

DOCTORAL THESIS

Multi-Objective Based Radio Resource Optimization for Massive Connectivity in 6G Networks

Aamir Latif

TALLINN UNIVERSITY OF TECHNOLOGY
DOCTORAL THESIS
45/2026

Multi-Objective Based Radio Resource Optimization for Massive Connectivity in 6G Networks

AAMIR LATIF



TALLINN UNIVERSITY OF TECHNOLOGY

School of Information Technologies

Thomas Johann Seebeck Department of Electronics

The dissertation was accepted for the defence of the degree of Doctor of Philosophy (Information and Communication Technology) on 16 July 2026.

Supervisor: Professor Muhammad Mahtab Alam,
Thomas Johann Seebeck Department of Electronics, School of Information Technologies,
Tallinn University of Technology
Tallinn, Estonia

Co-supervisor: Professor Yannick Le Moullec,
Thomas Johann Seebeck Department of Electronics, School of Information Technologies,
Tallinn University of Technology
Tallinn, Estonia

Opponents: Professor Olivier Berder,
IUT Lannion, University of Rennes,
Rennes, France

Professor Maurizio Magarini,
Politecnico di Milano,
Milan, Italy

Defence of the thesis: 26 August 2026, Tallinn

Declaration:

Hereby I declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology, has not been submitted for any academic degree elsewhere.

AAMIR LATIF

signature

Copyright: Aamir Latif, 2026
ISSN 2585-6898 (paperback)
ISBN 978-9916-80-529-9 (paperback)
ISSN 2585-6901 (PDF)
ISBN 978-9916-80-530-5 (PDF)
<https://doi.org/10.23658/taltech.45/2026>
Printed by EVG Print

Latif, A. (2026). *Multi-Objective-based radio resource optimization for Massive Connectivity in 6G Networks* [TalTech Press]. <https://doi.org/10.23658/taltech.45/2026>

TALLINNA TEHNIKAÜLIKOOL
DOKTORITÖÖ
45/2026

Raadioressursside mitme eesmärgiga optimeerimine massilise ühenduvuse saavutamiseks 6G-võrkudes

AAMIR LATIF



Contents

List of Publications	8
Author's Contributions to the Publications	9
Abbreviations.....	11
1 Introduction	13
1.1 Background and Motivation	13
1.2 Central Research Problem	16
1.2.1 Research questions.....	16
1.3 Research Methodology	18
1.4 Summary of Thesis Contributions	19
1.5 Thesis Organization	21
2 State of the Art and Related Works	23
2.1 Overview and Scope	23
2.2 Challenge 1 — Coordination: O-RAN Scheduling and Energy-Aware Design .	23
2.2.1 O-RAN Functional Architecture and Control Hierarchy	24
2.2.2 Scheduler-Level Energy Profiling in O-RAN.....	24
2.2.3 Inter-O-DU Coordination and Scheduler-Level Energy Profiling: Problem, Limitations, and Prior Approaches	25
2.2.4 Research Gap and Link to RQ1	26
2.3 Challenge 2 — Physical-Layer Realism: FBL-Aware Multi-Cell NOMA opti- mization	27
2.3.1 NOMA in Multi-Cell Dense Deployments; Cooperative and Non- Cooperative Interference	27
2.3.2 Joint EE-Connectivity optimization Under IFBL (Shannon)	28
2.3.3 FBL Theory, Short-Packet Transmission, and EE Impact	28
2.3.4 Joint EE-Connectivity optimization Under FBL (Short-Packet Regime)	29
2.3.5 Research Gap and Link to RQ2	31
2.4 Challenge 3 — Adaptability: Multi-Objective Reinforcement Learning for 6G Scheduling in Centralized and Distributed O-RAN	31
2.4.1 RL Taxonomy: On-Policy, Off-Policy, Model-Based, and Model-Free Learning.....	32
2.4.2 Single-Agent vs. Multi-Agent Architectures	33
2.4.3 MORL: Solution Concepts and State of the Art	34
2.4.4 Research Gap and Link to RQ3	37
2.5 Cross-Cutting Themes: Sustainability and O-RAN Alignment	37
2.6 Consolidated Gap Analysis and Mapping to Contributions	38
2.7 Chapter Summary	39
3 System Model	40
3.1 Overview	40
3.2 Unified Notation, Symbols, and Constraints	40
3.3 Uplink System Model	42
3.4 Non-Cooperative Multi-Cell Uplink Model	43
3.5 Cooperative Multi-Cell Uplink Model	44
3.6 Clustered Cooperative Multi-Cell Model for Learning-Based Control	45
3.7 Constraints and Underlying Assumptions	45
3.8 Underlying assumptions	47
3.9 Model Evolution.....	48
3.10 Constraints Reference Summary	49

4	Coordination and Scheduling in O-RAN Architecture	50
4.1	Chapter Overview and Motivation	50
4.2	Related Work and Positioning	50
4.3	Methodology: Inter-O-DU Coordination (Publication (I))	51
4.4	Methodology: Computational Energy Modeling (Publication (II))	52
4.5	Results and Discussion	52
4.6	Chapter Summary	55
5	Sustainable Connectivity Optimization under 6G Constraints	57
5.1	Overview, Motivation, and Chapter Scope	57
5.2	Related Work and Positioning	58
5.3	System Perspective and Analytical Problem Formulation	58
5.3.1	Non-Cooperative Fixed-ICI Model	59
5.3.2	Publication III Problem Formulation	59
5.3.3	Cooperative FBL-Aware Model	60
5.3.4	Publication IV Problem Formulation	60
5.3.5	Relationship Between the Two Formulations	61
5.4	Algorithm Design	61
5.4.1	EERS-MTCN	61
5.4.2	SRS-MTCN	62
5.4.3	Complexity Comparison	62
5.5	Results and Discussion	63
5.5.1	EERS-MTCN: EE, Connectivity, and Execution Time	63
5.5.2	SRS-MTCN: Interference, EE, Connectivity, and FBL Sensitivity	66
5.5.3	EE trends under cooperative FBL-aware scheduling	66
5.5.4	Discussion	70
5.6	Chapter Summary and Bridge to Chapter 6	70
6	Reinforcement Learning for Multi-Objective Optimization	72
6.1	Overview	72
6.2	Related Work	73
6.3	Positioning of the Contribution: Gap-to-Mechanism Mapping	75
6.4	Methodology of the Contributions	76
6.5	Experimental Setup and Baselines	80
6.6	Results and Discussion	83
6.7	Chapter Summary and Bridge to Chapter (7)	92
7	Conclusion and Future Work	93
7.1	Research Questions, Contributions, and Answers	93
7.2	Broader Implications for 6G and O-RAN	95
7.3	Limitations and Future Work	95
7.3.1	L1: Simplified O-DU coordination model and scalability.	95
7.3.2	L2: FBL model assumptions and coding complexity.	95
7.3.3	L3: Static user association and mobility.	96
7.3.4	L4: Hardware-in-the-loop validation of WiMO-CEPPO.	96
7.3.5	L5: Distributed and transferable learning.	96
7.3.6	L6: Network topology and heterogeneous access.	96
7.4	Concluding Remarks	96
	List of Figures	99
	List of Tables	100
	References	101

Acknowledgments	111
Abstract.....	112
Kokkuvõte	114
Appendix 1.....	117
Appendix 2	125
Appendix 3	133
Appendix 4	141
Appendix 5	159
Appendix 6	167
Curriculum Vitae	176
Elulookirjeldus.....	180

List of Publications

This PhD thesis is based on the following 6 publications that are referred to in the text by Roman numbers.

- I Latif, Aamir; Alam, Muhammad Mahtab; Le Moullec, Yannick, "Inter O-DUs Coordination for Scheduling of Massive Users in O-RAN," in *2023 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit 2023)*, pp. 1–6, 2023.
- II Latif, Aamir; Elgarhy, Osama; Le Moullec, Yannick; Alam, Muhammad Mahtab, "Energy Consumption Evaluation of NOMA-based Sustainable Scheduling in 6G O-RAN," in *2024 International Wireless Communications and Mobile Computing (IWCMC 2024)*, pp. 484–489, 2024.
- III Latif, Aamir; Elgarhy, Osama; Le Moullec, Yannick; Alam, Muhammad Mahtab, "Energy Efficiency and Connectivity Maximization for Massive Machine Type Communication in Multi-Cell NOMA," in *2024 19th Biennial Baltic Electronics Conference (BEC 2024)*, pp. 1–6, IEEE, 2024.
- IV Latif, Aamir and Elgarhy, Osama and Mahmood, Nurul Huda and Le Moullec, Yannick and Alam, Muhammad Mahtab, "Sustainable Massive Connectivity Maximization Scheduling Scheme for 6G," *IEEE Transactions on Green Communications and Networking*, pp. 2075–2090, January 2026.
- V A. Latif, Y. Le Moullec, and M. M. Alam, "Multi-objective policy design using reinforcement learning in 6g networks," in *2026 International Wireless Communications and Mobile Computing (IWCMC)*, pp. 1352–1357, 2026.
- VI A. Latif, Y. L. Moullec, and M. M. Alam, "Sustainable reward mechanism for multi-objective policy design using reinforcement learning in 6g networks," in *2026 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, pp. 527–533, 2026.

Author's Contributions to the Publications

- I **Publication I:** I served as the lead author of this article. I conceptualized and implemented the proposed inter-O-DU coordination architecture for Open Radio Access Network (O-RAN) deployments. I developed the mathematical formulation of the interference-minimization problem and designed the cooperative Non-Orthogonal Multiple Access (NOMA)-based scheduling algorithm. I was primarily responsible for implementing the simulation framework, conducting the performance evaluation, and interpreting the results in terms of throughput, energy consumption, and user connectivity. I also prepared the initial manuscript draft and incorporated feedback from co-authors to improve its technical depth and clarity.
- II **Publication II:** I served as the lead author of this article. I formalized the sustainable scheduling framework for energy-efficient O-RAN operation and developed analytical models for comparing Orthogonal Multiple Access (OMA) and NOMA techniques. I designed the EERS-MTCN algorithm and performed complexity and energy modeling using instruction-level power analysis on the RISC-V architecture. I also conducted detailed benchmarking against state-of-the-art schedulers and interpreted the results to highlight energy savings across different user densities. In addition, I structured the manuscript, integrated co-author feedback, and ensured alignment with 6G and O-RAN sustainability objectives.
- III **Publication III:** I served as the lead author of this article. I proposed and developed an energy-efficient resource-scheduling and power-allocation framework for ultra-massive machine-type communications (umMTC) in 6G NOMA networks under a non-cooperative multi-cell interference model. In this formulation, neighbouring cells were modeled as sources of constant maximum interference, and the achievable rate was characterized using the Shannon infinite-block-length assumption. I formulated the problem as a constrained multi-objective optimization task that jointly maximizes energy efficiency (EE) and connectivity under practical system constraints. I designed the EERS-MTCN algorithm, which decouples scheduling and power control to improve tractability and scalability, and led the simulation setup using a standardized 3GPP channel model. I evaluated the proposed method against benchmark algorithms, including RAAP-NOMA and ESAP-NOMA, and demonstrated up to a 42% improvement in EE. I also wrote the manuscript and coordinated its structure, figures, and technical validation.
- IV **Publication IV:** I served as the lead author and conceptualized a framework for sustainable massive connectivity maximization in ultra-dense 6G environments under a cooperative multi-cell interference model. Unlike the non-cooperative formulation in Publication III, this work incorporated cooperative interference awareness and finite-block-length (FBL) transmission to capture short-packet reliability constraints in mMTC scenarios. I designed an interference-aware user-pairing strategy integrated with FBL-aware EE-connectivity trade-off optimization. I developed and implemented the proposed SRS-MTCN algorithm, incorporated variable NOMA configurations with

2, 4, and 8 users per resource block (RB), and validated its convergence behavior across reliability and block-length settings. I led the end-to-end simulation design, conducted comparative evaluations against benchmark schemes, and demonstrated up to a 58% EE gain under realistic mMTC traffic loads. I also drafted and revised the manuscript to ensure scientific rigor and clarity.

V Publication V: I served as the lead author and designer of a single-cell, single-agent MORL framework for sustainable scheduling in ultra-dense 6G mMTC networks. This study included a foundational step by isolating the core joint scheduling and power-allocation behavior under a simplified interference abstraction. I formulated the Multi-Objective Reinforcement Learning (MORL) framework and designed a reward function incorporating EE, connectivity, reliability, and cross-link interference. I developed the complete simulation pipeline, implemented the PPO-based preference-conditioned optimization framework, benchmarked it against representative MORL baselines, and conducted comprehensive Pareto analyses. I authored the manuscript, prepared all figures and experimental results, and worked closely with co-authors to ensure methodological rigor and alignment with sustainability objectives.

VI Publication VI: I extended the single-cell formulation to a multi-cell cluster setting aligned with centralized RAN and O-RAN architectures. I led the adaptation of the WiMO-CEPPO framework to jointly address EE, connectivity, reliability, and interference under inter-cell coupling, while explicitly incorporating hardware and computational efficiency as deployment constraints. This extension addressed scalability and real-time feasibility challenges, particularly in distributed RAN scenarios where inference latency is a critical bottleneck. I implemented the enhanced framework, performed comparative evaluations against state-of-the-art MORL baselines, conducted Pareto analysis using HV and IGD metrics and ablation analysis, and prepared the manuscript and all experimental artifacts. This study positions WiMO-CEPPO as a deployable multi-objective scheduling framework for realistic 6G operational constraints.

Abbreviations

3GPP	3rd generation partnership project.
5G	Fifth Generation.
6G	Sixth Generation.
AI	Artificial Intelligence.
AP	Access point.
AVs	Autonomous vehicles.
B5G	Beyond 5th generation.
CCI	Co-channel interference.
COTS	Commercial Off-The-Shelf
CoMP	Coordinated multi-point.
CPM	Context Preference Mapper.
CQI	Channel quality indicator.
CS	Channel Statistics
DDPG	Deep Deterministic Policy Gradient.
EE	Energy Efficiency.
EERS-MTCN	Energy Efficient Resource Scheduling for Massive Machine Type Communication NOMA.
ESAP-NOMA	Exhaustive Search and Pairing NOMA.
eDRX	Extended discontinuous reception.
eMBB	Enhanced mobile broadband.
FBL	Finite block-length.
FDR	Full duplex relaying
HDR	Half duplex relaying.
HV	Hyper Volume.
ICI	Inter-cell interference.
IFBL	Infinite block-length.
IGD	Inverted generational distance.
IoT	Internet of Things.
KKT	Karush–Kuhn Tucker.
KPIs	Key Performance Indicators.
LTE-M	Long-term evolution MTC.
LoS	Line-of-Sight
LS	Least Squares
LSTM	Long Short-Term Memory
M2M	Machine-to-machine.
MAE	Mean Absolute Error
ML	Machine Learning.
MORL	Multi-objective reinforcement learning.
MSE	Mean Squared Error.
mMTC	Massive machine-type communications.
NB-IoT	Narrow band Internet of Things.

NLoS	Non-Line of Sight.
NN	Neural Networks.
NOMA	Non-orthogonal multiple access.
NR	New radio.
OMA	Orthogonal multiple access.
O-RAN	Open Radio Access Network.
PD	Power domain.
PPO	Proximal Policy Optimization.
QoS	Quality of service.
RAN	Radio Access Network.
RAAP-NOMA	Random Access and Pairing NOMA.
RB	Resource block.
RF	Radio Frequency.
RL	Reinforcement Learning.
RMSE	Root Mean Squared Error.
RNN	Recurrent Neural Network.
RQ	Research Question.
RRM	Radio Resource Management.
RSSI	Received Signal Strength Indicator.
SAC	Soft Actor Critic.
SC	Superposition coding.
SIC	Successive interference cancellation.
SRS-MTCN	Sustainable Resource Scheduling for Massive Machine Type Communication Networks.
TD3	Twin Delayed Deep Deterministic.
UEs	User equipments.

1 Introduction

1.1 Background and Motivation

Sixth-generation (6G) wireless networks, targeted for commercial deployment around 2030, are expected to support device densities of up to 10^8 devices/km², driven primarily by massive Machine-Type Communication (mMTC) and its ultra-massive evolution (umMTC) [7, 8]. Application domains include smart manufacturing, industrial IoT, autonomous transportation, precision agriculture, and immersive healthcare, all of which impose heterogeneous, often conflicting demands on the network infrastructure [9]. In response, international bodies (ITU-R and 3GPP) have defined expanded capability requirements for IMT-2030 that explicitly encompass both performance and sustainability, i.e., not merely *how many* devices can be connected, but *how efficiently and reliably* each connection can be sustained under stringent energy and interference constraints [10, 11, 12, 13].

Fig. 1 summarizes the main design challenges in 6G and illustrates the mMTC ecosystem with highly dense connectivity among heterogeneous devices and services. Within such environments, ensuring reliable, energy-efficient, and secure connectivity demands architectural solutions that can operate under stringent resource constraints while adapting to highly variable traffic and interference conditions[14].

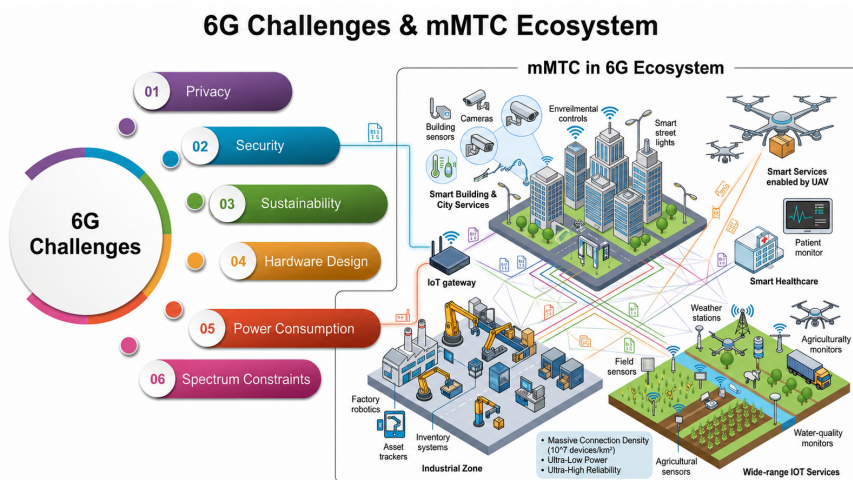


Figure 1: Overview of major 6G design challenges and the mMTC ecosystem. Key challenges that shape next-generation communication systems include sustainability constraints, advanced hardware design requirements, power-efficiency considerations, and increasing pressure on the radio spectrum. These issues intersect with essential concerns related to privacy and security, particularly as networks evolve to support dense, heterogeneous deployments of IoT devices (mMTC) and industrial, healthcare, agricultural, and UAV-enabled services.

This combination of scale, reliability, and energy efficiency makes the 6G mMTC problem fundamentally different from challenges found in earlier network generations. From 1G to 4G, cellular systems were designed to optimize essentially one performance dimension at a time, either voice quality, data rate, or spectral efficiency [15]. Fifth-generation (5G) NR took a step toward multi-dimensional design by defining three service classes

(eMBB, URLLC, and mMTC), but the mMTC implementation in 5G (realized through NB-IoT and LTE-M) remains constrained by signaling overhead, uncoordinated interference, and the absence of a joint framework for optimizing energy efficiency and connectivity simultaneously. The transition to 6G therefore requires a qualitatively different approach to radio resource management (RRM), i.e., one that treats EE and massive connectivity not as separate objectives to be traded off in a fixed way, but as jointly optimizable targets within a single, coherent framework [16].

Achieving this goal is not trivial because the objectives conflict. Serving more devices simultaneously requires more aggressive spectrum reuse, which amplifies inter-device interference and drives up power consumption. Increasing transmit power to improve coverage and reliability, especially at the cell edge, reduces EE. Allocating resources conservatively to save energy limits the number of devices that can be served. These issues cannot be resolved solely with better hardware; they are *fundamental trade-offs* embedded in the physics of wireless communication and become more pronounced at the high levels of device density envisioned for 6G [10, 12, 13]. Navigating these trade-offs efficiently is the defining challenge of 6G mMTC RRM [17]. RRM for massive connectivity must be designed explicitly as a *multi-objective* optimization problem that balances competing goals such as energy efficiency, connectivity, reliability, and cross-link interference.

Understanding why existing approaches fall short at handling the above requires looking at three specific challenges that arise when attempting to perform multi-objective optimization for RRM in a realistic 6G deployment.

The first challenge, i.e., interference, is the key physical-layer obstacle that makes this multi-objective problem especially challenging in dense deployments. In a multi-cell uplink, decisions made in one cell directly affect the interference experienced by neighboring cells; therefore, the performance of one cell cannot be optimized in isolation. This creates a coordination dilemma. In a *non-cooperative* setting, each cell schedules independently using only local information. This is simple and scalable, but as density grows, it can lead to *decision-dependent inter-cell interference*, i.e., neighboring schedulers unintentionally reuse the same resources and raise each other's interference floor, negatively affecting both connectivity and energy efficiency. On the other hand, in a *cooperative* (coordination-enabled) setting, cells exchange limited information (e.g., load, scheduling information, or inter-cell interference level) to reduce the impact of uncoordinated decisions, but this introduces real deployment constraints, i.e., signaling overhead, latency, and limited observability, especially in decentralized architectures such as O-RAN. The core difficulty is that coordination can improve performance, yet it is never free; any realistic solution must justify coordination by showing that the gains outweigh its overhead and remain feasible under near-real-time constraints [18, 19, 20].

The second challenge arises from the traffic characteristics of 6G mMTC systems, where the vast majority of devices transmit short data bursts—such as sensor readings, control commands, or event notifications—typically consisting of only tens to hundreds of bytes. Under such conditions, the classical Shannon capacity formula, which assumes asymptotically long codewords, significantly overestimates the achievable transmission rate. Finite block-length (FBL) information theory, formalized by Polyanskiy [21], establishes that

short-packet transmissions incur a non-negligible rate penalty proportional to the channel dispersion and inherently operate with non-zero decoding error probability. Consequently, EE and connectivity formulations that neglect FBL effects may yield allocations that appear optimal under infinite-block-length assumptions but violate practical reliability constraints in realistic mMTC deployments. This leads to systematic mismatches between predicted and realized reliability, as well as inflated energy consumption due to retransmissions or conservative power compensation. Incorporating FBL constraints is not a minor refinement but a prerequisite for ensuring physical-layer consistency and deployment credibility in ultra-dense 6G mMTC networks.

The third challenge concerns learning-based methods, especially RL, which offer a promising path because they can learn scheduling policies directly from interaction with dynamic environments, rather than relying on simplified models or repeatedly solving optimization problems. However, learning does not remove the multi-objective difficulty; it actually amplifies it. Real networks require policies that balance several conflicting objectives at once, and naively combining them into a single reward can hide trade-offs or produce behavior that only works for one operating condition. Multi-objective reinforcement learning (MORL) addresses this by treating performance as a vector of objectives and learning policies that can navigate trade-offs. Yet MORL introduces its own challenges, i.e., training can become unstable when objectives conflict, policies may fail to generalize across different operating preferences, and practical deployment constraints (limited coordination, control-loop latency, and execution feasibility) must still be respected [22, 23]. In other words, learning-based methods are not a shortcut; they are a powerful tool that must be designed carefully to produce stable, controllable, and deployable scheduling behavior.

These three challenges are not independent; rather, they form a sequential continuum in which each must be resolved before the next can be meaningfully addressed. Coordination constitutes the first step: without mechanisms to manage decision-dependent inter-cell interference, the uplink environment remains uncontrolled, and no optimization framework can reliably achieve predictable EE-connectivity trade-offs. Physical-layer realism follows as the second challenge: once interference is coordinated, the optimization framework must accurately reflect what the physical layer can deliver under short-packet, finite-block-length conditions, since classical Shannon-based models systematically overestimate achievable rates and reliability in 6G mMTC traffic. Adaptability emerges as the third challenge: after the analytical foundations are established, the framework must be extended to operate under non-stationarity, conflicting key performance indicators (KPIs), and real-time constraints in practical deployments. This ordered progression defines the structure of the research problem addressed in this thesis.

Despite the importance of this research problem, to the best of my knowledge, no prior work has jointly addressed all three challenges within a unified framework. As detailed in Section 2, existing studies typically focus on only one dimension—either coordination in isolation, or energy-efficiency optimization without finite-block-length constraints, or learning-based scheduling without grounding in analytical multi-objective trade-off characterization. The absence of a coherent, progressive framework that integrates coordination, analytical multi-objective optimization under realistic physical constraints, and policy-based

multi-objective scheduling represents the central research gap that this thesis aims to fill.

1.2 Central Research Problem

The three challenges identified above converge on a single, unified statement of the research problem this thesis seeks to address:

How can RRM in 6G mMTC networks be designed to jointly optimize energy efficiency and device connectivity, while operating under the realistic constraints of decentralized RAN architectures, finite-block-length short-packet transmission, and dynamically evolving network conditions?

The central research problem addressed in this work is inherently multi-faceted rather than a single isolated technical challenge. It encompasses three tightly coupled dimensions: (i) a *coordination challenge*, which involves mitigating inter-cell interference (ICI) in distributed O-RAN architectures; (ii) a *physical-layer realism challenge*, which requires the joint optimization of EE and connectivity under FBL constraints that fundamentally alter achievable rates and reliability; and (iii) an *adaptation challenge*, which focuses on designing a scheduling policy capable of responding to non-stationary network conditions while balancing multiple conflicting KPIs.

To systematically address these intertwined challenges, the thesis decomposes the overarching problem into three targeted research questions, each corresponding to a distinct layer of the challenge hierarchy as illustrated in Fig. 2.

1.2.1 Research questions

The thesis is structured around three research questions outlined here:

RQ1 –(Coordination Layer): *How can coordination mechanisms among distributed O-DUs in O-RAN architectures be designed to mitigate inter-cell interference (ICI) and enable scalable, energy-efficient scheduling for massive device connectivity?*

RQ2 – (Physical-Layer Realism): *How can multi-cell NOMA-based resource scheduling be optimized to jointly balance energy efficiency and connectivity under realistic constraints such as FBL transmission, dense interference coupling, and heterogeneous device capabilities?*

RQ3 – (Adaptability): *How can MORL be employed to learn a single, adaptive scheduling policy that generalizes across preference trade-offs and dynamically responds to time-varying network conditions in 6G systems?*

These research questions collectively follow a progressive design hierarchy, in which each layer addresses a limitation of the previous one: coordination mitigates interference, analytical optimization characterizes trade-offs, and MORL enables real-time adaptive decision-making. The three research questions are also aligned with the thesis map (Fig. 2, see below), as well as with the system-model ladder in Chapter (3).

Fig. 2 provides a structured overview of the thesis, illustrating how the central problem of multi-objective RRM in 6G is decomposed into three sequential and interdependent



Figure 2: PhD thesis master map illustrating the unified research framework, linking the central problem of multi-objective RRM for 6G massive connectivity with three sequential and interdependent challenges: coordination, physical-layer optimization, and adaptability. The figure shows how these challenges are decomposed into the corresponding research questions and addressed through the methodological approaches developed in this thesis. It also maps each challenge to the technical Chapters (4–6) and to the key contributions (I–VI).

challenges: coordination, physical-layer optimization, and adaptability. The figure follows a top-down organization, starting from the overarching goal of achieving scalable, energy-efficient, and reliable massive connectivity, and progressively narrowing to the core scheduling and power-allocation problem addressed in this work.

The foundational components of the thesis are established in Chapters (1–3), which define the problem context, review related work, and formalize the system model and assumptions. Building on this, Chapter (4) addresses **RQ1** (Coordination), focusing on mitigating inter-cell interference through inter-O-DU coordination in O-RAN environments. Chapter (5) addresses **RQ2** (Physical-layer realism), formulating and solving the joint EE-connectivity trade-off under realistic constraints such as interference coupling and finite block-length transmission. Chapter (6) addresses **RQ3** (Adaptability), where MORL is employed to design adaptive scheduling policies capable of operating under dynamic network conditions.

The contribution blocks associated with each chapter summarize the key research outputs and their alignment with the corresponding research questions. The directional flow in the figure highlights the progressive nature of the thesis, where each layer addresses a specific limitation of the previous one, collectively forming a unified framework for sustainable and adaptive 6G resource management.

1.3 Research Methodology

This thesis follows a *design-science* research methodology [24], centered on reproducible, simulation-based experimentation. Each contribution is developed through a *construct-validate* cycle: (i) specification of a formal system model (Chapter (3)) and problem formulation, (ii) design of an algorithm or policy grounded in this formulation, and (iii) validation through controlled simulations against well-defined baselines. This cycle is executed across three methodological layers—*coordination*, *analytical optimization*, and *reinforcement learning*—where each stage builds on the insights and limitations of the preceding one. Simulation serves as the primary experimental tool for performance evaluation.

The selection of methods is aligned with the research questions.

For **RQ1**, an inter-O-DU coordination mechanism is designed and evaluated against a non-cooperative baseline in which each O-DU schedules independently. Deployability is treated as a constraint; the coordination design assumes no joint transmission and no centralized channel knowledge, consistent with distributed O-RAN operation.

For **RQ2**, the problem is formulated as a non-convex mixed-integer multi-objective optimization of energy efficiency and connectivity. It is addressed through decomposition into tractable subproblems, solved using tailored heuristics and iterative updates. This analytical stage provides interpretability and explicit characterization of trade-offs (Pareto behavior) under controlled conditions.

For **RQ3**, the scheduling problem is reformulated as a multi-objective Markov decision process (MO-MDP) conditioned on a preference vector, and solved using MORL. MORL is employed because analytical solutions cannot be recomputed at near-real-time scheduling timescales; instead, it learns a policy that adapts to dynamic network conditions via a single inference step while preserving objective trade-offs.

All experiments are conducted in a unified simulation environment aligned with 3GPP

TR 38.913 [25] and ITU-R M.2412 [26]. The setup uses a seven-cell topology with 500 m inter-site distance and a 3GPP Urban Macro (UMa) channel model. Uplink transmit power is limited to 23 dBm per device. Device densities range from 16 to 400 active devices per cell, covering both mMTC and umMTC regimes. For short-packet scenarios (**RQ2**), finite-block-length performance is modeled using the normal approximation of Polyanskiy [21], with block-lengths $L \in [50, 600]$ and decoding error probabilities $\varepsilon \in [10^{-3}, 10^{-6}]$. Key parameters are held constant across all stages to ensure fair comparison.

Evaluation is performed using metrics aligned with each research question. The primary *wireless key performance indicators (KPIs)* include:

- **Energy Efficiency (EE)**: measured in bits/Joule,
- **Connectivity**: number of simultaneously connected devices,
- **Reliability**: probability of successful decoding (or SINR satisfaction),
- **Cross-Link Interference (CLI)**: residual inter-cell interference,
- **Quality-of-Service (QoS)**: rate or SINR satisfaction levels.

For **RQ3**, Pareto-quality metrics are additionally reported. The *hypervolume (HV)* measures the dominated objective-space volume, while the *inverted generational distance (IGD)* quantifies proximity to the reference Pareto front. All results are averaged over multiple independent random seeds to ensure statistical robustness. The scope of this thesis is simulation-based evaluation. Practical deployment aspects—including hardware-in-the-loop validation, inference latency, memory footprint, and real-time feasibility on O-RAN-compliant platforms—are identified as important directions for future work.

Table 1 maps the six publications to the research questions, methods, and thesis chapters. The respective summaries of the publications are provided in Section 1.4.

Table 1: Mapping of the research questions (RQs), methods, and thesis chapters

Publication	RQ	Method	Chapter
1	RQ1	Inter-O-DU coordination	Chapter (4)
2	RQ1	Instruction-level energy framework	Chapter (4)
3	RQ2	EERS-MTCN (Joint EE-conn., Shannon IFBL)	Chapter (5)
4	RQ2	SRS-MTCN (Coop., FBL-aware)	Chapter (5)
5	RQ3	WiMO-CEPPO (4-obj MORL, CPM, Single-cell to multi-cell C-RAN)	Chapter (6)
6	RQ3	WiMO-CEPPO (4-obj MORL, extended to multi-cell O-RAN)	Chapter (6)

1.4 Summary of Thesis Contributions

Given the research questions presented in Section 1.2.1 and their mapping to the thesis chapters in Table 1, the following summarizes the thesis contributions, one paper at a time.

- **[Publication (I)] Inter O-DUs Coordination for Scheduling of Massive Users in O-RAN:** The first contribution develops an inter-O-DU coordination framework for mMTC in 6G O-RAN systems. Unlike most prior studies, which treat each distributed unit (DU) independently, this framework uses open interfaces, and near-RT RIC functionality to exchange interference, traffic, and scheduling information across DUs. As a result, neighboring DUs can jointly adapt user pairing and resource allocation to improve throughput and EE in ultra-dense deployments.
- **[Publication (II)] Energy Consumption Evaluation of NOMA-based Sustainable Scheduling in 6G O-RAN:** The second contribution proposes an instruction-level energy profiling methodology for scheduling operations in O-RAN environments. Rather than evaluating energy consumption only at the network or cell level, this contribution uses the RISC-V instruction set architecture to quantify the execution cost of scheduler functions on the CU/DU platform. This approach enables direct and transparent comparison of scheduling strategies in terms of joules per bit.
- **[Publication (III)] Energy Efficiency and Connectivity Maximization for Massive Machine Type Communication in Multi-Cell NOMA:** The third contribution develops a joint optimization framework that maximizes EE and connectivity in multi-cell non-cooperative NOMA systems under practical constraints. The contribution unifies EE and connectivity in a single formulation and proposes the Energy-Efficient Resource Scheduling for Machine-Type Communication Networks (EERS-MTCN) algorithm to optimize both objectives simultaneously. Simulation results demonstrate substantial gains.
- **[Publication (IV)] Sustainable Massive Connectivity Maximization Scheduling Scheme for 6G:** The fourth contribution advances cooperative multi-cell NOMA scheduling under FBL conditions. It enables base stations to coordinate power allocation and user pairing across cells to minimize inter-cell interference while jointly satisfying connectivity and EE objectives. The evaluation demonstrates that this cooperative approach outperforms non-cooperative designs in both EE and scalability.
- **[Publication (V)] Multi-Objective Policy Design using Reinforcement Learning in 6G Networks:** The fifth contribution establishes *WiMO-CEPPO* in a *single-cell, single-agent* setting as a foundational step for sustainable 6G scheduling. It learns a *single preference-conditioned policy* using an *envelope-based update* to balance EE, connectivity, and reliability under simplified interference. The approach achieves Pareto-dominant performance over scalarized baselines. However, it assumes centralized training and does not fully address real-time or distributed deployment.
- **[Publication (VI)] Sustainable Reward Mechanisms for Multi-Objective Policy Design using Reinforcement Learning in 6G Networks:** The sixth contribution extends *WiMO-CEPPO* to a *multi-cell decentralized RAN* setting with *O-RAN deployment considerations*. It incorporates *hardware and computational efficiency constraints* to address latency-sensitive scheduling in distributed environments. The framework

maintains stable multi-objective learning and broad Pareto coverage, achieving superior performance over baselines. Nevertheless, training remains centralized with limited support for fully distributed, real-time adaptation.

Although the first contribution I establishes the foundation by addressing the coordination challenge, specifically, how inter-cell interference can be mitigated through inter-O-DU coordination and cooperative scheduling, the research trajectory does not unfold in a strictly linear manner. This is because the overarching problem of multi-objective radio resource optimization in 6G is inherently multi-dimensional: coordination, physical-layer realism, and adaptability are deeply interdependent and cannot be resolved through a single methodological lens. The initial coordination results do not, by themselves, resolve the joint energy-efficiency-connectivity trade-off under realistic constraints such as finite block-length, heterogeneous device capabilities, and dense interference coupling. This necessitates additional contributions that examine complementary aspects of the problem under different modeling assumptions—first through analytical multi-objective optimization (Contributions III–IV), and later through adaptive learning-based policy design (Contributions V–VI).

In this sense, the thesis follows a design-science research paradigm, where problem understanding, artifact construction, and iterative evaluation evolve through repeated construct-evaluate cycles rather than a one-directional sequence [24, 27]. The non-linear progression of contributions is therefore methodologically justified: each contribution isolates and resolves one layer of the overall research problem—coordination, physical-layer realism, and adaptability—before these layers are integrated into a unified framework for sustainable, scalable, and adaptive 6G radio resource management.

1.5 Thesis Organization

The main body of this thesis comprises seven chapters, i.e., Chapters (1–3) establish the conceptual and methodological foundations, while Chapters (4–6) present the core technical contributions. These are followed by six appendices (Appendix 1– Appendix 6), which contain the Publications [1]–[6] on which the thesis is based.

- **Chapter (1)** (this chapter) introduces the motivation, overarching challenges, and research scope for sustainable user scheduling in future mobile networks. It formulates the central research problem and states the three research questions that guide the thesis.
- **Chapter (2)** surveys and synthesizes the related scientific literature on sustainable user scheduling in 6G, with an emphasis on massive connectivity, interference-aware resource allocation, and emerging O-RAN control architectures. This chapter identifies the research gaps that motivate the contributions developed in the remainder of the thesis.
- **Chapter (3)** is included because the publications in the appendices were developed at different stages of the research journey and therefore employ slightly varying

system models and notational conventions. To ensure coherence across the main body of the thesis, this chapter consolidates the shared modeling elements into a unified system model and notation framework. Subsequent Chapters (4–6) refer back to this unified model to avoid redundancy and maintain a consistent and self-contained presentation of the technical contributions.

- **Chapter (4)** is based on Publications [1], and [2] which addresses the *coordination* challenge. It investigates inter-O-DU coordination and sustainable scheduler design within the CU-RIC architecture, proposing mechanisms to minimize interference and energy-aware scheduling in O-RAN deployments. The chapter evaluates how coordination impacts scalability, interference suppression, and power efficiency.
- **Chapter (5)** is supported by Publications [3], and [4] tackles the *physical-layer realism* challenge. It first advances the analysis using a non-cooperative multi-cell uplink NOMA model to jointly maximize energy efficiency and connectivity under dense interference. It then transitions to a cooperative framework that incorporates finite-block-length reliability constraints and introduces the SRS-MTCN algorithm for joint user grouping, interference mitigation, and power control while ensuring reliability under short-packet transmission.
- **Chapter (6)** investigates the *adaptation* challenge drawing on Publications [5], and [6]. It first introduces a single-cell, single-agent reinforcement learning framework for multi-objective policy design, demonstrating how a preference-conditioned agent can navigate conflicting KPIs and generalize toward centralized multi-cell operation. It then extends these principles to an O-RAN-compatible xApp that incorporates hardware sustainability and computational-efficiency constraints, illustrating the feasibility of deploying the proposed learning-based scheduler in practical near-real-time environments.
- Finally, **Chapter (7)** synthesizes the contributions across all chapters, summarizes the main findings, and outlines future research directions. These include deeper integration with real-time O-RAN interfaces, transfer learning for faster adaptation, and hardware-in-the-loop evaluation to enhance deployment realism.

2 State of the Art and Related Works

2.1 Overview and Scope

This chapter provides a systematic survey of the literature underpinning the three research questions introduced in Chapter (1). Its purpose is threefold: (i) to situate this thesis within the broader scientific discourse on radio resource management (RRM) for 6G massive machine-type communication (mMTC) networks; (ii) to identify, through rigorous comparative analysis, the specific limitations of existing work that directly motivate each research question; and (iii) to establish the conceptual and technical context within which the proposed contributions operate.

Rather than following a conventional technology-by-technology survey, the chapter is organized around the three challenges defined in the thesis introduction: *coordination*, *physical-layer realism*, and *adaptability*. This structure reflects the logical dependency among the research questions. Inter-cell coordination must first be addressed before any scheduling framework can manage decision-dependent interference in distributed O-RAN deployments. Once coordination is available, the physical-layer model must accurately capture the short-packet transmission conditions of mMTC, particularly finite-block-length (FBL) effects. Finally, because optimization-based schedulers cannot be repeatedly solved at near-real-time control cadences, adaptive learning-based policies are needed to operate under dynamic network conditions and changing performance preferences.

Section 2.2 reviews the state of the art on O-RAN scheduling architectures, inter-distributed-unit coordination, and scheduler-level energy profiling. Section 2.3 surveys FBL information theory and multi-cell NOMA resource allocation. Section 2.4 reviews multi-objective reinforcement learning (MORL) for wireless scheduling. Section 2.5 draws out two cross-cutting themes that connect all three challenges. Section 2.6 consolidates the gap analysis in a single table that maps each gap to the corresponding thesis contribution. Section 2.7 concludes the chapter and provides the narrative bridge to the technical chapters.

2.2 Challenge 1 — Coordination: O-RAN Scheduling and Energy-Aware Design

The first challenge concerns the management of decision-dependent inter-cell interference (ICI) in distributed O-RAN architectures. In dense multi-cell networks, interference is not merely a propagation effect; it is also a consequence of simultaneous scheduling and power-control decisions made by neighbouring cells. If O-DUs schedule users independently, the interference field becomes unstable, difficult to predict, and difficult to optimize. Consequently, no resource-allocation or learning framework can deliver reliable energy-efficiency (EE)-connectivity trade-offs unless the coordination problem is addressed at the architectural level.

This section reviews three tightly connected aspects of the coordination challenge: the O-RAN functional architecture and control hierarchy, the computational energy cost of software-defined scheduling in virtualized RANs, and the specific limitations of existing inter-O-DU coordination approaches.

2.2.1 O-RAN Functional Architecture and Control Hierarchy

The O-RAN Alliance has introduced a disaggregated, software-defined RAN architecture that decomposes the traditional monolithic base station into interoperable logical functions connected through open interfaces [28, 29]. The O-RAN Central Unit (O-CU) hosts higher-layer protocol functions, including Packet Data Convergence Protocol (PDCP) and Radio Resource Control (RRC). The O-RAN Distributed Unit (O-DU) implements time-sensitive MAC, RLC, and upper-PHY functions, while the O-RAN Radio Unit (O-RU) performs low-PHY and radio-frequency processing [30, 31]. The F1 interface connects the O-CU and O-DU, while fronthaul interfaces such as eCPRI connect the O-DU and O-RU.

O-RAN further introduces a two-tier intelligence layer. The Non-Real-Time RIC operates at timescales above one second and is responsible for policy management, model life cycle management, and long-term optimization through the A1 interface [32, 33, 34]. The Near-Real-Time RIC operates at approximately 10–100 ms control timescales and hosts xApps that implement near-real-time RRM functions by subscribing to E2 telemetry and issuing control decisions to O-RAN nodes [35, 36]. This hierarchy creates a programmable RAN control plane in which scheduling, interference management, slicing, and traffic steering can be exposed to externalized intelligence.

Polèse et al. [37] provide a comprehensive treatment of O-RAN interfaces, algorithms, and research challenges, emphasizing that the Near-RT RIC is the primary architectural anchor for adaptive near-real-time control. Alam et al. [38] extend this perspective to slicing-aware O-RAN deployments and show that AI-enabled resource allocation is most naturally positioned at the Near-RT RIC. Balasubramanian et al. [39] demonstrated the feasibility of RIC-based intelligent control, providing early evidence that learning-based schedulers can be implemented within RAN control-loop constraints.

However, the same architecture also introduces limitations that are central to this thesis. The E2 control loop is not designed for sub-millisecond per-slot global scheduling. Telemetry can be delayed or aggregated, and control decisions must satisfy strict latency constraints. The O-RAN has therefore investigated dApps for real-time RAN control as a complementary mechanism to xApps in latency-sensitive scenarios [18]. In addition, xApp conflict mitigation has emerged as a separate challenge because concurrently executing xApps may issue inconsistent control actions [35, 20]. These constraints imply that interference coordination in O-RAN cannot assume ideal centralized joint scheduling; it must operate within the observability, latency, and interface limitations of the O-RAN control hierarchy.

2.2.2 Scheduler-Level Energy Profiling in O-RAN

O-RAN changes the energy model of the RAN. In conventional radio access networks, scheduling is tightly coupled to vendor-specific hardware. In O-RAN, scheduling functions increasingly execute as software components on commercial off-the-shelf processors or virtualized cloud-native platforms. As a result, the energy consumed by the scheduling algorithm itself becomes part of the RAN energy budget.

Most energy-efficiency studies in wireless networks define EE as transmitted bits per joule of radio-side energy, usually incorporating transmit power and sometimes circuit power. This is necessary but incomplete for O-RAN, where scheduling computation, teleme-

try processing, xApp execution, container orchestration, and virtualized O-DU processing all contribute to energy consumption [7, 40]. The ITU-R IMT-2030 framework explicitly identifies sustainability as a central 6G design objective [10, 13]. Ahmad et al. [15] further argue that 6G sustainability must include not only energy efficiency but also reliability, fairness, adaptability, and environmental responsibility. The O-RAN green communication report similarly identifies energy-aware RAN operation as a key research priority and highlights the role of open interfaces in enabling software-defined energy management [41].

Kundu et al. [42] survey energy-efficient RAN techniques from standards to live deployments, and identify a gap between network-level energy-saving strategies and fine-grained energy profiling of RAN functions. TENORAN [43] proposes automated energy profiling for Open RAN systems at the function level, representing an important step toward reproducible benchmarking. Nevertheless, neither 3GPP nor O-RAN has standardized metrics at the granularity of individual scheduling operations, which means that two schedulers may be compared in terms of radio-side EE, while their computational energy costs remain unobserved [44].

Several related works reinforce this limitation. Wang et al. [45] studied energy minimization for IoT devices in O-RAN-based systems but focused on terminal-side energy rather than O-DU or RIC processing energy. Larsen et al. [46] surveyed green approaches for 5G-and-beyond RAN deployments and observed that virtualization transfers part of the energy responsibility from specialized hardware to software-defined control. Aslan et al. [47] considered fair resource allocation in virtualized O-RAN platforms and recognized that virtualization changes the resource-cost structure of scheduling. However, none of these works provides an instruction-level or scheduler-level methodology for quantifying computational energy on virtualized O-DU/RIC platforms.

This omission is significant for dense mMTC. Scheduling complexity grows with the number of users, resource blocks, objective functions, and coordination constraints. A scheduler that improves radio-side EE but requires excessive computation may not be sustainable when deployed at near-real-time cadence. Therefore, scheduler-level energy profiling is not an auxiliary concern; it is a necessary component of sustainable O-RAN scheduler design.

2.2.3 Inter-O-DU Coordination and Scheduler-Level Energy Profiling: Problem, Limitations, and Prior Approaches

In standard O-RAN operation, each O-DU schedules its own users using locally available channel, buffer, and service information. Neighbouring O-DUs may therefore allocate overlapping RBs to users whose uplink transmissions generate strong mutual interference. This gives rise to decision-dependent ICI: the interference experienced by one cell is determined not only by geometry and propagation, but also by the current resource-allocation decisions of adjacent cells [48]. Unlike fixed external interference, this interference field changes whenever neighbouring schedulers change their actions. Treating it as a static channel property is therefore insufficient.

The E2 interface provides a potential mechanism for exchanging telemetry and control information, but its 10–100 ms control-loop timescale is not compatible with ideal per-

slot joint scheduling in highly dynamic mMTC scenarios [48]. Coordination over O-RAN interfaces must therefore be robust to stale, partial, or aggregated information. It must also avoid excessive message exchange, since coordination overhead competes with the latency and computational budget available to the scheduler. This distinguishes O-RAN inter-O-DU coordination from fully centralized CoMP-style joint transmission: in the O-RAN setting, each O-DU remains responsible for serving its own users, and coordination should guide rather than replace local scheduling.

Prior work has approached this problem from several directions. Akhtar et al. [49] studied NOMA-based RRM for traffic steering in O-RAN and showed that non-orthogonal access can improve spectrum utilization. However, the work does not provide a general inter-O-DU coordination mechanism for dense multi-cell mMTC. Ranjbar et al. [50] examined cell-free massive MIMO support in O-RAN and argued that additional inter-O-DU coordination is necessary to realize distributed spatial-multiplexing gains. Zangoeei et al. [51] investigated flexible RAN slicing using constrained multi-agent RL and demonstrated that learning methods can respect heterogeneous QoS constraints, but their focus is slice-level resource management rather than scheduler-table or interference-footprint coordination among O-DUs. Toure et al. [52] proposed a MORL-based xApp capable of adapting to different throughput-fairness preferences without retraining, but the study remains focused on single-O-DU control. Lotfi et al. [53] studied O-RAN resource management using multi-agent RL with explicit coordination overhead constraints, confirming that coordination benefits must be weighed against latency and message-passing cost.

Mwakwata et al. [54] provide one of the most directly relevant external baselines by proposing a cooperative scheduler for massive connectivity in OMA and NOMA systems. Their results show that sharing scheduling information between neighbouring cells can improve connectivity by approximately 1.5 times compared with non-cooperative operation. However, the work is developed in a 5G NB-IoT context, and relies on Shannon IFBL capacity, and does not incorporate FBL reliability, O-RAN-specific interface constraints, or scheduler-level computational energy profiling.

Other energy-aware multi-cell works similarly, addressing only part of the problem. Rabee and Barhumi [55] used multi-agent DRL for uplink NOMA EE optimization and reported approximately 25% improvement, but the evaluation was limited to 30 users per cell. Rashid et al. [56] reported EE and throughput gains in uplink MC-NOMA heterogeneous small-cell networks, with scalability limited to 20 users per cell. Jameel et al. [57] used Markov decision processes for CoMP C-NOMA EE optimization but did not analyze large-scale connectivity. Zhao et al. [58] proposed a DDQN scheme for EE and connectivity in ultra-dense small cells, but omitted FBL constraints and explicit O-RAN inter-O-DU coordination. Blanco et al. [16] introduced a scalable and sustainable 6G framework at the architectural level, but did not provide a concrete O-DU-level scheduling and energy-profiling methodology.

2.2.4 Research Gap and Link to RQ1

The O-RAN scheduling and energy-modeling literature reveals three coupled gaps that motivate RQ1. First, existing O-RAN scheduling studies do not provide an architecture-compliant

inter-O-DU coordination mechanism that is scalable to dense mMTC and compatible with E2/F1 interface constraints. Second, existing O-RAN coordination studies generally evaluate interference, throughput, or connectivity under Shannon IFBL assumptions. Third, energy-aware O-RAN studies largely focus on radio-side energy and do not provide a reproducible methodology for profiling the computational energy consumed by scheduling algorithms on virtualized O-DU hardware.

2.3 Challenge 2 — Physical-Layer Realism: FBL-Aware Multi-Cell NOMA optimization

The second challenge concerns the physical-layer validity of the scheduling model. Coordination can reduce decision-dependent ICI, but the scheduler must still optimize against a rate model that accurately reflects what the wireless link can deliver. Much of the multi-cell NOMA literature relies on Shannon's IFBL capacity expression. This assumption is appropriate for long codewords but optimistic for the short packets characteristic of mMTC. In the short-packet regime, the achievable rate depends jointly on SINR, block-length, and decoding error probability. A physically credible scheduler must therefore incorporate FBL constraints into its resource-allocation and power-control decisions.

2.3.1 NOMA in Multi-Cell Dense Deployments; Cooperative and Non-Cooperative Interference

Power-domain NOMA allows multiple devices to share the same time-frequency resource by assigning different power levels and applying successive interference cancellation (SIC) at the receiver. This enables higher spectral efficiency and connection density than OMA, particularly when co-scheduled users have sufficiently distinct channel gains [59, 60]. In single-cell settings, the capacity and fairness advantages of NOMA have been extensively studied, and NOMA has been identified as a promising access technique for massive connectivity [61, 60].

Classical NOMA resource-allocation studies have addressed power allocation, user pairing, and scheduling under EE, throughput, and QoS objectives [62, 63, 64, 65]. Additional works have studied NOMA for massive IoT and mMTC under latency and power constraints [66, 67]. Collectively, these studies confirm that NOMA can increase device density compared with OMA at equivalent radio-resource budgets. However, they typically rely on IFBL rate expressions and focus on single-cell or weakly coupled multi-cell settings.

The transition from single-cell to multi-cell NOMA introduces a more difficult interference structure. In single-cell NOMA, SIC can partially remove intra-cell interference among users sharing the same RB. In multi-cell NOMA, each user also experiences inter-cell interference from co-channel users in neighbouring cells. This interference is coupled to their scheduling and power-control decisions. A resource allocation in one cell therefore affects not only local users, but also the feasible SINR, rate, and reliability of users in adjacent cells. This coupling prevents independent per-cell optimization and motivates cooperative or coordinated scheduling [68].

Liu and Zhang [69] survey the evolution of NOMA towards next-generation multiple access and identify scalability, SIC complexity, and dense-deployment interference as key

barriers to practical 6G adoption. Belmekki and Alouini [70] extend the NOMA discussion to non-terrestrial network settings and highlight the structural role of interference in power-domain multiplexing. These surveys make clear that single-cell NOMA is well understood, while dense multi-cell NOMA remains limited by interference coupling, coordination overhead, and short-packet reliability constraints.

Cooperative scheduling has been proposed to mitigate this limitation. Mwakwata et al. [54] showed that sharing scheduling information among neighbouring cells can improve massive connectivity in OMA and NOMA systems. Muhammed et al. [71] studied resource allocation for energy-efficient NOMA in CoMP networks, demonstrating that coordinated scheduling can improve EE under IFBL assumptions. Trabelsi et al. [72] provide a broad survey of interference management for 5G and beyond, confirming that combined intra-cell and inter-cell interference remains a dominant bottleneck in dense NOMA deployments. Nevertheless, these works generally optimize under Shannon-capacity models and therefore do not capture the FBL rate loss and decoding error constraints that dominate short-packet mMTC.

2.3.2 Joint EE-Connectivity optimization Under IFBL (Shannon)

Under Shannon IFBL assumptions, joint EE-connectivity optimization in multi-cell NOMA has been studied through mixed-integer optimization, successive convex approximation, graph-based methods, and learning-based control. Zeng et al. [73] investigated energy-efficient joint user-RB association and power allocation for uplink hybrid NOMA-OMA systems, showing that coordinated allocation can outperform OMA baselines. Cui et al. [74] studied QoE-based resource allocation for multi-cell NOMA networks and demonstrated that joint user assignment and power optimization improve user experience compared with independent per-cell scheduling. Song et al. [75] applied cooperative multi-agent RL to EE optimization for uplink mURLLC in multi-cell networks, while Zhao et al. [58] used DDQN for uplink resource allocation with QoS constraints in 6G ultra-dense systems.

These studies establish the value of joint scheduling and power allocation for multi-cell NOMA. They also show that EE and connectivity cannot be optimized independently: increasing the number of admitted users can increase interference and reduce per-user reliability, while aggressively minimizing transmit power can reduce connectivity and outage performance. However, because these works use IFBL rate models, their achievable rate and EE estimates are optimistic for short-packet mMTC. The Shannon expression implicitly assumes asymptotically long codewords and vanishing error probability, whereas mMTC devices often transmit short, sporadic packets under non-negligible decoding error probability. Thus, IFBL-based optimization provides a useful upper bound but not a physically complete model for 6G mMTC scheduling.

2.3.3 FBL Theory, Short-Packet Transmission, and EE Impact

The theoretical foundation of FBL communication was established by Polyanskiy, Poor, and Verdú [21], who derived normal approximations for the maximum achievable coding rate as a function of block-length, decoding error probability, and channel quality. In a generic wireless setting, the FBL achievable rate can be expressed as

$$R(L, \varepsilon, \gamma) = \ln(1 + \gamma) - \frac{Q^{-1}(\varepsilon)}{\sqrt{L}} \sqrt{1 - \frac{1}{(1 + \gamma)^2}} + \mathcal{O}\left(\frac{\ln L}{L}\right) \quad (1)$$

where L is the block-length, ε is the decoding error probability, γ is the SINR, and $Q^{-1}(\cdot)$ is the inverse Gaussian Q -function mapping the error probability to a reliability penalty factor. $\sqrt{1 - \frac{1}{(1 + \gamma)^2}}$ is square root of the channel dispersion (V), which quantifies how much the information density fluctuates over the short blocklength. It disappears only as $L \rightarrow \infty$, meaning that Shannon capacity systematically overestimates the rate of finite-length packets. $\mathcal{O}\left(\frac{\ln L}{L}\right)$ is the asymptotic error remainder, bounding the numerical accuracy of the second-order normal approximation.

This FBL penalty has direct implications for EE and connectivity. First, the same SINR supports a lower rate under FBL than under IFBL; therefore, an IFBL-based scheduler may admit more users than the physical layer can reliably serve. Second, reliability, latency, and energy are coupled: reducing the block-length reduces latency but increases the rate penalty and decoding error probability. Third, EE can be non-monotonic in block-length. Very short packets suffer from poor coding efficiency, while very long packets increase transmission and circuit energy. Therefore, an EE-maximizing scheduler must identify a block-length and power operating region rather than simply minimizing packet duration or maximizing SINR. The effective-capacity framework of Wu and Negi [76] provides an early connection between the physical-layer rate and QoS-constrained service guarantees.

Recent works have incorporated these ideas into short-packet resource allocation. Liu et al. [77] studied the relationship between FBL coding and EE, showing that the EE-optimal operating point depends jointly on channel conditions, target error probability, and block-length. Zhong et al. [78] developed minimum-BLER NOMA designs under FBL constraints and clarified the SINR-error-probability trade-off in short-packet NOMA. Liu et al. [79] studied energy-efficient scheduling for short-packet communications with FBL coding and confirmed that ignoring FBL leads to inaccurate EE conclusions. These works establish the need for FBL-aware scheduling, but they are largely single-cell or link-level and do not address dense multi-cell NOMA coordination.

2.3.4 Joint EE-Connectivity optimization Under FBL (Short-Packet Regime)

The joint optimization of EE and connectivity under FBL constraints in multi-cell NOMA remains an open problem because it requires four dimensions to be handled simultaneously: EE maximization, massive connectivity, finite-block-length reliability, and inter-cell interference coordination. Existing works typically address only a subset of these dimensions.

Gao et al. [80] analyzed finite-block-length massive unsourced random access in a single-cell MIMO setting, establishing fundamental connectivity bounds, but not optimizing EE or multi-cell interference. Sabuj et al. [81] studied multi-UAV-aided mMTC/URLLC services under FBL constraints, focusing on energy and latency performance in a single-cell architecture. Wang et al. [82] studied FBL-enhanced effective-capacity optimization for uplink NOMA, addressing delay-aware throughput and EE, but not multi-cell ICI. Zhao et al. [83] optimized EE for multi-node IoT networks with finite-block-length codes, but without inter-cell interference coupling.

Table 2: Comparison of selected related works on sustainable scheduling under FBL in mMTC/URLLC networks. “ICI-aware” indicates whether both intra-cell and inter-cell interference are explicitly considered.

Ref.	Environment	Technique	EE	Conn.	FBL	ICI-Aware
[80]	Single-cell MIMO URA	FBL information-theoretic bounds	×	✓	✓	×
[81]	Single-cell UAV mMTC/URLLC	Joint power optimization	✓	×	✓	Intra-cell only
[82]	Single-cell uplink NOMA	Delay-aware throughput optimization	✓	×	✓	Intra-cell only
[83]	Single-cell IoT uplink	Power and block-length optimization	✓	✓	✓	×
[84]	Single-cell STAR-RIS MU-MIMO	Fractional-programming EE optimization	✓	×	✓	×
[85]	Single-cell URLLC/eMBB	Joint optimization and DRL	✓	×	✓	Intra-cell only
[86]	Multi-cell UAV-NOMA	FBL outage and stochastic geometry	✓	×	✓	Partial
[87]	Cell-free massive MIMO URLLC	SCA-based power and AP association	✓	×	✓	Partial
[88]	Single-cell massive URLLC	Adaptive FBL multi-agent DQN	×	✓	✓	Intra-cell only
[58]	Multi-cell ultra-dense small cells	DDQN RB and power control	✓	✓	×	Intra-cell only
[54]	Multi-cell NOMA	Cooperative scheduling	×	✓	×	Intra+Inter
This thesis	Multi-cell NOMA mMTC	Joint EE-connectivity optimization under FBL	✓	✓	✓	Intra+Inter

Soleymani et al. [84] considered FBL-constrained EE in STAR-RIS-aided MU-MIMO, and Tran et al. [85] studied energy-efficient NOMA for heterogeneous URLLC/eMBB services using joint optimization and DRL. Thapliyal et al. [86] analyzed energy-harvesting UAV-NOMA under FBL outage constraints, while Huang et al. [87] addressed energy-efficient optimization in user-centric cell-free massive MIMO for URLLC with FBL communications. Zhang et al. [88] investigated adaptive FBL multiple access for ultra-low delay, and Zhao et al. [58] studied multi-cell EE and connectivity without FBL constraints.

Chen et al. [68] studied optimal resource allocation for multi-carrier NOMA in short-packet communications, providing important evidence that conventional NOMA pairing heuristics may become sub-optimal under FBL constraints. Zheng et al. [89] analyzed uplink NOMA short-packet reliability and derived performance expressions for decoding success probability. Yang et al. [90] studied RIS-assisted NOMA-based MEC networks in the FBL regime. These works demonstrate the importance of FBL in NOMA and short-packet communication, but they do not jointly optimize EE and connectivity under coordinated multi-cell interference.

Table 2 summarizes the positioning of representative external works against the four dimensions required by the thesis problem. The table highlights a clear structural gap. FBL-aware works are usually single-cell or link-level. Multi-cell works that address interference often rely on IFBL assumptions. Works that optimize EE often treat connectivity or reliability as constraints rather than coequal objectives. Conversely, works that focus on connectivity rarely incorporate FBL reliability and multi-cell coordination simultaneously. This confirms that the short-packet, multi-cell, EE-connectivity problem remains insufficiently addressed.

2.3.5 Research Gap and Link to RQ2

The NOMA and FBL literature reveals a precise gap: no existing work simultaneously addresses EE maximization, connectivity maximization, finite-block-length reliability constraints, and inter-cell interference coordination in a scalable multi-cell NOMA framework. FBL-aware studies are primarily single-cell or link-level, while multi-cell NOMA coordination studies generally rely on Shannon IFBL capacity. This absence of a unified multi-objective, FBL-aware, interference-coordinated framework motivates RQ2.

2.4 Challenge 3 — Adaptability: Multi-Objective Reinforcement Learning for 6G Scheduling in Centralized and Distributed O-RAN

The third challenge identified in Chapter 1 is *adaptability*: the ability of the scheduler to operate under time-varying channels, traffic loads, interference coupling, and service priorities without repeatedly solving a non-convex mixed-integer optimization problem at every scheduling interval. The optimization frameworks reviewed under the physical realism challenge provide a rigorous basis for characterizing EE, connectivity, reliability, and interference trade-offs. However, such formulations are NP-hard to solve from scratch within O-RAN near-real-time control loops, particularly when the scheduler must jointly decide user grouping, RB assignment, power allocation, and interference mitigation in dense multi-cell NOMA deployments.

Reinforcement learning (RL) offers a complementary paradigm: instead of solving a new

optimization instance at each control step, an agent learns a policy that maps observed network states to scheduling actions through a single inference pass. In O-RAN, this policy may be deployed centrally at the Near-RT RIC, locally at O-DUs, or through a hybrid architecture in which training and execution are split across centralized and distributed control planes. Since 6G scheduling involves multiple conflicting objectives rather than a single scalar target, multi-objective reinforcement learning (MORL) is particularly relevant. MORL represents the scheduler reward as a vector of KPIs and seeks Pareto-efficient policies capable of navigating trade-offs among EE, connectivity, reliability, latency, and interference suppression.

This subsection reviews the RL and MORL literature through the lens of O-RAN deployability. Section 2.4.1 summarizes the main RL algorithm families and their relevance to O-RAN control loops. Section 2.4.2 compares single-agent and multi-agent architectures, including centralized training with decentralized execution (CTDE). Section 2.4.3 reviews MORL solution concepts, including scalarization, multi-policy learning, preference conditioning, envelope updates, and Pareto-front discovery. Finally, Section 2.4.4 identifies the gap.

2.4.1 RL Taxonomy: On-Policy, Off-Policy, Model-Based, and Model-Free Learning

Model-free RL learns scheduling policies directly from interaction with an environment, without requiring an explicit analytical model of the network transition dynamics. This makes it attractive for RRM, where the channel variation, traffic burst, user mobility, and interference coupling are difficult to model exactly [91, 92]. In the O-RAN context, model-free policies can be trained using system-level simulators, digital twins, or logged telemetry, and then deployed as xApps or dApps, depending on the required control-loop latency. The main advantage is flexibility; the main limitation is that training may require substantial interaction data and careful constraint handling before deployment in a live RAN.

Within model-free RL, value-based methods such as DQN and DDQN learn action-value functions and select actions by maximizing estimated state-action values. These methods are typically off-policy and can reuse experience through replay buffers, which improves sample efficiency. However, their natural action space is discrete. This is suitable for problems such as RB selection or admission decisions, but less suitable for multi-cell NOMA scheduling, where transmit powers and NOMA power-domain coefficients are continuous or mixed discrete-continuous variables. DDQN has therefore been used for wireless resource allocation with QoS constraints, but continuous power control usually requires discretization or an additional action-parameterization layer [58]. In dense O-RAN scheduling, such as discretization, can reduce policy resolution and increase action-space dimensionality.

Policy-gradient and actor-critic methods optimize the policy directly and are therefore more naturally suited to continuous or hybrid scheduling actions. PPO, SAC, and TD3 are representative examples used in wireless resource allocation and network-control problems [93]. PPO is especially relevant for O-RAN scheduling because it is an on-policy method with a clipped surrogate objective that limits the magnitude of each policy update. This improves stability under nonlinear function approximation and makes PPO attractive when multiple KPIs generate conflicting learning signals. Schulman et al. [94] established

PPO as a stable policy-gradient method, and its constrained update structure aligns well with RAN control, where unstable policy changes can directly affect QoS and reliability.

Model-based RL learns or exploits a transition model and uses it for planning. In principle, this can improve sample efficiency because the agent can simulate future outcomes before acting. In wireless scheduling, however, accurate modeling of multi-cell interference, traffic dynamics, and user mobility is challenging. Model errors may accumulate during planning, and the additional model-evaluation burden can be incompatible with O-RAN Near-RT RIC control-loop deadlines [92]. For this reason, model-based RL is attractive for offline analysis and digital twin acceleration, but less straightforward for direct near-real-time scheduler execution.

Offline RL trains policies from logged datasets without further online exploration. This is highly relevant to O-RAN because E2 telemetry and RAN logs can, in principle, support policy training without exposing the live network to unsafe exploration. Offline RL is therefore a promising future direction for safe RAN intelligence, particularly when live trial-and-error learning is impractical [95]. Nevertheless, offline RL introduces distributional-shift challenges: a policy may select actions that are poorly represented in the dataset, which is problematic in safety- and QoS-constrained wireless systems.

A second taxonomy dimension is the on-policy versus off-policy distinction. On-policy methods, such as PPO, update the policy using data generated by the current policy. They tend to be more stable but less sample-efficient. Off-policy methods, such as DQN and SAC, reuse historical experience and can therefore be more data-efficient, but they are more sensitive to distributional shift and non-stationarity [91]. In O-RAN, this distinction has practical consequences. A Near-RT RIC scheduler must remain stable under evolving interference and traffic distributions, whereas a Non-RT RIC training pipeline may prioritise sample efficiency through replay and offline data reuse. A deployable learning architecture must therefore align the RL algorithm family with the control-loop location, available telemetry, action latency, and safety requirements.

2.4.2 Single-Agent vs. Multi-Agent Architectures

The choice between single-agent and multi-agent RL is not only an algorithmic decision; it is also an O-RAN architectural decision. A centralized single-agent controller assumes that a sufficiently broad network state is available at one decision point, such as a Near-RT RIC or a C-RAN-like controller. A distributed multi-agent design assigns separate agents to O-DUs, cells, slices, or user groups, allowing local execution with reduced state dimensionality. Between these two extremes lies CTDE, where agents are trained with access to global or regional information but execute using local observations.

A centralized single-agent architecture avoids the fundamental non-stationarity of independent multi-agent learning because one controller selects the joint scheduling action for a cluster of cells. This is attractive for interference-coupled NOMA scheduling: the controller can observe cross-cell channel and interference indicators and jointly assign RBs and powers to reduce ICI. In O-RAN terms, such a policy can be hosted at the Near-RT RIC as an xApp if the state collection, inference, action projection, and control delivery fit within the control-loop deadline. The limitation is scalability. The state and action

dimensions increase with the number of cells, UEs, and RBs, and the E2 interface may not provide fresh enough telemetry for fine-grained global scheduling. Thus, centralized control is most practical for limited interference clusters, not for network-wide real-time optimization.

Multi-agent RL (MARL) maps more naturally to the distributed structure of O-RAN. Each O-DU or cell can run a local agent, observe the local channel and load information, and choose scheduling actions with limited coordination. Recent surveys classify MARL for resource allocation into fully decentralized learning, CTDE, and communication-enabled coordination [96]. CTDE is particularly relevant for O-RAN: the Non-RT RIC or an offline training environment can train agents using global information, while the learned policies execute locally at O-DUs or through Near-RT RIC xApps using only local or partially shared observations. This matches the O-RAN separation between policy/model management and near-real-time execution.

O-RAN-specific studies show both the promise and difficulty of distributed learning. Lotfi et al. [53] investigated MARL for O-RAN resource management under coordination-overhead constraints and showed that multi-agent coordination must be evaluated jointly with the message-passing latency it introduces. Han et al. [97] studied multi-agent RL for uplink OFDMA in dense WLANs, demonstrating that distributed agents can coordinate resource usage under overhead constraints. Zhu et al. [98] provide a broader survey of multi-agent deep RL with communication and identify non-stationarity as a dominant obstacle: from the perspective of each agent, the environment changes as other agents simultaneously update their policies.

This non-stationarity is particularly problematic in dense multi-cell NOMA. The action of one O-DU changes not only its own users' SINR but also the interference experienced by neighbouring O-DUs. Independent learners may therefore chase a moving target, causing oscillatory power control or unstable RB reuse. Communication among agents can mitigate this, but it increases signalling overhead and may violate O-RAN latency constraints. Conversely, full centralisation reduces learning non-stationarity but increases observability and computational burden. Polese et al. [99] emphasise that O-RAN intelligence must respect interface and timing constraints, which means that the centralized/distributed design choice must be evaluated as part of the RAN architecture rather than as an abstract RL preference.

For the Adaptability challenge addressed in this thesis, the key issue is therefore not whether centralized or distributed RL is universally superior. The centralized perspective is needed to learn coherent trade-offs under explicit inter-cell interference coupling; the distributed perspective is needed for scalable O-RAN execution. A strong MORL scheduler should be compatible with both views: it should learn policies that exploit regional context when available, while remaining structured enough to support replication, decomposition, or CTDE-style deployment in distributed O-RAN.

2.4.3 MORL: Solution Concepts and State of the Art

MORL extends RL to vector-valued rewards, where each reward component corresponds to a distinct objective. In 6G scheduling, these objectives may include EE, connectivity,

reliability, throughput, latency, fairness, interference suppression, and computational feasibility. Unlike single-objective RL, MORL does not generally yield one universally optimal policy. Instead, the goal is to approximate a Pareto set of non-dominated policies or to learn a policy that can produce different Pareto-efficient actions when conditioned on an operating preference.

Foundational MORL work formalizes this problem through multi-objective Markov decision processes and Pareto dominance. Van Moffaert and Nowé [100] established the basis for learning sets of Pareto-dominating policies. Evaluation then requires metrics that measure both convergence and diversity. Hypervolume (HV) quantifies the region of objective space dominated by the learned solution set, while inverted generational distance (IGD) measures the distance between learned solutions and a reference Pareto front. These metrics are important for wireless scheduling because a policy that performs well for one operating preference may still be inadequate if it fails to cover reliability-oriented, energy-oriented, and connectivity-oriented regimes.

A first design family is scalarization. A vector reward is converted into a scalar reward using a weighted sum or another utility function. Static scalarization is simple and compatible with standard RL algorithms, but it fixes the trade-off during training. In a dynamic O-RAN environment, this is restrictive: the operator may prioritize EE during low load, reliability during critical service periods, or connectivity during dense access bursts. Multi-policy approaches address this by training separate policies for different weight vectors [101]. Abels et al. [102] showed that dynamic weights can improve Pareto coverage relative to fixed weights, but maintaining multiple policies remains costly in memory, training time, and deployment complexity. This is especially problematic for O-RAN, where policy life cycle management is already constrained by model validation, orchestration, and control-loop timing.

Single-policy preference-conditioned MORL provides a more deployable alternative. A single policy receives both the network state and a preference vector as inputs, allowing it to realize different operating points without retraining. Such methods are attractive for O-RAN because one xApp model can adapt to changing operator intent or network context. Yang et al. [103] developed preference-controllable RL and showed how conditioning can improve trade-off controllability across the preference simplex. Dynamic Preference Inference Networks further improve sample efficiency by estimating favorable preference directions during training [104, 22]. Ambadkar et al. [105] proposed decomposed and diversity-driven preference-conditioned MORL, demonstrating that explicit diversity objectives can improve Pareto-front coverage. Liu et al. [106] introduced C-MORL for efficient discovery of Pareto fronts, reinforcing the importance of coverage and constraint-aware optimization in high-dimensional objective spaces.

Envelope-based MORL is another important solution concept. Instead of learning a separate value function for each preference, envelope methods learn a value representation that approximates the upper envelope of multi-objective returns across the preference space. This can improve generalization and Pareto coverage because the policy is trained to represent multiple trade-offs within a unified value structure. Yan and Tabassum [107] applied envelope updates to URLLC-enabled vehicular networks and showed improved

Pareto performance over fixed-weight baselines. Terekhov and Gülcehre [108] further showed that value-function parameterization and loss design strongly affect Pareto quality and training stability. Their findings are especially relevant to wireless scheduling, where objective gradients may point in opposing directions.

Conflict-aware optimization is therefore a central requirement for MORL in 6G scheduling. EE, connectivity, reliability, and interference suppression are not merely different metrics; they can generate contradictory learning signals. Increasing transmit power may improve reliability and connectivity but reduce EE and increase ICI. Reducing interference may require more conservative RB reuse and thereby reduce throughput or connectivity. Standard scalarized policy-gradient updates can oscillate under such objective conflicts, leading to unstable training and poor Pareto coverage. Recent work on preference-conditioned and conflict-aware MORL, including PCRL-style formulations highlight the need for update mechanisms that explicitly account for preference diversity and gradient conflict [108, 23].

Wireless MORL is now emerging as a distinct research direction. Toure et al. [52] demonstrated multi-objective scheduling in wireless networks using DRL and showed that a learned controller can adapt to different throughput–fairness trade-offs in an O-RAN xApp setting. Heo et al. [109] studied dynamic preference MORL for network management and showed improved generalization compared with static preferences. Al-Habob and Tabassum [110] applied meta-MORL to age-optimal vehicular information dissemination, reducing per-preference retraining overhead. Li et al. [111] applied MORL to vehicle-to-vehicle resource allocation, while Chien et al. [112] addressed multi-objective optimization of 3D base-station deployment under performance and sustainability objectives. Guan et al. [113] employed multi-agent RL to optimize throughput and residual energy in IRS-assisted wireless-powered networks. Al-Tahmeesschi et al. [114] investigated multi-objective DRL for 5G base-station placement, and Balaji et al. [115] highlighted the weakness of static policies under non-stationary sensor-network conditions. Xu et al. [101] studied multi-objective resource allocation in cooperative vehicular networks, and Tan et al. [116] proposed adaptive distributed MARL for multi-objective intelligent transportation systems.

These works collectively demonstrate that MORL is a promising tool for adaptive wireless resource management. However, they also reveal several limitations relative to the thesis problem. Many studies optimize only two or three objectives, whereas sustainable 6G scheduling requires simultaneous control of EE, connectivity, reliability, and interference. Many studies assume preferences are externally specified, whereas O-RAN schedulers should infer priorities from context such as load, interference, QoS deficits, and reliability statistics. Many works are single-cell, vehicular, or deployment-planning oriented, and therefore do not validate policy learning under dense multi-cell NOMA interference. Finally, operational feasibility in O-RAN—including inference latency, action projection, computational cost, and centralized-versus-distributed deployment—is usually treated as an implementation detail rather than an explicit part of the learning problem.

2.4.4 Research Gap and Link to RQ3

The review of RL and MORL for wireless scheduling identifies four limitations that jointly motivate RQ3. First, existing wireless MORL formulations often rely on static scalarization, manually specified preferences, or multiple trained policies, which limit runtime adaptability under changing O-RAN conditions. Second, context-aware preference inference remains underdeveloped: current methods rarely infer the EE-connectivity-reliability-interference priority vector directly from real-time network state. Third, the centralized/distributed O-RAN deployment trade-off is not fully resolved; centralized MORL can model inter-cell coupling, but faces state-dimensionality and control-loop constraints, while distributed or CTDE-based MORL scales better but suffers from non-stationarity, coordination overhead, and incomplete observability. Fourth, no existing external work validates a single adaptive MORL scheduler under the combined conditions of dense multi-cell NOMA, FBL reliability constraints, decision-dependent inter-cell interference, and O-RAN near-real-time execution feasibility.

2.5 Cross-Cutting Themes: Sustainability and O-RAN Alignment

The three challenges reviewed above are connected by two cross-cutting themes: sustainability and O-RAN alignment.

First, sustainability in 6G cannot be reduced to transmit-power minimization. Ahmad et al. [15] argue that sustainable 6G must jointly account for energy, reliability, fairness, adaptability, and environmental impact. Li et al. [7] highlight the importance of green wireless design and carbon-aware network operation. Wang et al. [9] position energy efficiency and sustainability as co-equal design objectives alongside throughput, latency, coverage, and reliability. The ITU-R IMT-2030 framework similarly treats sustainability as a core capability rather than an auxiliary performance metric [10, 40, 13]. The O-RAN green communication report reinforces this view by emphasizing that virtualized, software-defined RAN control introduces new opportunities for energy management but also new responsibilities for computational energy accounting [41].

Second, sustainable scheduling must be cross-layer and architecture-aware. O-RAN provides a practical integration point because the Near-RT RIC can observe RAN telemetry through E2 and apply xApp-based control decisions. However, this observability is bounded by interface latency, telemetry granularity, and control-loop feasibility. Therefore, a scheduler must not only optimize physical-layer metrics such as SINR and rate; it must also account for service-level requirements, interference coupling, and software-execution constraints. Blanco et al. [16] provide an architectural vision for scalable and sustainable 6G networks, while Larsen et al. [46] show that green RAN design must include both radio and virtualization aspects. Li et al. [111] and Polese et al. [37] further support the need for context-aware control that bridges physical-layer state and higher-layer service objectives.

This thesis operationalizes sustainability and O-RAN alignment through three mechanisms. First, inter-O-DU coordination reduces decision-dependent ICI and lowers the transmit power required to satisfy SINR targets. Second, FBL-aware optimization prevents over-admission and over-optimistic EE estimates caused by IFBL capacity assumptions. Third, preference-conditioned MORL allows the scheduler to adapt the trade-off among

EE, reliability, connectivity, and interference according to the current network context. The dynamic preference motivation is consistent with Heo et al. [109], who show that adaptive preference mechanisms can improve generalization under changing conditions.

2.6 Consolidated Gap Analysis and Mapping to Contributions

The challenge-aligned review identifies three research gaps that form a progressive sequence. Gap 1 concerns coordination and scheduler-level energy accounting in O-RAN. Gap 2 concerns physical-layer realism through FBL-aware multi-cell NOMA optimization. Gap 3 concerns adaptability through preference-conditioned and conflict-stable MORL. Table 3 maps each gap to the corresponding research question, external motivating literature, thesis contribution, and technical chapter.

Table 3: Consolidated research gaps, motivating literature, and corresponding thesis contributions.

#	Research question	Key motivating works	Thesis contribution	Ch.
Gap1	RQ1: How can inter-O-DU coordination and scheduler-level energy profiling be designed within O-RAN interface constraints for massive mMTC?	[49, 50, 54, 18, 38, 42, 43, 44]	Publication I and Publication II:	Ch (4)
Gap2	RQ2: How can multi-cell NOMA scheduling be jointly optimized for EE and connectivity under FBL constraints and inter-cell interference?	[54, 21, 77, 78, 79, 68, 87]	Publication III and Publication IV:	Ch (5)
Gap3	RQ3: How can MORL learn a single adaptive policy that generalizes across preference trade-offs and dynamic conditions in 6G O-RAN?	[107, 108, 103, 104, 105, 109, 52, 23]	Publication (V) and Publication (VI):	Ch (6)

The dependency among the gaps is central to the thesis narrative. **Gap1** is the architectural prerequisite: without inter-O-DU coordination, the interference environment remains uncontrolled, and neither analytical optimization nor learning-based control can deliver predictable multi-cell performance. **Gap2** follows from **Gap1**: once interference coordination is introduced, the physical-layer model must be accurate for mMTC short-packet transmissions, which require replacing IFBL assumptions with FBL-aware reliability and rate constraints. **Gap3** follows from **Gap2**: even if the FBL-aware optimization problem is correctly formulated, it cannot be solved from scratch at near-real-time scheduling cadence, motivating adaptive MORL-based policy learning.

This progression also defines the bridge to the technical chapters. Chapter (4) addresses **Gap1** by developing inter-O-DU coordination and scheduler-level energy profiling. Chapter (5) addresses **Gap2** by formulating and solving FBL-aware EE-connectivity optimization for multi-cell NOMA mMTC. Chapter (6) addresses **Gap3** by introducing an adaptive MORL framework for preference-aware and conflict-stable 6G scheduling.

2.7 Chapter Summary

This chapter has provided a challenge-aligned synthesis of the state of the art in sustainable RRM for 6G mMTC networks. Rather than surveying the literature by technology domain, it has followed the logical sequence of the thesis: coordination, physical-layer realism, and adaptability.

The review leads to three central findings. First, existing O-RAN scheduling studies provide a programmable architecture for intelligent RAN control, but they do not jointly solve inter-O-DU coordination, FBL-consistent evaluation, and scheduler-level computational energy profiling. Second, existing NOMA and FBL studies show the importance of short-packet reliability and multi-cell interference, but no external work simultaneously addresses EE maximization, connectivity maximization, FBL constraints, and inter-cell interference coordination in scalable multi-cell NOMA. Third, existing wireless MORL studies demonstrate the promise of multi-objective policy learning, but they often rely on static preferences, multiple policies, or simplified interference models, and they rarely account for O-RAN deployment feasibility.

These findings map directly to **RQ1-RQ3**. Together, they motivate the thesis contributions as a coherent research program: first, coordinate the distributed O-RAN scheduler; second, make the physical-layer optimization short-packet aware, and third, learn adaptive multi-objective policies that can operate within the timing and computational constraints of future O-RAN-enabled 6G systems.

3 System Model

3.1 Overview

This chapter establishes the unified system model that underpins all technical contributions of this thesis. Three model families are developed: a baseline uplink model shared across all papers (Section 3.3); a non-cooperative multi-cell extension that isolates the cost of independent scheduling (Section 3.4); cooperative multi-cell realizations that introduce coordination and short-packet reliability (Sections 3.5); and a cluster-based model for near-real-time RL control (Section 3.6). The design rationale for this progression is discussed in Section 3.9.

Throughout, the uplink is modeled over a finite set of orthogonal resource blocks (RBs), indexed by $r \in \mathcal{R} \triangleq \{1, \dots, R\}$. Users are indexed by u (or i), and cells by s (or m), depending on the paper. Under NOMA, up to a maximum number of users can share the same RB; this parameter appears as M or N_m in different papers and is unified here as $K_{\max}$. The Shannon rate is used as the achievable rate model in all models using an SINR-based expression except the cooperative variant (Publications II), where it is replaced by a finite-block-length (FBL) normal-approximation expression

To avoid symbol fragmentation across the rest of the thesis, a unified notation is introduced in 3.2 (Table 4) and used consistently throughout. The design rationale connecting these model families to the three research questions is discussed in Section 3.9; a quick-reference constraint and notation summary is provided in Section 3.10

3.2 Unified Notation, Symbols, and Constraints

Table 4 provides a single unified notation used throughout this thesis and explicitly reconciles symbol differences across the papers (e.g., M vs. N_m , G vs. h , p vs. P). Unless otherwise stated, all users and base stations have a single-antenna. The technical contributions of this thesis are consolidated from multiple publications and are presented across Chapters (4–6). Because each publication introduces its own symbol choices and notation conventions (for example different indices for cells and users, different symbols for channel gains (e.g., G vs. h), and different representations of multiplexing limits (e.g., M vs. N_m)) directly reusing the original notations would make cross-chapter comparison difficult and increase the risk of ambiguity or inconsistent interpretation.

To ensure clarity and consistency, the thesis adopts a single unified notation that is used throughout the system model (Chapter (3)) and referenced consistently in later chapters. The mapping between the original chapter-specific notations used in Chapters (4–6) and the unified notation adopted in this thesis is provided in Table 4, whereas the constraints relevant to the models are listed in Table 5. This enables the reader to interpret all models, constraints, and results using a single coherent symbol set while preserving traceability to the original papers.

Table 4: Unified notation used in this thesis and reconciliation with symbols used across the papers.

Unified symbol	Meaning	Also written as
$\mathcal{S}, s$	Set of cells and cell index.	$\mathcal{M}, m$
$\mathcal{U}_s, u$	Set of users/UEs in cell s and UE index.	$\mathcal{N}_s, i$
$\mathcal{R}, r$	Set of RBs/resource block and RB index.	k, n
R	Number of RBs/resource block.	K
$K_{\max}$	Max. number of NOMA users multiplexed on a single RB (SIC/receiver constraint).	M, N_m
$a_{u,s}^{(r)}$	Binary RB assignment: $a_{u,s}^{(r)} = 1$ if UE u in cell s uses RB r .	$x_{i,m}^{(r)}$
$p_{u,s}^{(r)}$	UE transmit power allocated on RB r (continuous).	$P_{i,m}^{(r)}$
$p_{\max}$	Maximum UE transmit power constraint.	$P_{\max}$
$h_{u \rightarrow s}^{(r)}$	Channel coefficient/gain from UE u to serving cell s on RB r .	$G_{u,s}^{(r)}$
$ h_{u \rightarrow s}^{(r)} ^2$	Power gain used in SINR expressions (path loss, shadowing, fading).	$G_{u,s}^{(r)}$
B_0	Bandwidth of one RB/resource unit (e.g., 180 kHz in LTE-like numerology).	W
σ_N^2	Noise power on one RB (N_0 : noise spectral density).	$N_0 B_0$
$\gamma_{u,s}^{(r)}$	SINR of UE u in cell s on RB r .	$\text{SINR}_{i,m}^{(r)}$
$\gamma_{\min}$	Minimum SINR/QoS constraint for successful decoding.	$\Gamma_{\min}$
ε, L	FBL decoding error probability and block-length.	ε, n

Note: When the original paper uses gains G directly, interpret $G \equiv |h|^2$. When it uses channel coefficients h , interpret gains as $|h|^2$.

3.3 Uplink System Model

Network abstraction and scheduling locus

This thesis adopts a baseline uplink model aligned with the multi-cell RAN scheduling viewpoint Publication (I–VI). Scheduling decisions are taken at the cell side (e.g., gNB scheduler), over RBs $r \in \mathcal{R}$, serving a set of UEs $\mathcal{U}_s$ in cell s . For OMA, at most one UE occupies an RB in a given cell. For NOMA, up to $K_{\max}$ UEs may be multiplexed on the same RB in a cell.

Shannon Infinite Block length (IFBL) rate

For UE u in cell s scheduled on RB r , the Shannon-style uplink rate is

$$R_{u,s}^{(r)} = B_0 \log_2 \left(1 + \gamma_{u,s}^{(r)} \right), \quad (2)$$

where $\gamma_{u,s}^{(r)}$ is the corresponding SINR.

Generic SINR

The SINR experienced by a scheduled UE on a given RB depends on the strength of its own signal, intra-cell interference from co-scheduled NOMA users on the same RB, an inter-cell interference term $I_s^{(r)}$ (whose specific form varies across model families and is left unspecified here), and thermal noise. A unified expression covering all model families is

$$\gamma_{u,s}^{(r)} = \frac{a_{u,s}^{(r)} |h_{u \rightarrow s}^{(r)}|^2 p_{u,s}^{(r)}}{\underbrace{\sum_{v \in \mathcal{U}_s \setminus \{u\}} a_{v,s}^{(r)} |h_{v \rightarrow s}^{(r)}|^2 p_{v,s}^{(r)}}_{\text{intra-cell (NOMA) interference}} + \underbrace{I_s^{(r)}}_{\text{inter-cell interference model}} + \sigma_N^2}. \quad (3)$$

where $I_s^{(r)}$ is the aggregated uplink inter-cell interference received at gNB s on RB r from UEs scheduled in all other cells $t \neq s$ on the same RB. Depending on the paper, the SINR equation may include (i) intra-cell co-channel interference (NOMA), (ii) a simplified or bounded inter-cell interference term, and (iii) noise σ_N^2 . For OMA, the intra-cell interference term is zero because at most one UE is scheduled per RB in a cell, and can be defined as :

$$I_s^{(r)} \triangleq \sum_{\substack{t \in \mathcal{S} \\ t \neq s}} \sum_{w \in \mathcal{U}_t} a_{w,t}^{(r)} |h_{w \rightarrow s}^{(r)}|^2 p_{w,t}^{(r)}, \quad (4)$$

Channel modeling assumptions common to the thesis

Channel modeling is common across all system models because it drives (i) SINR and feasibility, (ii) NOMA user grouping (strong/weak users), and (iii) interference coupling strength across cells. A generic decomposition used throughout the thesis is:

$$h_{u \rightarrow s}^{(r)} = \sqrt{\beta_{u,s}} g_{u,s}^{(r)}, \quad (5)$$

where:

- $\beta_{u,s}$ is the **large-scale gain** capturing distance-dependent path loss (and optionally shadowing),
- $g_{u,s}^{(r)}$ is the **small-scale fading** term (e.g., Rayleigh/Rician as appropriate),
- $|h_{u \rightarrow s}^{(r)}|^2$ is used directly in the SINR expressions.

Large-scale propagation.

Across papers, large-scale propagation is modeled using a sub-6 GHz cellular path-loss model appropriate for macro/micro coverage (e.g., variants of the Okumura–Hata model in simulation-based studies). This choice provides realistic distance heterogeneity, which is essential for: (i) creating the near–far structure exploited by NOMA, and (ii) producing realistic interference graphs in multi-cell settings.

Small-scale fading and RB dependence.

Small-scale fading $g_{u,s}^{(r)}$ is modeled as i.i.d. across users and RBs (or as flat fading per RB, depending on the simulator resolution), with quasi-static behavior within each scheduling interval (Assumption A2). This balances realism with tractable Monte Carlo evaluation.

Noise model.

Noise power on RB r is modeled as $\sigma_N^2 = N_0 B_0$, consistent with the SINR definitions in Chapter (4). This is required for consistent feasibility evaluation under (C5) and for reliability-aware rate evaluation under FBL (C7).

The channel model above is the minimal structure needed to (i) capture near–far effects driving NOMA performance, (ii) quantify inter-cell interference coupling, and (iii) keep the optimization and learning problems tractable across dense user populations. It also ensures that comparisons across Chapters (4–6) reflect algorithmic differences rather than inconsistent propagation assumptions.

3.4 Non-Cooperative Multi-Cell Uplink Model

This non-cooperative model is used in relation to Chapter (4–5) and Publications (I–III). In a non-cooperative multi-cell setting, each cell schedules independently, without explicit coordination with neighboring cells. The central modeling feature is that *inter-cell interference is not controlled via cooperative decisions*: instead, each cell optimizes its own resources while treating interference from adjacent cells as a fixed background term.

Inter-cell interference approximation

When neighbouring cells do not share scheduling information, the exact inter-cell interference seen at gNB s on RB r given in Eq. (4) is unknown at scheduling time. To obtain a tractable problem that nevertheless reflects realistic interference scaling, the inter-cell interference is replaced by the following bounded, worst-case approximation:

$$I_s^{(r)} = K_{\max} N_{\text{adj}} p_{\max} \overline{G}_{\text{adj}}, \quad (6)$$

where N_{adj} denotes the number of adjacent interfering cells, and $\overline{G}_{\text{adj}}$ is an average (or representative) interfering-channel power gain. This abstraction captures interference scaling with the number of concurrent uplink transmissions in adjacent cells and their maximum power, while keeping the model analytically and algorithmically tractable.

This approximation is adopted for three reasons. *First*, it is analytically tractable: because $I_s^{(r)}$ no longer depends on the scheduling decisions of neighbouring cells, the optimization over $\{a_{u,s}^{(r)}, p_{u,s}^{(r)}\}$ decouples across cells and can be solved on a per-cell basis. This would not be possible if the exact decision-dependent ICI expression Eq. (4) were used.

Second, it is conservative: by assuming all K_{max} NOMA slots in every adjacent cell are simultaneously occupied at maximum power p_{max} , the approximation provides an upper bound on interference under fully loaded conditions. EE and connectivity gains reported under this model are therefore achievable lower bounds that hold even in worst-case loading scenarios.

Third, it isolates the research question of interest: as studied in Publications (I–III), the model quantifies the cost of *uncoordinated scheduling* decisions under realistic interference scaling, the reference point against which the cooperative models introduced in Section 3.5 are compared.

Non-Cooperative SINR

Substituting Eq. (6) into the generic SINR expression Eq. (3) gives the non-cooperative SINR used across Publications (I–III). This non-cooperative model is used in the analysis presented in Chapters (4–5). The following section introduces the cooperative extension, in which neighbouring cells exchange scheduling information to make inter-cell interference decision-dependent.

3.5 Cooperative Multi-Cell Uplink Model

The cooperative multi-cell uplink model addresses the central limitation of the non-cooperative setting: inter-cell interference is no longer treated as a fixed worst-case bound, but becomes *decision-dependent*, responding directly to the scheduling and power-control choices made by neighbouring cells. Two cooperative realizations are developed in this thesis, corresponding to Publications (IV) and (VI), respectively, and are presented sequentially below.

Cooperative Interference-aware Multi-Cell NOMA with FBL

In this cooperative model, interference is computed directly from neighboring-cell user schedules. Let $\mathcal{N}(s) \subseteq \mathcal{S} \setminus \{s\}$ denote the set of cells neighboring cell s . Then the inter-cell interference term on RB r is written as,

$$I_s^{(r)} = \sum_{t \in \mathcal{N}(s)} \sum_{j \in \mathcal{U}_t} a_{j,t}^{(r)} |h_{j \rightarrow s}^{(r)}|^2 p_{j,t}^{(r)}. \quad (7)$$

This definition makes inter-cell interference *decision-dependent*: if neighboring cells avoid co-scheduling strongly interfering users on the same RB, $I_s^{(r)}$ is reduced.

Multi-cell SINR with Intra- and Inter-cell interference

This model uses explicit intra- and inter-cell interference terms. For UE u in cell s on RB r ,

$$\gamma_{u,s}^{(r)} = \frac{a_{u,s}^{(r)} |h_{u \rightarrow s}^{(r)}|^2 p_{u,s}^{(r)}}{\underbrace{\sum_{v \in \mathcal{U}_s \setminus \{u\}} a_{v,s}^{(r)} |h_{v \rightarrow s}^{(r)}|^2 p_{v,s}^{(r)}}_{\text{intra-cell (NOMA)}} + \underbrace{\sum_{t \in \mathcal{S} \setminus \{s\}} \sum_{j \in \mathcal{U}_t} a_{j,t}^{(r)} |h_{j \rightarrow s}^{(r)}|^2 p_{j,t}^{(r)}}_{\text{inter-cell (multi-cell)}} + \sigma_N^2}. \quad (8)$$

FBL Rate Model

For short-packet mMTC transmissions, the classical Shannon capacity overestimates the achievable rate because it assumes packets of infinite length. In practice, mMTC devices transmit short bursts of data, and the finite number of channel uses per packet introduces a rate penalty and a non-negligible decoding error probability. The achievable rate under these conditions is modeled using the Gaussian normal approximation for the finite-block-length (FBL) regime,

$$R_{u,s}^{(r)} \approx B_0 \left[\log_2 \left(1 + \gamma_{u,s}^{(r)} \right) - \sqrt{\frac{V(\gamma_{u,s}^{(r)})}{L} \frac{Q^{-1}(\varepsilon)}{\ln 2}} \right], \quad (9)$$

where L is the block-length, ε is the target decoding error probability, $Q^{-1}(\cdot)$ is the inverse Gaussian Q -function, and $V(\gamma)$ is the channel dispersion term (as defined in the corresponding paper). This model supports reliability-aware connectivity definitions (e.g., successful decoding probability $1 - \varepsilon$). As $L \rightarrow \infty$, the penalty term vanishes, and the expression converges to the Shannon capacity; for small L , the achievable rate is strictly lower than the Shannon bound.

3.6 Clustered Cooperative Multi-Cell Model for Learning-Based Control

Publication (V) introduces a cluster-based model to keep learning-based control tractable in O-RAN. Cells are partitioned into clusters (e.g., of size $|\mathcal{C}| = 3$) and controlled at a near-real-time interval T_{ctrl} . In addition to SINR/rate, to support latency-aware scheduling in the RL setting, the system state is extended with per-user uplink queue dynamics to capture latency and stability considerations:

$$Q_{u,s}(t+1) = \max \left(0, Q_{u,s}(t) + A_{u,s}(t) - T_{\text{ctrl}} \sum_{r \in \mathcal{R}} a_{u,s}^{(r)}(t) R_{u,s}^{(r)}(t) \right), \quad (10)$$

where $A_{u,s}(t)$ denotes arrivals (bits) for UE u in cell s at time t . This cross-layer extension enables multi-objective policy design that accounts for throughput, energy, and delay.

3.7 Constraints and Underlying Assumptions

Constraints common to the papers

Although this thesis considers multiple system-model variants (baseline uplink, non-cooperative multi-cell, cooperative multi-cell, and learning-based clustered control), the

scheduling and power-control problems share a common constraint structure. These constraints ensure physical feasibility, decodability (especially under NOMA/SIC), and fair comparability across Chapters (4–6).

C1: Binary scheduling decision.

A resource assignment indicator is binary:

$$a_{u,s}^{(r)} \in \{0, 1\}, \quad \forall s \in \mathcal{S}, u \in \mathcal{U}_s, r \in \mathcal{R}. \quad (\text{C1})$$

C2: Per-RB multiplexing constraint (OMA/NOMA).

Each RB can serve at most one UE in OMA, and at most $K_{\max}$ UEs in NOMA:

$$\sum_{u \in \mathcal{U}_s} a_{u,s}^{(r)} \leq \begin{cases} 1, & \text{OMA} \\ K_{\max}, & \text{NOMA} \end{cases} \quad \forall s \in \mathcal{S}, r \in \mathcal{R}. \quad (\text{C2})$$

This constraint reflects the limits on receiver and SIC complexity and is therefore necessary for implementation.

C3: Per-UE allocation limit.

To bound control overhead and reflect practical uplink scheduling granularity, each UE can be restricted to a maximum number of RBs per scheduling interval:

$$\sum_{r \in \mathcal{R}} a_{u,s}^{(r)} \leq R_{u,s}^{\max}, \quad \forall s \in \mathcal{S}, u \in \mathcal{U}_s, \quad (\text{C3})$$

where $R_{u,s}^{\max} = 1$ is a common simplifying choice in massive access studies (and can be relaxed if needed).

C4: Transmit power feasibility.

Transmit power is non-negative and bounded (either per-RB or total per-UE, depending on the chapter/paper):

$$0 \leq p_{u,s}^{(r)} \leq p_{\max}, \quad \forall s, u, r, \quad (\text{C4a})$$

$$\sum_{r \in \mathcal{R}} a_{u,s}^{(r)} p_{u,s}^{(r)} \leq p_{\max}, \quad \forall s, u. \quad (\text{C4b})$$

The total-power form (C4b) is often preferred for battery-powered mMTC devices.

C5: QoS feasibility via SINR threshold.

Scheduled transmissions must satisfy a minimum SINR requirement. A compact way to write this is:

$$\gamma_{u,s}^{(r)} \geq a_{u,s}^{(r)} \gamma_{\min}, \quad \forall s, u, r, \quad (\text{C5})$$

so that the constraint is active only when $a_{u,s}^{(r)} = 1$.

C6: NOMA reception assumption (uplink SIC at the receiver).

In uplink power-domain NOMA, multiple UEs may share the same RB, and the receiver (gNB) is assumed to perform SIC. In this thesis, SIC is treated as a *receiver-side capability* rather than as an explicit optimization constraint; i.e., we do not model SIC stages, decoding-order constraints, or residual cancellation errors. Instead, NOMA feasibility is ensured through (i) the multiplexing bound $K_{\max}$ in (C2), (ii) power feasibility in (C4a)–(C4b), and (iii) QoS enforcement via the SINR threshold in (C5), together with the structured user grouping and power-control rules described in Chapters (4–5).

SIC feasibility enforced via (C2), (C4a)–(C5), and structured user grouping/power rules. (C6)

C7: Reliability constraint under FBL.

For short-packet mMTC modeling (Chapter (5)), reliability is incorporated via an error probability constraint:

$$\epsilon_{u,s}^{(r)} \leq \epsilon_{\max}, \quad \forall s, u, r, \quad (C7)$$

where $\epsilon_{u,s}^{(r)}$ is linked to the FBL rate model in Eq. (9). This constraint is necessary because Shannon-rate feasibility can be overly optimistic for short packets.

C8: RL feasibility (Chapter (6)).

In the RL/MORL setting, joint RB-assignment and power-allocation actions must satisfy the physical feasibility constraints C1–C5 at every control step. When the FBL reliability model is active, C7 is enforced in addition. Queue and backlog evolution follow Eq. (10).

$$\begin{aligned} \mathbf{a}_t &\text{ satisfies (C1)–(C5)} \\ Q_{u,s}(t) &\text{ evolves per Eq. (10).} \end{aligned} \quad (C8)$$

3.8 Underlying assumptions

The following assumptions are shared across the thesis system models to ensure tractability, comparability across contributions, and relevance to massive connectivity deployments.

(A1) Single-antenna uplink abstraction (SISO). Devices are modeled as single-antenna transmitters, and the receiver as a single-stream uplink endpoint per cell. This isolates the core bottlenecks of massive access-RB scarcity, interference coupling, and power control, keeping cross-paper comparisons consistent.

(A2) Quasi-static block fading within a scheduling interval. Channel gains are assumed to be constant within a scheduling interval (slot/sub-frame/control interval T_{ctrl}) and may change independently across intervals. This is required so that scheduling decisions based on current channel estimates remain valid for the full duration of one interval, and so that SINR feasibility constraint (C5) can be evaluated consistently.

(A3) Interference treated as additive and Gaussian. Intra-cell and inter-cell interference contributions are summed in the SINR denominator (Eqs. (3) and (8)). This standard abstraction is necessary to enable analytic optimization formulations and to support scalable simulation-based evaluation.

(A4) Scheduling time alignment across cells. Multi-cell models assume that neighbouring cells' scheduling intervals are aligned, so that RB collisions and ICI on RB r are well-defined. This does not require perfect synchronization in practice, but it is necessary to define interference coupling and coordination objectives in a tractable way.

3.9 Model Evolution

These model stages form a coherent ladder: from baseline uplink scheduling, to non-cooperative interference exposure, to coordinated and reliability-aware control capable of addressing the thesis objectives of EE and massive connectivity at scale. Chapters (4 –6) build upon this foundation, each employing the appropriate model family.

Stage 1: Uplink Multi-cell model to isolate scheduling trade-offs

The thesis begins from a baseline uplink multi-cell abstraction (Section 3.3) to quantify the fundamental trade-off between OMA and NOMA under massive connectivity: NOMA increases multiplexing by allowing multiple UEs per RB (bounded by $K_{\max}$), while requiring careful interference and power management to preserve SINR and EE. At this stage, inter-cell interference may be simplified or bounded to highlight the impacts of *intra-cell scheduling and power allocation*.

Stage 2: Non-cooperative multi-cell model to expose the cost of uncoordinated decisions

The second stage introduces explicit multi-cell effects without coordination (Section 3.4). The core question is: *how severely does massive uplink activity degrade performance when each cell schedules independently?* A tractable worst-case interference abstraction (Eq. (6)) makes it possible to evaluate connectivity and EE under realistic interference scaling, while still allowing algorithmic development and interpretation.

Stage 3: Cooperative multi-cell models to answer sustainability questions

The final stage shifts to cooperative multi-cell models (Section 3.5), because the thesis asks *What coordination mechanisms are effective in O-RAN to mitigate inter-cell interference while scaling to massive users?*

- The reliability-aware FBL model (Eq. (9)) is needed to evaluate *sustainable massive connectivity* under short-packet mMTC conditions where Shannon rates are optimistic and decoding error probability matters.
- The clustered RL model (Eq. (10)) is required to study *multi-objective control policies* in near-real-time O-RAN settings, where scalability and delay/queue dynamics are first-class constraints.

Together, the three model families form a coherent ladder: from baseline uplink scheduling, to non-cooperative interference exposure, to coordinated and reliability-aware control capable of addressing the thesis objectives of EE and massive connectivity at scale.

3.10 Constraints Reference Summary

To support quick cross-referencing in Chapter (4–6), Table 5 summarizes the common constraints used throughout this thesis. Unless otherwise stated, symbols follow Table 4, and constraints follow Section 3.7.

Table 5: Reference summary of common constraints used across Chapters (4–6). For symbol definitions see Table 4.

ID	Label	Statement / equation	Active in
C1	Binary scheduling	$a_{u,s}^{(r)} \in \{0, 1\}, \forall s, u, r. \quad (C1)$	Ch. 4–6
C2	Per-RB multiplexing	$\sum_u a_{u,s}^{(r)} \leq 1 \text{ (OMA) or } \leq K_{\max} \text{ (NOMA)}, \forall s, r. \quad (C2)$	Ch. 4–6
C3	Per-UE RB limit	$\sum_r a_{u,s}^{(r)} \leq R_{u,s}^{\max}, \forall s, u. \quad (C3)$	Ch. 4–6
C4	Power feasibility	$0 \leq p_{u,s}^{(r)} \leq p_{\max}, \forall s, u, r. \quad (C4a)–(C4b)$	Ch. 4–6
C5	QoS via SINR threshold	$\gamma_{u,s}^{(r)} \geq a_{u,s}^{(r)} \gamma_{\min}, \forall s, u, r. \quad (C5)$	Ch. 4–6
C6	NOMA/SIC decodability	Enforced via (C2)–(C5), and structured user grouping (C6).	Ch. 4–5
C7	FBL reliability	$\varepsilon_{u,s}^{(r)} \leq \varepsilon_{\max}, \forall s, u, r, \text{ linked to (9).} \quad (C7)$	Ch. 5 only
C8	RL feasibility	Actions satisfy (C1)–(C5) ; queue evolves per Eq. (10) (C8).	Ch. 6 only

Usage in later chapters: Chapters (4–6) refer to constraints (C1)–(C5) as the baseline feasibility set; Chapter (5) additionally uses (C7) for short-packet reliability; Chapter (6) uses (C8) to ensure learned policies remain physically feasible and comparable to analytical baselines.

4 Coordination and Scheduling in O-RAN Architecture

4.1 Chapter Overview and Motivation

This chapter is based on Publications (I) and (II). It instantiates the *baseline uplink O-RAN model* from Section 3. The *Inter-O-DU coordination* per-RB scheduling decisions and SINR/rate evaluation follow Sections 3.3 and 3.4 (Eqs. (2)–(3)), while O-DU coordination is realized via the explicit decision-dependent ICI of Section 3.5 (Eq. (7)).

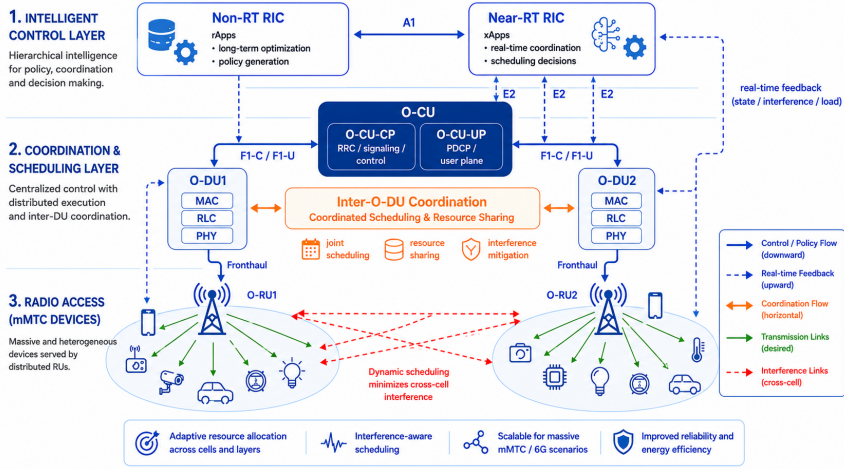


Figure 3: Conceptual O-RAN coordination and scheduling architecture for massive mMTC access. The figure illustrates hierarchical control across the Non-RT RIC, Near-RT RIC, O-CU, and distributed O-DUs, where inter-O-DU coordination supports scheduling and resource-sharing decisions to mitigate cross-cell interference. Real-time feedback from the radio access layer enables adaptive resource allocation, improved reliability, and energy-efficient operation in dense 6G deployments.

As illustrated in Fig 3, O-RAN disaggregates the traditional monolithic base station into an Open Distributed Unit (O-DU), an Open Central Unit (O-CU), and an Open Radio Unit (O-RU), interconnected through standardized open interfaces. The Non-RT RIC (rApps) and Near-RT RIC (xApps) add hierarchical intelligence via the A1 and E2 interfaces, operating on timescales above 1 s and between 10 ms and 1 s, respectively. This architecture enables programmable scheduling policies that were impractical in proprietary monolithic deployments. However, in ultra-dense mMTC scenarios where adjacent cells frequently co-schedule users on the same RB, each O-DU's independent scheduling decisions generate decision-dependent ICI that cannot be mitigated without cross-cell coordination. This chapter addresses **RQ 1** by developing and evaluating inter-DU coordination for uplink scheduling under massive connectivity, and by establishing that the computational cost of scheduling intelligence must be treated as a sustainability parameter alongside radio energy.

4.2 Related Work and Positioning

Chapter 2 (Sections 2.2.1) surveys the O-RAN architecture, programmable RIC-based control, and the energy-efficiency literature in detail. The present chapter builds on that foundation

and addresses two specific gaps left open by the surveyed works.

First, despite substantial progress in O-RAN specification and xApp deployment [117, 118, 37], no published work had proposed an inter-O-DU coordination mechanism that explicitly targets *decision-dependent* ICI in a multi-cell NOMA uplink while remaining compatible with O-RAN’s distributed architecture—that is, without joint transmission or centralized channel knowledge (Section 2.2.1). Second, as identified in Section 2.2.2, evaluations of NOMA-based scheduling in O-RAN had not accounted for the computational energy consumed by scheduling algorithms executing on edge hardware, leaving a benchmarking gap that prevents energy-aware algorithm selection for RIC-hosted schedulers.

Publication (I) fills the first gap by proposing an inter-O-DU architecture in which neighbouring O-DUs share forward-looking scheduling tables so that each scheduler can predict and reduce the interference it imposes on adjacent cells before finalizing its allocations. Publication (II) fills the second gap by developing an instruction-level complexity and energy measurement framework that decomposes total scheduling energy into computational and radio components, enabling a hardware-grounded sustainability comparison of OMA and NOMA schedulers.

4.3 Methodology: Inter-O-DU Coordination (Publication (I))

We consider a multi-cell uplink with massive device density and limited RBs. Each O-DU must jointly select binary RB assignments $a_{u,s}^{(r)}$ and, when applicable, transmit powers $p_{u,s}^{(r)}$, to meet QoS constraints (C5) under a power budget (C4b) while managing three coupled sources of performance degradation: (i) intra-cell interference from NOMA co-scheduling (bounded by $K_{\max}$, constraint C2); (ii) inter-cell interference from neighbouring O-DUs co-scheduling users on the same RB; and (iii) the computational energy cost of the scheduling algorithm itself, which must be bounded for near-real-time O-RAN execution. Crucially, ICI is decision-dependent in this setting: it is not a fixed background term but a quantity that improves as scheduling decisions across cells are coordinated. This coupling is precisely what makes inter-O-DU coordination valuable and what distinguishes this problem from the non-cooperative baseline established in Section 3.4.

Publication (I) proposes an inter-O-DU coordination architecture in which each O-DU shares its scheduling table with adjacent O-DUs before finalizing resource allocations. The scheduling table records each candidate user, the RB it would occupy, and the associated channel gain. Using the received tables from neighbors, each O-DU’s scheduler computes, for every candidate user-RB pair, the predicted inter-cell interference that the user would experience given the current candidate allocations in adjacent cells (Eq. (7)). User-RB combinations that generate lower predicted ICI are prioritized, while those that would impose severe mutual interference on neighbouring cells are deferred. Once allocations are finalized, power coefficients are assigned by user channel quality to ensure SIC decodability at the receiver: users with good channel conditions (close to the serving O-DU) are assigned 14 dBm, middling-channel users 20 dBm, and poor-channel users the full 23 dBm. This structured power assignment ensures sufficient received-signal separation for successive interference cancellation without requiring exhaustive power optimization.

The coordination mechanism differs from Coordinated Multi-Point (CoMP) in a critical

respect: each O-DU serves only its own users, and no joint transmission is performed. The shared information is limited to forward-looking scheduling, specifically, which users are candidates for which RBs in the next allocation interval, making the approach lightweight and compatible with the E2 interface latency budget.

4.4 Methodology: Computational Energy Modeling (Publication (II))

Publication (II) develops an energy consumption evaluation framework for scheduling algorithms running on edge hardware in an O-RAN-aligned setting. The motivation is that, as schedulers migrate from proprietary hardware to software running on commercial-off-the-shelf processors, the energy consumed by algorithm execution becomes non-negligible relative to radio transmission energy, particularly at the near-real-time control cadence of the Near-RT RIC.

The framework decomposes the total energy per scheduling interval into two additive components:

$$E_{\text{total}} = E_{\text{radio}} + E_{\text{compute}}, \quad (11)$$

where E_{radio} is the aggregate transmit energy across all scheduled users and RBs, and E_{compute} is derived from an instruction-level complexity analysis. For each scheduling algorithm, the number of floating-point operations and memory accesses per scheduling interval is counted, and an energy-per-instruction coefficient-calibrated to representative edge-processor specifications-is applied to obtain a joule-per-interval computational cost. This decomposition enables a hardware-grounded comparison of OMA, the proposed EERS-MTCN NOMA scheduler, and the exhaustive-search ESAP-NOMA baseline that would not be possible from radio KPIs alone.

An important and counter-intuitive finding of this methodology is that a simpler algorithm with lower per-interval computational energy is not automatically the better sustainable choice: the joint metric E_{total} must account for the radio EE improvement delivered by the more computationally demanding algorithm. This nuance is explored in the results below.

4.5 Results and Discussion

Results: Uplink Throughput

Fig. 4 shows the cumulative distribution function (CDF) of uplink throughput for coordinated and stand-alone O-DU scenarios under both OMA and NOMA. The coordination architecture consistently improves throughput at every percentile. At the 50th percentile, OMA with inter-O-DU coordination achieves approximately 115 kbps compared to roughly 80 kbps for stand-alone OMA, a gain of approximately 43%. At the same percentile, coordinated NOMA outperforms its stand-alone counterpart by approximately 75 kbps, with a maximum throughput gain of 30%. The figure also reveals that OMA achieves higher *per-user* throughput than NOMA in absolute terms, because OMA assigns each RB exclusively to one user, eliminating intra-cell NOMA interference and achieving a higher per-user SINR. However, OMA does so at the cost of serving fewer simultaneous users, which is the fundamental connectivity trade-off explored next.

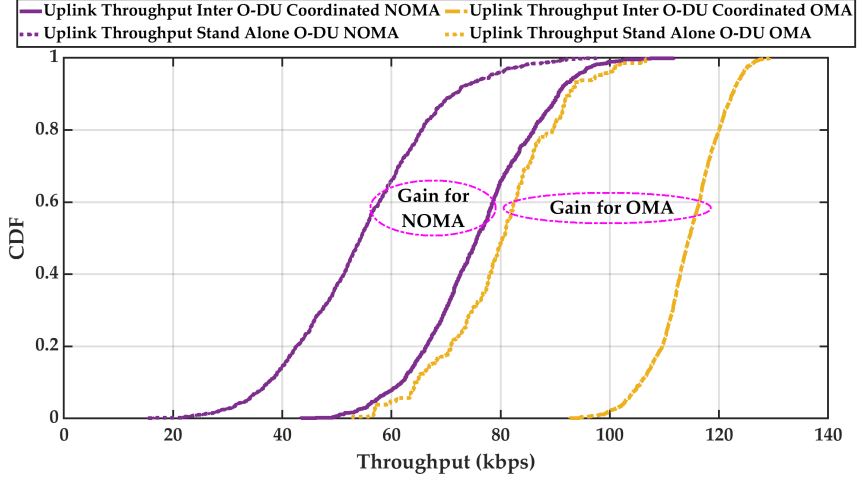


Figure 4: Cumulative distribution function (CDF) of uplink throughput for Inter O-DU coordinated and stand-alone O-DU scenarios under NOMA and OMA schemes. The figure highlights the throughput gains achieved through inter O-DU coordination and illustrates the performance advantage of NOMA over conventional OMA across different percentile levels. Reproduced from our (Publication I).

Energy consumption and connected devices:

Coordination also reduces energy consumption across both access schemes: at the median, coordinated OMA achieves approximately 40% lower energy expenditure than stand-alone OMA, while coordinated NOMA reduces energy consumption by approximately 10% relative to its uncoordinated counterpart. Regarding device connectivity, both coordinated and uncoordinated NOMA configurations support approximately 1.5 times as many simultaneously connected devices as OMA does, for both cell-center and cell-edge users. The smaller relative energy gain under NOMA compared to OMA reflects the additional interference management burden of intra-cell NOMA multiplexing, which coordination does not address. Energy CDF curves and the connected-device breakdown by user location are presented in Publication (I) as in (Appendix 1).

Computational energy and the OMA-NOMA trade-off:

Fig. 5 shows the combined computational plus radio energy consumption of OMA, RAAP-NOMA, EEPS-MTCN-NOMA, and ESAP-NOMA across user densities from 16 to 48 active devices per cell. The results reveal an important nuance that would be invisible if only radio KPIs were considered. OMA has a lower per-iteration computational energy than EERS-MTCN, because OMA's scheduling logic is simpler and requires fewer instructions per interval; this is not a flaw in the analysis but an accurate reflection of the lower algorithmic complexity of orthogonal scheduling.

However, EEPS-MTCN achieves a 27% EE improvement over RAAP-NOMA and delivers superior combined E_{total} across all evaluated densities, because its radio energy savings from interference-aware user grouping and power allocation more than offset the additional computational cost. Conversely, the ESAP-NOMA exhaustive-search baseline has an execution time approximately 10 times longer than EEPS-MTCN, making it computationally

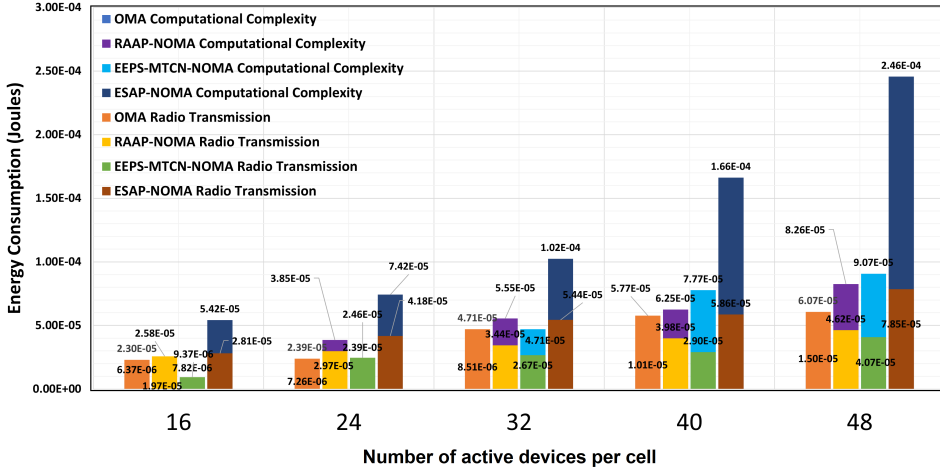


Figure 5: Combined energy consumption analysis of computational complexity and radio transmission for OMA, RAAP-NOMA, EEPs-MTCN-NOMA, and ESAP-NOMA schemes under varying user connection densities (16–48 active devices per cell). The results demonstrate the scalability impact of increasing user density on total energy expenditure, and highlight the relative efficiency of advanced NOMA-based approaches compared to conventional OMA. Reproduced from our Publication (II).

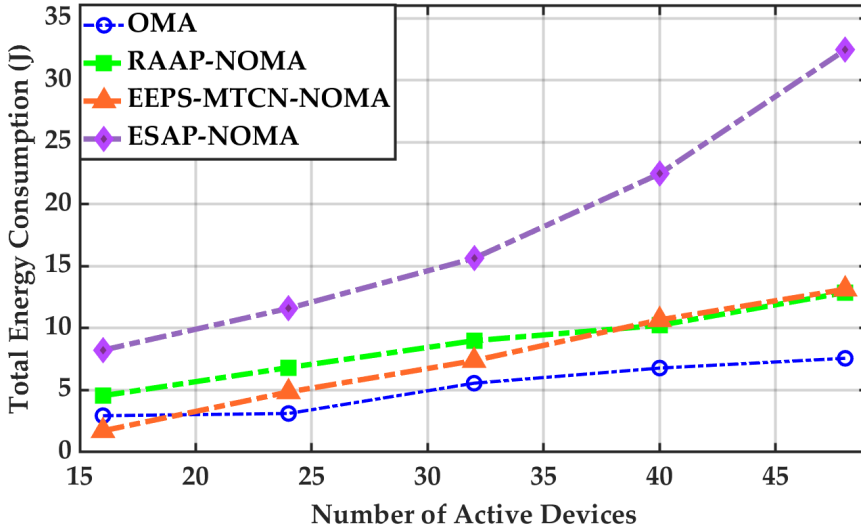


Figure 6: Total energy consumption, decomposed into computational complexity and radio transmission, for OMA, RAAP-NOMA, EEPs-MTCN-NOMA, and ESAP-NOMA as a function of the number of active devices per cell. The figure shows that total energy increases with network load for all schemes, while the exhaustive-search baseline incurs the highest cost because of its substantially larger computational burden. The results highlight the need to evaluate sustainability using the combined energy cost rather than radio performance alone.

impractical for near-real-time O-RAN deployment despite its marginally better radio EE (approximately 15% better than EERS-MTCN in radio EE alone). The crucial point is that sustainability evaluation in O-RAN must consider E_{total} Eq. (11), not radio metrics in isolation: a scheduler that appears optimal by radio KPIs may be unsustainable when computational execution cost is included.

Fig. 6 complements the radio-performance analysis by quantifying the full energy cost of each scheduling strategy, where total energy is obtained as the sum of computational complexity and radio transmission cost. This is important because a scheduler that appears favorable in radio-only terms may still be inefficient once algorithmic execution cost is included. Across all active-device populations, the total energy consumption increases monotonically, confirming that the scheduling problem becomes more demanding as the system load grows. OMA remains comparatively lightweight because of its simpler decision structure, whereas RAAP-NOMA and EEPS-MTCN-NOMA incur higher cost due to more involved user-grouping and power-allocation logic. ESAP-NOMA exhibits the steepest increase and the highest total energy consumption, which is consistent with the overhead of exhaustive search. Taken together, the figure supports the thesis argument that sustainable RAN design must be evaluated using the aggregate energy budget rather than radio metrics in isolation.

Discussion

The results support three conclusions relevant to RQ1. Even without joint transmission, interference-aware scheduling-table exchange between neighbouring O-DUs improves massive-uplink scheduling outcomes in O-RAN. The coordination gain is most pronounced for OMA, where the elimination of intra-cell interference leaves inter-cell ICI as the dominant residual impairment; for NOMA, both impairments are present, and coordination addresses only the inter-cell component. Algorithm selection in O-RAN changes not only radio KPIs but also computational energy and execution time, which directly affect practical deployment under Near-RT RIC latency constraints. A scheduling algorithm that achieves marginally better radio EE at 10 times higher execution cost, as illustrated by ESAP-NOMA, is not a viable candidate for real-time O-RAN deployment. Coordination reduces ICI but does not eliminate it, and the Shannon rate model used throughout this chapter assumes packet lengths long enough for the normal approximation to hold. In ultra-dense mMTC deployments, devices transmit short packets, making Shannon-based rate expressions overly optimistic and necessitating a reliability-aware model that accounts for finite-block-length (FBL) penalties and non-negligible decoding error probabilities. Chapter (5) introduces joint EE-connectivity optimization first under a non-cooperative interference model under Shannon IFBL and the cooperative interference model under FBL constraints, directly addressing this residual limitation.

4.6 Chapter Summary

This chapter developed two O-RAN-aligned contributions to address **RQ1**. Publication (I) demonstrated that lightweight, scheduling-table-exchange coordination between adjacent O-DUs reduces decision-dependent ICI without joint transmission: at the 50th percentile,

the coordinated OMA architecture achieves 43%, higher uplink throughput and 40% lower energy consumption than the stand-alone baseline, while coordinated NOMA delivers a 30% throughput gain and 10% energy reduction; both NOMA configurations support approximately 1.5 times more simultaneously connected devices than their OMA counterparts for the same spectrum resources.

Publication (II) showed that scheduling algorithm choice is a sustainability parameter in O-RAN: EERS-MTCN-NOMA achieves a 27% EE improvement over RAAP-NOMA and delivers 4 times lower computational energy than the exhaustive-search ESAP-NOMA baseline across all evaluated user densities (16–48 active devices per cell), while OMA-though computationally simpler-offers lower combined E_{total} only when radio efficiency is disregarded. Together, these results establish that sustainable massive-connectivity scheduling in O-RAN requires joint optimization of interference coordination, radio efficiency, and algorithmic execution cost. Chapter (5) builds on the cooperative interference model established here and extends the optimization framework to include short-packet reliability constraints, addressing the practical mMTC regime in which Shannon-capacity assumptions become invalid.

5 Sustainable Connectivity Optimization under 6G Constraints

5.1 Overview, Motivation, and Chapter Scope

This chapter develops sustainable resource scheduling methods for dense uplink NOMA networks based on Publications (III) and (IV). Publication (III) establishes the non-cooperative Energy-Efficient Resource Scheduling for Machine-Type Communications using NOMA (EERS-MTCN) framework. It considers a multi-cell uplink scenario in which each cell schedules its users independently while interference generated by neighboring cells is represented by a conservative fixed interference term. Publication (IV) extends this foundation through the Sustainable and Reliable Scheduling for Machine-Type Communications in NOMA (SRS-MTCN) framework, which explicitly models inter-cell interference coupling, finite blocklength (FBL) transmission, and decoding reliability in a cooperative multi-cell setting.

Chapter (4) demonstrated that coordination among distributed radio units can mitigate decision-dependent inter-cell interference. However, two limitations of that analysis motivate the present chapter. First, energy efficiency (EE) and connectivity were considered primarily as separate performance objectives rather than as jointly coupled design goals. Second, the earlier analysis relied on the Shannon-capacity rate model, which assumes arbitrarily long codewords and negligible decoding error. These assumptions are not appropriate for massive machine-type communication (mMTC), where devices generally transmit short packets under strict energy and reliability constraints.

Under FBL transmission, the achievable rate is lower than the corresponding Shannon rate and depends explicitly on block length, received SINR, and target decoding error probability. Consequently, maximizing the Shannon-based sum rate does not necessarily produce an energy-efficient or reliable scheduling decision for short-packet traffic. This chapter therefore investigates how EE and reliable connectivity can be jointly improved while accounting for both intra-cell and inter-cell interference.

The chapter addresses three related gaps. First, existing multi-cell NOMA studies frequently optimize EE while treating connectivity as a constraint, an indirect consequence of resource allocation, or a secondary metric. Second, many non-cooperative formulations approximate neighboring-cell interference as fixed background interference and therefore do not capture the decision-dependent coupling introduced by simultaneous scheduling and power-control decisions across cells. Third, only a limited number of studies jointly consider EE, connectivity, FBL reliability, and interference-aware multi-cell scheduling within one framework.

Publication (III) addresses the first gap and partially addresses the second by formulating EE and connectivity objectives under a fixed neighboring-cell interference approximation. Publication (IV) completes this progression by introducing explicit multi-cell interference coupling and cooperative FBL-aware scheduling and power refinement.

The chapter is organized as follows. Section 5.2 positions the two publications relative to the state of the art. Section 5.3 presents the unified system perspective and distinguishes the non-cooperative and cooperative configurations. Section 5.3.2– 5.3.4 introduces the analytical formulations used in the two publications and explains their relationship in

5.3.5. Section 5.4 describes EERS-MTCN and SRS-MTCN and compares their computational requirements. The corresponding results are discussed in Section 5.5, followed by the chapter summary and transition to multi-objective reinforcement learning (MORL).

5.2 Related Work and Positioning

The literature relevant to this chapter can be grouped into four categories: energy-efficient NOMA allocation, connectivity-oriented scheduling, and FBL communication, multi-cell interference coordination.

Energy-efficient resource allocation in multi-cell NOMA has been addressed through optimization and learning-based methods. For example, [56] studies EE-oriented resource allocation in heterogeneous small-cell NOMA, while [58] employs deep reinforcement learning for RB assignment and power control. These approaches improve EE, but connectivity is generally treated through QoS or admission constraints, and FBL reliability is not jointly incorporated.

Connectivity-oriented scheduling has also been examined in NB-IoT and mMTC. The cooperative scheduler in [54] improves massive connectivity through interference-aware scheduling across cells, but it does not formulate EE as a co-primary objective and does not account for FBL decoding reliability. Consequently, existing studies provide important partial solutions but do not simultaneously address EE, reliable connectivity, FBL transmission, and explicit multi-cell interference coupling.

The finite-blocklength framework introduced in [21] provides the theoretical basis for quantifying the rate penalty and non-zero decoding error probability associated with short packets. Subsequent studies have applied FBL theory to single-cell NOMA, short-packet IoT, URLLC, reconfigurable intelligent surface-assisted links, and cell-free massive MIMO. Examples include [82, 83, 89, 119]. Although these works capture short-packet reliability, most focus on rate, latency, outage, or EE in single-cell or weakly coupled settings and do not jointly optimize reliable connectivity and multi-cell scheduling.

To the best of our knowledge, prior work had not integrated these dimensions within a scalable cooperative uplink NOMA framework for dense mMTC. Publication (III) establishes a computationally tractable non-cooperative baseline under a conservative fixed-ICI approximation. Publication (IV) then introduces a network-wide cooperative formulation in which inter-cell interference is recomputed as a function of the scheduling and power decisions, while FBL rate and decoding-error constraints are incorporated directly into the optimization process.

5.3 System Perspective and Analytical Problem Formulation

The two publications share a common uplink NOMA setting but differ in how inter-cell interference and short-packet transmission are modeled. Consider a network containing a set of cells $\mathcal{S}$. Cell $s \in \mathcal{S}$ serves a set of users $\mathcal{U}_s$ over a common set of RBs $\mathcal{R}$. The binary variable $a_{u,s}^r \in \{0, 1\}$ indicates whether user u in cell s is assigned RB r , and $p_{u,s}^r$ denotes its uplink transmit power. At most M users may share one RB through power-domain NOMA. For each simulation scenario, a fixed number of active devices, denoted by N , is

randomly distributed within each cell. Throughout this thesis, N refers to the number of active devices per cell rather than a physical spatial density expressed in devices/m².

This chapter instantiates two configurations which are described below, from the system model in Chapter (3). Publication (III) uses the non-cooperative model (Section 3.4) with a bounded ICI approximation. Publication (IV) uses the cooperative FBL model (Section 3.5) with an explicit, ICI sum and FBL rate $R_{u,s}^{(r)}$ as defined in Eq. 9.

5.3.1 Non-Cooperative Fixed-ICI Model

Publication (III) considers a non-cooperative multi-cell configuration in which each cell performs scheduling and power allocation independently. Neighboring-cell interference is not jointly coordinated. Instead, it is represented by the fixed approximation

$$I_s^{r,\text{fix}} = MN_{\text{AC}}P_{\text{max}}G_{\text{AC}}, \quad (12)$$

where N_{AC} is the number of adjacent cells, P_{max} is the maximum UE transmit power, and G_{AC} is the mean interfering channel gain from neighboring cells.

The SINR of user u in cell s on RB r is therefore written as

$$\gamma_{u,s}^{r,\text{EERS}} = \frac{a_{u,s}^r |h_{u,s}^r|^2 p_{u,s}^r}{\sum_{\substack{v \in \mathcal{U}_s \\ v \neq u}} a_{v,s}^r |h_{v,s}^r|^2 p_{v,s}^r + I_s^{r,\text{fix}} + \sigma_N^2}. \quad (13)$$

This model captures the performance penalty produced by neighboring cells without implying that their scheduling decisions are exchanged or coordinated. Accordingly, EERS-MTCN is interference-aware but not inter-cell cooperative.

5.3.2 Publication III Problem Formulation

The system EE is

$$\eta_{\text{EE}}^{\text{EERS}} = \frac{\sum_{s \in \mathcal{S}} \sum_{u \in \mathcal{U}_s} \sum_{r \in \mathcal{R}} a_{u,s}^r B_r \log_2 (1 + \gamma_{u,s}^{r,\text{EERS}})}{P_c + \sum_{s \in \mathcal{S}} \sum_{u \in \mathcal{U}_s} \sum_{r \in \mathcal{R}} a_{u,s}^r p_{u,s}^r}. \quad (14)$$

Connectivity is represented by the number or density of scheduled users that satisfy the minimum SINR and rate requirements:

$$\kappa_{\text{EERS}} = \kappa_{\text{scale}} \sum_{s \in \mathcal{S}} \sum_{r \in \mathcal{R}} \sum_{u \in \mathcal{U}_s} a_{u,s}^r \mathbf{1} \{ \gamma_{u,s}^{r,\text{EERS}} \geq \gamma_{\text{th}}, R_{u,s}^r \geq R_{\text{min}} \}. \quad (15)$$

Publication (III) formulates EE and connectivity objectives under a common set of power, RB-assignment, multiplexing, and QoS constraints. The design intent can be summarized as

$$\max_{\mathbf{A}, \mathbf{P}} \{ \eta_{\text{EE}}^{\text{EERS}}, \kappa_{\text{EERS}} \}, \quad (16)$$

subject to

$$\text{Constraints (C1)-(C4b),} \quad (17)$$

The formulation is a mixed-integer non-convex problem because of binary RB assignments, the fractional EE objective, and intra-RB NOMA interference.

5.3.3 Cooperative FBL-Aware Model

Publication (IV) replaces the fixed neighboring-cell term with an explicit inter-cell interference summation. The SINR becomes

$$\gamma_{u,s}^{r,\text{SRS}} = \frac{a_{u,s}^r |h_{u,s}^r|^2 p_{u,s}^r}{\underbrace{\sum_{\substack{v \in \mathcal{U}_s \\ v \neq u}} a_{v,s}^r |h_{v,s}^r|^2 p_{v,s}^r}_{\text{intra-cell interference}} + \underbrace{\sum_{\substack{t \in \mathcal{S} \\ t \neq s}} \sum_{w \in \mathcal{U}_t} a_{w,t}^r |h_{w,t}^r|^2 p_{w,t}^r}_{\text{inter-cell interference}} + \sigma_N^2}. \quad (18)$$

Unlike (12), the inter-cell term in (18) depends directly on the RB assignments and transmit powers selected in all neighboring cells. The scheduling problem is therefore network-wide and cannot be decomposed into independent per-cell problem. Therefore, achievable FBL rate is modeled as

$$R_{u,s}^{r,\text{FBL}} = B_r \left[\log_2 (1 + \gamma_{u,s}^{r,\text{SRS}}) - \sqrt{\frac{V(\gamma_{u,s}^{r,\text{SRS}})}{L} \frac{Q^{-1}(\varepsilon_{u,s}^r)}{\ln 2}} \right], \quad (19)$$

where L is the blocklength, $\varepsilon_{u,s}^r$ is the decoding error probability, $V(\gamma)$ is the channel dispersion, and B_r is the RB bandwidth. As $L \rightarrow \infty$, the FBL penalty vanishes and (19) converges to the Shannon rate.

5.3.4 Publication IV Problem Formulation

The system EE is computed using the achievable FBL rate:

$$\eta_{\text{EE}}^{\text{SRS}} = \frac{\sum_{s,u,r} a_{u,s}^r R_{u,s}^{r,\text{FBL}}}{P_c + \sum_{s,u,r} a_{u,s}^r p_{u,s}^r}. \quad (20)$$

Connectivity is reliability weighted and is expressed as

$$\kappa_{\text{SRS}} = \kappa_{\text{scale}} \sum_{s \in \mathcal{S}} \sum_{r \in \mathcal{R}} \sum_{u \in \mathcal{U}_s} a_{u,s}^r (1 - \varepsilon_{u,s}^r), \quad (21)$$

Because EE and connectivity have different physical scales, normalized versions of the two objectives should be used in the scalarized formulation:

$$\bar{\eta}_{\text{EE}} = \frac{\eta_{\text{EE}}^{\text{SRS}}}{\eta_{\text{EE}}^{\text{ref}}}, \quad \bar{\kappa} = \frac{\kappa_{\text{SRS}}}{\kappa^{\text{ref}}}, \quad (22)$$

where $\eta_{\text{EE}}^{\text{ref}}$ and κ^{ref} are fixed normalization constants obtained from the simulation

configuration or feasible reference bounds.

The joint optimization problem is then

$$\begin{aligned} & \max_{\mathbf{A}, \mathbf{P}} \quad \alpha \bar{\eta}_{\text{EE}} + (1 - \alpha) \bar{\kappa} \\ & \text{subject to} \quad \begin{cases} \text{Constraints (C1)–(C4b),} \\ \text{Constraint C7} \in \text{Publication (IV)} \end{cases} \end{aligned} \quad (23)$$

where $\alpha \in [0, 1]$ determines the relative priority assigned to EE and connectivity.

5.3.5 Relationship Between the Two Formulations

The two formulations represent successive levels of modeling fidelity. Publication (III) establishes the underlying scheduling trade-off using Shannon rates, SINR-based connectivity, and a fixed neighboring-cell interference approximation. Publication (IV) retains the same fundamental resource-allocation variables but replaces the fixed-ICI term with an explicit network-wide interference model and replaces SINR-only feasibility with FBL rate and decoding-reliability constraints.

Accordingly, the SRS-MTCN problem reduces conceptually toward the EERS-MTCN problem when the following limiting assumptions are introduced:

$$L \rightarrow \infty, \quad \epsilon_{u,s}^r \rightarrow 0, \quad I_{u,s}^{r,\text{inter}} \rightarrow I_s^{r,\text{fix}}. \quad (24)$$

This relationship clarifies that SRS-MTCN is not an unrelated scheduler but a cooperative and FBL-aware extension of the non-cooperative baseline.

5.4 Algorithm Design

5.4.1 EERS-MTCN

The EERS-MTCN algorithm decomposes the joint scheduling problem into resource scheduling and power allocation. During resource scheduling, users are sorted according to their channel gains and partitioned into sectors. Candidate NOMA groups are then evaluated in terms of achievable rate, EE, and intra-RB interference under the fixed neighboring-cell interference approximation in (12). Feasible groups satisfying the SINR and minimum-rate requirements are assigned to RBs according to the interference- and EE-oriented selection rule.

During power allocation, the RB assignments are fixed, and the transmit powers are iteratively adjusted over the admissible power set. A power combination is retained only when all scheduled users satisfy the minimum SINR and rate constraints. This decomposition substantially reduces the search space relative to exhaustive joint pairing and power optimization.

The dominant operations are channel-gain sorting, candidate-group evaluation, and discrete power search. If M and the number of examined power levels are bounded, the resulting implementation has substantially lower complexity than exhaustive pairing. This conclusion is also supported by the measured execution-time results, in which ESAP-NOMA requires approximately one order of magnitude more computation than EERS-MTCN.

5.4.2 SRS-MTCN

The SRS-MTCN algorithm preserves the scheduling–power decomposition while introducing cooperative interference evaluation and FBL reliability. It contains three stages.

In the first stage, users are sorted by channel quality and organized into candidate NOMA clusters. Sectorization and opposite-sector grouping reduce the number of possible user combinations and promote channel diversity within each NOMA cluster. In the second stage, each candidate cluster–RB assignment is evaluated using the network-wide SINR in (18). The scheduler therefore accounts for the interference that an assignment produces in both its own cell and neighboring cells. Candidate assignments that violate the FBL minimum-rate or decoding-error constraints are discarded. In the third stage, transmit powers are iteratively refined across the cooperating cells. Following each update, both intra-cell and inter-cell interference are recomputed. The process terminates when the objective improvement becomes smaller than a convergence tolerance or when the iteration budget is exhausted.

The cooperative formulation assumes that the scheduling relevant channel, allocation, and interference information is available to a network-wide coordinating entity or is exchanged among the cooperating schedulers within the considered coordination region.

5.4.3 Complexity Comparison

Let N denote the number of active users, R the number of RBs, S the number of cooperating cells, M the maximum NOMA order, and I the number of power-refinement iterations.

The EERS-MTCN algorithm has time complexity $\mathcal{O}(N \log N + N \cdot M)$ per scheduling interval. This is significantly lower than the exhaustive-search ESAP-NOMA baseline ($\mathcal{O}(N^3)$), making it suitable for near-real-time execution.

A general upper bound for SRS-MTCN is

$$\mathcal{O}(SNR + RN \log N + ISRM). \quad (25)$$

For fixed bandwidth, cooperation region, NOMA order, and iteration budget, R , S , M , and I are bounded by system design. Under this practical assumption, the dominant dependence on the user population is

$$\mathcal{O}(N \log N). \quad (26)$$

If the number of RBs is allowed to scale linearly with the number of users, i.e., $R = \mathcal{O}(N)$, the conservative upper bound becomes

$$\mathcal{O}(N^2 \log N). \quad (27)$$

These asymptotic results demonstrate polynomial scalability. However, near-real-time feasibility should be established using measured execution latency on the intended computing platform rather than asymptotic complexity alone.

5.5 Results and Discussion

The results are organized according to the progression from the non-cooperative EERS-MTCN baseline to the cooperative FBL-aware SRS-MTCN framework. Because the two publications use different interference and rate models, their absolute EE values should not be interpreted as a direct algorithm-to-algorithm comparison. Instead, each set of results quantifies the performance gains relative to the baselines evaluated under the same system assumptions.

5.5.1 EERS-MTCN: EE, Connectivity, and Execution Time

Publication (III) addresses the following questions:

1. How does the number of active or admitted devices affect the achievable EE under non-cooperative multi-cell operation?
2. How does the NOMA multiplexing order affect the trade-off between spectral reuse, interference, and EE?
3. Can channel-aware and locally interference-aware scheduling improve energy efficiency compared with conventional random scheduling strategies while maintaining substantially lower computational cost?

These questions were investigated in publication (III) by comparing the proposed EERS-MTCN algorithm against ESAP-NOMA and RAAP-NOMA over a range of user connection densities and NOMA multiplexing configurations.

Figure 7 compares the achievable system energy efficiency of ESAP-NOMA, the proposed EERS-MTCN, and RAAP-NOMA as the number of multiplexed NOMA users per resource block increases. All three algorithms exhibit a gradual reduction in energy efficiency with increasing multiplexing order, reflecting the fundamental trade-off between spectral efficiency and energy efficiency in uplink NOMA systems. Although multiplexing more users on the same resource block improves resource utilization and network connectivity, it also increases intra-cell interference and transmit-power requirements to satisfy the QoS constraints. Consequently, the increase in power consumption outweighs the throughput gain, leading to a progressive decline in the achievable system energy efficiency.

Among the compared schemes, ESAP-NOMA achieves the highest energy efficiency because exhaustive search identifies the globally optimal user pairing and power allocation. The proposed EERS-MTCN consistently outperforms RAAP-NOMA by incorporating interference-aware user scheduling and QoS-constrained resource allocation, thereby achieving higher throughput with more efficient resource utilization. While its energy efficiency remains slightly below the exhaustive-search upper bound, EERS-MTCN provides a substantially lower computational complexity, resulting in a favorable trade-off between energy efficiency and implementation cost for practical multi-cell NOMA deployments.

Figure 8 illustrates the achievable system energy efficiency of the proposed EERS-MTCN algorithm under different active-device populations per cell. For all network loading conditions, the achievable energy efficiency decreases as the NOMA multiplexing order increases from one to eight users per resource block. This behavior reflects the fundamental

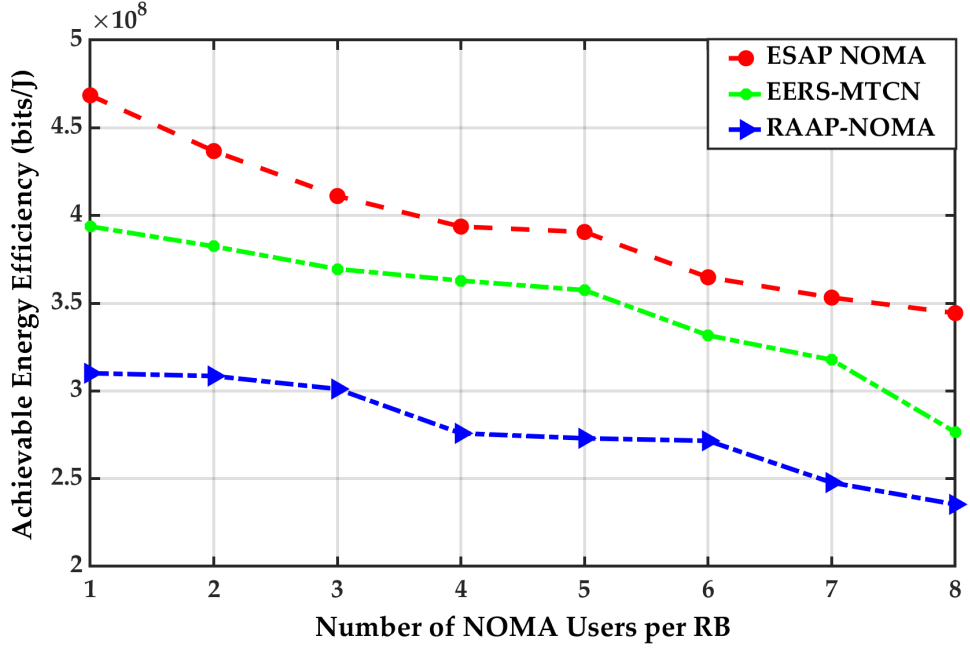


Figure 7: Comparison of achievable system energy efficiency versus the number of NOMA users multiplexed per resource block for ESAP-NOMA, EERS-MTCN, and RAAP-NOMA. ESAP-NOMA attains the highest energy efficiency, the proposed EERS-MTCN consistently outperforms RAAP-NOMA, and all schemes exhibit a decreasing trend as the NOMA multiplexing order increases.

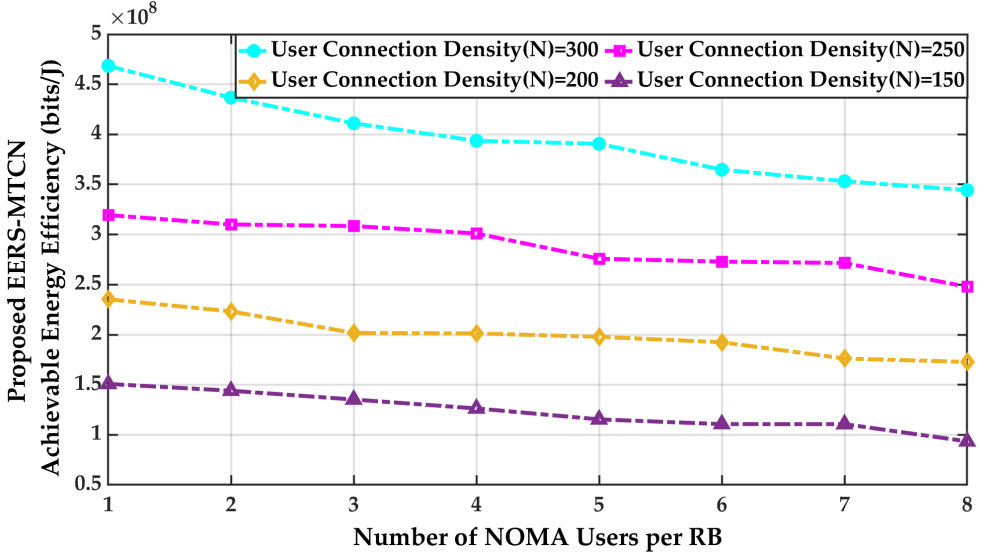


Figure 8: Achievable system energy efficiency of the proposed EERS-MTCN algorithm as a function of the NOMA multiplexing order for different active-device populations per cell ($N = \{150, 200, 250, 300\}$).

trade-off between multiplexing gain and interference in uplink NOMA systems. Although multiplexing more users on the same RB improves spectral utilization and connectivity, it also increases intra-cell interference, successive interference cancellation (SIC) complexity, and the transmit power required to satisfy the QoS constraints. Consequently, the additional power consumption grows more rapidly than the achievable throughput, resulting in a gradual reduction in system energy efficiency.

Nevertheless, larger active-device populations consistently achieve higher energy efficiency because the scheduler has access to a larger pool of candidate users, enabling more effective channel-aware user selection and interference-aware resource allocation. This additional scheduling flexibility improves resource utilization without proportionally increasing the total power consumption. These results demonstrate that the proposed EERS-MTCN algorithm remains scalable across different network loading conditions while effectively exploiting the increased scheduling diversity available in denser deployments.

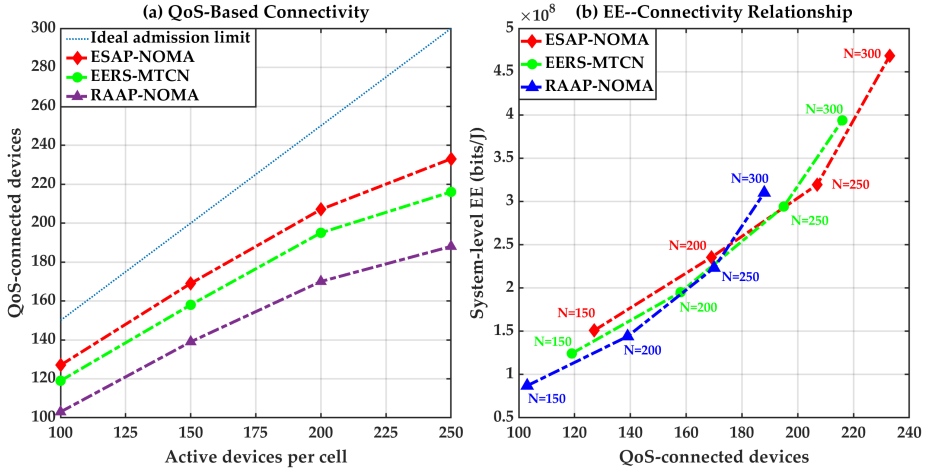


Figure 9: Illustrative visualization of the QoS-based connectivity and energy-efficiency relationship for Publication (III). (a) QoS-connected devices versus the number of active devices per cell. (b) Corresponding EE-connectivity trade-off for ESAP-NOMA, EERS-MTCN, and RAAP-NOMA, illustrating the relative behavior of the three scheduling algorithms.

Figure 9 provides an integrated conceptual interpretation of the joint energy-efficiency and connectivity objectives investigated in Publication (III). Whereas the published results evaluate these performance metrics separately under different active-device populations and NOMA loading levels, the illustrative visualization combines them to highlight their relationship. Figure 9(a) depicts the QoS-connected devices achieved by each scheduling algorithm as the active-device population per cell increases. The dotted line represents the ideal admission limit, corresponding to the hypothetical case in which all active devices satisfy the minimum QoS requirements. In practice, increasing network load together with the fixed inter-cell interference assumed in Publication (III) limits the number of devices that can be simultaneously admitted.

Figure 9(b) presents the corresponding energy efficiency-connectivity operating points. ESAP-NOMA defines the upper performance bound by exhaustively exploring all feasible

user-grouping and resource-allocation combinations, whereas RAAP-NOMA provides the lowest operating point because of its random allocation strategy. The proposed EERS-MTCN operates close to the ESAP-NOMA performance frontier while requiring substantially lower computational complexity, demonstrating that interference-aware scheduling and channel-aware user grouping provide an effective compromise between energy efficiency, QoS-based connectivity, and implementation complexity.

The execution-time results reported in Publication (III) provide a complementary assessment of the computational efficiency of the three scheduling strategies. Although ESAP-NOMA achieves the highest energy efficiency by exhaustively exploring all feasible user-pairing and resource-allocation combinations, it requires approximately one order of magnitude longer execution time than EERS-MTCN. In contrast, the proposed EERS-MTCN achieves energy-efficiency performance close to the exhaustive-search benchmark while substantially reducing computational complexity through interference-aware scheduling and power allocation. Although RAAP-NOMA incurs the lowest computational cost, its random resource allocation results in inferior energy efficiency and QoS-based connectivity. Collectively, these results demonstrate that EERS-MTCN provides a practical balance between scheduling performance and computational efficiency, making it well suited for dense uplink mMTC deployments.

5.5.2 SRS-MTCN: Interference, EE, Connectivity, and FBL Sensitivity

Publication (IV) extends the framework developed in Publication (III) along two complementary dimensions. First, it augments the non-cooperative scheduling framework by introducing cooperative multi-cell interference coordination, allowing neighboring cells to exchange scheduling-related information and jointly mitigate inter-cell interference. Second, it reformulates the resource-allocation problem using an FBL transmission model, thereby explicitly incorporating decoding error probability into the optimization process and accounting for the reliability limitations of short-packet communications in mMTC networks. Consequently, the proposed SRS-MTCN framework jointly optimizes cooperative interference management and reliability-aware resource allocation, providing a more realistic representation of practical 6G mMTC operation. From this perspective, Publication (IV) represents a natural progression of Publication (III): whereas the earlier work establishes the benefits of interference-aware scheduling under fixed inter-cell interference, the latter demonstrates how these gains can be further improved through cooperative interference coordination and FBL-aware scheduling under reliability-constrained short-packet transmission.

5.5.3 EE trends under cooperative FBL-aware scheduling

The performance of the proposed cooperative FBL-aware scheduling framework is evaluated in Figs. (10–12). Unlike the non-cooperative scenario investigated in the previous publication, where increasing connection density resulted in a continuous degradation of EE due to uncontrolled inter-cell interference, the cooperative framework exhibits a peak-and-decline trend. This behaviour demonstrates the ability of coordinated resource management to exploit additional multiplexing opportunities in moderate-density regimes

before interference accumulation and finite block-length constraints become dominant.

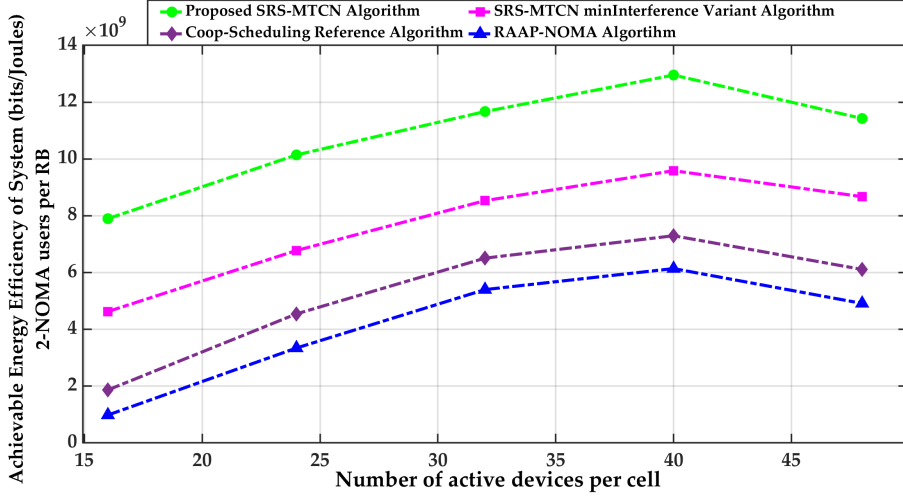


Figure 10: System-level EE comparison for 2 NOMA users per RB under varying user connection densities. The proposed EERS-MTCN (maxEE and minInterference variants) outperforms RAAP-NOMA and the reference algorithm across all density levels, demonstrating improved scalability and EE in moderately loaded scenarios.

The improved performance is attributed to the joint optimization of NOMA clustering, resource assignment, and power allocation across neighbouring cells. By incorporating inter-cell coordination, EERS-MTCN can better balance the benefits of increased user multiplexing against the resulting interference and reliability penalties introduced by FBL transmission. However, beyond a certain density threshold, the additional scheduling flexibility is insufficient to offset the increased interference and decoding complexity, leading to a reduction in system-level EE.

For the 2-user NOMA configuration shown in Fig. 10, EERS-MTCN achieves the highest EE throughout the considered user density range. The EE increases from approximately 8.0×10^9 bits/J at low density and reaches a maximum of approximately 1.3×10^{10} bits/J around 40 active devices per cell, followed by a moderate reduction at higher densities. The minInterference variant follows a similar trend but achieves lower EE, indicating that prioritizing interference suppression alone limits the exploitation of available multiplexing opportunities. In comparison, the cooperative reference scheme and RAAP-NOMA yield lower EE values, demonstrating the advantage of jointly considering energy efficiency and interference constraints in the scheduling.

A comparable trend is observed for the 4-user NOMA configuration in Fig. 11. The proposed EERS-MTCN scheme again provides the highest EE and achieves approximately 1.26×10^{10} bits/J at around 40 active devices per cell. At the same operating point, the minInterference variant, cooperative reference algorithm, and RAAP-NOMA achieve approximately 8.8×10^9 , 6.4×10^9 , and 4.8×10^9 bits/J, respectively. These results correspond to EE improvements of approximately 35–177% depending on the selected benchmark.

For the 8-user NOMA configuration in Fig. 12, the impact of aggressive multiplexing

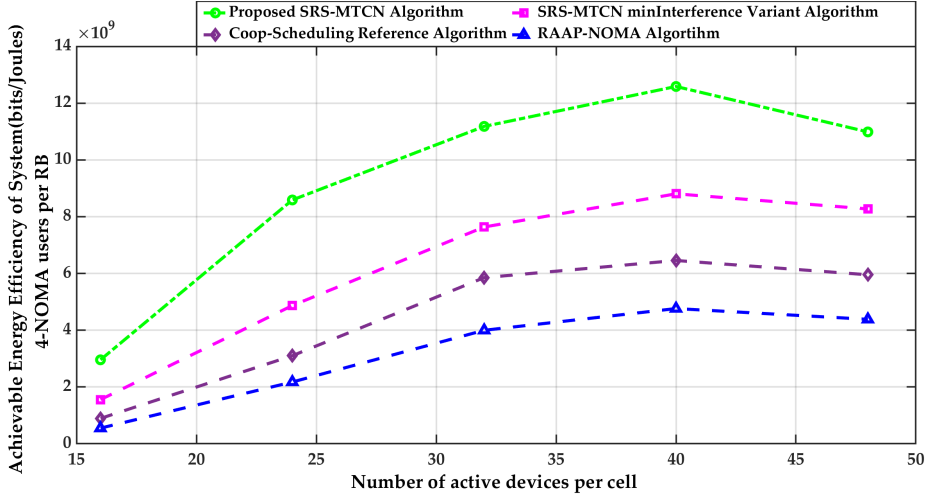


Figure 11: System-level EE comparison for 4 NOMA users per RB as a function of user connection density. Increasing the number of multiplexed NOMA users enhances EE performance for the proposed EERS-MTCN schemes, highlighting their robustness and superior interference management compared to baseline algorithms.

becomes more pronounced. Although cooperative scheduling continues to provide performance improvements compared with the benchmark algorithms, the achievable EE decreases compared with lower-order NOMA configurations. This reduction is caused by increased intra-cell interference, SIC decoding complexity, and the stricter reliability constraints associated with FBL transmission.

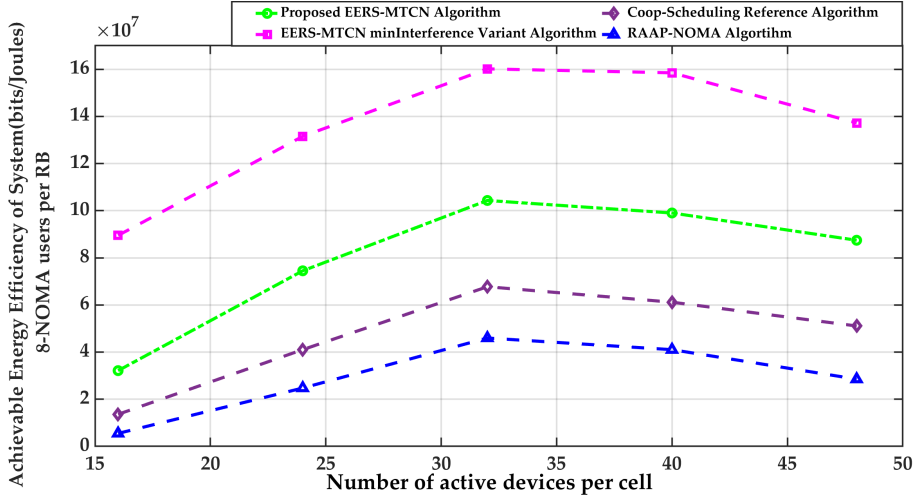


Figure 12: System-level EE performance with 8 NOMA users per RB under different user densities. The results illustrate the impact of higher NOMA multiplexing on EE and confirm that the proposed EERS-MTCN approaches maintain performance advantages over RAAP-NOMA and the reference method in dense mMTC deployments.

Interestingly, the minInterference variant becomes relatively more competitive in

this high-density regime, indicating that interference minimization becomes increasingly important when the number of simultaneously multiplexed users increases. These observations confirm that the optimal scheduling strategy depends on the operating regime: energy-efficiency-oriented optimization is advantageous under moderate loading, whereas stronger interference control becomes necessary under extreme multiplexing conditions.

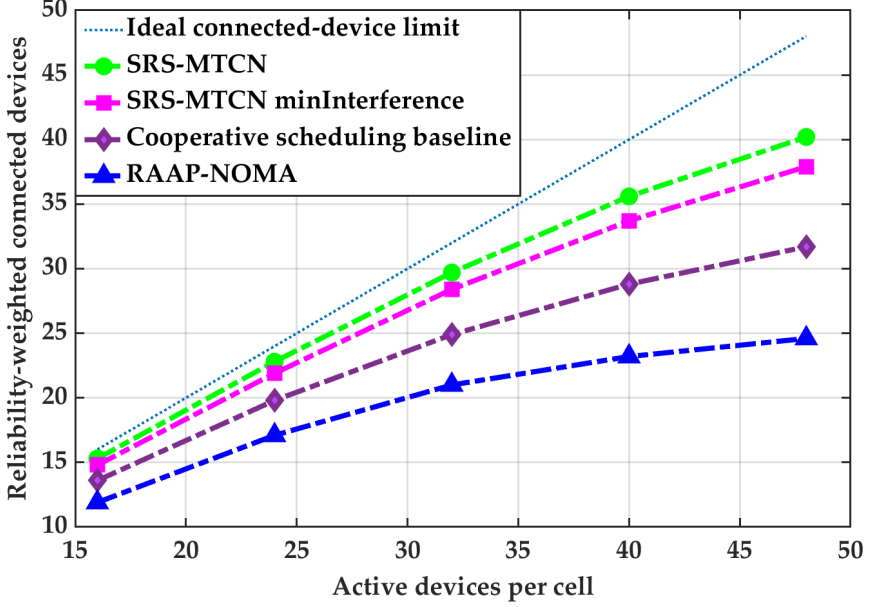


Figure 13: Reliability-weighted connectivity under finite-blocklength (FBL) transmission versus the number of active devices per cell. The dotted line denotes the ideal connected-device limit, while the proposed SRS-MTCN remains closest to this upper bound and outperforms the minInterference variant, the cooperative scheduling baseline, and RAAP-NOMA across all evaluated loads.

Fig. 13 shows the reliability-weighted connectivity achieved by the cooperative FBL-aware scheduling schemes as the number of active devices per cell increases. The metric counts the effective number of connected devices after weighting each scheduled transmission by its decoding-success probability, thereby capturing the reliability penalty introduced by finite-blocklength transmission. The proposed SRS-MTCN algorithm consistently remains closest to the ideal connected-device limit and outperforms the minInterference variant, the cooperative scheduling baseline, and RAAP-NOMA. The widening gap from the ideal line at higher device populations reflects the combined effect of limited RB availability, stronger intra- and inter-cell interference, SIC burden, and non-zero FBL decoding errors.

The FBL analysis further demonstrates the practical relevance of the proposed cooperative scheduling framework. Compared with the Shannon-based transmission model, the FBL regime results in an EE reduction of approximately 30–40% due to shorter transmission packets and enhanced reliability requirements of mMTC services. Nevertheless, EERS-MTCN maintains significant improvements over the baseline approaches. Particularly in the moderate-density range of approximately 16–48 active devices per cell, the pro-

posed method achieves EE improvements of approximately 35–58%. These results confirm that the proposed cooperative scheduler remains effective when realistic short-packet constraints are taken into account. Therefore, the fourth publication extends the previous non-cooperative framework by demonstrating that coordinated interference management, combined with FBL-aware resource allocation, can enable sustainable, energy-efficient operation for dense 6G mMTC networks.

5.5.4 Discussion

The cooperative results should therefore be interpreted in the context of a stronger system model than that used for EERS-MTCN. In Publication (IV), the scheduler is not only interference-coordinated across cells but also explicitly FBL-aware, meaning that the achieved gains reflect joint optimization of clustering, interference-aware resource-block assignment, and multi-cell power refinement under short-packet reliability constraints. This makes the SRS-MTCN results complementary rather than directly equivalent to the non-cooperative results: the former quantifies the benefit of coordination under practical reliability constraints, whereas the latter establishes the baseline performance of purely local interference-aware scheduling.

Unlike the non-cooperative case, where energy efficiency deteriorates monotonically with increasing density and multiplexing order, the cooperative model can exploit growing density more productively over a broader operating region. This is because the scheduler can suppress harmful inter-cell interactions before they accumulate, rather than merely reacting to them locally. The peak-and-decline pattern observed in Figs. (10–12) therefore reflects a more balanced trade-off between resource reuse, cooperation gain, and FBL-constrained reliability.

5.6 Chapter Summary and Bridge to Chapter 6

This chapter addressed **RQ2** by developing and evaluating joint EE and connectivity optimization frameworks for mMTC under multi-cell NOMA settings with interference coupling and short-packet reliability constraints. Publication (III) first demonstrated that substantial EE gains can be achieved under non-cooperative operation through interference-aware scheduling. In particular, the proposed EERS-MTCN framework achieved improvements of up to 160% and 55% over NOMA reference schemes at a connection density of 300 active devices per cell, depending on the multiplexing configuration.

Building on this foundation, Publication (IV) extended the study to cooperative and FBL-aware scheduling through the SRS-MTCN framework. The results showed that coordinated scheduling and power control can significantly reduce residual interference, achieving approximately 70–80% interference suppression and 35–58% EE improvement over state-of-the-art baselines in the 16–48 active devices per cell regime. These findings confirm that interference-aware and reliability-aware resource allocation is essential for sustainable massive connectivity in dense NOMA-based networks.

At the same time, the chapter reveals a methodological limitation of the optimization-based approaches developed so far. The algorithms in Chapters (4 and 5) operate under fixed optimization formulations and predetermined objective structures. As a result, they

must be reconfigured or re-optimized when traffic patterns, interference conditions, or service priorities change. This limitation becomes critical in practical 6G environments, where operating conditions are highly dynamic and non-stationary. This motivates the transition to Chapter 6, where the focus shifts from static optimization to adaptive policy learning. Specifically, Chapter 6 addresses **RQ3** by introducing multi-objective reinforcement learning for preference-conditioned scheduling, enabling online adaptation to changing performance trade-offs without repeated manual retuning or complete reformulation of the control problem.

6 Reinforcement Learning for Multi-Objective Optimization

6.1 Overview

The third challenge identified in Chapter 1—*adaptability*—requires that the scheduling framework respond to time-varying channel conditions, fluctuating interference loads, and shifting operator priorities without requiring repeated re-optimization. This chapter addresses that challenge through the following research question.

RQ3: How can MORL be employed to learn a single, adaptive scheduling policy that generalizes across preference trade-offs and dynamically responds to time-varying network conditions in 6G systems?

Chapters (4 and 5) established, respectively, a coordination mechanism for mitigating decision-dependent inter-cell interference and an analytically grounded optimization framework for joint energy-efficiency (EE)-connectivity under finite-blocklength (FBL) constraints. Both rely on deterministic, model-based solvers that must be re-invoked whenever the network state or problem parameters change. In practical 6G deployments, user arrivals, mobility-induced channel variations, shadowing, and neighbouring scheduling decisions alter the network state continuously; operator priorities may similarly shift between reliability, EE, and connectivity according to traffic class and service-level requirements. Resolving the full multi-objective program within the 10–100 ms control-loop budget of the O-RAN Near-RT RIC is therefore computationally impractical [18].

This chapter moves from fixed optimization formulations to *policy-level adaptation*. MORL provides the natural formalism for this shift: it represents competing KPIs as a vector-valued reward and learns Pareto-efficient policies rather than a single scalar optimum, and the design space of scalarization, decomposition, and preference-conditioned methods is now well documented [120, 121]. The proposed *WiMO-CEPPO* (*Wireless Multi-Objective Conditional Envelope PPO*) framework learns a single scheduling policy that adapts both to changing operating conditions *and* to shifting objective priorities. Adaptation to operating conditions follows from training the policy across a distribution of channel, load, and interference realizations; adaptation to objective priorities is achieved through a learned *Context-to-Preference Mapper* (CPM) that translates the observed network state into a preference vector at inference time. A single inference step, one forward pass through the CPM, followed by one pass through the policy network, suffices to produce a resource-allocation decision, keeping the approach compatible with Near-RT RIC latency constraints.

Publication (V) presents the core WiMO-CEPPO architecture for a single cell together with a sustainability-oriented reward formulation with a wider continuous-control baseline set, ablation, and KPI analysis. Publication VI broadens the evaluation across centralized multi-cell architectures for O-RAN with an explicit hardware/compute deployability constraint along a discrete-control baseline set, ablation, and KPI analysis. Together, these publications form the learning-based component of the thesis and provide the primary response to RQ3.

The remainder of this chapter is organized as follows. Section 6.2 revisits the subset of MORL literature most relevant to positioning the contribution. Section 6.3 maps each

identified gap to the specific WiMO-CEPPO mechanism that addresses it. Section 6.4 describes the MO-MDP formulation and the three-component architecture. Sections 6.5 and 6.6 present the experimental setup, baselines, and results.

6.2 Related Work

Chapter 2 surveyed the broader literature on MORL for wireless resource management, including scalarization-based methods, preference-conditioned policies, envelope architectures, and multi-agent formulations. This section revisits only the subset of prior work directly required to position Publications (V) and (VI).

Limitations in Wireless Objective Coverage

Most wireless MORL studies optimize two or three objectives, typically EE, throughput, or connectivity. In contrast, interference-coupled multi-cell uplink scheduling requires the joint treatment of four: EE, connectivity, reliability, and cross-link interference (CLI) mitigation [122, 113]. Toure et al. [122] showed that a single MORL agent can navigate throughput-fairness trade-offs in wireless scheduling without retraining. Still, their formulation is confined to two objectives and does not incorporate interference constraints or deployment feasibility. Static-scalarization methods (e.g., MO-PPO [102]) lock the policy to a single Pareto operating point; any change in network priorities requires full retraining, which is incompatible with the dynamic priority shifts seen in 6G mMTC scenarios. The limitation is also structural rather than merely practical: linear scalarization provably recovers only the convex portion of the Pareto front, so even an exhaustive weight sweep leaves solutions in concave regions unreachable [123]. The gap, therefore, is a policy that covers all four objectives at once and adapts to shifting priorities without re-optimization.

Limitations in Preference Adaptation

Preference-conditioned MORL lets a single policy represent multiple Pareto operating points by conditioning on a continuous preference vector [103, 104, 124]. Complementary lines of work reuse value information across preferences to improve coverage and transfer, for example, through successor-feature bases that support optimal policy transfer between weightings [125]. The critical limitation common to these approaches is that the preference vector is assumed to be supplied externally at inference time. In wireless scheduling, the appropriate trade-off depends on the current network context load, interference level, reliability deficit, and service requirement, none of which is constant. Without an internal mechanism for *inferring* preferences from observed state, such methods remain difficult to deploy when conditions change faster than manual operator updates can be issued [103].

Limitations in Conflict-Aware Learning and Envelope Methods

In 6G uplink scheduling, objective conflicts are structural: increasing transmit power simultaneously improves SINR while degrading EE and inter-cell interference; aggressive RB reuse improves connectivity while reducing SIC feasibility under NOMA superposition. During policy-gradient learning, such conflicts manifest as opposing gradient directions, causing oscillatory training and convergence to narrow Pareto arcs rather than broad

coverage [108, 106]; obtaining verifiably broad coverage generally requires dedicated front-discovery procedures [126]. One family of methods addresses the conflict at the gradient level, combining per-objective updates so that improving one objective does not destructively cancel another, whether by optimizing each objective in a distributional, scalarization-free manner [127] or by aggregating gradients in a conflict-aware way under constraints [128]. Envelope-based MORL methods [107] take a different route, improving coverage by learning a convex piecewise-linear approximation of the multi-objective value function; this approximation is static, however, fitted as a separate value representation that does not respond to how severely objectives conflict at a given update step. What remains absent in the prior literature is a mechanism that reacts to inter-objective gradient conflict *as it occurs*, damping updates only when objectives are locally incompatible and learning at full strength otherwise, without maintaining a separately parameterized envelope network.

Multi-Agent Architectures Across RAN Topologies

The choice of learning architecture is tightly coupled to the underlying RAN topology, and the difficulty grows as that topology becomes more distributed. In a conventional *single-cell* setting, a single agent observes one cell and treats neighbouring activity as external interference: learning is stationary, and no coordination is needed, but the agent is blind to the inter-cell coupling that dominates dense deployments. Extending to a conventional *multi-cell* RAN, in which a central controller such as a C-RAN BBU pool holds a global view, preserves a stationary single-agent formulation and restores coupling awareness, but concentrates all telemetry and computation at one point, so cost scales with the number of cells and the formulation assumes a fronthaul connectivity that is not always available. In a *distributed*, O-RAN, control is partitioned across O-DUs, which makes multi-agent RL (MARL) one autonomous agent per O-DU, the natural fit for the architecture.

The price of that distribution, however, is precisely the property that the single-cell and centralized formulations avoid. MARL agents face non-stationarity; each agent’s environment shifts as the others learn, an effect most severe in the high-density, high-interference regimes of interest, where the very coupling that motivates coordination also destabilizes the learning. Restoring stability typically requires repeated coordination messages and synchronized critic estimates among agents, introducing overhead that can exceed the $T_{\text{ctrl}} = 10\text{--}100\text{ ms}$ Near-RT RIC budget [18]. The architecture spectrum, therefore, presents a standing trade-off: single-cell agents are stable but coupling-blind; centralized multi-cell controllers are stable and coupling-aware but scale poorly and presume fronthaul connectivity that distributed RAN does not guarantee; and fully distributed MARL matches the O-RAN topology but reintroduces non-stationarity and per-step signaling cost. This leaves an open gap for an architecture that captures the dominant inter-cell coupling of a multi-cell deployment without per-step inter-agent signaling or non-stationary training dynamics, while remaining deployable within O-RAN latency limits.

This motivates the architectural extension in Publication (VI), where WiMO-CEPPO is evaluated beyond a narrow single setting and positioned against broader RAN-control baselines.

Deployment Feasibility

A further limitation of existing wireless MORL studies is the implicit treatment of computational cost. An xApp hosted on the Near-RT RIC must complete telemetry collection (t_{rx}), CPM inference and preference construction (t_{cpm}), policy inference (t_{inf}), action decoding (t_{dec}), and control delivery (t_{tx}) within the budget $t_{rx} + t_{cpm} + t_{inf} + t_{dec} + t_{tx} \leq T_{ctrl}$ [18]. Inference latency and per-decision computational cost are therefore operational constraints, not implementation details.

Accordingly, deployment feasibility is treated as part of the policy-design problem rather than as a post-training implementation concern.

6.3 Positioning of the Contribution: Gap-to-Mechanism Mapping

Table 6 maps the limitations identified in the related work to the corresponding WiMO-CEPPO mechanisms. Publication (V) introduces the base WiMO-CEPPO framework in a single-cell setting, including the preference-conditioned policy, CPM, conditional envelope, and sustainable multi-objective reward. Publication (VI) extends this framework from single-cell operation to a multi-cell O-RAN setting, where inter-cell coupling, centralized RAN control, Near-RT RIC feasibility, and broader baseline comparisons are explicitly considered.

Table 6: Gap-to-mechanism mapping for the WiMO-CEPPO framework.

Gap	WiMO-CEPPO mechanism	Pub.
Static scalarization: policy optimized for a fixed preference vector	Preference-conditioned policy with continuous $\mathbf{w}$ input enabling Pareto-aware policy adaptation	V; ext. VI
External preference specification: fixed objective priorities cannot adapt to network context	Context-to-Preference Mapper (CPM) generating dynamic preference vectors during policy execution	V; ext. VI
Gradient conflict: competing objectives create unstable policy updates	Conditional envelope mechanism regulating policy updates according to objective conflict	V
Limited architectural validation of MORL approaches	Evaluation across classical RAN and O-RAN scenarios with extended MORL baselines	V; ext. VI
Limited sustainability representation in wireless MORL	Sustainable four-dimensional reward formulation incorporating EE, connectivity, reliability, and interference	VI
Deployment feasibility neglected in MORL-based wireless control	O-RAN-aware policy deployment analysis including inference latency evaluation	VI
Single-cell classical RAN studies do not capture scalable network architectures	Extension of WiMO-CEPPO from classical RAN to O-RAN architecture for practical 6G control	VI

Taken together, the two publications develop WiMO-CEPPO from a base preference-

conditioned MORL framework into a deployment-oriented RAN-control architecture. Publication (V) establishes the continuous/hybrid-control evaluation against actor-critic MORL baselines, while Publication (VI) evaluates the O-RAN-oriented extension against a discrete value-based envelope comparator. The thesis-level contribution is the synthesis of these two evaluations: the same design principles—context-aware preference inference, conflict-aware policy updates, and constraint-aware action generation—remain valid across different baseline families and action-space parameterizations.

6.4 Methodology of the Contributions

Contextualization Within the Thesis Framework

Chapters (4 and 5) formulated the wireless scheduling problem through analytical and optimization-based models, including interference minimization, energy-consumption reduction, and joint energy-efficiency-connectivity optimization under finite-blocklength constraints. These formulations establish the feasible operating region under controlled assumptions, but they must be re-solved when the network state, traffic load, interference profile, or objective priorities change.

Publications (V and VI) extend this line of work by moving from fixed analytical optimization toward a preference-aware multi-objective policy-learning framework. Publication (V) introduces the base WiMO-CEPPO framework and evaluates its core preference-conditioned MORL mechanisms for balancing energy efficiency, connectivity, reliability, and interference mitigation. Publication (VI) extends the deployment perspective by emphasizing multi-cell coupling, O-RAN feasibility, and Near-RT RIC control-loop constraints. This progression provides the methodological bridge from sustainable wireless scheduling to deployment-aware multi-objective RAN control.

MO-MDP Formulation

The uplink scheduling problem is formulated as a preference-conditioned multi-objective Markov decision process (MO-MDP),

$$(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathbf{r}, \mathbf{w}, \pi),$$

where $\mathcal{S}$ is the state space, $\mathcal{A}$ is the action space, $\mathcal{P}$ is the transition kernel, $\mathbf{r}$ is the vector-valued reward, $\mathbf{w}$ is the preference vector, and π is the scheduling policy.

The state s_t contains channel quality indicators, inter-cell interference information, cell-load statistics, service deficits, and compact context indicators. In the base formulation, the policy observes the focal scheduling context and accounts for neighbouring-cell interference through measured or aggregated interference indicators. In the multi-cell O-RAN extension, the controller observes a joint state collected from multiple cells through RAN telemetry, enabling coupling-aware scheduling decisions under a centralized RAN-control view.

The action a_t contains resource-block (RB) assignment and transmit-power decisions. RB assignment is represented through a discrete or relaxed binary mask, while transmit power is represented as a continuous value constrained by the maximum UE power budget.

The reward is maintained as a four-dimensional vector,

$$\mathbf{r}_t = \left[r_t^{\text{EE}}, r_t^{\text{conn}}, r_t^{\text{rel}}, r_t^{\text{int}} \right]^\top,$$

where the components correspond to energy efficiency, connectivity, reliability, and interference mitigation, respectively. The preference vector

$$\mathbf{w}_t \in \Delta^3$$

weights these four objectives and is generated dynamically by the Context-to-Preference Mapper (CPM). The scalarized reward used for policy optimization is

$$r_t = \mathbf{w}_t^\top \mathbf{r}_t.$$

Hardware and computational sustainability are not treated as an additional reward dimension. Instead, they are handled as explicit deployment constraints through inference-latency and compute-cost terms. This preserves the interpretability of the four radio objectives while ensuring that the learned policy remains feasible for real-time RAN operation.

WiMO-CEPPO Architecture

Fig. 14 illustrates the WiMO-CEPPO framework. The architecture consists of three main components: the Context-to-Preference Mapper, the preference-conditioned PPO policy with a conditional envelope mechanism, and the constraint-aware action projection module.

Component 1 – Context-to-Preference Mapper. The CPM is a lightweight neural network

$$f_\phi : \mathcal{S} \rightarrow \Delta^3$$

that maps the observed state to a dynamic preference vector,

$$\mathbf{w}_t = f_\phi(s_t).$$

Its role is to infer objective priorities directly from the network context rather than relying on externally fixed weights. For example, high interference may increase the priority of interference suppression, while a reliability deficit may increase the weight assigned to the reliability component. This addresses the limitation of preference-conditioned MORL methods that assume the preference vector is externally supplied at inference time.

Component 2 – Preference-Conditioned PPO Policy. The actor is conditioned jointly on the observed state and the inferred preference vector:

$$\pi_\theta(a_t \mid s_t, \mathbf{w}_t).$$

The critic is also conditioned on $(s_t, \mathbf{w}_t)$, allowing the value estimate to adapt to different operating points on the Pareto front. This enables WiMO-CEPPO to learn a single reusable

policy over the preference simplex rather than requiring one separately trained policy for each trade-off configuration.

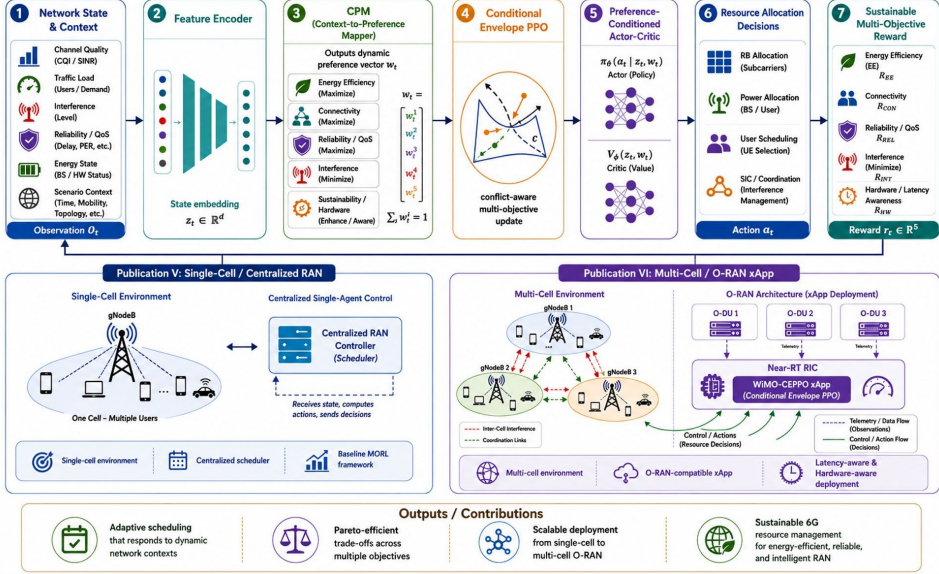


Figure 14: WiMO-CEPPO architecture. The CPM maps the observed network state s_t to a preference vector $\mathbf{w}_t$. The preference-conditioned policy $\pi_\theta(a_t | s_t, \mathbf{w}_t)$ generates candidate RB-assignment and power-control actions. The conditional envelope stabilizes PPO updates under inter-objective gradient conflict, while the feasibility projector maps raw actions to admissible scheduling decisions before execution.

Component 3 – Conditional Envelope Mechanism. The conditional envelope modifies the PPO advantage according to the degree of inter-objective conflict. Let $A_t^{(k)}$ denote the advantage estimate for objective k , and let $\bar{\rho}_t \in [0, 1]$ denote the normalized gradient-conflict measure. The envelope advantage is defined as

$$A_t^{\text{env}} = \left(\sum_{k=1}^4 w_{t,k} A_t^{(k)} \right) \max(0, 1 - \kappa \bar{\rho}_t),$$

where $\kappa > 0$ controls the attenuation strength. When the objectives are locally compatible, $\bar{\rho}_t$ is small and the policy update proceeds normally. When the objectives strongly conflict, the update is damped to prevent destructive gradient steps. This differs from static envelope methods because the attenuation is applied dynamically during policy learning and does not require a separately parameterized envelope network.

Component 4 – Constraint-Aware Action Generation. The raw policy output is mapped to a feasible RB-assignment and power-control decision through an action-projection step. The projection clips transmit powers to the admissible range, enforces the NOMA multiplexing limit per RB, and binarizes the RB-assignment mask:

$$(a_t, p_t) = \Pi_\Omega(\tilde{a}_t, \tilde{p}_t),$$

where $\Pi_{\Omega}(\cdot)$ denotes projection onto the feasible scheduling set. Thus, the hard structural constraints on transmit power, RB sharing, and binary assignment are enforced before the action is executed.

Rate, reliability, and hardware-cost requirements are handled through penalty terms during training and verified during evaluation. This separation is important: hard projection guarantees physical feasibility of the executed action, while soft penalties guide the policy toward QoS satisfaction and deployment feasibility without obscuring the four-objective reward structure.

Training Procedure

WiMO-CEPPO follows a three-stage training procedure.

In Stage 1, the preference-conditioned actor-critic is trained using preference vectors sampled from a Dirichlet distribution:

$$\mathbf{w} \sim \text{Dir}(\alpha_0).$$

This encourages broad Pareto-front coverage before the CPM is introduced.

In Stage 2, the CPM is fitted using state-preference pairs collected from trajectories. The target preference corresponds to the preference vector that achieves the strongest short-horizon scalarized return for a given state context. The CPM therefore learns to associate observable network conditions with suitable objective-priority vectors.

In Stage 3, the actor-critic and CPM are fine-tuned jointly. The actor receives the CPM output as a conditioning vector, while the CPM is updated through its own preference-alignment objective. A stop-gradient operation prevents unstable feedback between preference inference and policy-gradient updates. In the O-RAN-oriented extension, the training objective also includes deployment-aware penalties for inference latency and computational cost.

Computational Complexity and Deployment Feasibility

The complexity of WiMO-CEPPO is separated into training-time and deployment-time costs. During training, the algorithm computes per-objective advantages, estimates inter-objective gradient conflict, applies the conditional envelope, updates the PPO actor-critic, and fits the CPM. During deployment, however, only three operations are required: one CPM forward pass, one preference-conditioned actor forward pass, and one lightweight action projection step.

Let M denote the number of cells, U the number of UEs per cell, R the number of RBs, $N_{\max}$ the NOMA multiplexing limit, and $K_o = 4$ the number of objectives. For an MLP with input dimension d_{in} , output dimension d_{out} , hidden width H , and L hidden layers, one forward pass has complexity

$$\mathcal{C}_{\text{MLP}} = \mathcal{O}(d_{\text{in}}H + (L-1)H^2 + Hd_{\text{out}}).$$

The CPM inference complexity is

$$\mathcal{C}_{\text{CPM}} = \mathcal{O} \left(d_s H_c + (L_c - 1) H_c^2 + H_c K_o \right),$$

and the preference-conditioned actor inference complexity is

$$\mathcal{C}_\pi = \mathcal{O} \left((d_s + K_o) H_\pi + (L_\pi - 1) H_\pi^2 + H_\pi d_a \right),$$

where d_s and d_a denote the state and action dimensions, respectively.

In the base framework, the state and action dimensions scale approximately as $\mathcal{O}(RU)$. Therefore, the online complexity is

$$\mathcal{C}_{\text{online}}^{(1)} = \mathcal{O} \left(\mathcal{C}_{\text{CPM}}^{(1)} + \mathcal{C}_\pi^{(1)} + RU \log U \right),$$

where $RU \log U$ corresponds to per-RB user ranking during action projection. If top- N_{max} selection is implemented without full sorting, this term can be reduced to approximately $\mathcal{O}(RU)$.

In the multi-cell O-RAN extension, the joint state and action dimensions scale as $\mathcal{O}(MRU)$. The online complexity becomes

$$\mathcal{C}_{\text{online}}^{(M)} = \mathcal{O} \left(\mathcal{C}_{\text{CPM}}^{(M)} + \mathcal{C}_\pi^{(M)} + MRU \log U + M^2 R N_{\text{max}}^2 \right),$$

where the final term accounts for live inter-cell interference coupling among scheduled NOMA users. When the neural-network widths and depths are fixed by design and N_{max} is bounded by the NOMA multiplexing limit, the dominant non-neural deployment cost scales approximately as

$$\mathcal{O} \left(MRU \log U + M^2 R \right).$$

Importantly, the deployed controller remains a centralized single-agent inference model and therefore avoids per-step inter-agent signalling and non-stationary MARL dynamics.

The Near-RT RIC feasibility condition is

$$\tau_{\text{rx}} + \tau_{\text{inf}} + \tau_{\text{dec}} + \tau_{\text{tx}} \leq T_{\text{ctrl}},$$

where τ_{rx} is telemetry collection time, τ_{inf} includes CPM and actor inference, τ_{dec} is action decoding and projection time, and τ_{tx} is control-message delivery time. This directly addresses the deployment gap identified in the related work: inference latency and compute cost are treated as operational constraints rather than ignored implementation details.

6.5 Experimental Setup and Baselines

Publications (V) and (VI) use the same general uplink NOMA scheduling simulator and the same four-dimensional radio reward structure. The shared objective vector contains energy efficiency, connectivity, reliability, and interference mitigation. Hardware and computational sustainability are treated as deployment constraints, not as a fifth reward component. The two studies are nevertheless reported separately because they use

different baseline families, action-space parameterizations, and evaluation emphases.

Table 7: Evaluation roles of the two WiMO-CEPPO publications. The table clarifies how the studies jointly support RQ3 without treating their numerical results as one unified benchmark.

Study	Baseline family	Evaluation emphasis	Thesis-level interpretation
Publication (V)	MO-PPO, MO-SAC, MO-TD3, MO-DDPG	Continuous/hybrid actor-critic control for RB assignment and transmit-power adaptation	Tests whether CPM and conditional-envelope PPO improve learning over stochastic and deterministic continuous-control MORL baselines.
Publication (VI)	MO-PPO, MO-SAC, DDQN-Envelope	O-RAN-oriented evaluation with a value-based envelope baseline over a discretized RB/power action code book	Tests whether WiMO-CEPPO remains effective when compared with a discrete envelope/value-based MORL alternative under deployment-oriented constraints.

Training and evaluation protocol. All methods are trained for 4,000 episodes with an episode horizon of $T = 100$ time steps and $N_s = 5$ independent random seeds. During evaluation, exploration is disabled: actor-critic methods use deterministic actions, while value-based methods use greedy action selection. The same normalized reward components and feasibility-check procedure are applied within each publication-specific comparison.

Publication (V): Continuous or hybrid-control evaluation. Publication (V) evaluates the base WiMO-CEPPO framework against continuous-control MORL baselines: MO-PPO, MO-SAC, MO-TD3, and MO-DDPG. This baseline set is appropriate because the study involves transmit-power adaptation and hybrid RB/power-control decisions. MO-PPO is the closest on-policy comparator because it shares the PPO backbone but lacks CPM-based preference inference and conditional-envelope conflict attenuation. MO-SAC provides an off-policy maximum-entropy actor-critic comparison, while MO-TD3 and MO-DDPG provide deterministic actor-critic baselines for continuous transmit-power control.

Publication (VI): O-RAN-oriented DDQN-Envelope evaluation. Publication (VI) compares WiMO-CEPPO with MO-PPO, MO-SAC, and DDQN-Envelope. DDQN-Envelope is implemented over a discretized RB/power action code book because value-based DDQN methods require a finite action space. This makes it a useful value-based envelope comparator, but not a direct continuous-control equivalent of WiMO-CEPPO. The comparison therefore tests whether the proposed preference-conditioned PPO design remains effective relative to a discrete envelope/value-based MORL alternative under deployment-oriented constraints.

Evaluation metrics. Pareto quality is evaluated using hypervolume (HV, $\uparrow$) and inverted generational distance (IGD, $\downarrow$). Wireless KPIs include energy efficiency, connectivity, reliabil-

Table 8: Simulation Environment and deployment parameters used across the WiMO-CEPPO evaluations. Values follow the system model and publication-specific assumptions.

Symbol	Description	Value used	Applies to
<i>Environment and deployment parameters</i>			
M	Number of cells	As in system model	All methods
U	UEs per cell	As in system model	All methods
R	Resource blocks	As in system model	All methods
$N_{\max}$	NOMA multiplexing limit per RB	As in system model	All methods
$P_{\max}$	Maximum transmit power per UE	As in system model	All methods
$R_u^{\min}$	Minimum rate requirement per UE	As in system model	All methods
$\gamma_{\min}$	Minimum SINR / decoding threshold	As in system model	All methods
T_{ctrl}	RAN control-loop interval	O-RAN setting	Publication (VI)

ity, interference reduction, and constraint satisfaction. Training efficiency is assessed using episodes-to-target performance and area-under-curve (AUC) measures, while stability is examined using utility-loss behaviour and distributional KPI results across seeds and channel realizations.

Table 9: Training protocol and WiMO-CEPPO-specific configuration. Values follow the publication-specific simulation settings.

Symbol	Description	Value used	Applies to
<i>Training protocol</i>			
T	Episode horizon	100 time steps	All methods
N_{ep}	Training episodes	4,000	All methods
N_s	Independent random seeds	5	All methods
K_o	Number of radio objectives	4	All MORL methods
<i>WiMO-CEPPO-specific parameters</i>			
α_0	Dirichlet concentration for preference sampling	Simulation setting	WiMO-CEPPO
κ	Conditional-envelope attenuation coefficient	Simulation setting	WiMO-CEPPO
μ	CPM fitting rate / learning rate	Simulation setting	WiMO-CEPPO
λ_2	Rate-QoS penalty weight	Simulation setting	WiMO-CEPPO
λ_5	Hardware-cost penalty weight	Simulation setting	O-RAN extension

Ablation variants isolate the contribution of each WiMO-CEPPO component. The “without CPM” variant removes context-to-preference inference; the “without Envelope” variant removes conflict-conditioned update attenuation; and the fixed-weight variant uses a static preference vector. These variants test whether the observed gains arise from adaptive preference inference, conflict-aware learning, or merely from the PPO backbone.

6.6 Results and Discussion

The purpose of this section is not to repeat the complete result discussion already reported in Publications (V and VI). Instead, the results are synthesized to answer RQ3 at the thesis level: whether a single preference-conditioned policy can adapt across changing wireless conditions and objective priorities while remaining feasible for real-time RAN control.

Training-Efficiency Interpretation

Figs. 15 and 16 show that WiMO-CEPPO reaches the target performance in substantially fewer episodes than the corresponding baselines. In Publication (V), WiMO-CEPPO converges in 800 episodes, compared with 1200–1500 episodes for the baselines, which corresponds to an episode reduction of about 33%–47%. In Publication (VI), it reaches the target in 900 episodes versus 1300–1600 episodes for the baselines, corresponding to a reduction of about 31%–44%. These results quantify the sample-efficiency advantage of the proposed preference-conditioned policy across both continuous-control and discretized baseline settings.

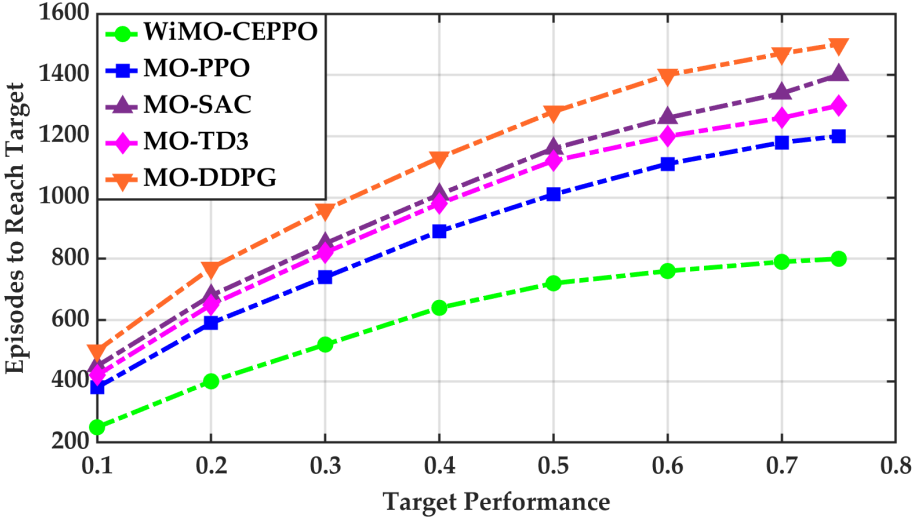


Figure 15: Sample-efficiency comparison of WiMO-CEPPO and baselines in terms of episodes required to reach target performance in Publication (V). Lower values indicate faster convergence, and WiMO-CEPPO is consistently the most sample-efficient.

In Publication (V), the proposed method converges faster than the continuous-control baselines MO-PPO, MO-SAC, MO-TD3, and MO-DDPG. In Publication (VI), it also maintains a clear episode advantage over MO-PPO, MO-SAC, and the envelope-based DDQN-Envelope. This pattern holds against both the actor-critic baselines and the envelope-based baseline, and the gap widening at higher target thresholds indicates that CPM-driven preference adaptation and conditional-envelope stabilization improve sample efficiency across both continuous and discretized action formulations.

Fig. 17 shows that WiMO-CEPPO achieves the highest cumulative reward over the 4,000-episode training horizon, indicating faster learning and stronger overall policy improvement.

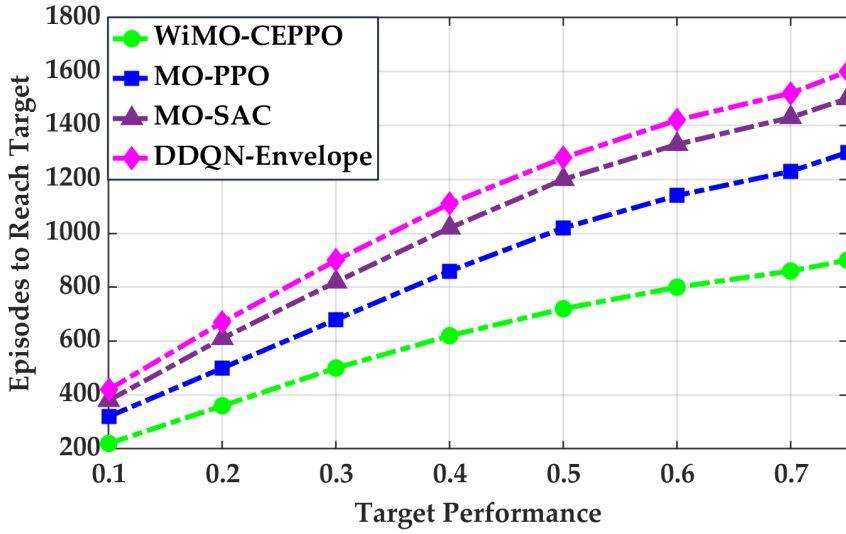


Figure 16: Sample-efficiency comparison of WiMO-CEPPO and baselines in terms of episodes required to reach target performance in Publication (VI). Lower values indicate faster convergence, and WiMO-CEPPO is consistently the most sample-efficient.

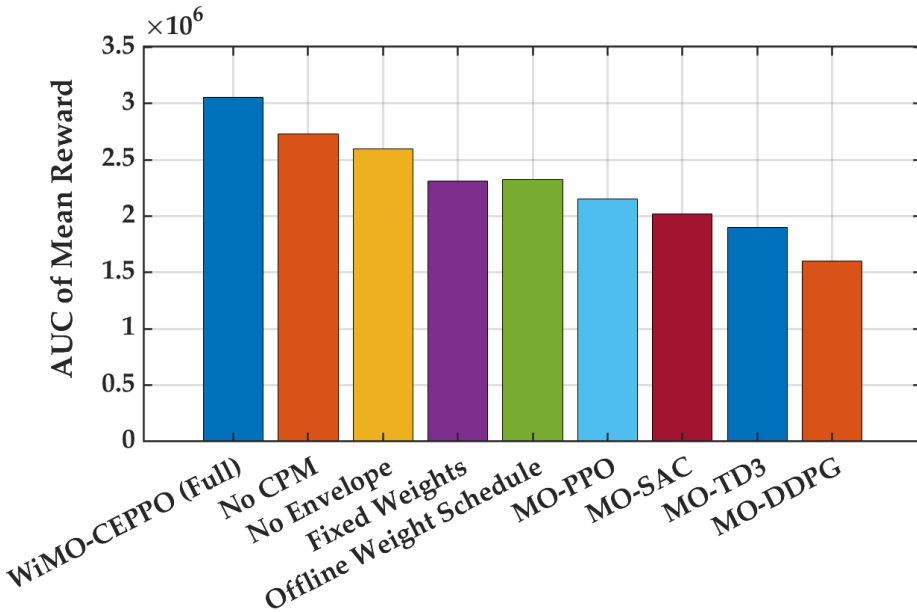


Figure 17: Mean reward AUC comparison for WiMO-CEPPO and baseline MORL methods over the 4,000-episode training horizon. Higher AUC indicates stronger cumulative reward acquisition over the training horizon.

The reward-AUC results provide a complementary view of training efficiency. WiMO-CEPPO achieves a mean-reward AUC of 2,722,715 and a shadowed-reward AUC of 2,558,387, exceeding MO-PPO, MO-SAC, and DDQN-Envelope by approximately 51%–67% and 51%–62%, respectively. This confirms that the proposed method accumulates higher reward throughout training, rather than only improving the final value.

Fig. 18 further shows that WiMO-CEPPO obtains the lowest utility-loss AUC, confirming that the learned preference-conditioned policy is not only reward-efficient but also more stable under conflicting objectives. Utility loss measures the gap between the reward achievable under the *ideal* preference vector for the current state and the reward achieved by the executed policy. The utility-loss AUC further supports this conclusion. WiMO-CEPPO attains a utility-loss AUC of 0.028, whereas the baselines lie in the range 0.051–0.060, implying a reduction of about 45%–53%. Since lower utility-loss AUC indicates smaller cumulative deviation from the context-conditioned operating point, this result shows that the proposed policy is not only more sample efficient but also more stable during training.

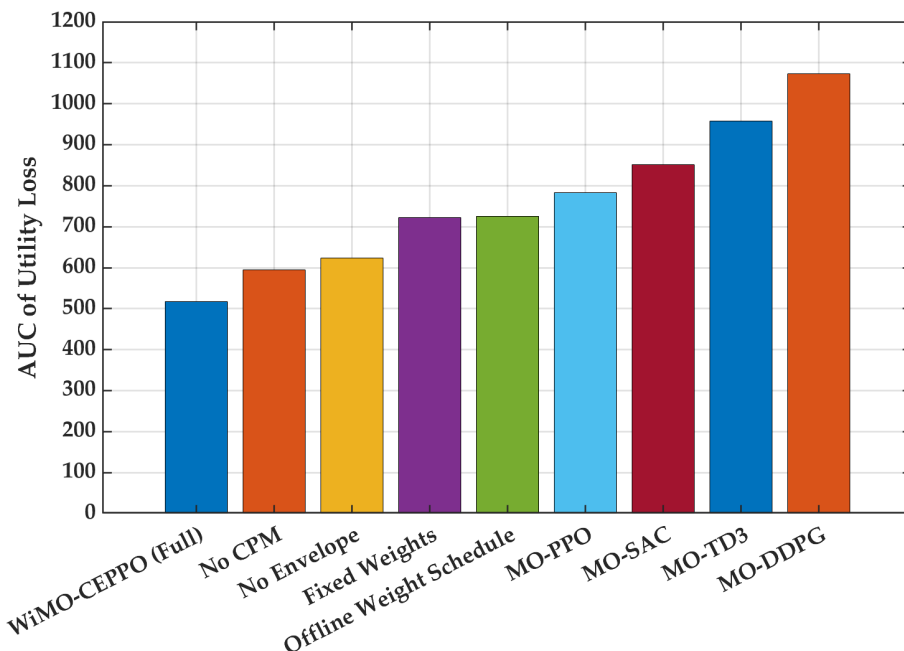


Figure 18: Area under the utility-loss curve for WiMO-CEPPO and baseline MORL methods over the 4,000-episode training horizon. Lower AUC indicates smaller cumulative deviation from the context-conditioned utility target and more stable preference-conditioned learning.

Together, these results suggest that the CPM improves the alignment between the policy and the current network context, while the conditional envelope reduces unstable updates when the objectives conflict. The ablation variants confirm this behaviour: removing the CPM, removing the envelope, or using fixed weights increases cumulative loss and weakens training efficiency.

Tables 10 and 11 report publication-specific performance anchors, reflecting the dif-

ferent baseline sets, action-space formulations, and deployment assumptions used in Publications (V and VI). Publication (V) evaluates continuous and hybrid actor-critic control, whereas Publication (VI) examines the O-RAN-aligned extension using a discretized DDQN-Envelope comparator. Consequently, variations in energy efficiency, connectivity, reliability, interference mitigation, constraint satisfaction, and episodes-to-target should be interpreted as context-dependent outcomes rather than repeated measurements of an identical experimental configuration.

Table 10: Performance anchors for Publication (V): continuous/hybrid action evaluation with MO-PPO, MO-SAC, MO-TD3, and MO-DDPG baselines. Results are averaged over $N_s = 5$ independent seeds. Interference is reported as normalized interference mitigation, where higher values indicate lower cross-link interference.

Algorithm	HV $\uparrow$	IGD $\downarrow$	EE $\uparrow$	Conn. $\uparrow$	Rel. $\uparrow$	Interf. Mit. $\uparrow$	Constr. $\uparrow$	Ep. $\downarrow$
WiMO-CEPPO (Full)	0.781	0.088	0.79	0.76	0.77	0.82	91.2%	800
WiMO-CEPPO (w/o CPM)	0.724	0.135	0.76	0.73	0.75	0.76	84.1%	–
WiMO-CEPPO (w/o Env.)	0.741	0.112	0.75	0.72	0.73	0.74	81.7%	–
WiMO-CEPPO (Fixed Weights)	0.698	0.161	0.73	0.70	0.71	0.70	79.3%	–
MO-PPO	0.582	0.205	0.74	0.70	0.72	0.68	86.7%	1200
MO-SAC	0.541	0.221	0.73	0.68	0.71	0.65	79.7%	1400
MO-TD3	0.518	0.218	0.70	0.67	0.69	0.63	74.4%	1300
MO-DDPG	0.535	0.230	0.68	0.65	0.68	0.61	64.7%	1500

Table 10 provides quantitative evidence that the full WiMO-CEPPO model achieves a superior operating point in Publication (V). Relative to the strongest baseline, MO-PPO, WiMO-CEPPO improves hypervolume from 0.582 to 0.781 (+34.2%) and reduces IGD from 0.205 to 0.088 (-57.1%), confirming a better and closer approximation to the reference Pareto front. These Pareto gains are reflected consistently in the system-level KPIs: EE increases from 0.74 to 0.79, connectivity from 0.70 to 0.76, reliability from 0.72 to 0.77, and interference mitigation from 0.68 to 0.82, while constraint satisfaction rises from 86.7% to 91.2%. The learning advantage is equally clear, since WiMO-CEPPO reaches the target in 800 episodes versus 1200–1500 episodes for the baselines. The ablation rows further isolate the contribution of each component: removing the CPM lowers HV to 0.724 and constraint satisfaction to 84.1%, removing the envelope yields HV = 0.741 and 81.7% constraint satisfaction, and fixing the weights degrades the results further to HV = 0.698 and 79.3%, showing that both preference adaptation and conflict-aware updates are required for stable multi-objective performance.

Table 11 shows that the same conclusion holds in the O-RAN-oriented evaluation of Publication (VI). WiMO-CEPPO again obtains the best Pareto quality, with HV increasing from 0.542 for MO-PPO to 0.752 (+38.7%) and IGD decreasing from 0.189 to 0.078 (-58.7%). The KPI values also remain consistently higher than the baselines: EE improves from 0.70 to 0.77, connectivity from 0.67 to 0.74, reliability from 0.70 to 0.75, and interference mitigation from 0.68 to 0.79, while constraint satisfaction increases from 85.9% to 90.4%. In terms of training efficiency, WiMO-CEPPO reaches the target in 900 episodes, compared with 1300 episodes for MO-PPO, 1500 for MO-SAC, and 1600 for DDQN-Envelope. The DDQN-Envelope baseline is particularly informative because it confirms that a discretized

Table 11: Performance anchors for Publication (VI): O-RAN-oriented DDQN-Envelope, evaluation with MO-PPO, MO-SAC, and DDQN-Envelope baselines. Results are averaged over $N_s = 5$ independent seeds. Interference is reported as normalized interference mitigation, where higher values indicate lower cross-link interference.

Algorithm	HV $\uparrow$	IGD $\downarrow$	EE $\uparrow$	Conn. $\uparrow$	Rel. $\uparrow$	Interf. Mit. $\uparrow$	Constr. $\uparrow$	Ep. $\downarrow$
WiMO-CEPPO (Full)	0.752	0.078	0.77	0.74	0.75	0.79	90.4%	900
MO-PPO	0.542	0.189	0.70	0.67	0.70	0.68	85.9%	1300
MO-SAC	0.493	0.202	0.67	0.64	0.67	0.65	78.8%	1500
DDQN-Envelope	0.488	0.190	0.64	0.62	0.62	0.62	74.4%	1600

envelope comparator does not achieve the same combination of Pareto quality, feasibility, and convergence speed as the proposed preference-conditioned PPO framework.

Table 12: Mechanism-level interpretation of the WiMO-CEPPO results. The table summarizes what each component contributes beyond the publication-level numerical comparisons.

Mechanism	Expected role	Evidence from the two evaluations
Context-to-Preference Mapper (CPM)	Infers objective priorities from the current network state instead of relying on fixed scalarization.	Removing CPM reduces reward performance and Pareto quality, showing that context-aware preference selection is necessary when load, interference, and service deficits vary over time.
Conditional envelope	Dampens policy updates when per-objective gradients conflict.	Removing the envelope reduces Pareto quality and increases instability, indicating that conflict-aware updates are important in interference-coupled scheduling.
Preference-conditioned policy	Represents multiple trade-off points with one trained policy.	The same policy architecture supports different operating points without retraining, directly addressing the adaptability requirement in RQ3.
Action projection and feasibility penalties	Ensure that learned actions remain compatible with power, RB-sharing, QoS, and deployment constraints.	Feasibility results show that Pareto gains are meaningful only when interpreted jointly with constraint satisfaction.
Action-space treatment	Determines whether a baseline can represent continuous transmit-power decisions accurately.	DDQN-Envelope is a useful value-based comparator, but its discretized action code book limits continuous-control fidelity compared with actor-critic methods.

Mechanism-Level Evidence

The ablation results are therefore used in this thesis primarily as mechanism validation rather than as another performance ranking. Removing the CPM tests whether preference

adaptation is necessary; removing the conditional envelope tests whether conflict-aware updates are necessary; and using fixed weights tests whether static scalarization is sufficient for the considered scheduling problem. The consistent degradation of the ablated variants supports the central thesis claim that adaptive multi-objective RAN control requires both context-aware preference inference and stable learning under objective conflict.

Ablation and Constraint-Feasibility Interpretation

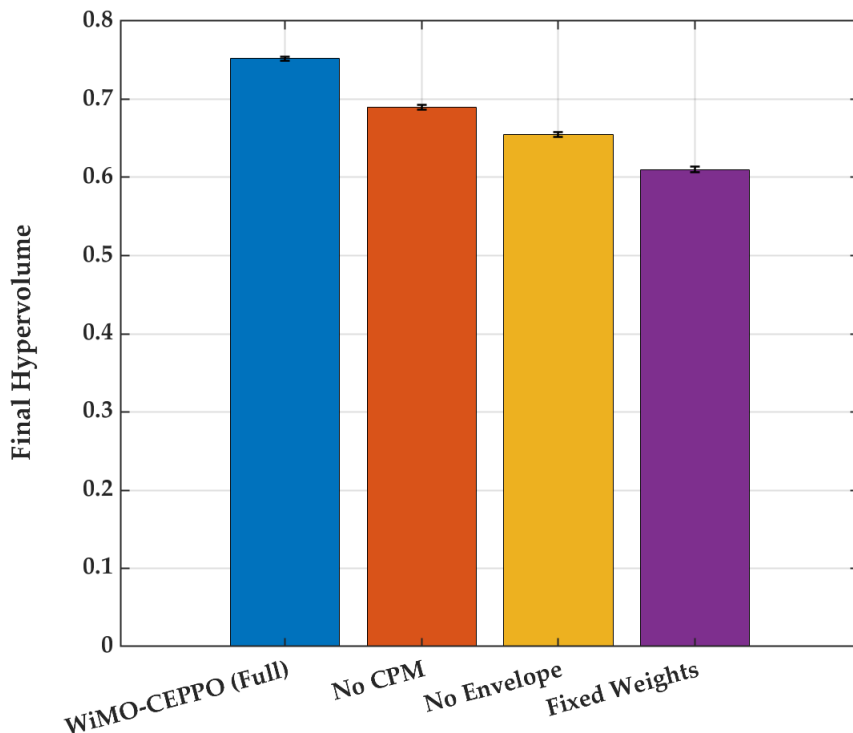


Figure 19: Ablation study of final hypervolume (HV). The full WiMO-CEPPO model achieves the highest HV, indicating the best Pareto-front coverage. Removing the CPM or the conditional envelope reduces performance, confirming the contribution of both mechanisms.

Figures 19 and 20 quantify the effect of removing individual WiMO-CEPPO mechanisms. The full model achieves the best Pareto quality in both evaluations, with the highest hypervolume and the lowest IGD, while every ablation degrades the result. In Publication V, the full model attains $HV = 0.781$ and $IGD = 0.088$, whereas removing the CPM reduces performance to $HV = 0.724$ and $IGD = 0.135$, removing the conditional envelope gives $HV = 0.741$ and $IGD = 0.112$, and using fixed weights further reduces the result to $HV = 0.698$ and $IGD = 0.161$. The same pattern holds in Publication VI, where the full model again provides the strongest Pareto approximation. This confirms that the gains do not come from PPO alone, but from the combination of context-aware preference inference and conflict-aware update stabilization. The CPM improves the policy’s ability to choose an operating point suited to the current network state, while the conditional envelope

prevents overly aggressive updates when the objectives conflict. The ablation figures therefore demonstrate that both mechanisms are necessary for broad Pareto coverage.

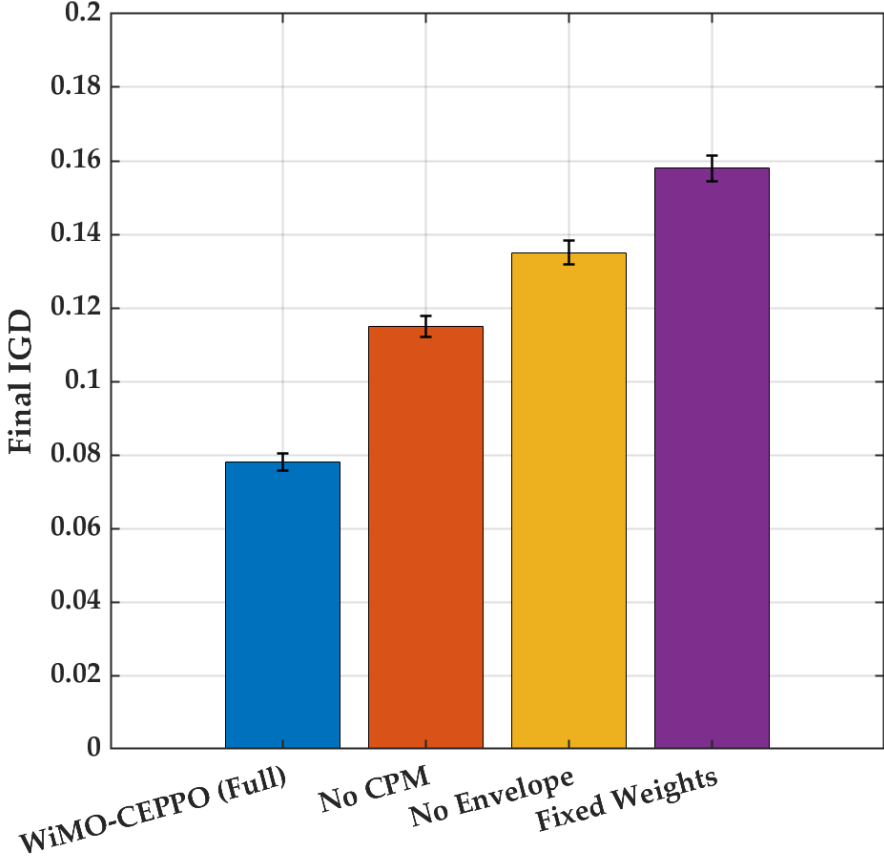


Figure 20: Ablation study of final inverted generational distance (IGD). The full WiMO-CEPPO model achieves the lowest IGD, indicating the closest approximation to the reference Pareto front. The ablated variants show larger deviation, confirming the role of context-aware preference adaptation and conflict-aware updates.

Fig. 19 shows that the complete WiMO-CEPPO framework attains the highest final hypervolume, reflecting superior Pareto coverage. The No-CPM and No-Envelope variants both degrade performance, while the fixed-weight case performs worst, indicating that adaptive preference inference and conflict-aware update stabilization are both important for multi-objective optimization.

Figures 21 and 22 introduce the feasibility perspective, which is crucial for a thesis-level interpretation. In Publication V, the complete WiMO-CEPPO model attains the highest level of constraint satisfaction at 91.2%, compared to 86.7% for MO-PPO, 79.7% for MO-SAC, 74.4% for MO-TD3, and 64.7% for MO-DDPG.

In Publication VI, the full model again delivers the best feasibility at 90.4%, surpassing MO-PPO at 85.9%, MO-SAC at 78.8%, and DDQN-Envelope at 74.4%. These findings confirm

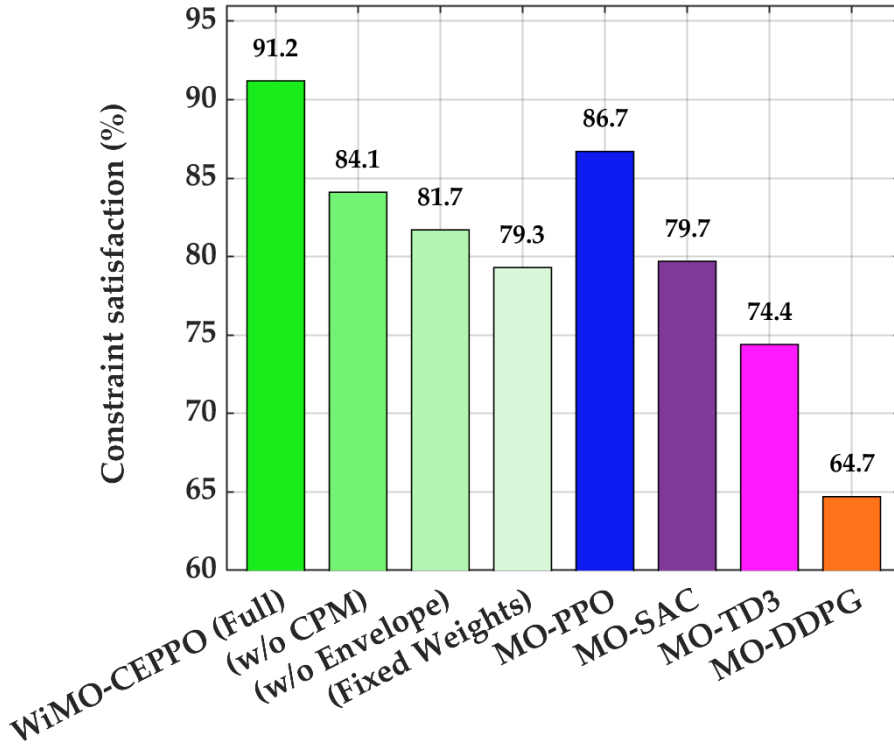


Figure 21: Constraint-satisfaction comparison for WiMO-CEPPO variants and baselines in Publication V. Higher values indicate a larger fraction of feasible scheduling decisions under power, RB-sharing, QoS, and deployment constraints. The full WiMO-CEPPO model again achieves the highest feasibility, while the ablations and baselines degrade progressively.

that the superior Pareto performance of WiMO-CEPPO is not obtained by loosening the scheduling constraints; instead, the learned policy remains feasible with respect to power limits, RB-sharing limits, QoS constraints, and deployment restrictions.

These findings confirm that the superior Pareto performance of WiMO-CEPPO is not obtained by loosening the scheduling constraints; instead, the learned policy remains feasible with respect to power limits, RB-sharing limits, QoS constraints, and deployment restrictions. This matters because even a highly performing policy is only practically valuable if its recommended actions can be implemented by the RAN controller. Overall, the combination of the ablation analysis and the constraint-feasibility results demonstrates that WiMO-CEPPO enhances both the quality of objective trade-offs and the operational executability, which is the central requirement for deployable multi-objective scheduling. The fixed-weight variant serves as the most direct control case: it retains the same learning pipeline but removes context-dependent preference adaptation. Its weaker Pareto coverage indicates that static scalarization is less suitable for wireless environments in which interference, load, and reliability deficits vary jointly over time. By contrast, the CPM improves the policy's ability to infer an operating point that matches the observed network context, while the conditional envelope mitigates destructive updates when the

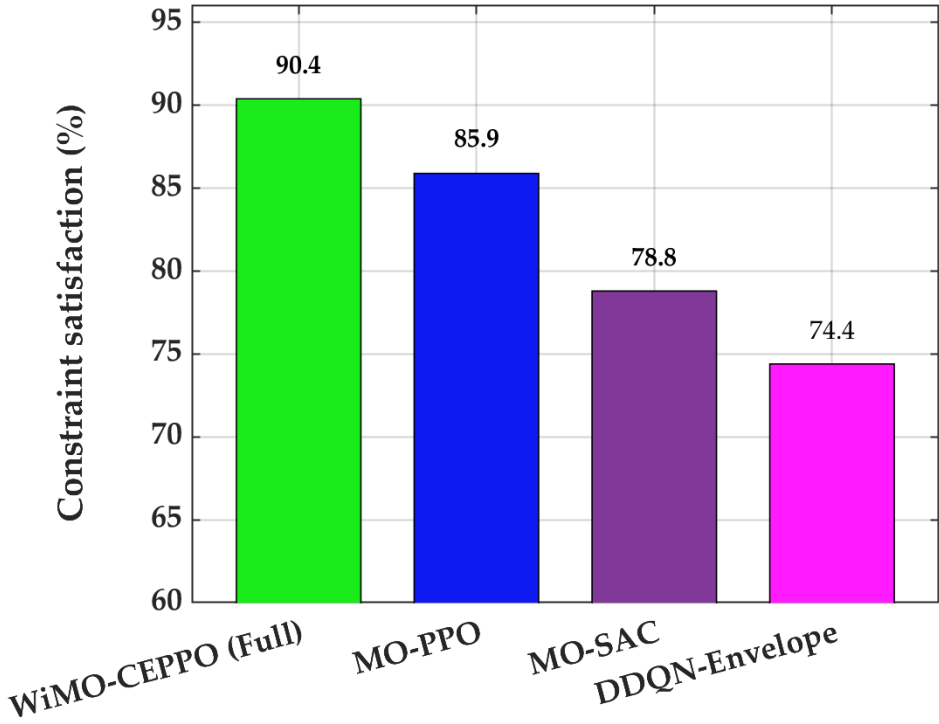


Figure 22: Constraint-satisfaction comparison for WiMO-CEPPO variants and baselines in Publication V. Higher values indicate a larger fraction of feasible scheduling decisions under power, RB-sharing, QoS, and deployment constraints. The full WiMO-CEPPO model achieves the highest feasibility.

objectives conflict.

In this sense, robustness should not be interpreted as invariance to channel or load variations. Rather, it indicates that the policy adapts to changing conditions by shifting toward a different but still feasible operating point instead of applying a fixed action pattern across all regimes.

Feasibility- and Action-Space-Aware Interpretation

A key thesis-level observation is that Pareto quality alone is insufficient for evaluating learning-based RAN control. A method may obtain favorable objective values while still producing actions that violate power, RB-sharing, SINR, or rate constraints. The results should therefore be interpreted jointly through Pareto metrics and constraint satisfaction, since only feasible actions are admissible in deployment-oriented RAN control.

The baseline families should also be interpreted in terms of their action-space assumptions. MO-TD3 and MO-DDPG are appropriate continuous-control baselines for transmit-power adaptation, whereas DDQN-Envelope operates over a discretized RB/power action code-book and therefore relies on a finite action space. The thesis treats these baseline families separately to avoid conflating algorithmic improvements with differences arising

from action-space discretization.

6.7 Chapter Summary and Bridge to Chapter (7)

This chapter completed the thesis argument for **RQ3** by showing that adaptive multi-objective policy learning can complement the model-based scheduling framework developed in the earlier chapters. Publications (V and VI) were kept as separate evaluations because they address different control settings: the former considers continuous and hybrid actor-critic control, whereas the latter examines the O-RAN-oriented extension with a discretized envelope-based comparator. Despite these differences, the two studies lead to the same conclusion: the proposed WiMO-CEPPO framework remains effective across distinct action-space formulations and deployment assumptions.

At the thesis level, the significance of the results lies in the role of the proposed mechanisms. The Context-to-Preference Mapper removes the need for externally fixed weights by adapting preferences to the observed network state, the conditional envelope stabilizes learning when objectives conflict, and the constraint-aware action generation step preserves feasibility at deployment. Together, these components explain why the learned policy can track changing operating conditions while maintaining both Pareto quality and implementable scheduling decisions.

Taken together with Chapters (4 and 5), this chapter completes the three-layer thesis narrative. Chapter (4) established coordination as a mechanism for reducing decision-dependent ICI. Chapter (5) quantified the EE-connectivity trade-off under finite-blocklength constraints. This chapter then showed that these trade-offs can be handled adaptively through reusable policy learning rather than repeated re-optimization. In this sense, the thesis progresses from interference-aware control, to constrained multi-objective optimization, and finally to adaptive learning-based RAN management.

Chapter (7) brings these strands together into unified answers to RQ1-RQ3 and places them in the broader context of sustainable 6G and O-RAN resource management. It also identifies the next technical steps, including scaling the coordination framework to larger O-RAN clusters under E2 latency uncertainty, incorporating richer finite-blocklength models and dynamic user association, extending WiMO-CEPPO to federated or multi-agent settings, and validating the full framework on hardware-in-the-loop O-RAN platforms.

7 Conclusion and Future Work

7.1 Research Questions, Contributions, and Answers

This thesis addressed the central problem of sustainable radio resource management (RRM) for 6G massive machine-type communication (mMTC): how to jointly improve energy efficiency, device connectivity, reliability, and interference control under realistic architectural and physical-layer constraints. The problem was structured around three interdependent challenges- coordination, physical-layer realism, and adaptability- addressed through a progressive sequence of six peer-reviewed contributions across three thesis chapters.

RQ1: Coordination and Computational Sustainability in O-RAN

RQ1: *How can coordination mechanisms among distributed O-DUs in O-RAN architectures be designed to mitigate inter-cell interference (ICI) and enable scalable, energy-efficient scheduling for massive device connectivity?*

Contributions I-II (Chapter (4)) showed that the lightweight selective exchange of scheduling information between adjacent O-DUs, operating over the E2 interface without joint transmission or centralized channel knowledge, substantially reduces decision-dependent inter-cell interference (ICI). The proposed coordination mechanism improved uplink throughput by up to 30%, reduced energy consumption by up to 10%, and increased the number of simultaneously connected devices by a factor of 1.5 relative to uncoordinated O-DU operation at ultra-dense device densities. The second contribution introduced a RISC-V instruction-level energy evaluation framework for O-RAN schedulers, demonstrating that the computational energy consumed by the scheduling algorithm itself is a non-negligible fraction of total O-DU energy consumption under virtualized deployment and cannot be omitted from a sustainability analysis.

Answer to RQ1: Inter-O-DU coordination is achievable within O-RAN's E2 and F1 interface constraints through selective scheduling-table exchange, without requiring joint precoding or centralized spectrum management. Sustainable O-RAN scheduling must account for both radio transmission energy and the computational energy of the scheduling operation, which requires instruction-level profiling on the CU/DU processing platform.

RQ2: Joint EE-Connectivity under FBL in Multi-Cell NOMA

RQ2: *How can multi-cell NOMA-based resource scheduling be optimized to jointly balance energy efficiency and connectivity under realistic constraints such as FBL transmission, dense interference coupling, and heterogeneous device capabilities?*

Contributions III-IV (Chapter (5)) formulated and solved the joint EE-connectivity optimization problem in multi-cell NOMA under both cooperative and non-cooperative interference conditions. The EERS-MTCN algorithm established the non-cooperative EE-connectivity baseline using channel-aware sectorization and showed 35% EE improvement over OMA baselines at moderate device densities. The SRS-MTCN algorithm introduced cooperative inter-cell scheduling, achieving 35–58% EE improvement and maintaining robust connectivity at device densities exceeding 40 active devices per cell. Both algorithms incorporate the FBL rate expression, which accounts for decoding error probability and

channel-dispersion penalties absent from Shannon capacity models. The FBL analysis further revealed that the EE-optimal blocklength lies in the range $L = 300\text{--}400$ channel uses for the evaluated multi-cell NOMA settings, confirming that IFBL-based optimizations overestimate achievable EE and connectivity and that FBL modeling changes the optimal scheduling operating point.

Answer to RQ2: Multi-cell NOMA scheduling under FBL constraints requires a joint formulation that couples EE, connectivity, decoding error probability, and inter-cell interference in a single multi-objective program. Shannon-capacity models are not an acceptable approximation for 6G mMTC; they overestimate achievable performance and produce sub-optimal scheduling decisions. Cooperative inter-cell coordination provides 35–58% EE gains relative to non-cooperative IFBL baselines.

RQ3: Adaptive Multi-Objective Policy Learning

RQ3: *How can MORL be employed to learn a single, adaptive scheduling policy that generalizes across preference trade-offs and dynamically responds to time-varying network conditions in 6G systems?*

Contributions V–VI (Chapter 6) introduced WiMO-CEPPO, a preference-conditioned MORL framework that learns a single scheduling policy navigating EE, connectivity, reliability, and interference trade-offs without repeated re-optimization. The Context-to-Preference Mapper (CPM) infers appropriate objective weights from observed network state, eliminating reliance on externally supplied preference vectors. The conditional envelope mechanism attenuates policy updates under strong inter-objective gradient conflict, improving Pareto coverage and training stability. Constraint-aware action generation enforces RB-assignment and power feasibility structurally.

WiMO-CEPPO achieves $HV = 0.752$ and $IGD = 0.078$ in the four-objective setting, improving HV by 38.7% and reducing IGD by 58.7% over the strongest baseline (MO-PPO), with a constraint-satisfaction rate of 91.2% and a 40% reduction in training episodes to target quality. Contribution 6 extended the framework to five objectives with a continuous action space and explicit deployment constraints, further improving HV to 0.781 (34.2% improvement over MO-PPO) and confirming compatibility with O-RAN control-loop latency requirements.

Answer to RQ3: A single preference-conditioned MORL policy, guided by context-inferred preference vectors and trained with conflict-aware gradient management, can generalize across multiple Pareto operating points and respond adaptively to changing network conditions without retraining. The CPM, conditional envelope, and constraint-aware action generation are all necessary for achieving competitive Pareto quality, feasibility, and distributional robustness.

Unified Answer to the Central Research Problem

Across the three chapters, the thesis develops a coherent methodological progression: coordination reduces the severity of the interference environment; analytical multi-objective optimization characterizes the achievable EE-connectivity frontier under FBL constraints; and adaptive MORL navigates that frontier in real time under non-stationary conditions.

Sustainable massive connectivity in 6G cannot be achieved by optimizing a single metric, nor by treating coordination, physical-layer realism, and adaptability as independent problems. Their joint treatment within a consistent RRM framework is necessary, and the proposed three-layer approach demonstrates that it is also technically achievable.

7.2 Broader Implications for 6G and O-RAN

The findings carry three broader implications for 6G scheduling and O-RAN control.

First, inter-cell interference in dense mMTC networks is decision-dependent and must be addressed at the scheduling and coordination layer, not only through physical-layer signal processing. Architectural mechanisms for inter-O-DU information sharing are therefore a necessary component of sustainable massive connectivity, not merely a performance enhancement.

Second, sustainable 6G scheduling must account for both radio and computational energy. Algorithms that improve transmission efficiency but incurs significant computational cost may not be deployable in virtualized O-DU environments where computation and cooling energy are co-managed. The instruction-level energy framework developed in Contribution 2 provides a methodology for this accounting.

Third, effective adaptive scheduling requires explicit mechanisms for preference control and conflict management. Static scalarization is incompatible with dynamic network environments, preference inference from observed context is necessary for deployment without manual operator intervention, and conflict-aware gradient management is necessary to achieve broad Pareto coverage under the structurally coupled objectives of 6G uplink scheduling.

7.3 Limitations and Future Work

The following limitations and corresponding future directions are aligned point-by-point.

7.3.1 L1: Simplified O-DU coordination model and scalability.

The coordination studies focus on a small cluster of adjacent O-DUs under idealized E2 interface conditions and do not model signaling delay, control-plane congestion, or asynchronous telemetry reporting. Extending the coordination framework to larger heterogeneous O-RAN deployments with realistic E2 latency distributions and partial observability [48] is a direct continuation. Multi-cluster hierarchical coordination, where a Non-RT RIC manages inter-cluster boundaries while Near-RT RICs manage intra-cluster scheduling, is a promising scalability path [38].

7.3.2 L2: FBL model assumptions and coding complexity.

The FBL analysis relies on the standard normal approximation [21], which may require refinement for very short packets ($L < 100$ channel uses) or more complex coding schemes such as polar codes or rate-compatible LDPC codes. Future work should investigate whether the EE-optimal blocklength result ($L = 300\text{--}400$ for the evaluated settings) generalizes under more accurate FBL bounds for structured codes, as studied for single-cell NOMA in

recent works [78, 79].

7.3.3 L3: Static user association and mobility.

The current analytical framework assumes structured NOMA grouping under controlled mobility. More general mobility models, heterogeneous traffic classes, and dynamic user association would extend the scope of the EE-connectivity analysis and test the robustness of the proposed algorithms under conditions closer to live deployments [87].

7.3.4 L4: Hardware-in-the-loop validation of WiMO-CEPPO.

WiMO-CEPPO is validated in simulation. Full deployment would require evaluating inference latency, memory footprint, telemetry noise tolerance, actuation delay, and interaction with concurrently executing xApps on an O-RAN-compliant Near-RT RIC platform [37, 35]. The instruction-level energy framework developed in Contribution 2 provides a partial methodology for this validation and could be extended to profile the CPM and policy network on candidate edge processors [43].

7.3.5 L5: Distributed and transferable learning.

WiMO-CEPPO uses a single centralized agent. Scaling to larger networks may require federated learning [95], multi-agent extensions, or transfer learning to reduce retraining cost across cells and deployment environments. Meta-learning and domain randomization are promising directions for improving robustness under topology and traffic shifts, particularly relevant for operators deploying across heterogeneous cell types.

7.3.6 L6: Network topology and heterogeneous access.

The framework currently assumes a terrestrial homogeneous cellular deployment. Extensions to heterogeneous radio access, including aerial access nodes, reconfigurable intelligent surfaces, and non-terrestrial network segments relevant to emerging 6G architectures [9, 13] would expose new trade-offs in cross-layer interference, mobility, and adaptive preference control. Such extensions are a natural continuation of the thesis framework once the terrestrial baseline is fully validated.

7.4 Concluding Remarks

This thesis establishes that sustainable massive connectivity in 6G requires the integration of three complementary capabilities: coordination to suppress decision-dependent inter-cell interference, physically grounded multi-objective optimization to characterize the EE-connectivity frontier under FBL constraints, and adaptive reinforcement learning to navigate that frontier in real time without repeated re-optimization. The proposed methods do not close all practical deployment gaps, but they provide a methodologically coherent, technically grounded, and experimentally validated foundation for future 6G RRM systems that are interference-aware, energy-conscious, reliability-aware, and adaptable to changing network objectives.

List of Figures

1	Overview of major 6G design challenges and the mMTC ecosystem. Key challenges that shape next-generation communication systems include sustainability constraints, advanced hardware design requirements, power-efficiency considerations, and increasing pressure on the radio spectrum. These issues intersect with essential concerns related to privacy and security, particularly as networks evolve to support dense, heterogeneous deployments of IoT devices (mMTC) and industrial, healthcare, agricultural, and UAV-enabled services.....	13
2	PhD thesis master map illustrating the unified research framework, linking the central problem of multi-objective RRM for 6G massive connectivity with three sequential and interdependent challenges: coordination, physical-layer optimization, and adaptability. The figure shows how these challenges are decomposed into the corresponding research questions and addressed through the methodological approaches developed in this thesis. It also maps each challenge to the technical Chapters (4–6) and to the key contributions (I–VI).	17
3	Conceptual O-RAN coordination and scheduling architecture for massive mMTC access. The figure illustrates hierarchical control across the Non-RT RIC, Near-RT RIC, O-CU, and distributed O-DUs, where inter-O-DU coordination supports scheduling and resource-sharing decisions to mitigate cross-cell interference. Real-time feedback from the radio access layer enables adaptive resource allocation, improved reliability, and energy-efficient operation in dense 6G deployments.....	50
4	Cumulative distribution function (CDF) of uplink throughput for Inter O-DU coordinated and stand-alone O-DU scenarios under NOMA and OMA schemes. The figure highlights the throughput gains achieved through inter O-DU coordination and illustrates the performance advantage of NOMA over conventional OMA across different percentile levels. Reproduced from our (Publication I).	53
5	Combined energy consumption analysis of computational complexity and radio transmission for OMA, RAAP-NOMA, EEPs-MTCN-NOMA, and ESAP-NOMA schemes under varying user connection densities (16–48 active devices per cell). The results demonstrate the scalability impact of increasing user density on total energy expenditure, and highlight the relative efficiency of advanced NOMA-based approaches compared to conventional OMA. Reproduced from our Publication (II).....	54

6	Total energy consumption, decomposed into computational complexity and radio transmission, for OMA, RAAP-NOMA, EEPS-MTCN-NOMA, and ESAP-NOMA as a function of the number of active devices per cell. The figure shows that total energy increases with network load for all schemes, while the exhaustive-search baseline incurs the highest cost because of its substantially larger computational burden. The results highlight the need to evaluate sustainability using the combined energy cost rather than radio performance alone.	54
7	Comparison of achievable system energy efficiency versus the number of NOMA users multiplexed per resource block for ESAP-NOMA, EERS-MTCN, and RAAP-NOMA. ESAP-NOMA attains the highest energy efficiency, the proposed EERS-MTCN consistently outperforms RAAP-NOMA, and all schemes exhibit a decreasing trend as the NOMA multiplexing order increases.	64
8	Achievable system energy efficiency of the proposed EERS-MTCN algorithm as a function of the NOMA multiplexing order for different active-device populations per cell ($N = \{150, 200, 250, 300\}$).	64
9	Illustrative visualization of the QoS-based connectivity and energy-efficiency relationship for Publication (III). (a) QoS-connected devices versus the number of active devices per cell. (b) Corresponding EE-connectivity trade-off for ESAP-NOMA, EERS-MTCN, and RAAP-NOMA, illustrating the relative behavior of the three scheduling algorithms.	65
10	System-level EE comparison for 2 NOMA users per RB under varying user connection densities. The proposed EERS-MTCN (maxEE and minInterference variants) outperforms RAAP-NOMA and the reference algorithm across all density levels, demonstrating improved scalability and EE in moderately loaded scenarios.	67
11	System-level EE comparison for 4 NOMA users per RB as a function of user connection density. Increasing the number of multiplexed NOMA users enhances EE performance for the proposed EERS-MTCN schemes, highlighting their robustness and superior interference management compared to baseline algorithms.	68
12	System-level EE performance with 8 NOMA users per RB under different user densities. The results illustrate the impact of higher NOMA multiplexing on EE and confirm that the proposed EERS-MTCN approaches maintain performance advantages over RAAP-NOMA and the reference method in dense mMTC deployments.	68
13	Reliability-weighted connectivity under finite-blocklength (FBL) transmission versus the number of active devices per cell. The dotted line denotes the ideal connected-device limit, while the proposed SRS-MTCN remains closest to this upper bound and outperforms the minInterference variant, the cooperative scheduling baseline, and RAAP-NOMA across all evaluated loads.	69

14	WiMO-CEPPO architecture. The CPM maps the observed network state s_t to a preference vector $\mathbf{w}_t$. The preference-conditioned policy $\pi_\theta(a_t s_t, \mathbf{w}_t)$ generates candidate RB-assignment and power-control actions. The conditional envelope stabilizes PPO updates under inter-objective gradient conflict, while the feasibility projector maps raw actions to admissible scheduling decisions before execution.	78
15	Sample-efficiency comparison of WiMO-CEPPO and baselines in terms of episodes required to reach target performance in Publication (V). Lower values indicate faster convergence, and WiMO-CEPPO is consistently the most sample-efficient.	83
16	Sample-efficiency comparison of WiMO-CEPPO and baselines in terms of episodes required to reach target performance in Publication (VI). Lower values indicate faster convergence, and WiMO-CEPPO is consistently the most sample-efficient.	84
17	Mean reward AUC comparison for WiMO-CEPPO and baseline MORL methods over the 4,000-episode training horizon. Higher AUC indicates stronger cumulative reward acquisition over the training horizon.....	84
18	Area under the utility-loss curve for WiMO-CEPPO and baseline MORL methods over the 4,000-episode training horizon. Lower AUC indicates smaller cumulative deviation from the context-conditioned utility target and more stable preference-conditioned learning.....	85
19	Ablation study of final hypervolume (HV). The full WiMO-CEPPO model achieves the highest HV, indicating the best Pareto-front coverage. Removing the CPM or the conditional envelope reduces performance, confirming the contribution of both mechanisms.	88
20	Ablation study of final inverted generational distance (IGD). The full WiMO-CEPPO model achieves the lowest IGD, indicating the closest approximation to the reference Pareto front. The ablated variants show larger deviation, confirming the role of context-aware preference adaptation and conflict-aware updates.	89
21	Constraint-satisfaction comparison for WiMO-CEPPO variants and baselines in Publication V. Higher values indicate a larger fraction of feasible scheduling decisions under power, RB-sharing, QoS, and deployment constraints. The full WiMO-CEPPO model again achieves the highest feasibility, while the ablations and baselines degrade progressively.....	90
22	Constraint-satisfaction comparison for WiMO-CEPPO variants and baselines in Publication V. Higher values indicate a larger fraction of feasible scheduling decisions under power, RB-sharing, QoS, and deployment constraints. The full WiMO-CEPPO model achieves the highest feasibility.	91

List of Tables

1	Mapping of the research questions (RQs), methods, and thesis chapters ...	19
2	Comparison of selected related works on sustainable scheduling under FBL in mMTC/URLLC networks. “ICI-aware” indicates whether both intra-cell and inter-cell interference are explicitly considered.	30
3	Consolidated research gaps, motivating literature, and corresponding thesis contributions.	38
4	Unified notation used in this thesis and reconciliation with symbols used across the papers.	41
5	Reference summary of common constraints used across Chapters (4–6). <i>For symbol definitions see Table 4.</i>	49
6	Gap-to-mechanism mapping for the WiMO-CEPPO framework.	75
7	Evaluation roles of the two WiMO-CEPPO publications. The table clarifies how the studies jointly support RQ3 without treating their numerical results as one unified benchmark.	81
8	Simulation Environment and deployment parameters used across the WiMO-CEPPO evaluations. Values follow the system model and publication-specific assumptions.	82
9	Training protocol and WiMO-CEPPO-specific configuration. Values follow the publication-specific simulation settings.	82
10	Performance anchors for Publication (V): continuous/hybrid action evaluation with MO-PPO, MO-SAC, MO-TD3, and MO-DDPG baselines. Results are averaged over $N_s = 5$ independent seeds. Interference is reported as normalized interference mitigation, where higher values indicate lower cross-link interference.	86
11	Performance anchors for Publication (VI): O-RAN-oriented DDQN-Envelope, evaluation with MO-PPO, MO-SAC, and DDQN-Envelope baselines. Results are averaged over $N_s = 5$ independent seeds. Interference is reported as normalized interference mitigation, where higher values indicate lower cross-link interference.	87
12	Mechanism-level interpretation of the WiMO-CEPPO results. The table summarizes what each component contributes beyond the publication-level numerical comparisons.	87

References

- [1] Latif, Aamir; Alam, Muhammad Mahtab; Le Moullec, Yannick, "Inter O-DUs Coordination for Scheduling of Massive Users in O-RAN," in *2023 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit 2023)*, pp. 1–6, 2023.
- [2] Latif, Aamir; Elgarhy, Osama; Le Moullec, Yannick; Alam, Muhammad Mahtab, "Energy Consumption Evaluation of NOMA-based Sustainable Scheduling in 6G O-RAN," in *2024 International Wireless Communications and Mobile Computing (IWCMC 2024)*, pp. 484–489, 2024.
- [3] Latif, Aamir; Elgarhy, Osama; Le Moullec, Yannick; Alam, Muhammad Mahtab, "Energy Efficiency and Connectivity Maximization for Massive Machine Type Communication in Multi-Cell NOMA," in *2024 19th Biennial Baltic Electronics Conference (BEC 2024)*, pp. 1–6, IEEE, 2024.
- [4] Latif, Aamir and Elgarhy, Osama and Mahmood, Nurul Huda and Le Moullec, Yannick and Alam, Muhammad Mahtab, "Sustainable Massive Connectivity Maximization Scheduling Scheme for 6G," *IEEE Transactions on Green Communications and Networking*, pp. 2075–2090, January 2026.
- [5] A. Latif, Y. Le Moullec, and M. M. Alam, "Multi-objective policy design using reinforcement learning in 6g networks," in *2026 International Wireless Communications and Mobile Computing (IWCMC)*, pp. 1352–1357, 2026.
- [6] A. Latif, Y. L. Moullec, and M. M. Alam, "Sustainable reward mechanism for multi-objective policy design using reinforcement learning in 6g networks," in *2026 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, pp. 527–533, 2026.
- [7] L. Li, "Research on future 6G green wireless networks," *Green Technologies and Sustainability*, vol. 3, p. 100156, 2025.
- [8] N. H. Mahmood *et al.*, "Six key features of machine type communication in 6G," in *Proc. 2nd 6G Wireless Summit*, pp. 1–5, 2020.
- [9] C.-X. Wang, X. You, *et al.*, "On the road to 6G: Visions, requirements, key technologies, and testbeds," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 2, pp. 905–974, 2023.
- [10] International Telecommunication Union, "Framework and overall objectives of the future development of IMT for 2030 and beyond," Recommendation M.2160-O, ITU-R, Nov 2023.
- [11] International Telecommunication Union, "IMT vision—framework and overall objectives of the future development of IMT for 2020 and beyond," Recommendation M.2083-O, ITU-R, Sep 2015.

- [12] 3GPP, "TR 38.913 V18.0.0: Study on scenarios and requirements for next generation access technologies (release 18)," tech. rep., 3rd Generation Partnership Project (3GPP), May 2024. Available via ETSI as ETSI TR 138 913 V18.0.0 (2024-05).
- [13] 6G-IA Vision Working Group, "European vision for the 6g network ecosystem, white paper, version 2.0," Nov. 2024.
- [14] N. H. Mahmood, H. Alves, O. A. López, M. Shehab, D. P. M. Osorio, and M. Latva-Aho, "Six key features of machine type communication in 6g," pp. 1–5, 2020.
- [15] I. Ahmad *et al.*, "Sustainability in 6G networks: Vision and directions," in *Proc. IEEE Conference on Standards for Communications and Networking (CSCN)*, pp. 202–208, 2023.
- [16] L. Blanco, E. Zeydan, S. Barrachina-Munoz, F. Rezazadeh, L. Vettori, and J. Mangues-Bafalluy, "A novel approach for scalable and sustainable 6G networks," *IEEE Open Journal of the Communications Society*, vol. 5, pp. 1673–1692, 2024.
- [17] Q. Wang, Y. Liu, Y. Wang, X. Xiong, J. Zong, J. Wang, and P. Chen, "Resource allocation based on radio intelligence controller for open ran towards 6g," *IEEE Access*, 2023.
- [18] O-RAN Alliance nGRG, "dapps for real-time RAN control: Use cases and requirements," Tech. Rep. RR-2024-10, O-RAN next Generation Research Group (nGRG), 2024.
- [19] A. Elyasi *et al.*, "O-RAN xapps: Survey and research challenges," 2025. Preprint (commonly circulated as "xApps_COMNET_2025.pdf").
- [20] C. Cinemre and N. H. Mahmoodi, "xapp conflict mitigation with scheduler for open RAN," 2025. arXiv preprint.
- [21] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.
- [22] Y. Liu, W. Zhou, *et al.*, "Dynamic preference inference network: Improving sample efficiency for multi-objective reinforcement learning by preference estimation," *Knowledge-Based Systems*, p. 112512, 2024.
- [23] Y. Yang *et al.*, "Preference controllable reinforcement learning with advanced multi-objective optimization," in *Proceedings of Machine Learning Research*, 2025.
- [24] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design science in information systems research," *MIS Quarterly*, vol. 28, no. 1, pp. 75–105, 2004.
- [25] 3rd Generation Partnership Project (3GPP), "Study on Scenarios and Requirements for Next Generation Access Technologies (Release 18)," Technical Report TR 38.913, V18.0.0, 3GPP, Sophia Antipolis, France, March 2023. Outlines use cases and requirements for 5G and pre-6G systems, including mMTC and URLLC enhancements.

- [26] ITU, "Report ITU-R M.2412-O: Guidelines for evaluation of radio interface technologies for IMT-2020," tech. rep., International Telecommunication Union – Radiocommunication Sector, 2017.
- [27] A. R. Hevner and S. Chatterjee, *Design Research in Information Systems: Theory and Practice*, vol. 22 of *Integrated Series in Information Systems*. Springer, 2010.
- [28] O-RAN Working Group 1, "O-RAN architecture description v06.00," Tech. Rep. O-RAN.WG1.O-RAN-Architecture-Description-v06.00, O-RAN Alliance, 2022.
- [29] ETSI, "ETSI TS 103 982 V8.0.0 (2024-01): Publicly available specification (pas); O-RAN architecture description," tech. rep., European Telecommunications Standards Institute (ETSI), Jan. 2024.
- [30] O-RAN Working Group 2, "O-RAN AI/ML workflow description and requirements v1.03," tech. rep., O-RAN Alliance, 2021.
- [31] ETSI / 3GPP, "ETSI TS 138 401 V18.2.0 (2024-05): 5g; NG-RAN; architecture description (3gpp ts 38.401 release 18)," tech. rep., European Telecommunications Standards Institute (ETSI), May 2024.
- [32] ETSI, "ETSI TS 103 983 V3.1.0 (2024-01): Publicly available specification (pas); A1 interface: General aspects and principles," tech. rep., European Telecommunications Standards Institute (ETSI), Jan. 2024.
- [33] ETSI, "ETSI TS 103 985 V1.1.0 (2024-01): Publicly available specification (pas); A1 interface: Use cases and requirements," tech. rep., European Telecommunications Standards Institute (ETSI), Jan. 2024.
- [34] ETSI, "ETSI TS 103 987 V4.3.0 (2025-05): Publicly available specification (pas); A1 interface: Application protocol," tech. rep., European Telecommunications Standards Institute (ETSI), May 2025.
- [35] A. Elyasi, A. Ashdown, K. M. Rumman, and F. Restuccia, "O-RAN xapps: Survey and research challenges." SSRN, 2025.
- [36] M. Hoffmann, S. Janji, A. Samorzewski, L. Kulacz, C. Adamczyk, M. Dryjański, P. Kryszkiewicz, A. Kliks, and H. Bogucka, "Open RAN xapps design and evaluation: Lessons learnt and identified challenges," *IEEE Journal on Selected Areas in Communications*, vol. 42, pp. 473–486, Feb. 2024.
- [37] M. Polese, M. Dohler, F. Dressler, M. Erol-Kantarci, R. Jana, R. Knopp, and T. Melodia, "Empowering the 6G cellular architecture with open RAN," *IEEE Journal on Selected Areas in Communications*, vol. 42, pp. 245–262, Feb. 2024.
- [38] K. Alam, M. A. Habibi, M. Tammen, D. Krummacker, W. Saad, M. Di Renzo, T. Melodia, X. Costa-Pérez, M. Debbah, A. Dutta, and H. D. Schotten, "A comprehensive tutorial and survey of O-RAN: Exploring slicing-aware architecture, deployment options, use cases, and challenges." arXiv preprint, 2024.

- [39] B. Balasubramanian *et al.*, "RIC: A RAN intelligent controller platform for AI-enabled cellular networks," *IEEE Internet Computing*, vol. 25, no. 2, pp. 7–17, 2021.
- [40] ITU, "Report ITU-R M.2160-0: Framework and overall objective and future development of IMT for 2030 and beyond," tech. rep., International Telecommunication Union – Radiocommunication Sector, 2023.
- [41] O-RAN Alliance nGRG, "Use case analysis related to green communication in O-RAN," Tech. Rep. RR-2024-08, O-RAN next Generation Research Group (nGRG), 2024.
- [42] L. Kundu, X. Lin, and R. Gadiyar, "Towards energy efficient RAN: From industry standards to trending practice." arXiv preprint, 2024.
- [43] R. Shirkhani, S. Maxenti, L. Bonati, N. Mohamadi, M. Elkael, U. Hashmi, *et al.*, "TENORAN: Automating fine-grained energy efficiency profiling in open RAN systems." arXiv preprint, 2026.
- [44] M. Alavirad, "Use case analysis related to green communication in O-RAN," Tech. Rep. RR-2024-08, O-RAN Alliance, Next Generation Research Group (nGRG), 2024.
- [45] L. Wang, J. Zhou, M. Ma, and X. Niu, "Minimizing energy consumption of iot devices for o-ran based iot systems," *Energy Reports*, vol. 9, pp. 379–388, 2023.
- [46] L. M. P. Larsen, H. L. Christiansen, S. Ruepp, and M. S. Berger, "Toward greener 5g and beyond radio access networks—a survey," *IEEE Open Journal of the Communications Society*, vol. 4, pp. 768–797, 2023.
- [47] F. Aslan, G. Iosifidis, J. A. Ayala-Romero, A. Garcia-Saavedra, and X. Costa-Perez, "Fair resource allocation in virtualized o-ran platforms," *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, vol. 8, no. 1, pp. 1–34, 2024.
- [48] S. Ebrahimi, F. Bouali, and O. Haas, "Resource management from single-domain 5G to end-to-end 6G network slicing: A survey," *IEEE Communications Surveys & Tutorials*, vol. 26, no. 4, pp. 2836–2866, 2024.
- [49] M. W. Akhtar *et al.*, "Exploiting NOMA for radio resource efficient traffic steering use-case in O-RAN," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, pp. 5771–5776, 2022.
- [50] V. Ranjbar *et al.*, "Cell-free mMIMO support in the O-RAN architecture: A PHY layer perspective for 5G and beyond networks," *IEEE Communications Standards Magazine*, vol. 6, no. 1, pp. 28–34, 2022.
- [51] M. Zangooei *et al.*, "Flexible RAN slicing in open RAN with constrained multi-agent reinforcement learning," *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 7, pp. 1875–1890, 2024.
- [52] B. Toure, D. Tsilimantos, and T. Giannakas, "Multi-objective scheduling in wireless networks with deep reinforcement learning," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, 2025.

- [53] M. Lotfi *et al.*, "Multi-agent reinforcement learning-based resource management in O-RAN for beyond 5G and 6G networks," *IEEE Open Journal of the Communications Society*, 2025.
- [54] C. B. Mwakwata, O. Elgarhy, M. M. Alam, Y. Le Moullec, S. Parand, K. Trichias, and K. Ramantas, "Cooperative scheduler to enhance massive connectivity in 5G and beyond by minimizing interference in OMA and NOMA," *IEEE Systems Journal*, vol. 16, no. 3, pp. 5044–5055, 2022.
- [55] A. Rabee and I. Barhumi, "Energy efficient resource allocation approach for uplink NOMA multi-cell systems based on multi-agent DRL," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, 2024.
- [56] B. Rashid *et al.*, "Energy efficient resource allocation for uplink MC-NOMA based heterogeneous small cell networks with wireless backhaul," *IEEE Transactions on Vehicular Technology*, vol. 72, pp. 3419–3429, 2023.
- [57] S. M. Jameel, A. K. Tareen, H. Jung, and S. Hassan, "Energy efficiency optimization of CoMP C-NOMA system using markov decision process," in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, 2024.
- [58] Y. Zhao, T. Peng, Y. Guo, Y. Niu, and W. Wang, "An intelligent scheme for energy-efficient uplink resource allocation with QoS constraints in 6G networks," *IEEE Transactions on Network and Service Management*, vol. 22, no. 1, pp. 255–269, 2025.
- [59] Z. Yang, Z. Ding, P. Fan, and N. Al-Dhahir, "A general power allocation scheme to guarantee QoS in downlink and uplink NOMA systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7244–7257, 2016.
- [60] S. M. R. Islam, N. Avazov, O. A. Dobre, and K.-s. Kwak, "Power-domain non-orthogonal multiple access (noma) in 5g systems: Potentials and challenges," *IEEE Communications Surveys Tutorials*, vol. 19, no. 2, pp. 721–742, 2017.
- [61] Shirvanimoghaddam, Mahyar *et al.*, "Massive Non-Orthogonal Multiple Access for Cellular IoT: Potentials and Limitations," *IEEE Communications Magazine*, vol. 55, no. 9, pp. 55–61, 2017.
- [62] F. Fang, H. Zhang, J. Cheng, S. Roy, and V. C. Leung, "Joint user scheduling and power allocation optimization for energy-efficient noma systems with imperfect csi," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 12, pp. 2874–2885, 2017.
- [63] D. Zhai, R. Zhang, L. Cai, B. Li, and Y. Jiang, "Energy-efficient user scheduling and power allocation for noma-based wireless networks with massive iot devices," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1857–1868, 2018.
- [64] S. Han, X. Xu, X. Tao, and P. Zhang, "Joint power and sub-channel allocation for secure transmission in noma-based mmTC networks," *IEEE Systems Journal*, vol. 13, no. 3, pp. 2476–2487, 2019.

- [65] Han Shujun *et al.*, "Energy-Efficient Short Packet Communications for Uplink Noma-Based Massive MTC Networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 12, pp. 12066–12078, 2019.
- [66] Liu Binghong *et al.*, "Resource Allocation for Energy-Efficient MEC in NOMA-Enabled Massive IoT Networks," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 4, pp. 1015–1027, 2021.
- [67] Zhai, Daosen *et al.*, "Delay Minimization for Massive Internet of Things with Non-Orthogonal Multiple Access," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 3, pp. 553–566, 2019.
- [68] J. Chen, L. Zhang, Y.-C. Liang, and S. Ma, "Optimal resource allocation for multicarrier NOMA in short packet communications," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 2, pp. 2141–2156, 2020.
- [69] Y. Liu and S. Zhang, "Evolution of NOMA towards next generation multiple access (NGMA) for 6G," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 4, pp. 1037–1071, 2022.
- [70] B. E. Y. Belmekki and M. Alouini, "Noma as the next-generation multiple access in nonterrestrial networks," *Proceedings of the IEEE*, 2024.
- [71] A. J. Muhammed, Z. Ma, Z. Ding, M. Xiao, P. Fan, W. Xu, G. Liu, and Z. Zhang, "Resource allocation for energy-efficient noma system in coordinated multi-point networks," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 2, pp. 1577–1591, 2021.
- [72] N. Trabelsi, L. C. Fourati, and C. S. Chen, "Interference management in 5g and beyond networks: A comprehensive survey," *Computer Networks*, vol. 239, p. 110159, 2024.
- [73] M. Zeng, A. Yadav, O. A. Dobre, and H. V. Poor, "Energy-efficient joint user-rb association and power allocation for uplink hybrid noma-oma," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 5119–5131, 2019.
- [74] J. Cui, Y. Liu, Z. Ding, P. Fan, and A. Nallanathan, "Qoe-based resource allocation for multi-cell noma networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 9, pp. 6160–6176, 2018.
- [75] Q. Song, F.-C. Zheng, and J. Luo, "Optimization of energy efficiency for uplink murllc over multiple cells using cooperative multiagent reinforcement learning," *IEEE Internet of Things Journal*, vol. 11, no. 9, pp. 16351–16363, 2024.
- [76] D. Wu and R. Negi, "Effective capacity: A wireless link model for support of quality of service," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 630–643, 2003.
- [77] Y. Liu, J. Lee, C. Sun, Y. Han, and W. Chen, "Energy efficient scheduling for short packet communications with finite blocklength coding," *IEEE Transactions on Green Communications and Networking*, vol. 8, no. 4, pp. 1645–1660, 2024.

- [78] T. Zhong, Y. Wang, J. Wang, X. Xia, and X. Gao, "Minimum BLER NOMA design with finite blocklength," *IEEE Transactions on Wireless Communications*, 2024.
- [79] Y. Liu, J. Lee, C. Sun, Y. Han, and W. Chen, "Energy efficient scheduling for short packet communications with finite blocklength coding," *IEEE Transactions on Green Communications and Networking*, vol. 8, no. 4, pp. 1645–1660, 2024.
- [80] J. Gao, Y. Wu, T. Li, and W. Zhang, "Energy efficiency of MIMO massive unsourced random access with finite blocklength," *IEEE Wireless Communications Letters*, vol. 12, no. 8, pp. 1408–1412, 2023.
- [81] S. R. Sabuj, S. Ahmed, and H.-S. Jo, "Multiple CUAV-enabled mMTC and uRLLC services: Review of energy efficiency and latency performance," *IEEE Transactions on Green Communications and Networking*, vol. 7, no. 3, pp. 1369–1382, 2023.
- [82] J. Wang, W. Cheng, and W. Zhang, "FBC-enhanced ε -effective capacity optimization for NOMA," 2024.
- [83] Y. Zhao, W. Gao, Y. Zhu, Y. Hu, and J. Gross, "Energy efficiency maximization for multi-node IoT networks operating with finite blocklength codes," in *Proc. 19th International Symposium on Wireless Communication Systems (ISWCS)*, 2024.
- [84] M. Soleymani, I. Santamaria, E. Jorswieck, R. Schober, and L. Hanzo, "Energy efficiency optimization of finite block length STAR-RIS-aided MU-MIMO broadcast channels," 2025.
- [85] D.-D. Tran, V.-N. Ha, S. K. Sharma, T. T. Nguyen, S. Chatzinotas, and P. Popovski, "Energy-efficient NOMA for 5G heterogeneous services: A joint optimization and deep reinforcement learning approach," *IEEE Transactions on Communications*, vol. 73, no. 4, pp. 2448–2465, 2025.
- [86] S. Thapliyal, R. Pandey, and C. Charan, "Evaluating energy harvesting UAV-NOMA network with random user pairing in the finite blocklength regime," *Vehicular Communications*, vol. 49, p. 100831, 2024.
- [87] Y. Huang, Y. Jiang, F.-C. Zheng, P. Zhu, *et al.*, "Energy-efficient optimization in user-centric cell-free massive MIMO systems for uRLLC with finite blocklength communications," *IEEE Transactions on Vehicular Technology*, pp. 1–14, 2024.
- [88] Y. Zhang, W. Cheng, and W. Zhang, "Multiple access integrated adaptive finite blocklength for ultra-low delay in 6G wireless networks," *IEEE Transactions on Wireless Communications*, vol. 23, no. 3, pp. 1670–1683, 2024.
- [89] J. Zheng, H. Jiang, L. Yu, S. Y. Khamaiseh, and M. Rahman, "On the performance of uplink NOMA short-packet communication system," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 12, pp. 18678–18687, 2024.

- [90] Y. Yang, Y. Hu, and M. C. Gursoy, "Energy efficiency of RIS-assisted NOMA-based MEC networks in the finite blocklength regime," *IEEE Transactions on Communications*, vol. 72, no. 4, pp. 2275–2291, 2024.
- [91] N. Kato *et al.*, "Ten challenges in advancing machine learning technologies toward 6G," *IEEE Wireless Communications*, vol. 27, no. 3, pp. 96–103, 2020.
- [92] M. Ghojagh *et al.*, "A modern tutorial on deep reinforcement learning," 2024.
- [93] R. Sun *et al.*, "A comprehensive survey of knowledge-driven deep learning for intelligent wireless network optimization in 6G," *IEEE Communications Surveys & Tutorials*, 2025.
- [94] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," 2017.
- [95] E. Eldeeb and H. Alves, "Offline reinforcement learning for wireless radio resource management: A survey," *IEEE Communications Surveys & Tutorials*, 2024.
- [96] A. Hady *et al.*, "Multi-agent deep reinforcement learning for resource allocation optimization: A survey," *IEEE Communications Surveys & Tutorials*, 2025.
- [97] M. Han, Z. Zhang, Z. Wei, S. Dang, and D. I. Kim, "Multi-agent reinforcement learning based uplink OFDMA for IEEE 802.11ax WLANs," *IEEE Transactions on Wireless Communications*, 2024.
- [98] C. Zhu *et al.*, "A survey of multi-agent deep reinforcement learning with communication," *Autonomous Agents and Multi-Agent Systems*, vol. 38, no. 4, 2024.
- [99] M. Polese, L. Bonati, *et al.*, "Understanding O-RAN: Architecture, interfaces, algorithms, security, and research challenges," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 2, pp. 1376–1411, 2023.
- [100] K. Van Moffaert and A. Nowé, "Multi-objective reinforcement learning using sets of Pareto dominating policies," *Journal of Machine Learning Research*, vol. 15, pp. 3483–3512, 2014.
- [101] Y. Xu *et al.*, "Deep reinforcement learning for multi-objective resource allocation in multi-platoon cooperative vehicular networks," *IEEE Transactions on Wireless Communications*, vol. 22, no. 9, pp. 6185–6198, 2023.
- [102] A. Abels, D. Roijers, T. Lenaerts, A. Nowé, and D. Steckelmacher, "Dynamic weights in multi-objective deep reinforcement learning," in *Proc. International Conference on Machine Learning (ICML)*, pp. 11–20, 2019.
- [103] Y. Yang *et al.*, "Preference controllable reinforcement learning," in *Proc. International Conference on Artificial Intelligence and Statistics (AISTATS)*, Proceedings of Machine Learning Research, 2025.

- [104] Y. Liu, Y. Zhou, Z. He, Y. Yang, Q. Han, and J. Li, "Dynamic preference inference network: Improving sample efficiency for multi-objective reinforcement learning by preference estimation," *Knowledge-Based Systems*, vol. 304, p. 112512, 2024.
- [105] A. Ambadkar *et al.*, "Improving pareto-front coverage and diversity in preference-conditioned MORL," 2026.
- [106] R. Liu, Y. Pan, L. Xu, L. Song, J. Bian, P. You, and Y. Chen, "C-MORL: Multi-objective reinforcement learning through efficient discovery of Pareto front," in *Proc. International Conference on Learning Representations (ICLR)*, 2025.
- [107] Z. Yan and H. Tabassum, "Generalized multi-objective reinforcement learning with envelope updates in URLLC-enabled vehicular networks," *IEEE Transactions on Vehicular Technology*, 2024.
- [108] K. Terekhov and Ç. Gülcehre, "In search for architectures and loss functions in multi-objective reinforcement learning," 2024.
- [109] D. Heo, D. N. Rim, and H. Choi, "Dynamic preference multi-objective reinforcement learning for internet network management," in *Proc. International Conference on Neural Information Processing (ICONIP)*, 2025.
- [110] A. A. Al-Habob and H. Tabassum, "Non-orthogonal age-optimal information dissemination in vehicular networks: A meta multi-objective reinforcement learning approach," *IEEE Transactions on Vehicular Technology*, 2024.
- [111] K. Li *et al.*, "Wireless resource allocation algorithm based on multi-objective deep reinforcement learning for vehicle-to-vehicle communications," *Journal of Computer Research and Development*, vol. 61, no. 9, pp. 2229–2245, 2024.
- [112] W.-C. Chien, "Multi-objective optimization of 3D cell deployment in sustainable B5G/6G networks: Balancing performance and sustainability," *IEEE Transactions on Network and Service Management*, 2025.
- [113] X. Guan, J. Xue, H. Jiang, and G. Tian, "Multiobjective optimization of wireless powered communication networks assisted by intelligent reflecting surface based on multiagent reinforcement learning," *IEEE Transactions on Wireless Communications*, 2024.
- [114] A. Al-Tahmeesschi *et al.*, "Multi-objective deep reinforcement learning for 5G base station placement to support localisation for future sustainable traffic," in *Proc. EuCNC/6G Summit*, 2024.
- [115] M. T. Balaji *et al.*, "Adaptive multi-objective evolutionary design for sustainable and efficient wireless sensor networks," *IETE Journal of Research*, pp. 1–13, 2025.
- [116] J. Tan *et al.*, "Multi-objective optimization using adaptive distributed multi-agent reinforcement learning," *IEEE Transactions on Intelligent Transportation Systems*, 2024.

- [117] ETSI, “TS 103 982 V8.0.0: O-RAN; O-RAN architecture description,” tech. rep., European Telecommunications Standards Institute (ETSI), Jan. 2024.
- [118] ETSI, “TS 103 983 V4.0.0: O-RAN; A1 interface: General aspects and principles,” tech. rep., European Telecommunications Standards Institute (ETSI), Apr. 2025.
- [119] Y. Huang, Y. Jiang, F.-C. Zheng, P. Zhu, *et al.*, “Energy-efficient optimization in user-centric cell-free massive mimo systems for urllc with finite blocklength communications,” *IEEE Transactions on Vehicular Technology*, pp. 1–14, 2024.
- [120] C. F. Hayes, R. Rădulescu, E. Bargiacchi, J. Källström, M. Macfarlane, M. Reymond, T. Verstraeten, L. M. Zintgraf, R. Dazeley, F. Heintz, *et al.*, “A practical guide to multi-objective reinforcement learning and planning,” *Autonomous Agents and Multi-Agent Systems*, vol. 36, no. 1, p. 26, 2022.
- [121] F. Felten, E.-G. Talbi, and G. Danoy, “Multi-objective reinforcement learning based on decomposition: A taxonomy and framework,” *Journal of Artificial Intelligence Research*, vol. 79, pp. 679–723, 2024.
- [122] B. Toure, D. Tsilimantos, and T. Giannakas, “Multi-objective scheduling in wireless networks with deep reinforcement learning,” in *IEEE Wireless Communications and Networking Conference (WCNC)*, IEEE, 2025.
- [123] H. Lu, D. Herman, and Y. Yu, “Multi-objective reinforcement learning: Convexity, stationarity and Pareto optimality,” in *The Eleventh International Conference on Learning Representations*, ICLR, 2023.
- [124] T. Basaklar, S. Gumussoy, and U. Y. Ogras, “PD-MORL: Preference-driven multi-objective reinforcement learning algorithm,” in *International Conference on Learning Representations*, ICLR, 2023.
- [125] L. N. Alegre, A. Bazzan, and B. C. da Silva, “Optimistic linear support and successor features as a basis for optimal policy transfer,” in *Proceedings of the 39th International Conference on Machine Learning*, vol. 162 of PMLR, pp. 394–413, 2022.
- [126] W. Röpke, M. Reymond, P. Mannion, D. M. Roijers, A. Nowé, and R. Rădulescu, “Divide and conquer: Provably unveiling the Pareto front with multi-objective reinforcement learning,” *arXiv preprint arXiv:2402.07182*, 2024.
- [127] A. Abdolmaleki, S. Huang, L. Hasenclever, M. Neunert, F. Song, M. Zambelli, M. Martins, N. Heess, R. Hadsell, and M. Riedmiller, “A distributional view on multi-objective policy optimization,” in *Proceedings of the 37th International Conference on Machine Learning*, vol. 119 of PMLR, pp. 11–22, 2020.
- [128] D. Kim, M. . Hong, J. Park, and S. Oh, “Conflict-averse gradient aggregation for constrained multi-objective reinforcement learning,” in *International Conference on Learning Representations*, ICLR, 2025.

Acknowledgments

First and foremost, I am profoundly grateful to Allah Almighty, whose mercy, strength, and guidance carried me through every stage of this journey and enabled me to complete this thesis. I also send heartfelt blessings upon Prophet Muhammad (peace be upon him), whose example continues to illuminate my path and inspire me to strive to be among his devoted followers.

I would like to express my sincere appreciation to my supervisors, **Prof. Muhammad Mahtab Alam** and **Prof. Yannick Le Moullec**, for their unwavering support throughout the course of my research. Their continuous guidance, insightful feedback, and inspirational mentorship have shaped both my academic journey and personal growth. I am truly thankful for everything you have done for me and for always believing in my potential.

My heartfelt thanks also go to **Osama Mohammad Elgarhy** for his constant encouragement and enthusiasm. During moments when negative results felt discouraging, his ability to find new perspectives and identify promising directions was invaluable. I deeply appreciate the many hours of discussions and the collaborative spirit we shared over the past three years.

I would also like to express my sincere thanks to **Alar Kuusik** for his support as Head of Department, and to **Anum, Junaid, Fredrick, Marika, and Priit** for their continuous support throughout this journey. Your help in resolving both technical and non-technical challenges made a meaningful difference, and I am truly grateful.

My appreciation also goes to **Eva, Annet, and Jane** for motivating me to learn and speak Estonian, and especially to **Eva** for organizing wonderful events and efficiently managing all related matters.

I am profoundly grateful to my parents for encouraging me to pursue my dreams and for standing by me through every challenge. My deepest thanks to **Noor** for her love and support, and to my little angels **Aroosh, Aarsal, and Muhammad**, whose smiles always lifted my spirits and filled my life with joy. I am equally thankful to my brothers and sisters, whose support has been a constant source of strength. Whatever success I achieve in life, I owe it to them.

Finally, I extend my warmest gratitude to all my friends for the unforgettable memories and shared experiences over the recent years. Your companionship added meaning and balance to this journey, and those moments will always remain close to my heart.

Abstract

Multi-Objective based Radio Resource Optimization for Massive Connectivity in 6G Networks

Sixth-generation (6G) wireless networks are expected to support device densities approaching 10^6 devices/km² while simultaneously satisfying stringent energy-efficiency (EE) requirements. Achieving this target is complicated by three tightly coupled challenges. The first is *coordination*: in Open Radio Access Network (O-RAN) architectures, scheduling decisions are made independently by distributed units (O-DUs), giving rise to decision-dependent inter-cell interference (ICI) that grows with device density and cannot be mitigated by physical-layer techniques alone. The second is *physical-layer realism*: massive machine-type communication (mMTC) services operate in the finite-blocklength (FBL) regime, where Shannon capacity models systematically overestimate achievable rates and the EE-connectivity trade-off is qualitatively different from the infinite-blocklength case. The third is *adaptability*: practical 6G networks require scheduling policies that respond to time-varying channel conditions, fluctuating interference, and shifting operator priorities within the 10–100 ms control-loop budget of the Near-Real-Time RAN Intelligent Controller (Near-RT RIC), making repeated re-optimization computationally impractical.

This thesis develops a three-layer framework, validated through six peer-reviewed publications, that addresses each challenge in turn.

The *first layer* (Contributions I and II) introduces an E2-interface-compliant inter-O-DU coordination mechanism in which adjacent O-DUs exchange scheduling information selectively to suppress decision-dependent ICI without joint transmission. The coordination mechanism improves uplink throughput by up to 30%, reduces energy consumption by up to 10%, and increases simultaneously connected devices by a factor of 1.5 relative to uncoordinated operation. A companion RISC-V instruction-level energy evaluation framework is developed to quantify the computational energy of scheduling decisions on virtualized O-DU hardware, establishing that the total energy cost of scheduling must include both radio and computational components.

The *second layer* (Contributions III and IV) formulates and solves the joint EE-connectivity optimization problem in multi-cell NOMA under FBL reliability constraints. Two algorithms are proposed: EERS-MTCN, a non-cooperative baseline exploiting channel-aware sectorization, and SRS-MTCN, a cooperative algorithm incorporating inter-cell scheduling exchange. EERS-MTCN achieves 35% EE improvement over OMA baselines; SRS-MTCN achieves 35–58% EE gains at device densities exceeding 40 active devices per cell while maintaining connectivity targets. The FBL analysis shows that the EE-optimal blocklength lies in the range $L = 300$ –400 channel uses, and that Shannon-based optimizations overestimate achievable EE and produce sub-optimal scheduling decisions in the short-packet regime.

The *third layer* (Contributions V and VI) introduces Wireless Multi-objective Conditional Envelope PPO (WiMO-CEPPO), a preference-conditioned multi-objective reinforcement learning (MORL) framework that learns a single adaptive scheduling policy capable of balancing EE, connectivity, reliability, and cross-link interference without retraining as

network conditions evolve. The framework integrates a Context-to-Preference Mapper (CPM) that infers objective priorities from observed network state, a conditional envelope mechanism that attenuates policy updates under strong inter-objective gradient conflict, and constraint-aware action generation that structurally enforces resource-block and power feasibility. WiMO-CEPPO achieves a hypervolume of 0.752 and an inverted generational distance of 0.078 in the four-objective setting, improving Pareto coverage by up to 38.7% and constraint-satisfaction rate to 91.2% relative to the strongest MORL baseline. The extended five-objective formulation further improves hypervolume to 0.781 and confirms compatibility with Near-RT RIC deployment constraints.

Together, the three layers demonstrate that sustainable massive connectivity in 6G requires the coordinated treatment of inter-cell interference, short-packet physical-layer realism, and multi-objective adaptability. Each layer addresses a distinct limitation of existing approaches, and their combination provides a methodologically coherent and practically grounded solution to the challenge of energy-efficient massive connectivity in future wireless networks.

Kokkuvõte

Raadioressursside mitme eesmärgiga optimeerimine massilise ühenduvuse saavutamiseks 6G-võrkudes

Väitekirja pealkiri: Raadioressursside mitme eesmärgiga optimeerimine massilise ühenduvuse saavutamiseks 6G-võrkudes

Kuuenda põlvkonna (6G) traadita võrgud peaksid toetama enneolematut massilise ühenduvuse taset, saavutades samal ajal ranged energiatõhususe nõuded. Eelkõige peavad tulevikuvõrgud toetama seadmete arvu, mis läheneb 10^6 seadmele/km², kuid ühenduvuse maksimeerimise ja energiatarbimise minimeerimise eesmärgid on põhimõtteliselt vastuolulised. Teenindatavate seadmete arvu suurendamine nõuab agressiivsemat raadioressursside taaskasutamist, mis suurendab häireid ja energiatarbimist, samas kui edastusvõimsuse vähendamine energiatõhususe parandamiseks piirab leviala ja halvendab ühenduvust. See kompromiss moodustab käesoleva väitekirja keskse uurimisprobleemi st, kuidas optimeerida energiatõhusust ja massilist ühenduvust 6G raadioressursside haldamisel realistlikes tingimustes.

Probleemi teevad veelgi keerulisemaks kolm peamist väljakutset. Esiteks, avatud mobiilsidevõrgu (O-RAN) arhitektuurides tehakse edastuse ajastamisotsuseid karge juhiseadmed vahel sõltumatult, mille tulemuseks on otsustest sõltuv kargevaheline interferents, mis kasvab koos seadmete arvuga. Teiseks, massilised masintüüpi sideteenused (massive machine type communication, mMTC) töötavad lühikeste pakettidega lõpliku plokikipikkuse režiimis, kus tavapärased Shannoni läbilaskevõime mudelid hindavad saavutatavat jõudlust üle ja dekodeerimisvigu ei saa tähelepanuta jätta. Kolmandaks, praktilised 6G keskkonnad on väga dünaamilised; kanali tingimused, kasutajate arv ja operaatorite prioriteedid muutuvad pidevalt, mistõttu staatilised optimeerimismeetodid ei sobi O-RAN-i reaalse lähedase raadiopääsuvõrgu intelligentse kontrolleri (Near-Real-Time Radio Access Network Intelligent Controller, Near-RT RIC) 10-100 ms edastuse ajastamisaja jaoks. Nende probleemide lahendamiseks arendab see väitekirja välja kolmekihilise raamistiku, mis on avaldatud kuues eelretsenseeritud publikatsioonis.

Esimene kiht tutvustab interferentsitundlikku hajutatud juhtseadmete (O-distributed unit, O-DU) vahelist koordineerimismehhanismi, milles naaberkärje juhtseadmed vahetavad edastuse ajastamisinfot, et vähendada kargevahelist interferentsi. See kiht arendab ka kasutasemel energiamudelit, et kvantifitseerida ajastamisotsuste arvutuslikku energiakulu.

Teine kiht formuleerib energiatõhususe ja ühenduvuse ühise optimeerimise mitmekärjelistes süsteemides nii koostööl põhinevate kui ka mittekoostööl põhinevate interferentsitingimuste korral, pöörates erilist tähelepanu lõpliku plokikipikkuse usaldusväärsuse piirangutele. Selle probleemi lahendamiseks pakutakse välja kaks kohandatud mitme eesmärgiga ajastamisalgoritmi.

Kolmas kiht liigub staatilisest optimeerimisest kaugemale, tutvustades eelistustest sõltuvat mitme eesmärgiga võimendatud õppe (reinforcement learning) raamistikku. Kavandatud agent õpib tundma ühte adaptiivset ajastamispoliitikat, mis on võimeline dünaamiliselt tasakaalustama konkureerivaid eesmärgi ilma ümberõpetamata, kui võrgutingimused

muutuvad.

Kavandatud raamistik saavutab märkimisväärse jõudluse kasvu kõigis uuritud stsenaariumides. Häireteadlik edastuse koordineerimine parandab üleslingi läbilaskevõimet kuni 43%, vähendab energiatarbimist 40% ja suurendab samaaegselt ühendatud seadmete arvu 1,5 korda võrreldes koordineerimata lähenemisviisidega. Analüütilised optimeerimismetodid parandavad energiatõhusust kuni 160% kasutajate arvul 400 aktiivset seadet lahtri kohta ja pakuvad 35-58% kasvu fikseeritud plokipikkuse korral. Lõpuks saavutab kavandatud mitme eesmärgiga võimendatud õppe raamistik Traadita mitme eesmärgiga tingimusliku ümbrikuga PPO (WiMO-CEPPO) hüpermahu 0,752 ja invertteeritud põlvkondade vahelise kauguse 0,078, edestades baasmeetodeid Pareto katvuses kuni 49,5% ja lõpliku keskmise tasu osas 60,5%.

Üldiselt näitab väitekirj, et interferentsi koordineerimine, analüütiline mitme eesmärgiga optimeerimine ja adaptiivne tugevdusõpe on teineteist täiendavad lähenemisviisid 6G raadioressursside jätkusuutlikuks haldamiseks. Koos pakuvad need sidusat ja praktiliselt põhjendatud lahendust energiatõhusa massilise ühenduvuse saavutamise väljakutsele tuleviku traadita võrkudes.

Appendix 1

Publication (I):

Latif, Aamir; Alam, Muhammad Mahtab; Le Moullec, Yannick, "Inter O-DUs Coordination for Scheduling of Massive Users in O-RAN," in *2023 Joint European Conference on Networks and Communications & 6G Summit (Eu-CNC/6G Summit 2023)*, pp. 1–6, 2023

Inter O-DUs Coordination for Scheduling of Massive Users in O-RAN

Aamir Latif, Muhammad Mahtab Alam, Yannick Le Moullec
Tallinn University of Technology, Tallinn
Email: (aamir.latif; muhammad.alam; yannick.lemoullec)@taltech.ee

Abstract—Open Radio Access Network (O-RAN) is a decentralized, intelligent, and open network architecture. RAN Network Functions (RNF) are split into Open Central Unit (O-CU), Open Distributed Unit (O-DU), and Open Radio Unit (O-RU), which are deployed on commercial hardware and cloud nodes as containerized Network Functions (CNFs). This paper presents a coordination architecture between O-RAN Distributed Units (O-DU) for scheduling massive number of users. We propose a Non-Orthogonal Multiple Access (NOMA) based access technique for interference avoidance and scheduling of such massive number of users through a cooperative strategy. This involves sharing individual scheduling tables to compute interference for subsequent User Equipment (UEs) transmissions in the network. The results of the interference calculations are then fed into O-DU schedulers. The scheduler of each O-DU shares the radio resources among the UEs to have minimum interference. The proposed scheme's effectiveness in terms of throughput, energy efficiency, and connectivity is examined through simulations. The results show that the proposed method reduces the effects of interference, making it better suited to the needs of dense deployments. Users operating under the proposed approach experience up to +30% and -10% improvements in achieved user data rates and energy consumption, respectively, when compared to the absence of a coordination architecture.

Keywords—6G, Beyond-5G, O-RAN, Orthogonal Multiple Access (OMA), Non-Orthogonal Multiple Access (NOMA), O-DU Coordination, Radio Resource Management(RRM)

I. INTRODUCTION

This paper proposes inter-O-DU coordination based on a cooperative scheduling algorithm that can help maximize throughput and energy efficiency in O-RAN. The interference minimization problem is formulated for a multi-cell scenario considering NOMA for massive users in O-RAN [4]. O-RAN, an evolution of Next-Generation (NG)-RAN, features disaggregation, intelligence, openness, and virtualization to handle future network growth and capacity needs [5]. The O-RAN alliance disaggregates the NG-RAN into logical nodes and RAN intelligent controllers (RICs) as articulated in [1] and illustrated in Fig (1).

The RIC has two components: the non-Real-Time RIC (non-RT-RIC) having a control loop on the timescale of greater than 10 s and the near-Real-Time RIC (near RT-RIC) that operates on tens or hundreds of ms [17]. The two intelligent controllers support external apps known as rApps and xApps, which provide controls of RAN components and associated resources [9], [16]. The intelligent controllers have an interface with logical network nodes through an open interface.

The O-CU handles higher-level functions such as control and data plane interfaces to the core network. The O-DU pro-

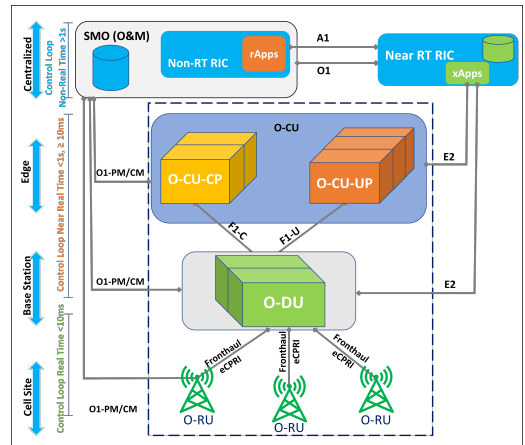


Fig. 1: Simplified O-RAN Architecture with control-loop latency with O1, A1, E1, E2, R1 interfaces between Non-RT RIC and Near RT RIC and network function elements O-CU, O-DU, O-RU. Based on [9].

vides High-level Physical (High-PHY), MAC, and Radio Link Control (RLC) functionalities, and the O-RU processes Low-level Physical (Low-PHY) operations [2]. The performance of a 5G network is tied to its transport infrastructure; eCPRI, which is necessary for an O-RAN architecture (between the O-DU and O-RU) [12]

5G is not a simple expansion toward a high-frequency band; instead, it is a confluence of new technological developments fueled by new services i.e. eMBB (enhanced Mobile Broadband), URLLC (Ultra Reliable Low Latency Communications), and mMTC (massive Machine Type Communications) [3], [13], [14]. Interference is a limiting factor that affects the performance of a wireless system. One of the ways to minimize interference is to employ adequate transmission powers, which guarantees excellent quality of service to users without causing interference to others [10]. Other methods include assigning orthogonal resources in a time-frequency grid to achieve high spectrum and throughput efficiency. The latest 5G deployments use Orthogonal Multiple Access (OMA) systems, which provide orthogonality in terms of frequency resources [21]. In this context, proactive scheduling and sophisticated multiple access mechanisms are required as proposed in our earlier study [7].

The main two differences between this paper and [7] are 1) an interference minimization problem for stand-alone O-DU for O-RAN is formulated and 2) the problem is further extended to a coordinated O-DU case. The power and resource scheduling for both stand-alone and coordinated O-DU is carried out for NOMA and compared with OMA. Some researchers believe that the NOMA strategy is the most promising method for boosting capacity to over 100,000 devices per cell [6]. In [8], the authors introduced cooperative NOMA, where strong devices can relay for weak devices with unfavorable channel circumstances, boosting diversity gain. Yet, increasing successive interference cancellation (SIC) processing raised computational and power needs; many SIC in-network with limited energy, computing, and time resources may increase energy costs for the users and there has been a significant amount of research on cooperative relaying in NOMA systems [15], [20].

In contrast to existing work [18] which only discusses that the Inter-O-DU Coordination Interface is required to enable the full potential of Cell-Free (CF) massive MIMO (Multiple Input Multiple Output); the inter-O-DU coordination itself has not yet been studied in O-RAN architecture which will help in the minimization of interference and maximization of throughput for different verticals.

Contributions: The contributions of this paper are the following:

- First, the inter-cell interference experienced from adjacent cell users is modeled via an interference minimization problem for the uplink use case only.
- Secondly, we extrapolate our findings to examine ICI's effect in a more generalized, i.e. multi-cell setting, and an efficient radio resource and power allocation for the aforementioned scenarios.
- Thirdly, we propose a cooperative scheduling algorithm to minimize massive interference of UEs by sharing the scheduling tables between the O-DUs to increase the overall network throughput and energy efficiency.
- Lastly, as a proof of concept, simulation results show up to 30% higher throughput and 10% lower energy consumption as compared to the non-coordinated case i.e., stand-alone O-DU case.

The rest of the paper is organized as follows: Section II presents the problem formulation for the considered scenario along with the proposed architecture and Section III discusses the algorithms for NOMA scheduling. Finally, Section IV and Section V provide the performance evaluation and conclusion of the work, respectively.

II. PROPOSED ARCHITECTURE OF INTER-O-DU COORDINATION

The main goal of this paper is to investigate how coordination between O-DUs in O-RAN could benefit massive connectivity. We propose an inter-O-DU coordinated architecture, as shown in Fig. 2 (discussed in this section), and a NOMA-based cooperative scheduler for managing massive numbers of users in an O-RAN, as illustrated in Fig. 3 (discussed in Section III)

with the aim that inter-O-DU coordination will significantly help in the mitigation of interference and efficient utilization of radio resource management. To the best of our knowledge, this is the first work that entails a cooperative interference avoidance strategy to optimize the radio resources in O-RAN. In [4] a NOMA-based radio resource management technique for small cell base stations for traffic steering use cases in O-RAN has been studied.

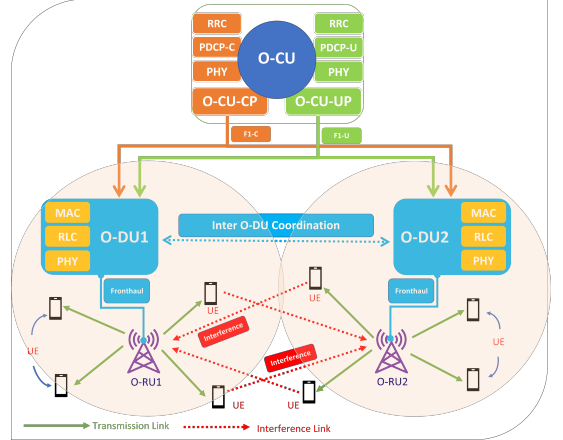


Fig. 2: Proposed Inter O-DU Coordination Architecture

A. Mathematical Formulation

We consider the NOMA case for an available set of resources and their indexes are $\mathcal{R} = \{1, 2, \dots, R\}$. We consider M to be a positive, maximum number of devices that can be supported per subcarrier. There are $\mathcal{I} = \{1, 2, \dots, I\}$ set of transmitting users UE_s belonging to cell s where the number of cells is $\mathcal{S} = \{1, 2, \dots, S\}$ and user i in the cell s will be denoted by i_s having resource i_s^r .

The achievable rate of user i belonging to cell s on a given resource unit r is denoted by $R_{i_s}^r = B_0 \log_2(1 + SINR_{i_s}^r)$ where $SINR_{i_s}^r$ is the signal to interference plus noise ratio experienced by NOMA user i belonging to cell s on a given resource unit r which can be written as:

$$SINR_{i_s}^r = a_{i_s}^r \left(\frac{G_{i_s,s}^r P_{i_s}^r}{\sum_{t \neq s, t \in \mathcal{S}} \sum_{j \in \mathcal{I}_t} G_{j_t,s}^r P_{j_t}^r a_{j_t}^r + Int_s^r + P_n} \right) \quad (1)$$

where $G_{i_s,s}^r$ is user i_s channel gain in cell s to its own base station on resource unit r . $P_{i_s}^r$ denotes the transmission power of user i_s to resource unit r , whereas in the denominator the interference term comes from another cell t , with a group of users $\mathcal{I}_t$ within the cell. We use j as the interference user index, thus a user j in cell t will be denoted as j_t . Moreover, $G_{j_t,s}^r$ denotes the channel gain of an interfering user, belonging to cell t on resource unit r to the base station of cell s , and $P_{j_t}^r$ denotes the interfering user j_t 's transmission power on resource unit r . The binary allocation variables are used for

scheduling: variable $a^r_{i_s}$ denotes the resource unit occupancy coefficient such that $a^r_{i_s}=1$ if the resource unit r is occupied by user i_s , and $a^r_{i_s}=0$ otherwise. P_n is the noise power and $Int^r_{i_s}$ is assumed as constant interference experienced by user i from the interfering users j from adjacent cells.

It is desirable to optimize the problem so that user i experiences as little interference from other users in neighboring cells as possible. The problem can be treated as the interference experienced by user i on a given resource unit r ; alternatively, in order to avoid using the two binary variables, the problem can be modeled as minimizing the interference on the resource units in cell s , where $Int^r_{i_s}$ on resource unit r in cell s can be written as :

$$Int^r_{j_t,s} = \sum_{i \neq j, i \in M} G^r_{j_t,s} P^r_{j_t} a^r_{j_t} \quad (2)$$

We want to reduce interference at user i caused by NOMA users in the same cell sharing the same resource unit. Therefore, we can write down the objective function as:

$$\min \sum_{s \in S} \sum_{r \in R} \left(\sum_{i \neq j, i \in M} G^r_{j_t,s} P^r_{j_t} a^r_{j_t} \right) \quad (3)$$

subject to:

$$a^r_{i_s} \left(\frac{G^r_{i_s,s} P^r_{i_s}}{G^r_{j_t,s} P^r_{j_t} a^r_{j_t} + Int^r_{i_s} + P_n} \right) \geq \gamma_{i_s,lim} \quad (4)$$

$$0 \leq P^r_{i_s} a^r_{i_s} \leq P_{max}, \forall s \in S, \forall i_s \in I, \forall r \in R \quad (5)$$

which is a limitation while trying to provide the minimum acceptable level of service (QoS), where $\gamma_{i_s,min}$ is the SINR threshold for acceptable results that user i_s can have to satisfy the QoS and P_{max} denotes the maximum allowable transmit power per U_E .

$$\sum_{i_s \in I} a^r_{i_s} \leq 1, \forall r \in R, s \in S \quad (6)$$

$$\sum_{i_s \in I} a^r_{i_s} \leq I, \forall i \in I \forall r \in R, s \in S. \quad (7)$$

It is evident that the objective function poses a complex combinatorial optimization problem. The following solution is proposed to this problem.

B. Solution

Initially, we assign an initial interference power to each user. Then in the second step, we manage the scheduling for every UE. Afterward, we set in power allocation to cut down the interference even further at the target receiver.

Hence power allocation follows channel assignment. We rephrase the optimization problem as we are working per resource unit since there is no interference from adjacent resource units, but we must include NOMA interference users in a specific resource unit. For NOMA, we have more than

one user per resource unit; hence, we define U_t as the group of NOMA users per cell per resource unit; i_s is a user in cell s that is a member of U_s , and we are not using any resource index. The optimization goals become:

$$\min \sum_{s \in S} \sum_{i_s \in U_t} \left(\sum_{t \neq s, t \in S} \sum_{j \in U_t} G^j_{t,s} P^j_t + \sum_{j \neq i_s, j \in U_t} G^r_{s,s} P^j_s \right) \quad (8)$$

where $G^j_{t,s}$ represents the channel gain of user j from cell t and the group of NOMA users U_s on cell s . The j user's power is P^j_t . The two terms refer to the inter-cell interference that occurs between different NOMA users of other cells. Regarding the NOMA aspect, the channel gain of NOMA user j that belongs to the same group U_t is denoted by $G^r_{s,s}$ and subject to

$$SINR^r_{i_s,NOMA} \geq \gamma_{s,lim} \quad (9)$$

$$0 \leq P^r_{i_s} \leq P_{max}, \forall s \in S. \quad (10)$$

$$SINR^r_{i_s,NOMA} = \left(\frac{G^r_{i_s,s} P^r_{i_s}}{\sum_{t \neq s} \sum_{j \in U_t} G^j_{t,s} P^j_t + P_n + \sum_{j \neq i_s, j \in U_t} G^r_{s,s} P^j_s} \right) \quad (11)$$

The $SINR^r_{i_s,NOMA}$ does not account for the effect of SIC on reducing interference caused by other NOMA users within the same cell and resource unit. To account for the SIC's impact, we just set the inequalities so that the NOMA interference users in the same cell as the primary user will have a value of zero after going through the SIC. The constraint in (9) is not linear; so we start by substituting the value of $SINR^r_{i_s,NOMA}$ from Equation (11) into (9), yielding (12) to (15).

$$\left(\frac{G^r_{i_s,s} P^r_{i_s}}{\sum_{t \neq s} \sum_{j \in U_t} G^j_{t,s} P^j_t + P_n + \sum_{j \neq i_s, j \in U_t} G^r_{s,s} P^j_s} \right) \geq \gamma_{i_s,lim} \quad (12)$$

$$G^r_{i_s,s} P^r_{i_s} \geq \gamma_{i_s,lim} * \left(\sum_{t \neq s, t \in S} \sum_{j \in U_t} G^j_{t,s} P^j_t \right) + \left(\sum_{j \neq i_s, j \in U_t} G^r_{s,s} P^j_s \right) + P_n \quad (13)$$

which can be further simplified as:

$$G_{i,s}^r P_{i,s}^r - \gamma_{i,s,lim} \left(\sum_{t \neq s, t \in S} \sum_{j \in U_t} G_{t,s}^j P_t^j \right) - \gamma_{i,s,lim} * \left(\sum_{j \neq i, j \in U_t} G_{s,s}^r P_s^j + P_n \right) \geq \gamma_{i,s,lim} * P_n. \quad (14)$$

equivalently we can write as :

$$-G_{i,s}^r P_{i,s}^r + \gamma_{i,s,lim} \left(\sum_{t \neq s, t \in S} \sum_{j \in U_t} G_{t,s}^j P_t^j \right) + \gamma_{i,s,lim} * \left(\sum_{j \neq i, j \in U_t} G_{s,s}^r P_s^j + P_n \right) \leq \gamma_{i,s,lim} * P_n. \quad (15)$$

Given that this is a binary linear programming problem, it is difficult to solve through a trivial MATLAB program; so, in order to solve (15) we use our proposed implementation of NOMA as presented in Algorithm 1 in Section III.

III. PROPOSED COOPERATIVE SCHEDULER FOR INTER-O-DU COORDINATION

We investigated O-DU coordination by sharing scheduling tables with UE channel attributes to be scheduled to better minimize the impact of interference, as illustrated in Fig. 3. The suggested scheduling strategy, in contrast to prior research, reduces the negative effects of inter-cell interference by sharing radio resources with neighboring O-DUs. In the NOMA technique, each O-DU divides UEs into three classes depending on their channel parameters: good, middling, and bad. We assume two main types of interference: NOMA interference from users simultaneously assigned to the same resource unit and time slot, and ICI from users broadcasting at the same resource unit but from neighboring cells.

The scheduling tables from each O-DU are shared with the cooperative scheduler. The scheduler selects one UE from each user group to be superimposed at a given resource unit. Four UEs can occupy a resource unit at any given time. The scheduler calculates the optimum combination of UEs for all available resource units before allocating slots to O-DU. The scheduler allocates power to reduce co-channel and ICI interference. Good channel users **Pcon=14 dBm**, modest channel users **Pcon=20 dBm**, and bad channel users **Pcon=23 dBm**. Different power coefficients are assigned to users to successively perform SIC at the receiving end. We presume good channel users are close to their serving O-DU and can be given lower power limits and the same is true for bad channel users. Fig. 3 illustrates the proposed cooperative strategy for NOMA.

In general, our suggested cooperative approach is different from the coordinated or joint processing that occurs in coordinated multipoint (CoMP) [11]. In CoMP, user equipment (UE) at the cell edge is served by two or more base stations in order to improve signal quality and boost throughput [19]. Nevertheless, in our approach, each O-DU serves only its

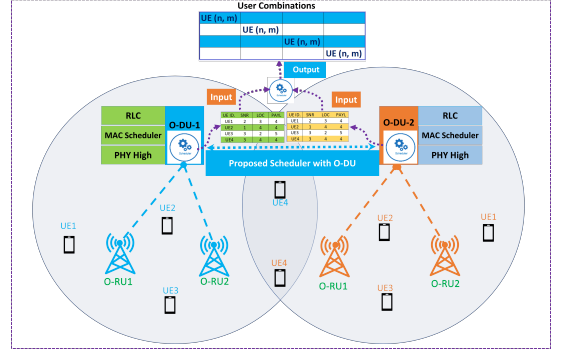


Fig. 3: Radio resource management: The two cooperating O-DUs (O-DU1 and O-DU2) pool their future transmission scheduling information. As a result, UEs with the lowest interference impact are given priority when it comes to resource allocation. Then, depending on the approach selected, the O-DU will employ either the OMA or NOMA protocol.

own users. The simulation has been carried out on a NOMA-based resource allocation algorithm and compared with the OMA one and the simulation parameters are listed in Table. I. The NOMA algorithm is presented in Algorithm 1. There are three key phases to the algorithm: the first is to generate UE parameters using the channel model and set the transmit power to zero, the second step is to allocate resource units, and the third optimizes the power allocation.

We begin by giving each user a certain amount of interference to contend with. Second, we perform scheduling for all users. Then power allocation will be carried out to further reduce the interference power. Initially, the interference power will be allocated in the same way as OMA.

IV. PERFORMANCE EVALUATION AND DISCUSSIONS

The performance metrics, i.e., uplink throughput, uplink energy consumption, and the number of connected devices, are analyzed, discussed, and compared for both the OMA and NOMA schemes for inter-O-DU coordination vs. stand alone (i.e., absence of coordination) cases. For the time being, these simulations have been performed for the sub-6GHz massive IoT use case. This will be further extended to cellular cases in future work.

To find the channel parameters, we run the simulation 1000 times with UEs spread randomly across each O-DU. At a given time slot, we select transmitting UEs from each O-DU and interfering UEs from a nearby O-DU. For the OMA system, interfering UEs are those with the same time slot from neighboring cells. However, in NOMA, the interfering set includes UEs broadcasting at the same time slot from nearby cells, as well as UEs transmitting at the same time slot but from the same cell.

Simulation parameters can affect findings. When the number of users per O-RU increases, the number of users who interfere with each other increases, which might reduce performance. In this instance, it is best to specify QoS criteria

Algorithm 1 :Proposed NOMA Scheme

procedure User equipment creation ▷

 $i_s \leftarrow G_{i_s,s}^r$
while $P_{i_s}^r \neq 0, j \neq i$ **do**
Equation (11)
return $SINR_{i_s,NOMA}^r$
procedure Share the scheduling tables

while $Int_{i_s}^r = \frac{G_{i_s,s}^r P_{i_s}^r}{\gamma_{i_s,lim}} - P_n$ **do**
compute the best combination of UEs
Divide the UEs in three groups
Superpose One UE from each group in a given sub-carrier
while $P_{i_s}^r \neq 0$ **do**
compute the best combination of UEs
 $0 < P_{i_s}^r \leq P_{max}, \forall c \in C$
return i_s
procedure Evaluate the resource

while $\frac{Int_{i_s}^r \gamma_{i_s,lim}}{G_{i_s,s}^r} \leftarrow p$ **do**
Calculate Rate R_i
Calculate Energy Consumption
return $R_{i_s}, EnergyConsumption$

TABLE I: Simulation Parameters

Sr. No	Simulation Parameters	
	Name	Value
A.	Transmit Power of Base Station, and UE (dBm)	35, 14, 20, 23
B.	Carrier Frequency (GHz)	3
C.	Receiver Thermal Noise density (dBm/Hz)	-174
D.	No. of UE per cell	1,10
E.	Modulation Scheme	BPSK
F.	No. of Cooperating O-RU per O-DU	2
G.	Interference Margin (dB)	0
H.	Channel Model	Extended Okumura Hata model
I.	Effective Noise Power (dBm)	$C + D + J + 10 \log(L)$
J.	Noise figure of Base Station, UE (dB)	9, 5
K.	Receiver Sensitivity	$H + I$
L.	Occupied System Bandwidth (kHz)	180
M.	Height of Base Station, UE (m)	30, 1
N.	Radius of Cell (km)	1
O.	Maximum Coupling Loss (MCL) (dB)	$A - P$
P.	Required/Calculated SINR (dB)	

for both the O-DU and its users. As the cell radius increases, so does the user's path loss; in this regard, it is recommended to utilize the telecom operator's distributed unit settings as the benchmark.

Fig. 4 depicts an overview of uplink throughput; as can be seen, our proposed coordination scheme performs better than in the absence of inter O-DU coordination. Our proposed system considers UEs in both good and bad channel situations by allocating various power coefficients to avoid interference. As a result, our suggested scheme allows UEs, especially those

using the OMA technique, to achieve comparatively better throughput. It should be emphasized that the OMA technique has fewer UEs per resource unit, resulting in less interference and higher achievable throughput.

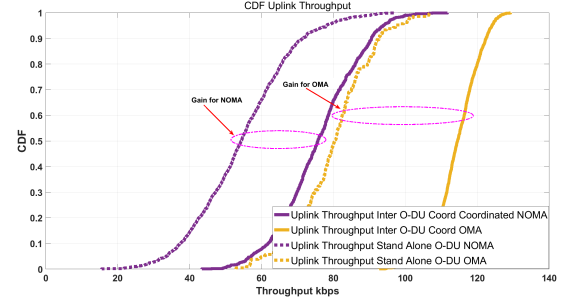


Fig. 4: Cumulative distribution function of Stand Alone and Inter O-DU Coordinated Uplink Throughput

More than 50% of UEs achieve above 115 kbps with our proposed coordinated architecture (and a maximum throughput gain of +43%), and 50% of NOMA UEs achieve around 75 kbps more with our proposed coordinated architecture (and a maximum throughput gain of +30%). These improvements are obtained thanks to the regulated impact of ICI and by restricting the number of UEs on a given subcarrier.

Fig. 5 entails the energy consumption for OMA and NOMA with the proposed architecture compared to the stand-alone case. Comparing the OMA and NOMA schemes, it is clear that both OMA schemes use less power than the NOMA schemes. Next, focusing on OMA schemes, with the proposed coordinated OMA, 50% of UEs consume around 40% less energy than with stand-alone OMA. Similarly, for NOMA schemes, the coordination architecture consumes approximately 10% less energy than the stand-alone mode. The optimal power allocation that eliminates the extra transmission power while providing the desired QoS at the transmitting users enables energy consumption improvements. In addition, higher SINR is achieved with fewer transmission repetitions at the transmitting UEs due to the diminished interference impact.

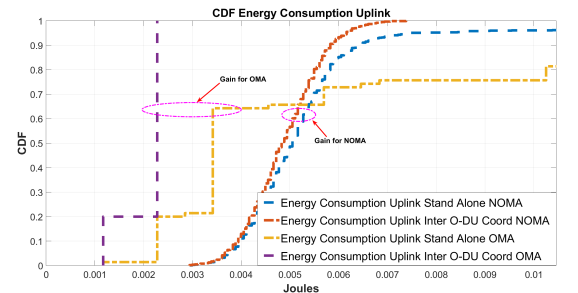


Fig. 5: Cumulative distribution function of stand alone and inter O-DU Coordinated Uplink Energy Consumption

Finally, Fig. 6 depicts OMA and NOMA device connections at the cell center and edge. The coordination architecture

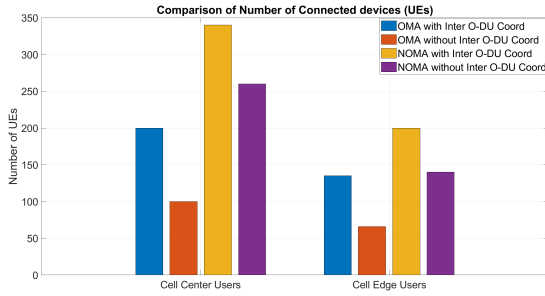


Fig. 6: Comparison of Connected Devices

and NOMA scheme outperform the OMA approaches by approximately 1.5 times for both cell-center and cell-edge users in terms of number of connected devices. These improvements are possible because cooperative scheduling between neighboring O-DUs reduces the impact of ICI. Superimposing numerous UEs in the same tone allows for improved connection in NOMA schemes, which is not possible in OMA schemes; for instance, only a single UE is allowed to use any particular uplink scheduling frame's allotted tones.

V. CONCLUSION

We proposed a coordination architecture between O-RAN Distributed Units (O-DU) for scheduling a massive number of users. We developed a NOMA-based approach to enhance the users' and cells' performance. We have analyzed the impact of interference due to massive connectivity in O-RAN; to achieve this, our inter-O-DU coordination and scheduling algorithm for massive users in O-RAN uses the information shared in advance by the distributed units (O-DUs) that are working together to calculate the interference weight and then schedule the successive transmissions of UEs that will cause the least amount of interference. The simulation results show that the proposed architecture and NOMA scheme are more spectrum efficient, as they support approximately more than 1.5 times the number of connected devices for the same number of available resources. The results also show that our proposed algorithm enhances the UEs' achievable throughput by up to 30% and reduces the energy consumption by up to 10% as compared to the stand-alone O-DU architecture. The future outlook involves further studies on the computation complexity of algorithms for real-time use scenarios, especially for a large number of UEs, dynamic traffic models, energy efficiency, latency, and mobility analysis for systems, along with machine learning techniques for resources and interference management.

ACKNOWLEDGMENT

This project has received funding from the European Union's Horizon Europe Research program under grant agreement No 101058505 - 5G-TIMBER. This research has also been supported in part by the Internet of Intelligent Things project of the Estonian IT Academy program and Estonian Research Council under Grant PUT-PRG424.

REFERENCES

- [1] O-RAN Working Group 1. O-RAN Architecture Description 6.00," O-RAN.WG1.O-RAN-Architecture-Description-v06.00. Technical Report, 2022.
- [2] O-RAN Working Group 2. O-RAN AI/ML Workflow Description and Requirements-v1.03. Technical Report, O-RAN Alliance, 2021.
- [3] Abedin, Sarder Fakhru and Mahmood, Aamir et al., "Elastic O-RAN slicing for Industrial Monitoring and Control: A Distributed Matching Game and Deep Reinforcement Learning Approach," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 10, pp. 10808–10822, 2022.
- [4] M. W. Akhtar and A. M. et al., "Exploiting NOMA for Radio Resource Efficient Traffic Steering Use-case in O-RAN," in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, 2022, pp. 5771–5776.
- [5] B. Balasubramanian and E. S. D. et al., "RIC: A RAN Intelligent Controller Platform for AI-Enabled Cellular Networks," *IEEE Internet Computing*, vol. 25, no. 2, pp. 7–17, 2021.
- [6] I. Budhiraja and S. T. et al., "Cross Layer NOMA Interference Mitigation for Femtocell Users in 5G Environment," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 4721–4733, 2019.
- [7] Collins Burton Mwakwata et al., "Cooperative scheduler to enhance massive connectivity in 5G and Beyond by minimizing interference in OMA and NOMA," *IEEE Systems Journal*, vol. 16, no. 3, pp. 5044–5055, 2021.
- [8] Z. Ding, M. Peng, and H. V. Poor, "Cooperative Non-Orthogonal Multiple Access in 5G Systems," *IEEE Communications Letters*, vol. 19, no. 8, pp. 1462–1465, 2015.
- [9] Dryjanski, Marcin and Kulacz, Lukasz et al., "Toward Modular and Flexible Open RAN Implementations in 6G Networks: Traffic Steering Use Case and O-RAN xApps," *Sensors*, vol. 21, no. 24, p. 8173, 2021.
- [10] O. Elgarhy and L. Reggiani, "Application of the water filling algorithm to the sum rate problem with minimum rate and power constraint," in *2018 Advances in Wireless and Optical Communications (RTUWO)*. IEEE, 2018, pp. 12–16.
- [11] Hua, Meng and Wu, Qingqing et al., "Intelligent Reflecting Surface-Aided Joint Processing Coordinated Multipoint Transmission," *IEEE Transactions on Communications*, vol. 69, no. 3, pp. 1650–1665, 2020.
- [12] C. P. R. Interface, "eCPRI Interface Specification, eCPRI Specification, V2. 0, May 2019," 2019.
- [13] Lorenzo Bertizzolo, and Tuyen Xuan Tran et al., "Streaming from the air: Enabling drone-sourced video streaming applications on 5G open-RAN architectures," *IEEE Transactions on Mobile Computing*, 2021.
- [14] Mahmood, Aamir and Beltramelli, Luca et al., "Industrial IoT in 5G-and-Beyond Networks: Vision, Architecture, and Design Trends," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 6, pp. 4122–4137, 2021.
- [15] W. Mei and R. Zhang, "Uplink Cooperative NOMA for Cellular-Connected UAV," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 3, pp. 644–656, 2019.
- [16] Niknam, Solmaz and Roy, Abhishek and Dhillon, Harpreet S et al., "Intelligent O-RAN for Beyond 5G and 6G Wireless Networks," in *2022 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2022, pp. 215–220.
- [17] Polese, Michele and Bonati, Leonardo et al., "Understanding O-RAN: Architecture, Interfaces, Algorithms, Security, and Research Challenges," *IEEE Communications Surveys Tutorials*, pp. 1–1, 2023.
- [18] V. Ranjbar and A. e. a. Girycki, "Cell-free mMIMO support in the O-RAN Architecture: A PHY Layer Perspective for 5G and Beyond Networks," *IEEE Communications Standards Magazine*, vol. 6, no. 1, pp. 28–34, 2022.
- [19] Tandra, Tahmina Khanom and Tajrian, Fehima et al., "Joint Transmission Coordinated Multipoint on Mobile Users in 5G Heterogeneous Network," in *2022 IEEE 2nd Conference on Information Technology and Data Science (CITDS)*. IEEE, 2022, pp. 273–278.
- [20] Z. Yang, Z. Ding, Y. Wu, and P. Fan, "Novel Relay Selection Strategies for Cooperative NOMA," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10 114–10 123, 2017.
- [21] Zhang, Haijun and Fang, Fang and Cheng, Julian et al., "Energy-Efficient Resource Allocation in NOMA Heterogeneous Networks," *IEEE Wireless Communications*, vol. 25, no. 2, pp. 48–53, 2018.

Appendix 2

Publication (II):

Latif, Aamir; Elgarhy, Osama; Le Moullec, Yannick; Alam, Muhammad Mahtab, "Energy Consumption Evaluation of NOMA-based Sustainable Scheduling in 6G O-RAN," in 2024 *International Wireless Communications and Mobile Computing (IWCMC 2024)*, pp. 484–489, 2024

Energy Consumption Evaluation of OMA vs NOMA for Sustainable Scheduling in 6G O-RAN

Aamir Latif, Osama Elgarhy, Yannick Le Moullec, Muhammad Mahtab Alam

Tallinn University of Technology. Email: (aamir.latif; osama.elgarhy; yannick.lemoullec; muhammad.alam)@taltech.ee

Abstract—In sustainable wireless communication networks, it is crucial to assess energy consumption during the scheduling of ultra-massive machine-type communications for the forthcoming 6G radio access network (RAN). We introduce an energy-efficient scheduling algorithm for O-RAN that ensures quality of service (QoS) and reduces complexity. Although 3GPP offers models for energy consumption and efficiency in RAN, these models do not consider the effects of radio access techniques. Therefore, it is essential to determine energy consumption cost of successful packet reception using OMA and NOMA, in line with the 6G sustainability goals and the O-RAN framework. We introduce a mathematical model for energy-efficient scheduler and its analysis in managing ultra-massive machine-type communication (umMTC) devices. The run-time and energy consumption of these algorithms are assessed by calculating and comparing their algorithmic complexity, along with the number of operations and data exchanges. These computations provide a basis for forecasting their run-time and energy consumption. The findings indicate that our proposed scheduling algorithm consumes four times less energy than the benchmark schemes ESAP-NOMA and RAAP-NOMA based on computational complexity, while handling the maximum number of users, and offers a 10% increase in energy efficiency with the maximum NOMA users.

Index Terms—6G, Beyond-5G, O-RAN, Orthogonal Multiple Access (OMA), Non-Orthogonal Multiple Access (NOMA), Power Consumption, Radio Resource Management(RRM).

I. INTRODUCTION

The significance of energy-efficient communication in wireless networks due to the rise in interconnected devices requiring high data rates and minimal latency. These demands must be met while maintaining high Quality of Service (QoS) and Quality of Experience (QoE) [1]. It highlights the need for sustainable scheduling techniques in network design to optimize performance and reduce energy consumption, focusing on a proposed sustainable scheduling algorithm in the Open Radio Access Network (O-RAN). The algorithm's effectiveness is evaluated by comparing its energy consumption in orthogonal multiple access (OMA) and non-orthogonal multiple access (NOMA) scenarios. The goal is to enhance network efficiency and environmental sustainability [2].

O-RAN revolutionizes traditional network architecture by decentralizing and modularizing base stations into components like O-CU, O-DU, and O-RU, which are deployed as Containerized Network Functions on commercial and cloud platforms [3]. This approach, which utilizes open interfaces and interoperability, enhances innovation, flexibility, and cost-

efficiency in managing wireless networks [4]. It incorporates software-defined networking and network function virtualization to improve dynamic scheduling and performance while reducing energy consumption [5]. However, there is a gap in developing a comprehensive power consumption model that fully addresses both hardware and software needs of O-RAN. Current models, such as the EARTH model, are inadequate, highlighting a significant research need for a standardized model specific to O-RAN [6]. Furthermore, in [7], a novel resource management technique using Q-learning has been introduced to optimize power and frequency allocation in small cell base stations coexisting with macro gNBs, aiming to minimize interference and maintain service quality, thereby enhancing spectral efficiency.

Yet, the energy implications of these scheduling techniques and processing overhead in O-RAN are not fully understood, indicating a need for further study. Recent advances in energy consumption within the Radio Access Network (RAN), highlight the role of global standardization efforts by bodies like the 3GPP and the O-RAN Alliance [8]. It emphasizes the use of key performance indicators (KPIs), particularly focusing on the *Joules(J)* metric, which **determines the amount of energy consumed for transmission of the minimum number of bits**. These KPIs are applicable at various network levels, aiding in the assessment and enhancement of energy efficiency in telecommunications networks [9]. The research in [10] addresses the optimization of energy consumption in IoT devices within O-RAN by applying the Successive Convex Approximation (SCA) algorithm to convert a non-convex problem into a convex one. This method not only aims to reduce energy usage by 20% but also enhances the Quality of Service (QoS). It proves more effective than traditional edge network designs in lowering both latency and energy consumption.

However, the study does not cover the computational complexity and the energy costs associated with scheduling algorithms for numerous users, including those for transmission and re-transmission. In [11], the computation offload problem was defined with the goal of minimizing energy usage by jointly determining the offload ratio, transmission power, and local processing speed. To simplify the optimization problem, it was divided into two sub-problems, which were iteratively solved to address the original problem. Simulation outcomes

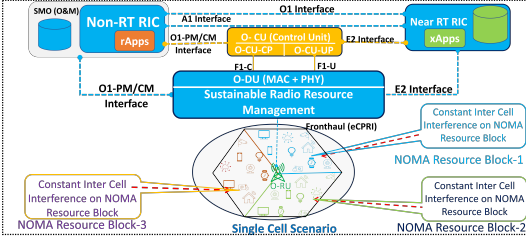


Figure 1. Simplified O-RAN Architecture with control-loop latency with O1, A1, E1, E2, R1 interfaces between Non-RT RIC and Near RT RIC and network function elements O-CU, O-DU, O-RU for single Cell Scenario based on [3].

demonstrate that the strategy obtained for the O-RAN-based IoT system proposed in the study can decrease the average energy consumption of all IoTs by over 20%. In [12], the O-DU coordination architecture using NOMA has been studied for scheduling massive machine-type users to minimize the energy consumption and maximize the connection density and throughput.

Contributions: The contributions of this paper are the following: The paper presents a sustainable scheduling algorithm for O-RAN that significantly reduces energy consumption. Key contributions include:

- A new energy consumption model comparing the proposed algorithm with existing methods such as OMA, ESAP-NOMA, and RAAP-NOMA.
- An analysis using instruction-level power modeling and computational complexity, mapped on the RISC-V Instruction Set Architecture (ISA).
- Simulation results that demonstrate up to four times lower energy usage than reference schemes and 10% more efficiency in handling maximum NOMA users.

II. SUSTAINABLE SCHEDULING IN O-RAN

A. Proposed Sustainable Scheduler System Model

What follows presents the system model and then our proposed solution for sustainable scheduling, that is, the EEPs-MTCN scheduling algorithm and its calculation of energy consumption along with time complexity in Section III.

We consider a conventional scenario with NOMA UEs and assume maximum interference from all other cells, where a central gNB serves N users, each equipped with a single antenna. It examines the impact of maximum interference from neighboring cells and the allocation of radio resources. Each user, denoted as u_s , belongs to a set U_s within cell s , part of a larger set S . The radio resources are indexed by r and range from 1 to R . The study also considers a maximum number, m , of devices that can be supported per resource block, influenced by receiver complexity and processing delays. NOMA implementation allows concurrent sharing of a single resource block by multiple umMTC UEs, enhancing connection capacity.

The achievable data rate measured in bits per second (bps) of the user u belonging to the cell s on a given resource block r is expressed as follows:

$$Rate_u^r = B_0 \log_2(1 + SINR_u^r), \forall u \in U, \forall r \in R \quad (1)$$

where $SINR_u^r$ is the signal-to-interference plus noise ratio experienced by NOMA user u on a given resource block r and B_0 is the bandwidth. SINR for NOMA users can be expressed as:

$$SINR_u^r = a_u^r \left(\frac{G_u^r P_u^r}{\sum_{u \neq v, v \in U} G_v^r P_v^r a_v^r + Int^r + P_n} \right) \quad (2)$$

where a_u^r is a binary allocation variable, G_u^r denotes channel gain and P_u^r the transmission power of the user u in cell s on resource block r . Furthermore, G_v^r is the channel gain and P_v^r the transmission power of a NOMA interfering user v in the resource block r . The term P_n is the noise power and Int^r is the inter-cell interference, assumed to be the constant interference experienced by the user u by the interfering users of the adjacent cell NOMA users. Here, we consider that:

B. Problem Formulation for Energy Efficiency Maximization

In this work, the achievable EE ($b/J/Hz$) of the proposed network is computed as the sum rate of the umMTC devices ($b/s/Hz$) divided by the total power consumption (W) of the network. The objective is to maximize the EE of the considered system while guaranteeing QoS for each user, that is, $Rate_u^r \geq Rate_u^{\min}$.

$$\eta^{EE} = \frac{\sum_{u \in U} \sum_{r \in R} \log_2(1 + SINR_u^r)}{p_c + \sum_{u \in U} \sum_{r \in R} P_u^r} \quad (3)$$

Given the SINR from Equation (2), the above equation becomes:

$$\eta^{EE} = \frac{\sum_{u \in U} \sum_{r \in R} \log_2(1 + a_u^r \left(\frac{G_u^r P_u^r}{\sum_{u \neq v, v \in U} G_v^r P_v^r a_v^r + Int^r + P_n} \right))}{p_c + \sum_{u \in U} \sum_{r \in R} P_u^r} \quad (4)$$

The EE maximization is given as:

$$\underset{p_u^r \geq 0}{\text{maximize}} \quad \eta^{EE} \quad (5)$$

Furthermore, we have defined the constraints outlined below.

- **Minimum SINR Constraint:** The minimum SINR requirement ensures the system's viability within the network, where $\gamma_{threshold}$ represents the minimum SINR level required for user u_s to meet the Quality of Service (QoS).

$$(C1) : SINR_u^r \geq \gamma_{threshold} \quad \forall r \in R \quad (6)$$

- **Minimum and Maximum Power Constraint** stands for the minimum and maximum transmit power constraint of each umMTC UE.

$$(C2) : 0 \leq P_u^r a_u^r \leq P_{max}, \quad \text{forall } r \in R, \forall u \in U \quad (7)$$

- **Binary Resource Constraint** indicates the binary constraint of resource block (RB) allocation. i.e. $a_u^r \in \{0, 1\}$, where $a_u^r = 1$ represents that the resource block is allocated to the device.

$$(C3) : a_u^r \in \{0, 1\}, \forall r \in R \quad (8)$$

- Resource Allocation guarantees that each RB is allocated to M umMTC devices .

$$(C4) : \sum_{u \in U} a_u^r \leq M \quad \forall r \in R \quad (9)$$

- Non-overlapping devices in the same cell requires that each umMTC device is allocated to one RB.

$$(C5) : \sum_{r \in R} a_u^r \leq 1 \quad \forall u \in U \quad (10)$$

C. Energy Efficiency Prediction Algorithm for Sustainable Scheduling

The proposed EEPS-MTCN algorithm designed to enhance energy efficiency in device networks by clustering devices and assigning them to specific resource blocks (see Algorithm1). This method ensures all devices in a cluster achieve necessary data rates and focuses on selecting umMTC devices to minimize interference and maintain connectivity. The algorithm also involves power allocation to maximize EE, reduce latency, and preserve connectivity, with the number of devices per block limited by interference thresholds to optimize resource access.

Algorithm 1 :Proposed Algorithm for Sustainable Scheduling

```

Require  $BW, M_{NOMA}, P_{max}, G_u^r, P_n, Area, ISD$ 
User Equipment creation  $\triangleright$  Set  $u_s \leftarrow G_{u_s, s}^r$ 
Initialize  $P_u^r = P_{max}$  and Compute CSI
Step 1: Resource Allocation  $STN^R$  as in (Eq.2)
while NumSector do
  for Utrs in NOMARB do
    Calculate  $IntU_M = Sum(IntU_M)$ 
    Calculate  $RtU_M = Sum(RtU_M)$ 
    Calculate  $EEU_M = Sum(EEU_M)$ 
  end for
  Best Combination  $IntU_{Mmin}, RtU_{Mmax}$ 
  Best Combination  $EEU_{Mmax}$ 
  Assign RB  $IntU_{Mmin}, RtU_{Mmax}, EEU_{Mmax}$ 
end while
Require  $IntU_{Mmin}, RtU_{Mmax}, EEU_{Mmax}$ 
for CHGain in  $IntU_{Mmin}$  do
  for Utr in NOMARB do
    Calculate  $IntU_M = Sum(IntU_M)$ 
    Calculate  $RtU_M = Sum(RtU_M)$ 
    Calculate  $EEU_M = Sum(EEU_M)$ 
  end for
  Best Combination  $IntU_{Mmin}, RtU_{Mmax}$ 
  Best Combination  $EEU_{Mmax}$ 
end for
Assign RB  $IntU_{Mmin}, RtU_{Mmax}, EEU_{Mmax}$ 
return  $IntU_{Mmin}, RtU_{Mmax}, EEU_{Mmax}$ 
=====
Require  $IntU_{Mmin}, RtU_{Mmax}, EEU_{Mmax}$ 
Step 2: Power Allocation Require  $P_u^r = PowMat$ 
while UtrChGain do
  for Power in PowMat do
    for Utr in NOMARB do
      if  $STN^R \geq STN^R_{min}$  then
        Calculate  $RtU_M = Sum(RtU_M)$ 
        Calculate  $EEU_M = Sum(EEU_M)$ 
      end if
    end for
    Calculate Power  $RtU_{Mmax}, EEU_{Mmax}$ 
    return  $RtU_{Mmax}, EEU_{Mmax}, Pwr_{alloc}$ 
  end while

```

III. ENERGY CONSUMPTION ESTIMATION AND MODELING

The energy consumption models and the estimation method are discussed to calculate the complexity and required number of operations for the proposed EEPS-MTCN, and the baseline ESAP-NOMA, RAAP-NOMA, and OMA are derived.

A. Network Energy Consumption

We estimate the energy consumption of an UE using the following model proposed and used in [13]. The Energy E consumed by UE is given as

$$\frac{E}{V_{sup}} = I_{act}t_{act} + I_{low}t_{low} + I_{trx}t_{trx} + I_{rx}t_{rx} + \sum I_{oth}t_{oth}, \quad (11)$$

where V_{sup} represents the power supply voltage, I_{act} and t_{act} denote the current consumption and processing time of the O-DU processor in active mode, I_{low} and t_{low} refer to the current consumption and processing time in low power mode, I_{trx} , t_{trx} , I_{rx} , and t_{rx} indicate the current consumption and processing times of the radio in transmit and receive modes, while I_{oth} and t_{oth} stand for the current consumption and processing time of other network components. By prioritizing the comparison and evaluation of algorithm performance, we simplify the equation to concentrate solely on the active, transmit, and receive states. We presume that the power consumption in the low-power state and other components remain consistent for both algorithms. It is essential to note that the computational time and the duration of transmission and reception vary for each UE. We assume that the computation for m NOMA users varies from 2,4 to 8 users per radio resource block.

$$\frac{E_n(m)}{V_{sup}} = I_{act}t_{act,n}(m) + I_{trx}t_{trx,n}(m) + I_{rx}t_{rx,n}(m) \quad (12)$$

where the energy consumption for UE n in iteration m is:

$$E_n(m) = V_{sup}(I_{act}t_{act,n}(m) + I_{trx}t_{trx,n}(m) + I_{rx}t_{rx,n}(m)) = E_{comp,n}(m) + E_{radio,n}(m).$$

UE n uses $E_n(m)$ energy per iteration m . $E_{radio,n}(m)$ is the communication energy consumption of node n and the computational energy consumed by node n at iteration m is $E_{comp,n}(m)$. The energy consumed by the network in one iteration m is given as:

$$E_{network}(m) = \sum_{n=1}^N E_n(m). \quad (13)$$

The energy consumed by the network during the observation period m_{obs} is given as:

$$E_{obs} = \sum_{m=1}^{m_{obs}} E_{network}(m). \quad (14)$$

B. Computational Energy Consumption

The energy consumption of algorithms is evaluated primarily through their algorithmic complexity and the total number of operations, which is a reliable metric for assessing energy use [14]. Furthermore, the energy usage of the RISC-V ISA processor differs with various operations as detailed in the instruction sets [15] and is further influenced by the execution time and operational demands as discussed in instruction level modeling (Section III-B1).

1) **Instruction Level Energy Model:** Instructions, such as NOP, ADD, and MUL, consume varying amounts of energy due to their use of different components within the processor's data path. It highlights that each instruction has a fundamental energy cost associated with the core processing required for its execution [16] for detailed energy consumption metrics.

$$E_m = \sum_m (B_m \times N_m) + \sum_{m,j} (O_{m,j} \times N_{m,j}) + \sum_n E_n \quad (15)$$

The total energy consumption of a program, denoted as E_m , is comprised of three main components. Firstly, the base energy cost of each instruction, represented by B_m for the base cost and N_m for the execution count of instruction m . Secondly, the energy cost due to instruction switching, where $O_{m,j}$ is the overhead for switching between instructions (m, j) and $N_{m,j}$.

is the frequency of these transitions. Lastly, E_n accounts for additional energy used due to resource constraints. The model assumes that the current I used for executing instruction m depends on the supply voltage (V_{sup}) and clock frequency (f), forming a basic framework for understanding energy consumption in programming.

$$E_{tot,n}(m) = I_n \times V \times t_n(m) \quad (16)$$

$$t_n(m) = f(N_{mul,k}(m)C_{mul} + N_{div,k}(m)C_{div} + N_{add,k}(m)C_{add} + N_{comp,k}(m)C_{comp}) \quad (17)$$

Equation 16 defines the time taken for the execution of an instruction, denoted as $t_n(m)$. The processor's frequency is symbolized by the variable f . The quantities $N_{mul,k}(m)$, $N_{div,k}(m)$, $N_{add,k}(m)$, and $N_{comp,k}(m)$ indicate the counts of multiplication, division, addition, and comparison instructions, respectively. The variables C_{mul} , C_{div} , C_{add} , and C_{comp} denote the number of clock cycles needed for multiplication, division, addition, and comparisons, respectively.

C. Time Complexity Analyses

1) Time Complexity Analysis for NOMA-based Proposed Sustainable Scheduling Algorithm: A three-step algorithm for optimizing signal-to-interference plus noise ratio (SINR) and energy efficiency in a network with non-orthogonal multiple access (NOMA) users. Initially, the algorithm calculates channel gains and SINR for each user, with a complexity of $O(n + n^m)$. Subsequently, it divides users into sectors, computes interference combinations for each resource block, and identifies optimal combinations through a linear search, resulting in a complexity of $O((n^m) + n^m \times m^2)$. Finally, it allocates power to these combinations based on channel gains to maximize efficiency, with a complexity of $O(n^m \times m^2)$. The overall computational complexity of the algorithm is thus detailed, highlighting its steps and computational demands.

$$O((n^m) + n + (n^m) \times m^2 + (n^m) + (n^m) \times m^2) \quad (18)$$

which simplifies to $O(n^m \times m^2)$. For comparison, we present the time complexity of the ESAP-NOMA, RAAP-NOMA, and OMA baseline algorithms namely Algorithms 2 and 3.

2) Time Complexity Analysis for NOMA-based Reference Algorithm: Algorithm(2) outlines a three-step process for optimizing user scheduling in NOMA systems. Initially, it calculates the SINR for each user with a complexity of $O(n)$. The second step, using the ESAP method, evaluates all possible user combinations to minimize interference, with a complexity of $O(m^2 + n^m + n \times m \times m)$. The final step involves optimal power allocation based on channel gains to improve energy efficiency, with a complexity of $O(n^m + 1 \times m \times m^3)$. Overall, the algorithm aims to enhance system performance through efficient user pairing and power distribution.

$$O(n^m \times m^2 + n \times m + n^m + 1 \times m^3) \quad (19)$$

3) Time Complexity Analysis for Random Sustainable Scheduling: In Algorithm(3), *Step1* involves computing channel gains and signal-to-interference plus noise ratio (SINR) for each user. The computational complexity of *Step 1* is $O(n)$, where n is the number of users. This is followed by *Step2* employing a random pairing of users to NOMA resource block. The computational complexity of *Step2* is based on the power allocation loop and is based on the search for the best pairs. The computational complexity of this step is $O((n^m) \times m^2)$. Thus, the overall time complexity of the RAAP-NOMA algorithm can be expressed as $O(n + n \times m^2)$.

Algorithm 2 :ESAP Algorithm for Sustainable Scheduling

```

Require  $BW, MNOMA, P_{max}, G_u^r, P_n, Area, ISD$ 
User Equipment creation  $\triangleright$  Set  $u_s \leftarrow G_{u,s}^r$ 
Initialize  $P_u^r = P_{max}$  and Compute CSI
Step 1: Resource Allocation Calculate  $SLNR$  as in (Eq.2)
while NumSector do
  for Usrs in TotUsr do
    for Usrs in NOMARB do
      Calculate  $IntU_s = M \times NAC \times P_{max} \times GAC$ 
      Calculate  $IntU_s = Sum(IntU_s)$ 
      Calculate  $RtU_s = Sum(RtU_s)$ 
      Calculate  $EEU_s = Sum(EEU_s)$ 
    end for
    Best Combination  $IntU_s min, RtU_s max$ 
    Best Combination  $EEU_s max$ 
    Assign RB  $IntU_M min, RtU_M max, EEU_M max$ 
  end while
  return  $IntU_M min, RtU_M max, EEU_M max$ 
=====
Require  $IntU_M min, RtU_M max, EEU_M max, LatU_M min$ 
Step 2: Power Allocation Require  $P_u^r = PowMat$ 
while UsrChGain do
  for Power in PowMat do
    for Usrs in NOMARB do
      if  $SLNR \geq SLNR_{min}$  then
        Calculate  $RtU_M = Sum(RtU_M)$ 
        Calculate  $EEU_M = Sum(EEU_M)$ 
      end if
    end for
    Calculate Power  $RtU_M max, EEU_M max$ 
  end while
  return  $RtU_M max, EEU_M max, Pwr_{alloc}$ 
end while

```

4) Time Complexity Analysis for OMA Sustainable

Scheduling: In the OMA Algorithm, *Step1* involves computing channel gains and Signal-to-Interference-plus-Noise Ratio (SINR) for each user. The computational complexity of *Step 1* is $O(n)$, where n is the number of users. This is followed by *Step2* employing resource block allocation to the user one at a time. The computational complexity of this step is $O(n)$. Thus, the overall time complexity of the OMA algorithm can be expressed as $O(n)$.

Algorithm 3 :RAAP-NOMA Scheduling Algorithm

```

Require  $BW, NumUsr, P_{max}, G_u^r, P_n, Area, ISD$ 
User Equipment creation  $\triangleright$  Set  $u_s \leftarrow G_{u,s}^r$ 
Initialize  $P_u^r = P_{max}$  and Compute CSI
Step 1: Resource Allocation Calculate  $SLNR$  as in (Eq.2)
for Usrs in NOMARB do
  Assign RB  $RandTup = UsrComb$ 
  Assign RB  $IntU_{RandM}, RtU_{RandM}, EEU_{RandM}$ 
end for
=====
Require  $IntU_{RandM}, RtU_{RandM}, EEU_{RandM}$ 
Step 2: Power Allocation  $P_u^r = PowMat$ 
while UsrChGain do
  for Power in PowMat do
    for Usrs in NOMARB do
      if  $SLNR \geq SLNR_{min}$  then
        Calculate  $RtU_M = Sum(RtU_M)$ 
        Calculate  $EEU_M = Sum(EEU_M)$ 
      end if
    end for
    Calculate Power  $RtU_M, EEU_M, LatU_M$ 
  end while
  return  $EEU_M, LatU_M, Pwr_{alloc}$ 
end while

```

D. Communication System Energy Consumption

In addition to the computation costs, let's calculate the energy consumption for radio communication $E_{radio,n}(m)$ at UE n during iteration m which can be expressed as

$$E_{radio,nk}(m) = V_{sup}(I_{rx}t_{rx,n}(m) + I_{tx}t_{tx,n}(m)) \quad (20)$$

The current consumed by the transmitter and receiver is denoted as I_{tx} and I_{rx} , respectively. Additionally, $t_{rx,n}(m)$ and $t_{tx,n}(m)$ represent the duration of communication of the receiver and transmitter at UE n in iteration m , respectively.

$E_{radio,n}(m) = (N_n(m) - 1)V_{sup}(I_{rx}t_{rx} + I_{tx}t_{tx})$, (21)
where $n_n(m) - 1$ is the number of neighbours for UE n . We can further write that the network radio consumption can be given as:

$$E_{radio}(m) = \left(\sum_{n=1}^M (N_n(m) - 1) \right) V_{sup}(I_{rx}t_{rx} + I_{tx}t_{tx})$$

$$= (N(m) - M)V_{sup}(I_{rx}t_{rx} + I_{tx}t_{tx}), \quad (23)$$

IV. RESULTS

A. Simulation Setup

The proposed and baseline algorithms were rigorously tested by running the simulation 1000 times with UEs randomly distributed across a single-cell network with a 500-meter inter-site distance. The connection density of the UEs in the cell was uniformly and randomly distributed between 16 to 50, with a maximum transmit power of 23 dBm in the uplink. We map our proposed sustainable scheduling algorithm (EEPS-MTCN) and baseline algorithms (ESAP-NOMA, RAAP-NOMA, and OMA) on the RISC-V architecture and used the architecture instruction set to perform instruction-level modeling and energy consumption calculations. We conducted a thorough analysis of the total energy consumption of the algorithms, comparing the baseline algorithm with the proposed sustainable scheduling algorithm. We also provided a detailed breakdown of the transmission and computational energy consumption to gain a comprehensive understanding of each algorithm's overall energy consumption.

B. Performance Evaluation and Discussion

In our simulation, the total number of umMTC devices range from 16 to 50, with the maximum number of umMTC devices per resource block being $M_{mux} = 8$. We consider an imperfect successive interference cancellation (SIC) case; the users experiencing interference from other users pair to the

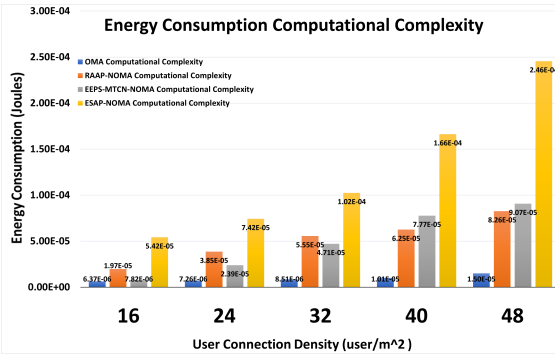


Figure 2. Energy Consumption (Joules) comparison for different Algorithms based on Computational Complexity based on Instruction level modeling with the different user connection density for the baseline ESAP-NOMA, baseline OMA, and the proposed EEPS-MTCN.

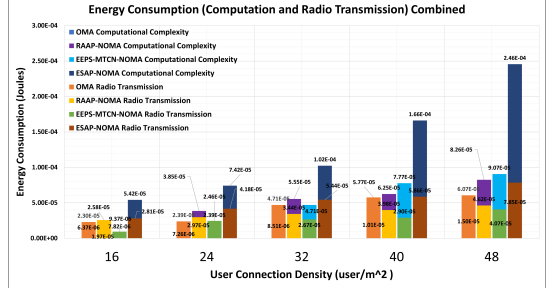


Figure 3. Total Energy Consumption(Joules) based on Computational and Radio transmission and comparison of proposed EEPS-MTCN Algorithm for different connection density $usersq/m$ with respect to baseline (ESAP-NOMA, RAAP-NOMA, and OMA).

same resource. We conducted 5000 iterations for the Signal-to-Interference-plus-Noise Ratio (SINR) to evaluate energy consumption based on the successful packet rate. The prediction was made using the receiver sensitivity value, and only users with SINR above the threshold were considered. Although this approach may affect user connection density, it ensures successful packet reception and a balanced trade-off between energy consumption and re-transmission. We conducted a comparison of the total energy consumption with respect to the user connection density $usersq/m$ for various algorithms based on the summation of computational complexity.

As illustrated in Figure 2, it is evident that the energy consumption of ESAP-NOMA significantly exceeds that of the proposed EEPS-MTCN. This is attributed to the notably higher computational complexity of ESAP-NOMA in contrast to that of EEPS-MTCN. It is also observable that the total energy consumption of OMA is lower than that of EEPS-MTCN. This can be attributed to the fact that the computational complexity of the OMA algorithm is $O(n)$ compared to EEPS-MTCN, albeit at the expense of a higher number of resource blocks, which is nearly eight times more than that of the proposed algorithm. Similarly, the computational complexity of RAAP-NOMA is lower, which plays a significant role in this comparison. It is evident from Figure(2) that the energy consumption of ESAP-NOMA significantly

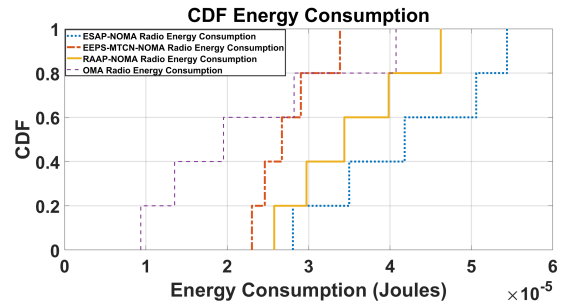


Figure 4. CDF for the comparison of Radio transmission energy consumption of proposed EEPS-MTCN Algorithm with respect to baseline (ESAP-NOMA, RAAP-NOMA, and OMA).

exceeds that of the proposed EEPS-MTCN (up to 29.68 less energy vs. ESAP i.e. 58 mJoules for 48 users/m²). This is attributed to the notably exhaustive search and computation complexity of ESAP-NOMA in contrast to EEPS-MTCN. It is also observable that the total energy consumption of OMA is lower than that of EEPS-MTCN. This can be attributed to the single-user power consumption compared to EEPS-MTCN, albeit at the expense of a higher number of resource blocks, which is nearly eight times more than that of the proposed algorithm. A comparison was also carried out on the overall energy usage (comprising computational and radio transmission costs) in relation to the user connection density $user(sq/m)$ for different algorithms.

As shown in Figure(3), the energy consumption of EEPS-MTCN for radio transmission outperforms ESAP-NOMA by 48%, RAAP-NOMA by 11.9%, and OMA by 32.9%. This indicates that the accurate prediction of Signal-to-Interference-plus-Noise Ratio (SINR) significantly influences the performance of this algorithm.

The energy consumption for successful packet transmission is illustrated in Figure 4, it is evident that EEPS-MTCN has less energy consumption as compared to other baseline NOMA algorithms but for higher connection densities it also surpass OMA and more energy efficient due to the inclusion of minimum interference while packing NOMA user in a resource block and prediction of SINR for successful packet transmission. The superiority of our proposed scheme over the baseline schemes stems from their following shortcomings. Although the ESAP NOMA algorithm explores all possible user pairs in the system to identify the optimal pairing that maximizes performance metrics, but it is very computationally intensive and time complexity is very high. It finds the pair for interference and latency considerations. In contrast, the RAAP-NOMA algorithm randomly pairs NOMA users, which generates a diverse range of user pairs but overlooks all the QoS constraints and pair without looking into interference and latency considerations. On the other hand, our proposed scheme minimizes interference and latency and specifically pairs devices with favorable channel gain and minimal interference and latency to enhance the EE of the system.

V. CONCLUSION

This paper proposed a sustainable scheduling algorithm for O-RAN that prioritizes energy efficiency, maintains QoS, while decreasing complexity. To realistically evaluate the energy consumption, we calculated the cost of successful packet reception with OMA and NOMA, considering 6G and O-RAN architecture and umMTC. Various factors associated with sustainable scheduling techniques that impact energy consumption in O-RAN systems have been explored including user density, radio resource and power allocation, and algorithm computation complexity. The simulation results show that the proposed scheduling algorithm requires four times less energy based on the complexity of the computation than the

baseline ESAP-NOMA and RAAP-NOMA, while scheduling the maximum number of users, and is 10% more energy efficient for the maximum NOMA users.

ACKNOWLEDGMENT

This project has received funding from the European Union's Horizon Europe Research Program under grant agreement No. 101058505 - 5G-TIMBER, and from the Estonian Education and Youth Board ÕF11 "AIoT*5G - Artificial intelligence, edge computing and IoT solutions in distributed systems".

REFERENCES

- [1] NGMN, "Ngmn, "green future networks: Network energy efficiency," white paper, v1.1, dec. 2021." NGMN, Tech. Rep., 2021.
- [2] L. K. *et al.*, "Towards energy efficient ran: From industry standards to trending practice," in *ICC 2024 - IEEE International Conference on Communications*, 2024.
- [3] O-RAN Working Group 1. O-RAN Architecture Description 6.00," O-RAN.WG1.O-RAN-Architecture-Description-v06.00. Technical Report, 2022.
- [4] Dryjanski, Marcin *et al.*, "Toward Modular and Flexible Open RAN Implementations in 6G Networks: Traffic Steering Use Case and O-RAN xApps," *Sensors*, vol. 21, no. 24, p. 8173, 2021.
- [5] t. Aslan, Fatih, "Fair Resource Allocation in Virtualized O-RAN Platforms," *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 2024.
- [6] Ayala-Romero, "Experimental evaluation of power consumption in virtualized base stations," in *2021-IEEE International Conference on Communications*, 2021, pp. 1–6.
- [7] M. W. Akhtar, A. Mahmood, S. F. Abedin, S. A. Hassan, and M. Gidlund, "Exploiting noma for radio resource efficient traffic steering use-case in o-ran," in *GLOBECOM 2022-2022 IEEE Global Communications Conference*. IEEE, 2022, pp. 5771–5776.
- [8] Arnaz, Azadeh *et al.*, "Toward Integrating Intelligence and Programmability in Open Radio Access Networks: A Comprehensive Survey," *IEEE Access*, 2022.
- [9] ETSI, "5G; management and orchestration end to end key performance indicators (kpi) (3gpp ts 28.554 version 17.7.0 release 17)," ETSI, Tech. Rep., 2017.
- [10] L. t. Wang, "Minimizing energy consumption of iot devices for o-ran based iot systems," *Energy Reports*, vol. 9, pp. 379–388, 2023.
- [11] Wang, Liping *et al.*, "Energy Conserved Computation Offloading for O-RAN based IoT Systems," in *IEEE ICC-2022*, 2022.
- [12] Latif, Aamir *et al.*, "Inter O-DUs Coordination for Scheduling of Massive Users in O-RAN," in *EuCNC-6G Summit*, 2023, pp. 1–6.
- [13] Dunkels, Adam *et al.*, "Software-based on-line energy estimation for sensor nodes," in *Proceedings of the 4th Workshop on Embedded Networked Sensors*, ser. EmNets '07. New York, NY, USA: Association for Computing Machinery, 2007, p. 28–32.
- [14] W. Wang, "An improved instruction-level power and energy model for riscmicroprocessors," University of SOUTHAMPTON, Tech. Rep., 2017.
- [15] G. Fang, "Instruction-level power consumption simulator for modeling simple timing and power side channels in a 32-bit risc-v micro-processor," MIT, Tech. Rep., 2021.
- [16] V. Tiwari, S. Malik, and A. Wolfe, "Power analysis of embedded software: A first step towards software power minimization," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 2, no. 4, pp. 437–445, 1994.

Appendix 3

Publication (III):

Latif, Aamir; Elgarhy, Osama; Le Moullec, Yannick; Alam, Muhammad Mahtab, "Energy Efficiency and Connectivity Maximization for Massive Machine Type Communication in Multi-Cell NOMA," in *2024 19th Biennial Baltic Electronics Conference (BEC 2024)*, pp. 1–6, IEEE, 2024

Energy Efficiency and Connectivity Maximization for Massive Machine Type Communication in Multi-Cell NOMA

Aamir Latif, Osama Elgarhy, Yannick Le Moullec, Muhammad Mahtab Alam
Tallinn University of Technology, Estonia.

Email: (aamir.latif; osama.elgarhy; yannick.lemoullec; muhammad.alam)@taltech.ee

Abstract—This paper proposes an energy-efficient resource scheduling and power-allocation framework for ultra-massive machine-type communication (umMTC) in uplink non-orthogonal multiple access (NOMA) networks. The considered non-cooperative multi-cell model accounts for neighboring-cell interference using a fixed-interference approximation and addresses limited radio resources, stringent device power budgets, and quality-of-service constraints. Energy-efficiency and connectivity optimization problems are formulated for joint user scheduling, resource-block assignment, and transmit-power control. Because the resulting problem is mixed-integer and non-convex, a low-complexity Energy-Efficient Resource Scheduling for Machine-Type Communications in NOMA (EERS-MTCN) algorithm is developed using channel-aware user grouping, interference-aware resource allocation, and iterative power refinement. Simulation results show that EERS-MTCN consistently outperforms random allocation and pairing NOMA (RAAP-NOMA), achieving approximately 27% higher energy efficiency at the maximum evaluated connection density of 300 users/m². Although exhaustive-search pairing NOMA (ESAP-NOMA) achieves approximately 16% higher energy efficiency, its execution time is about one order of magnitude greater. The proposed method therefore provides a favorable trade-off between energy efficiency, connectivity support, and computational complexity for dense uplink NOMA deployments.

Keywords—6G, Massive machine type communication (mMTC), Sustainable scheduling, Non-Orthogonal Multiple Access (NOMA), Radio Resource Management (RRM).

I. INTRODUCTION

Sixth-generation (6G) communication systems require a sophisticated approach to manage concurrent communication needs. By 2030, it is projected that 25 billion machine-type communication (MTC) devices will be in operation [1]. The proliferation of umMTC devices has introduced complexities in connection establishment, radio resource allocation, simultaneous signaling, and data transmission, potentially leading to significant congestion. Addressing these challenges remains an open research topic [2]. However, sustaining this growth in ultra-massive machine-type communication (umMTC) devices presents significant challenges to accommodate dense deployments and ensure reliable and energy-efficient connectivity [3].

Therefore, the key question is *How can we efficiently use limited spectrum resources to connect a large number of umMTC devices with minimal latency and high energy efficiency?*

In M2M communication, frequent access congestion can occur due to simultaneous connections by massive machine-

type communication devices (MTCs) to the core network. To mitigate this issue, non-orthogonal multiple access (NOMA) is suggested as a potential solution because it can serve multiple users on the same time-frequency channel resource. It has been demonstrated that NOMA can outperform orthogonal multiple access (OMA) in providing more equitable access opportunities for devices [4]. However, it is crucial to note that this benefit is conditional, and the increased cochannel interference caused by NOMA must be managed more effectively. Existing studies have explored power allocation, channel assignment [5], and other methods to manage cochannel interference.

Nonetheless, these studies assume that the network can support connections to all users simultaneously. This assumption may limit the applicability of current research in M2M communication systems, which typically accommodate far more devices than traditional networks. Therefore, user scheduling is necessary to enhance the performance of NOMA-based M2M communication networks. User scheduling schemes in the time domain have been investigated in [6], [7]. Specifically, the uplink user scheduling problem was examined in [6] to maximize connections in narrow-band IoT systems with NOMA. In [7], under the constraint of a maximum number of multiplexed users sharing the same sub-channel, the authors focused on the joint user scheduling and power allocation problem for NOMA systems.

In the MTC domain, research has been conducted to maximize network efficiency, resulting in dynamic user scheduling and power allocation techniques, as discussed in [8]. This work achieved an average energy efficiency of 4.5×10^{10} bits per joule for an average density of 150 MTC User Equipment's (UEs) /km², without considering the delay and arrival rate. The work by [9] introduces a power allocation technique demonstrating a 0.5 b/J/Hz enhancement over the standard scheme when considering up to 24 IoT devices across 12 sub-carriers. The author of [10] investigated dynamic user scheduling and power allocation by formulating a stochastic optimization problem to minimize the total power consumption of a massive IoT network.

In [11], the author investigated the enhancement of energy efficiency (EE) in narrow-band IoT systems by comparing different resource unit configurations (RUCs). They aimed to optimize both EE and latency and proposed a scheduler based on the shortest job first. The scheduler achieved a better balance between reducing latency and improving EE across a wider range of parameter configurations in multicellular layouts. Although the aforementioned works [6], [7], [11] have addressed the massive access problem by utilizing user

scheduling schemes, they do not consider the QoS requirements of users. Unfortunately, for the overloaded network we considered, the cochannel interference caused by NOMA may negatively impact the achievable rate of users and further reduce QoS satisfaction, which is a critical performance metric for M2M communication networks. Therefore, unlike the above strategies, we focus on designing a user scheduling scheme under QoS provisioning for umMTC.

To enhance connectivity, it is essential to tackle interference, scheduling delays, and the reliability of wireless transmission. These elements are constrained by the available bandwidth (BW) and the signal-to-noise ratio at the receiver [12]. For ultra-massive Machine Type Communication (umMTC), managing high user density, low transmission power, and limitations on BW and radio resources is crucial. From an energy efficiency standpoint, effective resource scheduling algorithms are vital for massive machine-type communication in 6G. These algorithms are key to supporting many users with minimal interference. Given the limited availability of radio spectrum resources, it is crucial to strategically allocate them based on the specific needs of the application and the current radio conditions.

Previous research [13] has focused primarily on coordination mechanisms that, while addressing certain challenges, often lead to increased network complexity and energy consumption, especially in high-density device scenarios. The diverse data rate requirements of ultra massive machine type communications (umMTC) have not been fully addressed in the literature, potentially causing QoS issues, especially for low data rate devices. Combinatorial optimization methods have been used to tackle the challenges of massive connectivity, but they are lacking in accommodating the maximum number of users in a single resource block with NOMA to enhance energy efficiency and connectivity.

A. Contribution

The contributions of this paper are threefold.

- We develop a non-cooperative multi-cell uplink NOMA system model for dense umMTC scheduling. Each cell allocates resources independently, while the interference originating from neighboring cells is incorporated through a conservative fixed inter-cell interference term. The model captures the impact of adjacent-cell activity on user SINR, achievable rate, and scheduling feasibility without assuming explicit inter-cell coordination.
- We formulate constrained energy-efficiency and connectivity optimization problems for joint user scheduling, resource-block assignment, and power allocation. The formulation incorporates binary allocation variables, maximum NOMA multiplexing order, per-user power limits, minimum SINR requirements, and one-RB-per-user constraints. Connectivity is promoted by scheduling user groups whose members satisfy the required QoS and minimum-rate conditions.
- We propose the Energy-Efficient Resource Scheduling for Machine-Type Communications in NOMA (EERS-MTCN) algorithm. The method combines channel-aware user grouping, interference-aware resource allocation, and power refinement to reduce intra-RB

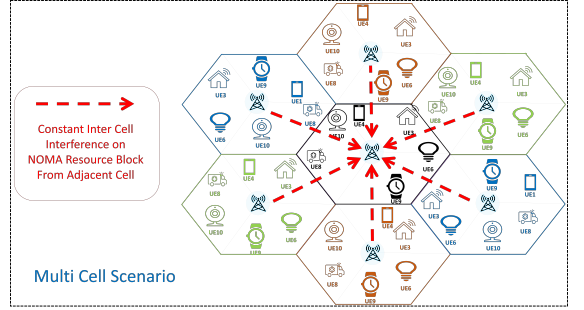


Figure 1: Multi-Cell Non-Cooperative Interference Scenario Consideration

interference while maintaining QoS. Compared with RAAP-NOMA, EERS-MTCN achieves approximately 27% higher EE at the maximum evaluated density. It remains approximately 16% below ESAP-NOMA in EE but requires about one-tenth of its execution time, providing a more practical EE-complexity trade-off.

The rest of this paper is organized as follows. Section II describes the system model and defines problem formulation for energy efficiency and connectivity maximization for umMTC. Section III presents the details of our proposed solution and the proposed algorithm for energy-efficient resource scheduling. Section IV presents an evaluation of the performance of our proposed solution through simulations. Finally, Section V concludes the paper with a summary of our findings and suggestions for future work.

II. PROBLEM FORMULATION FOR RESOURCE SCHEDULING FOR UMMTC SYSTEM

A. System Model

We consider a non-cooperative multi-cell scenario with NOMA UEs and assume maximum interference from all adjacent cells. Our primary focus is on the central cell (referred to as the multi-cell scenario as in Fig (1)). The gNB in the network serves N users. It is assumed that both the gNB and all umMTC devices are equipped with a single antenna. Let u_s^r denote the user index u in cell s , utilizing the resource block indexed by r , and belonging to the set of users U_s , where all cells are represented by $s \in \mathcal{S}$.

Hence total users served by the gNB are $N \in \mathcal{U}$ and the set of available radio resources is denoted by R , with indices $r = \{1, 2, \dots, R\}$. We consider $m \in \mathcal{M}$, representing the maximum number of NOMA UEs that can be supported per radio resource block (RB), where $M \leq N$, depending on the complexity of the receiver design and the delay associated with signal processing for successive interference cancellation. The deployment of NOMA facilitates the concurrent sharing of a single resource block by multiple umMTC UEs, thereby enhancing the capacity for additional connections. The achievable data rate measured in bits per second (bps) of the user u belonging to the cell s on a given resource block r is expressed as follows:

$$Rate_{u,s}^r = B_0 \log_2(1 + SINR_{u,s}^r), \forall r \in \mathcal{R}, \forall s \in \mathcal{S}, \forall u \in \mathcal{U} \quad (1)$$

where $SINR_{u,s}^r$ is the signal-to-interference plus noise ratio experienced by NOMA user u in cell s in a given resource

block r and B_0 is the bandwidth. SINR for NOMA users can be expressed as:

$$SINR_{u,s}^r = a_{u,s}^r \left(\frac{G_{u,s}^r P_{u,s}^r}{\sum_{u \neq v, v \in U} G_{v,s}^r P_{v,s}^r a_{v,s}^r + Int_s^r + P_n} \right) \quad (2)$$

where $a_{u,s}^r$ represents a binary allocation variable, $G_{u,s}^r$ signifies the channel gain, and $P_{u,s}^r$ is the transmission power of user u in cell s on resource block r . Furthermore, $G_{v,s}^r$ denotes the channel gain and $P_{v,s}^r$ the transmission power of a NOMA that interferes with the user v on the resource block r . The term P_n refers to the noise power and Int_s^r stands for intercell interference, which is assumed to be the constant interference experienced by the user u from the interfering users of adjacent cell NOMA users. Here, we consider the following.

$$Int_s^r = M \times N_{AC} \times P_{max} \times G_{AC} \quad (3)$$

where M represents the count of NOMA users, N_{AC} denotes the number of neighboring cells, P_{max} represents the peak transmission power and G_{AC} indicates the mean interference channel gain of neighboring cells.

B. Problem Formulation for Energy Efficiency Maximization

The objective is to maximize the EE of the network by identifying users that are multiplexed into the radio resource block and to maximize the guaranteed rate of the network. The achievable EE (b/J) of the proposed network is calculated as the sum of umMTC devices (b/s) divided by the total power consumption (W) of the network. The objective is to maximize the EE while guaranteeing QoS for each user, that is, $Rate_{u,s}^r \geq Rate_{u,s}^{\min}$.

$$\eta^{EE} = \frac{\sum_{u \in U} \sum_{r \in R} \log_2(1 + SINR_{u,s}^r)}{p_c + \sum_{u \in U} \sum_{r \in R} P_{u,s}^r} \quad (4)$$

where p_c is referred circuit power consumption. Given the SINR from Equation (2), the above equation becomes:

$$\eta^{EE} = \frac{\sum_{u \in U} \sum_{r \in R} \log_2(1 + a_{u,s}^r \left(\frac{G_{u,s}^r P_{u,s}^r}{\sum_{u \neq v, v \in U} G_{v,s}^r P_{v,s}^r a_{v,s}^r + Int_s^r + P_n} \right))}{p_c + \sum_{u \in U} \sum_{r \in R} P_{u,s}^r} \quad (5)$$

The EE maximization is given as:

$$P1 : \quad \underset{P_{u,s}^r, a_{u,s}^r \geq 0}{\text{maximize}} \quad \eta^{EE} \quad (6)$$

Furthermore, we have defined the constraints described below.

Minimum SINR constraint: The minimum SINR requirement ensures the system's viability within the network, where $\gamma_{threshold}$ represents the minimum SINR level required for user u_s to meet the Quality of Service (QoS).

$$(C1) : SINR_{u,s}^r \geq \gamma_{threshold} \quad \forall r \in R, \forall u \in U \quad (7)$$

Minimum and Maximum Power constraint stands for the minimum and maximum transmit power constraint of each umMTC UE.

$$(C2) : 0 \leq P_{u,s}^r a_{u,s}^r \leq P_{max}, \quad \forall r \in R, \forall u \in U \quad (8)$$

Binary resource constraint indicates the binary nature of resource block (RB) allocation, i.e. $a_{u,s}^r \in \{0, 1\}$, where $a_{u,s}^r = 1$ signifies that the resource block is allocated to the device.

$$(C3) : a_{u,s}^r \in \{0, 1\}, \quad \forall r \in R \quad (9)$$

Resource Allocation constraint guarantees that each RB is allocated to M umMTC devices

$$(C4) : \sum_{u \in U} a_{u,s}^r \leq M \quad \forall r \in R \quad (10)$$

Non-overlapping device constraint in the same cell require that each umMTC device is allocated to one RB.

$$(C5) : \sum_{r \in R} a_{u,s}^r \leq 1 \quad \forall u \in U \quad (11)$$

C. Problem Formulation for Connectivity Maximization

The objective function maximizes the total connectivity metric (e.g., packet success rate) across all connected devices, subject to resource and connectivity constraints. By solving this optimization problem, we can maximize connectivity for umMTC devices. Therefore, we can express the objective function as follows:

$$P2 : \text{Const} \times \max \left(\sum_{r \in R} \sum_{u \in U} a_{u,s}^r Rate_{u,s}^r \right) \quad (12)$$

subject to : (C1), (C2), (C3), (C4), (C5)

where

$$Const = \left(\frac{BW * IAT * M_{mux}}{PacketSize * BW_{perRB} * A} \right)$$

where BW refers to the bandwidth considered, IAT is the inter-arrival packet time, M_{mux} is the number of multiplexing users in a single resource block, $PacketSize$ is the total number of payload size. BW_{perRB} is the bandwidth of a single resource block, A is the area of consideration for a given simulated scenario.

Since the objective of the Optimization problems referred to in Eq (6) and (12) is to identify users that can be multiplexed in the same radio resource block by minimizing interference between other users in the cell and determine the optimal power allocation to maximize the sum rate of the network and hence the EE and connectivity. This twofold perspective will be the basis of the sub-optimal solution proposed in Section III-C.

III. PROPOSED SOLUTION FOR RESOURCE SCHEDULING

Previously, energy efficiency and connectivity have been defined with constraints that did not account for their interdependencies. For resource scheduling, it is now recognized that these factors are interlinked, and optimizing one in isolation may not yield the most effective overall solution. Optimizing energy efficiency and connectivity simultaneously complicates the optimization problem due to the inclusion of binary allocation variables and the non-convex characteristics of the signal-to-interference-plus-noise ratio (SINR) due to NOMA's non-orthogonality. Consequently, finding the optimal solution is extremely challenging. Inevitably, trade-offs and compromises must be made; for example, lowering energy consumption could lead to outages, whereas improving connectivity might require more energy.

To address this issue holistically, we propose a sub-optimal algorithm that decomposes the problem into two distinct components: Resource Scheduling and Power Allocation. This algorithm offers a viable solution to the optimization challenge. Initially, an interference power value is allocated to each user. Subsequently, each UE is scheduled, and power is distributed to UEs, hence mitigating interference at the target receiver. Power allocation is predicated on channel assignment and aims

to minimize interference among users within the same cell and those sharing the same radio resource block, considering an SINR threshold $\gamma_{threshold}$ and an optimal distribution of power and resources. We ensure that each user is assigned a resource only once within a radio frame and that the power remains within the range of zero to P_{max} .

A. Resource Scheduling

The energy efficiency (EE) of each UE is influenced by the presence of other NOMA users sharing the same resource block (RB) and by inter-cell interference.

In scenarios involving multiple users and RBs, it is imperative to optimize the EE system by pairing users and allocating RBs. The channel coefficients for NOMA users assigned to the same RB should be systematically arranged in ascending order, reflecting the proximity of users with stronger channel gains to those with weaker channel gains. The power allocated to all users should be constant for performing scheduling. We set the power of all users $P_{u,s}^r = P_{v,s}^r = P$, where P is equal to a maximum power limit, $P = P_{max}$.

$$G_1^r \leq G_2^r \leq \dots \leq G_m^r \leq \dots \leq G_M^r \quad (13)$$

Following channel assignment, power allocation is performed. This involves reformulating the optimization problem for each resource block, as interference from adjacent cells is negligible. However, it is necessary to incorporate the interference generated by NOMA users within a given resource block. Let M denote the set of NOMA users per resource block per cell and u represent a user in cell s that belongs to M . The optimization objective is formulated as follows.

$$\max \left(\left(\frac{\sum_{u \in U} \sum_{r \in R} \log_2(1 + SINR_{u,s}^r)}{p_c + \sum_{r \in R} \sum_{u \in U} P_{u,s}^r} \right) + \left(\sum_{r \in R} \sum_{u \in U} a_{u,s}^r Rate_{u,s}^r \right) \right) \quad (14)$$

This can be simplified as,

$$\max \left(\frac{\sum_{u \in U} \sum_{r \in R} \log_2(1 + a_{u,s}^r (\frac{G_{u,s}^r P_{u,s}^r}{\sum_{u \neq v, v \in M} G_{v,s}^r P_{v,s}^r a_{v,s}^r + Int_s^r + P_n}))}{p_c + \sum_{u \in U} \sum_{r \in R} P_{u,s}^r} \right) \quad (15)$$

The goal is to optimize the rate while minimizing energy consumption, leading to improved energy efficiency. The user must satisfy the minimum rate requirement to ensure connectivity. The

$$SINR_{u,s}^r \geq \gamma_{threshold_{s,lim}} \quad (16)$$

$$0 \leq P_{u,s}^r \leq P_{max}, \forall s \in S. \quad (17)$$

B. Power Allocation

Once the scheduling is completed, we can optimize the power for the users. Since there is no interference between users on different resources, we can rewrite the optimization problem as follows:

$$a_{u,s}^r \left(\frac{G_u^r P_{u,s}^r}{\sum_{u \neq v, v \in M} G_v^r P_{v,s}^r a_{v,s}^r + Int^r + P_n} \right) \geq \gamma_{threshold} \quad (18)$$

$$a_{u,s}^r G_{u,s}^r P_{u,s}^r \geq \gamma_{threshold} * \left(\sum_{u \neq v, v \in M} a_{v,s}^r G_{v,s}^r P_{v,s}^r \right) + \quad (19)$$

$$a_{u,s}^r G_{u,s}^r P_{u,s}^r - \gamma_{threshold} * \left(\sum_{u \neq v, v \in M} a_{v,s}^r G_{v,s}^r P_{v,s}^r \right) \geq \gamma_{threshold} * (Int_s^r + P_n) \quad (20)$$

After presenting the problem decomposition as in resource scheduling (Section III-A) and simplified SINR equations to meet the minimum rate constraint as presented in Eq.(18 - 20) in power allocation (Section III-B). The suboptimal EERS-MTCN algorithm is presented in Section III-C.

C. Proposed Energy Efficient Resource Scheduling Algorithm

The proposed EERS-MTCN algorithm aims to maximize the energy efficiency (EE) to support high connection density while minimizing the interference on the user u caused by NOMA users sharing the same resource block r within a cell. We assumed a constant interference from neighboring cells. The proposed scheduling algorithm has two steps (see Algo.1), i.e. UEs/Resource Scheduling and power allocation.

Two main steps of EERS-MTCN are depicted in Algorithm1. First, users are divided into various sectors according to their channel gains within a cell. The interference experienced by users within a cell is determined concerning adjacent cells (i.e., constant interference in the non-cooperative scenario considered here). Then, users with the minimum interference are paired together in a NOMA resource block. To maximize connectivity, a maximum number of 8 users per RB is packed to any given resource block so that all devices in that RB achieve the minimum data rate requirement.

In the subsequent step following resource allocation, power is assigned to NOMA users within a given resource block (RB) to maximize the network's energy efficiency (EE). This allocation must ensure that the interference caused to other NOMA users within the same RB is minimized, while still satisfying the minimum rate constraint and maintaining high network EE.

IV. PERFORMANCE EVALUATION AND DISCUSSION

A. Simulation Setup

The performance of the proposed framework is evaluated through Monte Carlo simulation over 1000 independent UE deployments. A reference cell with a radius of 500 m is considered, while the effect of adjacent cells is represented through the fixed inter-cell interference model. The number of active umMTC devices varies from 50 to 400, and the maximum NOMA multiplexing order is set to eight users per RB. All UEs are uniformly distributed within the reference cell and transmit with a maximum uplink power of 23 dBm. The remaining system parameters follow the 3GPP standard [14], the IMT-Advanced Report ITU-R [15], and are listed in Table I.

B. Results and Discussion

In our simulations, the total number of ultra-massive machine-type communication (umMTC) devices ranges from 50 to 400, with a maximum of $M_{mux} = 8$ umMTC devices per resource block. We consider a scenario with imperfect successive interference cancellation (SIC), where users experience interference from other users paired to the same resource.

Algo. 1 : Proposed Algorithm for Sustainable Scheduling

Input: $BW, m_{NOMA}, P_{max}, G_{u,s}^r, P_n, Area, ISD$
 User Equipment creation $\triangleright$ Set $u_s \leftarrow G_{u,s}^r$
 Initialize $P_{u,s}^r = P_{max}, RtU_m = RtU_{min}$
 Compute CSI and $SINR$ as in (Eq.2)

Step 1: Resource Allocation

```

while numSector do
  for Usr in rbNOMA do
    Calculate  $Int_s^r = M * N_{AC} * P_{max} * G_{AC}$ 
    calculate  $RtU_m = Sum(RtU_m)$ 
    calculate  $EEU_m = Sum(EEU_m)$ 
  end for
  best combination  $IntU_{min}^m, RtU_{max}^m, EEU_{max}^m$ 
  Assign RB  $IntU_{min}^m, RtU_{max}^m, EEU_{max}^m$ 
end while

```

Input: $IntU_{min}^m, RtU_{max}^m, EEU_{max}^m$
 for chGain in $IntU_{min}^m$ do
 for U_{sr} in rbNOMA do
 calculate $IntU_m = Sum(IntU_m)$
 calculate $RtU_m = Sum(RtU_m)$
 calculate $EEU_m = Sum(EEU_m)$
 end for
 best combination $IntU_{min}^m, RtU_{max}^m, EEU_{max}^m$
 end for
 assign RB $IntU_{min}^m, RtU_{max}^m, EEU_{max}^m$
 return $IntU_{min}^m, RtU_{max}^m, EEU_{max}^m$

Input: $IntU_{min}^m, RtU_{max}^m, EEU_{max}^m, P_{u,s}^r = PowMat$

Step 2: Power Allocation

```

while UsrChGain do
  for Power in PowMat do
    for Usr in rbNOMA do
      if  $SINR \geq SINR_{min}$  then
        calculate  $RtU_m = Sum(RtU_m)$ 
        calculate  $EEU_m = Sum(EEU_m)$ 
      end if
    end for
  end for
  calculate  $Pwr_{alloc}, RtU_{max}^m, EEU_{max}^m$ 
  return  $RtU_{max}^m, EEU_{max}^m, Pwr_{alloc}$ 
end while

```

The following results validate the effectiveness of the proposed techniques.

C. Results and Discussion

The proposed EERS-MTCN algorithm is compared with exhaustive-search pairing NOMA (ESAP-NOMA) and random allocation and pairing NOMA (RAAP-NOMA). The results should be interpreted in terms of both achievable EE and computational cost. Fig. 2 presents the achievable EE as the maximum number of NOMA users multiplexed on each RB increases. For all three algorithms, EE decreases with the NOMA multiplexing order. Although serving more users on the same RB increases spectral reuse and supports additional simultaneous connections, it also strengthens intra-RB interference and increases the power required to satisfy the minimum-SINR constraints. Consequently, the additional throughput does not fully compensate for the higher interference and power

Table I: Simulation Parameters

Parameters	Value
Transmit Power of Base Station, and UE (dBm)	46, 23
Carrier Frequency (MHz)	700
Receiver Thermal Noise Density (dBm/Hz)	-174
No. of UE per cell	10
Interference Margin (dB)	0
Channel Model	Extended Okumura Hata Model
Noise Figures of Base Station, UE (dB)	7, 5
Effective Noise Power (dBm)	Thermal Noise+ Noise Figure + $N_{UE} + 10 \log(\text{Receiver Sensitivity})$
Occupied System Bandwidth (kHz)	180
Radius of Cell (m)	500
Heights of the base station, UE (m)	25, 1.5
Antenna element gain of BS, UE (dBi)	8, 0
No. of antenna support per UE	1
Maximum Coupling Loss (MCL) (dB)	Trx Pwr + SINR

consumption at large multiplexing orders.

ESAP-NOMA achieves the highest EE because it exhaustively evaluates the available user-pairing and power-allocation combinations. The proposed EERS-MTCN consistently achieves the second-highest EE and outperforms RAAP-NOMA across the evaluated multiplexing levels. At the maximum evaluated connection density, EERS-MTCN achieves approximately 27% higher EE than RAAP-NOMA while remaining approximately 16% below ESAP-NOMA. This result confirms that EERS-MTCN approaches the exhaustive-search benchmark while avoiding its high computational burden. Table II compares the algorithms across connection densities of 150, 200, 250, and 300 users/m². Over this evaluated range, aggregate EE increases with connection density because the growth in successfully scheduled throughput exceeds the corresponding increase in total transmit power. Nevertheless, for each fixed density, increasing the number of NOMA users per RB reduces EE because of the stronger intra-RB interference.

Relative to RAAP-NOMA, EERS-MTCN improves EE by approximately 26.98%, 31.76%, 35.48%, and 42.84% at connection densities of 300, 250, 200, and 150 users/m², respectively. Relative to ESAP-NOMA, EERS-MTCN is lower by approximately 15.95%, 7.93%, 17.09%, and 17.58% at the same operating points. These results show that EERS-

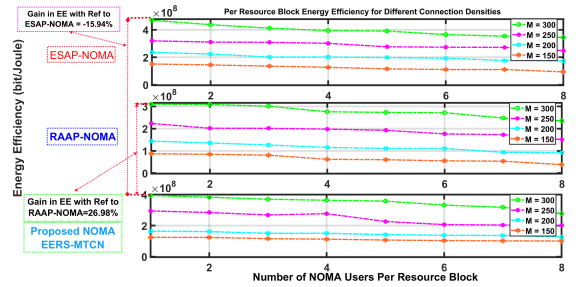


Figure 2: System Energy Efficiency with the different maximum number of umMTC devices M_{max} per resource block for the baseline ESAP-NOMA, baseline RAAP-NOMA and the proposed EERS-MTCN.

Table II: Energy Efficiency comparison among different algorithms in terms of connection density

Connection Density users/m^2	Proposed EERS-MTCN's EE (bits/J)	ESAP-NOMA's EE(bits/J)	RAAP-NOMA's EE(bits/J)
300	3.9375e+08	4.6847e+08	3.1008e+08
250	2.9413e+08	3.1947e+08	2.2324e+08
200	1.9513e+08	2.3534e+08	1.4403e+08
150	1.2429e+08	1.5080e+08	8.7012e+07

MTCN provides a robust performance advantage over random scheduling while remaining reasonably close to exhaustive search. The execution-time comparison in Fig. 3 demonstrates the practical advantage of EERS-MTCN. ESAP-NOMA requires approximately an order of magnitude more computation because it evaluates all feasible user-pairing combinations. RAAP-NOMA is computationally inexpensive but does not account for interference, QoS, or EE when forming NOMA groups. EERS-MTCN therefore provides the most favorable compromise between EE and execution time among the evaluated practical schemes.

Furthermore, it has been observed that when connection density and NOMA users per resource block are the same, the EE value obtained by the three algorithms decreases as P_{max} increases. This occurs because the total communication power consumption increases when the allocated transmit power of umMTC devices increases, leading to a reduction in the EE of umMTC devices. Therefore, compared to the baseline algorithms, the proposed ERRS-MTCN algorithm demonstrates superior performance in terms of effectiveness and energy efficiency.

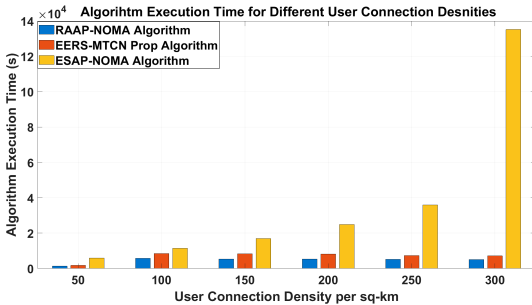


Figure 3: Algorithm execution time for baseline ESAP-NOMA, baseline RAAP-NOMA, and the proposed EERS-MTCN.

The superiority of our proposed scheme over the baseline schemes stems from the following shortcomings. Although the ESAP NOMA algorithm explores all possible user pairs in the system to identify the optimal pairing that maximizes performance metrics, it is very computationally intensive, and the time complexity is very high, as can be seen in Fig.(3). It finds the pair for minimum interference and maximum EE. In contrast, the RAAP-NOMA algorithm randomly pairs NOMA users, which generates a diverse range of user pairs but overlooks all the QoS constraints and pair without looking into interference and EE considerations. On the other hand, our proposed scheme minimizes interference and EE and specifically pairs devices with favorable channel gain and minimal interference to enhance the EE of the system.

V. CONCLUSION

This paper presented EERS-MTCN, an energy-efficient scheduling and power-allocation framework for dense uplink NOMA-based umMTC networks under a non-cooperative multi-cell model with fixed inter-cell interference. The proposed algorithm jointly performs channel-aware user grouping, interference-aware resource allocation, and power optimization to improve energy efficiency while satisfying QoS constraints. Simulation results show that EERS-MTCN consistently outperforms RAAP-NOMA, achieving up to 27% higher EE at a connection density of 300 users/m², while remaining within approximately 16% of the exhaustive-search ESAP-NOMA benchmark. Considering that ESAP-NOMA incurs nearly an order of magnitude higher execution time, EERS-MTCN offers a practical trade-off between energy efficiency and computational complexity for dense mMTC deployments. Future work will investigate dynamic inter-cell interference coordination, finite-blocklength transmission, and learning-based adaptive scheduling for time-varying wireless environments.

ACKNOWLEDGMENT

This project has received funding from the European Union's Horizon Europe Research Program under grant agreement No. 101058505 - 5G-TIMBER, and from the Estonian Education and Youth Board ÕÜ11 "AIoT*5G - Artificial intelligence, edge computing and IoT solutions in distributed systems".

REFERENCES

- [1] Nurul Huda Mahmood *et al.*, "Six Key Features of Machine Type Communication in 6G," in *2nd 6G Wireless Summit*, 2020, pp. 1–5.
- [2] N. Xia, H.-H. Chen, and C.-S. Yang, "Radio resource management in machine-to-machine communications—a survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 1, pp. 791–828, 2017.
- [3] Zhang, L. *et al.*, "6G Visions: Mobile Ultra-Broadband, Super IoT, and Artificial Intelligence," *China Communications*, vol. 16, pp. 1–14, 2019.
- [4] Wei, Zhiqiang *et al.*, "On the performance gain of noma over oma in uplink communication systems," *IEEE Transactions on Communications*, vol. 68, no. 1, pp. 536–568, 2020.
- [5] D. Zhai and J. Du, "Spectrum efficient resource management for multi-carrier-based noma networks: A graph-based method," *IEEE Wireless Communications Letters*, vol. 7, no. 3, pp. 388–391, 2017.
- [6] Mostafa Ahmed Elhany *et al.*, "Connectivity Maximization for Narrowband IoT Systems with NOMA," *IEEE Transactions on Wireless Communications*, vol. 18, no. 10, pp. 4708–4722, 2019.
- [7] F. Fang, H. Zhang, J. Cheng, S. Roy, and V. C. Leung, "Joint user scheduling and power allocation optimization for energy-efficient noma systems with imperfect esi," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 12, pp. 2874–2885, 2017.
- [8] Han Shujun *et al.*, "Energy-Efficient Short Packet Communications for Uplink Noma-Based Massive MTC Networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 12, pp. 12066–12078, 2019.
- [9] Khan, Wali Ullah *et al.*, "Efficient Power Allocation for NOMA-enabled IoT Networks in 6G era," *Physical Communication*, vol. 39, p. 101043, 2020.
- [10] D. Zhai, R. Zhang, L. Cai, B. Li, and Y. Jiang, "Energy-efficient user scheduling and power allocation for noma-based wireless networks with massive iot devices," *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1857–1868, 2018.
- [11] O. Elgarhy, L. Reggiani, M. M. Alam, A. Zoha, R. Ahmad, and A. Kuusik, "Energy efficiency and latency optimization for iot urllc and mmte use cases," *IEEE Access*, vol. 12, pp. 23 132–23 148, 2024.
- [12] Malik, Hassan *et al.*, "Radio Resource Management in NB-IoT Systems: Empowered by Interference Prediction and Flexible Duplexing," *IEEE Network*, vol. 34, no. 1, pp. 144–151, 2020.
- [13] Latif, Aamir *et al.*, "Inter O-DUs Coordination for Scheduling of Massive Users in O-RAN," in *EuCNC-6G Summit*, 2023, pp. 1–6.
- [14] ETSI, "5G, Study on Channel Model for Frequencies from 0.5 to 100 GHz 3GPP TR 38.901 Version 16.1.0 Release 16," ETSI, Tech. Rep., 2017.
- [15] ITU, "Report ITU-R M.2412-0 Guidelines for Evaluation of Radio Interface Technologies for IMT-2020," ITU-R, Tech. Rep., 2017.

Appendix 4

Publication (IV):

Latif, Aamir and Elgarhy, Osama and Mahmood, Nurul Huda and Le Moullec, Yannick and Alam, Muhammad Mahtab, "Sustainable Massive Connectivity Maximization Scheduling Scheme for 6G," *IEEE Transactions on Green Communications and Networking*, pp. 2075–2090, January 2026

Sustainable Massive Connectivity Maximization Scheduling Scheme for 6G

Aamir Latif¹, Graduate Student Member, IEEE, Osama Elgarhy², Member, IEEE, Nurul Huda Mahmood³, Senior Member, IEEE, Yannick Le Moullec⁴, Senior Member, IEEE, and Muhammad Mahtab Alam⁵, Senior Member, IEEE

Abstract—This paper proposes a scalable and energy-efficient scheduling framework tailored for massive machine-type communications (mMTC) in future 6G networks. Motivated by the dual challenges of massive connectivity and strict energy constraints, we introduce a new cooperative multi-cell uplink scheme that jointly optimizes user scheduling, power allocation, and interference management under finite block length (FBL) constraints. While prior studies have explored energy efficiency (EE) or connectivity separately, they often neglect FBL effects and lack scalable, interference-aware coordination across multi-cell NOMA deployments. Leveraging non-orthogonal multiple access (NOMA), our framework addresses both intra- and inter-cell interference through a multi-objective optimization problem that simultaneously maximizes EE and connectivity. To solve the problem, we develop the Sustainable and Reliable Scheduling for Massive Machine Type Communication using NOMA (SRS-MTCN) algorithm, integrating channel-aware user clustering, interference-aware resource allocation, and iterative power optimization. Simulation results demonstrate that the proposed approach performs significantly better than state-of-the-art baselines in dense deployments, achieving up to 35–58% EE improvement in medium-density regimes (16–32 devices/m²), while ensuring robust scalability and reliability in ultra-dense scenarios exceeding 40 devices/m². These results illustrate its suitability for sustainable 6G mMTC networks operating under FBL and energy constraints.

Index Terms—Sixth-generation (6G), energy efficiency (EE), finite blocklength (FBL), connectivity, massive machine-type communication devices (mMTC), non-orthogonal multiple access (NOMA), resource scheduling, uplink multi-cell.

I. INTRODUCTION

THE forthcoming sixth-generation (6G)¹ wireless networks are envisioned to support an unprecedented number of connected devices projected to exceed 10⁸ devices/km² driven by massive Machine Type Communication (mMTC)

Received 16 September 2025; revised 25 December 2025; accepted 22 January 2026. Date of publication 30 January 2026; date of current version 12 February 2026. This work was supported in part by the European Union's Horizon 2020 Research and Innovation Program under Grant 101058505 '5G-TIMBER'; in part by the Project "Increasing the Knowledge Intensity of Ida-Viru Entrepreneurship" co-funded by the European Union; in part by the NATO-SPS G7699 PROTECT Project; and in part by Telia Eesti AS through TalTech Development Fund. The editor coordinating the review of this article was K. Dev. (Corresponding author: Aamir Latif.)

Aamir Latif, Osama Elgarhy, Yannick Le Moullec, and Muhammad Mahtab Alam are with Tallinn University of Technology, 19086 Tallinn, Estonia (e-mail: aamir.latif@taltech.ee; osama.elgarhy@taltech.ee; yannick.lemoullec@taltech.ee; muhammad.alam@taltech.ee).

Nurul Huda Mahmood is with the 6G Flagship Program, University of Oulu, 90014 Oulu, Finland (e-mail: nurulhuda.mahmood@oulu.fi).

Digital Object Identifier 10.1109/TGCN.2026.3659720

¹Table 1 provides a list of the commonly used acronyms in this paper.

2473-2400 © 2026 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission.

See <https://www.ieee.org/publications/rights/index.html> for more information.

Authorized licensed use limited to: Tallinn University of Technology. Downloaded on May 11, 2026 at 12:52:58 UTC from IEEE Xplore. Restrictions apply.

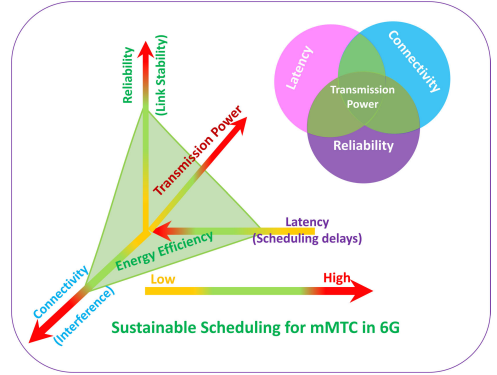


Fig. 1. Challenges in Sustainable scheduling, highlighting the optimal transmission power region for energy-efficient connectivity.

use cases in smart cities, industrial Internet of Things (IoT), and autonomous systems [1], [2], [3]. These devices are characterized by stringent energy, latency, and complexity constraints, which pose unique challenges for sustainable resource scheduling in dense environments [4]. A central goal is the design of *sustainable scheduling* techniques that jointly optimize EE and *massive connectivity*, while respecting power budgets and reliability constraints. Optimizing EE or connectivity in isolation is insufficient [5], [6], i.e., focusing solely on EE can limit access, whereas prioritizing connectivity without energy considerations can destabilize the network. Given the above, a comprehensive framework is needed to balance these objectives and ensure scalable energy-efficient access, which are two key issues in 6G networks [7].

This challenge is further intensified at the radio access network (RAN) layer, which accounts for over 70% of the total energy consumption in modern cellular networks [8]. As also depicted later in Fig. 1, sustainable scheduling in ultra-dense mMTC networks requires a balance among conflicting objectives: maximizing device connectivity, preserving reliable SINR, and minimizing energy consumption. These interlinked trade-offs make the problem inherently non-convex, where optimizing one metric often compromises another. However, mMTC devices typically transmit short packets, where this assumption does not hold.

Non-orthogonal multiple access (NOMA) has emerged as a promising solution, enabling simultaneous user transmissions

over shared time-frequency resources and thus improving spectral efficiency [9], [10]. However, in ultra-dense mMTC scenarios, even NOMA's performance is fundamentally limited by the short-packet transmissions and the constraints of finite block length (FBL) coding [11], [12]. Moreover, co-channel interference (CCI) in dense deployments further degrades FBL rates, which stresses the need for interference-aware, inter-cell scheduling [13]. This leads to motivation to carry out this research, i.e., how can multi-cell NOMA systems leverage sustainable scheduling to achieve massive connectivity and EE through inter- and intra-cell interference minimization under FBL constraints. To address this, we propose a cooperative multi-cell NOMA framework that jointly maximizes the EE-connectivity trade-off by minimizing interference while accounting for FBL limits. Our solution aligns with the efforts of 3GPP Release 18 [14], and bridges theoretical limits with practical deployment of 6G mMTC.

Existing studies primarily address subsets of key objectives such as EE, FBL transmission, connectivity, or interference coordination, but lack integrated multi-cell solutions. Energy-efficient resource allocation has been explored extensively, though often within simplified or limited scenarios. Rabee and Barhumi [15] used multi-agent deep reinforcement learning (DRL) for EE optimization in multi-cell uplink NOMA, achieving approximately 25% EE improvement but validated only up to 30 users per cell. Rashid et al. [16] investigated MC-NOMA in heterogeneous small-cell networks, reporting around 35% EE and 20% throughput enhancements, with scalability limited to 20 users per cell. Coordinated multi-point (CoMP) approaches, such as the method proposed by Jameel et al. [17], used Markov Decision Processes (MDP) and achieved EE improvements of (15–20)%, but lacked large-scale connectivity analysis. Recent studies have integrated EE with FBL constraints. For example, Yang et al. [18] explored reflecting intelligent surfaces (RIS)-assisted multi-access edge computing (MEC) networks under FBL, optimizing EE but not explicitly addressing scalability or multi-cell interference. Liu et al. [19] studied cell-free, full-duplex industrial IoT networks, demonstrating EE improvements but not explicitly considering extensive connectivity or inter-cell interference.

Specifically, Gao et al. [20] provide an FBL information-theoretic connectivity analysis in a single-cell MIMO un-sourced random access (URA) setting, but do not optimize EE nor address scheduling under inter-cell coupling. Zhang et al. [21] focused on single-cell massive URLLC scenarios using adaptive FBL-based multi-agent DRL for delay optimization, omitting explicit EE consideration. FBL-based reliability was the focus in works such as Zheng et al. [22], analyzing uplink NOMA reliability without explicit EE or interference management, and Hu et al. [23], who optimized EE for intelligent reflective surface (IRS)-assisted uncrewed aerial vehicle (UAV) communications under short-packet conditions, excluding scalability analysis. Cao et al. [13] proposed AoI-based optimization for multi-connectivity wireless networked control systems, achieving around 30% EE improvements, without explicitly addressing extensive connectivity or multi-cell interference. Tu and Lee [24] analyzed multi-hop MIMO full-duplex relay networks for short-packet

TABLE I
ACRONYMS

3GPP	3rd generation partnership project.
5G	Fifth-generation.
6G	Six-generation.
CCI	Co-Channel Interference.
CoMP	Coordinated multipoint.
CF-MIMO	Cell-Free Massive Input Massive Output.
DDQN	Double Deep Q-Network.
DRL	Deep Reinforcement Learning.
FBL	Finite Block Length.
IFBL	Infinite Block Length.
ICI	Inter Cell Interference.
IRS	Intelligent Reflective Surface.
IoT	Internet of Things.
MDP	Markov Decision Processes.
MEC	Multi-access Edge Computing.
mMTC	Massive Machine Type Communications.
NOMA	Non-Orthogonal Multiple Access.
QoS	Quality of Service.
RIS	Reflecting Intelligent Surface.
SCA	Successive Convex Approximation.
SIC	Successive Interference Cancellation.
SISO	Single Input Single Output.
STAR-RIS	Simultaneously Transmitting and Reflecting Reconfigurable Intelligent Surfaces.
UAV	Unmanned Aerial Vehicle.
UE	User Equipment.
URA	Un-sourced Random Access.
URLLC	Ultra-Reliable Low-Latency Communication.

URLLC, prioritizing reliability without EE or connectivity optimization.

In UAV-assisted multi-cell NOMA scenarios, Thapliyal et al. [25] combined stochastic geometry and FBL outage analysis, reporting EE gains of approximately (20 – 25)%, without explicit connectivity optimization. Huang et al. [26] addressed EE optimization in multi-cell cell-free massive MIMO through successive convex approximation (SCA) methods for joint power allocation and AP-user association, omitting detailed connectivity and interference analysis. Several works incorporate FBL modeling and EE-centric optimization in *single-cell* settings, e.g., joint power/throughput or delay-aware formulations for mMTC/URLLC and NOMA [27], [28], [29], [30], [31]; however, their designs are limited to intra-cell scheduling and do not model explicit inter-cell interference coordination.

TABLE II
COMPARISON OF SELECTED RELATED WORKS ON SUSTAINABLE SCHEDULING UNDER FBL IN mMTC NETWORKS FOR 6G

Reference	Environment	Supported Technique	EE	Connectivity	FBL	ICI/CCI-Aware UE Scheduling
[20]	Single-cell MIMO URA	FBL information-theoretic analysis	×	✓	✓	×
[21]	Single-cell massive URLLC	Multi-Agent DQN for delay minimization	×	✓	✓	Intra-cell only (✓)
[27]	Single-Cell Multi-UAV aided mMTC/URLLC	Joint mMTC-URLLC power optimization	✓	×	✓	Intra-cell only (✓)
[28]	Single-cell Uplink NOMA	Joint optimization delay-aware throughput	✓	×	✓	Intra-cell only (✓)
[29]	Single-cell IoT Uplink	Power and Block length optimization	✓	✓	✓	×
[30]	Single-cell MU-MIMO and STAR-RIS	Fractional-Programming framework to maximize Sum EE	✓	×	✓	×
[31]	Single-Cell Heterogeneous URLLC/eMBB	Multi-agent DRL Joint Power and Sub-channel allocation	✓	×	✓	Intra-cell only (✓)
[25]	Multi-cell UAV-assisted NOMA	FBL outage and Stochastic-geometry	✓	×	✓	Intra-cell only (✓)
[26]	Cell-Free Massive Input Massive Output (CF-MIMO) URLLC in Multi-Cell Networks	SCA Joint power-allocation and AP-user association	✓	×	✓	Intra-cell only (✓)
[32]	Multi-Cell Ultra-Dense Small Cells	Double-DQN Optimization for RB assignment and power control	✓	✓	×	Intra-cell only (✓)
[33]	Multi-Cell NOMA	Cooperative scheduler for NB-IoT	×	✓	×	Intra-cell and Inter-cell (✓)
Our Work	Multi-Cell NOMA mMTC	Joint EE-Connectivity maximization through Interference minimization	✓	✓	✓	Intra-cell and Inter-cell (✓)

Wang et al. [28] addressed delay-aware throughput and EE under FBL for uplink NOMA without explicit large-scale connectivity or interference management. Zhao et al. [29] optimized single-cell IoT uplink networks focusing on joint EE and connectivity, without interference coordination. Soleymani et al. [30] optimized EE in single-cell MU-MIMO with simultaneously transmitting and reflecting reconfigurable intelligent surfaces (STAR-RIS) under FBL constraints, but did not explicitly consider connectivity or interference. Tran et al. [31] employed multi-agent DRL for EE optimization in single-cell heterogeneous URLLC/eMBB scenarios without addressing scalability or interference explicitly. Zhao et al. [32] examined EE and connectivity in multi-cell ultra-dense small cells using double deep Q-network (DDQN) optimization but omitted FBL constraints and inter-cell interference.

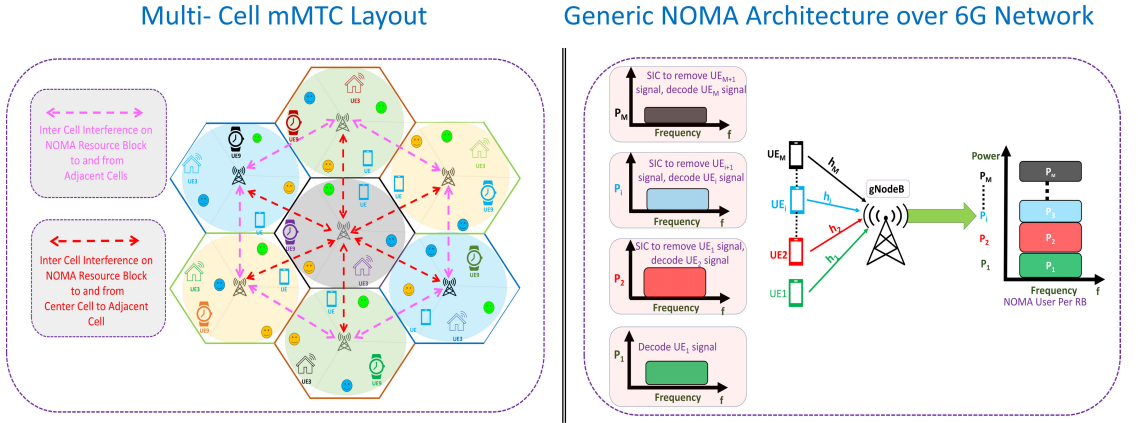
Table II highlights selected works among the above presented references that offer partial progress toward sustainable scheduling, but that do not simultaneously address sustainable scheduling, FBL constraints, and inter-cell interference (ICI) coordination in a scalable multi-cell NOMA framework.

Unlike conventional single-cell NOMA schemes, our proposed framework unifies these critical dimensions by jointly maximizing EE and connectivity through interference minimization within a cooperative multi-cell NOMA architecture. This approach addresses the scalability and sustainability requirements of future 6G mMTC networks. Moreover, while prior works have examined individual aspects such as EE, massive connectivity, or interference management, they often neglect the combined impact of FBL constraints and decoding error probability, which is essential for practical short-packet mMTC systems. Motivated by this gap, we also explicitly incorporate FBL conditions and their associated error probabilities into our optimization framework.

The key contributions of this paper are summarized as follows:

- We propose a sustainable resource scheduling framework for multi-cell uplink NOMA networks that jointly maximizes EE and massive connectivity via interference management. Unlike conventional isolated methods, our model incorporates FBL constraints and decoding error probabilities into power control, user pairing, and interference coordination-ensuring sustainable scalability and mitigating the traditional EE-connectivity trade-off.
- We formulate a multi-objective optimization problem that addresses the non-convex EE-connectivity trade-off through interference minimization. The problem accounts for both intra- and inter-cell interference, FBL constraints. Devices are dynamically grouped into NOMA sectors based on channel conditions, enabling adaptive user pairing that significantly improves scalability-supporting over devices per cell and achieving nearly three-fold gains over prior cooperative NOMA approaches [33].
- We develop the Sustainable and Reliable Scheduling for Massive Machine Type Communication using NOMA (SRS-MTCN) algorithm, comprising three stages: 1) channel-aware clustering, 2) interference-aware resource allocation, and 3) iterative power optimization with multi-cell interference coordination. The algorithm ensures reliable connectivity by explicitly incorporating decoding error, quality of service (QoS), and minimum rate constraints under FBL. Simulation results confirm the algorithm's robustness and superior performance compared to benchmarks, namely random allocation and pairing-based NOMA (RAAP-NOMA), and cooperative NOMA scheduling [33].

The remainder of this paper is structured as follows: Section II introduces the system model and formally defines



Multi-Cell NOMA Cooperative Scenario

Fig. 2. Left: Multi-cell cooperative scenario system model layout for mMTC devices in 6G. This layout illustrates the inter-cell interference on NOMA RBs from adjacent cells to the center cell and the adjacent cells. Right: Generic NOMA architecture over a 6G network. The NOMA system operates by superimposing multiple user equipment (UEs) within a single RB, utilizing distinct power coefficients.

the optimization problem. Section IV details the proposed joint EE-connectivity optimization framework and presents the SRS-MTCN algorithm; Section IV-D.1 analyzes the computational complexity of the algorithm. Simulation results and discussions are provided in Section V. Finally, Section VI concludes the paper and outlines potential future research directions.

II. SYSTEM MODEL

Consider a SISO multi-cell uplink network as illustrated in Fig. 2 (Left) with S cells, where each cell $s \in S$ serves U_s mMTC devices under EE and scalability constraints. Each UE has a single transmit antenna while each gNB is equipped with L receive antennas. The network prioritizes scalable, energy-efficient resource scheduling through dynamic power allocation and interference management between cells. Let R denote the set of orthogonal RBs. Each cell s may assign multiple distinct RBs from the common pool $\mathcal{R}$. The system operates on the principle of NOMA, as depicted in Fig. 2 (Right), wherein multiple mMTC devices share a single RB and are multiplexed using power-domain NOMA. The maximum number of mMTC devices multiplexed on a single RB is denoted by M , subject to the constraint $M \leq N$, the total number of users served by the network given as: $N = \sum_{s \in S} U_s$.

For ease of reference, the key symbols and notations used throughout the system model are summarized in Table III.

As illustrated in Fig. 1, achieving sustainable scheduling in ultra-dense mMTC networks entails balancing inherently conflicting objectives: maximizing device connectivity, ensuring reliable signal-to-interference-plus-noise ratio (SINR), and minimizing overall energy consumption. These interdependent performance metrics introduce a complex, non-convex optimization landscape, wherein improving one criterion may

inadvertently degrade another. Traditional communication models based on Shannon's capacity assume infinite block length (IFBL) and negligible error probabilities; these assumptions are invalid in mMTC scenarios. In contrast, mMTC devices predominantly operate with short-packet transmissions under stringent latency and reliability constraints, necessitating the adoption of FBL models to capture practical system behavior accurately. To characterize the complex relations between the transmission rate (R), block length (L) and decoding error probability (DEP) (ϵ) as a function of the SINR (γ), the maximum achievable transmission rate $R_s(L, \epsilon, \gamma)$ in SPC is expressed as [11] and [34].

$$R_s(L, \epsilon, \gamma) = \ln(1 + \gamma) - \frac{Q^{-1}(\epsilon)}{\sqrt{L}} \sqrt{1 - \frac{1}{(1 + \gamma)^2}} + \mathcal{O}\left(\frac{\ln L}{L}\right), \quad (1)$$

where L represents the finite block length, and γ denotes the SINR; $Q^{-1}(\cdot)$ is the inverse of the Gaussian Q-function with $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{t^2}{2}} dt$; ϵ is the DEP, and can be derived as $\epsilon = Q(f(\gamma, L, R_s(L, \epsilon)))$ and $f(\gamma, N, R) = \frac{\ln 2 \sqrt{N}}{\sqrt{V(\gamma)}} (\log_2(1 + \gamma) - R_s(L, \epsilon))$; moreover, since the transmission rate $R_s(L, \epsilon, \gamma)$ in (1) must not exceed Shannon's capacity, $R_s(L, \epsilon) \leq \log_2(1 + \gamma)$. This formulation reflects that as $L \rightarrow \infty$, $R_s(L, \epsilon, \gamma)$ converges to the Shannon rate. However, for small L in FBL, the second term penalizes the rate, and a non-negligible DEP is inevitable.

The SINR γ for user $u \in U_s$ considering inter-cell interference is defined as:

$$\gamma_{u,s}^r = \frac{a_{u,s}^r |h_{u,s}^r|^2 p_{u,s}^r}{\underbrace{\sum_{\substack{v \in U_s \\ v \neq u}} a_{v,s}^r |h_{v,s}^r|^2 p_{v,s}^r}_{\text{intra-cell}} + \underbrace{\sum_{\substack{t \in S \\ t \neq s}} \sum_{u_t \in U_t} a_{u_t,t}^r |h_{u_t,t}^r|^2 p_{u_t,t}^r + \sigma_N^2}_{\text{inter-cell}}} \quad (2)$$

TABLE III
MATHEMATICAL SYMBOLS

Mathematical Symbols	
$a_{u,s}^r$	Resource allocation indicator of user u in cell s on RB block r .
$a_{v,s}^r$	Resource allocation indicator of co-scheduled user v in cell s on RB r .
$a_{u,t}^r$	Resource allocation indicator of user w in cell t on RB r (interfering cell).
$h_{u,s}^r$	Channel gain of user u to BS s on RB r .
$h_{v,s}^r$	Channel gain of intra-cell co-user v to BS s on RB r .
$h_{u,t}^r$	Channel gain of user u_t in cell t to BS s on RB r .
$p_{u,s}^r$	Transmission power of user u in cell s on RB r .
$p_{v,s}^r$	Transmission power of user v in cell s on RB r .
$p_{u,t}^r$	Interfering power from user w in cell t on RB r .
P_{max}	Maximum allowable transmit power of a user.
$Q^{-1}(\cdot)$	Inverse of Gaussian Q-function.
$R_s(L, \epsilon, \gamma)$	Achievable transmission rate.
R_{min}	Minimum Rate.
$\epsilon_{u,s}^r$	Decoding error probability of user u in cell s on RB r .
ϵ_{th}	Reliability threshold.
$\gamma_{u,s}^r$	Signal-to-Interference-plus-Noise Ratio (SINR) of user u_s on RB block r .
$\gamma_{threshold}$	Minimum SINR threshold.
σ_N^2	Noise Power.

where each channel to gNB s on RB r is a vector $h_{u,s}^r$, and the effective scalar gain is $|h_{u,s}^r|^2$. $p_{u,s}^r$ is the UE transmit power and $a_{u,s}^r \in \{0, 1\}$ the RB indicator, and σ_N^2 is the noise power. To enforce sustainability, we bound $0 \leq p_{u,s}^r \leq P_{max}$, where P_{max} is set according to the UE power-class limits in 3GPP technical specification (e.g. Class 3 = 23 dBm) [35]. The choice of P_{max} trades off coverage and battery life: we select the class that meets cell-edge reliability targets via link-level simulations. Similarly, we cap the number of users per RB at M to avoid overload in ultra-dense mMTC; M is optimized by sweeping its value in system-level simulations to balance spectral efficiency against SIC complexity. Finally, this SINR is substituted into (1) to compute the FBL rate.

A. Multi-Cell Specific Challenges

In contrast to single-cell NOMA systems, the considered multi-cell scenario is intrinsically limited by strong inter-cell interference and tight coupling between neighboring base stations. As seen in the SINR expression in (2), each user's received SINR includes not only intra-cell interference from other users served by the same gNB, but also an explicit inter-cell interference term arising from transmissions in adjacent

cells. This coupling through the SINR denominator implies that resource allocation and power control decisions in one cell directly affect the feasible rates and reliability in other cells. Consequently, an independent per-cell optimization is no longer adequate; instead, a cooperative multi-cell design is required to coordinate scheduling and power allocation across cells and to avoid highly unfair or interference-limited operating points.

III. PROBLEM FORMULATION FOR JOINT OPTIMIZATION FOR MMTC SYSTEM: MULTI-CELL SCENARIO

EE is defined as the ratio of the total achievable FBL rate to the total power consumed across the entire multi-cell network:

$$\eta^{EE} = \frac{\sum_{s \in S} \sum_{u_s \in U_s} \sum_{r \in R} a_{u_s}^r R_s(L, \epsilon_{u_s}^r, \gamma_{u_s}^r)}{\sum_{s \in S} \sum_{u_s \in U_s} \sum_{r \in R} (p_{u_s}^r + p_c)} \quad (3)$$

where p_c is the circuit power. The DEP $\epsilon_{u_s}^r$ is not only a parameter of the rate function $R_s(L, \epsilon, \gamma)$, but also a direct measure of communication reliability. In mMTC, ensuring successful packet delivery is more critical than maximizing raw throughput. Therefore, connectivity $\mathcal{C}$ is naturally defined using $(1 - \epsilon)$, the probability of success of decoding.

Connectivity ($\mathcal{C}$) measures the number of reliably connected devices (devices successfully decoded), defined as:

$$\mathcal{C} = \kappa \cdot \left(\sum_{s \in S} \sum_{u_s \in U_s} \sum_{r \in R} a_{u_s}^r (1 - \epsilon_{u_s}^r) \right) \quad (4)$$

$$\kappa = \left(\frac{\text{BW} \cdot \text{IAT} \cdot M_{\text{mux}}}{\text{Packet Length} \cdot \text{BW}_{\text{per RB}} \cdot \text{Area}} \right)$$

Whereas κ scales massive connectivity with system bandwidth: BW (Hz), packet inter-arrival time per device: IAT(s), NOMA multiplexing (users per RB): M_{mux} , average payload size: PacketLength (bits), bandwidth of one Resource Block(RB): $\text{BW}_{\text{per RB}}$ (Hz), geographical area under consideration: Area (m^2).

A higher η^{EE} represents the network's ability to manage increased traffic while minimizing total power consumption, thus improving sustainability. When users exceed the min-rate R_{min} threshold, the network's scalability increases, enabling it to handle more connections effectively. However, increased user activity can increase interference and power consumption, revealing the challenge of balancing maximizing user massive connectivity and maintaining high EE.

A. Optimization Constraints for Joint EE and Connectivity Maximization

The joint optimization of EE and connectivity in a multi-cell uplink NOMA network involves various constraints necessary to reflect realistic and practical wireless scenarios and to ensure practical and scalable operation in a sustainable 6G uplink NOMA network; the following constraints are categorized under two key objectives: EE and scalability.

1) EE Driven Resource Constraints:

- **FBL Rate Constraint (SC1):** The achievable data rate under FBL constraints must meet the minimum rate constraint $R_{\min}$ that ensures connectivity while minimizing energy wastage for reliable operations.

$$(SC1) : R_s(L, \epsilon_{u_s}^r, \gamma) \geq R_{\min}, \quad \forall u_s \in U_s, r \in R, s \in S. \quad (5)$$

- **Transmit Power Constraint (SC2):** Transmit power levels are bounded to avoid excessive energy consumption and to maintain practical operation and device battery life; each mMTC device has a maximum power limit $P_{\max}$ to prevent energy overuse.

$$(SC2) : 0 \leq p_{u_s, s}^r a_{u_s, s}^r \leq P_{\max} \quad \forall u_s \in U_s, r \in R, s \in S. \quad (6)$$

- **Decoding Error Probability Constraint (SC3):** To ensure the minimum requirements of communication reliability under FBL regimes:

$$(SC3) : \epsilon_{u_s}^r \leq \epsilon_{\text{th}}, \quad \forall u_s \in U_s, r \in R, s \in S. \quad (7)$$

2) Scalability-Driven Resource Constraints:

- **Binary Resource Constraint (SC4):** To facilitate clear and efficient scheduling decisions that can scale with network size, resource allocation is made binary:

$$(SC4) : a_{u_s, s}^r \in \{0, 1\}, \quad \forall u_s \in U_s, r \in R, s \in S. \quad (8)$$

- **RB Allocation Constraint (SC5):** To prevent any single device from consuming disproportionate resources and to maintain fairness across a large number of devices, each user is limited to at most one RB:

$$(SC5) : \sum_{r \in R} a_{u_s, s}^r = 1, \quad \forall u_s \in U_s, s \in S \quad (9)$$

- **User Pairing and Multiplexing Constraint (SC6):** Each RB supports up to M devices multiplexed through power-domain NOMA. The total number of users sharing the same RB should not exceed the block capacity, ensuring efficient and fair use of resources even as the number of users increases. Thus, the number of users per RB must adhere to this constraint:

$$(SC6) : \sum_{u_s \in U_s} a_{u_s, s}^r \leq M, \quad \forall r \in R, s \in S. \quad (10)$$

B. Objective Optimization Problem Definition

The joint optimization problem comprehensively balances EE and massive connectivity, which are the core pillars of sustainable 6G networks. The objective function maximizes a weighted sum of these metrics."

$$(P0): \max_{\{a_{u_s, s}^r, p_{u_s, s}^r\}} \underbrace{\alpha \cdot \eta^{\text{EE}}}_{\text{Sustainability}} + \underbrace{\beta \cdot \mathcal{C}}_{\text{Scalability}}. \quad (11)$$

For the scalarized objective in (11), the weighting parameters are restricted to the feasible set $\alpha \in [0, 1]$ and $\beta = 1 - \alpha$. This choice ensures a normalized trade-off between energy efficiency and connectivity, where α controls the relative emphasis on sustainability (EE) and $(1 - \alpha)$ reflects the emphasis on massive connectivity. In the numerical evaluations, different values of α are selected to illustrate how the proposed

framework can flexibly operate in EE-oriented, connectivity-oriented, or balanced regimes, while always respecting the same underlying constraints and system model.

By inserting the closed-form expressions for EE (3) and connectivity (4) into the objective (11), we obtain a fully explicit mathematical formulation of the joint EE-connectivity optimization.

$$(P0) \max_{a_{u_s, s}^r, p_{u_s, s}^r} \alpha \cdot \left(\frac{\sum_{s \in S} \sum_{u_s \in U_s} \sum_{r \in R} a_{u_s, s}^r R_s(L, \epsilon_{u_s}^r, \gamma_{u_s}^r)}{\sum_{s \in S} \sum_{u_s \in U_s} \sum_{r \in R} (p_{u_s, s}^r + p_c)} \right) + \beta \cdot \left(\kappa \cdot \sum_{s \in S} \sum_{u_s \in U_s} \sum_{r \in R} a_{u_s, s}^r (1 - \epsilon_{u_s}^r) \right) \quad (12)$$

$$\text{s.t.: } R_s(L, \epsilon_{u_s}^r, \gamma_{u_s}^r) \geq R_{\min}, \quad \forall u_s, r, s, \quad (13a)$$

$$0 \leq p_{u_s, s}^r a_{u_s, s}^r \leq P_{\max}, \quad \forall u_s, r, s, \quad (13b)$$

$$\epsilon_{u_s}^r \leq \epsilon_{\text{th}}, \quad \forall u_s, r, s, \quad (13c)$$

$$a_{u_s, s}^r \in \{0, 1\}, \quad \forall u_s, r, s, \quad (13d)$$

$$\sum_{r \in R} a_{u_s, s}^r = 1, \quad \forall u_s, s, \quad (13e)$$

$$\sum_{u_s \in U_s} a_{u_s, s}^r \leq M, \quad \forall r, s, \quad (13f)$$

The optimization problem in (12) constitutes a joint EE-connectivity maximization framework that simultaneously determines the optimal set of active users, their NOMA groupings, RB assignments, and transmit power levels. Due to the coupling between binary scheduling variables, finite-blocklength reliability, and inter-cell interference, (P0) is a highly non-convex mixed-integer problem. To obtain a tractable solution, the problem is decomposed into two sub-problems. In the first stage, the resource scheduling sub-problem (P1), as discussed in detail in Section IV-C, optimizes the RB assignment matrix $a_{u_s, s}^r$ under a fixed power budget, enabling interference-aware NOMA pairing that satisfies FBL-based rate and reliability constraints. In the second stage, the Power allocation sub-problem (P2) in Section IV-D refines the transmit powers $p_{u_s, s}^r$ to further enhance EE while ensuring that the decoding error probability, rate, and maximum power constraints are met. These two stages provide an efficient alternating-optimization procedure that approximates the solution of the original joint EE-connectivity maximization problem.

It is important to emphasize that the above optimization is formulated *jointly across all cells* $s \in S$. Both the objective function and the constraints are written as network-wide sums over s , and the SINR terms contain inter-cell interference contributions. As a result, the problem cannot be decomposed into $|S|$ independent single-cell subproblems. In a single-cell formulation, the inter-cell interference summation in (2) would vanish, and each cell could optimize its own resources in isolation. In contrast, in the proposed multi-cell setting, all scheduling and power decisions are inherently coupled through the interference terms, which motivates the cooperative design adopted in this work.

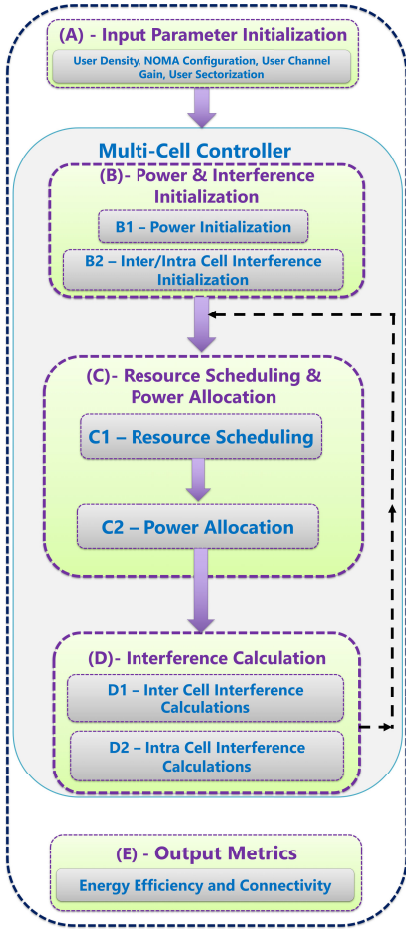


Fig. 3. Framework for joint optimization of EE and massive connectivity in multi-cell NOMA networks.

IV. PROPOSED JOINT OPTIMIZATION FRAMEWORK FOR A MULTI-CELL SCENARIO

In this section, we introduce a unified joint optimization framework explicitly tailored for multi-cell NOMA scenarios. This framework aims at simultaneously optimizing EE and massive connectivity through strategic interference management, resource allocation, and power control. The proposed algorithms address the critical trade-offs inherent in ultra-dense mMTC deployments characterized by FBL constraints.

A. Overview of the Joint Optimization Framework

The proposed framework, depicted in Fig. 3, integrates two complementary algorithms: the primary algorithm, SRS-MTCN, detailed in Algorithm 1, and its variant, SRS-MTCN minimum interference (minInterference Variant), described in Algorithm 2. While SRS-MTCN jointly maximizes EE

Algorithm 1 Proposed SRS-MTCN Algorithm for Multi-Cell Scenario

Require: $BW, M, P_{max}, G'_{u_s}, \sigma_N^2, Area, ISD, G_{AC}, G_{U_s}, \alpha, \beta$
Ensure: EEU_s (sum of values in the array)
User Equipment Initialization: Set $u_s \leftarrow G'_{u_s,s}$
 Sort users within each sector by ascending channel gain
Initialize: $p'_{u_s} = P_{max}, RtU_m = RtU_{min}$
Initialize: $IntU_s \leftarrow M \cdot N_{AC} \cdot P_{max} \cdot G_{AC}$
procedure STEP 1: INITIALIZATION AND PARAMETER COMPUTATION
 for $monteIter \leq numIter$ **do**
 for $cell \leq numCell$ **do**
 for each user u_s in NU **do**
 for each sector M_s in $nomaRB_{config}$ **do**
 Compute $SLNR$ using (2)
 Compute total interference $IntU_s \leftarrow \sum IntU_s$
 Compute total data rate $RtU_s \leftarrow \sum RtU_s$
 Compute total EE $EEU_s \leftarrow \sum EEU_s$
 end for
 Select Best User Combination:
 $[CHGain^{Int_{min}}, U_s^{Int_{min}}]$,
 $[CHGain^{Rt_{max}}, U_s^{Rt_{max}}]$,
 $[CHGain^{EE_{max}}, U_s^{EE_{max}}]$
 end for
 end for
 end procedure
procedure STEP 2: RESOURCE ALLOCATION (U_s, M_s)
Require: $IntU_s^{min}, RtU_s^{max}, EEU_s^{max}$
 for $U_s \in IntU_s^{min}$ **do**
 for $Users \in nomaRB_{config}$ **do**
 for $\alpha, \beta \in \{w_1, w_2\}$ **do**
 Compute total interference $IntU_M \leftarrow \sum IntU_M$
 Compute total data rate $RtU_M \leftarrow \sum RtU_M$
 Compute total EE $EEU_M \leftarrow \sum EEU_M$
 end for
 end for
 Select α and β maximizing EEU_M and RtU_M
 Assign RBs: $EEU_M^{max}, RtU_M^{max}, IntU_M^{min}$
 end for
 return $RtU_M^{max}, EEU_M^{max}, IntU_M^{min}$
end procedure
procedure STEP 3: POWER ALLOCATION (p'_{u_s})
Require: $p'_{u_s} = PowMat$
 while User Channel Gain Exists **do**
 for Power in $PowMat$ **do**
 for Users in $nomaU_{sr}$ **do**
 if $SLNR \geq SLNR_{min}$ **then**
 Compute $RtU_M^{max} \leftarrow \sum RtU_M$
 Compute $EEU_M^{max} \leftarrow \sum EEU_M$
 end if
 end for
 end for
 (i) **Intra-cell Interference Calculation:**
 Optimize power levels within each cell to minimize interference.
 : Assign power levels to NOMA pairs.
 (ii) **Inter-cell Interference Calculation:**
 Adjust power levels iteratively to minimize inter-cell interference.
 Apply constraints for maximum EE while ensuring QoS.
 Compute Final Power Allocation: RtU_M^{max}, EEU_M^{max}
 return $RtU_M^{max}, EEU_M^{max}, P_{wr_{alloc}}$
 end while
end procedure

Algorithm 2 Proposed SRS-MTCN Variant for Multi-Cell Scenario

Step 1: as in Algorithm 1
procedure STEP 2: RESOURCE ALLOCATION (U_s, M_s)
Require: $IntU_s^{min}, RtU_s^{max}, EEU_s^{max}$
for $U_s \in IntU_s^{min}$ **do**
 for $Users \in nomaRB_{config}$ **do**
 for $\alpha, \beta \in \{w_1, w_2\}$ **do**
 Compute total interference $IntU_M \leftarrow \sum IntU_M$
 Compute total data rate $RtU_M \leftarrow \sum RtU_M$
 Compute total EE $EEU_M \leftarrow \sum EEU_M$
 end for
 end for
 Select α and β minimizing $IntU_M$
 Assign RBs: $EEU_M^{max}, RtU_M^{max}, IntU_M^{min}$
 end for
return $RtU_M^{max}, EEU_M^{max}, IntU_M^{min}$
end procedure
procedure STEP 3: POWER ALLOCATION (p'_{u_s})
Require: $p'_{u_s} = PowMat$
 while User Channel Gain Exists **do**
 for Power in $PowMat$ **do**
 for Users in $nomaU_{sr}$ **do**
 if $SLNR \geq SLNR_{min}$ **then**
 Compute $RtU_M^{max} \leftarrow \sum RtU_M$
 Compute $EEU_M^{max} \leftarrow \sum EEU_M$
 end if
 end for
 end for
 (i) **Intra-cell Interference Calculation:**
 Optimize power levels within each cell to minimize interference.
 Assign power levels to NOMA pairs.
 (ii) **Inter-cell Interference Calculation:**
 Adjust power levels iteratively to minimize inter-cell interference.
 Apply constraints for minimum interference while ensuring QoS.
 Compute Final Power Allocation: RtU_M^{max}, EEU_M^{max}
 return $RtU_M^{max}, EEU_M^{max}, Pwr_{alloc}$
 end while
end procedure

and connectivity, minInterference Variant explicitly prioritizes interference reduction.

The SRS-MTCN algorithm operates through a three-step scheduling process. The first step calculates critical parameters such as user positions, channel gains, and user sectorization as illustrated in block “A” of Fig. 3. Initially, it partitions the search space into sectors based on users’ channel conditions, significantly improving computational efficiency without compromising performance. Users within each cell are categorized into sectors to facilitate efficient resource allocation, supporting configurations of 2, 4, and 8 NOMA users per RB. To further manage interference, the channel gains of users sharing an RB r within sector s are sorted in ascending order as:

$$|h_{u_1,s}^r|^2 \leq |h_{u_2,s}^r|^2 \leq \dots \leq |h_{u_{m,s}}^r|^2 \leq \dots \leq |h_{u_{M,s}}^r|^2. \quad (14)$$

This structured ordering underpins the design of interference-aware resource allocation and power control schemes. The pairing strategy selects users with optimal channel conditions from opposite sectors to minimize intra-cell interference, thus meeting QoS, data rate thresholds, and EE targets. Initial conditions are set with maximum inter-cell

interference and transmission power. SINR, EE, and data rates are calculated (blocks “B1” and “B2”), as detailed in Section IV-B.

In the second step, resource scheduling assigns users to NOMA RB pairs to minimize interference (block “C1” in Fig. 3 and Section IV-C). Optimized user pairing substantially enhances spectral efficiency and overall system performance.

The third step involves iterative power allocation and interference mitigation, detailed in Section IV-D (block “C2” in Fig. 3). After power allocation, updated inter- and intra-cell interference values are computed and adjusted according to the new power levels (blocks “D1” and “D2” in Fig. 3), effectively refining initial interference estimates.

Unlike conventional single-cell NOMA schedulers, the proposed SRS-MTCN algorithm operates in a fully multi-cell cooperative manner. User sectorization, NOMA pairing, and RB assignment (blocks “A” and “C1” in Fig. 3) are performed with explicit awareness of the interference generated towards neighboring cells, rather than treating other cells as fixed background noise. Subsequently, the power allocation stage (block “C2”) iteratively updates transmit powers across all cells and recomputes both intra-cell and inter-cell interference (blocks “D1” and “D2”), so that improvements in one cell do not unduly degrade users in adjacent cells. This joint, interference-aware operation enhances overall system performance and fairness across cells compared to non-cooperative single-cell algorithms that optimize each cell in isolation.

The proposed framework, through its integrated and iterative structure, offers a practical yet nearly optimal solution for sustainable and scalable multi-cell NOMA systems under realistic mMTC conditions.

B. Transmit Power Initialization

In the proposed framework, users are initially assigned a constant transmission power for the initial scheduling phase. The initial interference power is estimated based on the maximum interference from neighboring cells. This approach simplifies the initial resource allocation process, providing a foundation for subsequent power optimization and scheduling refinement. It effectively addresses interference and maximizes EE in multi-cell NOMA networks.

A key aspect of the proposed algorithm is the introduction of sectorization, which significantly minimizes the search space, improving computational efficiency without compromising system performance. User pairing is governed by two key constraints. First, a minimum SINR threshold, denoted as $\gamma = \gamma_{threshold}$, is applied to ensure that users possess sufficient channel quality to support successive interference cancellation (SIC) within a NOMA RB. Second, a FBL-based data rate constraint is enforced, requiring that each user’s achievable rate satisfies $R(L, \epsilon, \gamma) \geq R_{min}$. Resources are assigned to users only once per radio frame, with transmit power restricted within the range $[0, p_{max}]$. By jointly satisfying these constraints, the algorithm ensures that selected users are both decodable and energy-efficient, thereby enhancing reliability while minimizing unnecessary power consumption. This dual filtering mechanism promotes sustainable connectivity in ultra-dense mMTC environments.

C. Resource Scheduling: Multi-Cell Scenario

Once the initial transmission power and interference estimates are established in Step 1 of Algorithm 1, as described in Section IV-B, the scheduling framework proceeds with NOMA-based resource allocation. Specifically, the set M represents the group of users within cell s sharing a particular RB, while $u_m \in M$ denotes an individual user within that group.

At this stage, resource scheduling is carried out using a fixed and uniform power level, denoted as $\bar{p}_{u_s}^r$, assigned to each user. This simplification decouples power allocation from the initial scheduling process, enabling a tractable evaluation of user combinations based on connectivity and interference profiles.

Under this fixed-power setting, the optimization problem defined in (12) is reformulated to focus on the scheduling decision variables $a_{u_s}^r$, which determine the allocation of users to RBs. The resulting objective function evaluates the trade-off between EE and connectivity, providing a foundation for selecting the most suitable user groupings under current channel and interference conditions.

Resource Scheduling: Sub-Problem (P1)

$$\max_{a_{u_s}^r} \alpha \cdot \left(\frac{\sum_{s \in S} \sum_{u_s \in U_s} \sum_{r \in R} a_{u_s}^r R_s(L, \epsilon_{u_s}^r, \gamma_{u_s}^r)}{\sum_{s \in S} \sum_{u_s \in U_s} \sum_{r \in R} (p_{u_s}^r + p_c)} \right) + \beta \cdot \left(\kappa \cdot \sum_{s \in S} \sum_{u_s \in U_s} \sum_{r \in R} a_{u_s}^r (1 - \epsilon_{u_s}^r) \right) \quad (15)$$

$$\text{s.t.: } R_s(L, \epsilon_{u_s}^r, \gamma_{u_s}^r) \geq R_{\min}, \quad \forall u_s, r, s, \quad (16a)$$

$$\epsilon_{u_s}^r \leq \epsilon_{\text{th}}, \quad \forall u_s, r, s, \quad (16b)$$

$$a_{u_s}^r \in \{0, 1\}, \quad \forall u_s, r, s, \quad (16c)$$

$$\sum_{r \in R} a_{u_s}^r = 1, \quad \forall u_s, r, s, \quad (16d)$$

$$\sum_{u \in U_s} a_{u_s}^r \leq M, \quad \forall u_s, r, s, \quad (16e)$$

D. Power Allocation: Multi-Cell Scenario

Once the scheduling phase is completed and feasible user-resource assignments are obtained, the problem is reformulated to focus exclusively on power allocation. In this stage, the scheduling variables $a_{u_s}^r$ are fixed and omitted, and the transmission power $p_{u_s}^r$ becomes the primary optimization variable. Since users assigned to different RBs do not interfere with each other, the power allocation can be independently optimized across RBs. The resulting objective function (12) becomes:

1) Power Allocation: Sub-Problem (P2):

$$\max_{p_{u_s}^r} \alpha \cdot \left(\frac{\sum_{s \in S} \sum_{u_s \in U_s} \sum_{r \in R} a_{u_s}^r R_s(L, \epsilon_{u_s}^r, \gamma_{u_s}^r)}{\sum_{s \in S} \sum_{u_s \in U_s} \sum_{r \in R} (p_{u_s}^r + p_c)} \right) + \beta \cdot \left(\kappa \cdot \sum_{s \in S} \sum_{u_s \in U_s} \sum_{r \in R} a_{u_s}^r (1 - \epsilon_{u_s}^r) \right) \quad (17)$$

$$\text{s.t.: } R_s(L, \epsilon_{u_s}^r, \gamma_{u_s}^r) \geq R_{\min}, \quad \forall u_s, r, s, \quad (18a)$$

$$0 \leq p_{u_s}^r \leq P_{\max}, \quad \forall u_s, r, s, \quad (18b)$$

$$\epsilon_{u_s}^r \geq \epsilon_{\text{th}}, \quad \forall u_s, r, s, \quad (18c)$$

E. Algorithms for the Proposed SRS-MTCN and minInterference Variant

The operational procedures for the proposed **SRS-MTCN** algorithm and its **min-Interference SRS-MTCN** variant are summarized in Algorithm 1 and Algorithm 2, respectively. The first focuses on maximizing joint energy efficiency and connectivity through cooperative scheduling, while the second emphasizes minimizing interference to enhance multi-cell performance.

F. Computational (Time) Complexity Analysis of the Proposed Algorithm

The overall time complexity of the proposed SRS-MTCN algorithm is expressed as:

$$\mathcal{O}(N \log N) + \mathcal{O}(NM) + \mathcal{O}(\text{PowMat} \cdot N), \quad (19)$$

which scales as $\mathcal{O}(N \log N)$ with the number of active devices N , the NOMA cluster size (number of users per RB) M , and the number of discrete power levels explored during the power allocation stage PowMat (which is fixed by system design). Below, we provide the detailed complexity analysis that yields the above.

1) *Detailed Complexity Analysis:* The SRS-MTCN algorithm can be decomposed into three main steps:

- 1) Channel estimation and device sorting,
- 2) Interference-aware resource allocation,
- 3) Power allocation and interference coordination, and Iterative convergence evaluation

Let N denote the number of active devices, R the number of available resource blocks (RBs), S the number of cooperating cells, M the number of users per NOMA cluster (typically $M \in \{2, 4, 8\}$), and I the number of iterations performed by the inner resource allocation and power update loop. In dense mMTC scenarios, it is common to assume that the number of RBs grows linearly with the number of devices, i.e.,

$$R = cN, \quad c > 0, \quad (20)$$

which implies $R = \mathcal{O}(N)$ for worst-case asymptotic analysis. In practical systems, S , M , and I are upper-bounded by design constraints (e.g., limited cooperation region and maximum iteration budget).

a) *Channel Estimation and Device Sorting:* Each device estimates its effective channel gains over all R RBs and across S cooperating cells. This leads to a complexity proportional to the number of device-RB-BS combinations:

$$C_{\text{CE}} = \mathcal{O}(SNR). \quad (21)$$

For each RB, the N devices are sorted according to their effective channel quality to facilitate near-far pairing and interference-aware grouping. Sorting N devices has complexity $\mathcal{O}(N \log N)$, and this operation is repeated over all R RBs:

$$C_{\text{Sort}} = \mathcal{O}(RN \log N). \quad (22)$$

TABLE IV
COMPUTATIONAL COMPLEXITY COMPARISON
OF SCHEDULING ALGORITHMS

Algorithm	Complexity	Scalability	Remarks
SRS-MTCN (Proposed)	$\mathcal{O}(N^2 \log N)$	High	Real-time feasible.
Cooperative Scheduling [33]	$\mathcal{O}(N^2)$	Moderate	Greedy iterative allocation.
RAAP-NOMA [36]	$\mathcal{O}(N^3)$	Low	Exhaustive power-coefficient search over clusters.

b) Interference-Aware Resource Allocation: Once clusters are formed, the scheduler assigns clusters to RBs and updates interference levels across cells. For each iteration, these operations involve at most SRM cluster-RB combinations, yielding

$$C_{RA} = \mathcal{O}(SRM) \text{ per iteration.} \quad (23)$$

c) Power Allocation, Interference Coordination and Iterative Convergence Evaluation: The power control procedure updates the power coefficients of all users and evaluates convergence criteria over I iterations. The per-iteration complexity is on the order of the number of cluster-RB combinations, so the total cost of this stage can be expressed as:

$$C_{PACon} = \mathcal{O}(ISRM). \quad (24)$$

d) Aggregate Complexity: For a single scheduling interval, a natural upper bound on the total number of operations is obtained by summing the stage-wise contributions:

$$C_{Total} = \mathcal{O}(SNR + RN \log N + ISRM). \quad (25)$$

In *practical 6G mMTC deployments*, the system bandwidth, RB granularity, NOMA order, and iterations (R , S , M , and I) are expected to be bounded constants that do not scale with N . Under this regime, the dominant term in C_{Total} is $\mathcal{O}(RN \log N)$ and the complexity simplifies to

$$\mathcal{O}(N \log N). \quad (26)$$

In addition, for completeness, we also consider an *extreme worst-case* scenario in which the number of RBs grows linearly with the number of devices, i.e., $R = cN$ with $c > 0$. Substituting this into the expression above yields

$$\mathcal{O}(SN(cN) + (cN)N \log N + IS(cN)M) = \mathcal{O}(N^2 \log N), \quad (27)$$

assuming S , M , and I remain bounded. This $\mathcal{O}(N^2 \log N)$ bound is a conservative theoretical upper limit when both the number of users and the number of RBs are allowed to scale simultaneously, and does not reflect typical standardized mMTC deployments where R is fixed.

For benchmarking, the proposed complexity is compared with existing schemes as summarized in Table IV. The conventional Cooperative NOMA scheduler [33] and the Random Allocation and Pairing-based NOMA (RAAP-NOMA) [36] exhibit quadratic and cubic complexity, respectively, due to nested search operations and exhaustive interference computations. Overall, the analysis confirms that SRS-MTCN achieves polynomial-time scalability, ensuring energy-efficient and sustainable performance with manageable computational overhead in massive connectivity environments.

V. PERFORMANCE EVALUATION AND DISCUSSION

This section presents a comprehensive performance analysis of the proposed SRS-MTCN algorithm compared to several baseline scheduling schemes under both FBL and ideal Shannon IFBL assumptions. The analysis emphasizes three critical metrics: residual interference, system-level EE, and scalability with varying user densities and multiplexing levels.

A. Simulation Setup

The simulated scenario represents a multi-cell network comprising 7 cells with a 500-meter inter-site distance, implementing synchronous resource allocation for mMTC users within each cell. UEs are randomly and uniformly distributed within each cell, with uplink transmit power of 23 dBm [35]. The simulation framework adheres to the assumptions and parameters outlined in the standards of 3GPP's [14] and the IMT-Advanced Report ITU-R [37], to ensure compliance with industry standards.

Certain simulation parameters can substantially impact performance outcomes. For instance, increasing the user density within a cell amplifies inter-user interference, potentially degrading system performance relative to ideal expectations. Additionally, as the cell radius expands, users experience greater path loss due to increased distance from the base station. In such cases, it is recommended to adopt baseline base station configurations aligned with those used by telecom operators to ensure scenario relevance and benchmarking consistency. At each time slot, a subset of UEs is selected as active transmitters, while others from neighboring cells are designated as potential interferers. The interference originates both from UEs in neighboring cells and from co-scheduled UEs within the same cell. Based on these configurations, we compute the key performance indicators, and the corresponding results are presented in the following section.

The detailed simulation parameters are summarized in Table V. We compare EE with the maximum number of mMTC devices, N_{max} , per subcarrier for different algorithms at a given connection density. The results presented in the figures offer a comparative analysis of four different methods: i) the proposed Sustainable and Reliable Scheduling for Massive Machine Type Communication using NOMA (SRS-MTCN), ii) the SRS-MTCN Min-Interference Variant, iii) the Random Resource and Power Allocation Algorithm in NOMA (RAAP-NOMA), and iv) a baseline Cooperative Scheduling in NOMA method (Reference-CoopSch) described in [33].

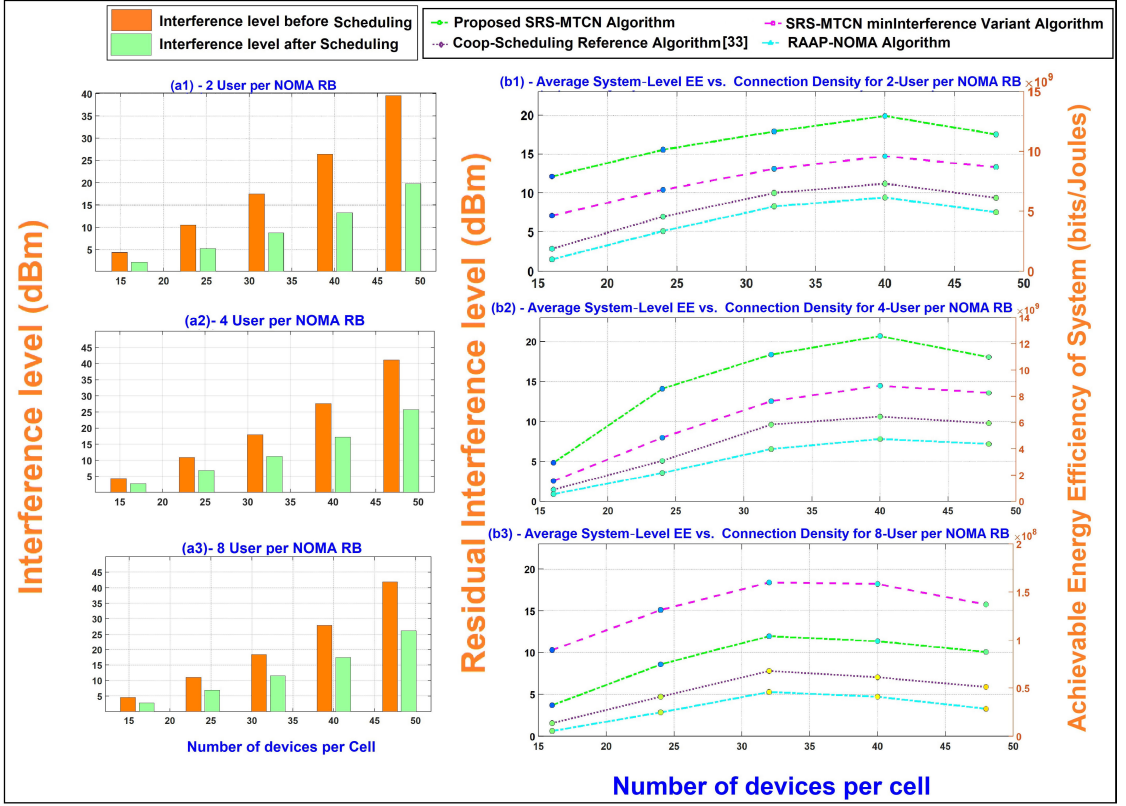


Fig. 4. Average system-level EE versus connection density under the ideal Shannon IFBL assumption for 2, 4, and 8 users per NOMA RB. Left (a1), (a2), (a3): Interference level before and after scheduling against the number of devices per cell, while managing 2-, 4-, and 8-user NOMA per RB, respectively. The orange bars represent the interference level before scheduling, which is substantially reduced by the proposed SRS-MTCN algorithm (green bars), resulting in a (70–80)% reduction across scenarios. This highlights the effectiveness of interference-aware user selection and scheduling for ultra-dense mMTC environments. Right (b1), (b2), (b3): Average EE (EE) versus number of devices per cell for the same 2-, 4-, and 8-user NOMA configurations. Each scatter point is annotated with its corresponding residual interference value, showing that lower interference levels (after scheduling) directly correlate with higher EE gains. The SRS-MTCN algorithm consistently outperforms benchmark methods, achieving up to 2×10^9 bits/J, and establishing the EE upper bound under Shannon capacity. Constraining these IFBL results with those for FBL in Fig. 5 further underscores the performance gap introduced by FBL and reliability constraints.

B. Results and Discussion

This analysis examines how EE varies with increasing user density under the ideal Shannon capacity assumption, serving as a theoretical upper bound for benchmarking practical schemes. Fig. 4 illustrates the system-level impact of user scheduling and interference minimization across varying NOMA configurations (2-, 4-, and 8-user multiplexing per RB).

Fig. 4's left-side subplots (a1-a3) are bar graphs of interference levels before and after scheduling for each NOMA configuration. The x-axis shows the number of devices per cell (user density), while the y-axis represents the interference level in dBm. The orange bars indicate the initial interference levels before scheduling, whereas the green bars represent the residual interference after applying the proposed SRS-MTCN algorithm. These results show a consistent reduction of approximately (70–80)% in interference across all scenarios,

validating the effectiveness of the scheduling strategy. This becomes increasingly important at higher device densities, where unmanaged interference severely impacts system scalability.

Fig. 4's right-side subplots (b1-b3) present the average system-level EE (in bits/Joule) versus the number of devices per cell, again under 2-, 4-, and 8-user NOMA configurations. The x-axis denotes device density, and the right-side y-axis quantifies EE. Notably, each scatter point is color-coded according to its associated post-scheduling interference level, as indicated by the color bar (z-axis) on the right. Blue-to-green shades represent lower interference, while orange-to-red tones indicate higher interference levels, allowing a visual correlation between interference suppression and EE.

The proposed SRS-MTCN algorithm consistently outperforms baseline methods. For instance, in the 4-user NOMA case with 40 users per cell, SRS-MTCN achieves a peak EE of approximately 1.3×10^9 bits/J, compared to

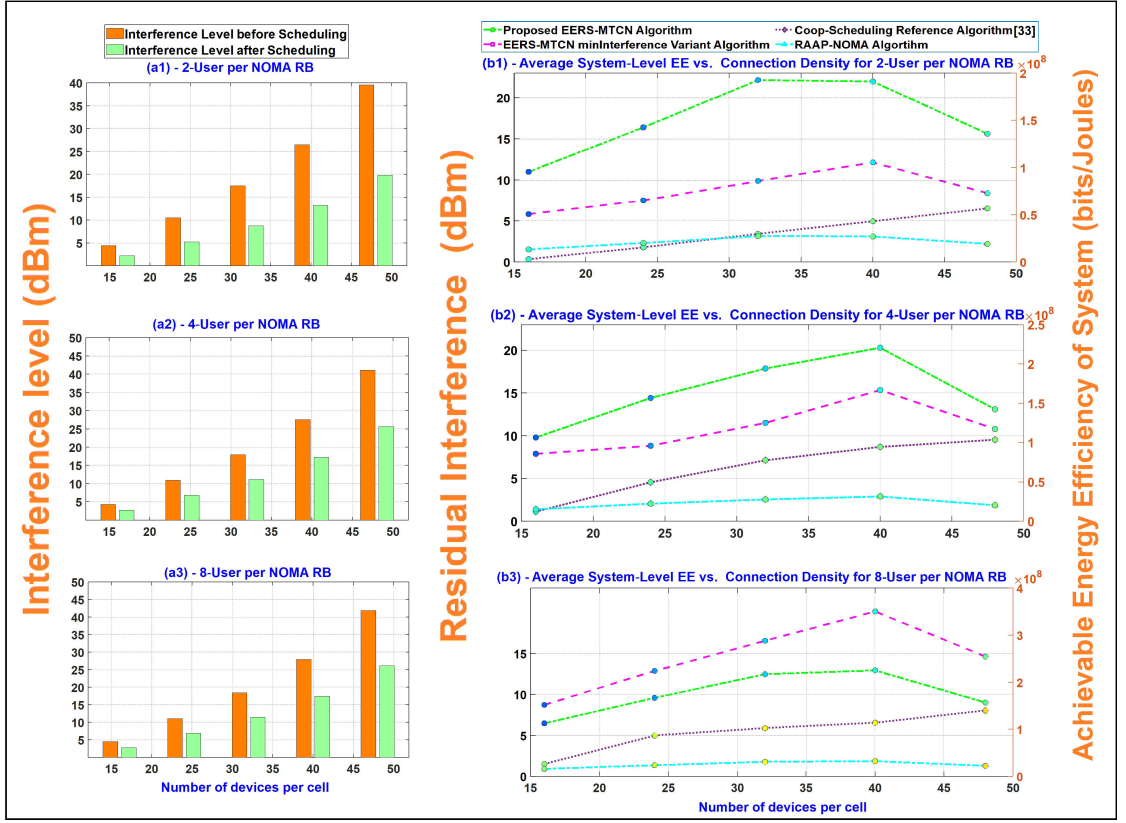


Fig. 5. Average system-level EE versus connection density under FBL regime for 2, 4, and 8 users per NOMA RB. Left (a1), (a2), (a3): Interference level before and after scheduling versus number of devices per cell, for 2-, 4-, and 8-user NOMA RB configurations, respectively. The orange bars represent pre-scheduling interference, while the green bars show the significant post-scheduling reduction achieved using the proposed SRS-MTCN algorithm. This demonstrates the algorithm's ability to reduce residual interference by approximately 60 – 75%, keeping post-scheduling interference consistently below 20-dBm across varying connection densities. Right (b1), (b2), (b3): Average system-level EE versus number of devices per cell, corresponding to the same 2-, 4-, and 8-user per RB scenarios. The EE performance trends are annotated with scatter markers showing the residual interference associated with each configuration. These results indicate that the SRS-MTCN algorithm not only minimizes interference but also enhances EE significantly, outperforming benchmark schemes by 35 – 58 % even under realistic FBL constraints. EE values under FBL constraints are lower than those achieved in the Shannon capacity case, Fig. 4, with EE peaking at approximately 1.3×10^9 bits/J.

8.8×10^8 (minInterference Variant), 6.4×10^8 (Coop-Scheduling), and 4.7×10^8 (RAAP-NOMA). Even under the more challenging 8-user NOMA setting, EE remains above 8.7×10^8 bits/J. These scatter-based patterns reinforce a key insight: lower residual interference correlates with higher EE, demonstrating the dual benefits of the proposed algorithm in suppressing interference while maximizing connectivity.

The above presented the results under an ideal Shannon IFBL regime. Now, Fig. 5 illustrates the performance of the proposed algorithm and benchmarks under FBL constraints, with a focus on EE vs. connection density. Fig. 5's left-side subplots (a1-a3) are bar charts that depict interference levels before and after scheduling across varying NOMA configurations (2-, 4-, and 8-user per RB). The x-axis shows the number of devices per cell, while the y-axis represents the interference level in dBm. Orange bars correspond to residual interference before scheduling, and green bars

reflect the post-scheduling interference levels after applying the SRS-MTCN algorithm. Across all NOMA groupings, the proposed scheme effectively reduces interference level, achieving up to 70% suppression, particularly at higher connection densities.

Fig. 5's right-side subplots (b1-b3) show the evaluation of the average system-level EE versus the number of devices per cell for the same NOMA configurations. The x-axis again represents the number of users per cell, while the right-side y-axis quantifies EE in bits/Joule. The scatter markers used in each curve are color-coded according to the corresponding post-scheduling interference level, with color mapping defined by the adjacent color bar (z-axis in dBm). Blue and cyan shades indicate lower interference levels, while yellow to red tones indicate higher residual interference. These visual cues reveal a direct relationship between reduced interference and higher energy efficiency.

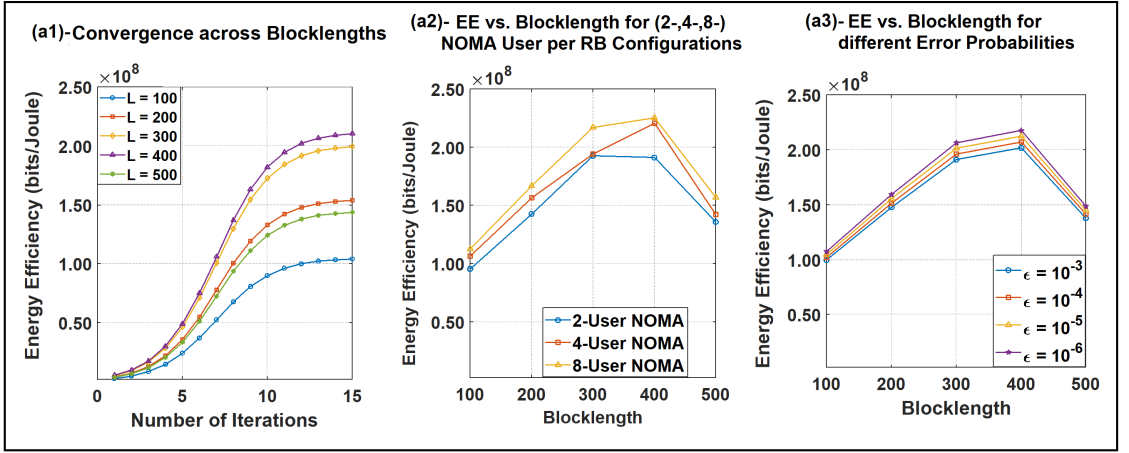


Fig. 6. Convergence and EE performance analysis of the proposed SRS-MTCN algorithm under FBL constraints. Left (a1): Convergence of the algorithm for different blocklengths ($L = 100$ to 500) across iterations. Middle (a2): EE versus blocklength for various NOMA configurations (2-, 4-, and 8-user per RB). Right (a3): EE versus blocklength under varying error probabilities $\epsilon \in \{10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}\}$. EE consistently improves as reliability (lower ϵ) increases, confirming the algorithm's robustness under stringent reliability constraints. All subplots reinforce that optimal blocklength and error probability combinations are critical to maximizing EE in FBL-based mMTC networks.

TABLE V
SIMULATION PARAMETERS FOR THE SEVEN-CELL
UPLINK NOMA SYSTEM

Parameters	Value
Transmit Power of Base Station; {UE} (dBm)	46; {23}
Noise Figure of Base Station; UE (dB)	7; 5
Carrier Frequency (MHz)	3500 MHz
Receiver Sensitivity	Channel Model + Effective Noise Power
Receiver Thermal Noise Density (dBm/Hz)	-174
Occupied System Bandwidth (kHz)	180
No. of UE per cell	100
Radius of Cell (m)	500
Interference Margin (dB)	0
Height of the base station; and UE (m)	25; 1.5
Channel Model	3GPP Urban Macro (UMa)
BS; UE antenna element gain (dBi)	8; 0
No. of antenna support per UE	1
Maximum Coupling Loss (MCL) (dB)	Trx Pwr + SINR
Effective Noise Power (dBm)	Thermal Noise + Noise Figure + $10\log_{10}(BW) + 10\log_{10}(N_{UE})$
SINR (dB)	To be calculated in the simulation

The SRS-MTCN algorithm consistently outperforms all benchmark methods (RAAP-NOMA, Coop-Scheduling, and its own minInterference variant) across all configurations. For instance, at 40 devices/m² with 4 users per RB, SRS-MTCN achieves an EE of approximately 2.2×10^9 bits/J, compared to 1.6×10^9 for the minInterference variant, 9.4×10^8 for the Coop-Scheduling approach, and only 3.1×10^8 for RAAP-NOMA. Even in the more interference-prone 8-user NOMA case, the SRS-MTCN approach sustains EE above 2.17×10^9 bits/J.

A direct comparison between Fig. 4 and Fig. 5 highlights the performance gap introduced by FBL constraints. While the Shannon-based results represent the ideal upper bound, the FBL plots confirm that the proposed SRS-MTCN algorithm retains strong efficiency even under realistic short-packet conditions. These results reinforce the importance of jointly optimizing power control and scheduling to maintain energy efficiency in ultra-dense FBL-constrained mMTC scenarios.

To further validate the effectiveness of the proposed algorithm, we evaluate its convergence behavior and EE performance across varying reliability requirements and block lengths. The findings reveal that the algorithm consistently achieves significant EE improvements under FBL constraints. The algorithm's behavior across three dimensions, convergence speed, EE sensitivity to blocklength, and reliability levels (error probability ϵ) has been analyzed. Our analysis captures scalability across multiplexing levels (2-, 4-, and 8-user NOMA per RB) and robustness to decoding reliability requirements. Fig. 6 presents a detailed assessment of the SRS-MTCN algorithm's performance under FBL conditions.

Overall, the above analysis demonstrates the scalability of the proposed algorithm across NOMA densities and its robustness against practical FBL and reliability constraints. These insights are vital for designing low-latency, energy-efficient 6G mMTC systems.

Finally, Fig. 6 illustrates the impact of the blocklength L , the NOMA user loading, and the target decoding error probability ϵ on the energy efficiency (EE) of the proposed SRS-MTCN scheme under finite-blocklength transmission.

(a1) - Convergence across blocklengths: Fig. 6 (a1) depicts the convergence of EE versus the iteration index of the SRS-MTCN algorithm for different blocklengths $L \in \{100, 200, 300, 400, 500\}$. For all considered values of L , the

EE increases monotonically with the number of iterations and reaches a stable plateau after approximately 12-15 iterations, which confirms the numerical stability of the proposed iterative procedure. Quantitatively, the final EE improves from about 1.0×10^8 bits/J for $L = 100$ to roughly 2.0×10^8 - 2.2×10^8 bits/J for $L = 300 - 400$, corresponding to an EE gain of almost a factor of two. When the blocklength is further increased to $L = 500$, the achieved EE decreases to approximately 1.4×10^8 bits/J. This non-monotonic trend is consistent with finite-blocklength information theory: for very short packets ($L = 100$), the coding rate and reliability are severely penalized, resulting in a low effective throughput per Joule. Increasing L to moderate values ($L = 300 - 400$) significantly improves the decoding reliability and the effective rate, which dominates the additional energy cost and thus increases EE. Beyond this regime, the reliability is already close to its target, and additional symbols mainly act as redundancy, causing the transmit energy to grow almost linearly with L while the useful throughput saturates, which results in a reduction in EE.

(a2) - EE vs. blocklength across NOMA configurations: Next, Fig. 6 (a2) compares the steady-state EE as a function of blocklength across three NOMA configurations with 2, 4, and 8 users per RB. All curves exhibit the same characteristic “rise-and-fall” shape: EE increases with L up to an optimal range around $L = 300$ -400 and then decreases for $L = 500$. For example, in the 2-user case the EE grows from about 0.95×10^8 bits/J at $L = 100$ to approximately 1.9×10^8 bits/J at $L = 300$, and then slightly decreases for $L = 400$ before dropping more noticeably at $L = 500$. The 4- and 8-user NOMA configurations exhibit the same behaviour, but with consistently higher EE values. At $L = 400$, the 8-user configuration attains about 2.25×10^8 bits/J, which is roughly 15-20% higher than the 2-user configuration.

These results lead to two key observations. First, for each NOMA loading, there exists an FBL operating point that maximizes EE; increasing the blocklength beyond this point improves reliability only marginally but reduces EE due to the increased energy consumption. Second, the proposed scheme can exploit the additional multiplexing gain provided by higher NOMA orders: when moving from 2- to 8-user NOMA, the EE gain at the optimal L is on the order of 20%, indicating that the joint scheduling and power-allocation procedure can sustain dense user loading while preserving energy efficiency.

(a3) - EE vs. blocklength for different error probabilities: Fig. 6 (a3) reports the EE as a function of L for several reliability targets $\epsilon \in \{10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}\}$. For all error-probability constraints, the same qualitative pattern is observed: EE increases with blocklength up to $L \approx 300 - 400$ and decreases thereafter. Tightening the reliability requirement (decreasing ϵ) shifts the curves slightly upward: in the EE-optimal region around $L = 400$, the curve for $\epsilon = 10^{-6}$ is about 7-8% higher than that for $\epsilon = 10^{-3}$. This indicates that the proposed resource-allocation scheme can leverage the longer blocklengths required by stricter reliability constraints to achieve higher bits per Joule, as long as L remains within the moderate regime. For very large blocklengths ($L = 500$), all curves exhibit a drop in EE irrespective of the error target. In this region, the reliability has essentially saturated close

to $1 - \epsilon \approx 1$, so the main effect of further increasing L is to consume more transmit and circuit energy per packet. The fact that the decrease in EE is observed for all ϵ confirms that energy efficiency and reliability are not fully aligned objectives: beyond a critical blocklength, any additional improvement in reliability comes at the expense of reduced EE. Fig. 6 (a1)-(a3) clearly demonstrate the fundamental trade-off between EE and connection reliability in the finite-blocklength regime. For small and moderate blocklengths, increasing L simultaneously improves reliability and EE; in this region, the proposed SRS-MTCN scheduler balances user clustering, power allocation, and NOMA loading to exploit the reliability gains. Beyond the EE-optimal blocklength, however, reliability improvements become marginal while the energy cost continues to grow, and the curves in Fig. 6 (a2)-(a3) show a consistent degradation in EE. The optimal operating region around $L = 300 - 400$, therefore, represents a practical compromise in which the system achieves near-ultra-reliable communication with high energy efficiency across different NOMA configurations and error-probability targets.

VI. CONCLUSION

This paper presented a joint EE and connectivity maximization framework for sustainable scheduling in dense 6G mMTC networks. By formulating a multi-objective optimization problem under FBL constraints and proposing the SRS-MTCN algorithm, we addressed the complex trade-offs among energy consumption, interference management, and user connectivity in cooperative multi-cell NOMA systems.

In terms of system-level performance under Shannon IFBL conditions, the proposed SRS-MTCN and its interference-aware variant consistently outperform the baseline algorithm across different user densities and NOMA configurations (2-, 4-, and 8-user per RB). The proposed framework achieves peak EE up to 1.9×10^9 bits/Joule with interference levels significantly reduced, validating the effectiveness of sectorized scheduling and dynamic power control.

In terms of the impact of FBL constraints, although absolute EE values are reduced by up to 40% (peaking at 1.12×10^9 bits/Joule for 4-user NOMA), the proposed scheme maintains a clear advantage. Specifically, it achieves a 35 - 58% EE improvement (up to 9.2×10^8 bits/Joule), compared to conventional benchmarks, and sustains lower residual interference even at higher device densities.

Experimental results further confirmed the algorithm's convergence behavior and sensitivity to critical FBL parameters. The proposed scheduling and power allocation scheme consistently converges within 10 - 12 iterations and achieves peak EE at moderate blocklengths ($L = 300 - 400$), with robustness to decoding reliability variations across a wide range of error probabilities ($\epsilon = 10^{-3} - 10^{-6}$). These results confirm the need for adaptive, interference-aware resource allocation strategies that account for short-packet constraints inherent in mMTC communications.

In summary, the SRS-MTCN framework enables scalable and energy-efficient user scheduling in ultra-dense networks, bridging theoretical performance limits and real-world deployment constraints. Future extensions may include

incorporating user mobility, dynamic traffic patterns, and hybrid grant-free access in non-terrestrial or vehicular network contexts.

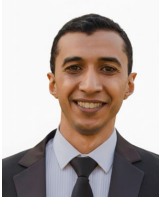
REFERENCES

- [1] L. Li, "Research on future 6G green wireless networks," *Green Technol. Sustainability*, vol. 3, no. 2, Apr. 2025, Art. no. 100156.
- [2] D. M. Soleymani and E. Roth-Mandutz, "6G energy efficiency and sustainability," Fraunhofer Inst. Integr. Circuits, Erlangen, Germany, White Paper, Version 1.0, Jan. 2023. [Online]. Available: <https://www.iis.fraunhofer.de/content/dam/iis/en/doc/1v/Whitepaper6GSustainability.pdf>
- [3] C.-X. Wang et al., "On the road to 6G: Visions, requirements, key technologies, and testbeds," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 2, pp. 905–974, 2nd Quart., 2023.
- [4] N. H. Mahmood, H. Alves, O. A. López, M. Shehab, D. P. M. Osorio, and M. Latva-Aho, "Six key features of machine type communication in 6G," in *Proc. 2nd 6G Wireless Summit (6G SUMMIT)*, Mar. 2020, pp. 1–5.
- [5] I. Ahmad et al., "Sustainability in 6G networks: Vision and directions," in *Proc. IEEE Conf. Standards Commun. Netw. (CSCN)*, Nov. 2023, pp. 202–208.
- [6] Y. Liu, J. Lee, C. Sun, Y. Han, and W. Chen, "Energy efficient scheduling for short packet communications with finite blocklength coding," *IEEE Trans. Green Commun. Netw.*, vol. 8, no. 4, pp. 1645–1660, Dec. 2024.
- [7] L. Blanco, E. Zeydan, S. Barrachina-Muñoz, F. Rezazadeh, L. Vettori, and J. Mangués-Bafalluy, "A novel approach for scalable and sustainable 6G networks," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 1673–1692, 2024.
- [8] *Framework and Overall Objective and Future Development of IMT for 2030 and Beyond*, document ITU-R M.2160-0, 2017.
- [9] B. E. Y. Belmekki and M.-S. Alouini, "NOMA as the next-generation multiple access in nonterrestrial networks," *Proc. IEEE*, vol. 112, no. 9, pp. 1303–1345, Sep. 2024.
- [10] Y. Liu et al., "Evolution of NOMA towards next generation multiple access (NGMA) for 6G," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 4, pp. 1037–1071, Apr. 2022.
- [11] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.
- [12] D. Wu and R. Negi, "Effective capacity: A wireless link model for support of quality of service," *IEEE Trans. Wireless Commun.*, vol. 24, no. 5, pp. 630–643, May 2003.
- [13] J. Cao, X. Zhu, S. Sun, E. Kurniawan, and A. Boonkajay, "Risk-aware and energy-efficient AoI optimization for multiconnectivity WNCs with short-packet transmissions," *IEEE Internet Things J.*, vol. 11, no. 12, pp. 21474–21485, Jun. 2024.
- [14] *5G; Study on Scenarios and Requirements for Next [G]eneration Access Technologies*, document TR 38.913, 3GPP, 2024.
- [15] A. Rabee and I. Barhumi, "Energy efficient resource allocation approach for uplink NOMA multi-cell systems based on multi-agent DRL," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2024, pp. 1–6.
- [16] B. Rashid, A. Ahmad, S. Ali, Z. Kaleem, A. Alkhayat, and C. Yuen, "Energy efficient resource allocation for uplink MC-NOMA based heterogeneous small cell networks with wireless backhaul," *IEEE Trans. Veh. Technol.*, vol. 72, no. 3, pp. 3419–3429, Mar. 2023.
- [17] S. M. Jameel, A. K. Tareen, H. Jung, and S. A. Hassan, "Energy efficiency optimization of CoMP C-NOMA system using Markov decision process," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2024, pp. 1–6.
- [18] Y. Yang, Y. Hu, and M. C. Gursoy, "Energy efficiency of RIS-assisted NOMA-based MEC networks in the finite blocklength regime," *IEEE Trans. Commun.*, vol. 72, no. 4, pp. 2275–2291, Apr. 2024.
- [19] B. Liu, P. Zhu, J. Wang, and X. You, "Cell-free network-assisted full-duplex enabled short-packet communications toward energy efficiency in IIoT networks," *IEEE Trans. Wireless Commun.*, vol. 24, no. 12, pp. 10244–10260, Dec. 2025.
- [20] J. Gao, Y. Wu, T. Li, and W. Zhang, "Energy efficiency of MIMO massive uncoordinated random access with finite blocklength," *IEEE Wireless Commun. Lett.*, vol. 12, no. 4, pp. 743–747, Apr. 2023.
- [21] Y. Zhang, W. Cheng, and W. Zhang, "Multiple access integrated adaptive finite blocklength for ultra-low delay in 6G wireless networks," *IEEE Trans. Wireless Commun.*, vol. 23, no. 3, pp. 1670–1683, Mar. 2024.
- [22] J. Zheng, H. Jiang, L. Yu, S. Y. Khamiseh, and M. Rahman, "On the performance of uplink NOMA short-packet communication system," *IEEE Trans. Veh. Technol.*, vol. 73, no. 12, pp. 18678–18687, Dec. 2024.
- [23] H. Hu, S. Zhao, Y. Huang, K. Yu, Q. Kang, and G. Cheng, "Energy efficiency maximization for IRS-assisted UAV short packet communication," *Sci. Rep.*, vol. 15, no. 1, p. 9124, Mar. 2025.
- [24] N. H. Tu and K. Lee, "Multihop MIMO full-duplex relay networks with short-packet URLLCs," *IEEE Syst. J.*, vol. 18, no. 4, pp. 1975–1986, Dec. 2024.
- [25] S. Thapliyal, R. Pandey, and C. Charan, "Evaluating energy harvesting UAV-NOMA network with random user pairing in the finite blocklength regime," *Veh. Commun.*, vol. 49, Oct. 2024, Art. no. 100831.
- [26] Y. Huang, Y. Jiang, F.-C. Zheng, P. Zhu, D. Wang, and X. You, "Energy-efficient optimization in user-centric cell-free massive MIMO systems for URLLC with finite blocklength communications," *IEEE Trans. Veh. Technol.*, vol. 73, no. 9, pp. 12801–12814, Sep. 2024.
- [27] S. R. Sabuj, S. Ahmed, and H.-S. Jo, "Multiple CUAV-enabled mMTC and URLLC services: Review of energy efficiency and latency performance," *IEEE Trans. Green Commun. Netw.*, vol. 7, no. 3, pp. 1369–1382, Sep. 2023.
- [28] J. Wang, W. Cheng, and W. Zhang, "FBC-enhanced ϵ -effective capacity optimization for NOMA," 2024, *arXiv:2410.11319*.
- [29] Y. Zhao, W. Gao, Y. Zhu, Y. Hu, and J. Gross, "Energy efficiency maximization for multi-node IoT networks operating with finite blocklength codes," in *Proc. 19th Int. Symp. Wireless Commun. Syst. (ISWCS)*, 2024, pp. 1–6.
- [30] M. Soleymani, I. Santamaria, E. Jorswieck, R. Schober, and L. Hanzo, "Energy efficiency optimization of finite block length STAR-RIS-aided MU-MIMO broadcast channels," 2025, *arXiv:2506.11594*.
- [31] D.-D. Tran, V. N. Ha, S. K. Sharma, T. T. Nguyen, S. Chatzinotas, and P. Popovski, "Energy-efficient NOMA for 5G heterogeneous services: A joint optimization and deep reinforcement learning approach," *IEEE Trans. Commun.*, vol. 73, no. 4, pp. 2448–2465, Apr. 2025.
- [32] Y. Zhao, T. Peng, Y. Guo, Y. Niu, and W. Wang, "An intelligent scheme for energy-efficient uplink resource allocation with QoS constraints in 6G networks," *IEEE Trans. Netw. Service Manage.*, vol. 22, no. 1, pp. 255–269, Feb. 2025.
- [33] C. B. Mwakwata et al., "Cooperative scheduler to enhance massive connectivity in 5G and beyond by minimizing interference in OMA and NOMA," *IEEE Syst. J.*, vol. 16, no. 3, pp. 5044–5055, Sep. 2022.
- [34] J. Chen, L. Zhang, Y.-C. Liang, and S. Ma, "Optimal resource allocation for multicarrier NOMA in short packet communications," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 2141–2156, Feb. 2020.
- [35] *5g-ur; User Equipment (UE) Radio Transmission and Reception; Part 2: Range 2 Standalone*, document TS 38.101-2, 3GPP, 2024.
- [36] Z. Ding, Z. Yang, P. Fan, and H. V. Poor, "On the performance of non-orthogonal multiple access in 5G systems with randomly deployed users," *IEEE Signal Process. Lett.*, vol. 21, no. 12, pp. 1501–1505, Dec. 2014.
- [37] *Guidelines for Evaluation of Radio Interface Technologies for IMT-2020*, document ITU-R M.2412-0, 2017.



Aamir Latif (Graduate Student Member, IEEE) received the M.S. degree in telecommunication engineering from the University of Bologna, Italy, in 2016. He is currently pursuing the Ph.D. degree in information and communication technology with the Thomas Johann Seebeck Department of Electronics, Tallinn University of Technology, Estonia. From 2016 to 2022, he was with Chiplogix (Pvt) Ltd., as an Embedded Design and Test Manager for onboard communication systems. His research focuses on resource allocation and multi-objective optimization

for massive connectivity in 6G systems.



Osama Elgarhy (Member, IEEE) received the B.Sc. degree in electrical and communication engineering in 2009, and the M.Sc. and Ph.D. degrees in telecommunication engineering from the Politecnico Di Milano, Milan, Italy, in 2016 and 2020, respectively. From 2009 to 2013, he was a Teaching Assistant with the Pyramids Higher Institute of Engineering, Cairo, Egypt. In 2016, he spent a research period with Azcom Technology, Milan, Italy. He is currently a Research Engineer with the Thomas Johann Seebeck Department of Electronics, Tallinn

University of Technology, Tallinn, Estonia. His current research interests include energy efficiency-latency optimization, resource allocation for cellular systems, NB-IoT, V2X, and device-to-device communications.



Yannick Le Moullec (Senior Member, IEEE) received the M.Sc. degree from the Université de Rennes I, Rennes, France, in 1999, and the Ph.D. and H.D.R. (accreditation to supervise research) degrees in electrical engineering from the Université de Bretagne Sud, Lorient, France, in 2003 and 2016, respectively. From 2003 to 2013, he was a Post-Doctoral Researcher, an Assistant Professor, and an Associate Professor with the Center for Embedded Software Systems and the Department of Electronic Systems, Aalborg University, Denmark. He was with

the Thomas Johann Seebeck Department of Electronics, Tallinn University of Technology, Tallinn, Estonia, as a Senior Researcher from 2013 to 2016, and has been on a professorship since 2017. He has supervised or co-supervised more than 60 M.Sc. students and 17 Ph.D. students. He has been involved in more than 20 research projects at the national and European levels. His research interests span embedded systems, edge machine learning, AI-IoT, wireless communication, and applications thereof.



Nurul Huda Mahmood (Senior Member, IEEE) was born in Bangladesh. He received the Ph.D. degree from NTNU, Norway, in 2013. He is currently a Senior Research Fellow with CWC, University of Oulu, Finland, and the Coordinator of wireless connectivity research with the Finnish 6G Flagship program. Before joining the University of Oulu in 2018, he was an Associate Professor with the Department of Electronics Systems, Aalborg University, Denmark. He has contributed to various international research projects, as a WP Leader in

Hexa-X-II—EU's flagship 6G research project. He has co-authored over 100 peer-reviewed publications. His current research interests include resilient communications for wireless networks.



Muhammad Mahtab Alam (Senior Member, IEEE) received the M.Sc. degree in electrical engineering from Aalborg University, Denmark, in 2007, and the Ph.D. degree in signal processing and telecommunication from the INRIA Research Center, University of Rennes 1, France, in 2013. From 2014 to 2016, he was a Post-Doctoral Researcher with Qatar Mobility Innovation Center, Qatar. In 2016, he joined as the European Research Area Chair and an Associate Professor with the Thomas Johann Seebeck Department of Electronics, Tallinn

University of Technology, where he was elected as a Professor in 2018 and a tenured Full Professor in 2021. Since 2019, he has been the Communication Systems Research Group Leader. He has over 15 years of combined academic and industrial multinational experiences while working in Denmark, Belgium, France, Qatar, and Estonia. He has several leading roles as a PI in multimillion Euros international projects funded by European Commission (Horizon Europe 5G-BALTICS, 5G-TIMBER, H2020 5GROUTES, NATO SPS (G5482), Estonian Research Council (PRG424), and Telia Industrial Grant). He is the author or co-author of more than 180 research publications. He is actively supervising several Ph.D. and post-doctoral researchers. He is also a contributor in two standardization bodies (ETSI SmartBAN and IEEE-GeenICT-EECH), including "Rapporteur" of work item: DTR/ SmartBAN-0014. His research focuses on the fields of wireless communications-connectivity, mobile positioning, 5G/6G services, and applications.

Appendix 5

Publication (V):

A. Latif, Y. Le Moullec, and M. M. Alam, "Multi-objective policy design using reinforcement learning in 6g networks," in *2026 International Wireless Communications and Mobile Computing (IWCMC)*, pp. 1352–1357, 2026

Multi-Objective Policy Design using Reinforcement Learning in 6G Networks

Aamir Latif, Yannick Le Moullec, Muhammad Mahtab Alam
Tallinn University of Technology, Estonia.

Email: (aamir.latif; yannick.lemoullec; muhammad.alam)@taltech.ee

Abstract—Sustainable 6G wireless networks are expected to deliver jointly optimized key performance indicators (KPIs) to meet growing demands for highly reliable, energy-efficient, and massively connected devices. Balancing these inherently conflicting objectives under stringent energy, latency, and computational constraints remains challenging. Multi-Objective Reinforcement Learning (MORL) enables the study of Pareto-efficient trade-offs; contemporary wireless solutions generally rely on static scalarization or multiple trained policies, and rarely accommodate rapid, context-sensitive preference adaptation. We propose *Wireless Multi-Objective Conditional Envelope PPO (WiMO-CEPPO)*, a centralized single-agent MORL framework that learns a single preference-conditioned policy and adapts online to varying network conditions without retraining. WiMO-CEPPO integrates: (i) a context-to-preference mapper that infers objective priorities from real-time network state, (ii) a conditional envelope mechanism that attenuates policy updates under strong inter-objective conflict, and (iii) a sustainable reward formulation that jointly accounts for energy efficiency, connectivity, reliability, interference, and hardware sustainability. Simulation results show that WiMO-CEPPO outperforms representative MORL baseline methods in Pareto trade-off quality and key wireless KPIs, and improves training stability under conflicting objectives.

I. INTRODUCTION AND RELATED WORK

The transition to sixth-generation (6G) wireless systems is driven by requirements that extend beyond prior generations: sustained support for ultra-reliable low-latency communication, pervasive connectivity for massive device populations, and substantial gains in energy efficiency (EE), all under hard operational bounds on compute, inference latency, and end-to-end sustainability [1]–[3]. The central difficulty is not optimizing any single KPI in isolation, but managing their inherent tension under joint optimization. For example, reliability-oriented mechanisms (e.g., conservative adaptation or redundancy) typically increase energy consumption and queuing delay; connectivity-maximizing policies in dense deployments amplify cross-link interference (CLI), which degrades reliability and spectral efficiency; and aggressive CLI suppression competes directly with throughput and QoS when spectrum is scarce [4], [5]. As a result, resource allocation in interference-coupled, time-varying networks is intrinsically Pareto-structured: improvements along one dimension often incur losses along others [6], [7].

Classical optimization approaches address this complexity through convex relaxation, tractable interference approximations, or decomposed cross-layer formulations. While these methods provide guarantees under their modeling assumptions, they degrade in the non-stationary regimes characteristic of practical 6G deployments (e.g., fluctuating load, heterogeneous QoS, and interference dynamics) [5], [8]. This motivates learning-based resource management, in which policies observe network state and adapt decisions over time to match the prevailing operating regime.

Multi-Objective Reinforcement Learning (MORL) provides a natural foundation by representing EE, connectivity, reliability, and interference mitigation as a vector-valued reward and learning Pareto-efficient policies rather than a single-point solution [6], [9]. Recent wireless MORL studies include MORL scheduling under varying operator preferences [4], multi-agent MORL for IRS-assisted wireless-powered networks trading throughput and residual energy [7]. Related work has also investigated learning-driven joint optimization of EE, connectivity, and reliability under complex dynamics [7], [10], [11]. A recurring limitation, however, is the reliance on static scalarization: a fixed weight vector collapses objectives into a scalar reward before training, confining the learned policy to a narrow region of the Pareto frontier and requiring retraining (or policy portfolios) as operating conditions change [12].

Preference-conditioned MORL mitigates this issue by learning a single generalized policy conditioned on a continuous preference vector, enabling different trade-off behaviors at inference time without retraining [13]. Dynamic preference estimation can further improve sample efficiency by steering training toward favorable preferences (e.g., DPIN) [14], and dynamic preference formulations have shown improved generalization in network management tasks [15]. Envelope-based MORL augments preference conditioning by optimizing a convex envelope of multi-objective value functions, providing a unified representation of the Pareto surface and demonstrating strong performance in URLLC-oriented vehicular settings [13]. Nevertheless, two deployment-critical challenges persist for 6G RAN control: *preference determination* (preference vectors are typically externally specified, despite being strongly context-dependent in practice) and *conflict-stable learning* (opposing per-objective gradients can induce instability and poor Pareto coverage) [12], [14], [15]. Moreover, operational sustainability—notably inference cost and latency at scheduling timescales—is often omitted, despite directly constraining deployability in real-time RAN control.

To address these gaps, we propose **WiMO-CEPPO** (*Wireless-Aware Multi-Objective Conditional Envelope PPO*), a context-aware, preference-conditioned MORL framework for multi-cell uplink resource management that jointly optimizes EE, connectivity, reliability, and CLI. Three targeted contributions distinguish WiMO-CEPPO from prior art. (i) **Autonomous preference inference**: a lightweight Context-to-Preference Mapper (CPM), $\mathbf{w}_t = f_{\vartheta}(s_t)$, translates real-time network state into a dynamic objective-priority vector, eliminating manual weight scheduling. (ii) **Conflict-stable optimization**: a conditional envelope mechanism embedded within PPO down-weights updates when per-objective advantage signals conflict, improving training stability and Pareto coverage without separate per-trade-off

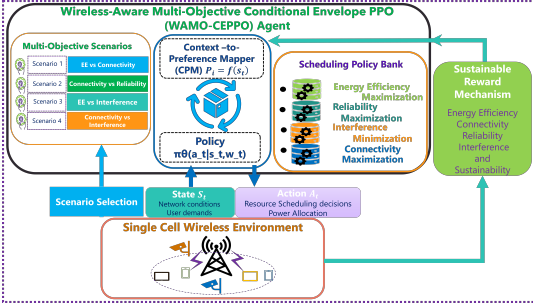


Figure 1: Single-cell WiMO-CEPPO framework for sustainable 6G resource allocation. A single RL agent leverages context-aware preference mapping to generate dynamic policies for joint scheduling and power allocation, optimized through a multi-objective reward integrating energy efficiency, connectivity, reliability, and interference minimization.

policies. (iii) **Operational sustainability:** normalized inference latency and per-decision compute cost are incorporated as explicit deployability constraints rather than implicit reward heuristics, ensuring feasibility on realistic RAN hardware [1], [2].

WiMO-CEPPO is evaluated against MO-PPO, MO-SAC, MO-TD3, and MO-DDPG across single-cell and multi-cell centralized and distributed settings, using hypervolume (HV), inverted generational distance (IGD), KPI radar analysis, 3D Pareto-front visualization, and ablation studies, consistently demonstrating superior multi-objective performance in interference-coupled wireless environments [5], [6].

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider an uplink multi-cell NOMA network of M cells indexed by $m \in \{0, \dots, M-1\}$, each served by a gNB supporting U UEs. The system bandwidth B_{tot} is partitioned into R orthogonal RBs of width $B_0 = B_{\text{tot}}/R$; at most N_m UEs may share one RB under NOMA. We evaluate WiMO-CEPPO against two architectures of increasing complexity.

In the **single-cell** baseline, the focal cell $m=0$ is optimized while neighboring cells contribute a constant per-RB interference aggregate I^r , estimated from long-run averages. This baseline is intentionally designed to isolate the core scheduling and power-allocation behavior of the agent while preserving the effect of neighboring-cell activity through a simplified interference abstraction. It also serves as a methodological bridge: once the core policy behavior is characterized in this setting, the formulation can be extended to multi-cell layouts in which the same policy is executed independently per cell, enabling scalable spatial decomposition. In the **centralized C-RAN** setting, a single BBU controller holds a global view of all M cells and treats inter-cell interference as a co-optimized decision variable.

A. Channel Model

Let $G_{u,m}^r$ denote the channel gain from UE u in cell m to its serving gNB on RB r , and $G_{w,t,m}^r$ the cross-cell gain from UE w in cell t to gNB m . Under an ideal SIC, the

uplink SINR is

$$\gamma_{u,m}^r = \frac{a_{u,m}^r G_{u,m}^r P_{u,m}^r}{\sum_{\substack{v \neq u, \\ v \in \mathcal{U}_m}} a_{v,m}^r G_{v,m}^r P_{v,m}^r + \mathcal{I}_m + \sigma_N}, \quad (1)$$

where $\sigma_N = N_0 B_0$ and the inter-cell term takes two forms:

$$\mathcal{I}_m^r = \begin{cases} I^r = \sum_{t \in \mathcal{N}(0)} \sum_w \bar{a}_{w,t}^r G_{w,t,0}^r \bar{P}_{w,t}^r, & \text{single cell,} \\ \sum_{t \neq m} \sum_{w \in \mathcal{U}_t} a_{w,t}^r G_{w,t,m}^r P_{w,t}^r, & \text{centralized.} \end{cases} \quad (2)$$

In the single-cell case, I^r is fixed during training and evaluation; in the centralized case, the same term is induced by the current joint action and is therefore co-optimized by the controller. The achievable rate on RB r is $R_{u,m}^r = B_0 \log_2(1 + \gamma_{u,m}^r)$.

B. Multi-Objective Problem Formulation

Let $\mathbf{a} = [a_{u,m}^r] \in \{0, 1\}$ and $\mathbf{P} = [P_{u,m}^r]$ denote the RB-assignment and power-allocation matrices (over cell 0 in the single-cell case; over all M cells in the centralized case). The four objectives, unified across both architectures, are

$$f_{\text{EE}} = \max_{\mathbf{a}, \mathbf{P}} \frac{\sum_{m,u,r} a_{u,m}^r \log_2(1 + \gamma_{u,m}^r)}{p_c + \sum_{m,u,r} P_{u,m}^r}, \quad (3)$$

$$f_{\text{Conn}} = \max_{\mathbf{a}, \mathbf{P}} \sum_{m,u} \mathbf{1}\{\sum_r a_{u,m}^r R_{u,m}^r \geq R_u^{\min}\}, \quad (4)$$

$$f_{\text{Rel}} = \max_{\mathbf{a}, \mathbf{P}} \sum_{m,u,r} a_{u,m}^r \mathbf{1}\{\gamma_{u,m}^r \geq \gamma^{\min}\}, \quad (5)$$

$$f_{\text{CLI}} = \min_{\mathbf{a}, \mathbf{P}} \sum_{m,r} \sum_{t \neq m} \sum_{w \in \mathcal{U}_t} a_{w,t}^r G_{w,t,m}^r P_{w,t}^r, \quad (6)$$

where sums over m run over $\{0\}$ (single cell) or $\{0, \dots, M-1\}$ (centralized), and p_c denotes circuit power per scheduled UE. The resulting vector optimization is

$$\max_{\mathbf{a}, \mathbf{P}} (f_{\text{EE}}, f_{\text{Conn}}, f_{\text{Rel}}, -f_{\text{CLI}}). \quad (7)$$

subject to

$$0 \leq P_{u,m}^r \leq P_u^{\max}, \quad \forall u, m, r, \quad (\text{SC1})$$

$$\sum_r a_{u,m}^r R_{u,m}^r \geq R_u^{\min}, \quad \forall u, m, \quad (\text{SC2})$$

$$\sum_{u \in \mathcal{U}_m} a_{u,m}^r \leq N_m, \quad \forall m, r, \quad (\text{SC3})$$

$$a_{u,m}^r \in \{0, 1\}, \quad \forall u, m, r, \quad (\text{SC4})$$

$$C_{\text{hw}} \leq 1, \quad \forall \tau. \quad (\text{SC5})$$

Constraints (SC1)–(SC3) apply for all u, m, r ; (SC5) bounds normalized inference latency and compute cost, thereby enforcing hardware sustainability as an explicit deployability constraint. In the centralized case, the state dimension grows as $\mathcal{O}(MRU)$, so (SC5) additionally imposes an architecture-level ceiling on policy-network width verified by hardware profiling. A static weighted sum $\alpha f_{\text{EE}} + \beta f_{\text{Conn}} + \delta f_{\text{Rel}} - \zeta f_{\text{CLI}}$ is not responsive to time-varying operating conditions; we therefore introduce a context-dependent preference vector $\mathbf{w}_t = [\alpha_t, \beta_t, \delta_t, \zeta_t] = f_\theta(s_t)$ via the CPM, converting (7) into a preference-conditioned MO-MDP.

C. MO-MDP Formulation

The centralized problem is modeled as a single-agent MO-MDP $(\mathcal{S}^C, \mathcal{A}^C, \mathcal{T}^C, \mathcal{W}, \mathbf{R}^C, \gamma)$. The per-cell state is

$$s_t^{(m)} = [\mathbf{G}_t^{(m)}, \mathcal{I}_t^{(m)}, \boldsymbol{\rho}_t^{(m)}, \mathbf{q}_t^{(m)}, \mathbf{h}_t^{(m)}], \quad (8)$$

where $\mathbf{G}_t^{(m)}$ holds local channel gains, $\mathcal{I}_t^{(m)}$ is the per-RB interference term from (2) (fixed I^r for the single-cell baseline; live for centralized), $\rho_t^{(m)}$ is cell load, $\mathbf{q}_t^{(m)}$ captures per-UE rate deficits, and $\mathbf{h}_t^{(m)}$ collects compact context features. For the **single-cell** MO-MDP, the state is $s_t^{(0)}$ with fixed $\mathcal{I}_t^{(0)}=I^r$; action $\mathbf{a}_t^{(0)}$ covers cell 0 only; reward $\mathbf{R}^{(0)}$ is per-cell. For the **centralized** MO-MDP, the **Joint state**:

$$\mathbf{s}_t^C = [s_t^{(0)}, s_t^{(1)}, \dots, s_t^{(M-1)}], \quad (9)$$

where each per-cell component $s_t^{(m)}$ has the form in (8) but with the fixed interference term replaced by the live per-RB inter-cell interference computed from the current joint action. The joint state is collected at the Near-RT RIC via E2 telemetry reports from all M cells at each control interval.

Joint action: $\mathbf{a}_t^C = (\tilde{\mathbf{a}}_t, \tilde{\mathbf{P}}_t)$ for all M cells. The projection step applies independently per cell:

$$(\mathbf{a}_t^{(m)}, \mathbf{P}_t^{(m)}) = \Pi_{\Omega_m}(\tilde{\mathbf{a}}_t^{(m)}, \tilde{\mathbf{P}}_t^{(m)}), \quad \forall m, \quad (10)$$

where Ω_m is the per-cell feasible set defined by (SC1)–(SC4).

Reward:

$$\mathbf{R}^C(\mathbf{s}_t, \mathbf{a}_t) = [R_{\text{EE}}^C, R_{\text{Conn}}^C, R_{\text{Rel}}^C, R_{\text{CLI}}^C]^\top, \quad (11)$$

where each component corresponds to (3)–(6) evaluated over the cluster. The CPM generates $\mathbf{w}_t = f_\vartheta(\mathbf{s}_t^C)$ from the joint state, and the scalarized reward is $r_t = \mathbf{w}_t^\top \mathbf{R}^C(\mathbf{s}_t, \mathbf{a}_t)$. Because the controller is centralized, the transition kernel remains stationary, which is a fundamental advantage of this architecture.

III. PROPOSED WiMO-CEPPO FRAMEWORK

The following components are shared across deployment architectures; architecture-specific instantiations are noted where relevant.

A. Constraint Handling: Projection and Penalty

WiMO-CEPPO uses a hybrid constraint-handling strategy. Hard feasibility constraints are enforced by projecting raw actor outputs onto the feasible set:

$$(\mathbf{a}_t^{(m)}, \mathbf{P}_t^{(m)}) = \Pi_{\Omega_m}(\tilde{\mathbf{a}}_t^{(m)}, \tilde{\mathbf{P}}_t^{(m)}), \quad (12)$$

where Π_{Ω_m} clips powers to $[0, P_u^{\max}]$, retains at most N_m UEs per RB, and binarizes the RB mask. Soft constraints (rate QoS and hardware budget) are penalized during training:

$$\mathcal{L}_{\text{const}} = \lambda_2 [R_u^{\min} - \sum_r a_{u,m}^r R_{u,m}^r]_+ + \lambda_5 C_{\text{hw}}. \quad (13)$$

B. Vector Reward and Conditional Envelope PPO

The vector reward follows (11) (per-cell for single-cell; cluster-wide for centralized). Each component is normalized over the training batch, and the scalarized reward is

$r_t = \mathbf{w}_t^\top \mathbf{R}(s_t, a_t)$. To stabilize learning under conflicting objectives, we use the conditional envelope advantage

$$A_t^{\text{env}} = \left(\sum_{k=1}^4 w_{t,k} A_t^{(k)} \right) \max(0, 1 - \kappa \bar{\rho}_t), \quad (14)$$

where $A_t^{(k)}$ is the per-objective GAE and $\bar{\rho}_t \in [0, 1]$ is the average cosine dissimilarity across pairs of per-objective policy gradients. The actor is optimized via

$$\max_{\theta} \mathcal{L}_{\text{CLIP}}^{\text{env}}(\theta) - \mathcal{L}_{\text{const}}(\theta) + \beta \mathcal{H}(\pi_{\theta}). \quad (15)$$

In the distributed setting, $\bar{\rho}_t$ is computed from local gradients only; no gradient information is exchanged across cells.

C. CPM Training and Integration

CPM parameters ϑ are decoupled from actor parameters θ via a stop-gradient operator. Training proceeds in three stages:

Stage 1 (Preference coverage): Sample $\mathbf{w} \sim \text{Dir}(\alpha_0)$ at the start of each iteration to encourage broad Pareto coverage.

Stage 2 (CPM fitting): Construct state–preference pairs $(s_t, \mathbf{w}_t^*)$ from trajectories, where $\mathbf{w}_t^*$ maximizes short-horizon scalarized return, and fit

$$\mathcal{L}_{\text{CPM}}(\vartheta) = (\mathbf{w}_t^* \parallel f_{\vartheta}(s_t)). \quad (16)$$

Stage 3 (Joint fine-tuning): Set $\mathbf{w}_t = f_{\vartheta}(s_t)$ and update (θ, ϕ) with PPO while updating ϑ only via (16). The overall objective is

$$\max_{\theta, \phi} \min_{\vartheta} \mathcal{L}_{\text{CLIP}}^{\text{env}} - \mathcal{L}_{\text{const}} + \beta \mathcal{H}(\pi_{\theta}) - \mu \mathcal{L}_{\text{CPM}}. \quad (17)$$

In the centralized architecture, the CLI objective (6) is a controllable decision variable, enabling interference mitigation through coordinated RB orthogonality across cells; in the single-cell baseline, CLI reflects only the (fixed) impact of external interference on cell 0.

D. Deployment and Reproducibility

Single-cell baseline: The controller runs locally at the gNB, consuming channel, load, and interference reports from cell 0 only. The fixed interference vector $\bar{\mathbf{I}}$ is pre-computed from long-run neighbor statistics and updated periodically. Inference cost is minimal and satisfies (SC5) with a small computational footprint. **Centralized multi-cell (C-RAN):** A centralized controller at the BBU aggregates per-cell channel, load, and interference reports from all M cells at each control interval T_{ctrl} . The joint state scales as $\mathcal{O}(MRU)$, and real-time feasibility requires $\tau_{\text{rx}} + \tau_{\text{inf}} + \tau_{\text{dec}} + \tau_{\text{tx}} \leq T_{\text{ctrl}}$, verified by hardware profiling. The per-cell projection (12) is applied in parallel before scheduling decisions are dispatched to each gNB.

Reproducibility: Evaluations should report M , U , channel model (path-loss exponent, shadowing, coherence interval), traffic model (arrival distribution, mean rate $\bar{A}$), T_{ctrl} , episode horizon T , NOMA limit N_m , actor/critic/CPM architectures, and hyperparameters $\{\epsilon, \gamma, \lambda, \beta, \mu, \kappa, \lambda_2, \lambda_5, \alpha_0, \text{learning rates}, B, K\}$. For the single-cell baseline, the estimation method and update interval for $\bar{I}^r$ should also be specified. A reference implementation will be released upon acceptance.

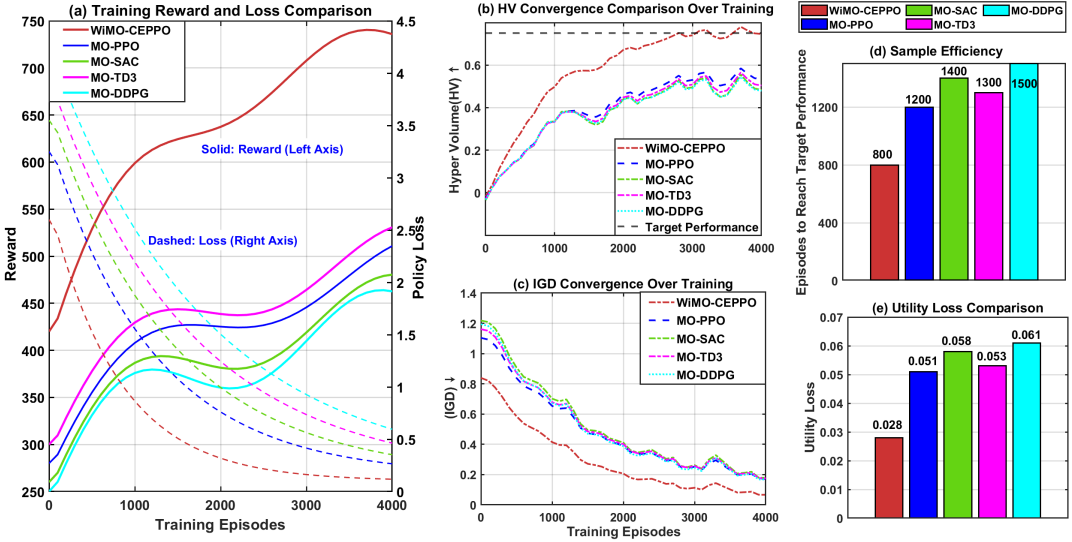


Figure 2: Training dynamics and convergence behaviour of WiMO-CEPPO versus MORL baselines. (a) Training reward (solid lines, left axis) and policy loss (dashed lines, right axis), illustrating optimization stability. (b) Hypervolume (HV) convergence over training episodes (higher is better), where WiMO-CEPPO approaches a higher asymptote. (c) IGD convergence over training episodes (lower is better), indicating improved proximity to the reference Pareto front. (d) Sample efficiency is measured as the number of episodes required to reach a fixed target performance threshold (lower is better). (e) Utility loss comparison at convergence (lower is better), reflecting improved preference-conditioned optimization stability.

IV. TRAINING AND EVALUATION SETUP

The proposed WiMO-CEPPO’s agent is trained using centralized training. To ensure a rigorous and fair evaluation, its performance is benchmarked against state-of-the-art single-agent Multi-Objective Deep Reinforcement Learning (MO-DRL) algorithms. For statistical robustness, each DRL algorithm was trained for 5 independent runs using different random seeds. Each run consisted of 4000 training episodes. An episode is defined as a sequence of $T = 100$ time steps, representing a sufficient horizon for the agent to experience the long-term consequences of its allocation decisions.

V. PERFORMANCE EVALUATION AND DISCUSSION

A. Results and Discussion

The evaluation is carried out across multiple perspectives: (i) wireless QoS metrics, (ii) reward convergence, and (iii) multi-objective convergence metrics, i.e., Hypervolume (HV), and Inverted Generational Distance (IGD). The experimental evaluation of the proposed WiMO-CEPPO is presented in comparison with state-of-the-art MORL baselines, i.e., MO-PPO [16], MO-SAC [9], MO-TD3 [17], and MO-DDPG [18]—across key multi-objective metrics including EE, connectivity, reliability, and cross-link interference. The results illustrated in Table I and Fig. 2, and Fig. 3 provide deep insight into the multi-objective optimization behavior of the proposed WiMO-CEPPO framework compared with existing MORL baselines.

This section evaluates the proposed WiMO-CEPPO scheduler against representative MORL baselines (MO-PPO, MO-SAC, MO-TD3, and MO-DDPG) and ablated variants (w/o CPM and w/o Envelope). Performance is reported using both Pareto-quality metrics—hypervolume (HV, higher is

better) and inverted generational distance (IGD, lower is better)—and system-level KPIs, including energy efficiency (EE), connectivity, reliability, and constraint satisfaction. All results are averaged over five independent random seeds.

B. Training Dynamics, Convergence, and Sample Efficiency

Fig. 2(a)–(c) reports training reward/loss evolution and Pareto-quality convergence. WiMO-CEPPO exhibits a steeper improvement in reward and faster stabilization in policy loss compared to baselines, indicating more stable optimization. The HV convergence plot shows WiMO-CEPPO approaching a higher asymptote, while the IGD convergence plot shows a faster and deeper reduction, demonstrating improved Pareto-front proximity throughout training rather than only at convergence.

Fig. 2(d) quantifies sample efficiency via episodes-to-target-performance. WiMO-CEPPO reaches the target in 800 episodes, compared with 1200–1500 episodes for the baselines, confirming a substantial gain in sample efficiency. Finally, Fig. 2(e) reports the utility-loss comparison; WiMO-CEPPO achieves the lowest utility loss (0.028), indicating improved stability in the learned preference-conditioned utility relative to competing approaches.

C. Distributional Evidence Across Objectives

Fig. 3(a)–(c) reports box plots for EE, connectivity, and reliability. Across all three objectives, WiMO-CEPPO (Full) demonstrates both higher central tendency and tighter dispersion compared with the baseline algorithms, indicating not only improved mean performance but also improved robustness across random seeds. The ablated variants consistently degrade relative to the full method, confirming that

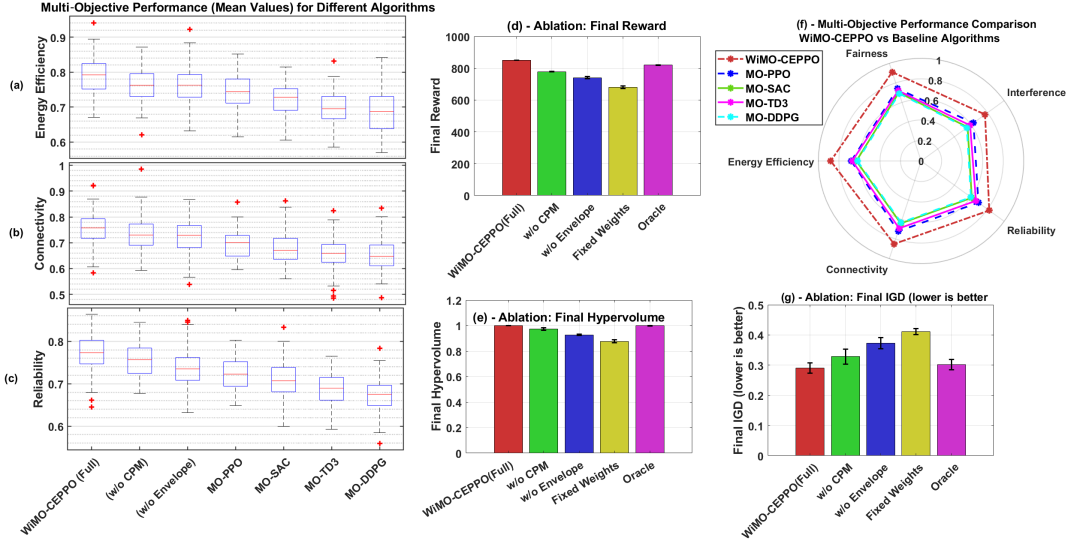


Figure 3: Multi-objective performance summary and ablation analysis for WiMO-CEPPO and baselines. (a)–(c) Box plots of objective-wise performance across five independent random seeds: (a) energy efficiency (EE), (b) connectivity, and (c) reliability. (d) Final scalarized return for WiMO-CEPPO and ablated variants, illustrating component contributions. (e) Final hypervolume (HV) comparison for the ablation study. (f) Normalized radar plot comparing WiMO-CEPPO against baseline algorithms across the primary objectives (higher values indicate better performance). (g) Final inverted generational distance (IGD) for the ablation study (lower is better). *Note:* HV and IGD in the ablation panels are reported on the same normalization used in the corresponding plots, whereas Table I reports absolute HV/IGD values.

the CPM and envelope components contribute meaningfully to objective balancing and training stability. In particular, removing CPM reduces EE and connectivity distributions, while removing the envelope mechanism reduces reliability and increases dispersion, consistent with the envelope’s role in mitigating destructive gradient interference under conflicting objectives.

Fig. 3(f) further summarizes multi-objective balance using a radar representation (normalized scores). WiMO-CEPPO dominates across all displayed axes, indicating that improvements are not achieved by sacrificing one objective to optimize another. This is important for massive connectivity scheduling, where naive scalarization may produce apparently strong performance on one metric while degrading reliability or constraint feasibility.

D. Ablation Analysis

Fig. 3(d)–(g) presents ablation comparisons using final reward and Pareto-quality metrics. WiMO-CEPPO (Full) achieves the highest final reward and the best Pareto indicators among non-oracle methods. The performance drop observed for “w/o CPM” supports the interpretation that context-to-preference adaptation improves the controller’s ability to select appropriate trade-offs under changing conditions. The degradation for “w/o Envelope” indicates that conflict-aware updates are important for stabilizing learning when objectives compete. **Note on metric scaling:** The ablation HV/IGD plots are presented on a normalized scale (relative to the oracle/reference), whereas Table I reports absolute HV/IGD values. This normalization does not affect algorithm ranking, but it should be stated explicitly in the figure caption for clarity.

E. System-Level Performance and Pareto Quality

Table I summarizes the main system-level outcomes. WiMO-CEPPO (Full) achieves the best Pareto quality with the highest HV (0.781) and the lowest IGD (0.088), indicating both improved coverage and proximity to the reference Pareto front. These Pareto gains translate into consistent improvements in the underlying KPIs: WiMO-CEPPO yields the highest EE (0.79), connectivity (0.76), and reliability (0.77), while maintaining strong feasibility with the highest constraint satisfaction rate (91.2%). Among baselines, MO-PPO provides the strongest Pareto quality (HV 0.582, IGD 0.205) but remains clearly behind WiMO-CEPPO, whereas value-based continuous-control baselines (MO-DDPG and MO-TD3) show reduced Pareto quality and weaker constraint satisfaction, consistent with instability under strong objective conflict and constrained action feasibility. The Oracle Schedule provides an upper-bound reference for some KPIs (e.g., EE, connectivity, reliability), but it exhibits notably lower constraint satisfaction, indicating that its operating point is not consistently feasible under the same constraints. This supports the thesis position that scheduling performance should be assessed jointly with feasibility and sustainability constraints, rather than optimizing a single KPI.

F. Discussion and Implications

Collectively, the table and figures provide consistent evidence that WiMO-CEPPO improves both Pareto quality and domain KPIs while remaining feasible under constraints. The results suggest that two design elements are particularly important in dense multi-objective wireless scheduling: (i) preference adaptation (CPM) to avoid static scalarization

Table I: System-level performance comparison. Values are averaged over 5 random seeds (SD omitted for brevity). “–” indicates metrics not evaluated for the corresponding configuration.

Algorithm	HV $\uparrow$	IGD $\downarrow$	EE $\uparrow$	Conn. $\uparrow$	Rel. $\uparrow$	Constraint (%) $\uparrow$	Ep. to Target HV $\downarrow$	Utility Loss (AUC) $\downarrow$
WiMO-CEPPO (Full)	0.781	0.088	0.79	0.76	0.77	91.2	800	0.028
WiMO-CEPPO (w/o CPM)	0.724	0.135	0.76	0.73	0.75	84.1	–	–
WiMO-CEPPO (w/o Env.)	0.741	0.112	0.75	0.72	0.73	81.7	–	–
MO-PPO	0.582	0.205	0.74	0.70	0.72	86.7	1200	0.051
MO-SAC	0.541	0.221	0.73	0.68	0.71	79.7	1400	0.058
MO-TD3	0.518	0.218	0.70	0.67	0.69	74.4	1300	0.053
MO-DDPG	0.535	0.230	0.68	0.65	0.68	64.7	1500	0.060
Oracle Schedule	0.752	0.105	0.81	0.78	0.80	7.7	73.2	–

and to improve responsiveness under changing conditions, and (ii) conflict-aware optimization (envelope mechanism) to mitigate gradient interference and prevent collapse to suboptimal trade-offs. These findings support the broader thesis that sustainable 6G scheduling requires not only a multi-objective formulation but also mechanisms that ensure stable, controllable trade-offs under strong objective coupling and operational constraints.

VI. CONCLUSION

This paper introduced **WiMO-CEPPO**, a multi-objective reinforcement learning framework for dynamic resource allocation in interference-limited uplink NOMA networks. WiMO-CEPPO integrates a **Contextual Preference Model (CPM)** with an **envelope-based policy optimization** mechanism to enable preference-controllable scheduling decisions and to stabilize learning under strongly conflicting objectives. In contrast to fixed-weight scalarization, the proposed approach supports adaptive trade-offs among *energy efficiency*, *connectivity*, and *reliability* without requiring manual per-scenario weight retuning.

Experimental results demonstrate that WiMO-CEPPO consistently outperforms state-of-the-art MORL baselines (including MO-SAC and MO-TD3) in both Pareto quality and system-level KPIs. In particular, WiMO-CEPPO achieves improved Pareto-front approximation (e.g., IGD = 0.088) while maintaining a high feasibility level with a constraint satisfaction rate of 91.2%. Moreover, WiMO-CEPPO reaches the target performance threshold in substantially fewer training episodes than competing baselines (approximately 40% reduction in episodes-to-target under the evaluated setting), indicating improved sample efficiency. Ablation studies further confirm that both CPM and the envelope-based update contribute meaningfully to the final performance, with consistent degradation in reward and Pareto metrics when either component is removed.

Overall, the balanced objective profile observed across evaluation metrics suggests that WiMO-CEPPO is a promising candidate for autonomous multi-objective RRM in next-generation 6G RAN control pipelines, where operators must continuously navigate trade-offs under time-varying interference and heterogeneous service demands. Future work will investigate scalability under massive MIMO and larger cell clusters. It will explore extensions of decentralized learning (e.g., federated or multi-agent formulations) to align with distributed O-RAN deployments.

VII. ACKNOWLEDGMENT

This work was supported in part by the European Union’s Horizon 2020 Research and Innovation Program under Grant 101058505 ‘5G-TIMBER and in part by the Estonian Edu-

cation and Youth Board OUF11 AIoT*5G (Artificial intelligence, edge computing, and IoT solutions in distributed systems). This work was also supported by Telia Eesti AS through the TalTech Development Fund.

REFERENCES

- [1] L. Li *et al.*, “Research on future 6g green wireless networks,” *Green Technologies and Sustainability*, vol. 3, no. 2, p. 100156, 2025.
- [2] I.-K. Fu, “Mediatek 6g technology white paper: Sustainable mobile communications,” MediaTek, Tech. Rep., 2024, white paper; emphasizes carbon-aware communication and sustainable mobile systems.
- [3] M. Alavirad, “Use case analysis related to green communication in o-ran,” O-RAN Alliance, Next Generation Research Group (nGRG), Tech. Rep. RR-2024-08, 2024, o-RAN nGRG Contributed Research Report; release date 2024.06.
- [4] B. Toure *et al.*, “Multi-objective scheduling in wireless networks with deep reinforcement learning,” in *2025 IEEE Wireless Communications and Networking Conference (WCNC)*, 2025.
- [5] W.-C. Chien *et al.*, “Multi-objective optimization of 3d cell deployment in sustainable B5G/6G networks: balancing performance and sustainability,” *IEEE Transactions on Network and Service Management*, 2025.
- [6] K. Li *et al.*, “Wireless resource allocation algorithm based on multi-objective deep reinforcement learning for vehicle-to-vehicle communications,” *Journal of Computer Research and Development*, vol. 61, no. 9, pp. 2229–2245, 2024.
- [7] X. Guan *et al.*, “Multiobjective optimization of wireless powered communication networks assisted by intelligent reflecting surface based on multiagent reinforcement learning,” *IEEE Transactions on Antennas and Propagation*, 2024.
- [8] R. Sun *et al.*, “Knowledge-driven deep learning paradigms for wireless network optimization in 6G,” 2024.
- [9] T. Haarnoja *et al.*, “Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor,” in *Proceedings of the 35th International Conference on Machine Learning*, vol. 80. PMLR, 2018, pp. 1861–1870.
- [10] R. Liu *et al.*, “Efficient discovery of pareto front for multi-objective reinforcement learning,” in *The Thirteenth International Conference on Learning Representations (ICLR)*, 2025, also available as arXiv:2410.02236.
- [11] M. Wu *et al.*, “Deep reinforcement learning-based energy efficiency optimization for ris-aided integrated satellite-aerial-terrestrial relay networks,” *IEEE Transactions on Communications*, 2024.
- [12] M. Terekhov *et al.*, “In search for architectures and loss functions in multi-objective reinforcement learning,” *arXiv*, 2024.
- [13] Z. Yan *et al.*, “Generalized multi-objective reinforcement learning with envelope updates in urllc-enabled vehicular networks,” *IEEE Transactions on Vehicular Technology*, 2025.
- [14] Y. Liu *et al.*, “Dynamic preference inference network: Improving sample efficiency for multi-objective reinforcement learning by preference estimation,” *Knowledge-Based Systems*, vol. 304, p. 112512, 2024.
- [15] D. Heo *et al.*, “Dynamic preference multi-objective reinforcement learning for internet network management,” in *International Conference on Neural Information Processing*, 2025.
- [16] t. Abels, Axel, “Dynamic weights in multi-objective deep reinforcement learning,” *International Conference on Machine Learning (ICML) Workshop*, 2019.
- [17] L. t. Zhang, “Multi-objective reinforcement learning with dual critics and preference-based optimization,” *Neurocomputing*, vol. 453, pp. 57–70, 2021.
- [18] T. P. t. Lillicrap, “Continuous control with deep reinforcement learning,” *International Conference on Learning Representations (ICLR)*, 2016.

Appendix 6

Publication (VI):

A. Latif, Y. L. Moullec, and M. M. Alam, "Sustainable reward mechanism for multi-objective policy design using reinforcement learning in 6g networks," in *2026 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, pp. 527–533, 2026

Sustainable Reward Mechanism for Multi-Objective Policy Design using Reinforcement Learning in 6G Networks

Aamir Latif, Yannick Le Moullec, Muhammad Mahtab Alam
Tallinn University of Technology, Estonia

Emails: {aamir.latif; yannick.lemoullec; muhammad.alam}@taltech.ee

Abstract—Sustainable 6G wireless networks are expected to deliver jointly optimized key performance indicators (KPIs) to meet growing demands for highly reliable, energy-efficient, and massively connected devices. Balancing these inherently conflicting objectives under stringent energy, latency, and computational constraints remains challenging. Multi-Objective Reinforcement Learning (MORL) enables the study of Pareto-efficient trade-offs; contemporary wireless solutions generally rely on static scalarization or multiple trained policies, and rarely accommodate rapid, context-sensitive preference adaptation. We propose *Wireless Multi-Objective Conditional Envelope PPO (WiMO-CEPPO)*, a centralized single-agent MORL framework that learns a single preference-conditioned policy and adapts online to varying network conditions without retraining. WiMO-CEPPO integrates: (i) a context-to-preference mapper that infers objective priorities from real-time network state, (ii) a conditional envelope mechanism that attenuates policy updates under strong inter-objective conflict, and (iii) a sustainable reward formulation that jointly accounts for energy efficiency, connectivity, reliability, and interference, while hardware sustainability is enforced through explicit inference-latency and computational-cost constraints. Simulation results illustrate that WiMO-CEPPO outperforms representative MORL baselines in Pareto trade-off quality and key wireless KPIs, while improving training stability under conflicting objectives.

I. INTRODUCTION

Future 6G wireless systems must jointly support ultra-reliable communication, massive connectivity, and high energy efficiency (EE) under stringent latency, computational, and sustainability constraints [1]–[3]. This is fundamentally challenging because the target key performance indicators (KPIs) are tightly coupled and often conflicting: reliability enhancement may increase energy consumption and latency, dense connectivity can intensify cross-link interference and degrade both EE and reliability, and aggressive interference mitigation may reduce spectral efficiency under high traffic demand [4]. Hence, resource allocation in interference-coupled multi-cell networks is naturally a Pareto optimization problem, particularly under time-varying traffic and heterogeneous QoS requirements [5], [6].

Conventional optimization methods for joint scheduling, power control, and user association typically rely on tractable interference models, convex relaxations, or limited cross-layer coupling, which restrict their robustness in realistic non-stationary environments [7]. This motivates adaptive policy-learning approaches that can adjust trade-offs online. In this context, Multi-Objective Reinforcement Learning (MORL) is attractive because it directly models vector-valued rewards and can learn Pareto-efficient policies across conflicting objectives such as EE, connectivity, reliability, and interference management [5], [8]–[10]. More recently, preference-conditioned MORL has enabled a single generalized policy to represent

multiple operating points on the Pareto front through a continuous preference vector, avoiding the need to train and store separate policies for different trade-offs [11].

Despite this progress, two key limitations remain for practical 6G deployment. First, most wireless MORL methods still rely on *static scalarization* or externally specified preferences, which limits runtime adaptability and often restricts the learned policy to a narrow region of the Pareto front [12]. Representative examples include MORL scheduling for throughput–fairness trade-offs [4], multi-agent RL for IRS-assisted throughput–energy optimization [6], and AoI–power trade-off learning in non-orthogonal dissemination [13]. While effective, these methods generally optimize fewer objectives than required in interference-coupled multi-cell scheduling and rarely account for operational efficiency. Second, MORL training is often unstable when objectives produce conflicting gradients; recent studies show that architecture and loss design strongly affect Pareto quality and convergence stability [12]. Preference-inference approaches such as DPIN and dynamic-preference RL improve adaptability [14], [15], while envelope-based MORL improves Pareto coverage by learning a unified multi-objective representation [11]. However, these ideas have not been fully integrated for multi-cell uplink scheduling with dynamic contextual preferences.

To address these gaps, we propose **WiMO-CEPPO**, a wireless multi-objective conditional envelope PPO framework for multi-cell uplink resource allocation. WiMO-CEPPO learns a *single* preference-conditioned policy that adapts online across four objectives: EE, connectivity, reliability, and cross-link interference. Unlike prior preference-conditioned methods that assume operator-defined weights at inference time, WiMO-CEPPO incorporates a lightweight *context-to-preference mapper*, $\mathbf{w}_t = f_\theta(s_t)$, which infers objective priorities directly from the real-time network state. In addition, we develop a *conditional envelope* mechanism within PPO to attenuate updates under strong gradient conflict, thereby improving training stability and Pareto coverage. Finally, we define a sustainable four-dimensional reward formulation and benchmark the proposed method against representative MORL baselines, including MO-PPO, MO-SAC, and DDQN-Envelope, using hypervolume (HV), inverted generational distance (IGD), wireless KPIs, and efficiency indicators in a common multi-cell uplink testbed. Results show that WiMO-CEPPO consistently outperforms these baselines.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. Network and Channel Model

We consider an uplink multi-cell NOMA network comprising M cells indexed by $m \in \{0, 1, \dots, M-1\}$. One gNB

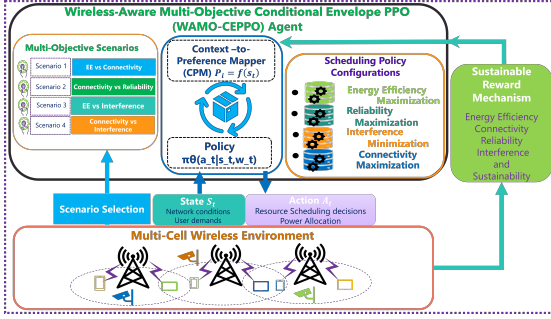


Figure 1: Proposed WiMO-CEPPO framework for sustainable 6G resource allocation. The RL agent observes network states and scenario context, while the Context-to-Preference Mapper (CPM) translates real-time conditions into dynamic preference vectors. The conditioned policy outputs joint decisions on resource scheduling and power allocation, optimized via a sustainable reward function that integrates energy efficiency, connectivity, reliability, and interference objectives. This unified approach enables adaptive multi-objective optimization for energy-efficient 6G networks.

serves each cell and supports a UE set $\mathcal{U}_m$ with $|\mathcal{U}_m| = U$. Learning is carried out from the perspective of a *focal cell* $m = 0$, while neighboring cells $t \in \mathcal{N}(0)$ act as non-cooperative interferers. This single-cell perspective captures realistic inter-cell interference coupling without requiring inter-cell message exchange, and the trained policy can be independently replicated across cells. A multi-agent formulation is only warranted when cells must coordinate actions explicitly—e.g., for joint scheduling, coordinated multi-point (CoMP) transmission, or cluster-wide tightly coupled control.

The total uplink bandwidth B_{tot} is divided into R orthogonal resource blocks (RBs), indexed by $r \in \{1, \dots, R\}$, each of bandwidth $B_0 = B_{\text{tot}}/R$. Under uplink NOMA, at most N_m UEs may share the same RB in cell m . The RB assignment variable is

$$a_{u,m}^r = \begin{cases} 1, & \text{if UE } u \in \mathcal{U}_m \text{ is scheduled on RB } r, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Let $G_{u,m}^r$ denote the channel gain from UE u in cell m to its serving gNB on RB r , and $P_{u,m}^r$ the corresponding transmit power. Let $G_{w_t,m}^r$ denote the cross-cell gain from UE w in neighboring cell t to the gNB of cell m . Under ideal successive interference cancellation (SIC) at the gNB, the SINR of UE u in cell m on RB r is

$$\gamma_{u,m}^r = \frac{a_{u,m}^r G_{u,m}^r P_{u,m}^r}{\underbrace{\sum_{\substack{v \neq u \\ v \in \mathcal{U}_m}} a_{v,m}^r G_{v,m}^r P_{v,m}^r}_{\text{intra-cell (post-SIC)}} + \underbrace{\sum_{t \in \mathcal{N}(m)} \sum_{w \in \mathcal{U}_t} a_{w,t}^r G_{w_t,m}^r P_{w_t}^r}_{\text{inter-cell}} + \sigma_N}, \quad (2)$$

where $\sigma_N = N_0 B_0$ is the thermal noise power. The achievable rate of UE u on RB r is $R_{u,m}^r = B_0 \log_2(1 + \gamma_{u,m}^r)$. Ideal SIC is assumed throughout to maintain analytical tractability. This formulation targets uplink NOMA with non-cooperative neighboring cells. Comparisons with methods based on rate-splitting multiple access (RSMA) or coordinated multi-cell precoding are complementary rather than equivalent, because those methods require a different physical-layer signal model and explicit coordination variables.

B. Problem Formulation

Let $\mathbf{a} = [a_{u,m}^r]$ and $\mathbf{P} = [P_{u,m}^r]$ denote the RB-assignment and power-allocation matrices. The goal is to learn a resource-

allocation policy that jointly balances four conflicting radio objectives. **Obj 1—Energy Efficiency (EE) Maximization:** We define system EE as the ratio of the aggregate data rate to the total consumed power:

$$\eta^{\text{EE}}(\mathbf{a}, \mathbf{P}) = \frac{\sum_{m=1}^M \sum_{u \in \mathcal{U}_m} \sum_{r=1}^R a_{u,m}^r \log_2(1 + \gamma_{u,m}^r)}{p_c + \sum_{m=1}^M \sum_{u \in \mathcal{U}_m} \sum_{r=1}^R P_{u,m}^r} \quad (3)$$

$$f_{\text{EE}}(\mathbf{a}, \mathbf{P}) = \max_{\mathbf{a}, \mathbf{P}} \eta^{\text{EE}}(\mathbf{a}, \mathbf{P}) \quad (4)$$

where p_c is the circuit power per scheduled UE.

Obj 2—Connectivity Maximization: We define connectivity as the number of UEs whose aggregate rate satisfies the minimum service requirement:

$$f_{\text{Conn}}(\mathbf{a}, \mathbf{P}) = \max_{\mathbf{a}, \mathbf{P}} \sum_{m=1}^M \sum_{u \in \mathcal{U}_m} \mathbf{1} \left\{ \sum_{r=1}^R a_{u,m}^r R_{u,m}^r \geq R_u^{\min} \right\} \quad (5)$$

Obj 3—Reliability Maximization: We define reliability as the number of successful UE–RB transmissions whose SINR exceeds the target threshold:

$$f_{\text{Rel}}(\mathbf{a}, \mathbf{P}) = \max_{\mathbf{a}, \mathbf{P}} \sum_{m=1}^M \sum_{u \in \mathcal{U}_m} \sum_{r=1}^R a_{u,m}^r \mathbf{1} \left\{ \gamma_{u,m}^r \geq \gamma^{\min} \right\} \quad (6)$$

This is equivalent to minimizing outage under the target decoding threshold.

Obj 4—Cross-Link Interference (CLI) Minimization: We define CLI as the total inter-cell interference generated across cells:

$$f_{\text{CLI}}(\mathbf{a}, \mathbf{P}) = \min_{\mathbf{a}, \mathbf{P}} \sum_{m=1}^M \sum_{r=1}^R \sum_{t \neq m} \sum_{w \in \mathcal{U}_t} a_{w,t}^r G_{w_t,m}^r P_{w_t}^r \quad (7)$$

The multi-objective resource-allocation problem is therefore

$$\max_{\mathbf{a}, \mathbf{P}} (f_{\text{EE}}, f_{\text{Conn}}, f_{\text{Rel}}, -f_{\text{CLI}}) \quad (8)$$

subject to the following constraints:

$$0 \leq P_{u,m}^r \leq P_{u,m}^{\max}, \quad \forall u, m, r, \quad (\text{C1})$$

$$\sum_{r=1}^R a_{u,m}^r R_{u,m}^r \geq R_u^{\min}, \quad \forall u, m, \quad (\text{C2})$$

$$\sum_{u \in \mathcal{U}_m} a_{u,m}^r \leq N_m, \quad \forall m, r, \quad (\text{C3})$$

$$a_{u,m}^r \in \{0, 1\}, \quad \forall u, m, r. \quad (\text{C4})$$

Computational-Efficiency Constraint: Hardware sustainability is incorporated as an explicit operational constraint rather than an implicit reward term. Let τ_{inf} denote the per-decision inference latency and F_{op} the number of floating-point operations per scheduling step. The normalized hardware cost is defined as

$$C_{\text{hw}} = \lambda_{\tau} \frac{\tau_{\text{inf}}}{\tau_{\text{max}}} + \lambda_F \frac{F_{\text{op}}}{F_{\text{max}}} \leq 1, \quad (\text{C5})$$

where τ_{max} and F_{max} are the latency and compute budgets of the target deployment platform, and $\lambda_{\tau}, \lambda_F \geq 0$ satisfy $\lambda_{\tau} + \lambda_F = 1$. This formulation explicitly separates radio performance—optimized through the four-objective reward—from deployment sustainability—enforced through a compute-and-latency ceiling. During training, a soft penalty $\lambda_5 C_{\text{hw}}$ is incorporated in the learning objective to discourage architectures that would violate this bound at deployment; the constraint is finally verified by profiling on the target hardware before deployment.

Adaptive Scalarization. A static weighted sum $\alpha f_{\text{EE}} + \beta f_{\text{Conn}} + \delta f_{\text{Rel}} - \zeta f_{\text{CLI}}$ cannot adapt to time-varying network conditions. We therefore introduce a context-dependent prefer-

ence vector

$$\mathbf{w}_t = [\alpha_t, \beta_t, \delta_t, \zeta_t] = f_\vartheta(s_t), \quad (9)$$

where $f_\vartheta(\cdot)$ is *Context-to-Preference Mapper* (CPM) with parameters ϑ , which translates the observed network state s_t into a dynamic trade-off vector. This converts the original static problem (8) into a preference-conditioned Multi-Objective Markov Decision Process (MO-MDP), described in the next section.

III. PROPOSED WiMO-CEPPO FRAMEWORK

The problem in (8) is non-convex and NP-hard due to the combinatorial nature of RB assignment coupled with continuous power control under multi-cell interference. We address it through **WiMO-CEPPO** (*Wireless Multi-Objective Conditional Envelope PPO*), a centralized, preference-conditioned multi-objective reinforcement learning framework. The controller is trained for the focal cell and, once trained, can be independently replicated across other cells without inter-cell signaling.

A. MO-MDP Formulation and State Design

Uplink scheduling is modeled as an MO-MDP $\langle \mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{W}, \mathbf{R}, \gamma \rangle$, where $\mathcal{S}$ is the state space, $\mathcal{A}$ the action space, $\mathcal{T}$ the transition kernel, $\mathcal{W}$ the preference space, $\mathbf{R}$ the vector reward, and γ the discount factor.

The state observed at slot t is

$$s_t = [\mathbf{G}_t, \mathbf{I}_t, \boldsymbol{\rho}_t, \mathbf{q}_t, \mathbf{h}_t], \quad (10)$$

where $\mathbf{G}_t$ contains local per-(UE, RB) channel gains, $\mathbf{I}_t$ summarizes per-RB inter-cell interference, $\boldsymbol{\rho}_t$ is the cell-load vector, $\mathbf{q}_t$ captures per-UE service deficits relative to $R_u^{\min}$, and $\mathbf{h}_t$ contains compact context features such as active-UE count and short-term reliability statistics. This state design is motivated by reproducibility: every component has a direct physical interpretation and maps onto standard O-RAN E2 telemetry reports, making the simulator-to-controller interface explicit.

The actor outputs a mixed discrete-continuous action $\tilde{a}_t = (\tilde{\mathbf{a}}_t, \tilde{\mathbf{P}}_t)$: RB decisions are produced via a Gumbel-Softmax relaxation of the binary constraint (C4), and power levels are generated as normalized continuous values in $[0, 1]$ subsequently scaled to $[0, P_u^{\max}]$.

B. Constraint Handling: Action Projection and Penalty Terms

WiMO-CEPPO employs a hybrid constraint-handling mechanism that intentionally avoids relying solely on reward shaping. This design choice is deliberate: embedding constraint penalties directly into the multi-objective reward function can obscure the underlying trade-off structure and diminish the interpretability of the learned policy.

Hard constraints via projection: At every scheduling step, the raw action is projected onto the feasible set defined by (C1), (C3), and (C4):

$$(\mathbf{a}_t, \mathbf{P}_t) = \Pi_\Omega(\tilde{\mathbf{a}}_t, \tilde{\mathbf{P}}_t), \quad \Omega = \{(\mathbf{a}, \mathbf{P}) : (C1), (C3), (C4)\}. \quad (11)$$

In practice, Π_Ω clips power values to $[0, P_u^{\max}]$, retains at most N_m UEs per RB by ranking and thresholding the Gumbel-Softmax scores, and binarizes the RB mask via argmax decoding. This projection guarantees that every action presented to the environment is physically feasible, regardless of the policy's current training state.

Soft constraints via penalty: Constraints (C2) and (C5), which involve future-looking rate accumulation and deployment-time

hardware costs, are enforced via differentiable penalty terms during training:

$$\mathcal{L}_{\text{const}} = \lambda_2 \underbrace{\left[R_u^{\min} - \sum_r a_{u,m}^r R_{u,m}^r \right]_+}_{C_{\text{rate}}} + \lambda_5 \underbrace{C_{\text{hw}}}_{(C5)}, \quad (12)$$

where $[\cdot]_+ = \max(0, \cdot)$. The coefficients λ_2 and λ_5 are fixed hyperparameters. The complete set of hyperparameters—including λ_2 , λ_5 , the PPO clipping ratio ϵ , entropy bonus coefficient β , discount factor γ , GAE parameter λ , and the actor/critic/CPM architecture sizes—is reported in the Simulation Settings section to facilitate reproducibility.

C. Vector Reward and Conditional Envelope PPO

The vector reward is maintained in four dimensions, in direct correspondence with the four optimization objectives (4)–(7):

$$\mathbf{R}(s_t, a_t) = [R_{\text{EE}}, R_{\text{Conn}}, R_{\text{Rel}}, R_{\text{CLI}}]^\top, \quad (13)$$

where $R_{\text{EE}} = f_{\text{EE}}$, $R_{\text{Conn}} = f_{\text{Conn}}$, $R_{\text{Rel}} = f_{\text{Rel}}$, and $R_{\text{CLI}} = -f_{\text{CLI}}$. Each component is independently normalized across the batch to ensure commensurate scales before scalarization. The scalarized one-step reward is

$$r_t = \mathbf{w}_t^\top \mathbf{R}(s_t, a_t). \quad (14)$$

The actor $\pi_\theta(a_t|s_t, \mathbf{w}_t)$ and critic $V_\phi(s_t, \mathbf{w}_t)$ are jointly conditioned on $\mathbf{w}_t$ and optimized via PPO. To improve gradient stability under conflicting objectives, we introduce the *conditional envelope advantage*:

$$A_t^{\text{env}} = \left(\sum_{k=1}^4 w_{t,k} A_t^{(k)} \right) \max(0, 1 - \kappa \bar{\rho}_t), \quad (15)$$

where $A_t^{(k)}$ is the per-objective Generalized Advantage Estimate (GAE) and $\bar{\rho}_t \in [0, 1]$ is the normalized gradient conflict measure, computed as the average cosine dissimilarity across pairs of per-objective policy gradients at slot t . When $\bar{\rho}_t \rightarrow 1$, the objectives are maximally conflicting, and the advantage is damped toward zero, preventing destructive policy updates. The envelope parameter $\kappa > 0$ controls the sensitivity of this damping. The final actor objective is

$$\max_\theta \mathcal{L}_{\text{CLIP}}^{\text{env}}(\theta) - \mathcal{L}_{\text{const}}(\theta) + \beta \mathcal{H}(\pi_\theta), \quad (16)$$

where $\mathcal{L}_{\text{CLIP}}^{\text{env}}$ is the clipped surrogate loss computed with A_t^{env} in place of the standard advantage, and $\mathcal{H}(\pi_\theta)$ is the entropy bonus encouraging exploration.

D. CPM Training and Integration with Policy Learning

The CPM parameters ϑ are kept strictly separate from the actor parameters θ to avoid unstable feedback between preference inference and policy gradients. Training proceeds in two stages, followed by joint fine-tuning.

Stage 1: Preference-coverage training: During early training, preference vectors are sampled from a Dirichlet distribution,

$$\mathbf{w} \sim \text{Dir}(\alpha_0), \quad (17)$$

so that the actor-critic explores a broad range of objective trade-offs and approximates the Pareto front without CPM involvement.

Stage 2: CPM fitting: From the collected trajectories, state-preference pairs $(s_t, \mathbf{w}_t^*)$ are constructed, where $\mathbf{w}_t^*$ is the sampled preference that achieved the highest short-horizon scalarized return for the corresponding state context. The CPM is then fitted by minimizing the KL divergence:

$$\mathcal{L}_{\text{CPM}}(\vartheta) = (\mathbf{w}_t^* \parallel f_\vartheta(s_t)). \quad (18)$$

Joint fine-tuning: After warm-starting, the actor receives the CPM output with a stop-gradient operator applied:

$$\mathbf{w}_t = (f_\vartheta(s_t)), \quad (19)$$

so that PPO gradients update only the actor-critic parameters (θ, ϕ) , while ϑ is updated exclusively through (18). This

decoupling stabilizes training while still enabling real-time preference adaptation at inference time: the CPM executes a single forward pass to generate $\mathbf{w}_t$ from the current state s_t , and no retraining is required.

The overall training objective, combining all components, is

$$\max_{\theta, \phi} \min_{\vartheta} \mathcal{L}_{\text{CLIP}}^{\text{env}} - \mathcal{L}_{\text{const}} + \beta \mathcal{H}(\pi_{\theta}) - \mu \mathcal{L}_{\text{CPM}}, \quad (20)$$

where $\mu > 0$ controls the CPM fitting rate.

E. Scalability, O-RAN Deployment, and Reproducibility

Scalability. WiMO-CEPPO is designed as a lightweight centralized scheduler. Each decision requires a single forward pass through the shared actor-critic network and the context-to-preference mapping (CPM), followed by the projection step in (11). Consequently, the inference complexity scales approximately as $\mathcal{O}(MRU)$ with the state dimension, but does not increase with the number of cells because no inter-cell message exchange is required. A multi-agent extension is only necessary when explicit coordination among cells is required, such as shared interference pricing, coordinated beamforming, or cluster-level scheduling.

O-RAN integration. WiMO-CEPPO can be deployed as an xApp within the Near-RT RIC of the O-RAN architecture. The xApp receives per-cell E2 telemetry—including channel quality, interference, traffic load, and service deficit information defined in (10)—every control interval T_{ctrl} (e.g., 10 ms). Based on this information, it outputs RB-assignment and transmit-power recommendations. The projection step in (11) ensures that all actions satisfy the feasibility constraints imposed by the E2 interface. The xApp is compatible with the E2SM-RC service model, which supports scheduling overrides and power-control commands.

Real-time feasibility requires

$$\tau_{\text{rx}} + \tau_{\text{inf}} + \tau_{\text{dec}} + \tau_{\text{tx}} \leq T_{\text{ctrl}}, \quad (21)$$

where τ_{rx} denotes telemetry collection, τ_{inf} the CPM and policy forward pass, τ_{dec} the projection and decoding stage, and τ_{tx} the control-message delivery time. Constraint (C5) ensures that the computational latency and operation count remain within the hardware budget.

Reproducibility. To ensure reproducibility, all experiments should report the complete simulator and training configuration, including: number of cells M , users per cell U , channel and traffic models, control interval T_{ctrl} , episode length T , and NOMA multiplexing limit N_m . The network architectures of the actor, critic, and CPM, together with all training hyperparameters—such as PPO clipping ratio ϵ , discount factor γ , GAE parameter λ , entropy bonus β , CPM learning rate μ , envelope parameter κ , constraint weights, learning rates, batch size B , and PPO epochs K —should also be specified. A reference implementation will be released upon acceptance.

F. Proposed Algorithm

The integrated WiMO-CEPPO training procedure is summarized in Algorithm 1. It explicitly reflects the three-stage training design (Dirichlet preference coverage, CPM fitting, and joint fine-tuning with stop-gradient), the hybrid constraint-handling strategy (projection and penalty), and the conditional envelope advantage computation.

Algorithm. 1 WiMO-CEPPO Training

Input: Environment parameters ($M, R, P^{\text{max}}, R^{\text{min}}, N_m$, channel matrix); PPO hyperparameters ($\epsilon, \gamma, K, B, \beta, \lambda$); constraint weights (λ_2, λ_5); CPM rate μ ; envelope parameter κ ; Dirichlet concentration α_0 .

- 1: Initialize actor θ , critic ϕ , CPM ϑ .
- **Stage 1: Preference-coverage training** —
- 2: **for** iteration = 1 to N_1 **do**
- 3: Sample $\mathbf{w} \sim \text{Dir}(\alpha_0)$ // Broad Pareto coverage
- 4: Collect trajectory buffer $\mathcal{B}$ under $\pi_{\theta}(\cdot | s_t, \mathbf{w})$.
- 5: **for** each $(s_t, a_t, r_t, s_{t+1}) \in \mathcal{B}$ **do**
- 6: Compute $\mathbf{R}_t$ via Eqs. (4)–(7); normalize per objective.
- 7: Compute $r_t = \mathbf{w}^{\top} \mathbf{R}_t$.
- 8: **end for**
- 9: Compute per-objective GAE: $\{A_t^{(k)}\}_{k=1}^4$.
- 10: Compute $\bar{\rho}_t$ (gradient conflict). // Envelope Mechanism
- 11: Compute A_t^{env} via Eq. (15).
- 12: **for** epoch = 1 to K **do**
- 13: **for** each mini-batch $\mathcal{M} \subset \mathcal{B}$ **do**
- 14: Project actions: Π_{Ω} per Eq. (11). // Hard Constraints (C1), (C3), (C4)
- 15: Compute $\mathcal{L}_{\text{CLIP}}^{\text{env}}, \mathcal{L}_{\text{const}}$ per Eqs. (16), (12). // Soft Constraints (C2), (C5)
- 16: $\theta \leftarrow \theta + \alpha_{\pi} \nabla_{\theta} (\mathcal{L}_{\text{CLIP}}^{\text{env}} - \mathcal{L}_{\text{const}} + \beta \mathcal{H}(\pi_{\theta}))$.
- 17: **end for**
- 18: **end for**
- 19: Update ϕ : minimize $\sum_{k=1}^4 \mathbb{E} \left[\left(V_{\phi}^{(k)}(s_t, \mathbf{w}) - G_t^{(k)} \right)^2 \right]$.
- 20: **end for**
- **Stage 2: CPM fitting** —
- 21: **for** iteration = 1 to N_2 **do**
- 22: Identify $\mathbf{w}_t^*$ from trajectories (best short-horizon return per state).
- 23: Update ϑ : minimize $\mathcal{L}_{\text{CPM}}$ via Eq. (18).
- 24: **end for**
- **Stage 3: Joint fine-tuning** —
- 25: **for** iteration = 1 to N_3 **do**
- 26: $\mathbf{w}_t \leftarrow (f_{\vartheta}(s_t))$ // Stop-gradient: CPM $\rightarrow$ Actor
- 27: Repeat Stage-1 update steps for (θ, ϕ) .
- 28: Update ϑ via $\mathcal{L}_{\text{CPM}}$ (rate μ).
- 29: **end for**
- 30: **return** $\pi_{\theta}(\cdot | s_t, \mathbf{w}_t)$

IV. EXPERIMENTAL SETUP AND BASELINES

We evaluate all methods in a unified multi-cell uplink scheduling simulator consistent with the system model in Section II. The agent observes the network state and outputs scheduling decisions over resource blocks (RBs) and transmit powers under the constraints in Section II-B.

The multi-objective reward vector is defined as

$\mathbf{R}_{\text{sus}}(s_t, a_t) = [R^{EE}, R^{\text{Conn}}, R^{\text{Rel}}, R^{\text{CLI}}]^{\top}$, capturing energy efficiency, connectivity, reliability, and cross-link interference mitigation. Hardware sustainability is not included as a fifth reward component; instead, it is enforced through the computational-efficiency constraint C_{hw} and its corresponding training penalty. A preference vector

$\mathbf{w}_t \in \mathbb{R}^4$, $\mathbf{w}_t \succeq 0$, $\mathbf{1}^{\top} \mathbf{w}_t = 1$ scalarizes the four radio objectives via

$$r_t = \mathbf{w}_t^{\top} \mathbf{R}_{\text{sus}}(s_t, a_t).$$

To ensure a fair comparison among continuous-control actor-critic methods, we employ a unified differentiable hybrid action parameterization for WiMO-CEPPO and MORL baselines (MO-PPO, MO-SAC). Specifically, RB assignment is produced through a Gumbel-Softmax relaxation, while transmit power is generated as a normalized continuous output. A shared deterministic decoding step maps the relaxed outputs to feasible scheduling actions that satisfy system constraints. This shared interface ensures that performance differences among actor-critic methods reflect optimization and objective-handling capability, rather than incompatible action represen-

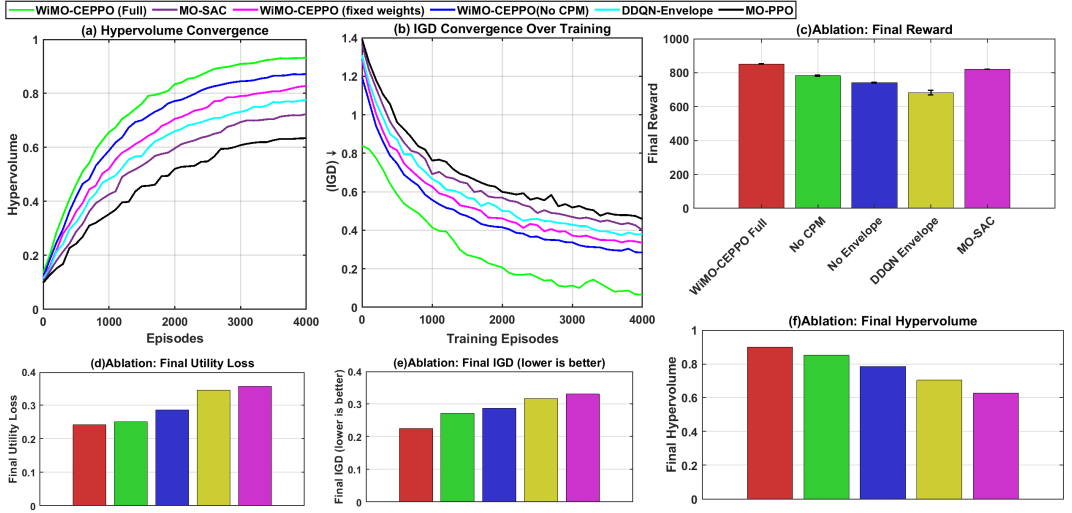


Figure 2: Pareto-quality convergence and ablation analysis. (a) Hypervolume (HV) convergence over training episodes (higher is better), where WiMO-CEPPO (Full) achieves the largest dominated volume, indicating superior Pareto coverage. (b) IGD convergence (lower is better), showing WiMO-CEPPO (Full) remains closest to the reference Pareto set throughout training. (c) Final scalarized reward for the full method, ablations, and baselines. (d) Final utility loss (lower is better) as an optimization-stability diagnostic. (e) Final IGD (lower is better) and (f) final HV (higher is better) for ablations, confirming that both the context-to-preference mapping (CPM) and the envelope mechanism contribute materially to Pareto quality and convergence.

tations or constraint implementations.

DDQN-Envelope is inherently value-based and requires a finite action space. We therefore implement DDQN-Envelope over a discretized RB/power action code book. This makes it a suitable value-based envelope comparator, but the discretization introduces quantization error and scaling limitations for continuous transmit-power control. Consequently, DDQN-Envelope is interpreted as a complementary discrete-action MORL baseline rather than a direct continuous-control equivalent of WiMO-CEPPO.

All methods are trained using the same interaction budget (total environment steps/episodes) and identical episode length. We use the same set of random seeds for initialization and environment stochasticity, and we report the mean $\pm$ standard deviation across seeds. Hyperparameters are tuned using a common protocol (e.g., grid search over a fixed budget) and then held fixed across scenarios. During evaluation, exploration is disabled (deterministic policy execution for actor-critic methods and greedy action selection for DDQN-Envelope). We compute Pareto metrics from the non-dominated set of objective vectors collected during evaluation and report wireless KPIs using consistent decoding and feasibility checks across all methods. We compare against representative actor-critic MORL baselines:

- **MO-PPO** [16]: a preference-conditioned PPO variant using multi-objective scalarization.
- **MO-SAC** [8]: a soft actor-critic formulation extended to handle multi-objective reward structures.
- **DDQN-Envelope** [11]: a discrete value-based MORL method that applies envelope-based scalarization over discrete actions.

To isolate the contribution of each module in WiMO-CEPPO, we evaluate the following ablation variants under identical

training budgets and evaluation protocols:

- **WiMO-CEPPO (Full)**: complete framework with context-to-preference mapping (CPM), conditional envelope conflict mitigation, and the full sustainable reward vector including R_{HW} .
- **WiMO-CEPPO (No CPM)**: removes the CPM module and uses externally specified preferences; $\mathbf{w}_t$ is sampled per episode (or held fixed), disabling context-driven preference adaptation.
- **WiMO-CEPPO (No Envelope)**: removes the conditional envelope modulation; policy updates use standard preference-weighted advantages without conflict attenuation.
- **WiMO-CEPPO (Fixed Weights)**: uses a constant preference vector $\mathbf{w}_t \equiv \mathbf{w}$ for all t , evaluating the impact of static scalarization within the same policy class.

For each ablation, we report (i) Pareto-quality metrics (HV/IGD and Pareto coverage), (ii) objective-level KPIs, and (iii) stability indicators (e.g., conflict metric trends and variance across seeds). This design enables causal attribution: improvements in adaptability are primarily attributed to CPM, improvements in stability and convergence are attributed to the conditional envelope mechanism, and improvements in operational efficiency and sustainability are due to single-agent architecture.

Fig. 2 indicates that WiMO-CEPPO (Full) delivers consistently superior multi-objective performance relative to actor-critic and value-based baselines. In Fig. 2(a), WiMO-CEPPO attains the highest hypervolume across training, implying broader Pareto coverage and a larger dominated region in objective space. Complementarily, Fig. 2(b) shows the lowest IGD trajectory, demonstrating improved proximity to the reference Pareto set and faster convergence under conflicting objectives.

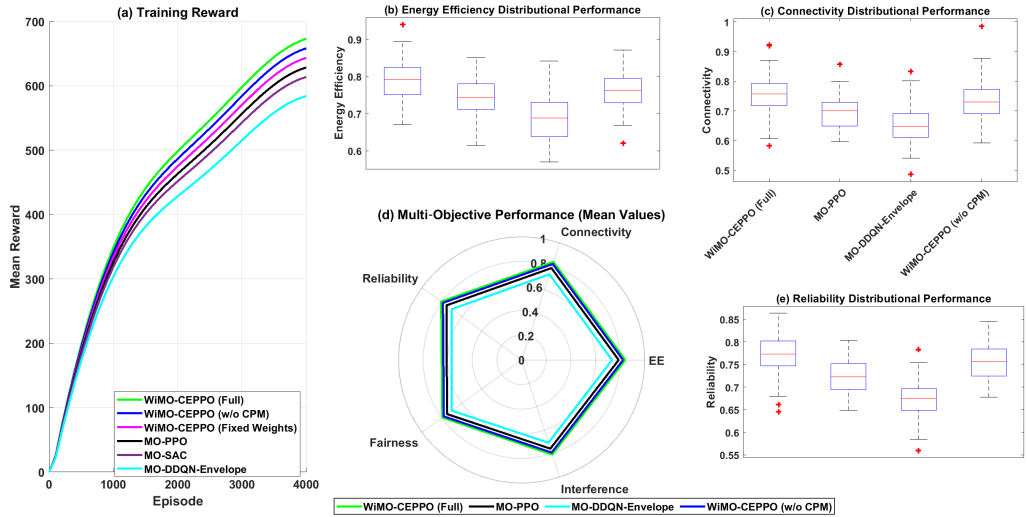


Figure 3: Training dynamics and practical KPI performance of WiMO-CEPPO and MORL baselines. (a) Mean training reward versus episodes, showing faster learning and higher asymptotic return for WiMO-CEPPO (Full) compared with ablations (w/o CPM, fixed weights) and baselines (MO-PPO, MO-SAC, MO-DDQN-Envelope). (b)–(c) Distributional performance (median, IQR, whiskers, outliers) for energy efficiency and connectivity over evaluation episodes/seeds, quantifying robustness beyond mean values. (d) Radar plot of mean multi-objective KPI profile (EE, connectivity, reliability, fairness, interference), illustrating the overall operating point induced by each method. (e) Distributional reliability, highlighting tail behavior under challenging channel/load realizations.

The ablation results in Fig. 2(c,e,f) further attribute these gains to the proposed components: removing CPM or the envelope mechanism degrades final reward and Pareto metrics (HV decreases, IGD increases), confirming that preference adaptation and conflict-aware optimization are both necessary for achieving high-quality trade-offs. Finally, the reduced utility loss in Fig. 2(d) suggests more stable optimization behavior for the full framework, consistent with its improved HV/IGD convergence.

Interpretation: Fig. 3 jointly connects learning behavior to system-level outcomes. In Fig. 3(a), WiMO-CEPPO (Full) exhibits the steepest reward growth and the highest final return, indicating better sample efficiency and a more effective resolution of competing objectives during training. The gap between WiMO-CEPPO (Full) and the ablations suggests that adaptive preference control (CPM) and dynamic conflict-aware updates (envelope/conditional weighting) are not cosmetic components: removing CPM reduces the attainable return, while fixed weights further constrain the policy to a narrow operating region, limiting long-run performance. In contrast, the MO-DDQN-Envelope curve saturates lower, which is consistent with discretization/quantization in the action space and reduced expressiveness in continuous power control, yielding weaker asymptotic performance.

Distributional KPI panels in Fig. 3(b,c,e) strengthen this conclusion by showing not only improved medians but also improved tail behavior. WiMO-CEPPO (Full) typically shifts the KPI distributions upward (higher medians) and reduces dispersion (tighter IQR/whiskers), implying a controller that is both *better on average* and *more robust* to stochastic fading, interference fluctuations, and traffic variability. Practically, tighter reliability and connectivity distributions indicate fewer outage-like events and more stable service in congested conditions, whereas wider boxes/outliers for baselines imply higher performance volatility and more frequent worst-case degradations.

Table I: Performance summary averaged over $N_s = 5$ independent training runs (different random seeds) per algorithm. For each seed, we report: (i) *final mean reward* and *final shadowed reward* at the end of training, (ii) *AUC* of the corresponding reward curves over the full training horizon (sample-efficiency/stability proxy), and (iii) Pareto quality metrics *HV* and *IGD* computed from the non-dominated set collected during evaluation. All algorithms are trained with the same environment realizations, interaction budget, and evaluation protocol; results are aggregated across seeds.

Algorithm	Final Mean Reward $\uparrow$	Final Shadowed Reward $\uparrow$	Mean Reward AUC $\uparrow$	Shadowed Reward AUC $\uparrow$	IGD $\downarrow$	HV $\uparrow$
WiMO-CEPPO	795.5	734.4	2722715	2558387	0.078	0.752
MO-PPO	532.3	513.2	1799749	1693498	0.189	0.542
MO-SAC	507.0	480.7	1684209	1580721	0.202	0.493
DDQN-Envelope	495.7	461.7	1632963	1528415	0.190	0.488

The DDQN-Envelope distributions tend to be lower and/or wider, aligning with its discrete action limitations, while MO-PPO/MO-SAC improve over DDQN but remain below WiMO-CEPPO, suggesting that preference-conditioning and conflict mitigation contribute beyond standard actor-critic learning.

Finally, the radar plot in Fig. 3(d) summarizes the multi-objective operating profile: WiMO-CEPPO (Full) achieves a more balanced expansion across all objectives, rather than over-optimizing one KPI at the expense of others. This shape-level view is practically important: a larger, more uniform radar polygon indicates that the learned policy offers a strong *compromise solution* suitable for deployment under variable network priorities. Conversely, smaller or skewed polygons for baselines reflect trade-off imbalance (e.g., sacrificing reliability or connectivity to gain EE), which can be unacceptable in 6G

scenarios where service-level constraints must be maintained. Overall, the figure supports that WiMO-CEPPO's architectural choices (preference-conditioned policy with CPM and conflict-aware updates) translate into both improved learning efficiency and more reliable, stable, and balanced network performance at inference time.

Table I reports simulation results averaged over $N_s = 5$ independent random seeds per algorithm. To ensure fair comparison, all methods are trained under identical interaction budgets, environment configurations, and evaluation procedures; each seed controls network stochasticity (e.g., channel realizations/traffic) and learning initialization. WiMO-CEPPO achieves the best end-of-training performance with a *final mean reward* of 795.5, improving over MO-PPO (532.3), MO-SAC (507.0), and DDQN-Envelope (495.7) by +49.5%, +56.9%, and +60.5%, respectively. For the conservative/tail metric, WiMO-CEPPO attains the highest *final shadowed reward* of 734.4, exceeding MO-PPO (513.2), MO-SAC (480.7), and DDQN-Envelope (461.7) by +43.1%, +52.8%, and +59.1%. Training efficiency is also superior: WiMO-CEPPO yields the largest *mean reward AUC* (2,722,715) and *shadowed reward AUC* (2,558,387), corresponding to improvements of approximately 51–67% over the baselines, which indicates faster learning and more stable training dynamics across the full horizon.

From a multi-objective standpoint, WiMO-CEPPO provides the highest Pareto quality with $HV = 0.752$ and the lowest $IGD = 0.078$. Relative to the strongest baseline (MO-PPO), WiMO-CEPPO improves HV by +38.7% (0.752 vs. 0.542) and reduces IGD by 58.7% (0.078 vs. 0.189), confirming that the learned non-dominated set is both more expansive (larger dominated volume) and closer to the reference front. The distributional plots further corroborate robustness: across energy efficiency, connectivity, and reliability, WiMO-CEPPO exhibits higher medians and a reduced lower tail (fewer low outliers) compared to baselines, consistent with the gains observed in shadowed reward and shadowed AUC. Finally, the radar chart shows that WiMO-CEPPO improves multiple objectives simultaneously (larger polygon area), indicating that the achieved performance is balanced rather than obtained by trading off a single objective aggressively.

V. CONCLUSION

This paper proposes *WiMO-CEPPO* for sustainable uplink resource allocation in multi-cell 6G-type networks, jointly optimizing energy efficiency (EE), connectivity, reliability, and inter-cell interference under practical feasibility constraints. WiMO-CEPPO learns a *single* preference-conditioned policy that adapts online to time-varying operator intent, while an *envelope-modulated* update mitigates destabilizing policy-gradient steps under strong objective conflicts. Combined with feasibility-aware action parameterization and multi-head value estimation, the approach enables stable learning across the preference simplex and yields a well-covered Pareto set.

Averaged over $N_s = 5$ seeds, WiMO-CEPPO achieves the strongest overall trade-off quality and learning efficiency, attaining the highest hypervolume ($HV = 0.752$) and the lowest inverted generational distance ($IGD = 0.078$), outperforming MO-PPO ($HV = 0.542$, $IGD = 0.189$), MO-SAC ($HV = 0.493$, $IGD = 0.202$), and DDQN-Envelope ($HV = 0.488$, $IGD = 0.190$). It also delivers higher end-of-training rewards and improved sample efficiency, with final mean/shadowed rewards of 795.5/734.4 and AUCs of 2,722,715/2,558,387, versus substantially lower AUCs for

the baselines (e.g., 1,799,749/1,693,498 for MO-PPO). These results align with convergence and distributional evaluations, indicating both faster convergence and more robust multi-objective performance.

Overall, the findings suggest that (i) preference conditioning to reuse experience across operating points without retraining, (ii) conflict-aware envelope advantages to reduce oscillations under competing objectives, and (iii) feasibility-aware action handling to limit infeasible exploration are key to reliable MORL for real-time RIC-driven scheduling. Future work will extend WiMO-CEPPO to distributed/multi-agent settings for tighter inter-cell coordination and further validate real-time feasibility via hardware-in-the-loop latency measurements and more pronounced traffic and channel non-stationarities.

VI. ACKNOWLEDGMENT

This work was supported in part by the European Union's Horizon 2020 Research and Innovation Program under Grant 101058505 '5G-TIMBER and in part by the Estonian Education and Youth Board OUF11 AIoT*5G (Artificial intelligence, edge computing and IoT solutions in distributed systems). This work was also supported by Telia Eesti AS through TalTech Development Fund.

REFERENCES

- [1] L. Li *et al.*, "Research on future 6g green wireless networks," *Green Technologies and Sustainability*, vol. 3, no. 2, p. 100156, 2025.
- [2] I.-K. Fu, "Mediatek 6g technology white paper: Sustainable mobile communications," MediaTek, Tech. Rep., 2024, white paper; emphasizes carbon-aware communication and sustainable mobile systems.
- [3] M. Alavirad, "Use case analysis related to green communication in o-ran," O-RAN Alliance, Next Generation Research Group (nGRG), Tech. Rep. RR-2024-08, 2024, o-RAN nGRG Contributed Research Report; release date 2024.06.
- [4] B. Toure *et al.*, "Multi-objective scheduling in wireless networks with deep reinforcement learning," in *2025 IEEE Wireless Communications and Networking Conference (WCNC)*, 2025.
- [5] R. Liu *et al.*, "Efficient discovery of pareto front for multi-objective reinforcement learning," in *The Thirteenth International Conference on Learning Representations (ICLR)*, 2025, also available as arXiv:2410.02236.
- [6] X. Guan *et al.*, "Multiobjective optimization of wireless powered communication networks assisted by intelligent reflecting surface based on multiagent reinforcement learning," *IEEE Transactions on Antennas and Propagation*, 2024.
- [7] R. Sun *et al.*, "Knowledge-driven deep learning paradigms for wireless network optimization in 6G," 2024.
- [8] T. Haarnoja *et al.*, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in *Proceedings of the 35th ICML*, vol. 80. PMLR, 2018, pp. 1861–1870.
- [9] F. Song *et al.*, "Energy-efficient trajectory optimization with wireless charging in uav-assisted mec based on multi-objective reinforcement learning," *IEEE Transactions on Mobile Computing*, vol. 23, no. 12, pp. 10 867–10 884, 2024.
- [10] M. Wu *et al.*, "Deep reinforcement learning-based energy efficiency optimization for ris-aided integrated satellite-aerial-terrestrial relay networks," *IEEE Transactions on Communications*, 2024.
- [11] Z. Yan *et al.*, "Generalized multi-objective reinforcement learning with envelope updates in urllc-enabled vehicular networks," *IEEE Transactions on Vehicular Technology*, 2025.
- [12] A.-T. *et al.*, "Multi-objective deep reinforcement learning for 5G Base station placement to support localisation for future sustainable traffic," in *EuNC/6G Summit*, 2024.
- [13] A. A. Habob *et al.*, "Non-orthogonal age-optimal information dissemination in vehicular networks: A meta multi-objective reinforcement learning approach," *IEEE Transactions on Mobile Computing*, vol. 23, no. 10, pp. 9789–9803, 2024.
- [14] Y. Liu *et al.*, "Dynamic preference inference network: Improving sample efficiency for multi-objective reinforcement learning by preference estimation," *Knowledge-Based Systems*, vol. 304, p. 112512, 2024.
- [15] D. Heo *et al.*, "Dynamic preference multi-objective reinforcement learning for internet network management," in *International Conference on Neural Information Processing*, 2025.
- [16] t. Abels, Axel, "Dynamic weights in multi-objective deep reinforcement learning," *International Conference on Machine Learning (ICML) Workshop*, 2019.

Curriculum Vitae

Name: Aamir Latif

E-mail address: aamir.latif@taltech.ee

ORCID: 0000-0001-7671-5377

Current address: Tallinn

1. Personal data

Name	Aamir Latif
Date and place of birth	9 March Rawalpindi, Pakistan
Nationality	Pakistani

2. Contact information

Address	Tallinn University of Technology, School of Information Technologies, Thomas Johann Seebeck Department of Electronics, Ehitajate tee 5, 19086 Tallinn, Estonia
Phone	+372 53711330
E-mail	aamir.latif@taltech.ee

3. Education

222-2026	Tallinn University of Technology, Thomas Johann Seebeck Department of Electronics, PhD studies
2013–2016	University of Bologna, Department of Electrical, Electronic and Information Engineering (DEI), MSc Telecommunication Engineering

4. Language competence

Urdu	native
Punjabi	fluent
English	fluent
Estonian	basic

5. Professional employment

2022– Present	Tallinn University of Technology, Early Stage Researcher
2016–2022	Chip Logix Design (Pvt.) Ltd., Islamabad, Senior RF & Embedded Systems Engineer
2006–2013	Advance Tech (Pvt.) Ltd., Islamabad, Embedded Hardware Design Engineer

6. Voluntary work

2019–Present	Secretary, Helping Hands, Islamabad
2008–2019	Secretary, Alfalah Welfare Trust, Wah Cantt

7. Computer skills

- Communications & 5G/6G Systems: LTE/NR PHY-MAC understanding, wireless system modeling and performance evaluation, link adaptation and interference-aware optimization, 3GPP KPI-oriented analysis and benchmarking, technical documentation.
- Simulation, AI & Programming: MATLAB (Advanced), Simulink (Advanced), NS-3, Python (Advanced), C/C++ (Intermediate), Pandas, PyTorch, MATLAB RL Toolbox, system-level simulation, statistical analysis, data analysis, 5G NR.
- RF & Embedded Systems: RF system integration, signal-chain validation, RF performance characterization, spectrum analysis, VNA calibration, oscilloscope operation, FPGA/Verilog-HDL design, embedded systems integration, DSP firmware development.
- Hardware Design & Signal Integrity: Altium Designer (Advanced), PCB design (mixed-signal integration), signal integrity analysis, hardware-in-the-loop validation techniques, RF calibration methods.

8. Honours and awards

- Oct 2024, Best Paper Award, IEEE BEC 2024.
- May 2023, Best Poster Award, Future Frontiers: PhD Conference on Emerging Technologies 2023, Tallinn, Estonia.
- Sep 2014, Erasmus+ Bologna to University of Glasgow, Sep 2014–July 2015.
- Sep 2014, UniBO Action II, University of Bologna, Italy
- Sep 2013, UniBO Action II, University of Bologna, Italy

9. Defended theses

- March 2016, **Interference Impact on Delay Tolerant Networking**, MSc, supervisor **Prof. Roberto Verdone**, University of Bologna, Department of Electrical, Electronic and Information Engineering (DEI), Italy

10. Field of research

- 6G system architecture research
- Standard-oriented wireless innovation
- PHY-layer and system-level trade-off analysis
- 3GPP KPI-based evaluation
- Algorithm development for future wireless systems
- Interference-aware resource optimization

- Evidence-based engineering for next-generation cellular systems

11. Scientific work

Papers

1. Latif, Aamir; Alam, Muhammad Mahtab; Le Moullec, Yannick, "Inter O-DUs Coordination for Scheduling of Massive Users in O-RAN," in *2023 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit 2023)*, pp. 1–6, 2023
2. Latif, Aamir; Elgarhy, Osama; Le Moullec, Yannick; Alam, Muhammad Mahtab, "Energy Consumption Evaluation of NOMA-based Sustainable Scheduling in 6G O-RAN," in *2024 International Wireless Communications and Mobile Computing (IWCMC 2024)*, pp. 484–489, 2024
3. Latif, Aamir; Elgarhy, Osama; Le Moullec, Yannick; Alam, Muhammad Mahtab, "Energy Efficiency and Connectivity Maximization for Massive Machine Type Communication in Multi-Cell NOMA," in *2024 19th Biennial Baltic Electronics Conference (BEC 2024)*, pp. 1–6, IEEE, 2024
4. Latif, Aamir and Elgarhy, Osama and Mahmood, Nurul Huda and Le Moullec, Yannick and Alam, Muhammad Mahtab, "Sustainable Massive Connectivity Maximization Scheduling Scheme for 6G," *IEEE Transactions on Green Communications and Networking*, pp. 2075–2090, January 2026
5. A. Latif, Y. Le Moullec, and M. M. Alam, "Multi-objective policy design using reinforcement learning in 6g networks," in *2026 International Wireless Communications and Mobile Computing (IWCMC)*, pp. 1352–1357, 2026
6. A. Latif, Y. L. Moullec, and M. M. Alam, "Sustainable reward mechanism for multi-objective policy design using reinforcement learning in 6g networks," in *2026 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, pp. 527–533, 2026

Conference presentations

1. A. Latif. **Sustainable Reward Mechanisms for Multi-Objective Policy Design using Reinforcement Learning in 6G Networks**, Oral Presentation at IEEE EUCNC & 6G Summit 2026, 4 June 2026, Malaga, Spain.
2. A. Latif. **Multi-Objective Policy Design using Reinforcement Learning in 6G Networks**, Oral Presentation at IEEE IWCMC 2026, 2 June 2026, Shanghai, China.
3. A. Latif. **Adaptive Policy Design using Reinforcement Learning in 6G Networks**, Oral Presentation at Future Frontiers: PhD Conference on Emerging Technologies 2025, 6–7 May 2025, Tartu, Estonia.
4. A. Latif. **Energy Efficiency and Connectivity Maximization for Massive Machine Type Communication in Multi-Cell NOMA**, Oral Presentation at IEEE BEC 2024, 1–4 October 2024, Tallinn, Estonia.

5. A. Latif. ***Energy Consumption Evaluation of NOMA-based Sustainable Scheduling in 6G O-RAN***, Oral Presentation at IEEE IWCMC 2024, 27–31 May 2024, Ayia Napa, Cyprus.
6. A. Latif. ***Inter O-DUs Coordination for Scheduling of Massive Users in O-RAN***, Oral Presentation at IEEE EUCNC & 6G Summit 2023, 6–9 June 2023, Gothenburg, Sweden.

Poster presentations

1. A. Latif. ***Sustainable Scheduling for Ultra Massive Machine Type Communication Using NOMA in 6G O-RAN***, Poster Presentation at IEEE EUCNC & 6G Summit 2024, June 2024, Antwerp, Belgium.
2. A. Latif. ***Inter O-DUs Coordination for Sustainable Scheduling of Massive Users***, Poster Presentation at PhD Technology Conference 2023, 5–6 May 2023, Tartu, Estonia.

Elulookirjeldus

1. Isikuandmed

Nimi	Aamir Latif
Sünniaeg ja -koht	9. märts, Rawalpindi, Pakistan
Kodakondsus	Pakistani

2. Kontaktandmed

Aadress	Tallinna Tehnikaülikool, Infotehnoloogia teaduskond, Thomas Johann Seebecki elektroonikainstituut, Ehitajate tee 5, 19086 Tallinn, Eesti
Telefon	+372 53711330
E-post	aamir.latif@taltech.ee

3. Haridus

2022–2026	Tallinna Tehnikaülikool, Thomas Johann Seebecki elektroonikainstituut, doktoriõpe
2013–2016	Bologna Ülikool, Elektrotehnika, elektroonika ja infotehnika osakond (DEI), magistrikraad telekommunikatsioonitehnika alal

4. Keeleoskus

Urdu	emakeel
Pandžabi	vabalt
Inglise	vabalt
Eesti keel	algtase

5. Töökogemus

2022–praegu	Tallinna Tehnikaülikool, nooremteadur
2016–2022	Chip Logix Design (Pvt.) Ltd., Islamabad, vanem RF- ja manussüsteemide insener
2006–2013	Advance Tech (Pvt.) Ltd., Islamabad, manustatud riistvara projekteerimise insener

6. Vabatahtlik tegevus

2019–praegu	sekretär, Helping Hands, Islamabad
2008–2019	sekretär, Alfalah Welfare Trust, Wah Cantt

7. Arvutioskused

- Kommunikatsioon ja 5G/6G süsteemid: LTE/NR PHY-MAC mõistmine, traadita süsteemide modelleerimine ja jõudluse hindamine, link-adaptatsioon ja häireteadlik optimeerimine, 3GPP KPI-põhine analüüs ja võrdlus, tehnilise dokumentatsiooni koostamine.
- Simulatsioon, tehisintellekt ja programmeerimine: MATLAB (edasijõudnud), Simulink (edasijõudnud), NS-3, Python (edasijõudnud), C/C++ (kesktase), Pandas, PyTorch, MATLAB RL Toolbox, süsteemitaseme simulatsioon, statistiline analüüs, andmeanalüüs, 5G NR.
- RF- ja manussüsteemid: RF-süsteemide integratsioon, signaalitee valideerimine, RF-jõudluse iseloomustamine, spektrianalüüs, VNA kalibreerimine, ostsillooskoobi kasutamine, FPGA/Verilog-HDL projekteerimine, manussüsteemide integratsioon, DSP püsivara arendus.
- Riistvaradisain ja signaali terviklikkus: Altium Designer (edasijõudnud), PCB projekteerimine (segasignaali integratsioon), signaali terviklikkuse analüüs, riistvara-in-the-loop valideerimismeetodid, RF-kalibreerimismeetodid.

8. Auhinnad ja tunnustused

- oktoober 2024, parima artikli auhind, IEEE BEC 2024.
- mai 2023, parima postri auhind, Future Frontiers: PhD Conference on Emerging Technologies 2023, Tallinn, Eesti.
- september 2014, Erasmus+ Bologna-Glasgow Ülikool, september 2014–juuli 2015.
- september 2014, UniBO Action II, Bologna Ülikool, Itaalia
- september 2013, UniBO Action II, Bologna Ülikool, Itaalia

9. Kaitstud lõputööd

- märts 2016, **Interference Impact on Delay Tolerant Networking**, magistritöö, juhendaja **prof Roberto Verdone**, Bologna Ülikool, Elektrotehnika, elektroonika ja infotehnika osakond (DEI), Itaalia

10. Teadusvaldkond

- 6G süsteemiarhitektuuri uurimine
- Standardikeskne traadita innovatsioon
- PHY-kihi ja süsteemitaseme kompromissanalüüs
- 3GPP KPI-põhine hindamine
- Algoritmide arendamine tulevaste traadita süsteemide jaoks
- Häireteadlik ressursside optimeerimine

- Töenduspõhine inseneritöö järgmise põlvkonna mobiilsüsteemidele

11. Teadustöö

Artiklid

1. Latif, Aamir; Le Moullec, Yannick; Alam, Muhammad Mahtab, "Multi-Objective Policy Design using Reinforcement Learning in 6G Networks," in 22nd International Wireless Communications & Mobile Computing Conference (IWCMC 2026).
2. Latif, Aamir; Le Moullec, Yannick; Alam, Muhammad Mahtab, "Sustainable Reward Mechanisms for Multi-Objective Policy Design using Reinforcement Learning in 6G Networks," in 2026 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit 2026).
3. Latif, Aamir; Elgarhy, Osama; Mahmood, Nurul Huda; Le Moullec, Yannick; Alam, Muhammad Mahtab, "Sustainable Massive Connectivity Maximization Scheduling Scheme for 6G," IEEE Transactions on Green Communications and Networking, pp. 2075-2090, January 2026.
4. Latif, Aamir; Elgarhy, Osama; Le Moullec, Yannick; Alam, Muhammad Mahtab, "Energy Efficiency and Connectivity Maximization for Massive Machine Type Communication in Multi-Cell NOMA," in 2024 19th Biennial Baltic Electronics Conference (BEC 2024), pp. 1-6, IEEE, 2024.
5. Latif, Aamir; Elgarhy, Osama; Le Moullec, Yannick; Alam, Muhammad Mahtab, "Energy Consumption Evaluation of NOMA-based Sustainable Scheduling in 6G O-RAN," in 2024 International Wireless Communications and Mobile Computing (IWCMC 2024), pp. 484-489, 2024.
6. Latif, Aamir; Alam, Muhammad Mahtab; Le Moullec, Yannick, "Inter O-DUs Coordination for Scheduling of Massive Users in O-RAN," in 2023 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit 2023).

Konverentsiettekanded

1. A. Latif. *Sustainable Reward Mechanisms for Multi-Objective Policy Design using Reinforcement Learning in 6G Networks*, suuline ettekanne IEEE EUCNC & 6G Summit 2026 konverentsil, 4. juuni 2026, Málaga, Hispaania.
2. A. Latif. *Multi-Objective Policy Design using Reinforcement Learning in 6G Networks*, suuline ettekanne IEEE IWCMC 2026 konverentsil, 6. juuni 2026, Shanghai, Hiina.
3. A. Latif. *Transfer Learning in Multi-Objective Policy Design using Reinforcement Learning in 6G Networks*, suuline ettekanne Future Frontiers: PhD Conference on Emerging Technologies 2026, 5.-6. mai 2026, Tartu, Eesti.
4. A. Latif. *Multi-Objective Policy Design using Reinforcement Learning in 6G Networks*, suuline ettekanne Future Frontiers: PhD Conference on Emerging Technologies 2025, 6.-7. mai 2025, Tartu, Eesti.

5. A. Latif. *Energy Efficiency and Connectivity Maximization for Massive Machine Type Communication in Multi-Cell NOMA*, suuline ettekanne IEEE BEC 2024 konverentsil, 1.–4. oktoober 2024, Tallinn, Eesti.
6. A. Latif. *Energy Consumption Evaluation of NOMA-based Sustainable Scheduling in 6G O-RAN*, suuline ettekanne IEEE IWCMC 2024 konverentsil, 27.–31. mai 2024, Ayia Napa, Küpros.
7. A. Latif. *Inter O-DUs Coordination for Scheduling of Massive Users in O-RAN*, suuline ettekanne IEEE EUCNC & 6G Summit 2023 konverentsil, 6.–9. juuni 2023, Göteborg, Rootsi.

Postriettekanded

1. A. Latif. *Sustainable Scheduling for Ultra Massive Machine Type Communication Using NOMA in 6G O-RAN*, posteriettekanne IEEE EUCNC & 6G Summit 2024 konverentsil, juuni 2024, Antwerpen, Belgia.
2. A. Latif. *Inter O-DUs Coordination for Sustainable Scheduling of Massive Users*, posteriettekanne PhD Technology Conference 2023 konverentsil, 5.–6. mai 2023, Tartu, Eesti.

ISSN 2585-6901 (PDF)
ISBN 978-9916-80-530-5 (PDF)