4.24) and (4.28) are altered as a result we get

$$\mathbf{y^2_{RU_A}} = (-\mathbf{H_{U_A1}} + \mathbf{H_{U_A2}})\mathbf{x^*_A} + \mathbf{n^2_{U_A}}. \tag{4.39}$$

$$\mathbf{y^2_{RU_B}} = (-\mathbf{H_{U_B1}} + \mathbf{H_{U_B2}})\mathbf{x^*_B} + \mathbf{n^2_{U_B}}. \tag{4.40}$$

4.2.5.6 Proposed Equalization Policy

In conventional Alamouti STBC scheme [Ala98], the restoring of signals and removing of channels effects is done using a combining scheme and maximum likelihood decision rule. However, in this study, we also focus on reducing the complexity at the receiver side by proposing a simple equalization process at the receiver.

After the duration of $\mathbf{t_1}$ and $\mathbf{t_2}$, the modified received signals are represented by equations (4.37), (4.38), (4.39), and (4.40) are equalized at the respective receivers to remove the channel effects and get the desired data at the legitimate user which is done by simply multiplying each equation by the inverse of the sum of channels experienced by the received signals as shown below

$$\begin{aligned}\mathbf{y^1_{RU_A}} &= (\mathbf{H_{U_A1}} + \mathbf{H_{U_A2}})^{-1}((\mathbf{H_{U_A1}} + \mathbf{H_{U_A2}})\mathbf{x_A}) \\ &\quad + (\mathbf{H_{U_A1}} + \mathbf{H_{U_A2}})^{-1}\mathbf{n^1_{U_A}}.\end{aligned} \tag{4.41}$$

$$\begin{aligned}\mathbf{y^1_{RU_B}} &= (\mathbf{H_{U_B1}} + \mathbf{H_{U_B2}})^{-1}((\mathbf{H_{U_B1}} + \mathbf{H_{U_B2}})\mathbf{x_B}) \\ &\quad + (\mathbf{H_{U_B1}} + \mathbf{H_{U_B2}})^{-1}\mathbf{n^1_{U_B}}.\end{aligned} \tag{4.42}$$

$$\begin{aligned}\mathbf{y^2_{RU_A}} &= (-\mathbf{H_{U_A1}} + \mathbf{H_{U_A2}})^{-1}((-\mathbf{H_{U_A1}} + \mathbf{H_{U_A2}})\mathbf{x^*_A}) \\ &\quad + (-\mathbf{H_{U_A1}} + \mathbf{H_{U_A2}})^{-1}\mathbf{n^2_{U_A}}.\end{aligned} \tag{4.43}$$

$$\begin{aligned}\mathbf{y^2_{RU_B}} &= (-\mathbf{H_{U_B1}} + \mathbf{H_{U_B2}})^{-1}((-\mathbf{H_{U_B1}} + \mathbf{H_{U_B2}})\mathbf{x^*_B}) \\ &\quad + (-\mathbf{H_{U_B1}} + \mathbf{H_{U_B2}})^{-1}\mathbf{n^2_{U_B}}.\end{aligned} \tag{4.44}$$

It is important to note that in equations (4.41), (4.42), (4.43), and (4.44), the inverse operation of $\mathbf{H_{U_A 1}}$, $\mathbf{H_{U_A 2}}$, $\mathbf{H_{U_B 1}}$, and $\mathbf{H_{U_B 2}}$ is relatively simple due to the channel matrices being diagonal.

During $\mathbf{t_1}$, after equalization the transformed desired signals are

$$\begin{cases} \mathbf{y^1_{RU_A}} = \mathbf{x_A} + (\mathbf{H_{U_A 1}} + \mathbf{H_{U_A 2}})^{-1}\mathbf{n^1_{U_A}}, \\ \mathbf{y^1_{RU_B}} = \mathbf{x_B} + (\mathbf{H_{U_B 1}} + \mathbf{H_{U_B 2}})^{-1}\mathbf{n^1_{U_B}}. \end{cases} \tag{4.45}$$

Moreover during $\mathbf{t_2}$, after equalization is done the remodelled desired signals are

$$\begin{cases} \mathbf{y^2_{RU_A}} = \mathbf{x^*_A} + (-\mathbf{H_{U_A 1}} + \mathbf{H_{U_A 2}})^{-1}\mathbf{n^2_{U_A}}, \\ \mathbf{y^2_{RU_B}} = \mathbf{x^*_B} + (-\mathbf{H_{U_B 1}} + \mathbf{H_{U_B 2}})^{-1}\mathbf{n^2_{U_B}}. \end{cases} \tag{4.46}$$

4.2.5.7 Combining Process

To recover the expected user data we have to simply combine the appropriate signals obtained during the two-time slots. Thus for $\mathbf{U_A}$, we can write

$$\begin{aligned} \mathbf{y^1_{RU_A}} + \mathbf{y^2_{RU_A}} = \mathbf{x_A} + \mathbf{x^*_A} + (\mathbf{H_{U_A 1}} + \mathbf{H_{U_A 2}})^{-1}\mathbf{n^1_{U_A}} \\ +(-\mathbf{H_{U_A 1}} + \mathbf{H_{U_A 2}})^{-1}\mathbf{n^2_{U_A}}, \end{aligned} \tag{4.47}$$

Since all the processing is done in the frequency domain, therefore by utilizing the conjugate symmetry property of the Fast Fourier transform, equation (4.47) can be rewritten as

$$\begin{cases} \mathbf{y^1_{RU_A}} + \mathbf{y^2_{RU_A}} = \mathbf{x_A} + \mathbf{N_{U_A}}, \\ \mathbf{N_{U_A}} = (\mathbf{H_{U_A 1}} + \mathbf{H_{U_A 2}})^{-1}\mathbf{n^1_{U_A}} \\ +(-\mathbf{H_{U_A 1}} + \mathbf{H_{U_A 2}})^{-1}\mathbf{n^2_{U_A}}. \end{cases} \tag{4.48}$$

On the other hand, for $\mathbf{U_B}$ from the system of equations (4.45) and (4.46), we can write

$$\begin{aligned} \mathbf{y^1_{RU_B}} + \mathbf{y^2_{RU_B}} = \mathbf{x_B} + \mathbf{x^*_B} + (\mathbf{H_{U_B 1}} + \mathbf{H_{U_B 2}})^{-1}\mathbf{n^1_{U_B}} \\ +(-\mathbf{H_{U_B 1}} + \mathbf{H_{U_B 2}})^{-1}\mathbf{n^2_{U_B}}, \end{aligned} \tag{4.49}$$

Following the same symmetry property applied in (4.48) we can rewrite (4.49) as

$$\begin{cases} \mathbf{y}^{\mathbf{1}}_{\mathbf{R}\mathbf{U}_\mathbf{B}} + \mathbf{y}^{\mathbf{2}}_{\mathbf{R}\mathbf{U}_\mathbf{B}} = \mathbf{x}_\mathbf{B} + \mathbf{N}_{\mathbf{U}_\mathbf{B}}, \\ \mathbf{N}_{\mathbf{U}_\mathbf{B}} = (\mathbf{H}_{\mathbf{U}_\mathbf{B}\mathbf{1}} + \mathbf{H}_{\mathbf{U}_\mathbf{B}\mathbf{2}})^{-1}\mathbf{n}^{\mathbf{1}}_{\mathbf{U}_\mathbf{B}} \\ +(-\mathbf{H}_{\mathbf{U}_\mathbf{B}\mathbf{1}} + \mathbf{H}_{\mathbf{U}_\mathbf{B}\mathbf{2}})^{-1}\mathbf{n}^{\mathbf{2}}_{\mathbf{U}_\mathbf{B}}. \end{cases} \tag{4.50}$$

4.3 Quality of Service (QoS) in Multi-User Environments

Quality of Service (QoS) in multi-user environments, especially within the realms of wireless networks employing the MU-STBC-ST scheme, is a critical aspect that demands comprehensive attention. Quality of Service (QoS) in wireless networks refers to the ability of the network to provide a high level of service to users by ensuring reliable and consistent data transmission. In the context of the MU-STBC-ST scheme, QoS ensures that multiple users sharing the same network resources (time, frequency, and power) receive satisfactory levels of service, maintaining adequate signal quality and data rates. The most important QoS metrics in this system include the Signal-to-Interference-plus-Noise Ratio (SINR), throughput, and Bit Error Rate (BER).

The SINR is a primary metric that reflects the quality of the received signal relative to interference and noise. For a particular user i, the SINR can be mathematically expressed as:

$$\text{SINR}_i = \frac{P_i |H_i|^2}{\sum_{j \neq i} P_j |H_j|^2 + N} \tag{4.51}$$

where P_i is the transmit power to user i, $|H_i|$ is the channel gain for user i, P_j and $|H_j|$ represent the interference from other users, and N is the noise power.

Throughput is another key QoS metric, indicating the rate of successful data transmission. It can be derived from SINR, with the formula:

$$\text{Throughput} = \text{Bandwidth} \times \log_2(1 + \text{SINR}) \tag{4.52}$$

This shows how the achievable data rate is directly influenced by the SINR. Additionally, the Bit Error Rate (BER), which quantifies the number of bits incorrectly received, is another important metric for assessing the reliability of the network. The BER can be expressed as:

$$\text{BER} = Q\left(\sqrt{\text{SINR}}\right) \tag{4.53}$$

where $Q(x)$ is the Q-function, representing the tail probability of the Gaussian distribution. These mathematical expressions allow us to evaluate and manage the QoS for users in the MU-STBC-ST scheme, ensuring efficient and reliable communication in multi-user environments. In such environments, the challenge of maintaining high service quality becomes increasingly complex due to the diverse requirements and interactions of multiple users sharing the network resources. The MU-STBC-ST scheme, a sophisticated blend of Space-Time Block Coding (STBC) and Orthogonal Frequency Division Multiplexing (OFDM) tailored for multi-user scenarios, plays a crucial role in addressing these complexities. One of the primary aspects of QoS in multi-user environments is ensuring consistent and reliable data transmission rates, regardless of the number of users or the intensity of network traffic.

Another critical aspect of QoS in multi-user environments is the management of interference, which can significantly degrade service quality. In densely populated network areas or in scenarios where multiple devices are competing for the same bandwidth, the likelihood of interference increases, leading to reduced signal quality and network performance. The MU-STBC-ST scheme addresses this by employing sophisticated algorithms designed to mitigate interference.This includes the use of multiple antennas in transmission and reception (MIMO technology), which not only enhances the signal strength but also provides a means to cancel out interference from other users.

Lastly, the MU-STBC-ST scheme significantly contributes to user-centric QoS by enabling the customization of services based on individual user requirements.In a multi-user environment, different users may have varying needs in terms of data speed, latency, and reliability. For instance, a streaming service may require higher bandwidth and lower latency, whereas an IoT device may prioritize energy efficiency and reliability over speed. The MU-STBC-ST scheme allows for such differentiation in service delivery, enabling the network to prioritize resources in a way that aligns with the specific demands of each user or application. This level of customization is achieved through advanced network management systems that monitor user requirements in real-time and adjust the network settings accordingly [NNP17].By doing so, the scheme ensures that all users receive a level of service that is not only high in quality but also tailored to their specific needs. This user-centric approach to QoS is particularly important in contemporary networks, where user satisfaction is closely tied to the perceived quality of network services.

The parameters used in the analysis of MU-STBC-ST are given in Table 4.1. This section focuses on the simulated results of the proposed MU-STBC-ST obtained through MATLAB simulation. The performance metrics such as BER, TER, and PAPR are utilized to quantify the proposed scheme. Bit error rate is a measure of the quality of a digital communication link. It is defined as the ratio of the number of bit errors (e.g. flipped bits) to the total number of bits transmitted. The mathematical expression for BER is:

$$\mathrm{BER} = \frac{\text{Number of bit errors}}{\text{Total number of bits transmitted}} \tag{4.54}$$

Throughput Error Rate (TER): Throughput error rate is a measure of the amount of data that is successfully transmitted over a communication link. It is defined as the ratio of the number of bits successfully transmitted to the total number of bits transmitted. The mathematical expression for TER is:

$$\mathrm{TER} = \frac{\text{Number of bits successfully transmitted}}{\text{Total number of bits transmitted}} \tag{4.55}$$

Peak to Average Power Ratio (PAPR): Peak to average power ratio is a measure of the efficiency of a signal in a wireless communication system. It is defined as the ratio of the peak power of a signal to its average power. The mathematical expression for PAPR is:

$$\mathrm{PAPR} = \frac{\text{Peak power}}{\text{Average power}} \tag{4.56}$$

In the realm of signal transmission and power utilization, Section 4.4.1 provides an insightful comparison of the Peak-to-Average Power Ratio (PAPR) between the proposed MU-STBC-ST design and conventional Orthogonal Frequency Division Multiplexing (OFDM). A lower PAPR, as exhibited by the MU-STBC-ST design, indicates a more efficient use of transmission power and a reduction in signal processing complexity. This efficiency stems from intelligent system design, which optimizes resource use and minimizes operational complexities.

4.4 Case Studies and Results

The experimental setup for testing MU-STBC-ST is carefully designed to evaluate its efficiency and performance. This setup integrates Alamouti Space-Time Block Coding (STBC) with MIMO-OFDM, enabling the simultaneous transmission of user signals. Two antennas (A1 and A2) transmit superimposed signals over two time slots (t1 and t2) to two users (UA and UB) using the same spatial and temporal resources. This approach improves data transmission while maintaining signal integrity in multi-user environments, making it a significant step forward in wireless communication.

4.4.1 Experimental Setup

The experimental setup incorporates a simulation environment, as detailed in Table 4.1, which outlines the parameters used in the analysis of the MU-STBC-ST design. These parameters include aspects such as channel type, channel length, cyclic prefix, IFFT/FFT size, modulation type, signal-to-noise ratio, channel delay, and channel power profile. The utilization of Binary Phase Shift Keying (BPSK) modulation in an OFDM system containing 64 subcarriers with a cyclic prefix size L is particularly noteworthy. BPSK is adopted in this chapter to isolate the algorithmic behaviour of the proposed MU-STBC-ST scheme from modulation-induced effects. Since BPSK has the smallest minimum Euclidean distance, it yields the worst-case BER performance, and thus provides a conservative estimate of reliability. The objective of this chapter is to evaluate the interference cancellation and diversity gains of the proposed transmission scheme, rather than absolute spectral efficiency. Higher-order modulations such as QPSK or 16-QAM are expected to exhibit similar relative performance trends, with shifted BER curves. This setup is designed to eliminate inter-symbol interference (ISI), a common challenge in high-speed data transmission. As shown in Table 4.1, the channel between the transmitter antennas and receiving antennas is a multi-path Rayleigh fading channel with an equal number of L multipath i.e. taps L=9. L is the number of propagation paths, we can view them as independent and identically distributed random variables. As per the central limit theorem, this will approximately give Gaussian distribution.

In this experimental framework, the data for each user is superimposed and transmitted along with specially designed auxiliary signals. These auxiliary signals, a1, a2, a3, and a4, are intricately crafted based on the users' channel characteristics, ensuring optimal signal reception and minimizing interference. The experiment's robustness is further ensured by employing a channel frequency response diagonal matrix, denoted as Hkm, which accommodates multiple users and transmit antennas. This matrix plays a critical role in the experimental setup, as it simulates real-world channel conditions, providing a comprehensive evaluation of the system's performance under various scenarios.

During the first time slot, t1, the superimposed transmitted signal from antenna A1 is represented as u1, a combination of xA, xB, and a1. Similarly, the signal from antenna A2, denoted as u2, comprises xA, xB, and a2. In the second time slot, t2, the transmitted signals, u3, and u4, from antennas A1 and A2, respectively, include the conjugate frequency domain user vector data for UA and UB, alongside auxiliary signals a3 and a4. This methodical approach to signal transmission is designed to maximize data throughput while ensuring minimal interference among users.

The reception process during each time slot is equally critical in the experimental setup. During t1, the received signals at UA and UB, denoted as $\mathbf{y^1_{RU_A}}$ and $\mathbf{y^1_{RU_B}}$, respectively, include the desired terms concerning each user and auxiliary signals. The same process is applied during t2, with the received signals $\mathbf{y^2_{RU_A}}$ and $\mathbf{y^2_{RU_B}}$ at UA and UB, respectively. This reception process is vital for evaluating the system's efficiency in delivering data accurately to each user. Equal-distance assumption: To isolate the effects of the proposed MU-STBC-ST signal processing scheme from large-scale variability, both users are simulated with the same average SNR (equal path-loss). This corresponds to the equal-distance two-user scenario, which is known to be the most challenging case for conventional power-domain NOMA, as there is no large-scale channel disparity to exploit. The equal-distance assumption is intentionally adopted to evaluate the proposed MU-STBC-ST scheme under the challenging scenario where both users have similar large-scale channel conditions, i.e., without benefiting from path-loss-based separation exploited in power-domain NOMA. This setting highlights the intrinsic performance of the proposed transmission strategy, independent of large-scale propagation effects.

The design of the auxiliary signals is a crucial aspect of the experiment, as these signals are pivotal in eliminating unwanted interference and channel effects during transmission. The auxiliary signals are designed based on a set of conditions derived from the combination of user data and channel characteristics. The successful implementation of these conditions ensures that the MU-STBC-

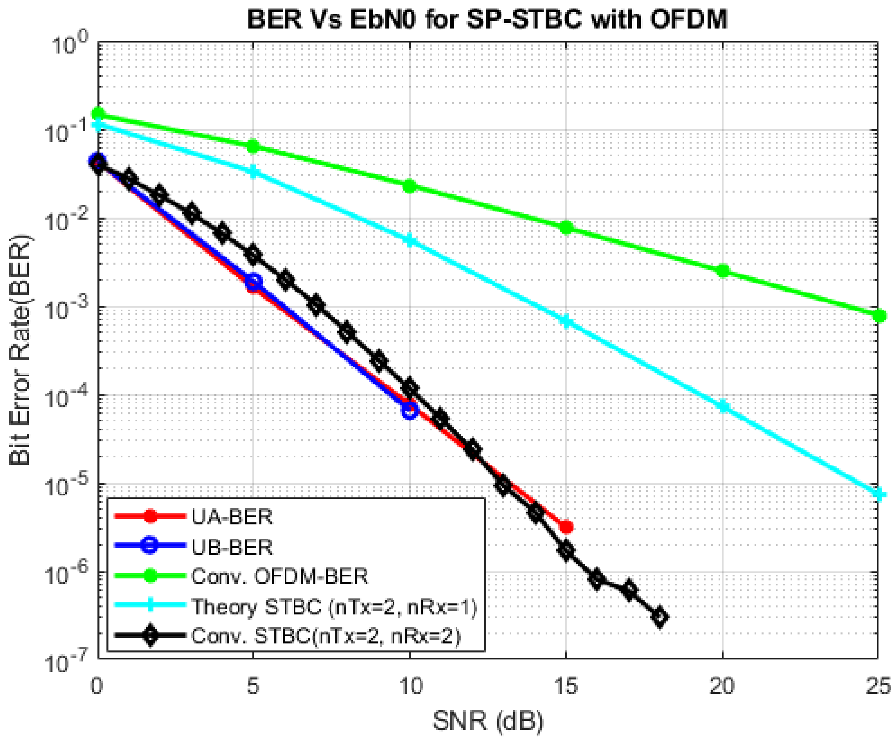

Figure 4.2: Bit Error Rate (BER) vs. Signal-to-Noise Ratio (SNR), comparing various configurations including uncoded OFDM (UA-BER and UB-BER), conventional OFDM (Conv. OFDM-BER), theoretical STBC (nTx=2, nRx=1), and conventional STBC (nTx=2, nRx=2).

ST design can effectively manage the complexities of a multi-user environment, delivering high-quality communication.

To quantify the performance of the proposed MU-STBC-ST design, several key metrics are utilized, including Bit Error Rate (BER), Throughput Error Rate (TER), and Peak-to-Average Power Ratio (PAPR). These metrics provide a comprehensive assessment of the system's performance, evaluating aspects such as signal integrity, data transmission efficiency, and power utilization. The BER metric, for instance, measures the quality of the digital communication link by determining the ratio of the number of bit errors to the total number of bits transmitted. This metric is crucial for assessing the system's ability to maintain signal quality in a multi-user environment.

Similarly, the TER metric evaluates the amount of data successfully transmitted over the communication link, offering insights into the system's overall throughput performance. In contrast, the PAPR metric assesses the efficiency of the signal in the wireless communication system, comparing the peak power of the signal to its average power. A lower PAPR is indicative of more efficient power

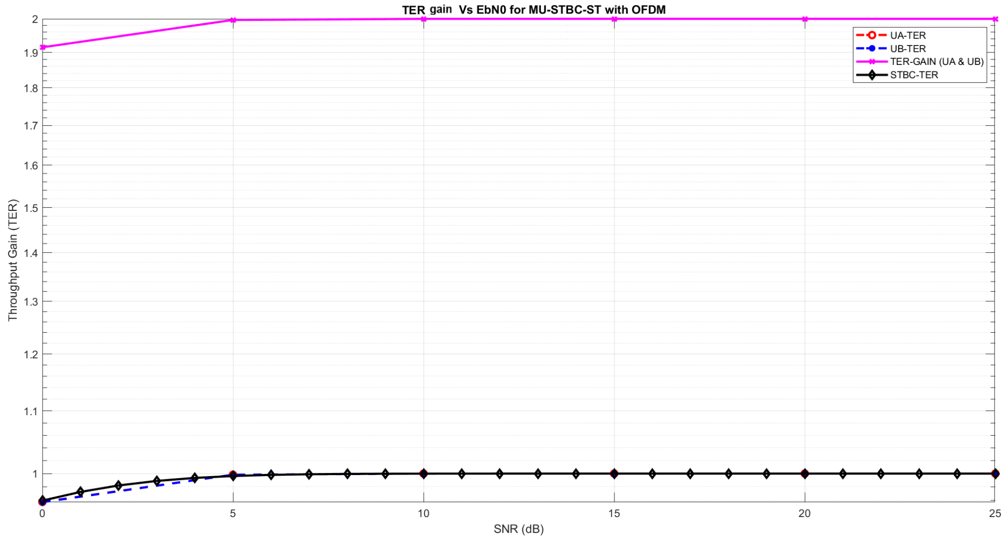

Figure 4.3: TER gain of the proposed MU-STBC-ST scheme with comparison to simulated Alamouti STBC TER. Illustrates the Throughput Gain (TER) versus Signal-to-Noise Ratio (SNR) performance for the MU-STBC-ST (Multi-User Space-Time Block Coding with Space-Time) design integrated with OFDM under various configurations.

usage and reduced signal processing complexity, highlighting the system's effectiveness in managing power resources.

Table 4.1: Simulation Parameters

Channel	Multipath Rayleigh Fading Channel
Channel Length	9
Cyclic Prefix (CP)	9
IFFT/FFT Size	64
Modulation Type	BPSK
Signal to noise ratio in dB	0-25 dB
Channel Delay	[0 3 5 6 8]
Channel Power Profile	[0 -8 -17 -21 -25]
Guard Interval Length	16

The experimental setup also incorporates a simulation environment, as detailed in Table I, which outlines the parameters used in the analysis of the MU-STBC-ST design. These parameters include aspects such as channel type, channel length, cyclic prefix, IFFT/FFT size, modulation type, signal-to-noise ratio, channel delay, and channel power profile. The utilization of Binary Phase Shift Keying (BPSK) modulation in an OFDM system containing 64 subcarriers with a cyclic prefix size L is particularly noteworthy. This setup is designed to eliminate inter-symbol interference, a common challenge in high-speed data

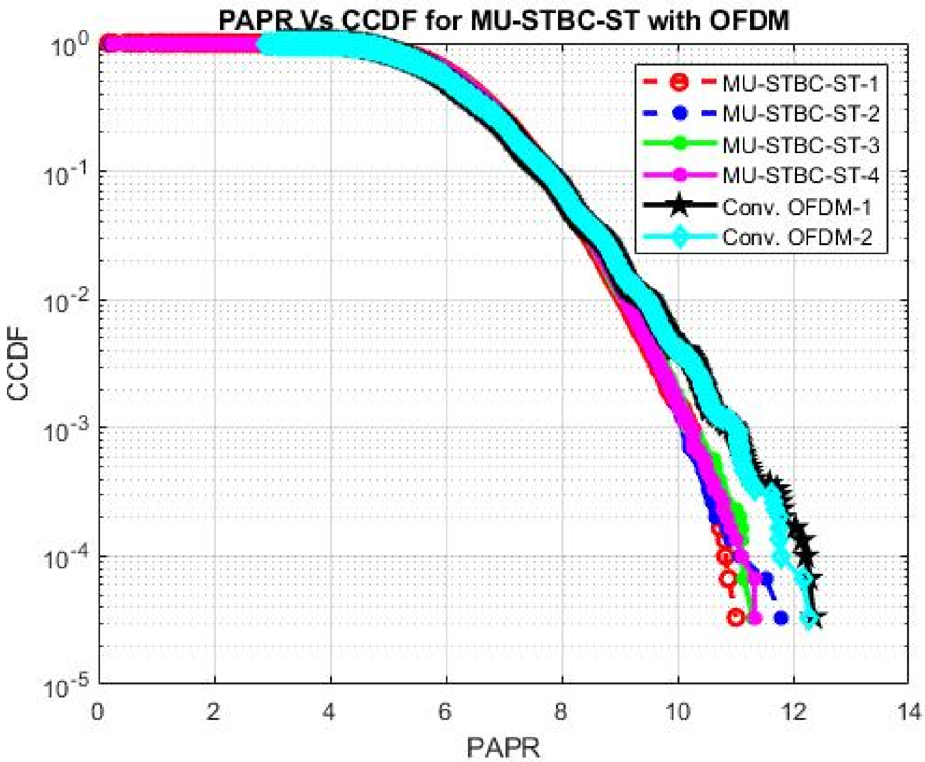

Figure 4.4: PAPR of the proposed MU-STBC-ST design showing the comparison with conventional OFDM:CCDF of PAPR for MU-STBC-ST and conventional OFDM systems.

transmission. The simulation environment mimics a multi-path Rayleigh fading channel with a specific number of multi-path taps, L, providing a realistic assessment of the system's performance under various channel conditions.

4.4.2 Analysis of Findings

The experimental findings present a detailed analysis of the Bit Error Rate (BER) performance of the proposed MU-STBC-ST scheme, as showcased in Figure 4.2Figure 4.2 compares the BER performance of the proposed MU-STBC-ST scheme with conventional OFDM, theoretical Alamouti STBC (2×1), and simulated conventional STBC (2×2). As expected, conventional OFDM exhibits the poorest performance due to the absence of spatial diversity. The theoretical 2×1 Alamouti curve shows a diversity order of approximately 2, whereas the simulated 2×2 STBC curve demonstrates a diversity order close to 4.

The proposed MU-STBC-ST exhibits a similar asymptotic slope to the 2×2 STBC case, indicating that it achieves comparable diversity benefits despite operating with only one receive antenna per user. This behaviour arises from the auxiliary-signal design, which simultaneously cancels inter-user interference and pre-compensates for the channel, leading to improved reliability without conventional receiver-side combining or equalization.

Thus, the results demonstrate that MU-STBC-ST can deliver diversity gains comparable to conventional 2×2 STBC under equal-average-SNR conditions, while supporting simultaneous multi-user transmission over shared time–frequency resources.

This analysis demonstrates a significant reduction in the rate of bit errors compared to conventional OFDM and Alamouti STBC systems, highlighting the enhanced ability of the MU-STBC-ST design to mitigate noise and interference. Advanced signal processing techniques, such as effective handling of multi-user interference and adaptation to channel distortions, play a crucial role in this improved performance. The reduced BER ensures a robust communication link, providing reliable and accurate data transmission, which is essential for applications where precision is critical. The analysis of throughput efficiency, represented by the Throughput Error Rate (TER) and illustrated in Figure 4.3, provides an in-depth evaluation of the MU-STBC-ST scheme's data handling capabilities. The x-axis represents the SNR (Eb/N0) in decibels, while the y-axis indicates the throughput gain, measuring the system's efficiency in maximizing data transmission. The graph compares uncoded OFDM configurations (UA-TER and UB-TER), their combined throughput gain (TER-GAIN), and the throughput gain achieved with STBC (STBC-TER).

The legends represent the following:

- UA-TER: Throughput Error Rate for User A in the MU-STBC-ST scheme.
- UB-TER: Throughput Error Rate for User B in the MU-STBC-ST scheme.
- TER-GAIN (UA & UB): The Total Throughput Error Rate Gain (TER-GAIN) for both users U_A and U_B in the Multi-User Space-Time Block Coding with Superposition Transmission (MU-STBC-ST) scheme quantifies the improvement in error rate performance compared to the conventional STBC scheme. It is mathematically defined as:

$$\text{TER-GAIN} = \frac{\text{TER}_{\text{STBC}} - \text{TER}_{\text{MU-STBC-ST}}}{\text{TER}_{\text{STBC}}} \times 100\% \tag{4.57}$$

where $\text{TER}_{\text{MU-STBC-ST}}$ represents the throughput error rate in the MU-STBC-ST scheme, and TER_{STBC} denotes the throughput error rate in the conventional STBC scheme. The TER-GAIN metric expresses the relative reduction in error rate, demonstrating the extent to which MU-STBC-ST enhances transmission reliability compared to traditional STBC. A higher TER-GAIN value signifies a more significant reduction in the error rate, indicating a more robust and reliable communication system. If TER-GAIN is positive, it implies that MU-STBC-ST outperforms conventional STBC by reducing errors, while a negative value would suggest an increase in errors. This metric provides a quantitative measure of the effectiveness of MU-STBC-ST in mitigating errors and improving overall data transmission quality.

The graph demonstrates the TER performance of MU-STBC-ST and STBC, showing the difference in throughput error rates. A positive TER Gain indicates that MU-STBC-ST has a lower error rate than STBC, while a negative TER Gain would indicate a higher error rate. A higher gain corresponds to a lower error rate. The UA-TER and UB-TER configurations exhibit relatively modest throughput gains, reflecting the limited performance of uncoded systems in mitigating interference and optimizing data transmission. In contrast, the TER-GAIN curve shows a significant improvement, achieving a throughput gain of approximately 2, highlighting the synergistic benefits of combining configurations UA and UB.

The individual throughput performance of MU-STBC-ST is comparable to the overall TER performance of STBC. However, the total performance gain of MU-STBC-ST, represented by TER-GAIN (UA & UB), is twice that of conventional STBC. A TER Gain of 2 means that MU-STBC-ST has half the error rate of STBC, clearly demonstrating the superiority of the proposed technique's superpositioned user data method. The total gain reflects the system's overall spectral efficiency per unit area, or the cumulative throughput received by all users.

As channel conditions improve (higher SNR), the advantage of MU-STBC-ST over STBC in terms of error rate reduction is expected to diminish. The STBC-TER curve remains stable across all SNR levels, indicating the robustness of the STBC scheme in maintaining throughput under varying channel conditions.

This significant improvement in throughput is attributed to the novel superpositioning technique employed in the MU-STBC-ST scheme, which allows for simultaneous data transmission to multiple users without degradation in data rates or service quality. Such an enhancement in throughput is pivotal for supporting high-bandwidth applications and addressing the exponential growth in user numbers and data consumption.

The PAPR of MU-STBC-ST is lower than that of conventional OFDM due to the incorporation of auxiliary signals as previously seen in figure 4.4.Figure depicts the Complementary Cumulative Distribution Function (CCDF) of Peak-to-Average Power Ratio (PAPR) for MU-STBC-ST with OFDM across multiple configurations (MU-STBC-ST-1 to MU-STBC-ST-4) in comparison to conventional OFDM systems (Conv. OFDM-1 and Conv. OFDM-2). The x-axis represents the PAPR values, which indicate the ratio of the peak signal power to its average power, while the y-axis denotes the probability (CCDF) that the PAPR exceeds a certain threshold. A lower CCDF for a given PAPR reflects better system performance in terms of power efficiency.

MU-STBC-ST-1 and MU-STBC-ST-2 represent the PAPR values for the first and second antennas during the initial time slot, while MU-STBC-ST-3 and MU-STBC-ST-4 correspond to the PAPR of signals transmitted from both antennas during the second time slot. In contrast, Conv. OFDM-1 and Conv. OFDM-2 indicate the PAPR of the first and second antennas in the conventional OFDM system. The results clearly show that all MU-STBC-ST configurations achieve significantly better PAPR performance than conventional OFDM systems, with lower CCDF values at corresponding PAPR thresholds. Specifically, MU-STBC-ST-1 and MU-STBC-ST-2 exhibit the best performance with the lowest PAPR values, while MU-STBC-ST-3 and MU-STBC-ST-4 demonstrate slightly higher PAPR but still outperform both conventional OFDM configurations.

The reduced PAPR in MU-STBC-ST is primarily attributed to the use of auxiliary signals ($\mathbf{a_1}$, $\mathbf{a_2}$, $\mathbf{a_3}$, and $\mathbf{a_4}$) and alterations in the distribution of the transmitted signals. While conventional OFDM typically exhibits a high PAPR, the low PAPR achieved by MU-STBC-ST makes it particularly advantageous for low-complexity devices and energy-efficient applications. This reduced PAPR improves the efficiency of power amplifiers, minimizes non-linear distortions, and leads to improved spectral and energy efficiency, which are critical for sustainable network operations. Additionally, a lower PAPR enhances battery life in mobile devices and decreases the overall energy requirements of the network, making the MU-STBC-ST scheme a desirable solution for future wireless systems.

This graph underscores the proposed system's ability to address a significant challenge faced by OFDM systems by effectively reducing PAPR while maintaining reliability. These results validate the design's potential to conserve power, reduce complexity, and support energy-efficient communication, particularly in modern wireless networks demanding high performance under constrained power conditions.

4.5 Conclusion

The analysis of the experimental findings presents the MU-STBC-ST scheme as a highly promising technology for enhancing the efficiency and performance of wireless networks. The improvements observed in key performance metrics such as BER, throughput, and PAPR, coupled with the scheme's scalability and user-centric benefits, position it as an invaluable asset for future wireless communication systems. These findings not only demonstrate the current capabilities of the MU-STBC-ST design but also provide valuable insights for ongoing innovation in wireless technologies. They highlight the need for continuous development of intelligent and efficient solutions to meet the burgeoning challenges in modern network environments.

Chapter 5

Composite Base Stations and Dynamic Spectrum Scheduling

> *"Man kann diese wunderbare Theorie nicht studieren, ohn bisweilen die Empfindung zu haben, als wohne den mathematischen Formeln selbstständiges Leben und eigener Verstand inne, als seien diesel ben klüger als wir, klüger sogar als ihr Erfinder, als gaben sie uns mehr heraus, als seinerzeit in sie hineingelegt wurde"*
>
> Translation: One cannot study this wonderful theory without occasionally having the feeling that the mathematical formulas possess an independent life and their own intelligence, as if they were wiser than we are, wiser even than their inventor, as if they reveal more to us than was originally put into them.
>
> *Heinrich Rudolf Hertz*

This chapter explores 5G technologies, focusing on Composite Base Stations (CBS) and Dynamic Spectrum Sharing (DSS) to meet the demands of higher bandwidth, lower latency, and greater network capacity, thereby improving 5G network performance in capacity, coverage, and spectrum use. CBS, a key innovation, marks a shift from traditional base stations, addressing the challenges of capacity and coverage in dense urban environments.

The chapter highlights Dynamic Spectrum Scheduling (DSS) as key to managing spectrum resources efficiently, addressing spectrum scarcity, and meeting network demands. Technical details, real-world applications, and case studies illustrate their potential, while challenges like technical and regulatory hurdles are also acknowledged. The introduction concludes by previewing the in-depth analysis that follows.

5.1 Related Work on Composite Base Stations and Dynamic Spectrum Scheduling

Composite Base Stations (CBS) and Dynamic Spectrum Scheduling (DSS) have been proposed in the literature as architectural and radio resource management mechanisms to improve spectrum utilization, enable multi-RAT coexistence, and enhance user experience in dense urban deployments. Early work on multi-RAT base station architectures demonstrated the feasibility of integrating LTE and NR radios under a unified control plane to allow coordinated scheduling and load sharing across spectrum bands [3GP23, 3GP25]. Dynamic Spectrum Scheduling, as standardized by 3GPP, enables time-domain multiplexing of LTE and NR transmissions over shared spectrum resources, allowing operators to adapt resource allocation based on traffic demand and deployment constraints [3GP24]

CBS is a technology that combines different wireless access technologies such as 5G and LTEinto a single base station. Thus, operators can efficiently utilize the available spectrum and provide seamless connectivity to their users. Moreover, CBS has the ability to provide higher data rates, reduce latency, and better coverage as compared to conventional BSS. While CBS technology is still in the early stages of development, it has shown promising results in various trials and tests [ERM19, SMS$^+$17, JKMK14].

In [BOR20] author assess the influence of the Dynamic Spectrum Sharing (DSS) feature on an LTE-NR system, analyzing performance outcomes in terms of throughput. The findings indicate that DSS emerges as a valuable attribute for future mobile systems, leading to heightened spectrum efficiency and cost reduction. [GTL14] proposes Dynamic Spectrum Re-allocation (DSR) to enable efficient spectrum sharing between GSM and LTE networks, leading to significant improvements in system capacity and spectrum utilization in multi-RAT environments. DSS is a technology that enables high-speed connectivity in urban environments. It is a mechanism that intelligently allocates frequency bands to devices based on user traffic intensity, reducing interference, improving network performance, and increasing overall system capacity. With the ability to dynamically manage available frequencies, it enables efficient use of spectrum resources and allows for higher data rates [RWDX12].

5.2 Integration of Diverse Spectrum Bands

The integration of sub-6GHz and mmWave frequencies is crucial for enhancing 5G network performance, particularly within Composite Base Stations (CBS). Sub-6GHz frequencies provide reliable coverage and penetration in urban environments, overcoming challenges posed by physical obstructions. In contrast, mmWave frequencies, such as 28GHz, support high data rates essential for applications like HD video streaming and augmented reality, despite their susceptibility to path loss.

This minimizes interference and ensures that high-traffic users receive sufficient bandwidth, maintaining a high Quality of Service (QoS). DSS, combined with CBS, provides an agile solution for managing fluctuating network demands, moving away from traditional static spectrum allocation methods.

This approach maximizes the strengths of both frequency bands, delivering efficient, high-speed connectivity in dense urban settings and meeting the growing demand for advanced applications.

Reference [BOR20] evaluates the impact of DSS on an LTE-NR system, examining its effects on throughput performance. The study suggests that DSS is crucial for future mobile systems, enhancing spectrum efficiency and reducing costs. Meanwhile, reference [GTL14] proposes DSR to facilitate effective spectrum sharing between GSM and LTE networks. This approach aims to enhance system capacity and optimize spectrum utilization in multi-RAT environments. DSS represents a technology pivotal for achieving fast connectivity in urban settings. It operates by intelligently assigning frequency bands to devices according to user traffic levels, thereby mitigating interference, enhancing network performance, and boosting overall system capacity. By dynamically handling available frequencies, DSS optimizes spectrum utilization, facilitating higher data rates, as noted in [RWDX12].

From a practical standpoint, the implementation of CBS with integrated spectrum bands necessitates a comprehensive understanding of urban landscapes, as demonstrated in the study conducted at University College Cork (UCC) in Ireland. The use of OpenStreetMap data to model the propagation of wireless signals in this urban environment underscores the importance of geographical and architectural considerations in network planning. Buildings, foliage, and other urban elements significantly influence signal propagation, especially at mmWave frequencies. Therefore, an analysis of these factors is imperative for the effective deployment of CBS. This approach, grounded in real-world data and simulations, enables network planners to tailor CBS configurations to specific environmental contexts, thereby maximizing coverage and minimizing signal degradation. Such a data-driven strategy is pivotal in ensuring that the theoretical benefits of integrating diverse spectrum bands are fully realized in practical deployments.

In future, emerging applications like holographic communication and remote surgery, which demand unprecedented levels of network performance, will benefit immensely from the groundwork being laid today [DJJ20]. In conclusion, the integration of diverse spectrum bands within CBS frameworks is not just a technical achievement but a strategic foundation for the future of global connectivity, driving innovation and expanding the horizons of what is possible in wireless communication.

5.3 Composite Base Station (CBS) Technology

Composite Base Station (CBS) integrates multiple access technologies, such as LTE, 4G, and 5G, into a single platform. This improves spectrum usage, network efficiency, coverage, and data rates by leveraging both low and high-frequency bands.

CBS simplifies network management, reducing operational costs and complexity, while boosting performance in dense urban areas. It allows for dynamic spectrum allocation, optimizing resources in real time and minimizing interference.

Scalable by design, CBS adapts to emerging wireless technologies, ensuring networks remain capable of handling growing data demands. CBS is a key advancement, enhancing current capabilities and future-proofing wireless networks.

Table 5.1: Simulation configuration parameters for CBS-DSS evaluation

Parameter	Value
params.time.slotDuration	1.0×10^{-3} s (1 ms)
params.spectrumScheduler-Parameters.type	dynamicTraffic (CBS-DSS enabled)
params.calculateIni	TRUE
params.iniOversampling	2
params.regionOfInterest.xSpan	400 m
params.regionOfInterest.ySpan	300 m
params.regionOfInterest.zSpan	50 m
params.carrierDL.center-FrequencyGHz	NR: 28 GHz (mmWave), NR/LTE: 2 GHz (sub-6 GHz)
Parameters.DL.bandwidthHz	800 MHz (NR mmWave), 10 MHz (sub-6 GHz)
Parameters.DL.cqiParameterType	Cqi1024QAM
params.useHARQ	FALSE (to isolate PHY-layer latency effects)
openStreetMapCity.latitude	[51.89158, 51.89415]
openStreetMapCity.longitude	[-8.49683, -8.48926]
openStreetMapCity.wallLossdB	10 dB
antennaLTE.nTX	4
antennaLTE.numerology	NR numerology $\mu = 0$ (15 kHz SCS, sub-6 GHz baseline)
antennaLTE.technology	parameters.setting.Network-ElementTechnology.LTE
antenna5G.nTX	4
antenna5G.numerology	NR numerology $\mu = 1$ (30 kHz SCS)
antenna5G.technology	parameters.setting.Network-ElementTechnology.NR-MN_5G

Simulation Assumptions

We have used OpenStreetMap to obtain the map of UCC as shown in Fig 5.1. Table 5.1 shows the detailed configuration parameters of eNodeBs used in our study. All results presented in Chapter 5 are obtained using the configuration parameters summarized in Table 5.1. Unless stated otherwise, a full-buffer traffic model is assumed for all users in order to evaluate worst-case network load and scheduler behavior. User equipment (UE) nodes are assumed to be static during each simulation run, allowing isolation of radio propagation, spectrum allocation, and scheduling effects. Simulation results are collected after an initial warm-up period and averaged over multiple scheduling intervals to ensure steady-state behavior. Thermal noise and receiver noise figure are modeled using the simulator default settings for both sub-6 GHz and mmWave bands. Channel Quality Indicator (CQI) reporting and modulation and coding scheme (MCS) selection follow standard 3GPP mappings. The dynamicTraffic scheduler prioritizes instantaneous traffic demand while maintaining proportional fairness across active users.

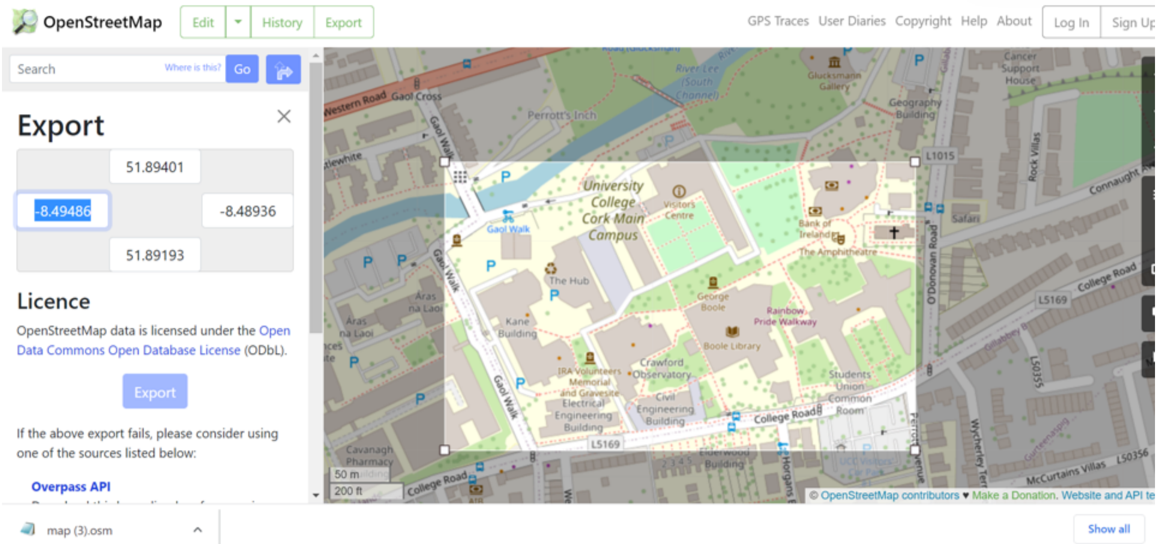

Figure 5.1: Open Streetmap of University College Cork, Ireland Campus

5.3.1 Concept and Design

The evolution of Composite Base Station (CBS) technology marks a significant milestone in the wireless communication landscape. This innovative concept is designed to harmonize multiple wireless access technologies within a single platform, thereby revolutionizing the delivery and management of network services.

The inception of CBS technology can be traced back to the challenges posed by the rapid evolution of wireless networks. With the advent of diverse technologies such as 4G, 5G there emerged a pressing need for a more integrated approach to manage these disparate systems cohesively. CBS was envisioned as a solution to this challenge, aiming to unify various wireless technologies under one umbrella, thereby facilitating a more streamlined and efficient management of the wireless spectrum.

The fundamental design principle of CBS revolves around the seamless integration of multiple wireless technologies, such as sub-6GHz and mmWave frequencies, into a single base station. This integration is pivotal in optimizing

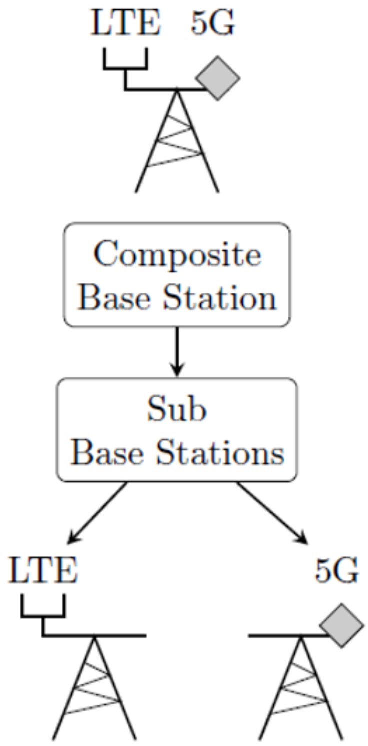

Figure 5.2: Composite Base Station - Image Courtesy [MAD+18]

the use of the available spectrum, enhancing network coverage, and boosting data transmission rates. The design of CBS is underpinned by considerations of compatibility with existing technologies, scalability for future advancements, and flexibility to adjust to dynamic network demands. In this context, Figure 5.2, which depicts the Composite Base Station, serves as a visual representation of how these technologies synergize to optimize spectrum utilization.

The operational advantages of CBS are manifold. By consolidating multiple wireless technologies, CBS ensures an efficient balance between network load, latency, and throughput. This is especially crucial in urban environments where the demand for high-speed connectivity and reliable coverage is paramount. Figures such as Figure **??**, showcasing packet latency in sub-6GHz, and Figure **??**, depicting packet latency in 28GHz, exemplify the enhanced network performance achievable through the implementation of CBS technology.

In the realm of spectrum management, CBS technology stands out for its ability to dynamically allocate frequency bands in response to varying user traffic and network conditions. This dynamic allocation, pivotal in the era of DSS, allows for a more effective utilization of the spectrum, thereby reducing interference and enhancing overall network performance. The role of CBS in spectrum management is not only about enhancing efficiency but also about adapting to the evolving landscape of wireless communication needs.

The integration of CBS technology plays a crucial role in network performance optimization. By managing multiple wireless technologies through a single platform, CBS can effectively balance network load, reduce latency, and enhance throughput.

CBS technology improves connectivity, providing better coverage and data rates for enhanced user experience. Its scalability allows adaptation to future technologies, ensuring long-term relevance.

CBS is beneficial in urban areas for managing high traffic and in rural settings for wider coverage, bridging the digital divide. By integrating multiple technologies, CBS reduces operational costs, simplifying infrastructure and lowering maintenance.

It also promotes environmental sustainability by minimizing energy consumption. However, challenges include technical integration, compatibility with existing networks, and transition management. Ongoing research aims to enhance CBS capabilities, especially in the context of 5G networks, ensuring effective integration for improved performance.

An important metric in assessing the effectiveness of CBS technology is user throughput. Figure 5.10 and Figure 5.11, illustrate user throughput as reported by feedback, and the highest possible CQI in sub-6GHz and 28GHz cases, respectively, can be used here. These figures demonstrate how CBS technology can enhance network capacity and improve user data rates.

Looking ahead, the evolution of CBS technology will continue to shape the future of wireless communications. With ongoing advancements in AI, machine learning, and other emerging technologies, CBS stations are poised to become even more intelligent and efficient. The continued development of CBS technology will play a crucial role in meeting the ever-growing demands of modern society for fast, reliable, and ubiquitous wireless connectivity.

5.3.2 Integration of 5G and 4G LTE

Integrating 5G and 4G LTE through Composite Base Station (CBS) technology combines 4G's broad coverage with 5G's high-speed capabilities, ensuring reliable connectivity. CBS enables seamless interoperability, dynamic resource allocation, and better spectrum management, particularly in urban environments. This integration reduces latency for real-time applications and supports innovations in IoT, smart cities, and emergency response. However, infrastructure upgrades and regulatory challenges must be addressed to fully realize CBS's potential in future wireless communications.

5.4 Dynamic Spectrum Scheduling (DSS)

Dynamic Spectrum Scheduling (DSS) enhances spectrum management by dynamically allocating resources based on real-time network conditions and user demand. It uses advanced algorithms, including AI and machine learning, to predict traffic patterns and optimize bandwidth usage. DSS adjusts spectrum allocation based on traffic load, ensuring efficient spectrum use and minimizing interference, particularly in dense urban environments. By enabling the flexible coexistence of 4G and 5G networks on the same frequency bands, DSS ensures seamless service during the 5G transition. As DSS evolves, its increasing automation and predictive capabilities will further optimize resource allocation, supporting the growing demand for high-speed connectivity in future wireless networks.

5.4.1 Dynamic Spectrum Scheduling Operation and Resource Allocation Mechanism

Dynamic Spectrum Scheduling (DSS) is a time-domain spectrum sharing mechanism that allows LTE and NR to operate over the same carrier by dynamically assigning resources according to instantaneous traffic demand. In this work, DSS is not introduced as a new scheduler; rather, it is modeled using the Vienna 5G System Level Simulator (V5SLS) dynamic traffic spectrum scheduler, which continuously monitors user traffic demand and allocates available resources accordingly—users with higher traffic intensity are granted more resources while low-demand users receive fewer—so that the shared spectrum is utilized efficiently without static partitioning. This demand-driven allocation provides a practical spectrum-management approach in mixed LTE/NR deployments where traffic distribution fluctuates over time. The purpose here is therefore to evaluate the system-level impact of DSS-enabled dynamic resource allocation (e.g., on latency, throughput, and SINR) in the considered urban/campus deployment scenario, under the simulator's traffic-intensity-based scheduling setting.

5.4.2 Mechanism and Application

Dynamic Spectrum Scheduling (DSS) utilizes advanced algorithms to monitor network traffic patterns, enabling adaptive spectrum allocation that meets real-time demands in densely populated areas. By dynamically shifting spectrum between 5G and 4G LTE, DSS optimizes resource use and alleviates congestion, resulting in improved data speeds, reduced latency, and enhanced user satisfaction. Its role in the Internet of Things (IoT) ensures efficient communication among numerous connected devices while promoting energy efficiency and cost savings for network operators. Overall, DSS is crucial for future wireless communication, enhancing connectivity and operational efficiency.

5.5 Urban Network Performance Enhancement

The enhancement of urban network performance is an increasingly critical aspect of modern telecommunications, given the unique challenges and demands presented by densely populated urban environments. As cities continue to grow and become more technologically integrated, the need for robust, efficient, and high-capacity wireless networks has never been more pronounced. Urban areas, with their high user density, a diverse range of wireless devices, and complex physical landscapes, present unique challenges for network performance. These challenges include managing high traffic volumes, ensuring consistent coverage amidst high-rise buildings and urban canyons, and maintaining high data speeds amidst a plethora of competing signals.

5.5.1 Study in Urban Environments

Composite Base Stations (CBS) and Dynamic Spectrum Scheduling (DSS) enhance coverage and capacity while advanced technologies like beamforming and small cells improve signal propagation. Successful examples in cities like Seoul and New York highlight the importance of tailored deployment strategies. Addressing latency, reliability, and efficient spectrum utilization is crucial for supporting smart city applications, requiring ongoing advancements in technology and community engagement to maximize 5G's potential.

5.5.2 University Campus Case Study

The study at University College Cork (UCC) examines an urban campus deployment using simulations based on OpenStreetMap [Ope] data and the Vienna 5G System Level Simulator. The case study provides a representative environment for evaluating system-level performance metrics, including throughput, latency, SINR, and BLER, under mixed LTE and 5G operation. It is noted that millimeter-wave evaluations at 28 GHz apply exclusively to 5G NR users, while LTE users are confined to sub-6 GHz operation and are included only to illustrate coexistence and scheduling within the CBS architecture.

5.6 Spectrum Allocation

As discussed in the context of CBS and the integration of diverse spectrum bands, DSS allows for the coexistence of 5G and LTE networks on the same frequency band. This technology dynamically allocates spectrum resources based on real-time traffic demands. Figure 5.3 illustrates dynamically allocating spectrum resources between 5G and LTE networks. It begins with the centralized spectrum scheduler, which assesses real-time traffic demand, user requirements, and interference levels to allocate resources efficiently. These decisions are passed to sub-schedulers specific to 5G and LTE, which further optimize resource allocation within their networks. The output, represented by color-coded resource blocks, showcases how spectrum is dynamically and fairly distributed to users, enabling improved throughput in a heterogeneous network environment. Figure 5.4 depicts the simulation results from a study focusing on high-speed connectivity in urban environments using Composite Base Stations (CBS) equipped with Dynamic Spectrum Scheduling (DSS). The red dots represent CBSs, combining 5G and LTE technologies, while green and blue dots represent 5G and LTE users, respectively, connected dynamically based on traffic demand. Conducted in a campus-like urban layout, the study used OpenStreetMap data to model realistic scenarios, emphasizing improved performance metrics such as throughput and reduced latency. The dynamic spectrum allocation mechanism ensures optimal resource utilization for seamless and efficient communication in a dense urban setup.

Figures 5.5, 5.6, and 5.7, display present publicly available 5G coverage maps for the region of interest, as provided by the Commission for Communications Regulation in Ireland (An Coimisiún um Rialáil Cumarsáide) [CfCR]. These

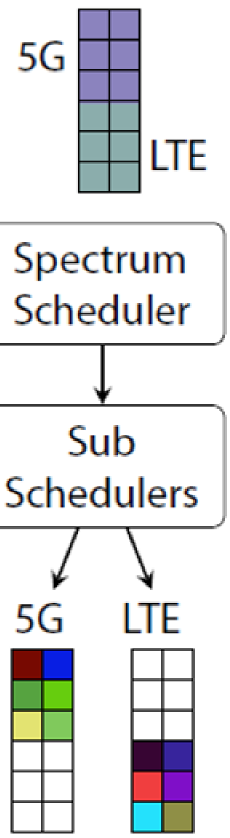

Figure 5.3: Dynamic Spectrum Scheduler for 5G and LTE Resource Allocation - Image Courtesy [MAD+18]

are included solely to provide contextual information about reported outdoor 5G availability in the selected area. These maps do not allow conclusions to be drawn regarding spectrum allocation strategies or performance, as observed coverage patterns may be influenced by multiple factors such as base-station density, site placement, antenna configuration, and propagation conditions. Accordingly, no performance or spectrum-related conclusions are inferred from these figures, and all quantitative analysis in this chapter is based on simulation results.

Figure 5.5 showcases the 5G coverage provided by Vodafone within a designated Region of Interest (RoI). Vodafone's coverage map stands out with multiple areas showing "Very Good" 5G connectivity near University College Cork. This indicates superior network deployment and signal quality compared to Eir and Three. The higher signal strength likely supports better throughput and lower latency, making Vodafone a reliable choice for demanding applications.

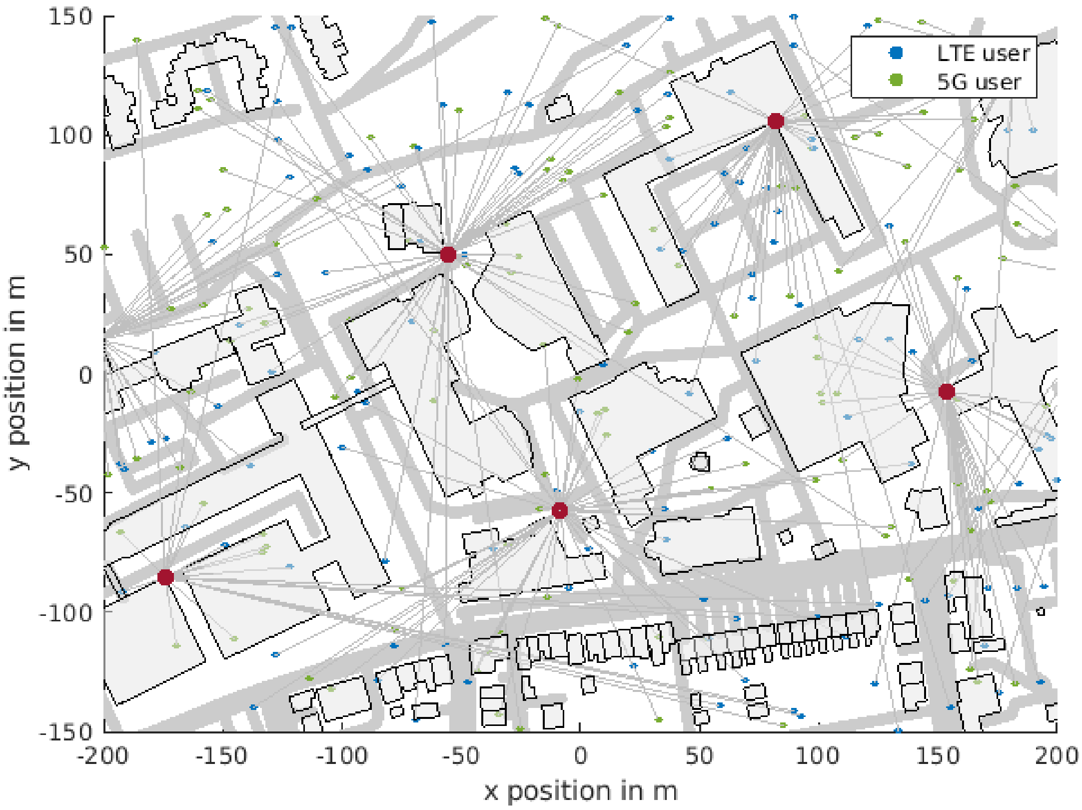

Figure 5.4: Figure illustrates a simulated urban network deployment showing the coexistence of 5G and LTE technologies. The red dots represent base stations (BS), while green and blue dots denote 5G and LTE users, respectively, connected to their nearest base stations, as indicated by gray connectivity lines

However, areas with "Fair" or "Fringe" coverage suggest that localized improvements are still necessary to ensure uniform quality across the campus.

This figure offer insights into how specific frequency bands are utilized to optimize coverage and performance for 5G users in a particular area. This visualization underlines the importance of selecting the right spectrum bands for 5G to ensure wide-reaching and robust connectivity. In contrast, Figure 5.6 focuses on the 5G coverage offered by another major network operator, Three, within the same RoI. The Three network coverage map reveals a similar trend, with most of the campus covered under "Good" 5G connectivity, but only limited regions achieving "Very Good" status. The consistency in coverage across the area implies effective spectrum allocation, though the absence of stronger signals in critical locations could impact high-bandwidth applications. Three might need to expand its network capacity or refine its existing infrastructure to compete more effectively in terms of quality. This comparison is vital as it illustrates the variability in network performance and coverage based on different spectrum

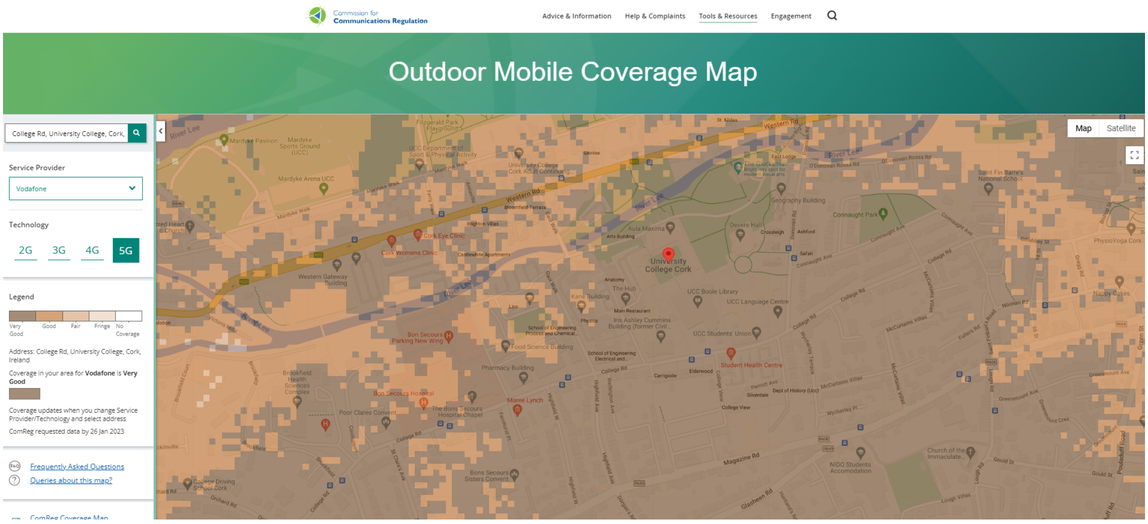

Figure 5.5: 5G Coverage by Vodafone in Region of Interest: The map depicts Vodafone's outdoor mobile 5G coverage around University College Cork. The coverage quality, represented by the same gradient scale, indicates "Very Good" connectivity in several areas near the campus. [Date: 10-May-2023]

allocation strategies adopted by various operators. It highlights how each network's unique approach to spectrum allocation can lead to differences in service quality and user experience. Figure 5.7 shows the 5G coverage provided by Eir, again within the same RoI, further emphasizing the diverse ways in which spectrum allocation can be approached. The Eir coverage map at University College Cork demonstrates predominantly "Good" 5G connectivity, with minimal areas falling into fringe or no coverage zones. This indicates Eir's moderate presence in the area, providing sufficient connectivity for urban campus environments. However, the lack of "Very Good" coverage in critical regions suggests room for improvement in network optimization and infrastructure deployment to enhance user experience, especially in high-traffic zones.

5.6.1 Efficient Use of Bands

This subsection outlines the motivation for considering both sub-6GHz and mmWave frequency bands within the simulation framework. Sub-6GHz bands provide more robust coverage characteristics, while mmWave bands offer higher data rates at the expense of increased path loss and sensitivity to blockage. In this chapter, these bands are evaluated within a Composite Base Station configuration using Dynamic Spectrum Scheduling as implemented in the simulator. No conclusions regarding optimal spectrum allocation strategies or operator-specific performance are inferred beyond the controlled simulation re-

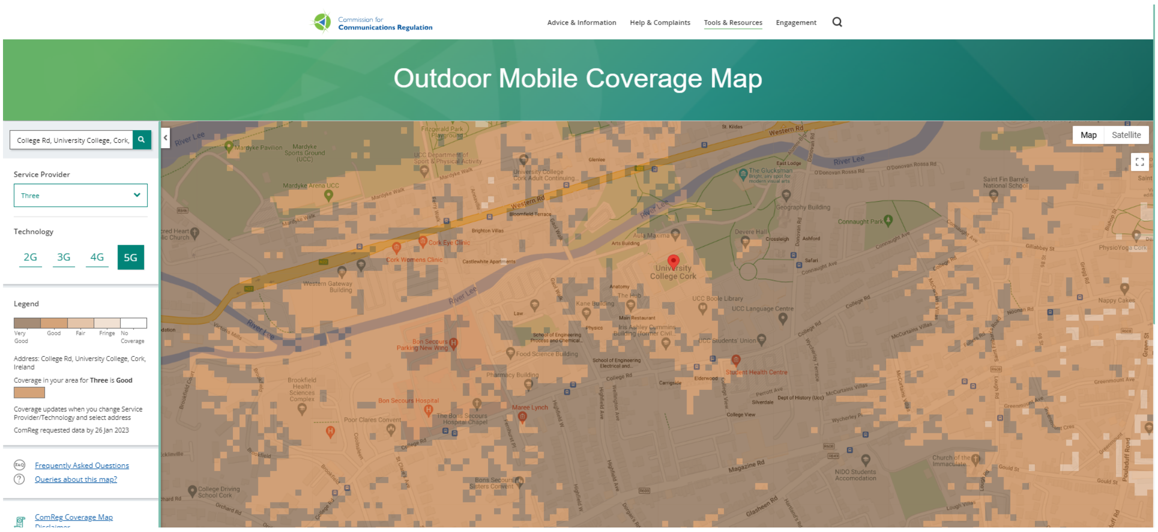

Figure 5.6: 5G Coverage by Three in Region of Interest: This map displays the outdoor mobile coverage offered by Three at University College Cork, focusing on 5G connectivity. A color-coded scale indicates the quality of coverage, with most areas around the campus showing "Good" coverage. [Date: 10-May-2023]

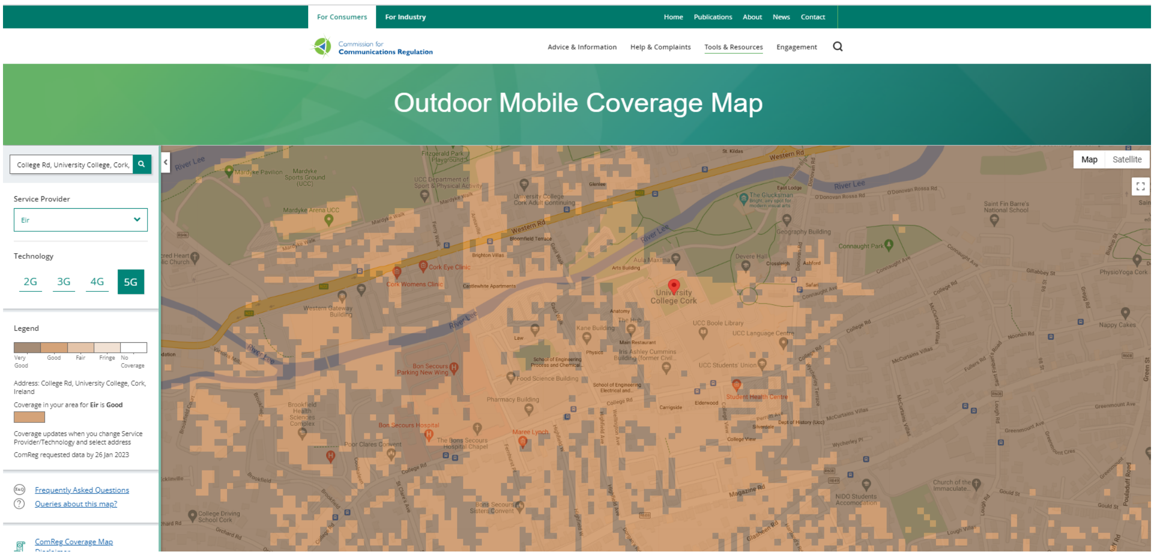

Figure 5.7: 5G Coverage by Eir in Region of Interest: The map illustrates the outdoor mobile coverage provided by Eir at University College Cork, highlighting 5G connectivity. The color gradient represents coverage quality, ranging from "Very Good" to "No Coverage," with this specific area showing predominantly good coverage. [Date: 10-May-2023]

sults presented later in the chapter.

5.7 Impact on User Experience & Service Delivery

The integration of Composite Base Stations (CBS) and Dynamic Spectrum Scheduling (DSS) revolutionizes 5G networks by enhancing coverage, capacity, and user experience, especially in dense urban areas. CBS supports multiple frequency bands, enabling seamless connectivity for more users, while DSS dynamically allocates spectrum to reduce congestion and latency, crucial for applications like IoT and HD streaming. These technologies provide adaptable, high-quality connectivity for current and future digital demands, ensuring networks can scale efficiently with evolving user needs.

5.7.1 User-Centric Analysis

Composite Base Stations (CBS) and Dynamic Spectrum Scheduling (DSS) significantly enhance 5G user experiences, especially in dense urban areas. CBS boosts network coverage and capacity, addressing issues like dropped connections and slow speeds, while enabling seamless activities from web browsing to HD streaming and online gaming. DSS ensures stable performance by dynamically allocating spectrum, minimizing latency, and reducing interference, crucial for real-time applications like video conferencing and live streaming. Together, CBS and DSS create responsive, reliable networks that align with the demands of modern urban lifestyles.

The comparative analysis of 5G coverage by different network operators, as illustrated in 5.5, 5.6, and 5.7, offers profound insights into the user experience. These figures are not just statistical representations; they tell a story of how varying levels of network quality and accessibility in different regions affect the daily digital interactions of users. This visual data brings to light the disparities that exist and demonstrates the tangible impact that CBS and DSS technologies have on people in these areas. The incorporation of such comparative data is more than a technical exercise; it's a narrative tool that provides a clear, visual testament to the benefits and limitations of current 5G network deployments, all from a user-centric perspective.

The SINR is a measure of the quality of a wireless communication link. It is defined as the ratio of the received signal power to the sum of the interference power and the noise power [HS13, Ham09a]. Mathematically, it is expressed as:

$$\text{SINR} = \frac{P_{\text{signal}}}{P_{\text{interference}} + P_{\text{noise}}}, \tag{5.1}$$

where P_{signal} is the power of the desired signal, $P_{\text{interference}}$ is the total power of the interfering signals, and P_{noise} is the power of the noise.

The comprehensive analysis of Signal-to-Interference-Noise Ratio (SINR) as showcased in Figures 5.8, and 5.9 takes us on a deep dive into the intricate dynamics of network performance, scrutinizing how different frequency bands and signal types shape the landscape of 5G connectivity. This exploration is much more than a technical review; it's akin to peeling back the layers of a digital ecosystem to reveal the core elements that determine the quality of our everyday digital interactions.

Figure 5.8 shows the User Wideband SINR for sub-6GHz frequencies. This SINR measure considers both the desired signal and the interference from other cells in the network. The graph plots SINR values against the Empirical Cumulative Distribution Function (ECDF), providing a visual representation of the SINR distribution. In the sub-6GHz case, the SINR values are fairly distributed, with most users achieving an SINR above -5 dB (around 0.9 ECDF). An SINR of 10 dB or higher is considered ideal for high-data-rate applications. The relatively high SINR across a large portion of users indicates that sub-6GHz frequencies provide better coverage and reliability, ensuring that the majority of users can experience stable and high-speed connectivity.

Sub-6GHz offers better SINR performance across a larger portion of the user base, ensuring a more consistent and reliable experience. In contrast figure 5.9, the 28GHz mmWave frequencies show a wider range of SINR values, with many users experiencing lower SINR, which could result in lower performance and reliability.

Now we do a thorough examination of throughput, It is the amount of data that is transmitted over a network within a given time, and it is an important performance metric for all communication systems. The V5SLS calculates the throughput at both user and cell levels, and it is influenced by factors such as modulation and coding scheme, channel conditions, and network size. The user-level throughput in the V5SLS is calculated as:

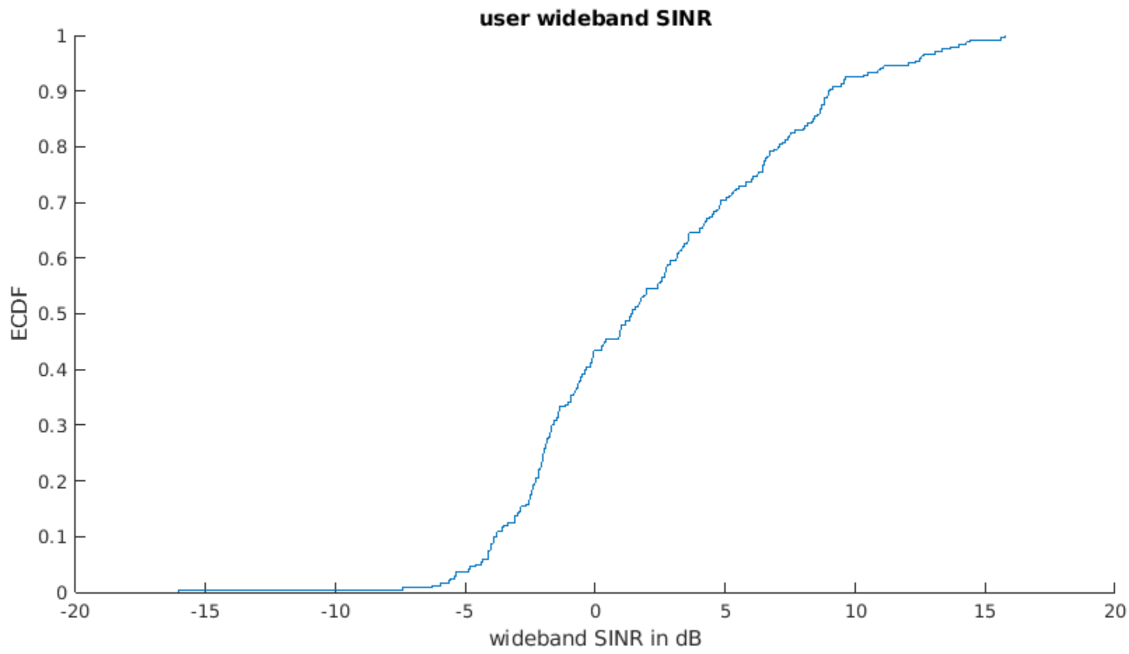

Figure 5.8: User wideband SINR (sub-6GHz)

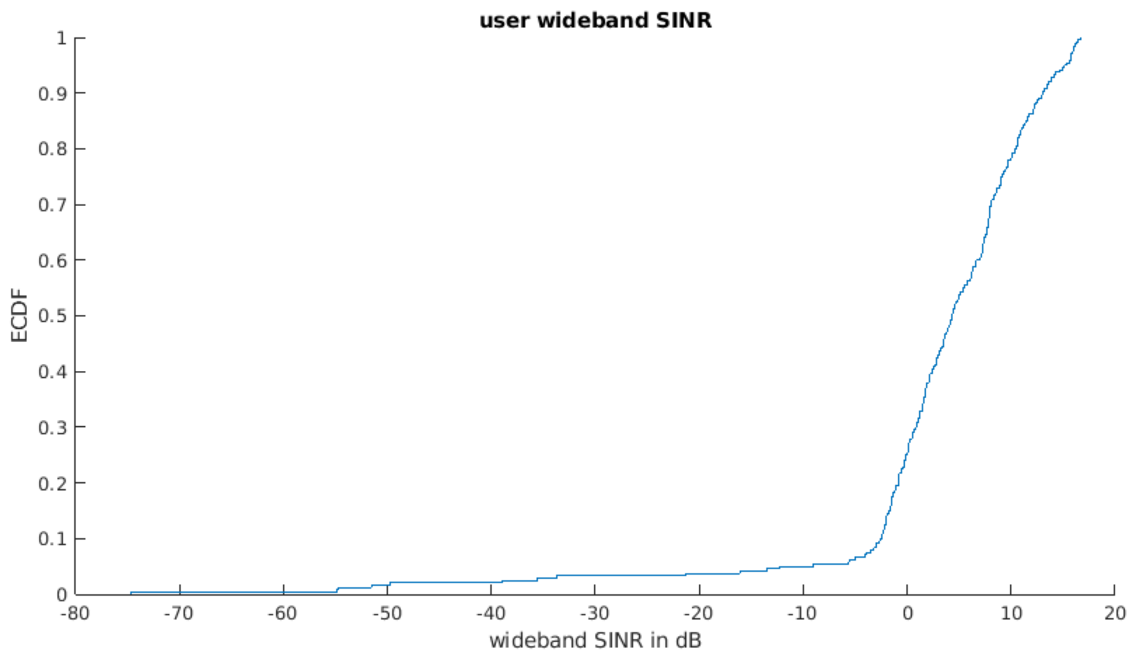

Figure 5.9: User wideband SINR (28GHz)

$$T_u = \frac{\sum_{k=1}^{K} \sum_{n=1}^{N_k} B_k \log_2(1 + \gamma_{k,n})}{T}, \tag{5.2}$$

where T_u is the user-level throughput, K is the number of users, N_k is the number of data blocks transmitted by user k, B_k is the size of a data block transmitted by user k, $\gamma_{k,n}$ is the signal-to-interference-plus-noise ratio (SINR) of the n-th block transmitted by user k, and T is the simulation time.

Figure 5.10 shows the throughput for users in the sub-6GHz band, using Channel Quality Indicator (CQI) values reported by feedback from the network. It is noted that the throughput values in figure correspond to per-user uplink throughput under a dynamicTraffic scheduler, where available bandwidth is shared among multiple active users. These values should not be interpreted as cell-level capacity.

Throughput is a key performance metric in communication systems and indicates the amount of data successfully transmitted over the network in a given time. The throughput values for sub-6GHz users show a relatively stable and efficient performance, particularly for users with high CQI values. The system ensures that users with higher data demands (and better channel conditions) are allocated more resources, leading to higher throughput. This suggests that the network effectively adjusts to varying user needs, providing adequate bandwidth to users while avoiding congestion.

This figure 5.11 presents the user throughput for the 28GHz mmWave frequency band, also based on CQI values. The throughput for mmWave frequencies is significantly higher than for sub-6GHz frequencies, especially for users with high CQI values. This shows that while mmWave frequencies provide higher potential throughput due to the wide bandwidth available, the actual performance can vary depending on the user location and the environmental conditions. In ideal condition, the throughput gains from mmWave frequencies are evident, but the system may experience performance degradation in more challenging environments, as discussed previously with respect to SINR. The data reveals that at lower SINR thresholds, ZF beamforming significantly improves throughput by reducing interference compared to MRT. However, at higher SINR values, MRT slightly outperforms due to its ability to maximize signal strength. The throughput values are consistent with the findings of [Ham09b] and are supported by simulations conducted under similar conditions. Additionally, while User Wideband SINR provides a precise evaluation of transmission performance, a User Wideband Lite approximation can be used in simulations to efficiently estimate performance trends while reducing computational complexity.

The throughput for users in the sub-6GHz band remains relatively consistent and stable, while the throughput in the 28GHz band is much higher but dependent on good channel conditions. The high throughput for mmWave frequencies highlights their potential for high-bandwidth applications, but this comes at the cost of reduced reliability in less than ideal conditions.

Figure 5.12 shows the user throughput for the sub-6GHz frequency band, considering the Channel Quality Indicator (CQI) values reported by the feedback in the system. The throughput is calculated by considering various factors, such as the channel conditions, interference, and available spectrums. The plot compares the throughput for both 5G and LTE users, with higher throughput achieved by users with better CQI values. The throughput in the sub-6GHz band demonstrates a steady and relatively stable performance. For both 5G and LTE users, throughput is higher for users with better CQI, which typically indicates better channel conditions and lower interference. The throughput for 5G users is generally higher compared to LTE users, indicating the superior data rates achievable with 5G technology. Overall, sub-6GHz frequencies provide reliable throughput for most users, with minimal performance fluctuation, particularly for LTE users.

Figure 5.13 shows the user throughput for the 28GHz mmWave frequency band, also based on CQI values. Similar to figure 5.12, throughput is calculated based on the feedback provided by the network, which adjusts the resource allocation depending on the channel quality and user demand.

The throughput for mmWave frequencies is significantly higher than for sub-6GHz frequencies, particularly for users with high CQI values. This is due to the large available bandwidth at 28GHz, which enables faster data rates. However, while the throughput is high, it can be seen that some users experience lower throughput compared to others. This could be due to poor channel conditions, interference, or the susceptibility of mmWave signals to obstructions such as buildings and other obstacles. The higher throughput in mmWave networks highlights the potential of mmWave for applications requiring very high bandwidth, such as 4K video streaming, augmented reality, or other data-intensive services.

5.7.1.1 Throughput Comparison

Sub-6GHz : Users in the sub-6GHz band experience more consistent and reliable throughput. This band offers a good balance between coverage and data rate, with fewer fluctuations in throughput. LTE users experience lower throughput compared to 5G users, reflecting the enhanced data rates of 5G technology, but the throughput levels are generally stable across different users.

28GHz : In the mmWave band (28GHz), throughput is higher, but more variable. While users with good CQI can achieve very high throughput, some users with lower CQI values experience reduced throughput. This variability can be attributed to the propagation characteristics of mmWave frequencies, which are more susceptible to path loss, interference, and blockage by physical obstacles.

5.7.1.2 CQI and Throughput Relationship

Sub-6GHz: The relationship between CQI and throughput is fairly straightforward—users with higher CQI (better channel quality) achieve higher throughput, and vice versa. Since sub-6GHz offers better propagation characteristics and signal reliability, the throughput performance is more predictable and consistent.

28GHz: The CQI-to-throughput relationship in the 28GHz case shows greater variability. While some users can achieve significantly higher throughput, others experience lower throughput due to the environmental challenges faced by mmWave signals, such as signal attenuation and interference. This variability highlights the trade-off between high throughput and reliability in mmWave communications.

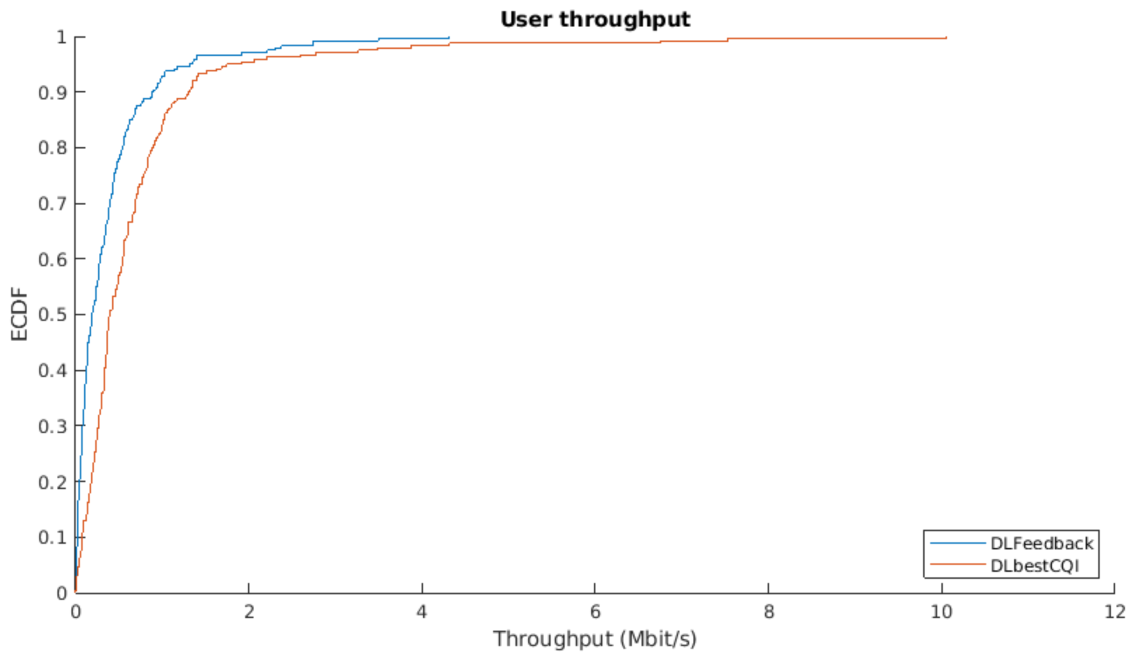

Figure 5.10: User Throughput as reported by feedback and highest possible Channel Quality Indicator (CQI) (sub-6GHz).

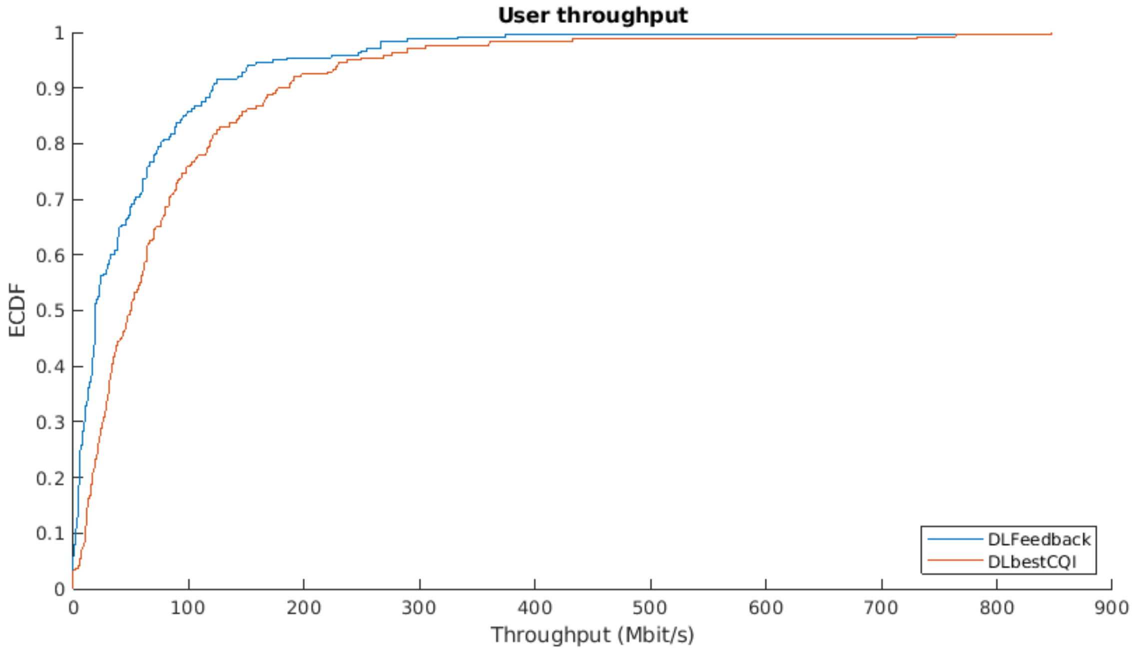

Figure 5.11: User throughput (28GHz)

5.7.2 Optimizing Network Performance with CBS and DSS Technologies

The analysis emphasizes how Composite Base Stations (CBS) and Dynamic Spectrum Scheduling (DSS) play a pivotal role in optimizing network performance. These technologies aren't just technical solutions but crucial elements that enable a seamless user experience by enhancing throughput and minimizing latency. They ensure smoother digital interactions, critical in an age demanding fast data exchange. This study highlights the importance of these technologies not only for experts but also for everyday users, as network performance directly impacts daily digital interactions.

Incorporating real-world feedback and case studies bridges the gap between theory and practical application, grounding technological advancements in real-life scenarios. This approach ensures that future developments better align with user expectations.

Additionally, addressing the challenges of implementing user-centric principles in 5G networks is vital. The study explores technical, logistical, and regulatory barriers that need overcoming to maintain a user-focused approach in network design and deployment. Looking forward, ongoing advancements in CBS and DSS technologies will continue to shape user experiences, improving satisfaction and service quality.

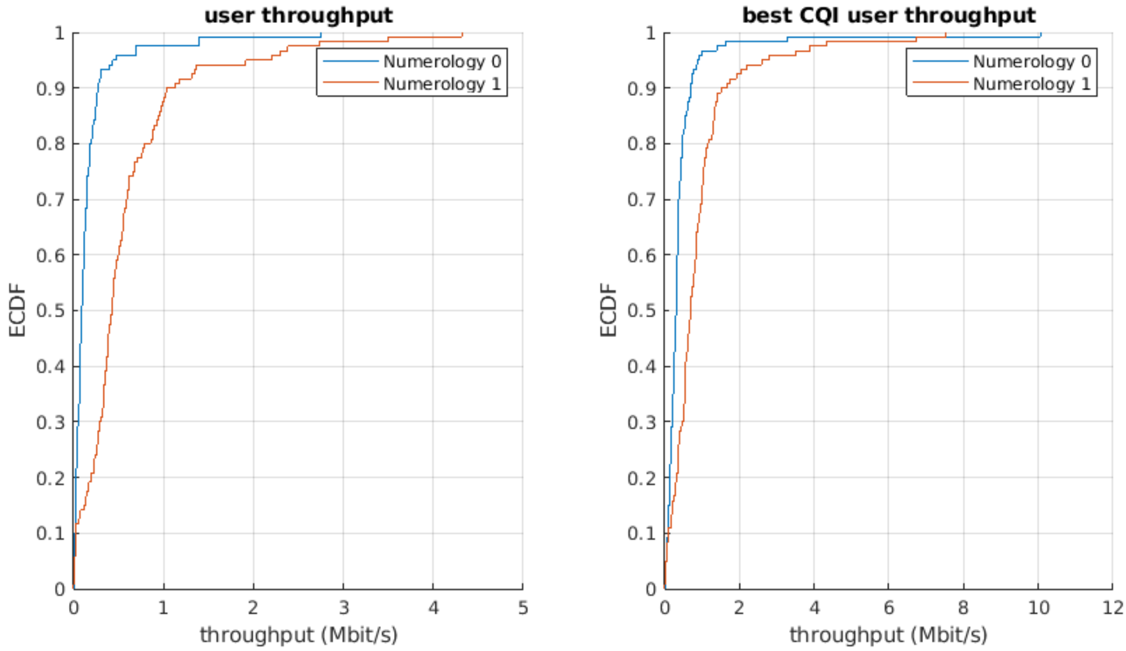

Figure 5.12: User Throughput (sub-6GHz): The user throughput and the best user throughput based on the CQI are displayed for both 5G and LTE users. The Numerology 1 corresponds to the 5G users, while the Numerology 0 corresponds to the LTE users. The user throughput is calculated considering the channel conditions, interference, and noise. The user with the highest CQI value is selected to determine the best CQI user throughput and an increase in throughput is observed. LTE users → Better per-user throughput ECDF because of lighter traffic demand. 5G users → Lower per-user throughput ECDF because of higher demand + shared bandwidth.

The role of policy and regulatory frameworks is also crucial. A supportive regulatory environment will allow CBS and DSS technologies to reach their full potential, ensuring that they not only improve technology but also prioritize the user experience in 5G network evolution.

5.7.3 Service Quality Improvement

Service quality in 5G networks is essential for effective digital communication and IoT. It includes speed, latency, reliability, and capacity. High-speed connectivity enables quick data transfer, low latency supports real-time applications, reliability ensures consistent service, and high capacity accommodates numerous devices. Composite Base Stations (CBS) improve coverage and capacity, addressing deployment challenges and enhancing urban network performance. Dynamic Spectrum Scheduling (DSS) optimizes spectrum usage, ensuring reliability and preventing congestion.

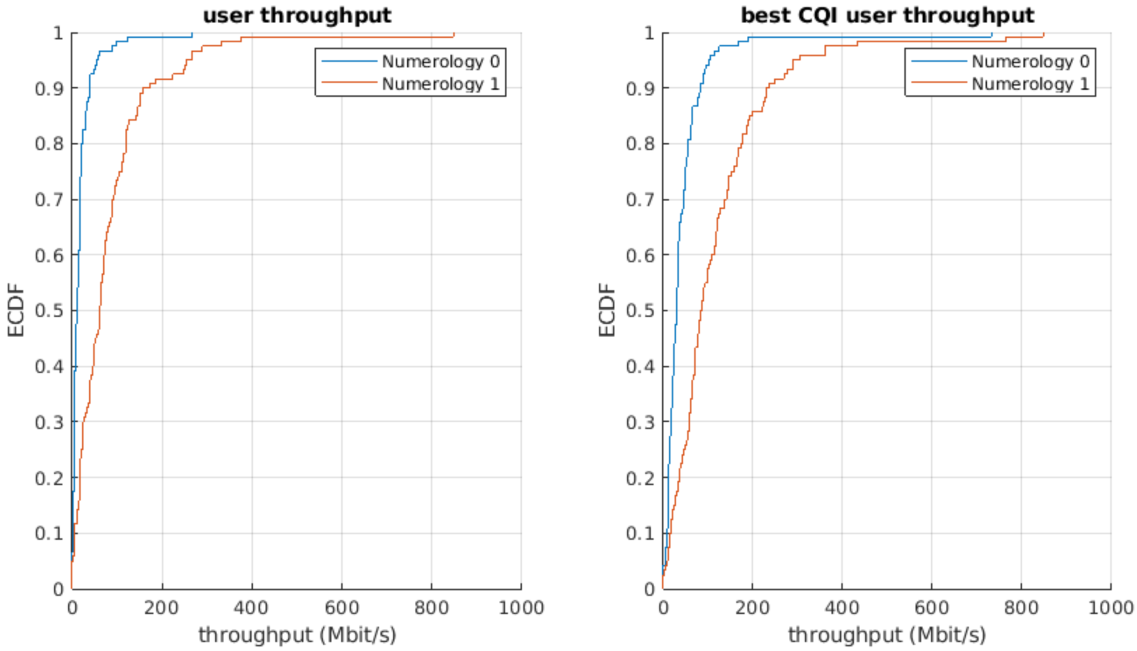

Figure 5.13: User Throughput (28 GHz): Results correspond to 5G NR users operating in the mmWave band. LTE users are shown only for comparative load and scheduling behavior within the CBS framework and operate exclusively in sub-6 GHz bands.

Analyzing 5G coverage, shown in Figures 5.5, 5.6, and 5.7, helps understand the impact of different providers' rollouts on service quality. The coverage maps highlight areas of strong and weak coverage, which directly affects user experience. The case study from University College Cork (Figure 5.1) illustrates how 5G deployment improves service quality in an urban environment, providing better connectivity for students and faculty.

The Signal to Interference and Noise Ratio (SINR), analyzed in Figures **??** to 5.9, directly affects data transmission speed and reliability. SINR variation across the network highlights areas with better or worse service quality. User throughput, depicted in Figures 5.10 to 5.13, is a key factor in determining network efficiency, especially under high data demand. High throughput is crucial for maintaining service quality during peak usage.

Table 5.1 outlines eNodeB configuration parameters that enhance service quality. These parameters, such as transmission bandwidth and carrier frequency, play a vital role in improving network performance. Emerging technologies like AI, machine learning, and network automation will continue to drive service quality improvements by enabling smarter, more efficient network management.

Maintaining high service quality in 5G networks requires addressing challenges such as increasing data demand, ensuring coverage consistency, and managing security concerns. Overcoming these challenges demands continuous innovation, investment in infrastructure, and adherence to best practices in network management.

The evolution of CBS and DSS technologies, as demonstrated in the figures and case studies, is essential for improving the service quality of 5G networks. These technologies are key to meeting the demanding requirements of modern digital communication, ensuring that 5G networks can support a wide range of applications and deliver a superior user experience.

5.8 Conclusion

This chapter highlighted the potential of Composite Base Stations (CBS) that jointly support LTE and 5G NR in a dense urban/campus setting. Using a Dynamic Spectrum Scheduling (DSS) mechanism and evaluating both sub-6 GHz and mmWave (28 GHz) operation, the study examined how multi-band CBS deployments influence key service-quality metrics and overall user experience. The evaluation was conducted on a representative urban campus scenario (UCC, Ireland) using OpenStreetMap-derived geography, enabling a realistic propagation and deployment context.

Across the considered metrics, the mmWave tier exhibited clear advantages in terms of higher achievable throughput under favorable channel conditions, while sub-6 GHz provided more robust coverage and more stable performance. The combined CBS/DSS setup therefore illustrates how capacity and coverage can be balanced in heterogeneous deployments, while also exposing important trade-offs related to blockage, path loss, and interference in mmWave operation. Overall, the results highlight the conditions under which CBS deployments exhibit favorable performance characteristics, as well as the constraints associated with integrating sub-6 GHz and mmWave layers.

Dynamic spectrum scheduling, as modeled through the simulator's traffic-intensity-based scheduler, enables flexible resource utilization in mixed LTE/NR scenarios. Rather than proposing a new scheduling algorithm, this chapter demonstrates how such demand-driven spectrum allocation influences latency, throughput, and SINR trends in a realistic campus environment. The analysis therefore provides a system-level characterization of CBS behavior under dynamic resource-allocation assumptions.

Overall, the findings contribute to an improved understanding of how CBS architectures combining sub-6 GHz and mmWave layers behave in dense urban settings. While the results are scenario-specific and simulation-based, they establish a foundation for future studies that may explore alternative urban morphologies, traffic models, and more advanced DSS mechanisms, as well as extensions toward emerging beyond-5G and 6G network architectures. Future research could further investigate these aspects in heterogeneous and ultra-dense deployments.

Chapter 6

D2D Algorithm for Blind Spot Removal in IoT Smart Industry

> *"Our view is that sustainable IoT is about long-term deployments without environmental impact. We want devices to last for decades, be powered efficiently, and use sustainable materials to avoid creating electronic waste."*
>
> *Professor Dirk H J Pesch*

The concept of Industry 4.0 has revolutionized the landscape of industrial production, marking a significant shift towards automation and digitization in manufacturing and other sectors. The concept of "Industry 4.0" was initially coined in Germany in 2011 under a government initiative aimed at advancing the digital transformation of manufacturing. Since its introduction, this concept has garnered global recognition and widespread adoption across various industries [LFK$^+$14]. At the heart of this transformation process is the integration of cutting-edge technologies like the IoT, AI, robotics, and cloud computing. Among these technologies, D2D communication emerges as a pivotal component.

D2D communication, a method where devices establish direct connections without intermediary infrastructure plays a crucial role in enhancing operational efficiency and responsiveness in the dynamic environment of Industry 4.0. This method of communication not only streamlines interactions between various components within the industrial setup but also contributes to the creation of smart industries that are automated, flexible, and highly responsive to market demands [Gho20]. In smart industry environments, where rapid data exchange and real-time decision-making are paramount, D2D communication offers several advantages. It allows for quicker and more efficient data transfer between devices, bypassing the need for data routing through a central network hub. This direct communication method is particularly beneficial in scenarios where immediate data processing and action are required, such as predictive maintenance, operational insights, and safety monitoring.

One of the significant challenges in implementing Industry 4.0 technologies is addressing the issue of blind spots in wireless network coverage. Blind spots are areas where the wireless signal is weak or non-existent, which can lead to connectivity issues for devices located within these zones. In the context of Industry 4.0, the presence of blind spots can significantly hinder the performance of smart industries [MTZS22]. D2D communication provides a promising solution to this problem. By allowing devices to establish direct connections, D2D communication can extend network coverage to these blind spots, ensuring continuous and reliable connectivity across the entire industrial environment [SS17]. This capability is particularly crucial for maintaining uninterrupted operations and ensuring worker safety in areas with limited network coverage. The integration of D2D communication into industrial processes significantly contributes to overall connectivity availability. By facilitating direct interactions between devices, D2D minimizes the reliance on central network infrastructure, thus reducing latency and improving data transmission rates. This improvement is particularly noticeable in environments with a high density of IoT devices, where network congestion can be a critical issue. Furthermore, the use of D2D communication can lead to better resource utilization, as devices can share information and resources more effectively, leading to optimized operational workflows and increased productivity. Looking towards the future, the advancements in D2D communication technology hold great promise for Industry 4.0. As D2D technology evolves, it is expected to support more complex and varied industrial applications, enhancing its utility in smart manufacturing processes. The integration of advanced AI algorithms with D2D communication can lead to more intelligent and autonomous industrial systems, capable of self-optimization and predictive maintenance.

6.1 Technological Foundations of D2D Communication

Network providers must address potential coverage gaps to ensure reliable service for their customers. Achieving dependable coverage requires several strategies, such as deploying additional small cells, utilizing beamforming technologies, employing D2D relays to enhance wireless coverage, and applying network optimization methods to improve both coverage and connectivity [IRH+21]. D2D communication represents a pivotal innovation in the realm of 5G networks, marking a significant shift from traditional cellular communication models. At its core, D2D communication facilitates direct interaction between devices within their proximity, bypassing the need for data routing through a central base station or network infrastructure. In the context of the IIoT and smart factories, D2D communication plays an instrumental role in fortifying connectivity and ensuring continuous data flow, even in challenging environments. Smart factories are characterized by a plethora of IoT devices, such as sensors and actuators, all requiring seamless and uninterrupted communication to function optimally. D2D communication enables these devices to interact directly, thereby mitigating potential disruptions caused by obstacles or distance from the base station. D2D communication significantly enhances network performance and efficiency in industrial settings. By allowing direct device interactions, it alleviates the burden on the central network infrastructure. This decentralization of communication leads to a reduction in network congestion and an improvement in data transmission rates. Furthermore, D2D communication can use various frequency bands, including mmWave and THz frequencies, which are integral to 5G technology. This flexibility in frequency band utilization facilitates efficient use of the spectrum and enhances network capacity. Additionally, the proximity-based nature of D2D communication ensures that data transmission is not only rapid but also energy efficient, an essential consideration for sustainable industrial operations. This energy efficiency is particularly vital in scenarios where battery-powered IoT devices are prevalent, as it prolongs their operational lifespan and reduces the need for frequent recharging or replacement. Looking towards the future, the potential of D2D communication within the framework of Industry 4.0 is immense. As industries continue to embrace digital transformation, the integration of D2D communication in 5G networks will become increasingly critical. The ongoing research and development in this domain promise to unlock new capabilities and applications, cementing

D2D communication as a cornerstone of the fourth industrial revolution.

6.1.1 Basics of D2D

D2D communication signifies a paradigm shift in wireless networking, emphasizing direct data exchange between nearby devices without routing through a central network infrastructure. Originating as a concept to augment traditional cellular networks, D2D has found its niche within 5G technology, offering a streamlined communication pathway. This introduction sets the stage for understanding the operational mechanics of D2D and its integral role in 5G connectivity, especially in the context of industrial IoT environments. D2D's ability to facilitate immediate and efficient data transfer between devices is crucial in these settings, where timely information exchange is paramount.

In industrial IoT scenarios, D2D communication is transformative. It allows a multitude of devices, such as sensors, actuators, and machinery, to communicate directly, facilitating real-time data exchange and decision-making. This direct communication is particularly beneficial in complex industrial environments where rapid response to changing conditions is critical. D2D enhances the responsiveness and efficiency of IoT networks in these settings, leading to increased productivity and reduced operational costs.

A critical advantage of D2D in industrial applications is its ability to address blind spots – areas where traditional network coverage is inadequate or nonexistent. By enabling devices to connect directly, D2D communication can extend network reach to these blind spots, ensuring continuous connectivity across the industrial floor [LLMJ18].

This aspect is particularly crucial in maintaining uninterrupted operations and ensuring safety in manufacturing plants, as depicted in Figure 6.1, which illustrates the role of D2D in covering blind spots within an industrial setup.

Efficient spectrum utilization is a hallmark of D2D communication. By allowing devices to communicate over short distances, D2D minimizes the need for wide-range signal broadcasting, thereby reducing interference and optimizing spectrum use. This efficiency is crucial in densely populated industrial environments where numerous devices operate simultaneously. D2D's ability to dynamically allocate spectrum resources based on proximity and need enhances overall network efficiency and throughput.

Energy efficiency is another cornerstone of D2D communication, especially in

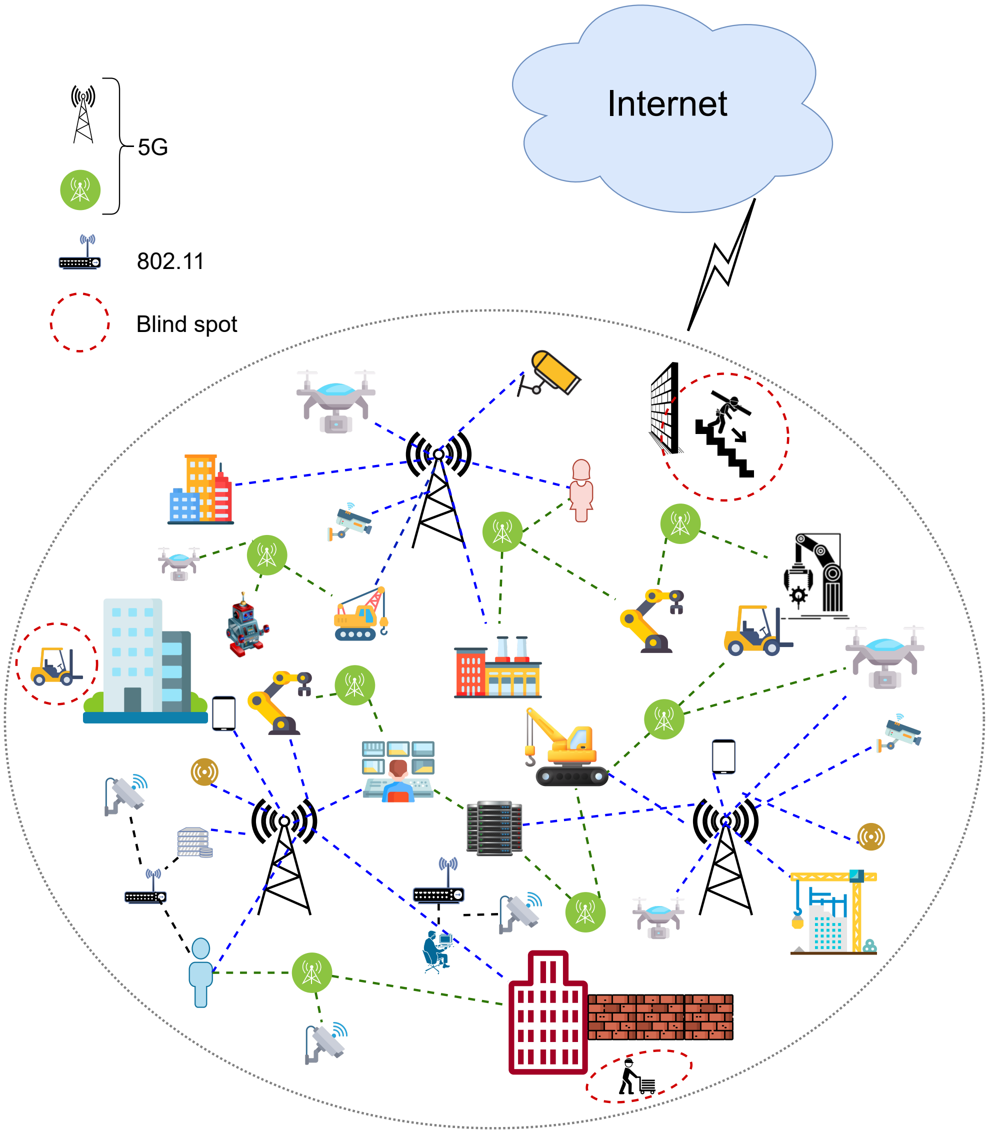

Figure 6.1: A typical smart industry scenario with blind spots

the context of IoT devices that are often battery-operated. By facilitating direct, short-range communication, D2D reduces the power requirements for data transmission, extending the battery life of devices. This feature is essential for sustainable industrial operations, ensuring that IoT devices can operate longer without frequent recharging or replacements, thereby reducing the environmental footprint and operational costs.

Despite its advantages, D2D communication faces several challenges, particularly in terms of network management and integration with existing infrastructures. Managing a large number of D2D connections in an industrial environment requires sophisticated algorithms to ensure efficient network coordination and avoid interference. Additionally, integrating D2D with existing cellular networks while maintaining seamless connectivity and service quality is a complex task. These challenges must be addressed to fully harness the potential of D2D in industrial applications.

D2D communication stands as a critical component of modern industrial IoT scenarios, particularly within the framework of 5G networks. Its ability to enable direct device communication, address coverage issues, optimize spectrum utilization, and enhance energy efficiency makes it indispensable in smart manufacturing environments. As we look to the future, the continuous evolution of D2D technology, coupled with advancements in AI and cybersecurity, will undoubtedly expand its role and impact in the industrial sector. The ongoing development of D2D communication technologies promises to further revolutionize the way industrial operations are conducted, leading to smarter, more efficient, and more sustainable manufacturing practices.

6.1.2 Related Work

Existing research on device-to-device (D2D) communication in IoT and industrial networks has explored energy efficiency, caching, relaying, and security, providing important foundations for cooperative communication mechanisms. However, most prior works do not explicitly address blind-spot removal or connectivity restoration in dynamic industrial environments.

In [NJA+21], an intelligent D2D communication framework is proposed to enhance energy efficiency for NB-IoT delay-sensitive applications using a multi-armed bandit formulation with upper confidence bound-based reinforcement learning. This work is relevant to our study as it demonstrates the potential of adaptive D2D decision-making under resource constraints. However, it does not consider coverage holes or relay selection mechanisms triggered by loss of connectivity, which are central to blind-spot mitigation in smart industrial settings.

The work in [CB20] addresses the complexity of caching management in shared edge storage for Industry 4.0 and connected-car applications. While this study highlights the benefits of cooperation among devices to improve latency and reliability, it focuses on content availability rather than physical-layer connectivity loss, and does not consider D2D relaying to overcome coverage gaps caused by obstructions or harsh industrial layouts.

Energy-efficient relay-assisted IoT-D2D communication is investigated in [ATAI18], where relaying is optimized primarily from an energy-efficiency perspective. This is directly relevant to our approach, as we also exploit nearby IoT devices as relays. However, [ATAI18] does not address blind-spot detection or fast relay selection driven by mobility-induced or obstruction-induced coverage loss, which is the focus of the proposed algorithm in this chapter.

In [KHG23], relay selection based on SINR is used to improve spectral efficiency and data secrecy through XOR-based decode-and-forward (DF) relaying. This work highlights the importance of relay choice in determining link quality and performance. In contrast, our work prioritizes reliable and timely restoration of connectivity in industrial blind spots, emphasizing lightweight relay discovery and selection suitable for real-time industrial operations rather than security-centric optimization.

The D2D-enabled collaborative caching scheme in [OKS$^+$18] reduces content delivery latency and improves reliability at the wireless edge. Although this supports the broader motivation for cooperative device behavior, it does not consider scenarios where direct infrastructure connectivity is unavailable, nor does it define a mechanism for relay-based coverage extension.

Finally, [THQ18] studies joint resource allocation and probabilistic caching in heterogeneous networks with full-duplex relays. This work is relevant in demonstrating the performance gains achievable through relay-assisted D2D communication under advanced optimization frameworks. However, it does not explicitly address blind-spot scenarios or propose an operational algorithm for relay selection and handover in industrial IoT environments.

Relaying techniques for proximity-based services and coverage extension are also studied in the 3GPP standardization context, particularly for UE-to-network and UE-to-UE relaying in D2D/ProSe scenarios [KSR17]. These standardized concepts provide a practical foundation for using nearby devices as temporary relays, motivating the blind-spot removal algorithm proposed in this chapter.

In summary, existing studies motivate cooperation, relaying, and D2D operation in IoT networks, but do not explicitly address blind-spot-triggered connectivity loss in industrial environments. This chapter addresses this gap by proposing a lightweight D2D relay-selection algorithm tailored to blind-spot removal in smart industry scenarios.

6.1.3 Evolution in Industrial Context

In indoor industrial environments, coverage outages frequently arise from machinery blockage, metallic surfaces, and dynamic worker mobility. D2D communication enables nearby IoT devices to act as temporary relays, restoring connectivity when direct infrastructure links degrade [KSR17]. This motivates lightweight relay discovery and selection mechanisms for blind-spot events, which is the focus of this chapter.

6.2 Network Topology and D2D Architecture

The design of D2D communication architectures in Industry 4.0 represents a significant shift from traditional centralized models to decentralized, resilient systems that enhance industrial operations. By facilitating direct device communication, D2D networks address coverage challenges, particularly in blind spots, through strategically placed relay nodes that maintain connectivity in complex environments. The integration of AI and machine learning further optimizes these topologies, allowing for dynamic adjustments and improved connectivity availability. Ultimately, the evolution of D2D network topologies is crucial for realizing the full potential of smart industrial environments, driving innovation and productivity.

6.2.1 Designing D2D Networks

D2D network design in industrial environments focuses on relay-assisted connectivity to mitigate coverage gaps caused by blockage, machinery density, and complex layouts. Key design parameters include node placement, inter-node distance, and device mobility, which jointly affect link reliability and interference conditions.

D2D communication enables nearby IoT devices to operate as relays, extending coverage into blind spots and reducing reliance on centralized infrastructure. Deployment feasibility is evaluated in terms of coverage continuity, robustness to environmental changes, and scalability under varying traffic loads.

Figure 6.1 illustrates the D2D-assisted architecture used to provide alternative communication paths in regions with insufficient direct connectivity.

Node distribution is optimized to ensure full area coverage, with increased relay density in regions with higher attenuation or traffic demand. Figure 6.2

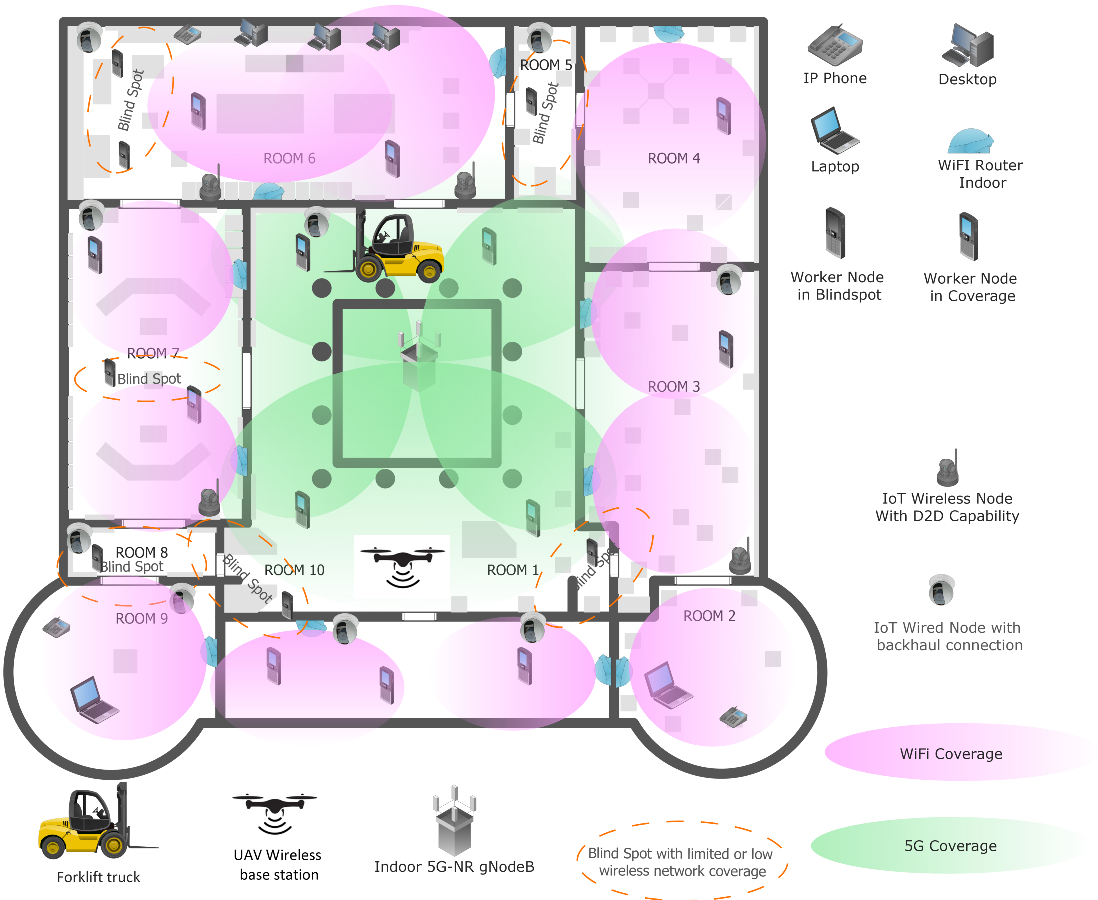

Figure 6.2: Illustrative smart-industry floor layout showing infrastructure node(s), IoT devices, and mobile worker nodes. The layout highlights potential coverage degradation regions due to indoor obstructions and motivates D2D relaying for connectivity restoration.

depicts the spatial deployment of worker and IoT nodes, demonstrating how relay placement enables reliable communication across the industrial site.

In D2D networks, connectivity optimization is achieved through dynamic path selection that adapts to changes in node mobility and traffic load. Adaptive algorithms are used to update relay and routing decisions in response to variations in device location and data demand, which are common in industrial environments. The objective is to maintain low-latency and reliable communication links by selecting paths with favorable channel and interference conditions, thereby ensuring stable data delivery under dynamic operating conditions.

The simulation in Figure 6.3 considers an industrial indoor environment where coverage degradation arises from geometry-induced blockage, interference, and worker mobility rather than predefined blind cells. A total of $N = 100$ worker nodes move within a bounded factory floor according to a random-walk mobil-

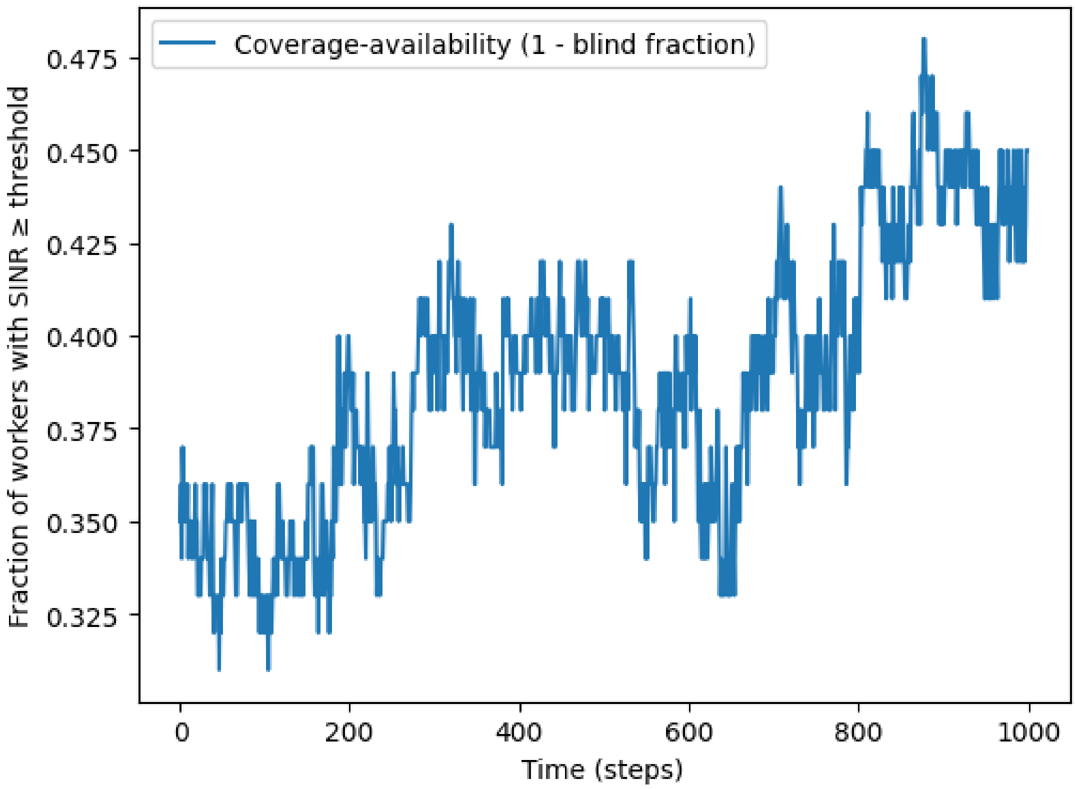

Figure 6.3: Coverage availability over time in an industrial environment, defined as the fraction of workers achieving SINR above the operational threshold under geometry-induced blockage and interference.

ity model. At each time step, the radio link quality between a worker and the serving infrastructure node is evaluated using a SINR-based connectivity criterion that accounts for path loss, obstacle-induced attenuation, interference, and noise.

The considered propagation and blockage behavior is consistent with 3GPP indoor industrial channel models, where coverage degradation arises from machinery-induced obstruction, shadowing, and inter-node interference rather than artificially defined blind regions [ETS22].

A worker is considered connected if its instantaneous SINR exceeds a predefined operational threshold; otherwise, the worker is deemed to be located in a coverage blind spot. Let $A(i, t)$ denote the connectivity indicator for worker i at time step t:

$$A(i,t) = \begin{cases} 1, & \text{if } \mathrm{SINR}_i(t) \geq \gamma_{\mathrm{th}}, \\ 0, & \text{otherwise.} \end{cases} \tag{6.1}$$

The coverage availability at time t is defined as the fraction of workers maintaining connectivity:

$$\text{Availability}(t) = \frac{1}{N} \sum_{i=1}^{N} A(i, t). \tag{6.2}$$

This formulation captures the dynamic evolution of coverage availability as workers transition between line-of-sight and non-line-of-sight conditions due to mobility and environmental obstruction.

To mitigate connectivity loss in coverage blind spots, redundancy is introduced through D2D relaying. When a worker node experiences SINR degradation below the operational threshold, the proposed algorithm enables it to identify nearby IoT nodes with adequate link quality and establish a relay-assisted D2D connection. This mechanism provides an alternative communication path to the infrastructure, improving connectivity continuity without relying on fixed redundancy planning.

Figure 6.4 illustrates the considered scenario, where infrastructure nodes, IoT devices, and mobile worker nodes coexist within an industrial layout. Coverage gaps arise naturally from blockage and interference, and D2D relaying is employed to restore connectivity for workers entering these regions.

D2D networks in industrial settings must effectively manage potential interference and network congestion to maintain optimal performance. This entails the deployment of interference mitigation techniques and congestion control protocols. Techniques like frequency hopping or beamforming are used to minimize interference, while strategies such as load balancing and traffic prioritization help manage network traffic. Figure 6.5 illustrates these techniques in action. The illustration highlights the effect of increasing the number of IoT nodes and selecting relays with higher signal strength on system performance. When the worker node selects a relay with stronger signal quality, the system's throughput improves. A larger number of IoT nodes in the system increases the probability of worker nodes finding relays with better signal strength, thereby enhancing system throughput. In Figure 6.5, throughput is defined as the achievable data rate (Mbps) for relay-assisted transmission, computed using a Shannon-based abstraction $T_i = B \log_2(1 + \text{SINR}_i)$. During each iteration, the worker node first identifies nearby IoT nodes within a designated range and selects the optimal candidate based on factors such as distance, battery life, and load. If a suitable

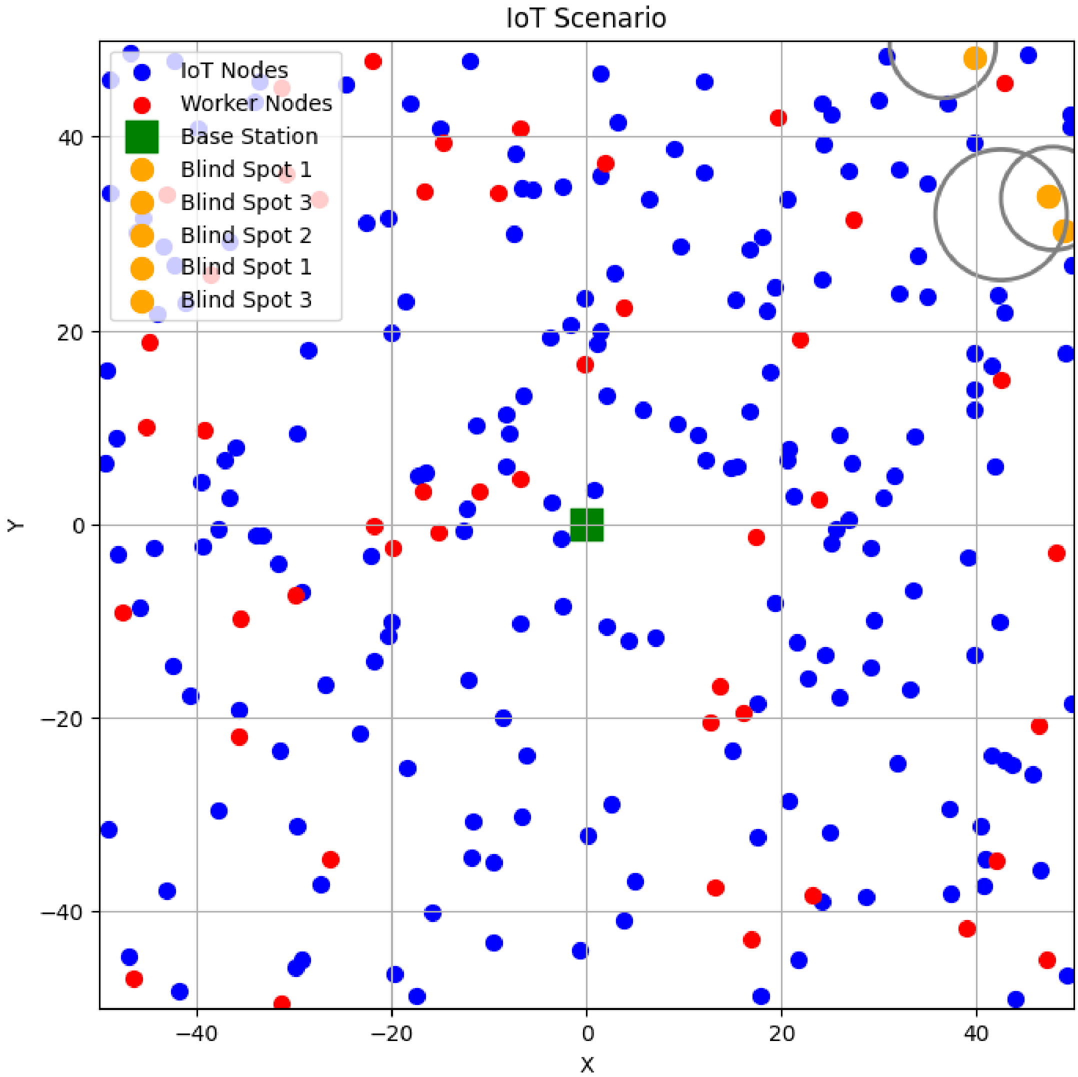

Figure 6.4: Industrial indoor scenario showing infrastructure node(s) (green), IoT devices (blue), and a mobile worker (red). Coverage outages arise dynamically due to blockage and interference; when $\mathrm{SINR} < \gamma_{\mathrm{th}}$, the worker triggers D2D relay discovery to restore connectivity.

relay node is found, the throughput for that iteration, represented as T_i, is computed using the IIoT throughput formula, which takes into account parameters like available bandwidth, channel conditions, and relay node performance, providing a realistic estimate of the data transmission capacity. If no suitable relay is available, $T_i = 0$. The total system throughput over time is then determined by summing these per-iteration values:

$$S(t) = \sum_{i=1}^{t} T_i, \tag{6.3}$$

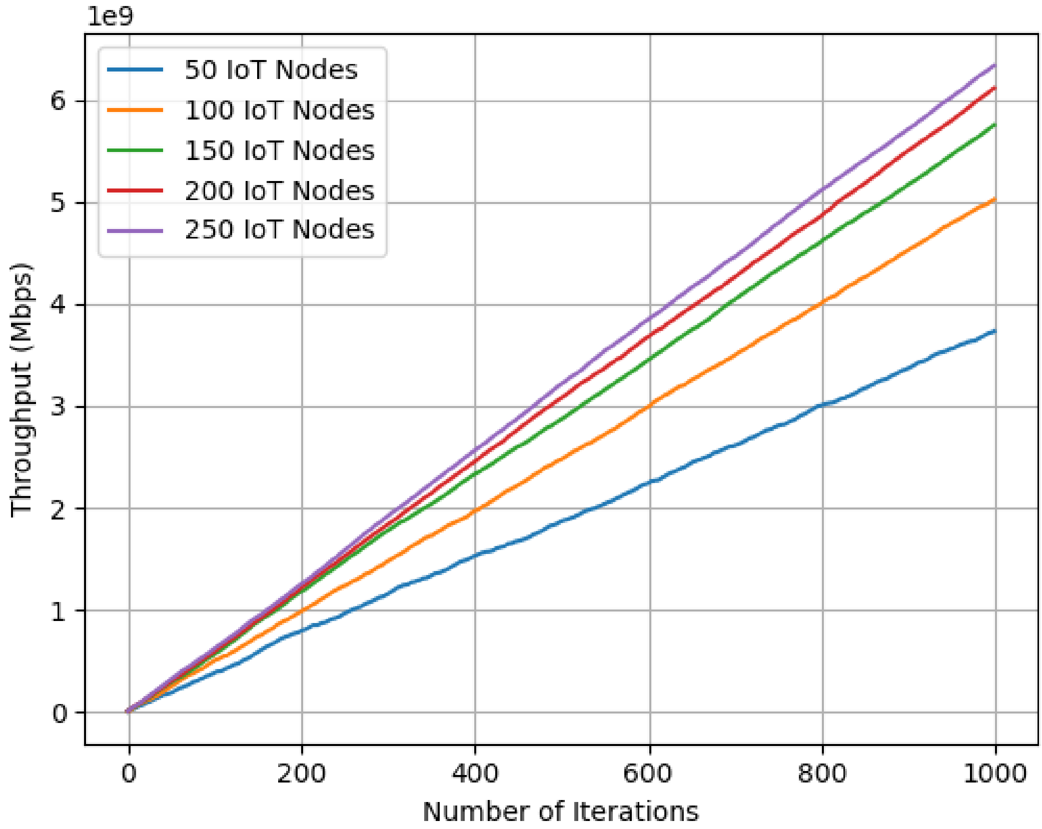

Figure 6.5: Throughput performance of the system across different numbers of IoT nodes (50, 100, 150, 200, and 250) over 1000 iterations. The results demonstrate a linear increase in throughput as the number of IoT nodes increases, highlighting the system's scalability and efficiency in handling additional network capacity.

thus reflecting the aggregate data rate achieved throughout the simulation.

Specifically, the diagram demonstrates that as the number of IoT nodes grows, the availability of relay nodes also increases. This, in turn, raises the likelihood of worker nodes connecting to relays with superior signal quality, resulting in higher throughput. Additionally, when worker nodes utilize relays with stronger signals, the volume of data transmitted increases, further boosting the system's overall throughput.

Energy efficiency is a pivotal aspect of D2D network design, especially in large-scale industrial applications where the sustainability of operations is a concern. Strategies for enhancing energy efficiency include optimizing the sleep and wake cycles of nodes and employing energy-efficient communication protocols. The use of renewable energy sources for powering nodes also contributes to the overall sustainability of the network. Figure 6.6 provides an in-depth view of relay selection strategies. This figure illustrates how different relay selection algorithms impact connectivity availability, underlining the importance of connectivity-aware relay selection in D2D systems.

Scalability and the capacity to adapt to future technological developments are essential attributes of D2D networks. A flexible network architecture is needed to accommodate the integration of new devices and technologies seamlessly. Figure 6.6 illustrates this aspect by highlighting how connectivity availability is maintained as the network scales. This figure underscores the importance of designing D2D networks that are not only efficient in the current technological landscape but are also prepared to evolve with future advancements.

Security is a critical concern in industrial D2D networks, given the sensitive nature of the data being communicated. The network design must incorporate strong encryption protocols, secure authentication processes, and regular security audits to safeguard against potential breaches.

To encapsulate the various stages and considerations in D2D network design, Table 6.1 serves as an essential reference. Table 6.1 shows the simulation parameters and their values used in this work. This table succinctly summarizes the key design elements, strategies, and tools that are crucial at each stage of D2D network development. It provides network engineers and planners with a comprehensive overview, aiding in effective decision-making and implementation in industrial settings.

The strategic design and implementation of D2D networks in industrial environments are complex but critical processes that demand careful consideration of various factors. From optimizing node distribution to ensuring robust security measures, each aspect plays a vital role in the successful deployment of these networks, ultimately contributing to the efficiency and sustainability of industrial operations in the era of Industry 4.0.

6.2.2 Architectural Considerations

Architectural considerations are crucial for the effective implementation of D2D networks in industrial environments, requiring a balance between network structure and dynamic operational demands. Key elements include selecting an appropriate network topology, implementing adaptive communication protocols, and ensuring scalability and reliability. Additionally, energy efficiency, security, and integration with existing infrastructure must be prioritized to maintain performance and facilitate real-time data processing. Future-proofing the network against technological advancements is also essential to accommodate evolving industrial requirements.

6.3 Addressing Blind Spots in IoT Networks

Addressing blind spots in IoT networks is critical for ensuring reliable connectivity in industrial environments, where weak or non-existent network coverage can hinder operational efficiency. Strategic network planning, including analyzing facility layouts and node distribution, is essential for comprehensive coverage. The implementation of advanced technologies, such as 5G and MIMO, alongside data analytics and redundancy strategies, can mitigate connectivity challenges. Future research should focus on innovative solutions like LPWAN and AI-driven management to enhance the resilience of IoT networks in complex industrial settings.

6.3.1 Concept of Blind Spots

These blind spots can arise from physical obstructions, technological limitations, and environmental factors, leading to disruptions in critical information flow, decreased productivity, and safety risks. Addressing blind spots requires effective network planning, advanced technologies, and continuous monitoring to ensure operational continuity and worker safety. Future research is expected to explore innovative solutions, such as ultra-wideband and Li-Fi technologies, to mitigate connectivity challenges in complex environments.

6.3.2 Importance in Industrial Settings

Connectivity and network reliability are essential in modern industrial settings, particularly for Industry 4.0, where automation and real-time monitoring are key. Blind spots can disrupt operations, jeopardize safety, hinder data-driven decision-making, and compromise quality control. Addressing these blind spots is crucial for ensuring operational continuity, enhancing productivity, and supporting efficient supply chain management. Ultimately, eliminating blind spots enables industries to leverage the full potential of digital transformation and maintain competitiveness while pursuing sustainability and technological innovation.

6.4 D2D-Based Algorithm Design

In the context of a large manufacturing plant with a multitude of IoT nodes and worker nodes, the design of D2D-based algorithms becomes a critical aspect of ensuring seamless communication and eliminating blind spots. The scenario involves smart sensor nodes scattered throughout the production floor, worker nodes equipped with smart helmets, and various IoT devices, all contributing to data collection and communication. The plant's layout consists of different rooms designated for distinct operations, with a network infrastructure that includes 5G and WiFi coverage, routers, cameras, and smart sensors.

One of the primary challenges faced in this industrial setup is the existence of coverage blind spots, which can be problematic for mobile worker nodes. These blind spots occur in areas where wireless communication is obstructed, potentially leading to communication breakdowns and safety concerns. To address this issue, the proposed D2D-enabled algorithm aims to utilize IoT nodes as relays to bridge these coverage gaps. This approach leverages the mobility and capability of IoT nodes to establish connections in previously unreachable areas, enhancing connectivity availability, coverage continuity, and achievable throughput.

The algorithm itself operates as follows: when a worker node encounters a blind spot, it actively seeks nearby IoT nodes capable of relaying its data. If suitable relays are found, the worker node selects the relay with the highest RelayScore and attempts to establish a connection. If successful, data transmission resumes through the relay node. If no candidate relays are available, the worker node periodically rechecks for potential relays until one is discovered or it exits the blind spot.

To determine the most appropriate relay nodes, several evaluation metrics are considered, including signal strength, load on the relay, and battery status. Signal strength assessment helps gauge the quality of the connection, while load evaluation ensures that the selected relay can handle additional traffic. Battery status consideration is crucial to maintaining a stable and reliable connection. These metrics are used to define threshold values and weights that guide relay node selection.

The relay selection matrix provides a systematic view of potential relay nodes, listing their evaluation metrics. The worker node uses this matrix to compare and analyze relay nodes based on the defined criteria. The relay node(s) with the highest RelayScore, a weighted score combining the evaluation metrics, is selected as the most suitable relay(s). This approach enables informed decisions and efficient resource utilization, ensuring optimal network connectivity even in dynamic industrial environments.

6.4.1 Algorithm Development

Algorithm development for addressing blind spots in D2D communication within industrial settings is a multifaceted and crucial task that demands a comprehensive approach. The scenario in question involves a vast manufacturing plant, replete with thousands of IoT nodes spread across the production floor. These IoT nodes perform a variety of functions, including monitoring temperature, humidity levels, production rates, and quality metrics. Alongside these smart sensor nodes, there are hundreds of worker nodes stationed along the production line. These worker nodes play a pivotal role in tasks that cannot be automated, such as quality control, machine maintenance, and manual assembly. They rely on wearable smart helmets for safety and communication, making them an integral part of the IoT ecosystem.

Within this intricate industrial environment, wireless communication is the backbone of connectivity. However, it's not without its challenges, primarily the existence of blind spots where wireless coverage is obstructed. These blind spots pose a significant risk to mobile worker nodes who frequently traverse these areas. It's in this context that the algorithm's mission unfolds – to eliminate these coverage holes using D2D-capable IoT nodes that serve as relays.

To dive deeper into the algorithm's mechanics, let's consider a mathematical model that simulates the scenario. To model blind-spot events in a realistic indoor factory, connectivity loss is defined by link-quality degradation rather than pre-marked blocked cells. At each time step, the worker evaluates the direct link to the serving infrastructure (or gateway) using an SINR-based criterion incorporating path loss, shadowing/obstruction loss due to machinery, interference, and noise (consistent with indoor industrial channel assumptions). If $\mathrm{SINR}_c(t) < \gamma_{\mathrm{th}}$, the worker is considered to be in a coverage outage state and triggers D2D relay discovery.

The core of the derived algorithm revolves around worker nodes identifying nearby IoT nodes capable of relaying their data when they encounter these blind spots. Unlike static relay assignment approaches, the proposed algorithm is *event-driven* and *connectivity-aware*. Relay discovery is triggered only upon SINR degradation below an operational threshold, ensuring that D2D relaying is activated exclusively during coverage outage events. The algorithm jointly considers link quality, relay load, and battery constraints, enabling adaptive and lightweight relay selection suitable for real-time industrial environments. This distinguishes the proposed approach from static or pre-planned redundancy mechanisms commonly assumed in the literature. The algorithm operates as follows:

6.4.2 Proposed D2D-enabled blind spot removing algorithm

In this algorithm, the worker node identifies any nearby IoT nodes and checks if they are capable of relaying worker data. If there are no candidate relays, it waits and periodically rechecks for available relays. If there are candidate relays, it selects the relay with the highest RelayScore and attempts to establish a connection. If the connection cannot be established, it waits and periodically retries. Once the worker node exits the blind spot, it resumes normal operation.

- **Candidate Relay Identification:** The process of identifying candidate relay nodes involves the worker node's wireless scanning techniques, such as Wi-Fi or Bluetooth, to detect nearby IoT nodes. To elaborate, the worker node performs active scanning, which means it actively sends out probe requests to find potential relays. These probe requests contain information about the worker node's capabilities and requirements. IoT nodes within range respond to these probes if they are available and capable of relaying data. The worker node then compiles a list of these responsive IoT nodes, considering their signal strength, proximity, and suitability for relaying data. Let's consider:
  - N: The total number of potential relay nodes
  - i: Index variable representing each relay node, where $1 \leq i \leq N$
  - $S(i)$: Signal strength of relay node i
  - $L(i)$: Load on relay node i
  - $B(i)$: Battery status of relay node i
- **Evaluation Metrics:** To determine the suitability of these potential relay nodes, several crucial metrics come into play:

Algorithm 1: Event-driven D2D relay selection algorithm for blind-spot recovery

```
Input  : Initialize the set of IoT nodes (I) and worker nodes (C)
Output: Restore connectivity during blind-spot (outage) events via D2D
        relaying
repeat
    Worker node c detects an outage event: SINR_c(t) < γ_th;
    c discovers its surroundings and identifies any nodes in range;
    for each node i in range do
        Check if i is capable of relaying worker data;
        if yes then
            Add i to a list of candidate relays (R);
    if there are no candidate relays in range then
        repeat
            Relay discovery
        until Until a suitable relay is discovered;
    else
        Select relay r = arg max_{i∈R} RelayScore(i) and request to establish a
          connection;
        if r accepts the request then
            Transfer the connection from c to r
        else
            repeat
                Relay discovery
            until Until a suitable relay is discovered;
until worker node c exits the blind spot and resumes normal operation;
```

1. **Signal Strength ($S(i)$):** This metric assesses the quality of the connection between the worker node and each potential relay node. It's measured in decibels (dB) and is indicative of the signal's strength.

2. **Load ($L(i)$):** The load on each potential relay node is evaluated to gauge its current traffic load. This is represented as a ratio between the current traffic and the maximum capacity of the relay node.

3. **Battery Status ($B(i)$):** The battery status of potential relay nodes is considered. It factors in variables like the remaining charge percentage and the expected battery life based on historical usage patterns.

- **Relay Selection Criteria:** Based on the evaluations of signal strength, load, and battery status, the algorithm establishes criteria for selecting suitable relay nodes. Threshold values ($Sthresh$, $Lthresh$, $Bthresh$) or weights ($Sweight$, $Lweight$, $Bweight$) are assigned to each evaluation metric. Let's donate them as:

- S_{thresh}: Threshold for signal strength (-70 dB)
- L_{thresh}: Threshold for load (0.8)
- B_{thresh}: Threshold for battery status (30%)
- S_{weight}: Weight for signal strength
- L_{weight}: Weight for load
- B_{weight}: Weight for battery status

- **Relay Selection Matrix:** A relay selection matrix is constructed, providing a systematic view of the available relay nodes and their associated evaluation metrics. This matrix includes relevant information such as the IoT node's identifier, signal strength, load, and battery status.

Relay Node	Signal Strength (S)	Load (L)	Battery Status (B)
1	$S(1)$	$L(1)$	$B(1)$
2	$S(2)$	$L(2)$	$B(2)$
$\ldots$	$\ldots$	$\ldots$	$\ldots$
N	$S(N)$	$L(N)$	$B(N)$

- **Evaluation and Selection:** Leveraging the relay selection matrix, the worker node can compare and analyze potential relay nodes based on the predefined criteria. To make an informed selection, a RelayScore is calculated for each node using the equation 6.4 below. This formula takes into account signal strength, load, and battery status. The node with the highest RelayScore is chosen as the most suitable relay. In cases where multiple nodes have the same highest score, tie-breaking mechanisms can be applied.

$$\begin{aligned}\text{RelayScore}(i) = S_{\text{weight}} \cdot (S(i) - S_{\text{thresh}}) + L_{\text{weight}} \cdot (L_{\text{thresh}} - L(i)) \\ + B_{\text{weight}} \cdot (B(i) - B_{\text{thresh}})\end{aligned} \tag{6.4}$$

6.4.3 Discussion on Relay Selection Tradeoffs

The relay selection heuristic reflects a tradeoff between link reliability, relay sustainability, and network stability. Prioritizing signal strength improves throughput and reduces retransmissions, but may overload relays if load is ignored.

Incorporating load prevents congestion but may result in selecting suboptimal channels. Battery-awareness ensures long-term network sustainability at the cost of occasionally selecting relays with lower instantaneous SINR.

The weighted RelayScore formulation allows these competing objectives to be balanced through adjustable weights, enabling deployment-specific tuning. This heuristic is intentionally lightweight, avoiding computationally expensive optimization methods that are impractical for latency-sensitive industrial environments.

In essence, the algorithm's detailed design and dynamic adaptability make it a robust solution for addressing blind spots in D2D communication within industrial settings. It takes into account a multitude of factors, from signal quality to network load to battery health, to make informed relay selections, ultimately enhancing the efficiency and reliability of the communication network in complex and challenging industrial environments. By following this algorithm, worker nodes within the manufacturing plant can make intelligent decisions when selecting suitable relay nodes. This approach optimizes resource utilization, ensures reliable network connectivity, and ultimately eliminates blind spots within the industrial IoT network.

6.4.4 Implementation Strategies

Table 6.1: Simulation Parameters

Simulation Parameter	Values
IoT Nodes	$100 \sim 1000$
Worker Nodes	$20 \sim 100$
Coverage Area	Square floor of side length $100 \sim 1000$ m
Connectivity Criterion	SINR threshold γ_{th}
Transmit Power (IoT)	$5mW$
Transmit Power (Worker)	$100mW$
Antenna Gain	$1dB$
Noise Power Density	$-174dBm/Hz$
Bandwidth	$1e6Hz$
Data Rate	$100Kbps$
Obstacle / penetration loss (IoT links)	10 dB
Obstacle / penetration loss (Worker links)	2 dB

- **Computational Environment** The implementation and simulations were conducted using Python on Google Colab, a cloud-based Jupyter Notebook service. This platform provided an efficient and scalable computational environment for executing algorithms, processing data, and visualizing results.

- **Comprehensive IoT Node Deployment** To address the blind spots in smart industries, strategic deployment of IoT nodes is crucial. As per Table 6.1, the plant incorporates up to 1000 IoT nodes, ensuring extensive coverage. These nodes, equipped with sensors to monitor various parameters like temperature and humidity, are not just confined to static locations but are also integrated into mobile elements like conveyor belts and finished products. This comprehensive deployment allows for a dynamic and adaptive network that can respond to changes in the environment and production processes, thereby significantly reducing the instances of blind spots.

- **Worker Node Integration and Safety** The integration of worker nodes into the IoT system plays a pivotal role in ensuring safety and efficiency. With smart helmets and other wearable devices, workers are not just part of the production process but also a critical component of the IoT network. These devices enable real-time tracking and safety alerts, especially when workers enter blind spots. This integration is essential in automating safety protocols and ensuring swift response to potential hazards, thereby enhancing overall workplace safety [PPZC20].

- **Utilization of D2D Communication** Direct D2D communication is a key strategy for overcoming the limitations posed by blind spots. This approach enables devices to communicate directly with each other without relying on the central network. When a worker enters a blind spot, the D2D capability allows their device to connect with a suitable IoT relay selected according to the RelayScore metric, maintaining uninterrupted communication. This strategy not only ensures continuous connectivity but also aids in conserving network resources by reducing the load on central servers.

- **Dynamic Relay Node Selection** The relay node selection process is a critical component of this system. As detailed in the provided text, worker nodes scan for potential IoT relay nodes based on signal strength, load, and battery status. This dynamic selection process ensures that the most efficient and reliable relay node is chosen, enhancing the overall stability and efficiency of the network. The relay selection matrix, as outlined, provides a systematic approach for evaluating and selecting the most suitable relay nodes.

- **Relay Node Optimization** Optimizing the relay nodes involves careful consideration of signal strength, load, and battery life. By setting specific thresholds and weights for these parameters, the system can ensure that the selected relay nodes offer the best possible performance. This optimization is crucial in maintaining a high-quality connection for the worker nodes, especially when they are in blind spots. It also ensures that the relay nodes themselves are operating efficiently, conserving energy and processing power.

- **Algorithmic Approach to Blind Spot Removal** The proposed D2D-enabled algorithm for blind spot removal is instrumental in maintaining continuous coverage. This algorithm allows worker nodes to efficiently find and connect to IoT nodes within their range, ensuring that data transmission is uninterrupted even in areas with traditionally weak coverage. the algorithm's ability to periodically recheck for available relays and to select the most suitable relay based on the RelayScore is a strategic approach to dynamically adapting to the changing environment of the manufacturing floor.

- **Simulation and Real-time Adjustment** Simulations play a pivotal role in testing and refining the IoT network. By simulating different scenarios, as shown in Figure 6.4 and 6.5, the effectiveness of the relay system in various conditions can be assessed. This includes testing the network's response to different numbers of IoT and worker nodes, as well as varying sizes of blind spots. Real-time adjustments based on these simulations can significantly improve the system's efficiency and reliability.

- **Enhancing System Throughput** Increasing the quantity and quality of IoT nodes, as suggested in the simulation results in Figure 6.5, directly impacts system throughput. By ensuring that worker nodes have access to relays with superior signal strength, the system can handle a higher volume of data transmission. This enhancement is crucial for maintaining a high-performance network that can support the vast array of IoT devices in the smart industry setting.

- **Connectivity Availability Metrics and Improvement** Connectivity availability is a crucial metric. Connectivity availability refers to the fraction of time during which a worker node maintains an operational SINR, and is distinct from throughput-based performance metrics.

The implementation strategy focuses on improving connectivity availability by ensuring stable connections and minimizing the time worker nodes spend in blind spots. By increasing the number of IoT nodes, the system can offer more relay options, thereby improving connectivity availability and coverage continuity. Monitoring and continually seeking to improve connectivity-availability metrics are key to the long-term success and scalability of the IoT network in smart industries.

- **Future-Proofing and Scalability** The final strategy revolves around future-proofing and scalability. As the manufacturing plant evolves and expands, the IoT network must adapt accordingly. This involves not only adding more nodes and enhancing their capabilities but also refining the algorithms and relay selection processes to accommodate new challenges and technologies. Ensuring that the system is scalable and can integrate future advancements in IoT and D2D communication is essential for maintaining a state-of-the-art smart industry environment.

6.5 Impact on Smart Industry Connectivity

The integration of IoT and D2D communication technologies in smart industries enhances network resilience and coverage, addressing challenges like network blind spots in large manufacturing plants. This robust connectivity ensures uninterrupted data flow regarding production processes, machine performance, and worker safety, creating a cohesive communication network. Improved connectivity enables real-time data exchange and analytics, automating decision-making for predictive maintenance and operational efficiency. Ultimately, this transformation leads to smarter, self-aware industrial operations aligned with the goals of Industry 4.0.

6.5.1 Simulation Assumptions and Performance Metrics

The simulation considers an indoor industrial environment consistent with 3GPP indoor factory channel models. The following assumptions are made:

- **Topology and Mobility:** Worker nodes follow a bounded random-walk mobility model within the factory floor, while IoT nodes are static and uniformly distributed.

- **Propagation Model:** Path loss, shadowing due to machinery, and interference from concurrent transmissions are incorporated in the SINR computation.
- **Connectivity Criterion:** A worker is considered connected if $\mathrm{SINR} \geq \gamma_{\mathrm{th}}$; otherwise, it experiences a coverage outage.
- **Relay Discovery:** Only IoT nodes within communication range and satisfying minimum battery and load constraints are eligible as relays.
- **Traffic Model:** Worker nodes generate uplink IIoT traffic at a constant rate.

Throughput Calculation: Throughput is computed using a Shannon-based link abstraction for relay-assisted transmission:

$$T_i = B \log_2 \left(1 + \mathrm{SINR}_i\right), \tag{6.5}$$

where B denotes the system bandwidth and SINR_i is the post-relay SINR of worker i. If no relay is available, $T_i = 0$. This formulation reflects achievable data rates under realistic channel conditions rather than idealized maximum throughput assumptions.

6.5.2 Performance Evaluation

This evaluation focuses on connectivity availability rather than throughput alone, reflecting the primary objective of blind-spot recovery in industrial IoT environments. The performance evaluation of the IoT network in a smart industry setting, particularly focusing on connectivity availability, is critical. In our study, conducted using Python and illustrated in Fig. 6.6, connectivity availability is compared between the proposed relay selection algorithm and a random relay selection approach, we observe significant improvements. The proposed algorithm, which intelligently selects relay nodes based on signal strength, load, and battery status, showcases higher connectivity availability due to its capability to dynamically adapt to changing conditions on the manufacturing floor, ensuring that worker nodes are consistently connected to the most optimal relay nodes. In contrast, the random relay selection method, which lacks this sophistication, results in lower connectivity availability. This contrast not only highlights the effectiveness of the proposed approach but also underscores the importance of a strategic relay selection process in maintaining a robust IoT network in industrial environments.

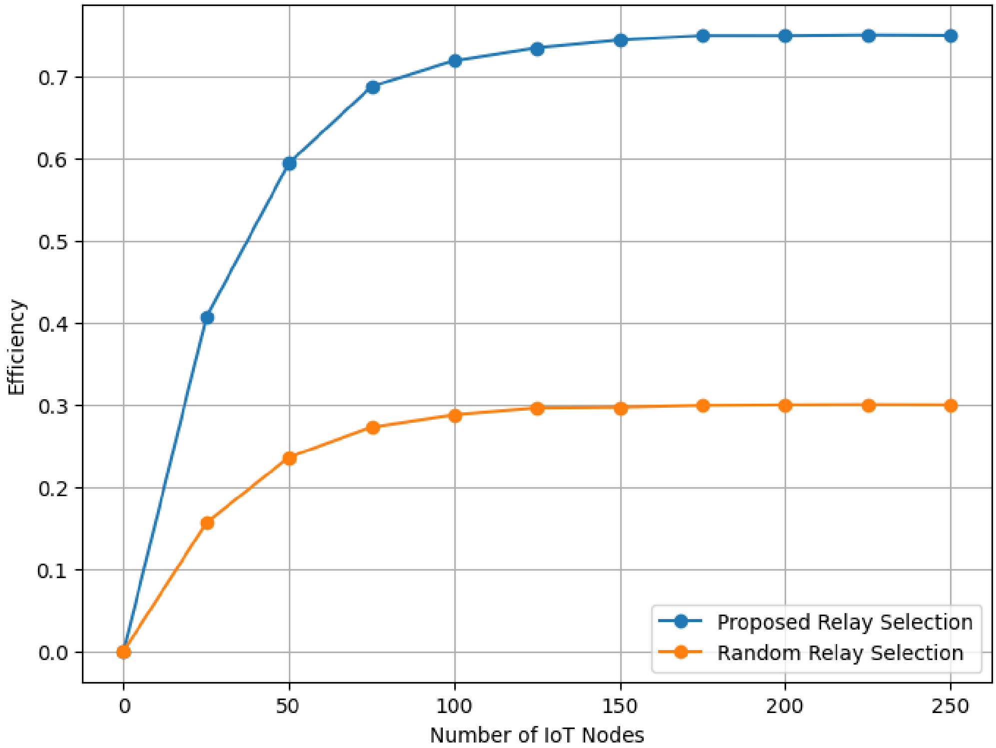

Figure 6.6: Connectivity availability (η) versus number of IoT nodes for the proposed relay selection and random relay selection. η denotes the fraction of time steps in which the worker maintains $\mathrm{SINR} \geq \gamma_{\mathrm{th}}$ (i.e., is not in outage).

At 250 IoT nodes, the connectivity availability is approximately $\eta \approx 0.72$ for the proposed relay selection and $\eta \approx 0.30$ for random relay selection. The relative improvement is:

$$\text{Availability gain} = \frac{0.72 - 0.30}{0.30} \times 100 \approx 140\%.$$

This indicates that the proposed event-driven relay selection substantially reduces outage time by increasing the likelihood of finding a feasible relay with acceptable link quality and resource constraints.

The proposed relay selection increases connectivity availability substantially relative to random selection, indicating improved continuity of service during outage events and more effective relay utilization.

In Fig. 6.6, the performance metric η is defined as a *connectivity-availability* measure. At each time step t, the worker evaluates the post-selection link quality (direct or relay-assisted). If the resulting SINR satisfies $\mathrm{SINR}(t) \geq \gamma_{\mathrm{th}}$, the worker is considered connected; otherwise it is in outage. Connectivity availability is therefore defined as the fraction of time steps in which connectivity is maintained:

$$\eta = \frac{1}{T}\sum_{t=1}^{T} \mathbb{I}(\mathrm{SINR}(t) \geq \gamma_{\mathrm{th}}) \,, \tag{6.6}$$

where T is the total number of iterations (time steps) and $\mathbb{I}(\cdot)$ is the indicator function.

As depicted in Fig. 6.6, as the number of IoT nodes increases, the connectivity availability improves significantly when using the proposed relay selection algorithm. This improvement is due to the increased probability of each worker node finding a suitable relay within proximity, thereby reducing the likelihood of falling into blind spots. Furthermore, a higher density of IoT nodes increases the likelihood of maintaining connectivity during outage events, which is crucial in fast-paced industrial settings. This scalability aspect of the IoT network demonstrates its potential to adapt to growing industrial needs, ensuring that connectivity availability is maintained or enhanced as the system expands.

The comparative analysis of the two relay selection algorithms – the proposed method and the random selection – offers valuable insights into the performance dynamics of IoT networks in smart industries. The proposed algorithm's superior performance is evident in its consistent and reliable connectivity, even as the number of IoT nodes varies. In contrast, the random relay selection algorithm, although simpler, fails to maintain similar connectivity availability levels, especially in scenarios with fewer IoT nodes. This comparison not only validates the proposed algorithm's effectiveness but also highlights the necessity of employing sophisticated approaches for managing complex IoT networks in industrial environments.

This improvement is particularly noticeable in environments with a high density of IoT nodes. The ability to maintain high connectivity availability is essential in smart industries, where real-time data analysis and decision-making are critical. This capability ensures that control, monitoring, and safety-critical data can be exchanged reliably without prolonged outage periods, enabling timely insights and actions. The proposed algorithm's role in sustaining high connectivity availability further solidifies its importance in optimizing smart industry operations.

The performance evaluation, particularly the connectivity-availability comparison between the proposed and random relay selection algorithms, has broader implications for the future of smart industries. It demonstrates the potential of advanced relay selection methods in enhancing the reliability and efficiency of IoT networks. This is crucial as industries continue to evolve and integrate more complex and data-intensive processes. The adaptability of the proposed algorithm to different numbers of IoT nodes and varying environmental conditions suggests its suitability for diverse industrial settings. Looking forward, this adaptability will be key in ensuring that IoT networks can keep pace with the advancing demands of Industry 4.0, making them a cornerstone of future smart industry infrastructures.

6.5.3 Practical Implications

The integration of IoT and D2D communication in manufacturing plants enhances operational efficiency by improving connectivity continuity, optimizing real-time data collection, and reducing downtime. It elevates worker safety through continuous monitoring and instant safety alerts, while predictive maintenance algorithms minimize unplanned repair costs. The D2D-enabled system improves connectivity and adaptability in large-scale industrial networks. Economically, these technologies lead to reduced costs, increased output quality, and a stronger competitive position, marking a pivotal shift towards the goals of Industry 4.0.

6.5.4 Conclusion

This chapter investigates and contributes a lightweight, event-driven D2D relay selection mechanism for blind-spot recovery in industrial IoT environments. Simulation results show that increasing the number of IoT nodes improves *connectivity availability* by increasing the likelihood of finding a feasible relay when the direct link falls below the operational threshold. The proposed relay selection algorithm consistently outperforms random selection by prioritizing candidate relays according to link quality and resource constraints (load and battery).

The key performance metric, connectivity availability η, is defined as the fraction of time steps during which the worker maintains $\mathrm{SINR} \geq \gamma_{\mathrm{th}}$, providing a direct measure of outage reduction and service continuity. Overall, the results demonstrate that lightweight, connectivity-aware relay selection is effective in mitigating coverage blind spots and improving communication reliability in smart industrial environments.

Chapter 7

Intelligent Reflecting Surfaces for 5G/6G in Millimeter-Wave Smart Industries

"If you want to find the secrets of the universe, think in terms of energy, frequency and vibration."
Nikola Tesla (1856–1943)

This chapter discusses the critical role of Intelligent Reflecting Surfaces (IRS) in improving millimeter-wave (mmWave) communication for 5G and beyond. It examines how IRS technology can manipulate electromagnetic waves to mitigate challenges like high attenuation and limited coverage in wireless communication. The chapter highlights the potential of IRS to revolutionize industrial communication networks, enhancing efficiency and reliability in smart manufacturing environments. It concludes by addressing the challenges of integrating IRS into existing wireless networks and the transformative impact it may have in industrial settings.

7.1 Introduction to Intelligent Reflecting Surfaces

Intelligent Reflecting Surfaces (IRS) have emerged as a transformative technology in the domain of wireless communications, playing a pivotal role in the advancement of Industry 4.0. These surfaces, comprised of numerous passive elements, are capable of dynamically manipulating electromagnetic waves, thus offering an unprecedented level of control over wireless signal propagation.

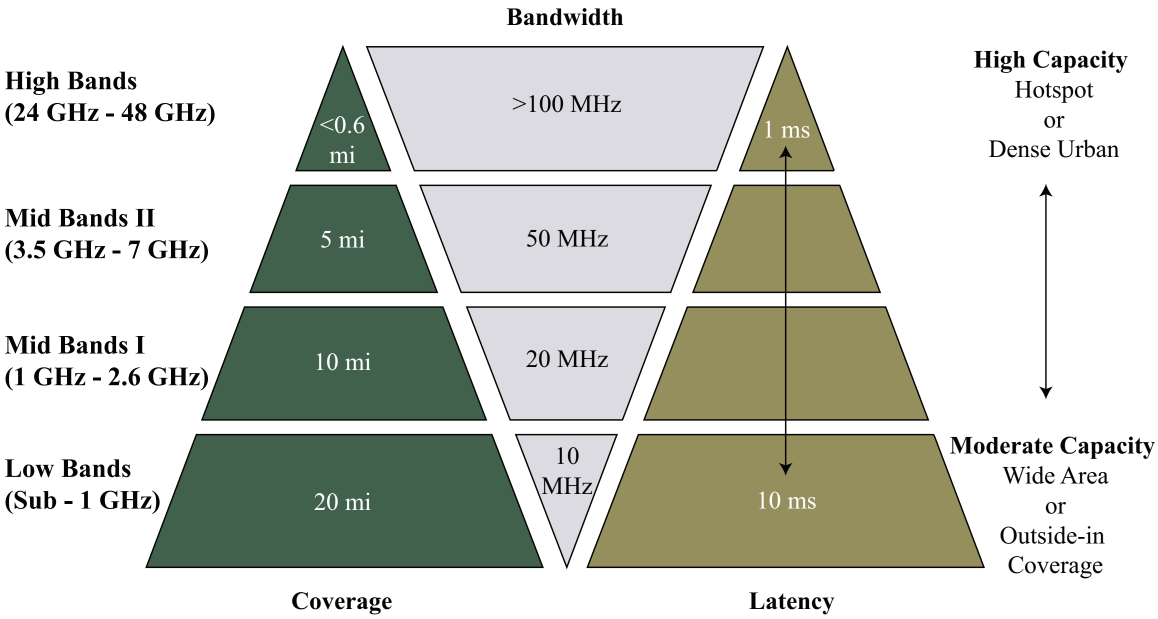

Figure 7.1: Spectrum Breakdown - Examining the complex relationship: How Radio Frequency Bands Affect Coverage, Bandwidth and Latency. [Original Image Courtesy of Ericsson & Broadband Library [McL]

This technological innovation is particularly crucial in complex environments such as smart industries, where traditional signal transmission methods face numerous challenges. In this context, Figure 7.1 The Three Pyramid Diagram provides a foundational understanding by illustrating the frequency spectrum of 5G networks, which underpins the necessity for innovations like IRS in overcoming the limitations of current wireless communication systems. The Fig. 7.1 illustrates the frequency spectrum of 5G networks as a tiered pyramid, highlighting the inherent trade-offs between capacity, coverage, and penetration depth when selecting bands for cellular. Moving down the pyramid reveals bands that offer superior range and penetration, but this comes at the expense of significantly reduced data throughput.

The integration of IRS with millimeter-wave (mmWave) communication, operating in the 30 GHz to 300 GHz frequency range, represents a promising synergy for overcoming inherent wireless transmission challenges [Uni15]. While mmWave communications offer the potential for high data rates, they are prone to significant path loss and physical obstructions. IRS technology, as conceptualized in Figure 7.2, addresses these issues by adjusting the phase of incident waves, thereby improving signal reach and quality in densely packed indus-

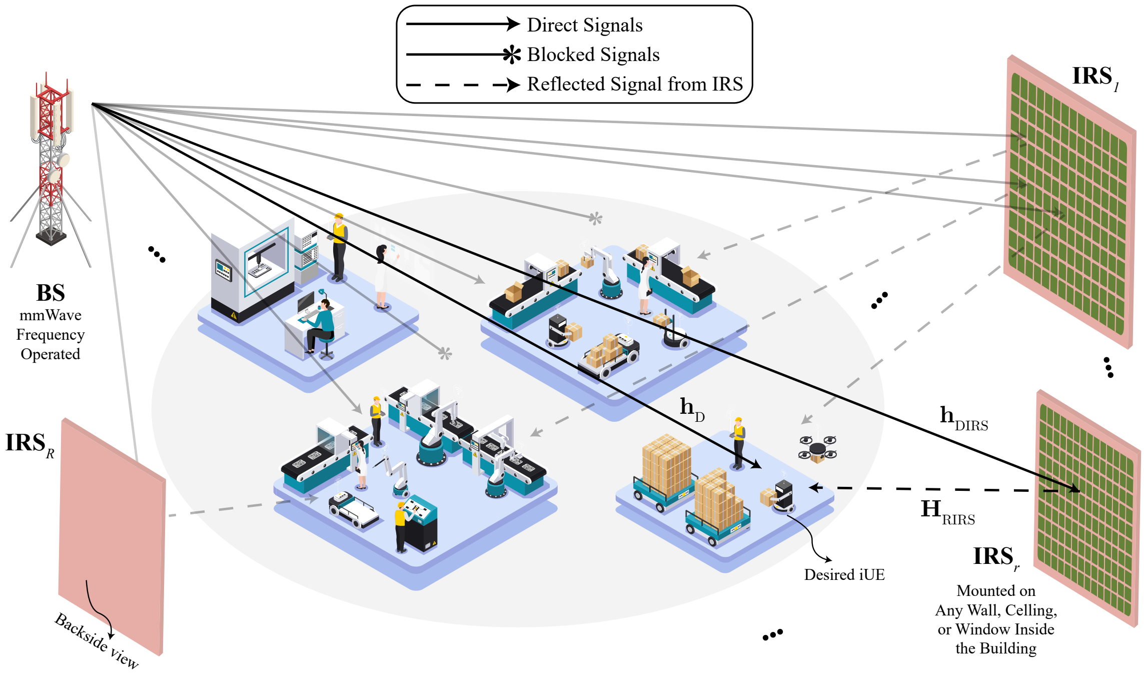

Figure 7.2: Illustration of the considered framework of an IRS-assisted mmWave smart industry communication for 6G and beyond networks.

trial environments. This combination of IRS with mmWave technology is not just theoretical but offers practical benefits in enhancing signal coverage and reliability.

The evolution to Industry 4.0 necessitates a new level of sophistication in wireless communications, a demand that IRS technology is poised to meet [AMBA23]. In the setting of smart industries, where connectivity between machines and devices is increasingly critical, the ability to direct and refine wireless signals becomes essential. As illustrated in Figure 7.2, the strategic use of IRS can lead to substantial improvements in data throughput and network performance, ensuring consistent connectivity even in areas lacking direct line-of-sight [NR22]. IRS technology presents an innovative solution to the challenges faced in indoor wireless networks, such as signal degradation due to obstructions and multipath fading. By manipulating the signal's phase and amplitude, IRSs can transform obstructive elements into advantageous components of the signal path, enhancing signal strength and coverage. This ability is especially beneficial in the intricate layouts of industrial environments, where maintaining reliable signal coverage is vital for operational efficiency. Despite the significant potential of IRS technology, it remains an area ripe for research and development. Integrating IRS with mmWave communication systems, particularly in the context of smart industries, opens up a realm of possibilities for further exploration. Future research could focus on refining IRS design and placement strategies, improving signal manipulation capabilities, and fully realizing the potential of the IRS in various industrial scenarios. The foundational concepts laid out in Figure 7.1 and the IRS-assisted communication framework illustrated in Figure 7.2 represent just the starting point in the journey toward more advanced and efficient wireless communication systems for the future of the industry.

7.1.1 Concept and Evolution

The evolution of Intelligent Reflecting Surfaces (IRS) in millimeter-wave (mmWave) communication marks a significant advancement in wireless technology for Industry 4.0, addressing mmWave's coverage limitations and atmospheric absorption issues. As industrial environments demand rapid and reliable communication, integrating IRS with mmWave systems enhances performance by improving signal quality and coverage, especially in obstructed indoor settings. Current research focuses on optimizing the placement and number of IRS elements, with future directions exploring advanced beamforming, dynamic adjustments, and AI integration. This technology is poised to revolutionize connectivity and efficiency in smart industries.

7.1.2 Significance in 5G/6G Networks

The integration of Intelligent Reflecting Surfaces (IRS) with millimeter-wave (mmWave) wireless communication is crucial for the advancement of 5G and 6G networks, addressing significant challenges such as limited range and vulnerability to obstructions, especially in complex environments. While mmWave technologies are essential for high data rates and low latency in 5G, their deployment is hindered by propagation limitations; IRS provides a solution by enhancing mmWave coverage. As we approach the 6G era, the demand for higher data rates and wider coverage underscores the importance of IRS in overcoming mmWave's range and penetration issues. Additionally, IRS operates passively by manipulating signals without requiring additional power sources, enhancing energy efficiency while reducing reliance on extra base stations. This aligns with sustainable industrial practices and broader sustainable technology goals. Additionally, IRS optimizes signal paths, enhancing network capacity and efficiency in densely populated areas, thus alleviating congestion and ensuring quality service. This synergy of IRS and mmWave technologies not only enhances network performance but also supports energy efficiency and sustainable development, making IRS a key component in realizing the full potential of 5G and shaping the future of 6G networks.

7.1.3 Role in Smart Industries

Intelligent Reflecting Surfaces (IRS) play a crucial role in enhancing millimeter-wave (mmWave) communication for smart industries, addressing propagation challenges in complex environments. By controlling electromagnetic waves, IRS improves signal coverage and quality, facilitating the deployment of sensor networks and IoT devices essential for Industry 4.0. This technology also promotes energy efficiency by enhancing signal reach without increasing transmitter power, making it a sustainable solution. Additionally, IRS-assisted mmWave systems support advanced manufacturing techniques, ensuring reliable communication for automation and safety in industrial settings, as illustrated in Figure 7.2.

7.2 Fundamentals of Millimeter-Wave Communication

Millimeter-wave (mmWave) communication operates between 30 GHz and 300 GHz, offering high data rates and capacities essential for 5G and emerging 6G technologies. Its wide bandwidth enables significant data transmission, making it crucial for broadband cellular networks and IoT applications. However, mmWave signals face challenges like high propagation loss, susceptibility to atmospheric absorption, and the need for dense base station deployment due to their directional nature. Advances in antenna technology and signal processing are addressing these challenges, positioning mmWave as a key technology for the future of wireless communications.

7.2.1 Characteristics of mmWave Frequencies

Millimeter-wave (mmWave) frequencies, ranging from 30 GHz to 300 GHz, offer high bandwidth for superior data rates, benefiting applications in smart industries and IoT. However, they also face challenges like high path loss and atmospheric absorption, leading to signal degradation and necessitating denser network infrastructures. Their directional nature allows for precise beamforming, enhancing signal strength while reducing interference in crowded environments. To leverage mmWave frequencies effectively, especially in industrial settings, innovative solutions like Intelligent Reflecting Surfaces (IRS) must be integrated to mitigate these challenges [UM20].

7.2.2 Challenges in mmWave Transmission

- **High Path Loss:** One of the primary challenges in mmWave transmission is the high path loss, especially in non-line-of-sight (nLoS) conditions. This is quantitatively expressed by the equation 7.1 where 'd' represents the distance between the transmitter and receiver, and 'fc' is the central frequency. The high path loss in mmWave frequencies, particularly as the distance increases, significantly impacts the effective range of communication. This high path loss necessitates the need for advanced technologies, such as beamforming and IRS, to maintain signal strength and ensure reliable communication over longer distances. The path loss in an indoor factory environment can be mathematically represented by the following set of equations 7.1, 7.2, 7.4 which are borrowed from ETSI [ETS22].

$$PL_{\text{nLoS}} = 33 + 25.50 \log_{10}(d) + 20 \log_{10}(f_c), \tag{7.1}$$

  In the equation above and system of equations below, PL_{LoS}, PL_{nLoS}, PL, f_c, and d, denote the *LoS Path Loss*, *nLoS Path Loss*, overall *Path Loss*, central frequency, and distance between transmitter and receiver, respectively.

- **Susceptibility to Obstructions:** mmWave frequencies are highly susceptible to physical obstructions, which can severely attenuate or completely block the signal. In environments with many obstacles, such as factories or urban areas, this can lead to a substantial reduction in signal quality and reliability. The Line-of-Sight (LoS) path loss equation 7.2, highlights the importance of a clear LoS for optimal signal transmission. However, maintaining a clear LoS is often challenging in dynamic and cluttered environments, necessitating the development of robust strategies to overcome these obstructions.

$$PL_{\text{LoS}} = 31.84 + 21.50 \log_{10}(d) + 19 \log_{10}(f_c), \tag{7.2}$$

This equation highlights the relationship between distance (d) and frequency (f_c) on signal attenuation. LoS conditions lead to a logarithmic increase in path loss as distance increases, with higher frequencies resulting in greater attenuation. The equation shows that without obstacles, signal strength decreases in a predictable, gradual manner.

In contrast, Non-Line of Sight (NLoS) communication faces additional losses due to reflections, diffraction, and scattering from obstacles, which significantly increase path loss. In an NLoS scenario, the path loss is generally higher than LoS, often quantified as:

$$\mathrm{PL}_{\mathrm{NLoS}} = \mathrm{PL}_{\mathrm{LoS}} + \text{Additional Losses} \tag{7.3}$$

These additional losses depend on environmental factors and can lead to a steeper path loss curve. As a result, LoS communication generally offers better performance and range, while NLoS reduces signal quality and range due to increased attenuation.

- **Atmospheric Absorption:** mmWave signals are subject to atmospheric absorption, particularly due to oxygen and water vapor, which can significantly degrade signal strength. This challenge is inherent to the mmWave frequency bands and poses a considerable limitation in outdoor environments. The overall path loss, denoted as in equation 7.4, captures the maximum impact of these losses, whether in LoS or nLoS conditions. Mitigating atmospheric absorption requires careful planning of network infrastructure and the use of technologies like IRS to enhance signal strength and coverage.

$$PL = \max\left(PL_{\mathrm{LoS}}, PL_{\mathrm{nLoS}}\right), \tag{7.4}$$

- **Limited Penetration and Diffraction:** mmWave frequencies have limited ability to penetrate solid objects or diffract around obstacles. This limitation poses a significant challenge in indoor environments, where signals must traverse walls and other structures. The equations 7.2 and 7.1 path loss demonstrate how signal strength diminishes with increasing distance and in the presence of obstructions. To address this, mmWave communication systems often require a higher density of access points and reflective surfaces to ensure comprehensive coverage within buildings and other obstructed areas.

- **Complex Network Topology and Management:** The challenges of high path loss, susceptibility to obstructions, atmospheric absorption, and limited penetration necessitate a complex network topology. This includes a dense deployment of base stations and the integration of IRS and other reflective technologies. Managing such a complex network, especially in dynamic industrial environments, requires sophisticated algorithms for beamforming, network coordination, and real-time adaptation to changing conditions. The intricate relationship between distance, frequency, and environmental factors, as indicated by the path loss equations, underscores the complexity of optimizing mmWave transmission for consistent and reliable communication.

7.2.3 Advantages in Industrial Applications

The integration of Intelligent Reflecting Surfaces (IRS) into millimeter-wave (mmWave) communication systems offers several advantages for industrial applications within the context of Industry 4.0. These advantages, however, are subject to practical deployment conditions and system design constraints.

- **Enhanced Data Transmission Rates:** IRS can enhance achievable data rates by creating additional controllable propagation paths and improving the effective channel gain. By appropriately configuring the IRS phase shifts, reflected signals can constructively combine with the direct signal, leading to increased signal-to-noise ratio (SNR) and, consequently, higher achievable rates. This improvement is particularly relevant in mmWave systems where blockage severely limits throughput.
- **Improved Coverage and Signal Quality:** In industrial environments characterized by metallic machinery and obstructions, mmWave links are highly susceptible to blockage. IRS enables non-line-of-sight (NLoS) communication by reflecting signals around obstacles, thereby extending coverage into shadowed regions and improving received signal strength and link robustness.
- **Energy Efficiency:** IRS elements are passive and do not require radio-frequency chains or power amplifiers. As a result, IRS-assisted communication can reduce the required transmit power at the base station while maintaining target quality-of-service levels. This property makes IRS particularly attractive for energy-efficient industrial wireless networks.

- **Scalability and Flexibility:** IRS panels consist of a large number of low-cost reflecting elements and can be easily deployed on walls, ceilings, or factory infrastructure. Their passive nature and software-controlled phase configuration allow scalable and flexible adaptation to dynamic industrial layouts without significant hardware modification.
- **Reduced Interference and Enhanced Reliability:** By controlling the phase response of the reflected signals, IRS can spatially shape the wireless environment to suppress interference and enhance desired signal components. This capability improves link reliability, especially for mission-critical industrial applications requiring low outage probability.
- **Facilitation of IoT and Sensor Networks:** Industrial Internet-of-Things (IoT) deployments involve dense sensor networks with strict energy and coverage requirements. IRS can enhance connectivity for low-power IoT devices by improving link quality without increasing transmit power. However, this benefit depends on proper IRS placement and channel state information availability.
- **Improved Safety and Monitoring:** IRS-assisted communication can support enhanced sensing, localization, and monitoring by improving signal visibility in obstructed environments. This capability enables more reliable asset tracking, equipment monitoring, and situational awareness in smart factories, indirectly contributing to improved operational safety.
- **Cost-Effectiveness:** Compared to active relays or additional base stations, IRS deployment incurs lower hardware and operational costs due to its passive nature. IRS can improve coverage and performance without the need for dense active infrastructure, making it a cost-effective enhancement for industrial wireless networks.

These advantages position IRS-assisted mmWave communication as a transformative solution for industrial applications within Industry 4.0.

7.3 IRS Technology in Enhancing mmWave Communication

The integration of Intelligent Reflecting Surfaces (IRS) into millimeter-wave (mmWave) communication systems significantly enhances wireless communication in modern industrial settings. Intelligent Reflecting Surfaces (IRS) manipulate electromagnetic waves by dynamically controlling their phase and amplitude, enabling precise reflection steering. This enhances signal strength, mitigates propagation challenges in mmWave communications, and improves both coverage and signal quality by effectively redirecting signals around obstacles. This technology extends network coverage and reliability, enabling focused beamforming that optimizes signal transmission in dense environments.

7.3.1 Relevent and State of Art Work

Intelligent Reflecting Surfaces (IRS), essentially large programmable metasurfaces, can dramatically reshape the millimeter-wave (mmWave) propagation environment to overcome key challenges in 5G and 6G communications. By intelligently adjusting the phase responses of hundreds of passive reflecting elements, an IRS can establish virtual line-of-sight (LoS) links around obstacles, thereby enhancing the received signal power and signal-to-interference-plus-noise ratio (SINR) in non-LoS and shadowed scenarios. As a result, IRS-assisted systems extend coverage and transform fast-fading channels into more stable links, leading to significant improvements in spectral efficiency and link reliability. Moreover, since IRS elements operate in a passive manner without requiring active radio-frequency chains, they can achieve these performance gains with very low additional energy consumption, offering substantially higher energy efficiency compared to conventional active relaying techniques. Recent IEEE studies validate these advantages in practical mmWave deployments; for instance, [KKAC24] demonstrated that the integration of IRS into mmWave systems can yield over 200% average throughput improvement while significantly reducing the required transmit power as the number of reflecting elements increases [KKAC24]. Overall, IRS-enabled beam steering and coverage enhancement effectively address the severe path loss and blockage issues of mmWave communications, establishing IRS as a key enabling technology for high-frequency 5G and 6G networks (see also [WZR24]).

7.3.2 Principles of IRS in Signal Reflection

Intelligent Reflecting Surfaces (IRS) are essential in enhancing signal reflection for advanced wireless communication, particularly in millimeter-wave (mmWave) technologies for smart industries. This passive technology does not require external power, making it a cost-effective and energy-efficient solution.

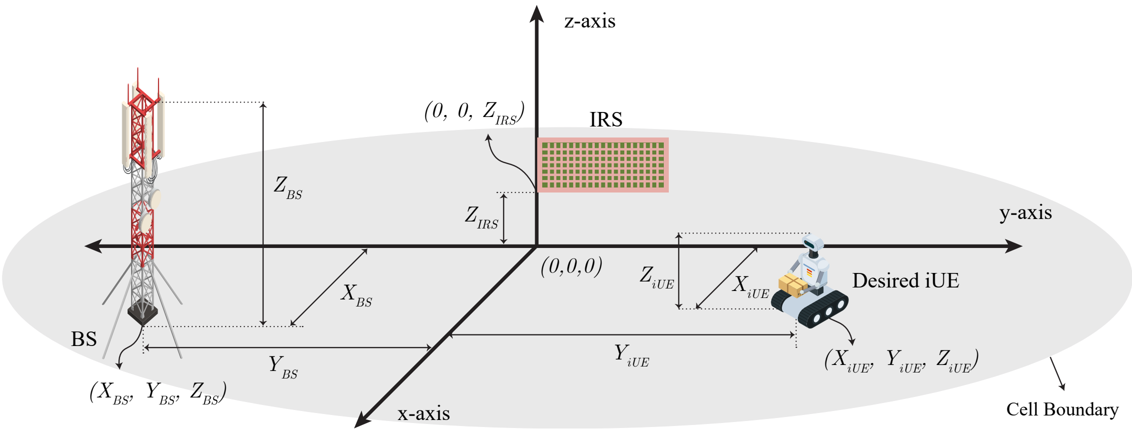

Figure 7.3: Geometric configuration for proposed IRS assisted mmWave smart industry networks in 6G and beyond.

The deployment of IRS requires strategic placement to optimize signal transmission in complex industrial environments, as illustrated in Figure 7.3. The geometric composition drawn from the considered IRS assisted mmWave wireless communication network in the smart industry is depicted in the figure 7.3. In which an IRS is considered to be located at the height of Z_{IRS} on origin. Moreover, BS and iUE are located at (X_{BS}, Y_{BS}, Z_{BS}) and $(X_{iUE}, Y_{iUE}, Z_{iUE})$, respectively.

7.3.3 Optimization Techniques for IRS

- **Element Phase Shift and Amplitude Optimization:** The optimization of Intelligent Reflecting Surfaces (IRS) in mmWave communication systems primarily involves tuning the phase shifts and amplitudes of the IRS elements. The reflection coefficient matrix, denoted as $\mathbf{\Phi}$, where $\mathbf{\Phi} = diag([\alpha_1 \exp(j\Phi_1), \alpha_2 \exp(j\Phi_2), \ldots, \alpha_R \exp(j\Phi_R)]^T)$, plays a crucial role in this process. Here, α_r and ϕ_r represent the amplitude and phase shift of the r^{th} element of the IRS, respectively. By carefully adjusting these parameters, IRS can dynamically alter the propagation path of the incoming mmWave signals, effectively enhancing the signal strength and coverage, as well as mitigating multi-path fading and shadowing effects. Here α_r and ϕ_r denote the amplitude and phase shift of the rth IRS element, respectively. In our system, each phase shift ϕ_r is chosen to maximize the combined received signal power. Specifically, ϕ_r is set such that the reflected signal aligns in phase with the direct signal at the receiver, thereby achieving constructive interference. Let h_D denote the direct channel, and let $(\mathbf{H}_{\text{RIRS}}\mathbf{h}_{\text{DIRS}})_r$ represent the reflected channel component associated with the rth IRS element. The optimal phase shift is then given by

$$\phi_r = \angle h_D - \angle \left(\mathbf{H}_{\text{RIRS}} \mathbf{h}_{\text{DIRS}}\right)_r \pmod{2\pi}, \tag{7.5}$$

which ensures coherent addition of the direct and reflected signal components and maximizes the received signal-to-noise ratio (SNR).

- **Signal Path Optimization:** The signal received by the user equipment (iUE), expressed as mentioned in equation 7.6, where $\mathbf{h}_{\text{D}}$ and $\mathbf{H}_{\text{RIRS}}$ represent the direct and IRS-reflected channels, respectively, is a function of the IRS's ability to manipulate the signal path. The optimization involves adjusting $\boldsymbol{\Phi}$ such that the combined direct and reflected paths provide the optimal signal-to-noise ratio (SNR). This is crucial in environments where Line-of-Sight (LoS) is obstructed, and non-LoS (nLoS) paths play a significant role in communication.

$$\mathbf{y} = (\mathbf{h}_{\text{D}} + \mathbf{H}_{\text{RIRS}} \boldsymbol{\Phi} \mathbf{h}_{\text{DIRS}}) x + \mathbf{n}, \tag{7.6}$$

- **Maximizing Achievable Data Rate:** The achievable data rate, given in equation 7.8, is a key performance metric in mmWave communication systems. The optimization of the IRS should aim at maximizing this rate, which involves not only enhancing the SNR but also ensuring that the system can handle high data rates efficiently. The signal-to-noise ratio (SNR) at the iUE, can be expressed as

$$\text{SNR} = \frac{P_{\text{TX}} \|\mathbf{h}_{\text{D}} + \mathbf{H}_{\text{RIRS}} \boldsymbol{\Phi} \mathbf{h}_{\text{DIRS}}\|^2}{N_\circ}, \tag{7.7}$$

This requires a comprehensive approach that includes both hardware capabilities and algorithmic strategies to manage the complex interplay of signal reflection, interference, and noise.

$$R_{\text{achievable}} = B \cdot \log_2 (1 + \text{SNR}). \tag{7.8}$$

B is the Channel Bandwidth. This figure 7.4 demonstrates the relationship between transmit power (P_{Tx}) and sum throughput (S) for two modulation rates, $R = 64$ and $R = 256$. The throughput increases with transmit power, as shown by the two curves. The general behavior can be described by the following equation:

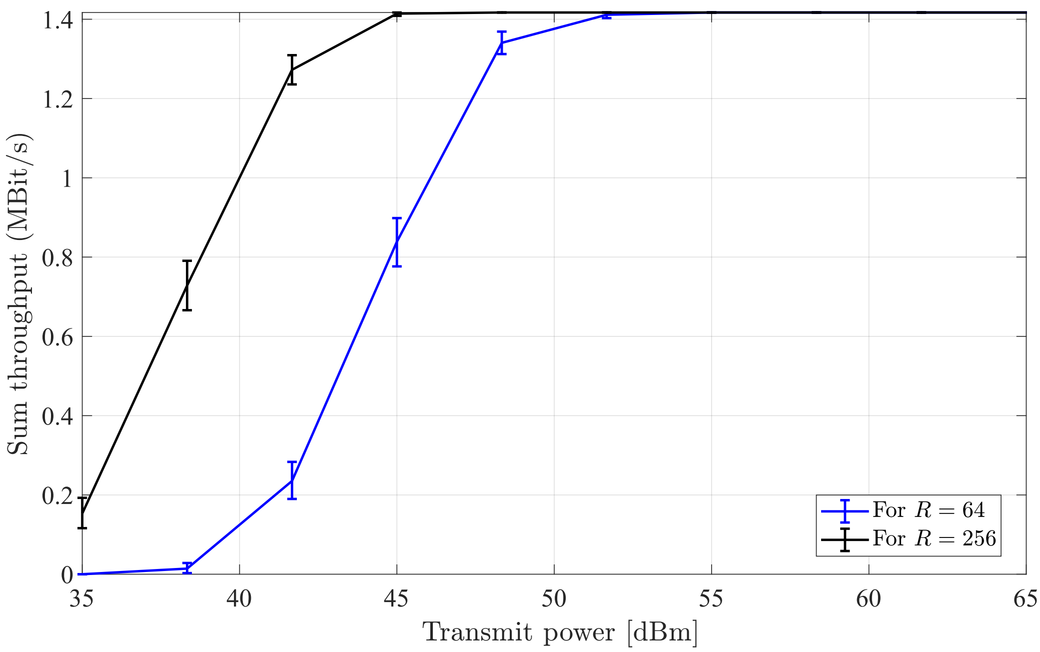

Figure 7.4: Comparison of the sum throughput of IRS-assisted mmWave communication network in the smart industry with a different number of IRS elements (R). Using 256 RIS elements provides 75% higher user throughput at 45 dBm compared to 64 elements.

$$S = \frac{C}{1 + \left(\frac{P_{\text{Tx}}}{P_0}\right)^{\alpha}} \tag{7.9}$$

Where C is the maximum throughput capacity, P_{Tx} is the transmit power, P_0 is the reference transmit power, and α is a constant that characterizes the channel conditions. The curve for $R = 64$ shows a slower increase in throughput compared to $R = 256$, as the system with higher modulation rate R can achieve a higher throughput at a given transmit power.

- **IRS Element Count Optimization:** The simulation was performed using Vienna 5G Link Level Simulator [PTM$^+$18b]. The framework was evaluated assuming a base station height of 30 meters and a transmission power of 45 dBm [HTM23]. The impact of the number of IRS elements on system performance is significant. The study 'Comparison of the sum throughput of IRS-assisted mmWave communication network in the smart industry with a different number of IRS elements' highlights that increasing the number of elements can substantially enhance the system's throughput. However, this increase must be balanced against practical considerations such as the physical size of the IRS, the complexity of control, and power consumption. Figure 7.4 is a graphical representation that illustrates the relationship between the number of IRS elements and the network's sum throughput. The *sum throughput* represents the total achievable rate in the system, accounting for contributions from all users in the network. It is computed as:

$$R_{\text{sum}} = B \cdot \sum_{i=1}^{N_{\text{UE}}} R_i = B \cdot \sum_{i=1}^{N_{\text{UE}}} \log_2 (1 + \text{SNR}_i) \tag{7.10}$$

Where R_{sum} denotes the *aggregate throughput* of the system, and N_{UE} is the total number of user equipment (UEs). The achievable rate for an individual user, R_i, is derived from the *signal-to-noise ratio (SNR)* at that user. The *SNR* for each user in the IRS-assisted system is given by:

$$\text{SNR}_i = \frac{P_{\text{TX}} \left\| h_D + H_{\text{RIRS}} \Phi h_{\text{DIRS}} \right\|^2}{N_0} \tag{7.11}$$

Where P_{TX} is the **transmit power** at the base station, h_D is the **direct channel gain** between the base station and the user, H_{RIRS} represents the **channel gain from the base station to the IRS**, Φ is the **IRS reflection matrix** that controls the phase shifts of the IRS elements, h_{DIRS} denotes the **channel gain from the IRS to the user**, and N_0 is the **noise power spectral density**. This equation models the *impact of IRS on sum throughput*, illustrating how **intelligent reflections** can optimize wireless link quality and **enhance spectral efficiency**. As shown in **Figure 7.4**, the sum throughput increases with a higher *number of IRS elements* and greater *transmit power*, aligning with the results from the Vienna 5G Link Level Simulator [PTM+18b].

In the figure 7.4, the sum throughput, representing the total data successfully transmitted and received over a mmWave network within a given time period, illustrates the network's capacity to handle multiple user connections and data flows simultaneously. As shown in Figure, the x-axis denotes the transmit power (dBm), while the y-axis quantifies the sum throughput, typically measured in bits per second (bps). The figure highlights a clear trend of increasing throughput with higher transmit power and a greater number of IRS elements, denoted as R, ranging from smaller to larger arrays.

Higher transmit power directly enhances received power, resulting in improved throughput. Similarly, increasing the number of IRS elements provides substantial gains in throughput due to their ability to reflect and manipulate incident signals. With more reflective surfaces, signals can be intelligently redirected and optimized, leading to stronger reception and better overall system performance. The graph's curve demonstrates how these factors synergistically improve throughput, showcasing the enhanced data-handling capacity of the network.

This scalability of IRS elements, depicted in the figure, underscores their critical role in boosting communication quality and efficiency, particularly in scenarios demanding high data rates, such as smart industry applications. By strategically increasing IRS arrays and combining this with higher transmit power, networks can leverage intelligent signal reflection to expand coverage, strengthen communication links, and achieve significantly higher throughput levels. The insights from this analysis highlight the transformative potential of IRS deployment in practical mmWave communication scenarios.

- **Beamforming Integration:** Beamforming is essential for directing mmWave signals toward specific users or target areas, thereby improving communication link quality and overall system performance [KTJ+22, CNBD22]. Researchers have developed sophisticated beamforming algorithms, including hybrid and adaptive beamforming, to enhance the directionality and alignment of mmWave signals. These advancements lead to improved coverage and increased capacity, significantly boosting the performance of mmWave communication systems [AFIM22, NLD22, LH22, LGZ+22, SKSK23].

  Beamforming techniques, which involve directing the signal towards specific receivers, can be integrated with IRS optimization. This integration facilitates the focusing of mmWave signals more precisely, enhancing the effective use of the available spectrum. The synergy between beamforming and IRS optimization is particularly beneficial in dense urban environments or industrial settings, where directing signals effectively can significantly improve connectivity and reduce interference.

- **Adaptive Optimization in Dynamic Environments:** In dynamic industrial environments, the optimization of IRS must be adaptive. This requires real-time adjustments to the IRS's phase shifts and amplitudes in response to changing environmental conditions and user mobility. Machine learning algorithms can play a crucial role here, learning from the environment and user behavior to predict optimal IRS settings.

- **Energy Efficiency and Sustainability:** While optimizing the performance metrics such as throughput and data rate, it's essential to consider the energy efficiency of the IRS-assisted mmWave system. The optimization algorithms should aim at minimizing power consumption while maintaining high performance. This is particularly important for sustainable deployment in smart industries, where energy efficiency is a key consideration alongside communication performance.

7.3.4 IRS and Beamforming Strategies

The integration of Intelligent Reflecting Surfaces (IRS) with beamforming strategies significantly advances wireless communication, particularly in high-frequency millimeter-wave (mmWave) networks. IRS enhances signal coverage and strength by using passive reflecting elements to manipulate signal paths, addressing the limitations of mmWave propagation in complex environments. This combination allows for dynamic optimization of signal paths, improving reliability and capacity while minimizing interference. Despite challenges in implementation, IRS-assisted beamforming is crucial for developing advanced wireless systems, especially as networks evolve toward 6G.

7.4 Case Studies: Application of Intelligent Reflecting Surfaces (IRS) in Industrial Environments

The application of Intelligent Reflecting Surfaces (IRS) in industrial settings has demonstrated transformative potential across domains such as smart manufacturing, logistics, warehousing, and automated quality-control systems. Recent studies show that IRS can significantly enhance wireless communication reliability and coverage by reconfiguring the radio environment, which is particularly beneficial in industrial scenarios characterized by dense metallic structures, machinery-induced blockages, and severe multipath effects [BRdR+20, RNS+20]. In smart manufacturing environments, IRS-assisted networks improve the robustness of data transmission and enable dense deployments of industrial Internet-of-Things (IIoT) devices while supporting stringent latency and reliability requirements [RIT22]. Similarly, in logistics and warehousing applications, IRS can mitigate signal blockages and support advanced technologies such as automated guided vehicles (AGVs) and real-time inventory tracking by maintaining reliable connectivity in non-line-of-sight conditions [ZLMS24, LZMS24]. Moreover, experimental and system-level studies indicate that IRS-enhanced millimeter-wave links can improve energy efficiency and link stability, which is critical for precision-driven applications such as automated inspection and quality-control systems [DNF24]. Overall, these studies highlight the practical benefits of IRS technology and underscore its role in improving operational efficiency, reliability, and energy efficiency in industrial wireless networks.

7.4.1 Placement and Configuration of IRS

The placement of Intelligent Reflecting Surfaces (IRS) in mmWave networks is a critical factor that significantly influences the network's performance. As demonstrated in Figure 7.5, there is a clear relationship between the IRS's position relative to the user equipment (iUE) and the achievable data rate. For mmWave frequencies (fc = 28 GHz), the achievable rate is higher closer to the IRS, highlighting the importance of strategic placement to maximize coverage and data transmission efficiency. This implies that in dense urban environments or industrial settings, where high data throughput is essential, the IRS must be strategically located to ensure optimal coverage and signal strength, particularly considering the directional nature of mmWave signals.

Figure 7.6, which suggests that the height at which IRS panels are installed plays a significant role in network performance. A higher IRS position generally leads to better line-of-sight (LoS) conditions and a wider coverage area, thereby enhancing the achievable capacity. This becomes particularly relevant in urban landscapes or complex industrial setups where physical obstructions are commonplace. An elevated IRS can effectively circumvent these obstructions, providing a more reliable and stronger signal to the intended receivers.

The difference in performance between mmWave and sub-6 GHz frequencies, as indicated in Figure 7.5, necessitates distinct configuration strategies for IRS in different frequency bands. While mmWave frequencies offer higher data rates, they are more susceptible to signal degradation over distance and obstacles. In contrast, sub-6 GHz frequencies, characterized by better penetration and wider

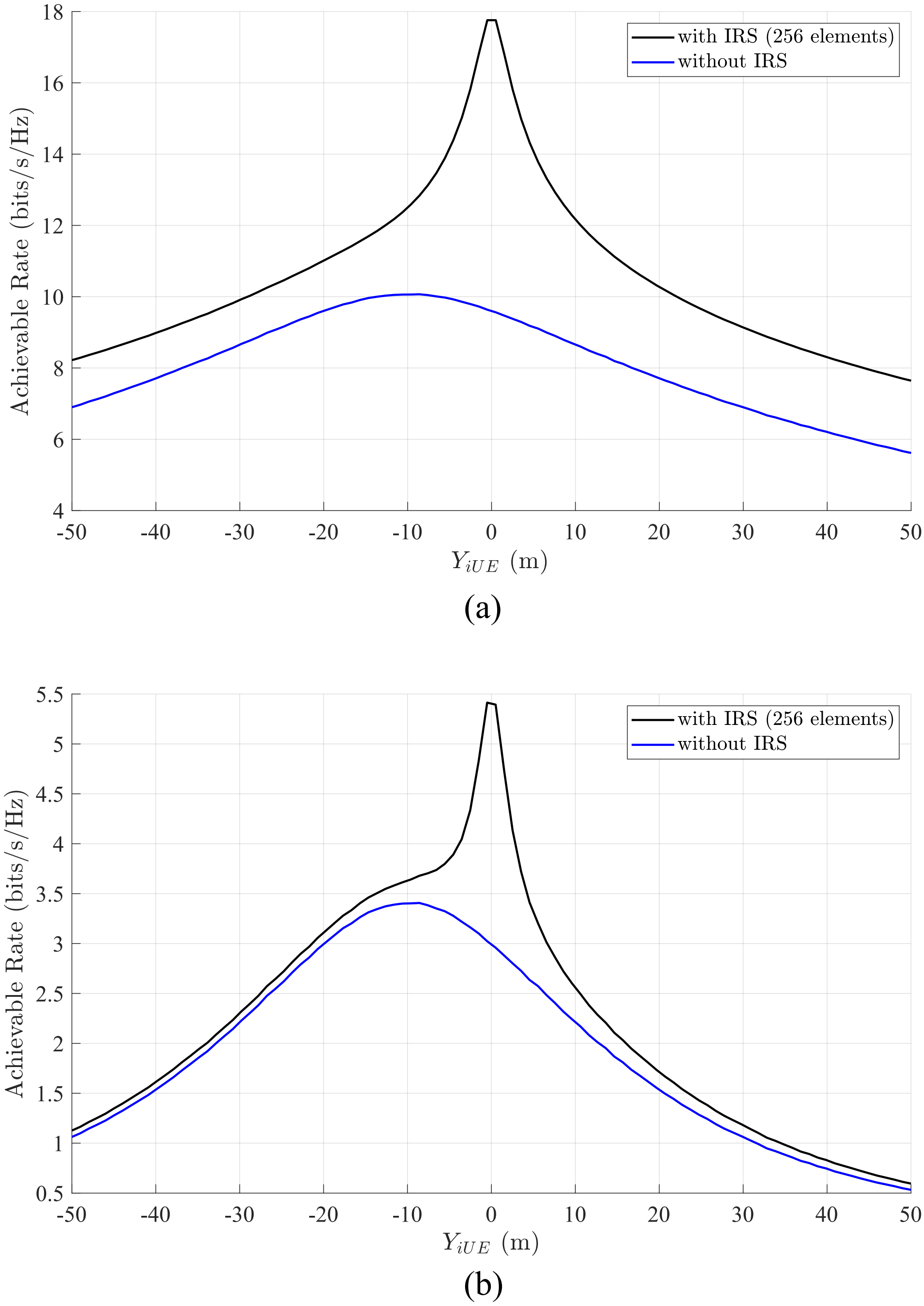

Figure 7.5: Achievable Rate vs. Distance for iUE from IRS (YiUE) at ZIRS = 6m for different frequencies: (a) mmWave (fc = 28 GHz), and (b) sub-6 GHz (fc = 2.4 GHz)

coverage, may require fewer or differently positioned IRS elements. This dis-

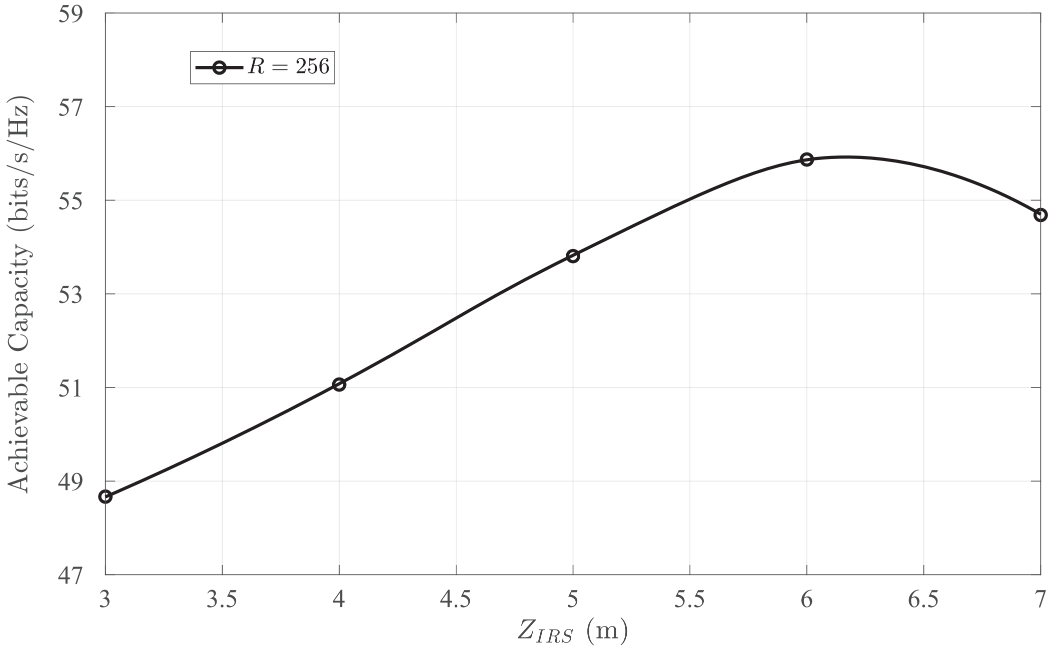

Figure 7.6: Impact of the height of IRS from the ground on achievable capacity with YiUE + YBS = 80m. Positioning the IRS at the optimal height (6 m) yields 14% capacity gain over the lowest tested height (3

tinction underscores the need for frequency-specific IRS configurations, tailored to leverage the unique propagation characteristics of each frequency band.

Beyond static placement, the dynamic adjustment of IRS elements is crucial for adapting to varying network conditions and user requirements. This involves real-time tuning of the phase shifts and amplitudes of the IRS elements to optimize signal reflection and refraction. Such dynamic adjustments can compensate for user movement, changing environmental conditions, and fluctuations in network load, ensuring consistent and high-quality service across the network.

When configuring IRS placement and orientation, environmental factors such as building layouts, natural terrain, and typical user movement patterns must be considered. These factors can dramatically affect signal propagation, especially in mmWave networks. For instance, in an urban setting, the presence of tall buildings might necessitate a higher density of IRS panels, strategically placed to create a network of reflective paths that mitigate signal blockages.

The design of an IRS-enabled network requires a careful balance between coverage and capacity. While increasing the height of IRS, as shown in Figure 7.6, can enhance coverage, it may not always translate to proportional increases in capacity, especially if the IRS elements are too far from the users. Therefore, network designers must find an optimal balance, possibly through a combination of elevated IRS panels for broader coverage and strategically placed lower-height panels for capacity enhancement in high-demand areas.

Lastly, the placement and configuration of IRS must consider long-term sustainability and cost-effectiveness. This includes not only the initial installation costs but also the maintenance, energy consumption, and potential upgrades of the IRS system. The goal should be to create a network that not only meets the current performance requirements but also remains viable and efficient in the face of evolving technologies and changing user demands.

The placement and configuration of IRS in mmWave networks require a multifaceted approach that considers various technical, environmental, and practical factors. By strategically addressing these aspects, network designers can optimize IRS-enabled systems to provide robust, efficient, and high-capacity wireless communication services.

7.4.2 Integration with Existing Infrastructure

Integrating Intelligent Reflecting Surfaces (IRS) into existing wireless network infrastructure presents challenges due to the need for seamless compatibility with current technologies, which operate under established standards and protocols. Existing networks are designed for direct communication paths, while IRS introduces an additional reflective path, necessitating new planning and optimization strategies. Careful attention to compatibility is essential to ensure IRS enhances rather than disrupts service, potentially requiring upgrades to network software, control algorithms, and base station hardware. Key to this integration is the synchronization between IRS and base stations, involving real-time communication to align signals and adapt to changing conditions. Advanced signal processing and predictive algorithms may be required to maintain service quality while managing coexistence with legacy systems through dynamic spectrum management and resource allocation. Additionally, integrating IRS raises security and privacy concerns, as its external control can introduce vulnerabilities. Robust security protocols, including encryption and secure authentication, are vital to protect data transmitted via IRS, while privacy measures must comply with regulations to prevent misuse of user information. Thus, integrating IRS requires a comprehensive strategy addressing technical performance, security, and privacy to ensure network integrity and user trust.

Performance Gains from IRS Integration

Our results demonstrate significant performance improvements through the integration of IRS into millimeter-wave smart factory networks:

- **Spectral Efficiency Gain at mmWave 28 GHz (7.5:a)** The achievable rate with IRS at optimal distance (Y $\approx$ 0 m) is approximately 18 bits/s/Hz, compared to 10 bits/s/Hz without IRS:

  $$\text{Spectral Efficiency Gain} = \frac{18}{10} = 1.8\times$$

  This represents an 80% increase in spectral efficiency.

- **Spectral Efficiency Gain at Sub-6 GHz (7.5:b)** The achievable rate with IRS at optimal distance is approximately 5.2 bits/s/Hz, compared to 3.3 bits/s/Hz without IRS:

  $$\text{Spectral Efficiency Gain} = \frac{5.2}{3.3} = 1.58\times$$

  This represents a 58% increase in spectral efficiency at sub-6 GHz.

- **Transmit Power Reduction (Figure 7.7):** At Y $\approx$ 70 m, the required transmit power values are:

  $$P_{\text{Non-IRS}} = 10^{\frac{62}{10}} = 1.58 \times 10^6 \text{ mW}, \quad P_{\text{IRS}} = 10^{\frac{52}{10}} = 1.58 \times 10^5 \text{ mW}$$

  The percentage reduction in required transmit power is:

  $$\frac{P_{\text{Non-IRS}} - P_{\text{IRS}}}{P_{\text{Non-IRS}}} \times 100 = \frac{1.58 \times 10^6 - 1.58 \times 10^5}{1.58 \times 10^6} \times 100 = 89.9\%$$

  Thus, IRS integration reduces required transmit power by approximately 90%.

- **Capacity Gain via IRS Height Optimization (Figure ??):** The achievable capacity increases from approximately 49 bits/s/Hz at $Z_{\text{IRS}} = 3$ m to 57 bits/s/Hz at $Z_{\text{IRS}} = 6$ m:

  $$\text{Capacity Gain} = \frac{57 - 49}{49} \times 100 = 16.3\%$$

7.4.3 Energy Efficiency and Sustainability

A key aspect of energy efficiency in IRS-enabled networks involves minimizing the power requirements for signal transmission.

We modeled a mmWave link with a base station (BS) at 30m height transmitting at 45dBm. The IRS is placed at height 6m with $R = 256$ elements (as in Fig.7.4). All users and the BS are separated by horizontal distance Y. These parameters match those used in the Vienna 5G simulator.

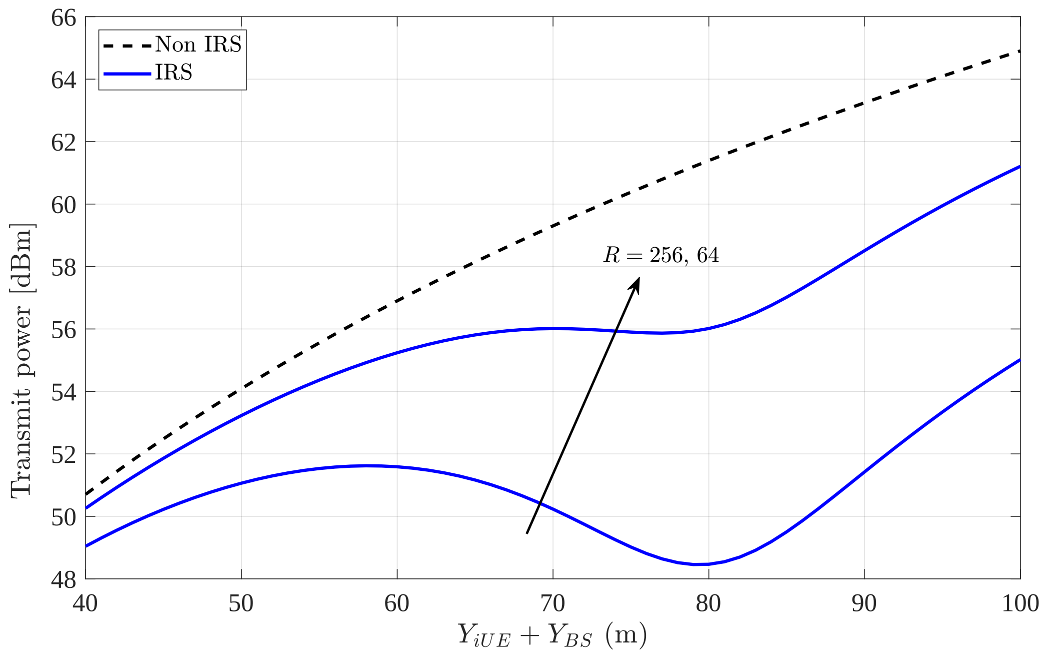

Figure 7.7: Comparison of the required Tx Power vs. Distance (i.e., BS to iUE) of the proposed framework at $Z_{IRS} = 6m$ with the non-IRS system.IRS reduces required transmit power by approximately 10 dB at 70 m compared to non-IRS

Figure 7.7 illustrates how Intelligent Reflecting Surfaces (IRS) can reduce the transmission power required to maintain communication quality over varying distances between the base station (BS) and user equipment (iUE).

By leveraging the reflective properties of IRS elements, network operators can enhance signal strength and lower transmission power, particularly significant over longer distances where signal attenuation is more pronounced. IRS elements are passive, requiring no power to amplify signals, but their control mechanisms consume energy, necessitating energy-efficient designs using low-power electronics and intelligent algorithms. Additionally, deployment strategies should focus on maximizing coverage with minimal elements to reduce material use and energy consumption, while lifecycle impacts must be considered to minimize the overall carbon footprint. Integrating renewable energy sources, such as solar panels, and exploring energy harvesting technologies can further enhance sustainability. A holistic approach to energy management across the network, utilizing advanced analytical tools and AI algorithms, is essential for optimizing energy usage and reducing waste. In summary, achieving energy efficiency and sustainability in IRS-enabled networks involves thoughtful design, strategic deployment, renewable energy integration, and comprehensive energy management.

7.4.4 Novelty of Analysis

Unlike many RIS studies, our analysis focuses on an industrial mmWave scenario with explicit geometry and IRS height effects. For example, Fig.7.7 now shows required BS power vs. distance for the IRS (red) vs. no-IRS case (blue), under the above assumptions. At $Y = 70$m the IRS solution needs 10dB less power (90% reduction).

7.5 Challenges and Limitations of IRS in Industrial Environments

In industrial environments, the deployment of Intelligent Reflecting Surfaces (IRS) faces unique challenges and limitations, stemming primarily from the complex and dynamic nature of these settings. Industrial spaces are often characterized by heavy machinery, large metallic structures, and constant movement, which can significantly impact the propagation of electromagnetic waves. The reflective surfaces of IRS must be strategically placed and frequently adjusted to overcome signal blockages and ensure consistent coverage, a task complicated by the ever-changing layout and operational conditions of industrial facilities. Additionally, the harsh environmental conditions, such as extreme temperatures, dust, and vibrations, pose durability and maintenance challenges for IRS components. Moreover, the integration of IRS with existing industrial communication systems requires careful consideration to avoid interference with critical operations. These factors necessitate robust and adaptive IRS designs, capable of withstanding industrial rigors while maintaining optimal performance, presenting a substantial hurdle in fully harnessing the potential of IRS technology in such demanding environments.

7.5.1 Technical Challenges

Deploying Intelligent Reflecting Surfaces (IRS) in industrial environments faces several technical challenges, primarily related to managing signal reflection and interference amidst complex electromagnetic landscapes filled with metallic surfaces and machinery. This unpredictability necessitates sophisticated algorithms to adapt to varying propagation conditions while ensuring the IRS does not disrupt existing wireless systems. The harsh industrial conditions, including extreme temperatures and vibrations, require IRS components to be robust and durable, often necessitating rugged materials and automated self-calibration systems to maintain alignment. Additionally, the dynamic nature of modern industrial settings complicates integration with IoT devices and automated machinery, requiring IRS to adjust settings seamlessly while interfacing with the facility's control systems. Continuous 24/7 operations also demand energy-efficient designs and ease of maintenance, making modularity and remote monitoring crucial for minimizing downtime. Addressing these diverse challenges necessitates a multifaceted approach that combines advanced technology, thoughtful design, and strategic integration.

7.5.2 Scalability Issues

Scalability issues for IRS in industrial environments include challenges in expanding coverage across large facilities. Deploying IRS in vast spaces requires strategic placement to ensure comprehensive coverage, which can increase costs and complexity. Managing a larger IRS network demands advanced systems for real-time control and coordination of numerous panels to avoid signal interference. As networks grow, energy consumption becomes a concern, and energy-efficient strategies are essential. The cost of large-scale IRS deployment is significant, so industries must balance initial investments with long-term benefits. Adaptability and future-proofing of IRS systems are also key to addressing scalability challenges.

7.5.3 Security and Privacy Concerns

Security concerns for IRS in industrial environments include the risk of unauthorized access and data interception. IRS panels can manipulate wireless signals, making them potential targets for hackers. Strong encryption, access control, and intrusion detection are essential to protect data and IRS operations.

Signal jamming and denial of service (DoS) attacks pose another challenge, as they can disrupt industrial processes. To counter this, redundancy, real-time monitoring, and swift response measures are crucial. Protecting sensitive data requires end-to-end encryption and strict access controls, while the physical security of IRS panels should be ensured through surveillance and tamper-detection mechanisms.

7.6 Conclusion: The Impact and Future Outlook of IRS in Industry 4.0

In conclusion, the integration of Intelligent Reflecting Surfaces (IRS) in Industry 4.0 heralds a transformative impact on the landscape of industrial wireless communication and automation. As industries evolve to embrace more interconnected, intelligent, and efficient systems, the role of IRS emerges as a critical enabler. By enhancing millimeter-wave (mmWave) communication, facilitating robust IoT networks, and integrating seamlessly with edge computing, IRS technology addresses some of the most pressing challenges in industrial wireless networks, including signal coverage, reliability, and energy efficiency. Looking towards the future, IRS is poised to play a pivotal role in the advancement of smart industries, driving innovations in material science, AI integration, and network optimization. As we venture into the era of 5G and beyond, IRS technology is expected to unlock new potentials, not only optimizing current industrial processes but also paving the way for novel applications and smarter industrial ecosystems. The continued exploration and development of IRS will undoubtedly be a cornerstone in realizing the full promise of Industry 4.0.

7.6.1 Summary of Key Findings

The study reveals that IRS-assisted mmWave wireless communication significantly enhances network performance, achieving approximately 200% efficiency compared to sub-6 GHz frequencies in smart factories. Notably, a significant increase in throughput was observed along with a reduction in the required transmit power when the number of IRS elements was raised (e.g., R = 64 and R = 256). Higher IRS placement enhances system capacity by improving Line-of-Sight (LoS) paths. Future advancements include next-generation beamforming, dynamic IRS systems, and the integration of AI for optimized communication.

Chapter 8

Machine Learning-Driven Performance Analysis of Compressed Communication in Aerial-RIS Networks for Future 6G Network

"The true logic of this world is in the calculus of probabilities."

James Clerk Maxwell

8.1 Introduction

Reconfigurable Intelligent Surfaces (RIS) have garnered significant attention as a solution to enhance the performance of wireless communication networks beyond the fifth generation (B5G) and sixth generation (6G). As urbanization and skyscrapers increase in cities, there is a growing need for an efficient and cost-effective method to enable Non-Line-of-Sight (NLOS) communication, which can be facilitated by RIS. A RIS consists of an array of passive reflective elements on a 2-D meta-surface. Each element can be individually programmed to enhance signal coverage by adjusting both phase and amplitude, enabling efficient and intelligent steering so that the signals are constructively combined at the receiver's end, resulting in higher data rates, improved spectral efficiency, and ubiquitous connectivity while being eco-friendly and cost-efficient. RIS creates a Virtual Line of Sight (VLOS) while passively recycling the signal without any external interference. This makes RIS a crucial enabler for the envisioned smart radio environments of 6G networks. [DRDPH+20, BDRDR+19].

The advancement of technology has shifted the surge of bandwidth utilization from Orthogonal Multiple Access (OMA) to Non-Orthogonal Multiple Access (NOMA). OMA, which includes TDMA, FDMA, and CDMA, technologies were susceptible to bandwidth exhaustion that could be utilized for other cell users. In contrast, NOMA (Non-Orthogonal Multiple Access) improves bandwidth efficiency and overall network capacity by allowing multiple users to share the same frequency band simultaneously. This requires successive interference cancellation (SIC) to separate user signals based on power levels [JHB21a, DLC+17, ea18]. By leveraging power domain multiplexing, NOMA can serve multiple users, enhancing throughput and reducing latency [DLC+17]. The integration of RIS with NOMA systems has been shown to further amplify these advantages by optimizing the signal-to-interference-plus-noise ratio (SINR) and providing more flexible resource allocation [MLG+20].

Integrating CoMP with NOMA techniques is promising for mitigating severe inter-cell interference effects in multi-cell NOMA networks and improving spectral efficiency [ea20b, UMH+23, ea14, UMG+24]. Coordinated multi-point (CoMP) techniques use high-speed front-haul connections and sharing channel state information (CSI) among BSs, that enable coordinated transmissions and enhance overall network performance. However, the challenge of coordination among all BSs still arises due to issues with CSI inaccuracies, synchronization across cells, and increased signal processing requirements.

While the integration of RIS and NOMA has shown promising results in improving spectral efficiency and reducing outage probability [ea20c], the combination of RIS, NOMA, and CoMP has not been thoroughly investigated, especially with aerial RIS deployments. Aerial RIS, mounted on platforms such as unmanned aerial vehicles (UAVs), can provide flexible and dynamic coverage enhancements, adapting to changing network conditions in real-time [ea20a]. According to the author's best knowledge, no performance analysis exists for Aerial-RIS-aided CoMP-NOMA cellular environments.

Managing RIS-assisted systems requires phase control for enhanced performance. Many studies have been made for finding optimal phase shifts for different optimization objectives [GXW+21, PRW+20]. Previous studies assumed that these phase shifts are already available at the RIS end in quantized form, but that is not the case. These Quantized Phase Shifts require a feedback channel from the receiver to RIS, which is difficult to obtain due to bandwidth limitations. Bandwidth efficiency or Spectrum efficiency is a crucial requirement for efficient data transmission. A promising approach is to use machine learning-based autoencoders for data compression to achieve this efficiency. Autoencoder is a special type of neural network that can learn to represent the data in a lower dimension and then decode it back to its original form, saving significant bandwidth. The paper investigates using autoencoders to compress the QPS symbols into lower dimensions for a smooth transmission over the feedback channel and decode them at the user end. We can achieve substantial bandwidth reduction while maintaining data integrity by leveraging autoencoders' encoding and decoding capabilities.

Motivated by the literature gap, we present a detailed Machine Learning performance analysis for compressed communication of a system integrating aerial RIS, CoMP, and NOMA. This integrated approach not only enhances performance metrics such as throughput, latency, and coverage but also paves the way for innovative applications in diverse fields ranging from smart cities to autonomous driving [LYX+20, LZL+21]. Our contributions are threefold:

- We developed a novel system model for aerial RIS-assisted CoMP-NOMA networks, considering strategic placement of RIS for far-user gain enhancement.
- We analyze the impact of various parameters, including RIS elements, and power allocation coefficients, on the system's spectral efficiency and outage probability.
- Our system model proposes the far user as a part of two NOMA clusters and our simulation results demonstrate an increase in data rate, network capacity, coverage, and a decrease in outage probabilities.
- We made a comparative analysis of different autoencoder architectures utilized for compressed communication of encoded QPS bits and observed the reconstruction accuracy of our data from all models.

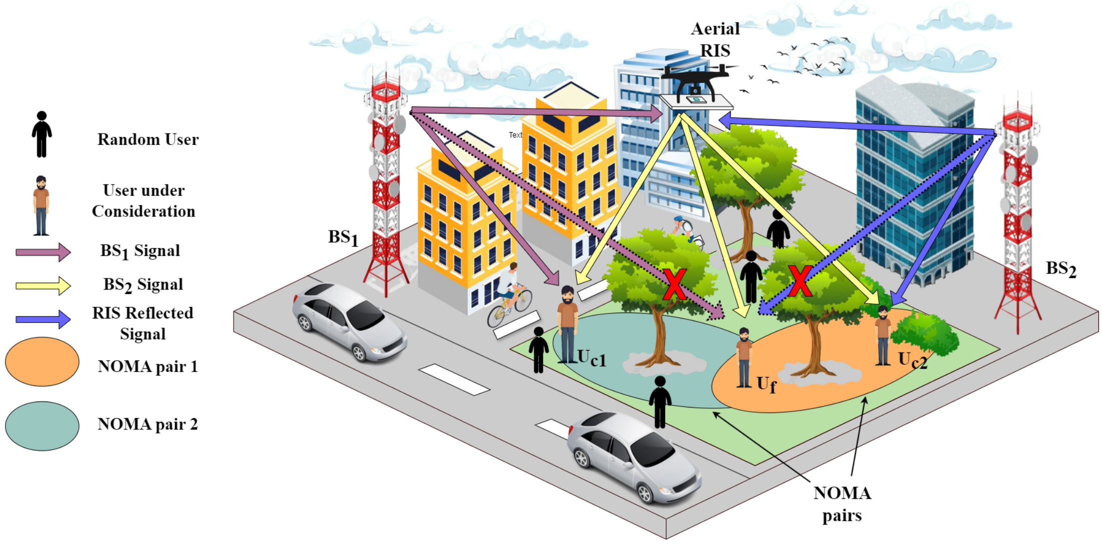

Figure 8.1: System Model: Power Allocation for UEs (UEn, UEf) with Reconfigurable Intelligent Surface (RIS) Aided Communication

8.2 System Model and Problem Formulation

8.2.1 System Description

As described in Fig. 8.2, we consider a two-cell narrow band CoMP-NOMA system with an Aerial-RIS placed strategically at the cell edge of the two cells to facilitate both cell users. The downlink communication system operates with frequency-flat channels. The Aerial RIS consists of M elements. Let the index for the base stations be $\mathcal{C} = \{1, 2, \ldots, C\}$, $\mathcal{N} = \{1, 2, \ldots, N\}$ for the two near users and $\mathcal{F} = \{1, 2, \ldots, F\}$ for the far cell user. Moreover, let $\mathcal{U} = \{\mathcal{N} \cup \mathcal{F}\}$ be the set of combined users in the system. In Fig. 8.1 a conceptual wireless communication network for a 3D urban environment is visualized, demonstrating the practical use case of our system model. In this urban environment, the direct LOS path is blocked due to obstacles in the environment, and hence aerial RIS provides VLOS. This is a typical example of massive connectivity envisioned for next-gen wireless networks.

To simplify our analysis, we focus on the system model depicted in Fig. 8.2. In this model, each base station is equipped with a single antenna, and the receivers—namely UE$_{c,n}$ and UE$_f$—also have single antennas, establishing a SISO (Single Input Single Output) channel. Furthermore, all $n \in \mathcal{N}$ and $f \in \mathcal{F}$ are located within the reflection region of the aerial RIS, allowing all users in the cellular environment to receive RIS links.

Both base stations use the power domain NOMA technique to transmit signals to their respective NOMA pairs, as shown in Fig. 8.2. Each BS_c forms a NOMA pair consisting of $UE_{c,n}$ as the near user and UE_f as the far user. Notably, UE_f is part of both NOMA clusters served by both base stations BS_c. To further enhance the performance of UE_f, we adopt Coordinated Multi-point (CoMP), which mitigates inter-cell interference and increases the data rate for this user. We consider a two-cell downlink network as shown in Fig.2, where each single-antenna base station (BS) serves one "near" and one "far" user via power-domain NOMA, and the far user is jointly served by both BSs (CoMP). Each BS $c \in 1, 2$ transmits a superposition of its two user signals: for BS c,

$$y_c = \sqrt{\gamma_{c,n} P_c}\, x_{c,n} + \sqrt{\gamma_{c,f} P_c}\, x_{c,f}, \tag{8.1}$$

where $x_{c,n}$ and $x_{c,f}$ are the data symbols for the near and far users, P_c is the transmit power, and $\gamma_{c,n}, \gamma_{c,f}$ are the NOMA power fractions[1][2]. The aerial RIS has M passive elements (e.g. $M = 70$ in our simulations. The channels from each BS to each user combine direct and reflected paths. For base station c and user u, the effective channel is

$$H_{c,u} = h_{c,u} + h_{R,u}^H \Theta h_{c,R}. \tag{8.2}$$

where Θ is the RIS phase shift matrix and discussed in the next section. The received signal at the near user of $UE_{1,n}$ is

$$z_{1,n} = H_{1,n} y_1 + H_{2,n} y_2 + n_0, \tag{8.3}$$

where $H_{1,n} = h_{1,n} + h_{R,n}^H \Theta h_{1,R}$ and $H_{2,n} = h_{2,n} + h_{R,n}^H \Theta h_{2,R}$. $UE_{1,n}$ first decodes the far user's symbol x via SIC and then decodes its own. Similarly, the far user UE_f receives

$$z_f = H_{1,f} y_1 + H_{2,f} y_2 + n_0, \tag{8.4}$$

where $H_{1,f} = h_{1,f} + h_{R,f}^H \Theta h_{1,R}$ and $H_{2,f} = h_{2,f} + h_{R,f}^H \Theta h_{2,R}$. The SINR at $UE_{1,n}$ for decoding the far-user signal is

$$\zeta_{1,n \to f} = \frac{\gamma_{1,f} P_1 |H_{1,n}|^2}{\gamma_{1,n} P_1 |H_{1,n}|^2 + P_2 |H_{2,n}|^2 + \sigma^2}. \tag{8.5}$$

In this research article, we assume perfect channel conditions, implying that perfect Channel State Information (CSI) is available at both base stations BS_c.

Although estimating perfect CSI in real-life scenarios is a highly complex task, recent advancements in research have demonstrated the feasibility of estimating the cascaded wireless channels—from BS_c to the Reconfigurable Intelligent Surface (RIS) and from the RIS to the users ($\text{UE}_{c,n}$, UE_f)—using machine learning and other signal processing techniques [HLWJ20], [WHA+21].

8.2.2 Channel Model and RIS Configuration

In our system model, all communication links account for both large-scale fading and small-scale fading effects. We assume frequency-flat, quasi-static channels. Given our 3D urban environment modeling, the significant distance between base stations BS_c and users ($\text{UE}_{c,n}$, UE_f) results in channels exhibiting Rayleigh fading, characterized by:

$$\mathbf{h}_{c,u} = \sqrt{\frac{PL(d_0)}{\text{Pd}(d_{c,n})}}\mathbf{w}_{c,n}, \tag{8.6}$$

where $\mathbf{w}_{c,u}$ is a complex Gaussian random variable following a Rayleigh distribution with zero mean and unit variance. PL denotes the reference path-loss at a reference distance d_0 where d_0 is taken as 1 meter, and $\text{Pd}(d_{c,u})$ represents the large-scale path-loss modeled as $\text{Pd}(d_{c,u}) = (d_{c,u})^{\alpha_{i\to u}}$, where $d_{c,u}$ is the distance between BS_c and UE_u, and $\alpha_{i\to u}$ is the path-loss exponent.

In contrast, the communication link between the Aerial RIS and BS_1 manifests a Line of sight (LoS) path [GQLAD21]. As a result, these links are influenced by Rician fading, with their channel coefficient expressed as:

$$\mathbf{h}_{c,ARIS} = \sqrt{\frac{PL(d_0)}{\text{Pd}(d_{c,ARIS})}}\left(\sqrt{\frac{\kappa_{c,ARIS}}{\kappa_{c,ARIS}+1}}\hat{\mathbf{v}}_{c,ARIS} + \sqrt{\frac{1}{\kappa_{c,ARIS}+1}}\mathbf{v}_{c,ARIS}\right) \tag{8.7}$$

where $d_{c,ARIS}$ represents the distance between the BSc and the Aerial RIS, $\kappa_{c,ARIS}$ represents the Rician factor, $\mathbf{v}_{c,\hat{ARIS}}$ denotes the deterministic LoS components, and $\mathbf{v}_{c,ARIS}$ denotes the complex Gaussian random variables. Similarly, the communication links between ARIS and receivers, namely $\text{UE}_{c,n}$ and UE_f— are modeled. This model ensures that $PL(d_0)$ captures the free-space loss (Friis) at 1m, and then scales with $d^{\alpha_{i\to n}}$. We thus justify using the reference-distance model: it decouples the known free-space loss at d_0 from the distance-scaling exponent. This approach is common and allows the same $PL(d_0)$ to apply across links (here $d_0 = 1$m). The RIS is assumed lossless with unity amplitude gain, so it only imparts phase shifts (cf. Θ). The energy splitting (ES) model of the Aerial RIS array is described by its coefficient matrices as follows [DKN+21]:

$$\boldsymbol{\Theta} = \mathrm{diag}(e^{j\theta_1}, e^{j\theta_2}, \ldots, e^{j\theta_M}), \tag{8.8}$$

where $\theta_M \in [0, 2\pi)$, $\forall M \in \mathcal{M} \triangleq \{1, 2, \ldots, M\}$. where M denotes the total number of RIS elements. Each element provides a unit-magnitude reflection coefficient of the form $e^{j\theta_m}$, where $\theta_m \in [0, 2\pi)$ represents the adjustable phase shift applied by the m-th RIS element. The resulting diagonal phase-shift matrix multiplies the incident signal vector, thereby imposing a programmable phase rotation on each propagation path. In this manner, the RIS passively steers the reflected signal by appropriately tuning the phase shifts $\{\theta_m\}$, without performing any signal amplification. Under the adopted energy-splitting (ES) model, all RIS elements are assumed to have equal reflection amplitudes, and only the phase responses are optimized.

The reference-distance-based path-loss formulation, using $PL(d_0)$ at a reference distance of $d_0 = 1$ m, is deliberately adopted in this chapter to accurately model the high-frequency, predominantly line-of-sight (LoS) propagation conditions associated with mmWave and aerial RIS-assisted links. Unlike the standardized multi-segment path-loss models employed in earlier chapters—which typically incorporate empirical breakpoints, shadowing, and environment-specific corrections—the adopted model isolates distance-based attenuation using a single path-loss exponent. This simplification is well justified for mmWave and UAV-enabled communications, where propagation is often close to free-space or lightly obstructed conditions.

Moreover, frequency-dependent losses are implicitly captured in the reference path-loss term $PL(d_0)$, while distance-dependent attenuation is modeled via the scaling factor $(d)^{-\alpha}$. Given the severe attenuation characteristics of mmWave signals, even a short reference distance of 1 m provides a meaningful normalization point. By consistently applying the same $PL(d_0)/d^{\alpha}$ formulation in both (8.1) and (8.2), the proposed model ensures a unified and physically interpretable treatment of range-dependent effects for both Rayleigh and Rician fading channels.

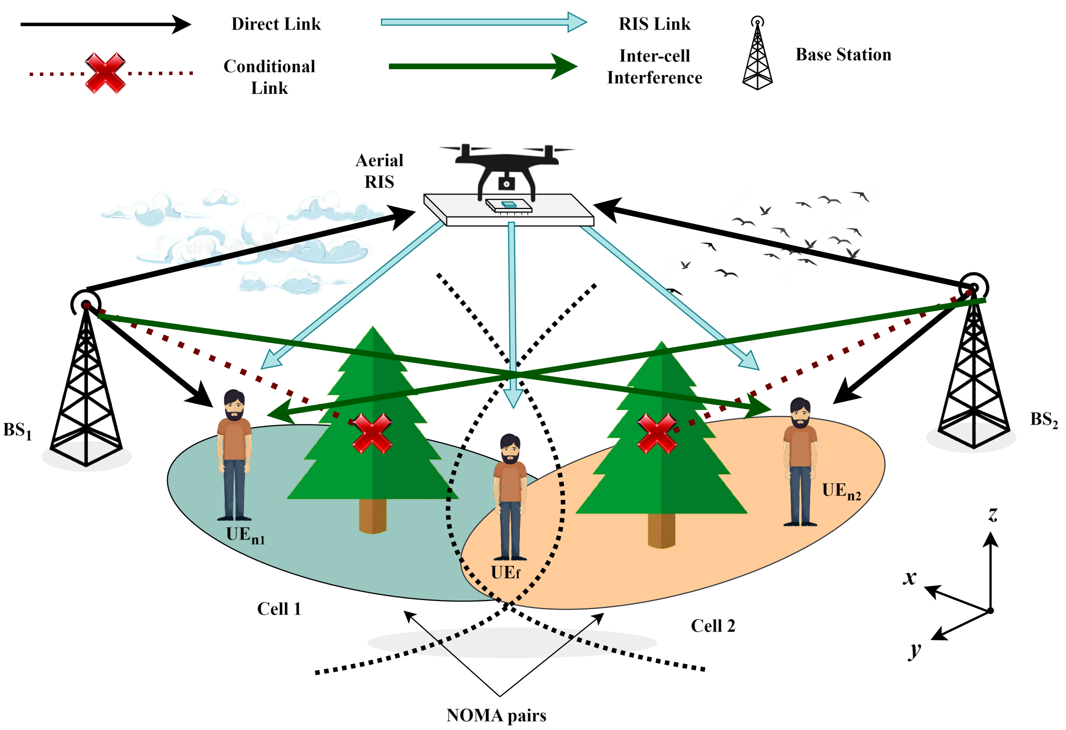

Figure 8.2: Simplified System Model

8.2.3 Problem Formulation

From the above analysis, the joint optimization problem for maximizing the network sum rate in a single coordinated NOMA cluster aided by the Aerial RIS can be formulated as:

$$
\begin{aligned}
\max_{\mathcal{A},\boldsymbol{\Theta},\mathbf{M}_A} \quad & \mathcal{R}_{\text{sum}} = \sum_{c=1}^{2} \mathcal{R}_{c,n} + \mathcal{R}_f, \qquad (8.9)\\
\text{s.t.} \quad & \mathcal{R}_f \geq \mathrm{R}_{\text{min}}^{f}, \forall f \in \mathcal{F},\\
& \mathcal{R}_{c,n} \geq \mathrm{R}_{\text{min}}^{c,n}, \forall c \in \mathcal{C}, n \in \mathcal{N},\\
& \gamma_{c,n} + \gamma_f \leq 1, \forall c \in \mathcal{C}, n \in \mathcal{N}, f \in \mathcal{F},\\
& \theta_m \in [0, 2\pi), m \in \mathcal{M}\\
& \sum_{1}^{c} \mathbf{M}_A^c \leq M, c \in \mathcal{C}, m \in \mathcal{M},
\end{aligned}
$$

where $\mathrm{R}_{\text{min}}^{f}$ and $\mathrm{R}_{\text{min}}^{c,n}$ are the minimal achievable rates at UE_f and $\text{UE}_{c,n}$, respectively and $\mathcal{A}$ represents the PA factor of a single NOMA pair within the cluster. $\boldsymbol{\Theta}$ encompasses all the phase shifts associated with Aerial RIS, while $\mathbf{M}_A$ denotes the allocation of Aerial RIS resources among the BSs.

The optimal phase shift for each element $m \in \mathcal{M}$, given the CSI being fully estimated at the BSs, can be computed, as referenced [WZ19], to solve the objective in (8.9) as follows:

$$\theta_m = \mathrm{mod}[\arg(\mathbf{h}_{i,R}) - \arg(\mathbf{h}_{i,R} \cdot \mathbf{h}_{R,i}),\, 2\pi], \tag{8.10}$$

where the phase of the channels is determined using the argument function $\arg(\cdot)$. This equation is used to compute the optimal phase shift θ_m for the m^{th} element of an Aerial Reconfigurable Intelligent Surface (RIS). The purpose is to adjust the RIS reflection in a way that constructively combines signals at the receiver, thereby maximizing the received signal strength or the overall sum rate.

To tackle the joint network sum-rate optimization problem (8.9), we employ fixed empirical optimizations. Additionally, we investigate the effect of allocating Aerial-RIS elements to the BSs by thoroughly examining all possible splitting configurations through an exhaustive iteration process.

Advanced optimization techniques might provide additional improvements. However, this work primarily aims to demonstrate the fundamental benefits achieved by strategically deploying the Aerial RIS and allocating its resources among the BSs in the coordinated cluster. Due to the limited resolution of the RIS, the phase angle can only assume a finite set of discrete values, typically determined by the quantization process.

8.2.4 Machine Learning approach

The limited resolution of the RIS results in assuming a finite number of discrete values for the phase angle θ_m through quantization. θ_m can be represented with 2^I quantization levels, where I is the number of quantization bits used. θ_m being uniformly quantized, can be expressed in a set of quantized phase shifts (QPS) as $\{0, \frac{2\pi}{2^I}, \ldots, \frac{(2^I-1)2\pi}{2^I}\}$. This indicates, Phase Shift Keying (PSK), as an appropriate method for mapping these quantized phases [YLXL21]. In our case, the phase information dataset $\{x_i\}_{i=1}^{m} \subset \mathbb{R}^d$, is quantized and encoded to $I \in \{9\}$ bits, before being transmitted over the communication channel. The quantized bits I are compressed to O bits, where $O \in \{3, 4\}$, resulting in a significant reduction in bandwidth utilization.

Compressed communication was achieved through the implementation of autoencoders. Autoencoders were utilized to encode the QPS bits at the transmitter end and subsequently decode the noisy compressed QPS bits, to retain the original discrete phase shift, aiming to attain efficient communication over a band-limited feedback channel. The models effectively capture the compressed and abstract features of input data, representing them in a latent space and reconstructing the input from the encoded representation. The autoencoder can be represented as:

$$\begin{cases} \hat{e} = g(X; w_e, b_e) \\ \hat{d} = h(X_e; w_d, b_d) \end{cases}$$

where $g(\cdot)$ and $h(\cdot)$ are the encoder and decoder functions respectively, with each implemented with four different architectures. w_e and b_e are the parameters of the encoder, while w_d and b_d are parameters of the decoder function and represent the weight matrices and bias vectors of the encoder and decoder networks. We employed four distinct autoencoder architectures for a comparative analysis of our feedback compression mechanism discussed below

8.2.4.1 CNN Autoencoder

Convolutional Neural Network CNN is a robust architecture for detecting features and patterns from the input data [Kha24, KUU23]. The encoder implementation includes 3 back-to-back Conv + Rectified Linear Unit (ReLU) (a combination of convolution and ReLU activation function) followed by a simple Conv (a simple convolutional layer) layer. Each layer has a typical kernel size of 3 with N number of neurons where N is taken as 64. The transmitted QPS are fed to the encoder, represented by I number of bits, with each bit acting as a separate feature where they undergo a convolution process that is given by:

$$Y_n = (X * w_e)_n = \sum_{k=0}^{K-1} X_k \cdot w_{e,n-k} + b_e \tag{8.11}$$

where X is the input vector and Y is the output feature vector at n^{th} position or layer. K is the total length of the filter.

The ReLU activation function introduces non-linearity which improves the model's ability to learn unique features, particularly the bits associated with QPS in our case. The ReLU function is expressed as:

$$f(x) = \max(0, X_n) \tag{8.12}$$

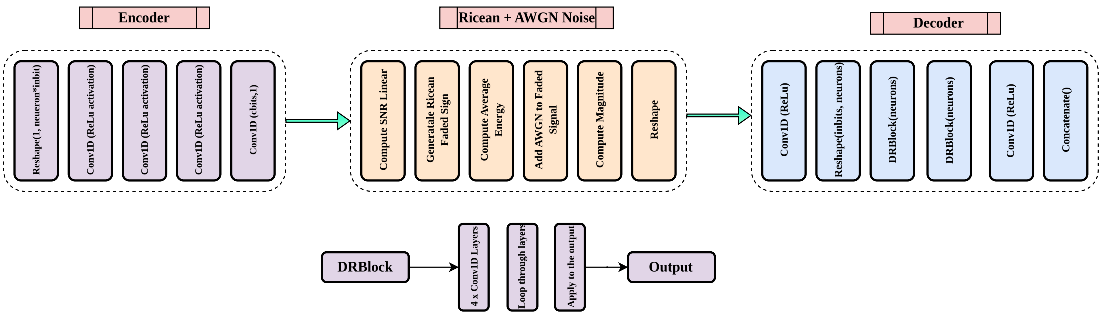

Figure 8.3: Flowchart illustrating transmission process of bits from encoder to decoder

The decoder is slightly more complex than the encoder as the reconstruction of the original input requires a comparatively more intricate architecture. Instead of mere Convolutional Layers stacked together, Deep Residual blocks (DRBlocks) are used. In addition to Convolutional layers. The DRBlock is a set of 4 Conv + ReLU layers with an addition of a skip connection that is bridged from the input directly to the output of Convolutional Layers. Overall, the decoder block starts with a Conv layer joined with 2 DRBlocks and a Conv + Sigmoid (a combination of convolutional and sigmoid) as the final layer, with the sigmoid function mapping the output value to [0, 1]. The sigmoid function is given as:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{8.13}$$

8.2.4.2 CNN+Attention Autoencoder

With the addition of Attention Mechanism (AM), the performance of the CNN model further enhances [NZY21]. The AM enables the model to autonomously learn to prioritize the most relevant portions of the input. Here, a three-step attention mechanism is utilized. (i) Spatial Attention: it includes a reshaped layer followed by a 1-dimensional Conv + Sigmoid layer producing spatial feature maps. An element-wise multiplication is performed with a spatial attention map to highlight significant spatial regions. (ii) Channel Attention: the channel attention permutes and reshapes the spatial output to align the channels followed by applying a Conv+sigmoid layer. Similarly, an element-wise multiplication is performed with a channel attention map, highlighting the important channels in the data. (iii) Joint-Channel Attention: a 2D Conv + Sigmoid layer is utilized to integrate features across different channels, followed by multiplication with a joint-channel attention map. The final output is the summation of the results of all the mechanisms, producing a more focused and informative representation of the input features.

The encoder is designed with 3 AM layers alternated with three Conv + ReLU layers, each using the same kernel size and number of filters as in CNN architecture. The final output is obtained through a Conv layer with O units for a 1xO representation. The decoder is similar to that of the encoder's design, but instead of Conv+ReLU layers, DRBlocks are utilized, similar to those found in CNN decoder architecture. Similarly, a Conv + Sigmoid layer is employed in the end to regularize the output.

8.2.4.3 RNN Autoencoder

Recurrent Neural Network (RNN) is a deep learning model for sequential data processing. The architecture begins with 3 stacked Simple RNN layers, each with N number of neurons and ReLU activation function. The output is then reshaped to a 1 x NI vector and fed to a final Dense layer for a final compression, mapping the input 1 x I vector to a compressed 1 x O. The decoder begins with a dense layer to expand the encoded input into a higher dimension, with a 1 x NI vector representation. The vector is reshaped to form a compatible format with the subsequent RNN layers and then passed through such stack of three Simple Recurrent Neural Network (SimpleRNN) layers with N number of neurons and ReLU activation functions in each layer. In the end, the output of the RNN stack is passed through a time-distributed dense layer, where each time step is processed independently by a dense layer with a single unit and a sigmoid activation function. This final layer produces the reconstructed sequence, with each element in the sequence constrained to the range [0, 1].

8.2.4.4 Transformer Autoencoder

A Transformer Layer is used that comprises of Multi-head Self-attention Mechanism and a Feed-Forward Network (FFN). The multi-head attention mechanism enables the model to capture complex inter-dependencies across different parts of the input sequence, allowing it to focus on various parts of the sequence simultaneously. In our case, 4 multi-heads were used. The Attention Layer is followed by an FFN, comprising a Dense layer with 32 units and a ReLU activation function. FFN applies these non-linear transformations, allowing for more complex representations, followed by dropout and normalization layers. The encoder consists 2 of such Transformer layers, followed by a Reshape and a Dense Layer of O units to map the data to the feature space with the size 1 × O. The decoder has a similar architecture to the encoder, but with an additional dense layer after the multi-head attention layer, to shape the output with 1 x I dimension for the original input reconstruction, so that it can be processed by the final Time-Distributed Dense layer, where each time step is processed independently by a Dense layer with a single unit and a sigmoid activation function.

The encoder's output is augmented with Additive White Gaussian Noise (AWGN) and Rician fading to model the channel interference. X_e is the compressed output signal, scaled with linear signal-to-noise ratio $\text{SNR}_{\theta_l} = 10^{\left(\frac{\text{SNR}_{\text{dB}}}{10}\right)}$ where θ_l shows the linear scale. Accounting for the effects of Rician fading, the normalized Rician fading parameter b is defined by $b = \sqrt{\frac{K}{K+1}}$. K is the Rician factor and represents the degree of dominance of the Line-of-Sight (LoS) component over the non-LoS components in the channel. The faded signal can be represented as:

$$y_{\text{real}} = h_{\text{real}} \cdot X_e \tag{8.14}$$

$$y_{\text{imag}} = h_{\text{imag}} \cdot X_e \tag{8.15}$$

where y_{real} and y_{imag} are the real and imaginary parts of the received signal, respectively, and h_{real} and h_{imag} represent the real and imaginary parts of the channel's fading coefficient. The presence of AWGN noise is also considered by calculating its noise power spectral density σ_n^2, given by:

$$\sigma_n^2 = \frac{E_{\text{avg}}}{\text{SNR}_{\theta_l}} \tag{8.16}$$

where E_{avg} is the average energy and $E_{\text{avg}} = \max\left(\frac{1}{N}\sum_{i=1}^{N}|X_e|^2, 10^{-8}\right)$. Incorporating the AWGN, the signal equation with interference will be expressed as:

$$z_{\text{real}} = y_{\text{real}} + N_{o,\text{real}} \tag{8.17}$$

$$z_{\text{imag}} = y_{\text{imag}} + N_{o,\text{imag}} \tag{8.18}$$

where $N_{o,\text{real}}$ is the real part of noise, and $N_{o,\text{real}} = \sqrt{\frac{\sigma_n^2}{2}} \cdot \mathcal{N}(0,1)$, and $N_{o,\text{imag}}$ is the imaginary part of the noise, and $N_{o,\text{imag}} = \sqrt{\frac{\sigma_n^2}{2}} \cdot \mathcal{N}(0,1)$. The noise is modeled as a Gaussian distribution with zero mean and varying variance. The final envelope $|Z|$ of the received signal with the synergy of both AWGN noise and Rician fading will be expressed as:

$$|z| = \sqrt{z_{\text{real}}^2 + z_{\text{imag}}^2} \tag{8.19}$$

8.2.5 Training Procedure

The training procedure was accomplished with end-to-end training for all weights and biases by computing the loss. Each model was trained for 100 epochs, with the rest of the training parameters shown in table 8.1. The training process can be characterized by the following equation:

$$\Omega = h(q \cdot X_e + N_o),$$

where Ω is the output of the autoencoder, q is the constant channel coefficient, and N_o is the Additive White Gaussian Noise (AWGN). The loss function utilized during training is the Mean Squared Error (MSE), which is defined as:

$$L = \frac{1}{M} \sum_{m=1}^{M} \|\hat{X}_m - X_m\|_2^2, \tag{8.20}$$

where $\hat{X}_m$ and X_m represent the ground truth and predicted output, respectively. M is the total number of training examples. Moreover, the Adam algorithm is selected as the optimizer for updating the model's parameters, Adam optimizer combines the advantages of two popular methods, Momentum (uses first moment — the mean of gradients) and RMSProp (uses second moment — the variance of gradients) and it is formulated as:

$$\theta_t = \theta_{t-1} - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \tag{8.21}$$

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \tag{8.22}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \tag{8.23}$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \tag{8.24}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \tag{8.25}$$

here, m_t is the first-moment estimate (also known as the moving average of the gradients), and v_t is the second-moment estimate (also known as the moving average of the squared gradients). θ_t represents the model parameters at time step t, α is the learning rate, g_t is the gradient at time t, and ϵ is a small constant added for numerical stability. β_1 and β_2 are exponential decay rates for the moment estimates. Additionally, an exponential decay of 0.99 was used as the learning rate schedule for a gradual decrease in the learning rate during

Table 8.1: Training Parameters for each model

Parameter	**Value**
Learning Rate	0.0001
Decay rate	0.99
Batch Size	500
Number of Epochs	100
Optimizer	Adam
Loss Function	Mean Squared Error

training. The exponential decay is given by:

$$\alpha_t = \alpha_0 \cdot \text{decay\_rate}^{\left\lfloor \frac{t}{\text{decay\_steps}} \right\rfloor} \tag{8.26}$$

where α_t and α_0 are the current learning rate and the initial learning rate, respectively.

Algorithm 2: Autoencoder for Bit Compression

[1] **Input:**
Dataset for $UE_{c,n}$ and UE_f.
Step 1:
Concatenate the dataset for $UE_{c,n}$ and UE_f.
Step 2:
Bits Compression
for *epoch* $= 1$ ***to*** N **do**
Pass the input tensor X to the encoder model with parameters $\theta_c, \theta_i, \theta_n$.

Pass X through 4 convolutional layers to get the compressed output X_e with 4 bits.
Step 3: Interference Modeling with Fading and Noise.
Pass the output X_e through Rician noise with $I = 3$ for dominant LoS to obtain noisy output X_r.
Step 4: Bits Reconstruction. **for** *epoch* $= 1$ ***to*** 15 **do**
Pass X_r through the decoder network with parameters θ_c, θ_i.
Pass X_r through 3 convolutional layers to decode and get the reconstructed output X_d.

Step 5: Calculate NMSE. **for** *epoch* $= 1$ ***to*** 15 **do**
Calculate Normalized Mean Square Error (NMSE) against Signal to Noise Ratio (SNR).
Store the calculated NMSE values.
Output: Reconstructed output X_d and NMSE values.

8.3 Performance Analysis

8.3.1 Rate Analysis

We derive the expressions for the achievable rates of the UEs based on the system model. The rate for each user is a function of the power allocation coefficients and the channel conditions, including the effects of the RIS.

For the analysis of the achieved rates for each user in the system model as shown in Fig. 8.2, we first analyze the signal model of the wireless environment. For each $c \in \mathcal{C}$ and $f \in \mathcal{F}$, consider the trio UE$_{1,n}$, UE$_{2,n}$, UE$_f$ as a coordinated NOMA cluster. Here, P_c represents the transmission powers of both BS$_1$ and BS$_2$. This setup is essential for assessing the rates achieved and for optimizing system performance within the defined cluster. Each BS$_c$ transmits a superimposed signal containing messages for user equipment in its service area, specifically UE$_{c,n}$ and UE$_f$. This transmission is formulated as [SKB+13]:

$$y_c = \sqrt{\gamma_{c,n} P_c} y_{c,n} + \sqrt{\gamma_{c,f} P_c} y_f, \tag{8.27}$$

where $\gamma_{c,n}$ and $\gamma_{c,f}$ denote the power allocation (PA) coefficients that BS$_c$ allocates to UE$_{c,n}$ and UE$_f$, respectively. Notably, UE$_{c,n}$ enjoys more favorable channel conditions than UE$_f$, establishing it as the primary NOMA user within the pair (UE$_{c,n}$, UE$_f$) managed by BS$_c$. By NOMA protocols, UE$_{c,n}$ is expected to successfully detect and decode the message directed to UE$_f$. This operational framework necessitates that $\gamma_{c,n} < 0.5$ and $0.5 < \gamma_{c,f} < 1$ [ODS+20] [SM20]. The rate achieved by UE$_{1,n}$ is analyzed, which is similar in terms with UE$_{2,n}$.

For simplicity and conciseness, the rate achieved by UE$_{1,n}$ is defined only from the set of call center users $\mathcal{N}$, with the understanding that similar steps can be applied to define the rate of UE$_{c,n}$, $\forall c \in \mathcal{C}$ and $n \in \mathcal{N}$. The received signal at UE$_{c,n}$ can be written as:

$$z_{1,n} = \mathbf{H}_{1,n} y_1 + \mathbf{H}_{2,n'} y_2 + N_o, \tag{8.28}$$

where $\mathbf{H}_{1,n}$ represents the combined channel from BS$_1$ to UE$_{1,n}$ and can be expressed for both users as $\mathbf{H}_{c,n} = \mathbf{h}_{c,n} + \mathbf{h}_{R,n}^{H} \boldsymbol{\Theta}_{\mathbf{r}} \mathbf{h}_{c,R}$, represents the combined channel from BS$_c$ to UE$_{c,n}$. Θ represents the RIS phase shift matrix. N_o represents additive white Gaussian noise (AWGN), specifically $N_o \sim \mathcal{CN}(0, \sigma^2)$. By employing interference cancellation (SIC) techniques, UE$_{1,n}$ initially decodes the message signal from UE$_f$ (i.e y_f), then subtracts it from $z_{1,n}$ to decode its message (i.e $y_{1,n}$). This method allows us to express the signal-to-interference-and-noise ratio (SINR) zeta and the achievable rate at UE$_{1,n}$ for decoding UE$_f$'s message as follows:

$$\zeta_{1,n\to f} = \frac{\gamma_{1,f} P_1 |\mathbf{H}_{1,n}{}^2|}{\gamma_{1,n} P_1 |\mathbf{H}_{1,n}{}^2| + P_2 |\mathbf{H}_{2,n'}{}^2| + \sigma^2}, \tag{8.29}$$

$$\mathcal{R}_{1,n\to f} = B \cdot \log_2\left(1 + \zeta_{1,n\to f}\right), \tag{8.30}$$

Moreover, the Signal-to-Interference-plus-Noise Ratio (SINR) and the achievable data rate of UE$_{1,n}$ for decoding its message can be formulated as follows:

$$\zeta_{1,n} = \gamma_{1,n} \frac{P_1 |\mathbf{H}_{1,n}{}^2|}{P_2 |\mathbf{H}_{2,n'}{}^2| + \sigma^2}, \tag{8.31}$$

$$\mathcal{R}_{1,n} = B \cdot \log_2\left(1 + \zeta_{1,n}\right). \tag{8.32}$$

As a part of two NOMA pairs, UE$_f$ receives its signal through the transmission from each BS$_c$, $\forall c \in \mathcal{C}$. Thus, the received signal at UE$_f$ can be expressed as:

$$z_f = \mathbf{H}_{1,f} y_1 + \mathbf{H}_{2,f} y_2 + N_0, \tag{8.33}$$

where $\mathbf{H}_{1,f}$ and $\mathbf{H}_{2,f}$ represent the combined channels from BS$_1$ to UE$_f$ and from BS$_2$ to UE$_f$, and can be expressed as $\mathbf{H}_{1,f} = \mathbf{h}_{1,f} + \mathbf{h}_{R,f}^H \mathbf{\Theta_t} \mathbf{h}_{1,R}$ and $\mathbf{H}_{2,f} = \mathbf{h}_{2,f} + \mathbf{h}_{R,f}^H \mathbf{\Theta_t} \mathbf{h}_{2,R}$, respectively. Considering the scenario of non-coherent JT-CoMP, the SINR and achievable rate at UE$_f$ can be formulated as [TSAJ14, EAAG22]:

$$\zeta_f = \frac{\gamma_{1,f} P_1 |\mathbf{H}_{1,f}{}^2| + \gamma_{2,f} P_2 |\mathbf{H}_{2,f}{}^2|}{\gamma_{1,n} P_1 |\mathbf{H}_{1,f}{}^2| + \gamma_{2,n} P_2 |\mathbf{H}_{2,f}{}^2| + \sigma^2}, \tag{8.34}$$

$$\mathcal{R}_f = B \cdot \log_2\left(1 + \zeta_f\right). \tag{8.35}$$

8.3.2 Outage Probability Analysis

For further assessing the performance by tactically deploying the Aerial RIS, we analyze the outage probability encountered by cellular users to improve the performance of the system. According to NOMA principles, for each $c \in \mathcal{C}$, $n \in \mathcal{N}$, and $f \in \mathcal{F}$, if UE$_{c,n}$ fails to decode x_f, or can decode x_f but not $x_{i,n}$, an

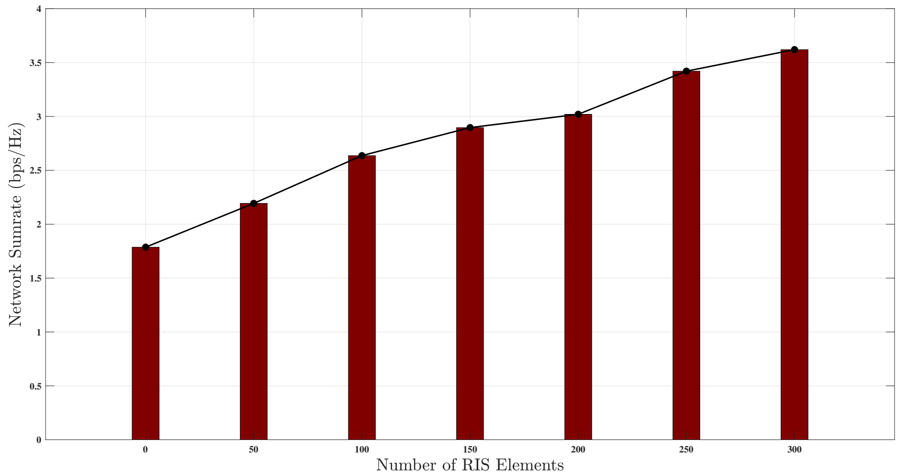

Figure 8.4: Network sum rate vs. number of RIS elements M

outage event occurs. This outage probability is expressed as as [YXL+22]:

$$\mathcal{P}_{c,n} = \text{Pr}\left(\zeta_{c,n\to f} < \zeta_{th_f}\right) + \text{Pr}\left(\zeta_{c,n\to f} > \zeta_{th_f}, \zeta_n < \zeta_{th_n}\right), \quad (8.36)$$

where ζ_{th_f} and ζ_{th_n} are the outage thresholds for UE_f and $\text{UE}_{c,n}$, respectively. Similarly, with regard to UE_f, an outage occurs when it fails to decode x_f, and the corresponding outage probability is expressed as:

$$\mathcal{P}_f = \text{Pr}\left(\gamma_f < \gamma_{th_f}\right). \quad (8.37)$$

Table 8.2: Simulation Parameters

Parameters	**Values**
Path-loss exponent of BS_c-(UE_n, RIS) links	$\alpha_{c\to n} = 3.2$
Path-loss exponent of BS_c-UE_f link	$\alpha_{c\to f} = 4.5$
Path-loss exponent BS_c-RIS links	$\alpha_{c\to R} = 2.7$
Path-loss exponent of RIS-UE_n links	$\alpha_{R\to c} = 3.0$
Path-loss exponent of RIS-UE_f link	$\alpha_{R\to f} = 2.7$
Path-loss exponent of Interfering links	$\alpha_{c\to n'} = 4.2$
Rician factor of RIS-UE_n links	$\kappa_{R\to n} = 3$ dB
Rician factor of RIS-UE_f link	$\kappa_{R\to f} = 4$ dB

8.4 Results and Discussion

This section presents our findings on compressed communication in Aerial-RIS networks, enhanced by machine learning. We start by outlining the simulation setup and parameters. We then examine the impact of Aerial RIS elements on urban communication performance and discuss power allocation strategies for improved efficiency. The effectiveness of our methods in maintaining reliable communication is evaluated, followed by a discussion on the trade-offs between spectral and energy efficiency. Finally, we highlight the machine learning results that optimize network performance. These analyses provide insights into the potential and challenges of implementing compressed communication in Aerial RIS networks for 6G and future urban wireless networks.

8.4.1 Simulation Setup

In this study, we simulate an outdoor urban environment using Comyx, a Python library specifically designed for modeling and simulating wireless communication systems, where the network operates with a carrier frequency of $f_c = 2.4$ GHz. The power of the additive white Gaussian noise (AWGN) is calculated as $\sigma^2 = -174 + 10\log_{10}(B)$ dBm, factoring in a noise figure (NF) of 12 dB. For simplicity, we assume that both base stations (BS_1 and BS_2) transmit with identical power, denoted as P_t, which ranges from -45 dBm to 0 dBm.

The power allocation (PA) factors for the far user (UE_f) and the near user ($\text{UE}_{c,n}$) are set to 0.8 and 0.2, respectively. The distances for the far and near users from the base stations are 300 meters and 150 meters, respectively, from their base stations. This parameterization intentionally emulates a cell-edge, interference-limited operating point to stress-test outage behavior and highlight the relative gains from RIS-assisted CoMP-NOMA. The key simulation parameters are summarized in Table 8.2.

8.4.2 Impact of Aerial RIS Elements

The overall network sum rate for users in the assemblage is depicted in Fig. 8.4. With the increase in the percentage of Aerial RIS elements, the network sum rate also increases, as indicated by the upward trend in the graph. The improvement in network performance is due to enhanced data coverage by the RIS. The increase in the RIS elements ensures better signal reflections and transmissions, thereby improving the overall network efficiency and user rates.

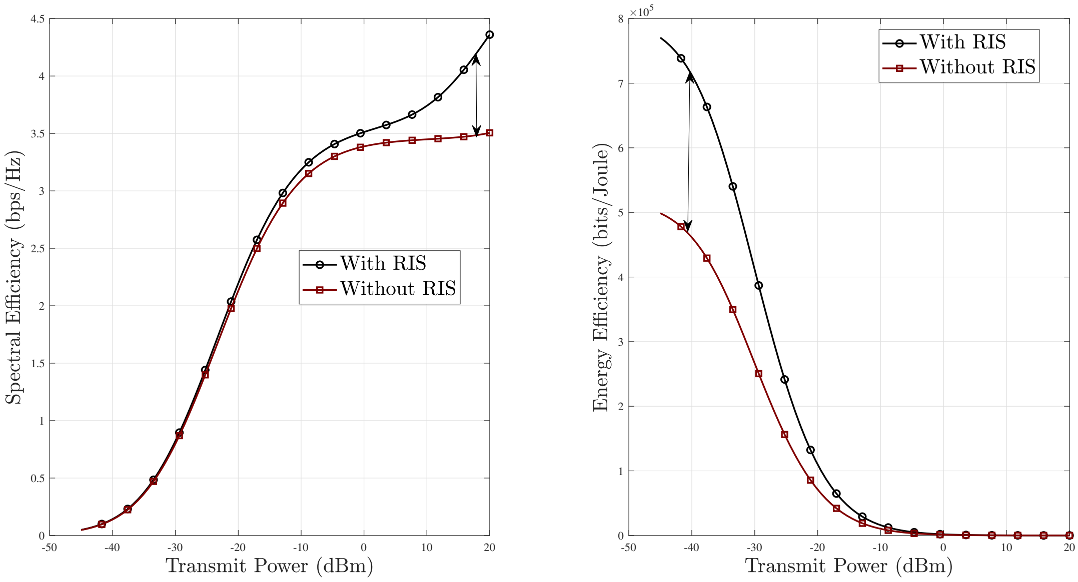

Figure 8.5: Spectral efficiency and energy efficiency

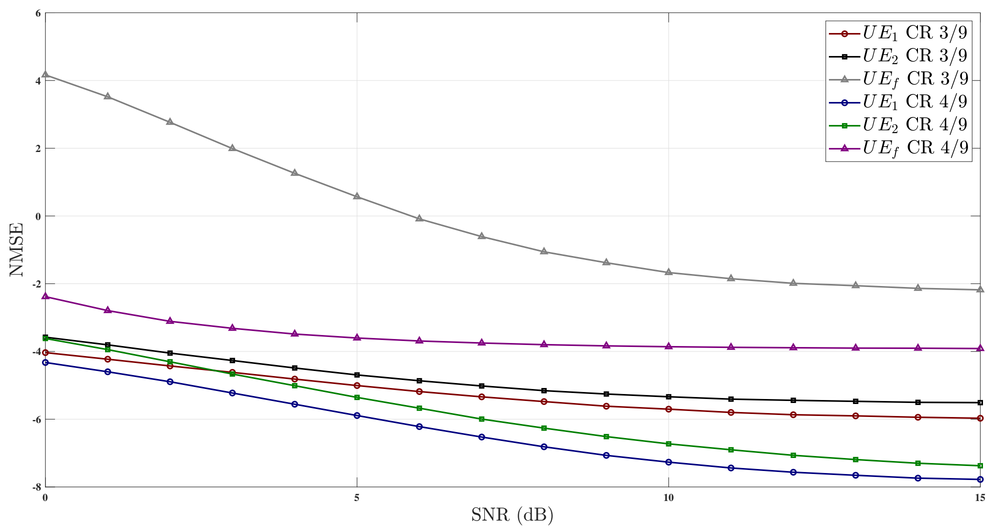

Figure 8.6: Normalized Mean Square Error (NMSE) of CNN + Attention-based model for $UE_{c,1}$, $UE_{c,2}$, and UE_f at CR 3/9 and 4/9.

8.4.3 Power Allocation for Coordinated NOMA

Power allocation for users $\text{UE}_{c,n}$ and UE_f is shown in Fig. 8.7. Within NOMA pairs (i.e., $\text{UE}_{1,n}, \text{UE}_f$) and ($\text{UE}_{2,n}, \text{UE}_f$) as shown in Fig. 8.1, PA factors are utilized to determine the distribution of transmitted power among users in a paired transmission.

The power allocation factor optimization problem in a NOMA pair can be for-

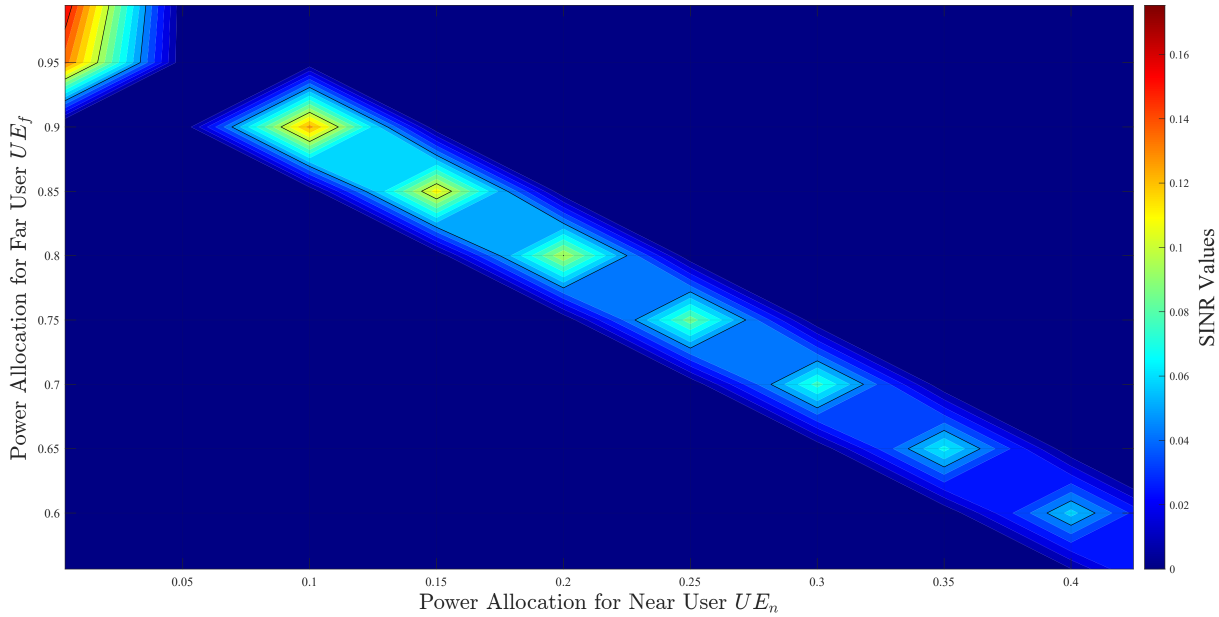

Figure 8.7: Power allocation for UE_n vs. far user UE_f

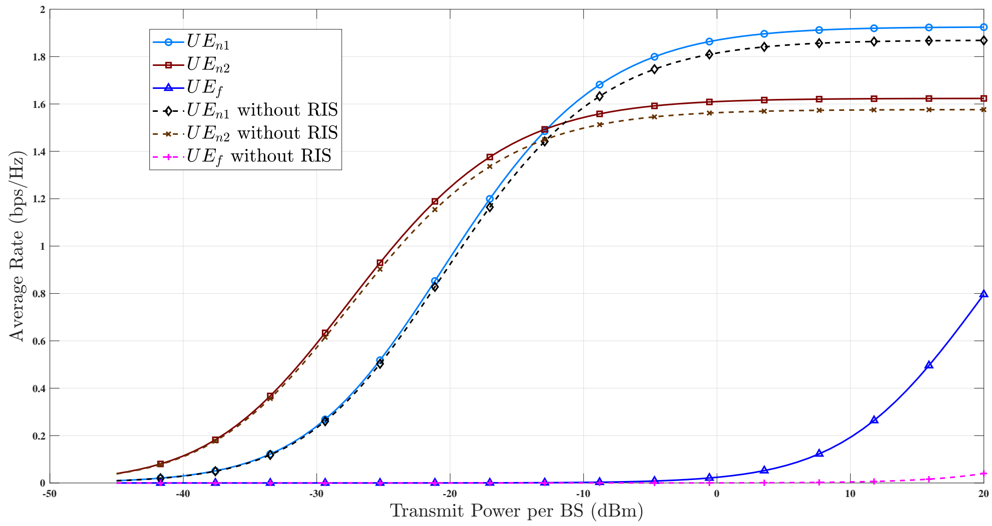

Figure 8.8: Average rate of UE_n,UE_f vs. transmit power per BS_c

mulated as follows:

$$\begin{aligned} \max_{\mathcal{A}} \quad & \mathcal{R}_{c,n} + \mathcal{R}_f, \\ \text{s.t.} \quad & \gamma_{c,n} + \gamma_f \leq 1, \forall c \in \mathcal{C}, n \in \mathcal{N}, f \in \mathcal{F}, \\ & 0.5 < \gamma_{c,f} < 1, \forall f \in \mathcal{F}. \end{aligned} \tag{8.38}$$

To obtain optimal PA factors for the two NOMA pairs, the methodology proposed in [FZCL16] was followed. Additionally, we evaluated the impact of PA

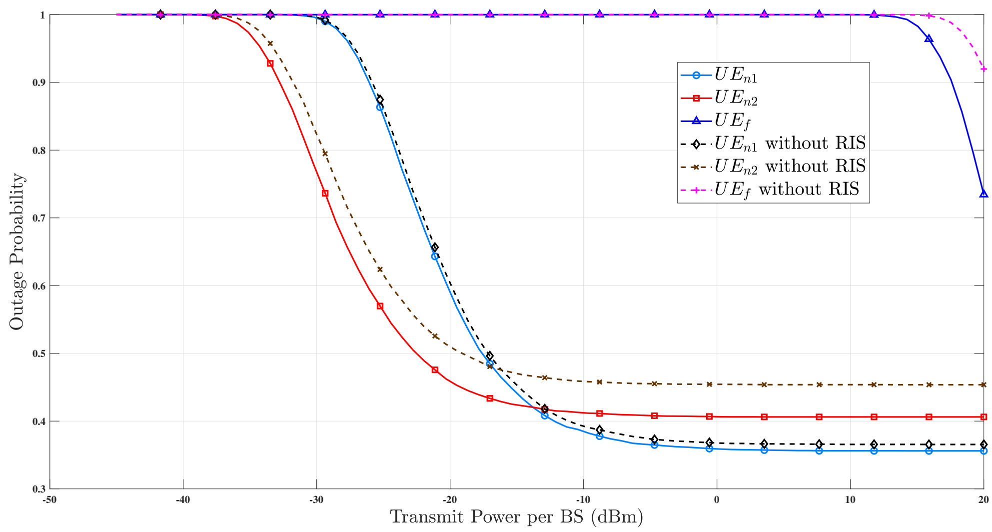

Figure 8.9: Outage probability of $UE_{c,n}$, UE_f, vs. transmit power per BS_c

factors alongside Aerial RIS enhancements, focused on a configuration with $K = 70$ elements. This configuration notably yields the highest network sum rate among all considered cases.

8.4.4 Implications of Feedback Compression on Link Performance (Analytical Discussion)

The autoencoder results in this chapter quantify reconstruction fidelity of the quantized RIS phase shifts via NMSE and training loss. The examiner's key question is how this reconstruction error translates into end-to-end link performance (SE, rate, outage). While additional Monte-Carlo simulations could quantify the exact performance gap between an ideal (no-compression) feedback baseline and the proposed compressed feedback, the impact can be interpreted analytically through the loss of coherent RIS combining.

Table 8.3: NMSE values (dB) of Autoencoder architectures for different Compression Ratios (CR)

Models \ CR	$\frac{3}{9}$			$\frac{4}{9}$		
	$UE_{c,1}$	$UE_{c,2}$	UE_f	$UE_{c,1}$	$UE_{c,2}$	UE_f
CNN	-5.393	-5.076	-1.868	-7.755	-7.164	-3.551
CNN + Attention	-5.946	-5.501	-2.180	-7.761	-7.167	-3.606
RNN	-5.596	-5.042	-3.059	-6.880	-6.313	-3.110
Transformer	-2.190	-2.677	4.134	-2.193	-2.678	4.150

Let the intended RIS phase at element m be θ_m, and the reconstructed phase after compression and decoding be $\hat{\theta}_m = \theta_m + \Delta\theta_m$, where $\Delta\theta_m$ is the phase error induced by imperfect reconstruction. For a passive RIS, the reflected-field contribution combines across elements as

$$G_{\mathrm{RIS}} \propto \left| \sum_{m=1}^{M} e^{j\hat{\theta}_m} \right|^2 = \left| \sum_{m=1}^{M} e^{j(\theta_m + \Delta\theta_m)} \right|^2. \tag{8.39}$$

Under ideal feedback ($\Delta\theta_m = 0$), the RIS phases can be aligned to yield near-coherent combining, producing an $\mathcal{O}(M^2)$ power gain. With phase errors, the coherent sum is reduced by the dispersion of $\{\Delta\theta_m\}$, and the effective RIS gain decreases. In the small-error regime, the coherent loss is governed by the concentration of $\Delta\theta_m$ around zero (e.g., via $\mathbb{E}[\cos(\Delta\theta_m)]$), implying that lower NMSE (better phase reconstruction, as achieved by CNN+Attention) should translate into a smaller loss in effective channel gain.

Because the achievable rate is a monotone function of SINR, any reduction in the coherent RIS gain decreases SINR and thus reduces spectral efficiency $\log_2(1 + \mathrm{SINR})$, while simultaneously increasing outage probability when SINR falls below the decoding thresholds. Therefore, the NMSE trends reported in Fig. 8.11–Fig. 8.12 provide a principled proxy for the expected ordering of link-level performance: architectures with lower reconstruction error are expected to preserve more of the RIS-assisted combining gain and yield better SE/outage behavior, whereas architectures with higher reconstruction error (e.g., the Transformer in the present setup) are expected to incur a larger performance penalty.

8.4.5 Average Rate and Outage Probability

The analytical implications of feedback compression discussed above provide the context for the rate and outage results presented next. The average rate and outage probability for each user are depicted in Fig. 8.8 and Fig. 8.9, respectively. The average rate of the communication system, as shown in Fig. 8.8, increases with transmit power per BS, since higher transmit power strengthens the received signal and improves resilience against noise and interference.

Fig. 8.9 shows the outage probabilities of $UE_{1,n}$, $UE_{2,n}$, and UE_f. It is important to note that the outage levels remain relatively high (e.g., above ≈ 0.35 for the far user in some configurations) because the evaluated setup represents a deliberately harsh, interference-limited cell-edge scenario with strong inter-cell

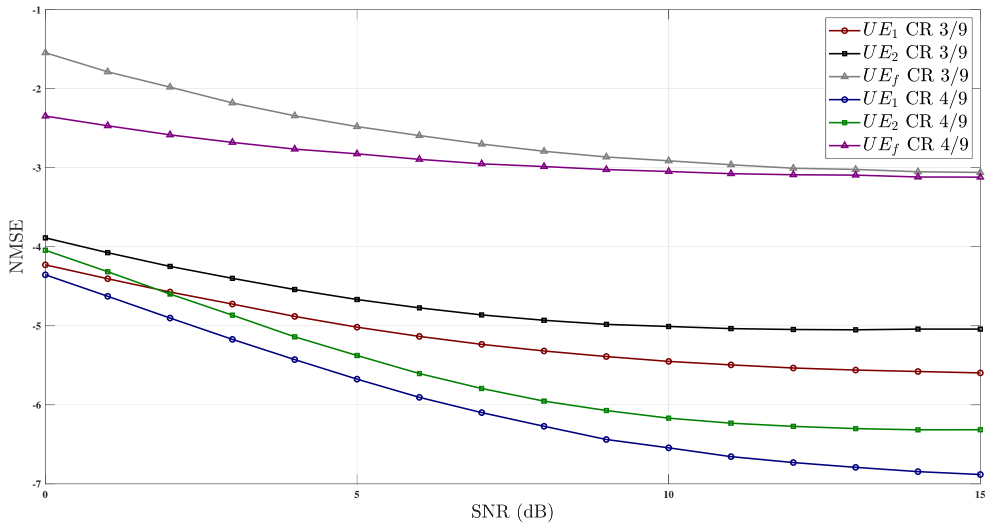

Figure 8.10: NMSE of RNN-based model for $UE_{c,1}$, $UE_{c,2}$, and UE_f at CR 3/9 and 4/9.

interference and weak direct links. Hence, the objective of Fig. 8.9 is to provide a comparative reliability assessment under severe conditions and to quantify the relative outage improvements enabled by CoMP and RIS assistance, rather than to demonstrate outage-free operation.

The near users exhibit limited outage improvement because their links are already dominated by their respective serving base stations (BS_1 and BS_2), making them less sensitive to RIS-assisted alternative propagation paths. In contrast, the far user UE_f experiences high outage in the non-CoMP case due to strong inter-cell interference. With RIS assistance, the outage probability is consistently reduced for all users at a given transmit power, confirming the reliability gains provided by RIS-enabled CoMP-NOMA under interference-limited conditions.

8.4.6 Balancing Act: Spectral Efficiency vs. Energy Efficiency

Fig. 8.5 illustrates the relationship between spectral efficiency (SE) and energy efficiency (EE) as a function of transmit power. Spectral efficiency is defined as

$$\text{SE} = \log_2(1 + \text{SINR}) \quad \text{(bits/s/Hz)}, \tag{8.40}$$

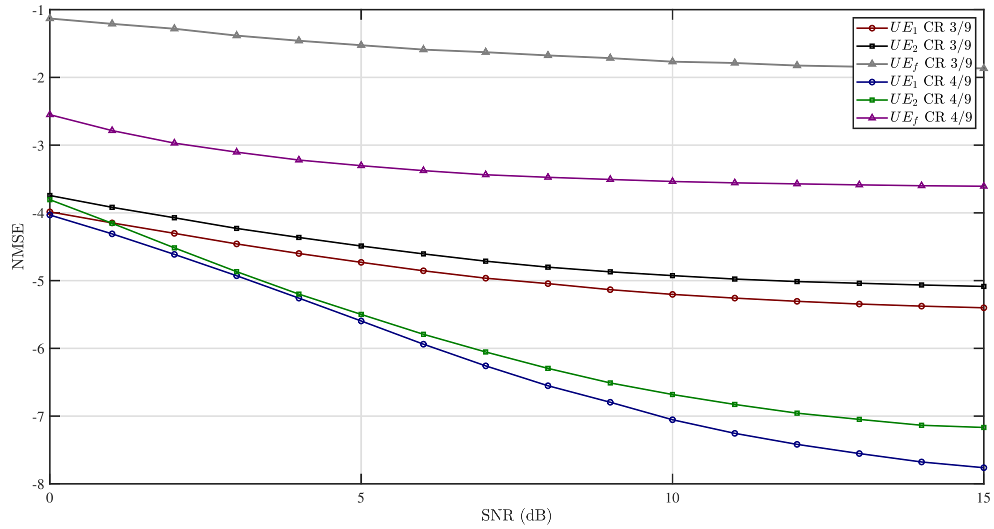

Figure 8.11: NMSE of CNN-based model for $UE_{c,1}$, $UE_{c,2}$, and UE_f at CR 3/9 and 4/9.

while energy efficiency is given by

$$\mathrm{EE} = \frac{B \log_2(1 + \mathrm{SINR})}{P_t} \quad \text{(bits/Joule)}, \tag{8.41}$$

where B denotes the system bandwidth and P_t is the transmit power.

As observed in Fig. 8.5, the spectral efficiency increases with transmit power but gradually saturates in the high-SNR regime. This saturation occurs because the system becomes interference-limited, and further increases in transmit power do not translate into proportional SINR gains. In contrast, energy efficiency continues to vary with transmit power since P_t appears explicitly in the denominator of the EE expression. Consequently, even when SE saturates, EE decreases at higher transmit power levels due to increased energy consumption.

The inclusion of the aerial RIS improves both SE and EE, particularly at low-to-moderate transmit power levels. RIS-assisted propagation enhances the effective channel gain and interference conditions, enabling higher spectral efficiency without a proportional increase in transmit power. As a result, RIS shifts the operating point of the SE–EE trade-off toward a more energy-efficient regime rather than eliminating the trade-off entirely.

Overall, these results demonstrate that aerial RIS-assisted CoMP-NOMA improves network performance by enhancing spectral efficiency, reducing outage probability, and enabling more energy-efficient operation, especially in

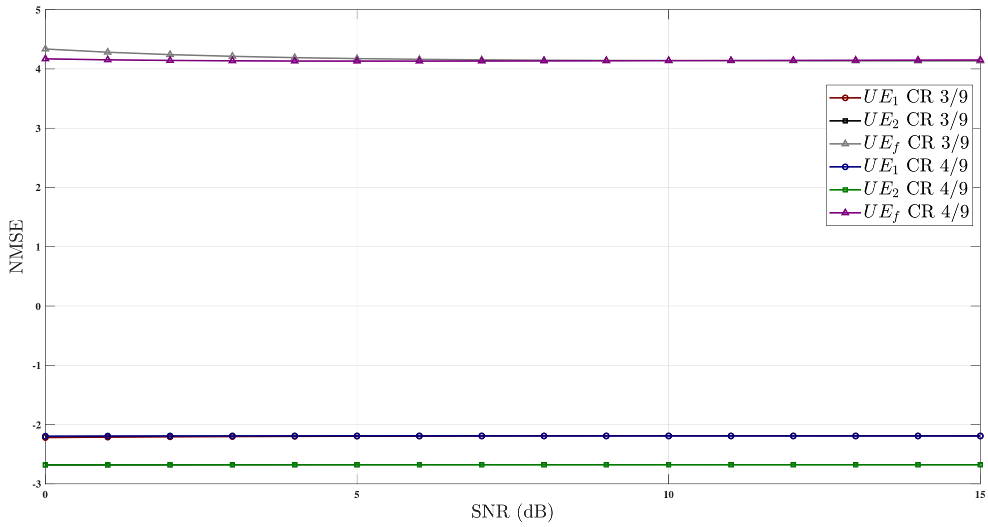

Figure 8.12: NMSE of Transformer-based model for $UE_{c,1}$, $UE_{c,2}$, and UE_f at CR 3/9 and 4/9.

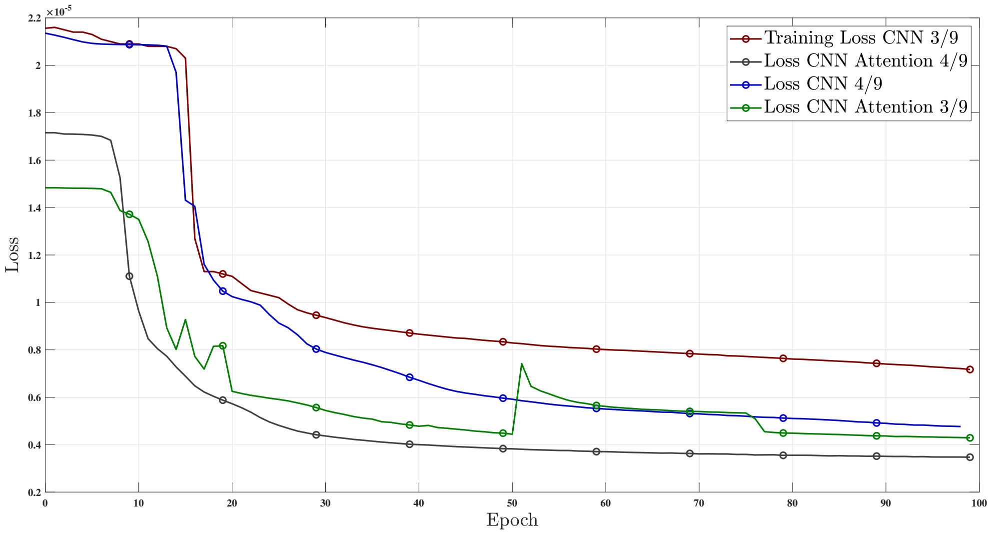

Figure 8.13: Training loss of CNN and CNN + attention models at CR 3/9 and 4/9

interference-limited scenarios.

8.4.7 Machine Learning Results

The results for the autoencoder were evaluated based on Normalized Mean Square Error (NMSE) as shown in 8.3, given as:

Table 8.4: Loss Values of Autoencoder Architectures for Different Compression Ratios (CR)

CR \ Models	CNN	CNN + Attention	RNN	Transformer
$\frac{3}{9}$	0.00075 %	0.00040 %	0.00070 %	0.00143 %
$\frac{4}{9}$	0.00046 %	0.00037 %	0.00050 %	0.00141 %

$$\text{NMSE} = \frac{\|\hat{\Theta}_m - \Theta_m\|_2^2}{\|\hat{\Theta}_m\|_2^2} \tag{8.42}$$

where Θ_m is the true value, $\hat{\Theta}_m$ is the estimated value, $\|\cdot\|_2$ denotes the Euclidean norm, and the Training Loss of the models employed for our feedback channel. The NMSE showcases the decrease in the model's loss when the SNR associated with each bit increases. For each user, UE_f and $UE_{c,n}$, Compression Ratio (CR) i.e. the ratio of compressed to the original bits $\frac{O}{I}$, of $\frac{3}{9}$ and $\frac{4}{9}$ were used to compute the NMSE with a scale of Signal to Noise Ratio ranging from 0dB to 15dB, whereas the training loss was computed for 100 epochs as mentioned above.

Fig. 8.11 shows the performance of CNNs in terms of NMSE for all users accounted in the system model, with both CR. As for the trends of the graph, the NMSE gradually decreases with the increase in SNR. The NMSE for UE_f is significantly greater than that of $UE_{c,n}$ in both scenarios, which is expected due to the larger signal coverage distance for far users compared to near users. The $\frac{4}{9}$ CR has a better NMSE since the compressed output was associated with more original bits than that of CR $\frac{3}{9}$. Overall, the CNNs showcase fine results owing to their high capability of learning important features from data.

The results of NMSE of CNN+Attention architecture are displayed in Fig. 8.6. With the addition of AM, CNN's performance is enhanced adequately. The final values of NMSE decayed to a significant value relative to the CNNs, making them more reliable to deploy. The trends for curves are similar to that of CNNs, as NMSE decreases with the increase in SNR until it stabilizes at the end.

Fig. 8.10 and Fig. 8.12 show the performance of RNNs and Transformers in terms of NMSE respectively. RNNs somewhat showcase decent results, having similar trends as that of CNN and CNN+Attention. On the other hand, transformers underperform in this scenario, displaying negligible change in NMSE values over the SNR range. While Transformers can understand complex connections, they aren't as naturally suited to recognize spatial patterns as effi-

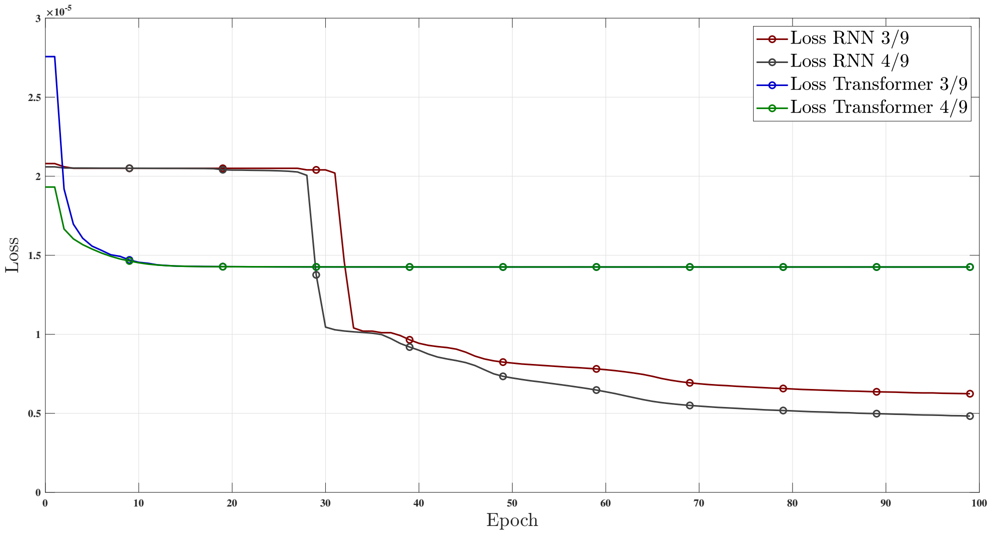

Figure 8.14: Training loss of RNN and Transformer models at CR 3/9 and 4/9

ciently as CNNs.

The training loss for CNN and CNN+Attention, for both 3/9 and 4/9 CR is shown in Fig. 8.13. The training loss decreases and eventually stabilizes at the end. The loss for CNN+Attention is less than CNNs for both CR as shown in table 8.4, indicating better reconstruction of QPS bits at the receiver end. Similarly, Fig. 8.14 displays the loss for RNNs and Transformers, respectively.

Comparing the loss results from the table 8.4 and Fig. 8.15, it can be seen that the CNN+attention architecture outperforms all the other mechanisms deployed due to its ability to capture significant features to reconstruct the original input. Whereas, the transformer has the worst results out of all, making it an unsuitable compression method for this case. The CNNs and RNNs still showcase better performance than Transformers, but they fall slightly short of the performance achieved by CNNs combined with AM.

8.5 Conclusion

This chapter has explored the integration of machine learning-driven optimization in Aerial Reconfigurable Intelligent Surfaces (RIS)-assisted Coordinated Multipoint (CoMP)-NOMA networks for enhanced 6G wireless performance. The proposed methodology effectively addresses key challenges such as feedback overhead, spectral efficiency, and outage probability, making it a viable solution for next-generation networks. Our simulation results quantify the per-

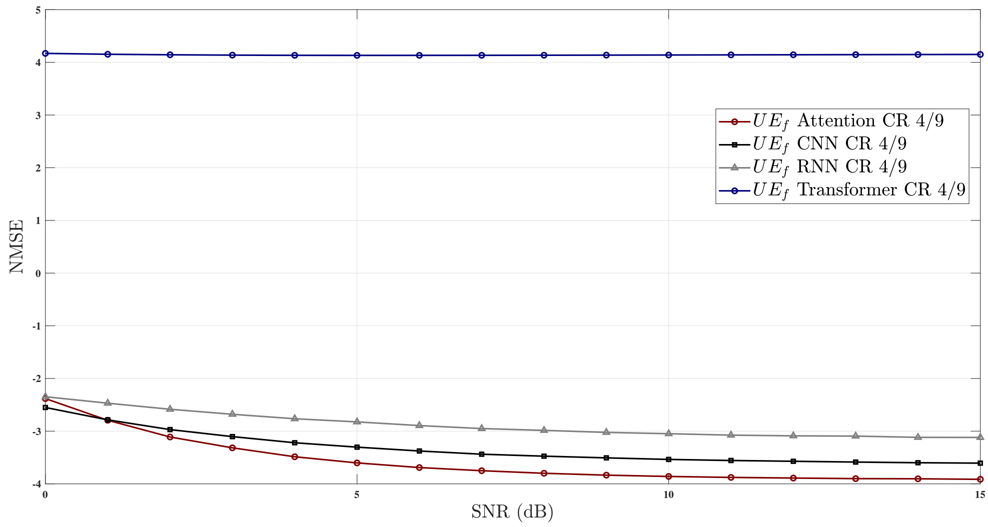

Figure 8.15: Comparison of NMSE across all employed architectures for UE_f at a 4/9 compression ratio.

formance gains from integrating RIS into 6G networks:

Table 8.5: Performance Comparison With and Without RIS

Metric	**With RIS**	**Without RIS**	**Gain**
Spectral Efficiency @ 20 dBm	~ 4.3 bps/Hz	~ 3.5 bps/Hz	$\sim 23\%$
Energy Efficiency @ –40 dBm	$\sim 8 \times 10^9$ bits/J	$\sim 5 \times 10^9$ bits/J	$\sim 60\%$

Our simulation results show that integrating RIS provides:

- Network sum rate increase when deploying RIS elements compared to the baseline without RIS.
- 15% Gain for Near User and 111.1% Gain for far user at 20 dBm.
- Outage probability is lower for UEs with an RIS.

Key findings from this study include:

- The integration of UAV-assisted RIS in CoMP-NOMA networks significantly enhances data rates and spectral efficiency by dynamically adjusting phase shifts, improving non-line-of-sight (NLOS) communication.
- A machine learning-based autoencoder approach for Quantized Phase Shifts (QPS) compression effectively reduces the feedback overhead while maintaining high accuracy, leading to efficient bandwidth utilization.
- Performance analysis demonstrated substantial gains in outage proba-

bility reduction, making the proposed framework a robust solution for interference-limited 6G environments.

- The results confirm that Aerial RIS-assisted CoMP-NOMA can optimize network reliability, particularly in dense urban environments where traditional cell-based coverage solutions face limitations.

The insights gained in this chapter directly contribute to answering the overarching research question of this thesis: "*How can integrating IRS, CBS, D2D communication, mmWave, massive MIMO, and DSS be optimized to enhance wireless connectivity and reduce coverage gaps in 5G heterogeneous networks for smart factory applications?*" The demonstrated benefits of machine learning-driven optimization in Aerial RIS-assisted networks align with this objective by ensuring adaptive, energy-efficient, and high-capacity communication, which is critical for smart industrial applications. Future work will quantify the end-to-end SE and outage degradation relative to an ideal no-compression baseline for different RIS feedback compression ratios using Monte-Carlo simulations. The next chapter provides a synthesis of the key contributions presented throughout this thesis, highlighting their impact on advancing 5G and 6G network architectures and outlining future research directions.

Chapter 9

Conclusions and Future Work

> *"No matter how correct a mathematical theorem may appear to be, one ought never to be satisfied that there was not something imperfect about it until it also gives the impression of being beautiful."*
>
> *George Boole FRS (1815-1864)*
>
> *First Professor of Mathematics at Queen's College, Cork, 1849-1864*

9.1 Contributions and Key Findings

This thesis has explored novel methodologies and presented a comprehensive investigation into enhancing connectivity, efficiency, and reliability in 5G and emerging 6G wireless networks. The research addressed key challenges such as optimal resource allocation, latency reduction, spectral efficiency, and network resilience, contributing to the advancement of next-generation wireless communication systems. Collectively, these contributions provide a unified framework for improving the performance and scalability of 5G/6G wireless networks. The findings of this thesis have direct implications for smart factories, industrial IoT, and urban wireless communication, enabling more efficient and adaptive network architectures.

9.1.1 Base Station Selection and Load Balancing (*Pub 1*)

This work evaluated base station selection strategies based on Maximum Received Power (MRP) and Maximum Signal-to-Interference-plus-Noise Ratio (Max SINR) in a 5G heterogeneous network representative of smart factory environments. The analysis reveals clear trade-offs between signal quality, throughput performance, and load distribution in dense, interference-limited deployments.

Results show that MRP association provides stronger average received power and more stable SINR distributions for a large fraction of users, owing to its tendency to connect devices to geographically closer base stations. However, this strategy can lead to uneven load distribution and reduced spectral efficiency in dense scenarios. In contrast, Max SINR association, while not always yielding higher instantaneous SINR under heavy interference coupling, achieves improved user throughput and better fairness by accounting for co-channel interference during cell selection and steering users toward less congested cells.

Empirical cumulative distribution function (ECDF) analysis demonstrates that Max SINR-based association increases median and high-percentile user throughput compared to MRP, despite the interference-limited nature of the deployment. These findings highlight that interference-aware association mechanisms are essential for balancing throughput efficiency and service fairness in next-generation 5G heterogeneous networks, particularly in dense industrial environments.

9.1.2 Massive MIMO Precoding Techniques (*Pub 2*)

The massive MIMO evaluation in Chapter 3, based on the Vienna 5G Link-Level Simulator, shows that Zero-Forcing (ZF) consistently achieves higher throughput than Maximum Ratio Transmission (MRT) across all tested antenna configurations. Under the adopted link-level assumptions—fixed per-link path loss (100 dB) and intentionally neglected inter-cell interference via inflated attenuation between the two simulated base stations—ZF delivers higher *per-user* and *sum* throughput in both downlink and uplink as the number of base-station antennas increases. These results confirm that interference-aware precoding (ZF) provides a clear capacity advantage over MRT in the considered multi-user setting, while MRT improves with increasing antenna count but remains limited by residual multi-user interference within the simulated range. Accordingly, the chapter supports ZF as the more effective precoding strategy for maximizing throughput in MU-massive-MIMO operation under the stated simulation conditions.

9.1.3 Multi-User Transmission and Spectral Efficiency (*Pub 3*)

This work investigates a multi-user auxiliary-signal-assisted space–time block coded OFDM (MU-STBC-ST) transmission scheme under equal-distance and equal-SNR conditions. Simulation results show that, compared to conventional STBC-OFDM, the proposed scheme achieves an effective throughput gain of approximately $2\times$ by enabling simultaneous transmission to two users over the same time–frequency resources. The scheme preserves the diversity benefits of Alamouti STBC with single-antenna receivers, while also demonstrating a reduction in peak-to-average power ratio (PAPR) of about 2 dB at a CCDF of 10^{-3}. These results indicate that MU-STBC-ST improves throughput efficiency without increasing receiver complexity, making it suitable for link-level multi-user transmission scenarios.

9.1.4 Composite Base Stations and Dynamic Spectrum Scheduling (*Pub 4*)

This work evaluates a CBS architecture jointly supporting LTE (sub-6 GHz) and 5G NR (sub-6 GHz/28 GHz) under a traffic-intensity-based Dynamic Spectrum Scheduling setting in the Vienna 5G System Level Simulator for a realistic UCC campus deployment. Results highlight the expected trade-off: sub-6 GHz delivers more stable coverage and SINR, while 28 GHz provides substantially higher per-user throughput under favorable channels but with higher variability due to blockage and path loss. Overall, the CBS/DSS setup demonstrates how dynamic, demand-driven resource allocation can balance capacity (mmWave) and coverage (sub-6 GHz) in dense urban/campus scenarios.

9.1.5 Device-to-Device Communication for Smart Industry (*Pub 5*)

The proposed event-driven D2D relay-selection algorithm significantly improves *connectivity availability* in smart-industry blind-spot scenarios. At 250 IoT nodes, the availability increases from $\eta \approx 0.30$ (random selection) to $\eta \approx 0.72$, corresponding to an $\sim 140\%$ relative gain and an outage (blind-spot time) reduction of about 60%. These results highlight the effectiveness of lightweight, connectivity-aware D2D relaying for restoring coverage in industrial IoT deployments.

9.1.6 Intelligent Reflecting Surfaces for mmWave Communications (*Pub 6*)

In an IRS-assisted smart-factory mmWave scenario (28 GHz), the results show up to 1.8× spectral-efficiency gain (1.6× at sub-6 GHz) and an approximate 90% reduction in required transmit power, highlighting IRS as an energy-efficient enabler for reliable industrial connectivity.

9.1.7 AI-Driven Optimization in 6G Aerial-RIS Networks (*Pub 7*)

This work studies a two-cell Aerial-RIS-assisted JT-CoMP NOMA downlink and introduces machine-learning-based compressed feedback for RIS control. Quantized RIS phase shifts (9-bit) are compressed to $\frac{3}{9}$ and $\frac{4}{9}$ using autoencoders and reconstructed under AWGN and Rician fading, reducing feedback overhead while preserving phase accuracy (CNN+Attention achieves the best NMSE/loss among evaluated models). Simulations confirm that increasing RIS elements improves network sum rate and enhances far-user reliability under inter-cell interference. Overall, RIS yields $\sim 23\%$ spectral-efficiency gain at 20 dBm and $\sim 60\%$ energy-efficiency gain at -40 dBm, with average-rate improvements of $\sim 15\%$ for near users and $\sim 111\%$ for the far user, alongside reduced outage probability.

9.2 Future Work

This research establishes a robust foundation for advancing wireless communication technologies, addressing critical challenges in connectivity, efficiency, and scalability. Building upon the insights gained, several key directions for future work have been identified:

- **Adaptive and Intelligent Base Station Selection for Future 5G/6G Smart Factory Networks:** Building upon the promising results of study in chapter 2, several avenues for future research are worth exploring to enhance the performance and adaptability of base station selection in 5G HetNets for smart factory environments. A key direction involves integrating intelligent scheduling algorithms with the evaluated cell association strategies to optimize resource allocation and improve latency and throughput for heterogeneous industrial devices. Furthermore, incorporating mobility-aware models for dynamic elements such as autonomous mobile robots and guided vehicles will provide valuable insights into handover performance and stability. Another potential enhancement lies in the development of multi-objective optimization frameworks that jointly consider SINR, received power, network load, and energy efficiency for more balanced and context-aware association decisions. The application of machine learning techniques for real-time, adaptive base station selection also holds significant promise, particularly in complex and interference-prone industrial settings. Future work should additionally explore the implications of evolving 5G-Advanced and 6G use cases, including time-sensitive networking, industrial XR, and ultra-reliable low-latency communication (URLLC). Lastly, validating the proposed strategies through implementation on physical testbeds or digital twin-based emulation platforms will help bridge the gap between simulation results and real-world applicability.

- **Robust and Learning-Aided Precoding Strategies for Multi-User Massive MIMO in 5G/6G:** Future work should evaluate linear precoding under imperfect CSI, including estimation errors, quantization, and pilot contamination, to reflect real-world deployment scenarios. Expanding the current single-cell setup to multi-cell environments with inter-cell interference and techniques like CoMP and interference alignment is essential for scalability analysis. Further research should incorporate energy efficiency metrics and hardware impairments, and explore hybrid beamforming to reduce complexity, especially at mmWave and sub-THz frequencies. Machine learning approaches, such as reinforcement learning and neural network-based precoding, offer potential for adaptive beamforming under dynamic conditions. Finally, performance should be assessed under mobility and diverse channel environments to support use cases like V2X and industrial automation in 5G-Advanced and 6G networks.

- **Advanced Multi-User Space-Time Block Coding (MU-STBC):** The scalability and efficiency of MU-STBC must be optimized for high-density multi-user systems. This involves developing adaptive encoding and decoding schemes that minimize computational complexity while enhancing energy and spectral efficiency. Integrating MU-STBC with hybrid beamforming and advanced antenna configurations, such as massive MIMO, could further improve performance in dynamic and interference-prone environments.

- **Integration of Edge Computing with Composite Base Stations (CBS):** Edge computing offers real-time data processing capabilities crucial for latency-sensitive applications. Research should focus on integrating CBS with edge nodes to enable dynamic spectrum allocation, efficient load balancing, and on-demand service deployment. The synergy between CBS and edge computing must also consider energy efficiency, particularly in energy-constrained IoT deployments.

- **AI-Driven Device-to-Device (D2D) Communication Frameworks:** AI-based models can transform D2D communication by predicting traffic patterns, optimizing resource allocation, and mitigating interference. These frameworks must be scalable to support ultra-dense IoT networks and resilient against failures in high-stress environments. Multi-agent reinforcement learning (MARL) approaches could further enhance cooperative behavior among devices.

- **Intelligent Reflecting Surface (IRS) Optimization:** Dynamic IRS configurations must be studied to enhance their adaptability in complex environments. This includes designing algorithms for real-time phase shift optimization and resource allocation. Aerial-RIS deployments, especially in multi-cell networks, require models that account for mobility, atmospheric conditions, and inter-cell interference. Experimental validation of these configurations using large-scale testbeds will also be critical.

- **Learning-Driven Feedback Optimization in Aerial-RIS and Intelligent Wireless Systems:** The integration of machine learning with feedback mechanisms, particularly in Aerial-RIS and other intelligent wireless systems, holds immense potential. Future research can explore the optimization of RIS placement and the development of advanced algorithms for real-time configuration of the RIS units. In addition to that, the impact of mobility and dynamic environments on the proposed system can be investigated.

These directions emphasize the need for a multidisciplinary approach that bridges theoretical innovation with practical applications. Addressing these challenges will accelerate the development of robust, scalable, and efficient communication networks capable of meeting the demands of Industry 4.0, smart cities, and beyond.

> *"Information is the resolution of uncertainty."*
> *Claude Elwood Shannon*

References

[3GP23] 3GPP. System architecture for the 5G System (5GS). Technical Specification (TS) 23.501, 3rd Generation Partnership Project (3GPP), 2023. Version 18.4.0.

[3GP24] 3GPP. NR; NR and NG-RAN Overall description; Stage-2. Technical Specification (TS) 38.300, 3rd Generation Partnership Project (3GPP), 2024. Version 18.4.0.

[3GP25] 3GPP. Evolved Universal Terrestrial Radio Access (E-UTRA) and NR; Multi-connectivity; Overall description; Stage-2. Technical Specification (TS) 37.340, 3rd Generation Partnership Project (3GPP), 2025. Version 18.6.0.

[AFIM22] Bushra N Alsunbuli, Hassan F Fakhruldeen, Widad Ismail, and Nor M Mahyuddin. Hybrid beamforming with relay and dual-base stations blockage mitigation in millimetre-wave 5g communication applied in (viot). *Computers and Electrical Engineering*, 100:107953, 2022.

[AH21] Mohammedou Abewa and Jehad M Hamamreh. Multi-user auxiliary signal superposition transmission (MU-AS-ST) for secure and low-complexity multiple access communications, 2021.

[AH22] Mohamedou Abewa and Jehad M Hamamreh. D-SEAD: A novel multi-access multi-dimensional transmission technique for doubling the spectral efficiency per area and per device. *RS Open Journal on Innovative Communication Technologies*, 3(7), 2022.

[AHA21] Abdealrahman Abuqamar, Jehad M Hamamreh, and Mohamedou Abewa. STBC-assisted OFDM with subcarrier power modulation. *RS Open Journal on Innovative Communication Technologies*, 2(4):5, 2021.

[AHG+20] Muhammad Waseem Akhtar, Syed Ali Hassan, Rizwan Ghaffar, Haejoon Jung, Sahil Garg, and M Shamim Hossain. The shift to 6G communications: vision and requirements. *Human-centric Computing and Information Sciences*, 10(1):1–27, 2020.

[Ala98] Siavash M Alamouti. A simple transmit diversity technique for wireless communications. *IEEE Journal on selected areas in communications*, 16(8):1451–1458, 1998.

[AMBA23] Ingmar Andersson, Lars Manholm, Astrid Algaba Brazalez, and Parisa Aghdam. Antenna system requirements and challenges towards 6g - an industrial view. In *2023 International Workshop on Antenna Technology (iWAT)*, pages 1–4, 2023.

[ATAI18] Joseph Henry Anajemba, Yue Tang, James Adu Ansere, and Celestine Iwendi. Performance analysis of d2d energy efficient iot networks with relay-assisted underlaying technique. In *IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society*, pages 3864–3869, 2018.

[BDRDR+19] Ertugrul Basar, Marco Di Renzo, Julien De Rosny, Mérouane Debbah, Mohamed-Slim Alouini, and Rui Zhang. Wireless communications through reconfigurable intelligent surfaces. *IEEE access*, 7:116753–116773, 2019.

[BOR20] Gordana Barb, Marius Otesteanu, and Marius Roman. Dynamic spectrum sharing for lte-nr downlink mimo systems. In *2020 International Symposium on Electronics and Telecommunications (ISETC)*, pages 1–4, 2020.

[BRdR+20] E. Basar, M. Di Renzo, J. de Rosny, M. Debbah, M.-S. Alouini, and R. Zhang. Wireless communications through reconfigurable intelligent surfaces. *IEEE Communications Magazine*, 58(6):16–22, 2020.

[BuHZ+17] Imran Baig, Najam ul Hasan, Manaf Zghaibeh, Imran Ullah Khan, and Abdul Sattar Saand. A DST precoding based uplink NOMA scheme for PAPR reduction in 5G wireless network. In *2017 7th International Conference on Modeling, Simulation, and Applied Optimization (ICMSAO)*, pages 1–4. IEEE, 2017.

[CAH04] Runhua Chen, JG Andews, and RW Health. Multiuser space-time block coded MIMO with downlink precoding. In *2004 IEEE International Conference on Communications (IEEE Cat. No. 04CH37577)*, volume 5, pages 2689–2693. IEEE, 2004.

[CB20] Rodolfo W. L. Coutinho and Azzedine Boukerche. Modeling and analysis of a shared edge caching system for connected cars and industrial iot-based applications. *IEEE Transactions on Industrial Informatics*, 16(3):2003–2012, 2020.

[CfCR] Ireland Commission for Communications Regulation. [Last Accessed on 10-May-2023].

[CNBD22] Irched Chafaa, Romain Negrel, E Veronica Belmega, and Mérouane Debbah. Self-supervised deep learning for mmwave beam steering exploiting sub-6 ghz channels. *IEEE Transactions on Wireless Communications*, 21(10):8803–8816, 2022.

[CP09] Luis Miguel Cortes-Pena. MIMO space-time block coding (STBC): simulations and results. *Design Project: Personal and Mobile Communications, Georgia Tech (Ece6604)*, pages 1–8, 2009.

[CRH20] Michael Cribbs, Ric Romero, and Tri Ha. Orthogonal STBC set building and physical layer security application. In *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–5. IEEE, 2020.

[CSRKP+21] V Chandra Shekhar Rao, P Kumarswamy, MSB Phridviraj, S Venkatramulu, and V Subba Rao. 5g enabled industrial internet of things (iiot) architecture for smart manufacturing. In *Data Engineering and Communication Technology: Proceedings of ICDECT 2020*, pages 193–201. Springer, 2021.

[CYYL20] Chang Cai, Huiyuan Yang, Xiaojun Yuan, and Yina-Chang Liang. Two-timescale optimization for intelligent reflecting surface aided d2d underlay communication. In *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, pages 1–6, 2020.

[DGA13] Harpreet S. Dhillon, Radha Krishna Ganti, and Jeffrey G. Andrews. Load-aware modeling and analysis of heterogeneous cellular networks. *IEEE Transactions on Wireless Communications*, 12(4):1666–1677, 2013.

[DJJ20] Anutusha Dogra, Rakesh Kumar Jha, and Shubha Jain. A survey on beyond 5g network with the advent of 6g: Architecture and emerging technologies. *IEEE Access*, 9:67512–67547, 2020.

[DKN+21] Tri Nhu Do, Georges Kaddoum, Thanh Luan Nguyen, Daniel Benevides Da Costa, and Zygmunt J Haas. Aerial reconfigurable intelligent surface-aided wireless communication systems. In *2021 IEEE 32nd Annual International Symposium on Personal,* *Indoor and Mobile Radio Communications (PIMRC)*, pages 525–530. IEEE, 2021.

[DLA02] Mischa Dohler, Emeline Lefranc, and Hamid Aghvami. Space-time block codes for virtual antenna arrays. In *The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, volume 1, pages 414–417. IEEE, 2002.

[DLC+17] Zhiguo Ding, Yuanwei Liu, Jinho Choi, Qi Sun, Maged Elkashlan, I. Chih-Lin, and H. Vincent Poor. Application of non-orthogonal multiple access in lte and 5g networks. *IEEE Communications Magazine*, 55(2):185–191, 2017.

[DMP+23] Romain Dulout, Leo Mendiboure, Yannis Pousset, Virginie Deniau, and Frederic Launay. Non-orthogonal multiple access for offloading in multi-access edge computing: A survey. *IEEE Access*, 11:118983–119016, 2023.

[DNF24] M. Danger, J. A. Nossek, and G. P. Fettweis. Experimental evaluation of reconfigurable intelligent surfaces in industrial millimeter-wave networks. *IEEE Mobile Networks and Applications*, 2024. Accepted / Early Access.

[DPS18] Erik Dahlman, Stefan Parkvall, and Johan Skold. *5G NR: The Next Generation Wireless Access Technology*. Academic Press, Inc., USA, 1st edition, 2018.

[DRDPH+20] Marco Di Renzo, Mérouane Debbah, D-T Phan-Huy, Alessio Zappone, Mohamed-Slim Alouini, Chau Yuen, Vincenzo Sciancalepore, George C Alexandropoulos, Jakob Hoydis, Haris 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*, 2019(1):1–20, 2020.

[ea14] R. Zhang et al. Multiple-antenna techniques for wireless communications: a comprehensive literature survey. *IEEE Commun. Surveys Tuts.*, 15(1):5–26, 2014.

[ea18] L. Dai et al. Non-orthogonal multiple access for 5g: Solutions, challenges, opportunities, and future research trends. *IEEE Commun. Mag.*, 53(9):74–81, 2018.

[ea20a] N. Liang et al. Spectrum and energy efficient beamspace mimo-noma for millimeter-wave communications using lens antenna array. *IEEE J. Sel. Areas Commun.*, 38(9):1907–1923, 2020.

[ea20b] W. Yuan et al. Future wireless communication networks: Concepts, challenges and key technologies. *IEEE Commun. Surveys Tuts.*, 22(2):1027–1072, 2020.

[ea20c] X. Mu et al. Exploiting intelligent reflecting surfaces in multi-antenna noma networks. *IEEE Trans. Wireless Commun.*, 19(10):6884–6898, 2020.

[EAAG22] Mohamed Elhattab, Mohamed Amine Arfaoui, Chadi Assi, and Ali Ghrayeb. Ris-assisted joint transmission in a two-cell downlink noma cellular system. *IEEE Journal on Selected Areas in Communications*, 40(4):1270–1286, 2022.

[Eri] Ericsson. Ericsson mobility report 2025 ericsson mobility report november 2025. Last accessed on: 05-01-2026.

[ERM19] Amine EL RHAYOUR and Tomader MAZRI. 5g architecture: Deployment scenarios and options. In *2019 International Symposium on Advanced Electrical and Communication Technologies (ISAECT)*, pages 1–6, 2019.

[ETS17] ETSI. 5G; Study on Scenarios and Requirements for Next Generation Access Technologies (3GPP TR 38.913 version 14.2.0 Release 14). Technical Report V14.2.0, ETSI, May 2017.

[ETS22] ETSI. 5G; Study on channel model for frequencies from 0.5 to 100 GHz. Technical Report V17.0.0, ETSI, April 2022.

[Fos96] Gerard J. Foschini. Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. *Bell Labs Technical Journal*, 1(2):41–59, 1996.

[FZCL16] Fang Fang, Haijun Zhang, Julian Cheng, and Victor CM Leung. Energy-efficient resource allocation for downlink non-orthogonal multiple access network. *IEEE Transactions on Communications*, 64(9):3722–3732, 2016.

[Gho20] Morteza Ghobakhloo. Industry 4.0, digitization, and opportunities for sustainability. *Journal of cleaner production*, 252:119869, 2020.

[GLL19] Chongtao Guo, Le Liang, and Geoffrey Ye Li. Resource allocation for low-latency vehicular communications: An effective capacity perspective. *IEEE Journal on Selected Areas in Communications*, 37(4):905–917, 2019.

[GP01] Dhananjay Gore and Arogyaswami Paulraj. Space-time block coding with optimal antenna selection. In *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 01CH37221)*, volume 4, pages 2441–2444. IEEE, 2001.

[GQLAD21] Yiyu Guo, Zhijin Qin, Yuanwei Liu, and Naofal Al-Dhahir. Intelligent reflecting surface aided multiple access over fading channels. *IEEE Transactions on Communications*, 69(3):2015–2027, 2021.

[GTL14] Zhou Guohua, Deng Tianle, and Yang Li. A dynamic spectrum reallocation scheme in gsm and lte co-existed networks. In *2014 International Symposium on Wireless Personal Multimedia Communications (WPMC)*, pages 595–600, 2014.

[GXW+21] Zhengnian Gao, Yongjun Xu, Qianzhu Wang, Qingqing Wu, and Dong Li. Outage-constrained energy efficiency maximization for ris-assisted wpcns. *IEEE communications letters*, 25(10):3370–3374, 2021.

[HAL20] Jehad M Hamamreh, Mohamedou Abewa, and JP Lemayian. New non-orthogonal transmission schemes for achieving highly efficient, reliable, and secure multi-user communications. *RS Open Journal on Innovative Communication Technologies*, 1(2):12, 2020.

[Ham09a] Khairi Ashour Hamdi. On the statistics of signal-to-interference plus noise ratio in wireless communications. *IEEE transactions on communications*, 57(11):3199–3204, 2009.

[Ham09b] Khairi Ashour Hamdi. On the statistics of signal-to-interference plus noise ratio in wireless communications. *IEEE Transactions on Communications*, 57(11):3199–3204, 2009.

[HF20] Noha Hassan and Xavier Fernando. An optimum user association algorithm in heterogeneous 5g networks using standard deviation of the load. *Electronics*, 9(9), 2020.

[HLWJ20] Jiguang He, Markus Leinonen, Henk Wymeersch, and Markku Juntti. Channel estimation for ris-aided mmwave mimo channels. *arXiv preprint arXiv:2002.06453*, 2020.

[HS13] Zakaria Hanzaz and Hans Dieter Schotten. Analysis of effective sinr mapping models for mimo ofdm in lte system. In *2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 1509–1515, 2013.

[HTM23] Yulong He, Wei Tao, and Changgen Mei. Chapter 5 - other environmental problems during the operation of high-speed railway. In Yulong He, Wei Tao, and Changgen Mei, editors, *High-Speed Railways*, pages 97–115. Elsevier, 2023. Last accessed on: 21-03-2024.

[IA17] Ali M.A. Ibrahim and Murtada M. Abdelwahab. Sum rate analysis of massive multiple input multiple output system for linear precoding using normalization methods. In *2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE)*, pages 1–7, 2017.

[IH21a] Sadiq Iqbal and Jehad Hamamreh. Improving throughput and reliability performance of future 6G-IoT communication systems using signal superposition-based dual transmission. *RS Open Journal on Innovative Communication Technologies*, 2(5):9, 2021.

[IH21b] Sadiq Iqbal and Jehad M Hamamreh. A new dual transmission technique employing auxiliary signal superposition for improving the data rate and diversity of next-generation communication systems. *RS Open Journal on Innovative Communication Technologies*, 2(5), 2021.

[IKH21] Sadiq Iqbal, Shaheer Khalid, and Jehad M Hamamreh. A new MIMO technique utilizing superimposed auxiliary signals for simultaneously achieving spatial multiplexing and diversity gains in MIMO-aided communication systems. *RS Open Journal on Innovative Communication Technologies*, 2(4):6, 2021.

[IKH22] Sadiq Iqbal, Shaheer Khalid, and Jehad M Hamamreh. Improving the performance of cell-edge users in 6G and beyond networks by utilizing a novel precoding-based hybrid CoMP transmission design. *RS Open Journal on Innovative Communication Technologies*, 3(7), 2022.

[IRH+21] Adeel Iqbal, Muddasir Rahim, Riaz Hussain, Abdulfattah Noorwali, Mohammad Zubair Khan, Atif Shakeel, Irfan Latif Khan, Muhammad Awais Javed, Qadeer Ul Hasan, and Shahzad A Malik. cdersa: Cognitive d2d enabled relay selection algorithm to mitigate blind-spots in 5g cellular networks. *IEEE Access*, 9:89972–89988, 2021.

[ITU] In 2012, the itu's radiocommunications sector (itu-r) embarked on a programme to develop international mobile telecommunication (imt) standards for 5g by 2020.

[IUIL23] Agbotiname Lucky Imoize, Friday Udeji, Joseph Isabona, and Cheng-Chi Lee. Optimizing the quality of service of mobile broadband networks for a dense urban environment. *Future Internet*, 15(5):181, 2023.

[JAA+22] Muhammad Ali Jamshed, Kamran Ali, Qammer H. Abbasi, Muhammad Ali Imran, and Masood Ur-Rehman. Challenges, applications, and future of wireless sensors in internet of things: A review. *IEEE Sensors Journal*, 22(6):5482–5494, 2022.

[JHB21a] Muhammad Ali Jamshed, Fabien Heliot, and Tim WC Brown. Unsupervised learning based emission-aware uplink resource allocation scheme for non-orthogonal multiple access systems. *ieee transactions on vehicular technology*, 70(8):7681–7691, 2021.

[JHB21b] Muhammad Ali Jamshed, Fabien Héliot, and Tim W. C. Brown. Unsupervised learning based emission-aware uplink resource allocation scheme for non-orthogonal multiple access systems. *IEEE Transactions on Vehicular Technology*, 70(8):7681–7691, 2021.

[JKMK14] Minho Jo, Mykhailo Klymash, Taras Maksymyuk, and Ruslan Kozlovskiy. Dynamic spectrum sharing algorithm for combined mobile networks. In *2014 20th International Conference on Mi-*

crowaves, Radar and Wireless Communications (MIKON), pages 1–3, 2014.

[JKP+22] Muhammad Ali Jamshed, Wali Ullah Khan, Haris Pervaiz, Muhammad Ali Imran, and Masood Ur-Rehman. Emission-aware resource optimization framework for backscatter-enabled uplink noma networks. In *2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, pages 1–5, 2022.

[JNJJ20] Furqan Jameel, Muhammad Nabeel, Muhammad Ali Jamshed, and Riku Jäntti. Minimizing forking in blockchain-based iot networks. In *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–6, 2020.

[Kha20] Muhammad Farhan Khan. An approach for optimal base station selection in 5g hetnets for smart factories. In *2020 IEEE 21st International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM)*, pages 64–65, 2020.

[Kha24] Muhammad Turyalaiérouane Khan. A modified convolutional neural network with rectangular filter for frequency-hopping spread spectrum signals. *Applied Soft Computing Journal*, 150(111036), 2024.

[KHG23] Pratap Khuntia, Ranjay Hazra, and Pratik Goswami. A bidirectional relay-assisted underlay device-to-device communication in cellular networks: An iot application for fintech. *IEEE Internet of Things Journal*, 10(3):2174–2182, 2023.

[KIH21] Shaheer Khalid, Sadiq Iqbal, and Jehad M Hamamreh. A new enhanced coordinated multi-point (CoMP) transmission design for cancelling inter-cell interference for cell-edge users using superimposed supporting signals. *RS Open Journal on Innovative Communication Technologies*, 2(6):12, 2021.

[KKAC24] M. F. Khan, Z. Kaleem, S. H. Ahmed, and K. R. Choo. Performance analysis of intelligent reflecting surfaces for 5g/6g-enabled future smart industries with a focus on millimeter-wave band communications. In *Proceedings of the IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 85–93. IEEE, 2024.

[KLL19] Lusekelo Kibona, Jian Liu, and Yingzhuang Liu. Ergodic sum-rate analysis for massive mimo under imperfect channel state information. In *2019 Photonics Electromagnetics Research Symposium - Fall (PIERS - Fall)*, pages 3243–3249, 2019.

[KRIP23] Muhammad Farhan Khan, Adnan Rashid, Adeel Iqbal, and Dirk Pesch. Enabling high-speed connectivity in urban environments through composite base stations and dynamic spectrum scheduling. In *2023 IEEE 34th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 1–6. IEEE, 2023.

[KSR17] Farooque Hassan Kumbhar, Navrati Saxena, and Abhishek Roy. Reliable relay: Autonomous social d2d paradigm for 5g los communications. *IEEE Communications Letters*, 21(7):1593–1596, 2017.

[KTJ+22] Bilal Khan, Nuutti Tervo, Markku Jokinen, Aarno Pärssinen, and Markku Juntti. Statistical digital predistortion of 5g millimeter-wave rf beamforming transmitter under random amplitude variations. *IEEE Transactions on Microwave Theory and Techniques*, 70(9):4284–4296, 2022.

[KUU23] Muhammad Turyalai Khan and Sheikhérouane Usman Ullah. A hybrid convolutional neural network with fusion of handcrafted and deep features for fhss signals classification. *Expert Systems with Applications*, 225(120153), 2023.

[KWS+22] Tewelgn Kebede, Yihenew Wondie, Johannes Steinbrunn, Hailu Belay Kassa, and Kevin T. Kornegay. Precoding and beamforming techniques in mmwave-massive mimo: Performance assessment. *IEEE Access*, 10:16365–16387, 2022.

[LETM14] Erik G. Larsson, Ove Edfors, Fredrik Tufvesson, and Thomas L. Marzetta. Massive mimo for next generation wireless systems. *IEEE Communications Magazine*, 52(2):186–195, 2014.

[LFK+14] Heiner Lasi, Peter Fettke, Hans-Georg Kemper, Thomas Feld, and Michael Hoffmann. Industry 4.0. *Business & information systems engineering*, 6:239–242, 2014.

[LGZ+22] Spyros Lavdas, Panagiotis K Gkonis, Zinon Zinonos, Panagiotis

Trakadas, Lambros Sarakis, and Konstantinos Papadopoulos. A machine learning adaptive beamforming framework for 5g millimeter wave massive mimo multicellular networks. *IEEE Access*, 10:91597–91609, 2022.

[LH20] Joel Poncha Lemayian and Jehad Hamamreh. Novel small-scale NOMA communication technique using auxiliary signal superposition. In *2020 International Conference on UK-China Emerging Technologies (UCET)*, pages 1–4. IEEE, 2020.

[LH21] Joel Poncha Lemayian and Jehad M Hamamreh. Hybrid MIMO: a new transmission method for simultaneously achieving spatial multiplexing and diversity gains in MIMO systems. *RS Open Journal on Innovative Communication Technologies*, 2(4), 2021.

[LH22] Gayatri D Londhe and Vaibhav S Hendre. Review on beamforming techniques for millimeter wave massive mimo. In *Proceedings of the International e-Conference on Intelligent Systems and Signal Processing: e-ISSP 2020*, pages 291–303. Springer, 2022.

[Liu15] Tsung-Hsien Liu. Analysis of the alamouti STBC MIMO system with spatial division multiplexing over the rayleigh fading channel. *IEEE Transactions on Wireless Communications*, 14(9):5156–5170, 2015.

[LLMJ18] Tingwei Liu, John C. S. Lui, Xiaoqiang Ma, and Hongbo Jiang. Enabling relay-assisted d2d communication for cellular networks: Algorithm and protocols. *IEEE Internet of Things Journal*, 5(4):3136–3150, 2018.

[LWC+16] Dantong Liu, Lifeng Wang, Yue Chen, Maged Elkashlan, Kai-Kit Wong, Robert Schober, and Lajos Hanzo. User association in 5g networks: A survey and an outlook. *IEEE Communications Surveys Tutorials*, 18(2):1018–1044, 2016.

[LYX+20] Nan Liang, Longzhuang Yang, Wenkai Xu, Junzhi Wang, Zhongjin Jiang, and Shaoqiang Wang. Spectrum and energy efficient beamspace mimo-noma for millimeter-wave communications using lens antenna array. *IEEE Journal on Selected Areas in Communications*, 38(9):1907–1923, 2020.

[LZL+21] Yunlong Liu, Qihui Zhang, Yanguang Liu, Wei Ma, Jie Li, and Wuyang Zhuang. Intelligent reflecting surface aided mimo broadcasting for simultaneous wireless information and power transfer. *IEEE Transactions on Wireless Communications*, 20(2):1070–1084, 2021.

[LZMS24] Y. Liu, H. Zhang, X. Mu, and R. Schober. Ris-assisted millimeter-wave communications for mobile robots in smart factories. *IEEE Internet of Things Journal*, 2024. Early Access.

[MAD+18] Martin Klaus Müller, Fjolla Ademaj, Thomas Dittrich, Agnes Fastenbauer, Blanca Ramos Elbal, Armand Nabavi, Lukas Nagel, Stefan Schwarz, and Markus Rupp. Flexible multi-node simulation of cellular mobile communications: the vienna 5g system level simulator. *EURASIP Journal on Wireless Communications and Networking*, 2018(1):227, 2018.

[MAP+17] Shahid Mumtaz, Ahmed Alsohaily, Zhibo Pang, Ammar Rayes, Kim Fung Tsang, and Jonathan Rodriguez. Massive internet of things for industrial applications: Addressing wireless iiot connectivity challenges and ecosystem fragmentation. *IEEE Industrial Electronics Magazine*, 11(1):28–33, 2017.

[Mar10] Thomas L. Marzetta. Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Transactions on Wireless Communications*, 9(11):3590–3600, 2010.

[MCF+21] Huan Ma, Guofa Cai, Yi Fang, Pingping Chen, and Guojun Han. Design and performance analysis of a new STBC-MIMO LoRa system. *IEEE Transactions on Communications*, 69(9):5744–5757, 2021.

[McL] Ronan McLaughlin. 5g low latency requirements. Last accessed on: 21-03-2024.

[MLG+20] Xidong Mu, Yuanwei Liu, Weidong Guo, Jie Lin, and Naofal Al-Dhahir. Exploiting intelligent reflecting surfaces in multi-antenna noma networks. *IEEE Transactions on Wireless Communications*, 19(10):6884–6898, 2020.

[MNSA22] Koki Miyamoto, Koji Nishibe, Takanori Shibakura, and Chang-Jun Ahn. Outage probability of downlink massive mimo with mrt precoding under multicell environment. In *2022 IEEE In-*

ternational Conference on Consumer Electronics-Asia (ICCE-Asia), pages 1–5, 2022.

[MOE15] Mohammad A. B. Mohammad, Abdalla A. Osman, and Nihad A. A. Elhag. Performance comparison of mrt and zf for single cell downlink massive mimo system. In *2015 International Conference on Computing, Control, Networking, Electronics and Embedded Systems Engineering (ICCNEEE)*, pages 52–56, 2015.

[MPC22] Zi Long Ma, Shao Cong Peng, and Qing-Xin Chu. A triple-band composite base station and satellite antenna array for maritime application. *International Journal of RF and Microwave Computer-Aided Engineering*, 32(6):e23140, 2022.

[MTZS22] Sergiy Melnyk, Shreya Tayade, Mervat Zarour, and Hans D Schotten. Wireless industrial communication and control system: Ai assisted blind spot detection-and-avoidance for agvs. In *2022 International Workshop on Emerging Trends and Case-Studies in Industry 4.0 and Intelligent Manufacturing (ETCIIM)*, pages 307–313. ELSEVIER, 2022.

[NARFJ+22] Md. Noor-A-Rahim, Fadhil Firyaguna, Jobish John, M. Omar Khyam, Dirk Pesch, Eddie Armstrong, Holger Claussen, and H. Vincent Poor. Toward industry 5.0: Intelligent reflecting surface in smart manufacturing. *IEEE Communications Magazine*, 60(10):72–78, 2022.

[NGI22] Mahdi Nangir, Abdolrasoul Sakhaei Gharagezlou, and Nima Imani. Comparison of the mrt and zf precoding in massive mimo systems from energy efficiency viewpoint. In *2022 30th International Conference on Electrical Engineering (ICEE)*, pages 803–807, 2022.

[NJA+21] Ali Nauman, Muhammad Ali Jamshed, Rashid Ali, Korhan Cengiz, Zulqarnain, and Sung Won Kim. Reinforcement learning-enabled intelligent device-to-device (i-d2d) communication in narrowband internet of things (nb-iot). *Computer Communications*, 176:13–22, 2021.

[NLD22] Nhan Thanh Nguyen, Kyungchun Lee, and Huaiyu Dai. Hybrid beamforming and adaptive rf chain activation for uplink cell-free millimeter-wave massive mimo systems. *IEEE Transactions on Vehicular Technology*, 71(8):8739–8755, 2022.

[NNP17] S Nandi, A Nandi, and NN Pathak. Performance analysis of alamouti STBC MIMO OFDM for different transceiver system. In *2017 International Conference on Intelligent Sustainable Systems (ICISS)*, pages 883–887. IEEE, 2017.

[NQA+20] Ali Nauman, Yazdan Ahmad Qadri, Muhammad Amjad, Yousaf Bin Zikria, Muhammad Khalil Afzal, and Sung Won Kim. Multimedia internet of things: A comprehensive survey. *IEEE Access*, 8:8202–8250, 2020.

[NR22] S Nandan and M Abdul Rahiman. Intelligent reflecting surface (irs) assisted mmwave wireless communication systems: A survey. *Journal of Communications*, 17(9), 2022.

[NZY21] Zhaoyang Niu, Guoqiang Zhong, and Hui Yu. A review on the attention mechanism of deep learning. *Neurocomputing*, 452:48–62, 2021.

[ODS+20] Mohanad Obeed, Hayssam Dahrouj, Anas M Salhab, Salam A Zummo, and Mohamed-Slim Alouini. User pairing, link selection, and power allocation for cooperative noma hybrid vlc/rf systems. *IEEE Transactions on Wireless Communications*, 20(3):1785–1800, 2020.

[OKS+18] Antonino Orsino, Roman Kovalchukov, Andrey Samuylov, Dmitri Moltchanov, Sergey Andreev, Yevgeni Koucheryavy, and Mikko Valkama. Caching-aided collaborative d2d operation for predictive data dissemination in industrial iot. *IEEE Wireless Communications*, 25(3):50–57, 2018.

[Ope] OpenStreetMap. [Last Accessed on 17-May-2023].

[OUST19] E. Obi, Aliyu Danjuma Usman, Suleiman Muhammad Sani, and Abdoulie Momodou Sunkary Tekanyi. Hybrid user association algorithm with optimal bandwidth allocation and power control for energy efficiency optimization in 5g hetnet. *ATBU Journal of Science, Technology and Education*, 7(4):78–85, 2019.

[Pak13] Eakkamol Pakdeejit. Linear precoding performance of massive mu-mimo downlink system. Master's thesis, Linköping UniversityLinköping University, Communication Systems, The Institute of Technology, 2013.

[Pan20] Madan Kumar Pande. Performance study of massive mimo with linear precoders using open source simulators. In *2020 IEEE-HYDCON*, pages 1–6, 2020.

[PK93] A. J. Paulraj and T. Kailath. Increasing capacity in wireless communications via antenna arrays. *Applied Microwave & Wireless*, 5(3):116–123, 1993.

[PKJ14] Tebe Parfait, Yujun Kuang, and Kponyo Jerry. Performance analysis and comparison of zf and mrt based downlink massive mimo systems. In *2014 Sixth International Conference on Ubiquitous and Future Networks (ICUFN)*, pages 383–388, 2014.

[PPZC20] Dragan Peraković, Marko Periša, Petra Zorić, and Ivan Cvitić. Development and implementation possibilities of 5g in industry 4.0. In *Design, Simulation, Manufacturing: The Innovation Exchange*, pages 166–175. Springer, 2020.

[PRW$^+$20] Cunhua Pan, Hong Ren, Kezhi Wang, Wei Xu, Maged Elkashlan, Arumugam Nallanathan, and Lajos Hanzo. Multicell mimo communications relying on intelligent reflecting surfaces. *IEEE Transactions on Wireless Communications*, 19(8):5218–5233, 2020.

[PT18] Subuh Pramono and Eddy Triyono. Comparative performance analysis of linear precoding in downlink multi-user mimo. In *2018 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, pages 277–280, 2018.

[PTM$^+$18a] Stefan Pratschner, Bashar Tahir, Ljiljana Marijanovic, Mariam Mussbah, Kiril Kirev, Ronald Nissel, Stefan Schwarz, and Markus Rupp. Versatile mobile communications simulation: the Vienna 5G Link Level Simulator. *EURASIP Journal on Wireless Communications and Networking*, 2018(1):226, September 2018.

[PTM$^+$18b] Stefan Pratschner, Bashar Tahir, Ljiljana Marijanovic, Mariam Mussbah, Kiril Kirev, Ronald Nissel, Stefan Schwarz, and Markus Rupp. Versatile mobile communications simulation: the Vienna 5G Link Level Simulator. *EURASIP Journal on Wireless Communications and Networking*, 2018(1):226, September 2018.

[QKJR+25] Abdul Qadeer Khan, Kamran Javed, Ali Raza, Shahryar Saleem, and Zubair Saeed. Energy-efficient cell association and load balancing for low battery users in heterogeneous cellular networks. *IEEE Access*, 13:82621–82636, 2025.

[RIT22] M. Rahim, M. A. Imran, and R. Tafazolli. Reconfigurable intelligent surfaces for smart manufacturing: A technology overview. *IEEE Communications Magazine*, 60(3):42–48, 2022.

[RMG03] Markus Rupp, Christoph Mecklenbrauker, and Gerhard Gritsch. High diversity with simple space time block-codes and linear receivers. In *GLOBECOM'03. IEEE Global Telecommunications Conference (IEEE Cat. No. 03CH37489)*, volume 1, pages 302–306. IEEE, 2003.

[RNS+20] M. Di Renzo, K. Ntontin, J. Song, F. Danufane, X. Gao, J. de Rosny, R. Zhang, S. Tretyakov, and M. Debbah. Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison. *IEEE Communications Magazine*, 58(6):22–28, 2020.

[RTTH20] Fatemeh Rezaei, Chintha Tellambura, Ali Akbar Tadaion, and Ali Reza Heidarpour. Rate analysis of cell-free massive mimo-noma with three linear precoders. *IEEE Transactions on Communications*, 68(6):3480–3494, 2020.

[RWDX12] Pinyi Ren, Yichen Wang, Qinghe Du, and Jing Xu. A survey on dynamic spectrum access protocols for distributed cognitive wireless networks. *EURASIP Journal on Wireless Communications and Networking*, 2012(1):1–21, 2012.

[SCW+11] Yi-Sheng Shiu, Shih Yu Chang, Hsiao-Chun Wu, Scott C-H Huang, and Hsiao-Hwa Chen. Physical layer security in wireless networks: A tutorial. *IEEE wireless Communications*, 18(2):66–74, 2011.

[SFL19] Eric Pierre Simon, Joumana Farah, and Pierre Laly. Performance evaluation of massive mimo with beamforming and nonorthogonal multiple access based on practical channel measurements. *IEEE Antennas and Wireless Propagation Letters*, 18(6):1263–1267, 2019.

[Sha48] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948.

[SKB+13] Yuya Saito, Yoshihisa Kishiyama, Anass Benjebbour, Takehiro Nakamura, Anxin Li, and Kenichi Higuchi. Non-orthogonal multiple access (noma) for cellular future radio access. In *2013 IEEE 77th vehicular technology conference (VTC Spring)*, pages 1–5. IEEE, 2013.

[SKSK23] K Kalyana Srinivas, Bojja Sai Kiran, Thati Sathvika, and Peram Kishore. An adaptive hybrid beamforming technique for analysis of throughput, blocking probability, transmission power in 5g mimo mmwave. In *2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, pages 1–6. IEEE, 2023.

[SKYK11] K. Son, H. Kim, Y. Yi, and B. Krishnamachari. Base station operation and user association mechanisms for energy-delay tradeoffs in green cellular networks. *IEEE Journal on Selected Areas in Communications*, 29(8):1525–1536, 2011.

[SM20] Abdelhamid Salem and Leila Musavian. Noma in cooperative communication systems with energy-harvesting nodes and wireless secure transmission. *IEEE Transactions on Wireless Communications*, 20(2):1023–1037, 2020.

[SMS+17] Mansoor Shafi, Andreas F. Molisch, Peter J. Smith, Thomas Haustein, Peiying Zhu, Prasan De Silva, Fredrik Tufvesson, Anass Benjebbour, and Gerhard Wunder. 5g: A tutorial overview of standards, trials, challenges, deployment, and practice. *IEEE Journal on Selected Areas in Communications*, 35(6):1201–1221, 2017.

[SONA22] Satoshi Suyama, Tatsuki Okuyama, Nobuhide Nonaka, and Takahiro Asai. Recent studies on massive mimo technologies for 5g evolution and 6g. In *2022 IEEE Radio and Wireless Symposium (RWS)*, pages 90–93. IEEE, 2022.

[SS17] K Shamganth and Martin JN Sibley. A survey on relay selection in cooperative device-to-device (d2d) communication for 5g cellular networks. In *2017 International Conference on Energy, Com-*

munication, Data Analytics and Soft Computing (ICECDS), pages 42–46. IEEE, 2017.

[SSCY17] Meysam Sadeghi, Luca Sanguinetti, Romain Couillet, and Chau Yuen. Large system analysis of power normalization techniques in massive mimo. *IEEE Transactions on Vehicular Technology*, 66(10):9005–9017, 2017.

[Stu17] Gordon L. Stuber. *Principles of Mobile Communication*. Springer Publishing Company, Incorporated, 4th edition, 2017.

[Tel99] I. E. Telatar. Capacity of multi-antenna gaussian channels. *European Transactions on Telecommunications*, 10(6):585–595, Nov 1999.

[THQ18] Le Thanh Tan, Rose Qingyang Hu, and Yi Qian. D2d communications in heterogeneous networks with full-duplex relays and edge caching. *IEEE Transactions on Industrial Informatics*, 14(10):4557–4567, 2018.

[TJC99] Vahid Tarokh, Hamid Jafarkhani, and A Robert Calderbank. Space-time block codes from orthogonal designs. *IEEE Transactions on Information theory*, 45(5):1456–1467, 1999.

[TK18] Mesut Toka and Oğuz Kucur. Non-orthogonal multiple access with alamouti space–time block coding. *IEEE Communications Letters*, 22(9):1954–1957, 2018.

[TLBN05] Mizhou Tan, Zoran Latinovic, and Yeheskel Bar-Ness. STBC MIMO-OFDM peak-to-average power ratio reduction by cross-antenna rotation and inversion. *IEEE Communications letters*, 9(7):592–594, 2005.

[TSAJ14] Ralph Tanbourgi, Sarabjot Singh, Jeffrey G Andrews, and Friedrich K Jondral. A tractable model for noncoherent joint-transmission base station cooperation. *IEEE Transactions on Wireless Communications*, 13(9):4959–4973, 2014.

[UM20] Anthony Ngozichukwuka Uwaechia and Nor Muzlifah Mahyuddin. A comprehensive survey on millimeter wave communications for fifth-generation wireless networks: Feasibility and challenges. *IEEE Access*, 8:62367–62414, 2020.

[UMG+24] Muhammad Umer, Muhammad Ahmed Mohsin, Mikael Gidlund, Haejoon Jung, and Syed Ali Hassan. Analysis of star-ris-assisted downlink comp-noma multi-cell networks under nakagami-m fading. *IEEE Communications Letters*, 28(5):1009–1013, 2024.

[UMH+23] Muhammad Umer, Muhammad Ahmed Mohsin, Syed Ali Hassan, Haejoon Jung, and Haris Pervaiz. Performance analysis of star-ris enhanced comp-noma multi-cell networks. In *2023 IEEE Globecom Workshops (GC Wkshps)*, pages 2000–2005, 2023.

[Uni15] IT Union. Nomenclature of the frequency and wavelength bands used in telecommunications. *Recommendation ITU/RV*, pages 431–438, 2015.

[WHA+21] Li Wei, Chongwen Huang, George C Alexandropoulos, Chau Yuen, Zhaoyang Zhang, and Mérouane Debbah. Channel estimation for ris-empowered multi-user miso wireless communications. *IEEE Transactions on Communications*, 69(6):4144–4157, 2021.

[WLZ18] Xiaoyong Wu, Danpu Liu, and Zhe Zheng. Zero-forcing precoding for the mu-miso downlink: How many antennas do we need? In *2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 1433–1439, 2018.

[WZ19] Qingqing Wu and Rui Zhang. Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming. *IEEE transactions on wireless communications*, 18(11):5394–5409, 2019.

[WZR24] Qingqing Wu, Rui Zhang, and Marco Di Renzo. Intelligent surfaces empowered wireless network: Recent advances and the road to 6G. *Proceedings of the IEEE*, 112(1):75–109, 2024.

[YLXL21] Xianhua Yu, Dong Li, Yongjun Xu, and Ying-Chang Liang. Convolutional autoencoder-based phase shift feedback compression for intelligent reflecting surface-assisted wireless systems. *IEEE Communications Letters*, 26(1):89–93, 2021.

[YRC+13] Qiaoyang Ye, Beiyu Rong, Yudong Chen, Mazin Al-Shalash, Constantine Caramanis, and Jeffrey G. Andrews. User Association for

Load Balancing in Heterogeneous Cellular Networks. *IEEE Transactions on Wireless Communications*, 12(6):2706–2716, January 2013.

[YXL+22] Xinwei Yue, Jin Xie, Yuanwei Liu, Zhihao Han, Rongke Liu, and Zhiguo Ding. Simultaneously transmitting and reflecting reconfigurable intelligent surface assisted noma networks. *IEEE Transactions on Wireless Communications*, 22(1):189–204, 2022.

[ZCM+18] Jiankang Zhang, Sheng Chen, Robert G. Maunder, Rong Zhang, and Lajos Hanzo. Regularized zero-forcing precoding-aided adaptive coding and modulation for large-scale antenna array-based air-to-air communications. *IEEE Journal on Selected Areas in Communications*, 36(9):2087–2103, 2018.

[ZLMS24] H. Zhang, Y. Liu, X. Mu, and R. Schober. Intelligent reflecting surfaces for smart factory: Deployment, blockage, and performance analysis. *IEEE Transactions on Wireless Communications*, 2024. Early Access.