Intelligent Resource Allocation for Knowledge-Driven Semantic Communication Networks

Kairong Ma

Submitted in fulfilment of the requirements for the
Degree of Doctor of Philosophy

School of Engineering
College of Science and Engineering
University of Glasgow

November 2025

Abstract

The future communication system is undergoing a paradigm shift from "transmitting bits" to "conveying meaning," with semantic communication (SemCom) emerging as its core technology. SemCom's ultra-high efficiency relies on a shared knowledge base (KB) between sender and receiver. However, in dynamic environments, the KB inevitably becomes outdated, causing a sharp decline in semantic efficiency. To restore system performance, the KB must be updated and synchronize. The core insight of this thesis is that the maintenance process of the KB and the transmission of semantic data itself compete for the same limited wireless resources. This creates a novel and fundamental dynamic trade-off between the immediate utility of transmission and the long-term benefits of KB updates. Existing literature often overlooks this coupling effect, failing to balance the immediate utility of data transmission against the long-term reliability of KB maintenance.

First, to address the KB staleness trade-off, this thesis models knowledge obsolescence as a quantifiable state variable. Based on this, we formulate the dynamic tradeoff between semantic transfer utility and knowledge consensus cost as a 0-1 Mixed-Integer Non-Linear Programming (MINLP) problem. We propose an online scheduling algorithm based on model predictive control (MPC) and iterative marginal cost allocation (IMCA) to efficiently solve this tradeoff. Second, the aforementioned resource allocation problem, along with other wireless networking applications, mathematically manifests as an NP-hard 0-1 MINLP. Traditional optimization solvers struggle to scale due to exponential complexity, while pure reinforcement learning (RL) methods suffer from inefficient search due to blind exploration. Thus, we aim to address the aforementioned challenges by constructing a unified resource optimization framework. Specifically, the main research contribution is proposing a unified 0-1 mixed optimization framework, which models discrete decision processes as Markov Decision Processes (MDPs). Its core innovation lies in solving the continuous relaxation of the original problem as guidance and theoretically proving that the neighborhood of this relaxed solution defines a high potential zone (HPZ), thereby transforming the RL agent's exploration from blind trial-and-error to efficient guided search. Third, regarding the KB synchronization mechanism itself, existing consensus protocols are designed for wired networks with deterministic fault models, making them overly conservative and inefficient for probabilistic wireless environments. To bridge this gap, this thesis builds an availability-robustness analysis framework for consensus protocols, for consensus protocols

used in tasks like KB maintenance, this paper introduces an innovative dual-metric reliability model. This model quantifies the inherent tradeoff between availability and robustness, guiding the optimal design of the quorum through solving a constrained optimization problem. Finally, this thesis designs and builds a multi-user SemCom physical platform based on non-orthogonal multiple access (NOMA). We define semantic throughput (STU) as the optimization objective and propose an improved watering algorithm to address the non-convex semantic-aware power allocation problem. Experimental results validate the significant performance gains of the proposed approach at the semantic level.

University of Glasgow

College of Science & Engineering

Statement of Originality

Name: Kairong Ma

Registration Number: xxxxxxx

I certify that the thesis presented here for examination for a PhD degree of the University of Glasgow is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it) and that the thesis has not been edited by a third party beyond what is permitted by the University's PGR Code of Practice.

I declare that the thesis does not include work forming part of a thesis presented successfully for another degree.

I declare that this thesis has been produced in accordance with the University of Glasgow's Code of Good Practice in Research.

I acknowledge that if any issues are raised regarding good research practice based on review of the thesis, the examination may be postponed pending the outcome of any investigation of the issues.

Signature: .....

Date: .....

List of Publications

Journals

1. **Kairong Ma**, Yao Sun, Shuheng Hua, Muhammad Ali Imran, and Walid Saad, "A Unified Learning-based Optimization Framework for 0-1 Mixed Problems in Wireless Networks," in IEEE Transactions on Transactions on Communication, doi: 10.1109/T-COMM.2025.3618171.

2. **Kairong Ma**, Hanaa Abumarshoud, Shuheng Hua, Muhammad Imran, and Yao Sun, "Design and Demonstration of a Visible Light Based Semantic Communication System," IEEE Transactions on Vehicular Technology. (Major revision)

3. **Kairong Ma**, Yao Sun, Yixuan Fan, Shuheng Hua, and Lei Zhang, "Balance Availability and Robustness in Probabilistic Consensus: A Quorum Optimization Framework," IEEE Transactions on Vehicular Technology. (Under Review)

4. **Kairong Ma**, Yao Sun, Xinyi Lin, Shuheng Hua, Muhammad Ali Imran, "Resource Allocation in Semantic Communication: A Trade-off Between Transmission Utility and the Opportunity Cost of Knowledge Consensus," to be submitted to IEEE Transactions on Wireless Communication. (In preparation)

5. Shuheng Hua, Yao Sun, **Kairong Ma**, Lei Feng, Mingkai Chen, and Zhaohui Yang, "Bandwidth Management in Semantic Communications: A Tradeoff Between Data Sensing and Transmission," in IEEE Transactions on Vehicular Technology, doi: 10.1109/TVT.2025.3590642.

6. Lin Xinyi, Yixuan Fan, Lei Zhang, **Kairong Ma**, Yao Sun, Anvar Tukmanov, Qammer Abbasi, and Muhammad Ali Imran. "Resource Allocation for RIS-aided mmWave System with Cooperative and Non-cooperative Base Stations," in IEEE Transactions on Vehicular Technology (2024).

7. Yixuan Fan, **Kairong Ma**, Lei Zhang, and Xianbin Wang, "Outcome Uncertainty Analysis and Quorum Redesign in Probabilistic Consensus," IEEE Transactions on Emerging Topics in Computing. (Major revision)

8. Zhixiang Qiao, Yao Sun, **Kairong Ma**, Yixuan Fan, Runze Cheng, Chengsi Liang, Lei Zhang, and Muhammad Ali Imran. “Channel Assignment for A Polar Code-Based Semantic Communication System," IEEE Transactions on Vehicular Technology. (Major revision)

9. Shuheng Hua, Yao Sun, **Kairong Ma**, Dusit Niyato, and Muhammad Ali Imran, "A Mathematical Framework of Semantic Communication based on Category Theory," IEEE Transactions on Mobile Computing. (Under Review)

Conference

1. **Kairong Ma**, Hanaa Abumarshoud, Shuheng Hua, Muhammad Imran, and Yao Sun, "Power Allocation for Throughput Maximization in NOMA-Based Semantic Communication System," ICC 2025 - IEEE International Conference on Communications, Montreal, QC, Canada, 2025, pp. 4288-4293, doi: 10.1109/ICC52391.2025.11161358.

2. Shuheng Hua, Yao Sun, **Kairong Ma**, Swash Rafiq, Wallizada Mohibullah, and Zhaohui Yang, "Optimizing Spectral Efficiency through Bandwidth Management in Semantic Communication Systems," 2024 IEEE International Conference on Communications Workshops (ICC Workshops), Denver, CO, USA, 2024, pp. 1635-1640, doi: 10.1109/ICCWorkshops59551.2024.10615639.

3. Zhixiang Qiao, Yao Sun, **Kairong Ma**, Runze Cheng, Yixuan Fan, Chengsi Liang, and Muhammad Ali Imran. "Channel Assignment for Image Transmission in Polar Code Based Semantic Communication," IEEE Globecom 2025 (accepted).

Contents

Abstract i

Statement of Originality iii

List of Publications iv

List of Acronyms xiii

Acknowledgements xv

1 Introduction 1

1.1 Research Background: From Transmitting Bits to Conveying Meaning 1

1.2 Core Challenge: The Knowledge Lifecycle and Resource Conflict 3

1.2.1 The Weakness of SemCom: Obsolescence of Knowledge 3

1.2.2 Resource Conflict: Transmission Utility and Consensus Cost 4

1.3 Key Derived Challenges 5

1.3.1 The Optimization Methodology Gap 5

1.3.2 The Consensus Reliability Gap 6

1.3.3 Hardware Demonstration for SemCom 7

1.4 Research Objectives and Key Contributions 8

1.5 Dissertation Organization 9

2 Background and Literature Review 11

2.1 Semantic communication (SemCom): From Theory to Knowledge Management 11

2.1.1 The Emergence of SemCom and Its Core Dependency 11

2.1.2 The Challenge of Knowledge Management: Obsolescence and Maintenance 12

2.2 A 0-1 Mixed-Integer Optimization Method for Knowledge Update 12

2.2.1 Traditional optimization methods 12

2.2.2 Heuristic and Meta-Heuristic Algorithms 13

2.2.3 ML and Reinforcement Learning Methods 13

2.3 Distributed Consensus and Reliability Models 13

2.3.1 Classic Deterministic Fault Tolerance Models (BFT/CFT) 14
2.3.2 Emerging probabilistic models 14
2.4 Physical Layer Applications: Semantic Optimization in VLC and NOMA . . . 14
2.4.1 VLC's Opportunities and Challenges 15
2.4.2 NOMA: VLC's Bit Layer Enhancement Technology 15

3 Resource Allocation in SemCom: A Trade-off Between Transmission Utility and the Opportunity Cost of Knowledge Consensus 16
3.1 Introduction 16
3.1.1 Related Works 17
3.1.2 Motivation 18
3.1.3 Contributions 19
3.2 System Model and Problem Formulation 20
3.2.1 knowledge base (KB) Staleness and Stochastic Updates 20
3.2.2 Semantic Transmission Utility Model 22
3.2.3 Knowledge Consensus Model 23
3.2.4 The 0-1 Mixed-Integer Optimization Problem 24
3.3 Solution 25
3.3.1 Fixed-Horizon Resource Allocation 25
3.3.2 Iterative Marginal Cost Allocation 27
3.3.3 Overall Controller Planning Algorithm 28
3.4 Simulation Results and Analysis 29
3.5 Conclusions 32

4 A Learning-based Optimization Framework for 0–1 Mixed Problem in Knowledge-Driven Semantic Communication 33
4.1 Introduction 33
4.1.1 Related Works 34
4.1.2 Contributions 36
4.2 A General 0-1 Mixed Optimization Problem for SemCom 36
4.2.1 Problem Description and Formulation 37
4.2.2 Challenges of Solving Problem (4.1) in SemCom 38
4.3 A Learning-based Optimization Approach for Solving 0-1 Mixed Optimization Problems 39
4.3.1 Transformation to Sequential Decision Problems 39
4.3.2 Relaxed Solutions as Priors for RL 41
4.3.3 Exploration Strategy in RL 43
4.4 Extensions to Non-convex Scenarios 47
4.4.1 Non-Convex Objective Function Problems 47

4.4.2 Non-Convex Constraints Problems 49
4.5 Simulation Results and Analysis 50
4.5.1 Simulation Results for Convex Relaxed Problems 51
4.5.2 Simulation Results for Non-Convex Relaxed Problems 53
4.6 Conclusions 58

5 Reliability Analysis and Optimization of Knowledge Consensus in Wireless Semantic Networks 60
5.1 Introduction 60
5.1.1 Related Work 61
5.1.2 Motivation 62
5.1.3 Contributions 63
5.2 Probabilistic Node Behavior and Consensus Reliability 64
5.2.1 Node Modeling 64
5.2.2 Consensus Performance Metrics 64
5.3 Analysis of Crash Failure 66
5.3.1 Availability-Robustness Trade-off under Quorum Size 66
5.3.2 Exact and Approximate Calculation Methods 67
5.4 Analysis of Malicious Behavior 68
5.4.1 Malicious Nodes Behavior 68
5.5 Problem Formulation and Proposed Solution 71
5.5.1 Multi-Objective Optimization of Q 72
5.5.2 Extension to Reputation Weighted Quorum 74
5.6 Simulations and Discussions 75
5.7 Conclusions 81

6 Case Study: Design and Demonstration of a Visible Light Based SemCom System 83
6.1 Introduction 83
6.1.1 Related Works 84
6.1.2 Motivations and Challenges 85
6.1.3 Contributions and Organization 86
6.2 System Model 86
6.2.1 VLC Channel Model 87
6.2.2 NOMA Link Modeling 88
6.2.3 Semantic Channel Capacity Modelling 89
6.3 Power Allocation in VL-SemCom 91
6.3.1 Problem Formulation 91
6.3.2 Proposed Solution 92
6.3.3 Complexity Analysis 95

6.4 Hardware Demonstration platform of visible light based SemCom (VL-SemCom) 95
6.4.1 Transmitter Design 95
6.4.2 Receiver Design 96
6.4.3 Semantic Transmission over VLC with Diffuser-Based Reconstruction . 97
6.5 Results and Discussions 99
6.5.1 Demo Results of the VL-SemCom System 99
6.5.2 Performance Evaluation of Power Allocation Algorithm 103
6.6 Conclusions 105

7 Conclusions and Future Trends 106
7.1 Conclusions 106
7.2 Future Trends 108
7.2.1 Synergistic Integration of Optimization Frameworks 108
7.2.2 Distributed Knowledge Maintenance and Multi-Agent Consensus . . . 108
7.2.3 Refinement of Semantic Metrics and Staleness Models 108
7.2.4 Hardware Prototyping and Semantic Security 109

List of Tables

3.1 Simulation Parameter Settings 29

4.1 Simulation Parameters 53

5.1 The Four Quadrants of Consensus Outcomes 65

6.1 Comparison of Image Quality Metrics 101
6.2 Simulation Parameters 103

List of Figures

1.1 Framework of a SemCom network 2

3.1 The system model of the knowledge-driven semantic communication. 21
3.2 A comparison of the evolution of cumulative semantic utility over time. 30
3.3 The impact of knowledge stale drift rate (δ_d) on total cumulative semantic utility. 31
3.4 The impact of KB update workload (D_{KB}) on total cumulative semantic utility. 31

4.1 Flowchart of the proposed approach. 40
4.2 Running time vs. normalized objective value ($N = 8, M = 4$). 52
4.3 Running time vs. normalized objective value ($N = 20, M = 5$). 52
4.4 Comparison of normalized objective value for different values of N under $M = 5$. 53
4.5 Comparison of normalized objective value for different values of N under $M = 10$. 54
4.6 Comparison of running time for different values of N under $M = 10$. 54
4.7 Running time vs. normalized objective value ($N = 10, M = 5$). 55
4.8 Comparison of normalized objective value for different values of N under $M = 10$ and $Q = 0.1$. 55
4.9 Comparison of running time for different values of N under $M = 10$ and $Q = 0.1$. 56
4.10 Comparison of normalized objective value for different values of N under $M = 10$ and $Q = 0.5$. 57
4.11 Comparison of running time for different values of N under $M = 10$ and $Q = 0.5$. 57

5.1 Normalized performance versus quorum size. 76
5.2 A comparison of availability (η_{A}) under varying malicious node strategies ($\bar{p}_f$) across simulation, dynamic analysis, and norm-based approximation, $N = 35$, $f = 8$. The theoretically valid range for the quorum size Q is $[22, 27]$, derived from (5.13) and (5.14) conditions. We therefore analyze three representative configurations. 77
5.3 Pareto frontiers illustrating the availability-robustness trade-off under different system parameters and quorum mechanisms. The subplots show the optimal performance for varying numbers of malicious nodes (f=3,5,7). 79
5.4 Normalized performance versus quorum size. 80

5.5 Normalized performance versus quorum size. 81

6.1 A Knowledge-Assisted Semantic Communication Framework for Multi-User NOMA-VLC Networks. 87
6.2 Example of $\partial STU/\partial P_k$ vs. P_k. 93
6.3 Schematic of the VLC transmitter circuit. 96
6.4 Schematic of the VLC receiver circuit. 97
6.5 The hardware demonstration platform for algorithm validation. 98
6.6 Visual results of semantic image transmission using the proposed framework. . 98
6.7 Bit Error Rate vs. Total available power P_{MAX}. 100
6.8 Sut Error Rate vs. Total available power P_{MAX}. 101
6.9 Image transmission and recovery results for users under different allocation strategies. 102
6.10 Sut throughput vs. Distance of U_3. 103
6.11 Bit throughput vs. Distance of U_3. 104
6.12 Fairness of sut rate vs. Total available power P_{MAX}. 105

List of Acronyms

1G	first-generation
5G	fifth-generation
BER	bit error rates
XR	extended reality
SemCom	semantic communication
KB	knowledge base
VLC	visible light communication
LED	light-emitting diode
B&B	Branch and Bound
GA	genetic algorithms
ML	machine learning
DRL	deep reinforcement learning
V2X	Vehicle-to-Everything
BFT	Byzantine Fault Tolerance
CFT	Crash Fault Tolerance
NOMA	Non-Orthogonal Multiple Access
MPC	model predictive control
IMCA	iterative marginal cost allocation
STU	semantic Throughput
DL	Deep Learning
DeepJSCC	Deep Joint Source-Channel Coding
SNR	Signal-to-Noise Ratio
DP	Dynamic Programming
SA	simulated annealing
PSO	particle swarm optimization
SIC	Semantic Interference Cancellation
PA	Power allocation
IRS	intelligent reflecting surfaces
RA	resource allocation
FL	federated learning

BS	Base Station
AP	Access Point
UE	user equipments
RL	reinforcement learning
VR	virtual reality
NCRL	Neural combinatorial deep reinforcement learning
LSTM	Long Short-Term Memory
MDP	Markov decision process
QoS	quality-of-service
HPZ	high potential zone
FLP	Fischer-Lynch-Paterson
TP	True Positive
FN	False Negative
FP	False Positive
TN	True Negative
CLT	central limit theorem
BitCom	bit communication
FPA	fixed power allocation
GRPA	gain-ratio-power-allocation
CNN	convolutional neural network
RF	radio frequency
PWM	pulse-width-modulation
CSI	channel state information
LoS	line-of-sight
PD	photo-detector
AWGN	additive white Gaussian noise
SINR	signal-to-interference-plus-noise ratio
PSNR	Peak Signal-to-Noise Ratio
SSIM	Structural Similarity Index
LPIPS	Learned Perceptual Image Patch Similarity
IoT	Internet of Things
AI	Artificial Intelligence

Acknowledgements

First, I would like to express my deepest gratitude to my supervisor, Dr. Yao Sun, for your unwavering support and invaluable guidance during the past three years of my PhD learning. Your expertise and encouragement have been instrumental in shaping the direction of my research. I am incredibly grateful for the guidance and freedom you have provided.

I want to acknowledge the assistance and cooperation of all my colleagues and friends who provided valuable insights and support during the course of this study. Their camaraderie has been a source of inspiration.

I am deeply thankful to my girlfriend, Shuheng Hua. I simply could not have completed this journey without you endless patience, emotional support, and the countless sacrifices you made. You were my rock.

Finally, I offer my heartfelt gratitude to my family. To my parents, Tielian Ma and Yindi Bian, thank you for your unconditional love and for always believing in my potential. Your sacrifices and encouragement have been my driving force.

This thesis would not have been possible without the support from any of them.

Chapter 1

Introduction

This chapter serves as the foundation for the thesis, introducing the paradigm shift from traditional bit-oriented communications to Knowledge-Driven Semantic Communications (SemCom). It establishes the research context by highlighting the critical role of the shared Knowledge Base (KB) and the challenges arising from KB staleness in dynamic wireless environments. The chapter identifies the research gaps in existing literature regarding resource allocation for KB synchronization and consensus. Finally, it outlines the thesis objectives, main contributions, and the structural organization of the subsequent chapters.

1.1 Research Background: From Transmitting Bits to Conveying Meaning

The development of wireless communication technology stands at a fundamental turning point. Over the past decades, from first-generation (1G) to fifth-generation (5G) mobile communications, Shannon's information theory has remained the core driving force behind its evolution [1]. The ultimate goal of this theoretical framework is to solve the physical layer problem of communication: how to accurately and completely replicate the bit stream generated by the source at the other end with higher rates, lower latency, and smaller bit error rates (BER). This data-oriented paradigm has achieved tremendous success, making global interconnection and high-definition video streaming a reality [2].

However, as we look ahead to the future communication networks, the research community widely recognizes that future networks will carry a mission far beyond traditional connectivity. It is not merely connecting things, but rather connecting intelligence [3]. Future applications, such as holographic telepresence, immersive extended reality (XR), autonomous vehicle fleet coordination, industrial digital twins, and distributed artificial intelligence, will no longer demand the perfect reconstruction of bits. Instead, their shared objective will be the successful execution of a specific task.

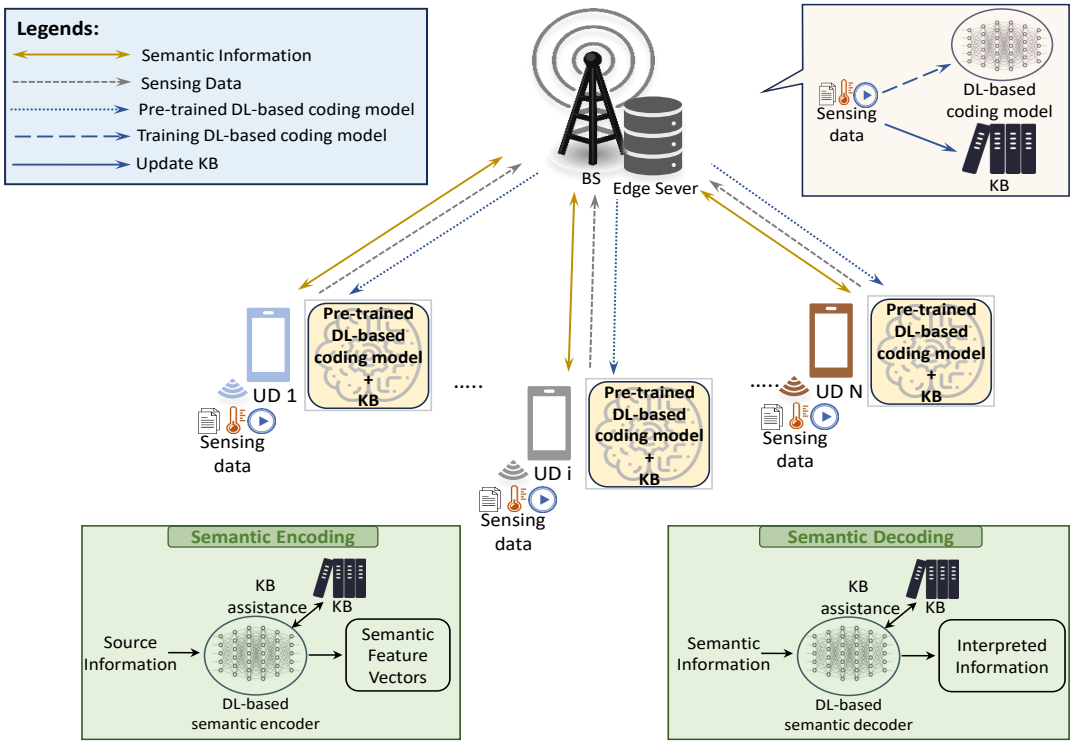


Figure 1.1: Framework of a SemCom network

For example, an autonomous driving system doesn't need a bit-perfect video stream from its LiDAR or camera sensors; all it requires is a timely and reliable answer "Is a pedestrian crossing my path?" This fundamental shift from data-driven to task-oriented or goal-oriented approaches exposes the inherent limitations of the Shannon paradigm [4]. In an era of increasingly scarce spectrum resources and exponentially growing demands for intelligent tasks, consuming precious wireless resources to transmit vast amounts of raw data, which artificial intelligence (AI) agents may ultimately deem redundant and discard is profoundly inefficient [5].

To address this challenge, SemCom has emerged as a revolutionary paradigm. As shown in Fig. 1.1, a SemCom network typically consists of semantic encoder and decoder, and the KB [6]. The core concept of SemCom is to shift focus away from bit-level precision reproduction toward extracting, compressing, and efficiently transmitting the "meaning" or "semantics" embedded within data [7]. The implementation mechanism of SemCom relies heavily on a shared KB between the sender and receiver [8]. This KB can take any form of shared context, such as pre-trained deep learning (DL) models, shared feature extractors, semantic vocabularies, or environment-specific databases [5]. Leveraging this shared KB, the sender can encode high-dimensional source data into low-dimensional, task-critical semantic features. The receiver then utilizes the identical KB to interpret these features, comprehend their meaning, and reconstruct the information required for the task. SemCom is regarded as the perfect candidate technology for overcoming physical bandwidth limitations. For instance, in the field of visible light communication (VLC), despite possessing vast amounts of unlicensed spectrum resources, its data

throughput is severely constrained by the limited modulation bandwidth of commercial light-emitting-diode (LED)s [9]. By significantly reducing the volume of data requiring transmission, SemCom provides an ideal solution for VLC to break through this physical bottleneck.

1.2 Core Challenge: The Knowledge Lifecycle and Resource Conflict

While extensive studies have focused on the architecture of SemCom [10], the management of the underlying Knowledge Base has received limited attention. Existing literature primarily assumes a static or perfectly shared KB [5]. However, model mismatches can lead to severe performance degradation. Based on this, we derive the research challenge of KB Synchronization, identifying a gap where traditional wireless resource allocation has not yet been adapted to account for the overhead of frequent model updates.

1.2.1 The Weakness of SemCom: Obsolescence of Knowledge

In the non-static real world, KBs must evolve with the environment or they will quickly become stale. Take an autonomous vehicle fleet operating in a city, for example. It shares a KB optimized for the urban street environment (e.g., containing models of pedestrians, traffic lights, and buildings) with cellular base stations. This enables efficient semantic transmission. However, when the fleet leaves the city and enters a rural road environment, this KB instantly becomes stale [6]. This knowledge staleness immediately triggers a catastrophic collapse in SemCom efficiency. The outdated KB lacks the concepts necessary to describe new environments (e.g., wildlife, debris on highways, different road signs). It is important to note that the KB is susceptible to staleness not only when switching between different scenarios but also due to temporal evolution within the same environment [5]. For example, in an autonomous driving scenario within a specific city district, the 'Knowledge' required for semantic encoding includes traffic patterns and road layouts. These features change dynamically due to temporary road closures, new construction, or variable weather conditions. If the KB is not updated to reflect these local changes, the semantic encoder may fail to recognize new obstacles, leading to safety-critical errors. At this point, the system faces two adverse consequences:

- **Semantic distortion and task failure:** The system continues to encode using outdated KBs. The encoder fails to find semantic concepts matching new features, resulting in severe semantic distortion or errors during transmission. At the receiving end, this may lead to catastrophic task failure (e.g., misclassify debris as shadows).
- **Semantic efficiency reduced to zero:** The intelligent encoder detects insufficient KB capacity, forcing it to abandon semantic compression. Instead, it transmits high-volume

raw sensor data to ensure the receiver can comprehend the unknown environment. While this behavior guarantees safety, it completely negates all efficiency gains from SemCom, degrading it to traditional communication.

Therefore, to maintain the high efficiency of SemCom, the KB itself must be a dynamic entity capable of evolving with the environment.

1.2.2 Resource Conflict: Transmission Utility and Consensus Cost

The dynamic nature of KBs introduces the core challenge addressed in this thesis: the maintenance process of KBs competes with data transmission itself for the same limited wireless resources. To address the issue of staleness, the system employs Knowledge Consensus. In the context of this thesis, knowledge consensus is defined as the distributed procedure by which the transmitter and receiver exchange update packets to reach an agreement on the current state of the Knowledge Base. Similar to database synchronization in distributed systems, these protocols ensure that the semantic decoder utilizes the exact same model parameters or context information as the encoder, thereby minimizing semantic ambiguity. In a highly saturated network (e.g., where mission-critical data is perpetually queued for transmission), the total bandwidth constitutes a fixed, scarce resource [6]. At any decision moment, the network controller faces a zero-sum game:

- **Allocating resources for semantic transmission:** The system uses the current (potentially outdated) KB to transmit task data. The utility of this action is a function of allocated bandwidth and the current knowledge obsolescence. As obsolescence increases, the semantic utility per hertz of bandwidth continuously declines.

- **Allocating resources for knowledge update:** involves the system pausing or reducing data transmission to execute a protocol synchronizing a new, more accurate KB. In many real-world architectures (e.g., when a vehicle in a connected fleet first detects a new environment), knowledge discovery is distributed. Base stations must coordinate network-wide updates. This process must be reliable and secure (e.g., preventing a single failed node from "poisoning" the KB across the entire network) while ensuring consistency among all relevant nodes. In distributed computing, this is termed "consensus."

However, executing the Knowledge Consensus protocol incurs substantial control signaling overhead, consuming bandwidth that could otherwise be used for data transmission. Consequently, this action immediately generates a significant opportunity cost. This creates a complex dynamic trade-off: immediate transmission yields instant but diminishing utility while bearing the risk of future knowledge obsolescence; executing consensus incurs present opportunity costs in exchange for long-term efficiency gains from future KB updates. This resource allocation problem must be resolved online and in real time to navigate the randomness of wireless channels and the evolution of knowledge states.

1.3 Key Derived Challenges

The core conflict identified in this thesis lies in the competition for limited wireless resources (bandwidth and energy) between two essential processes: Semantic Transmission and Knowledge Maintenance. On one hand, allocating resources to transmit semantic data yields immediate user utility. On the other hand, allocating resources to execute knowledge consensus (updating the KB) yields no immediate data throughput but is crucial for maintaining the long-term accuracy of future transmissions. High-frequency updates ensure KB freshness but consume bandwidth that could have been used for data; conversely, neglecting updates maximizes available bandwidth but degrades decoding accuracy due to KB staleness. This creates a coupled resource allocation problem that traditional methods fail to address. The core conflict outlined above further gives rise to three interrelated key challenges that collectively define the scope of this thesis. We must not only resolve this central issue but also simultaneously address the methodological, reliability, and physical layer foundations required to achieve it.

1.3.1 The Optimization Methodology Gap

How can we model and solve the dynamic trade-off between transmission and consensus? Mathematically, this is an extremely complex optimization problem. At each decision moment t, the controller needs to: 1. Make a binary (0-1) decision $x(t)$: Should the "knowledge consensus" protocol be initiated? 2. Make a continuous decision $b(t)$: If initiated, what proportion of bandwidth should be allocated to the consensus protocol?

This problem is thus formulated as a 0-1 Mixed-Integer Non-Linear Program (0-1 MINLP). Such problems are NP-hard, and their solution faces a significant methodological gap [11]. In fact, 0-1 MINLP problems are not confined to this scenario but are endemic across various resource management scenarios in wireless networks [12]. For instance, in user association, it is necessary to decide whether user i should connect to base station j (binary variables); in task offloading, it involves deciding whether to offload a task (binary variables) and allocating computational resources (continuous variables); in power control, it involves channel selection (binary variables) and power allocation (continuous variables).

Existing solution methodologies exhibit significant shortcomings:

- **Traditional optimization methods:** such as Branch and Bound (B& B) or dynamic programming, can theoretically find optimal solutions [13]. However, their computational complexity grows exponentially with problem scale (number of users, number of base stations). This makes them unsuitable for solving large-scale networks and incapable of meeting the real-time requirements of wireless communications.

- **Heuristic Algorithms:** Such as genetic algorithms (GA) or greedy algorithms, while fast, they typically lack performance guarantees and are highly prone to getting stuck in local

optima [14–16].

- **Pure machine learning (ML) approaches:** In recent years, deep reinforcement learning (DRL) has been regarded as a black-box solver. However, the essence of RL is "trial-and-error." In the vast, combinatorial discrete action space of 0-1 MINLP (e.g., M^N possible associations), this "blind" exploration leads to extremely low sample efficiency, slow convergence, and its black-box nature makes it difficult to trust for mission-critical networks [17–19].

Therefore, the academic community urgently requires a gray-box approach that can integrate the mathematical insights of traditional optimization—such as leveraging the structure of relaxation problems with the adaptive search capabilities of ML.

1.3.2 The Consensus Reliability Gap

The core of the transmission and consensus tradeoff assumes we have an available knowledge consensus protocol. However, this protocol itself must be reliable. If the consensus protocol used to maintain the KB frequently fails, the tradeoff will never deliver the expected long-term benefits, and the entire semantic system will collapse. However, in the open, dynamic environments (such as Internet of Things (IoT)) and vehicle-to-everything (V2X), ensuring the reliability of consensus presents unique challenges [20]. Traditional consensus protocols, such as Byzantine Fault Tolerance (BFT) or Crash Fault Tolerance (CFT), are designed for controlled environments like data centers [21, 22]. They rely on deterministic failure models, where nodes are either fully correct or fully failed.

This deterministic model often proves overly rigid and pessimistic in the real world:

- **Real-world failures are unpredictable:** In wireless environments, node failures are typically both benign, temporary, and probabilistic. They are caused by unpredictable factors such as transient hardware glitches, software errors, or unstable communication links, rather than deterministic malicious attacks.

- **Deterministic models lead to excessive conservatism:** BFT/CFT imposes extremely high, fixed redundancy requirements to withstand worst-case scenarios. In the aforementioned uncertain benign failure environment, such configurations may lead to excessive conservatism, limiting performance and flexibility.

On the other hand, some emerging probabilistic consensus models often reduce system reliability to a single, holistic success probability. This one-size-fits-all metric is inadequate because it fails to distinguish between two fundamentally different failure modes with vastly different costs:

- **Failure Mode 1 (Loss of Availability):** The system is overly conservative, preventing the achievement of a quorum and hindering the advancement of legitimate KB updates (i.e., False Negative).
- **Failure Mode 2 (Loss of Robustness):** The system is overly aggressive, leading to the erroneous acceptance of a flawed or even toxic KB update (i.e., False Positive).

In KB maintenance scenarios, the consequences of these two failures are starkly different. Therefore, we require an entirely new reliability framework that must explicitly model and optimize the trade-off between availability and robustness.

1.3.3 Hardware Demonstration for SemCom

Finally, even if we theoretically resolve the transmission and consensus trade-off and ensure consensus reliability, the ultimate value of SemCom must still be validated in real, resource-constrained physical systems. As discussed in previous section, VLC serves as an ideal platform for validating SemCom. It simultaneously exhibits the contradictory characteristics of abundant spectrum and severe bandwidth bottlenecks. SemCom was born precisely to resolve this contradiction.

However, applying SemCom to VLC introduces a new physical layer challenge: multi-user semantic resource allocation. Within a VLC coverage area, multiple users typically require simultaneous communication [9]. Non-orthogonal multiple access (NOMA) is the established multi-user access technique for VLC [23]. It maximizes spectral efficiency through power-domain multiplexing, leveraging the inherent characteristic of VLC systems: significant variations in user channel gains. Under the paradigm of SemCom, these traditional metrics lose their meaning. A channel that is noisy at the bit level may be perfect at the semantic level, and vice versa. It remains unclear how a semantic-aware NOMA-VLC system should operate. We must address:

- What new metrics should we optimize? (i.e., how should we define semantic throughput?)
- Is the traditional bit rate maximization power allocation strategy still effective under the goal of semantic throughput maximization?
- How can we design new resource allocation algorithms to optimize semantic performance?
- How can we validate all this on a real hardware platform?

1.4 Research Objectives and Key Contributions

In response to the three major challenges outlined above, this thesis aims to construct a comprehensive resource allocation framework for SemCom networks, spanning from reliability theory to optimization methodology and culminating in physical layer experimentation. The core contribution lies in proposing and validating a hybrid optimization-learning paradigm alongside a reliability model balancing availability-robustness, applying these to address two critical SemCom issues: the knowledge lifecycle and physical layer bandwidth bottlenecks.

The specific research objectives and four core contributions are outlined below:

- This thesis is the first to model and solve scheduling conflicts between knowledge obsolescence and knowledge consensus. The objective is to formally model and solve the dynamic trade-off between knowledge obsolescence and knowledge consensus costs in SemCom, addressing the challenges of core conflicts. This thesis introduces the first mathematical model of knowledge obsolescence in SemCom, treating it as a quantifiable, time-varying state variable. Based on this, we formulate the dynamic trade-off between semantic transmission utility and knowledge consensus cost as a 0-1 MINLP optimization problem. To enable online control, we propose a low-complexity scheduling algorithm based on model predictive control (MPC). The MPC's internal planner employs a novel heuristic iterative marginal cost allocation (IMCA) which leverages economic principles to rapidly solve 0-1 MINLP subproblems within fixed time domains by iteratively balancing marginal costs. Simulation results demonstrate that the proposed MPC-IMCA framework significantly outperforms benchmark strategies such as greedy transmission and periodic updates in terms of long-term cumulative semantic utility.

- This thesis proposes a unified relaxed guided reinforcement learning 0-1 hybrid optimization framework. The objective is to develop a universal, scalable 0-1 MINLP solving framework that addresses the methodological divide by integrating the strengths of traditional optimization and reinforcement learning. This thesis proposes a novel optimization-learning hybrid framework. The core idea is to leverage the prior knowledge from optimization theory to guide the posterior search of RL. We first relax the binary variable $x \in \{0,1\}$ to $\tilde{x} \in [0,1]$ and solve this (typically convex) relaxed problem. This thesis theoretically demonstrates that the neighborhood of this relaxed solution $\tilde{x}^*$ constitutes a HPZ, where the expected value (or lower bound on the probability) of high-quality integer solutions is significantly higher than in other regions. Subsequently, we utilize this relaxation solution as prior knowledge to initialize the value network or search policy of the RL agent. This relaxation-guided search significantly enhances the efficiency and performance of solving large-scale combinatorial optimization problems, achieving a 20% improvement in objective value compared to pure RL and reducing convergence time by 30% compared to Branch and Bound (B&B).

- This thesis establishes a consensus reliability framework for the usability-robustness tradeoff. To address challenges in consensus reliability, this thesis establishes a reliability analysis framework for the knowledge consensus protocol required in the preceding objective, transcending traditional deterministic models. Abandoning conventional deterministic failure models, we instead develop a novel dual-metric reliability framework reliability and robustness for systems with uncertain (probabilistic) node failures. We explicitly define the availability: the system's ability to correctly accept and process valid requests; and the robustness: the system's ability to correctly identify and reject invalid requests. This thesis demonstrates a fundamental trade-off between these two metrics and guides the optimal design of quorum size by solving a constrained optimization problem. Furthermore, the framework is extended to a reputation-weighted quorum mechanism, with simulation results verifying that this mechanism can enhance system robustness by several orders of magnitude.

- This thesis designs, optimizes, and builds a VLC multi-user SemCom demonstration system. To address the challenges of physical layer implementation, we design, optimiz, and build a real-world multi-user SemCom physical system, resolving the critical resource allocation issue. This thesis designs and implements a physical demonstration system for multi-user SemCom based on NOMA. To optimize this system, we rigorously define "Semantic Throughput" (STU) as a new objective function, based on the semantic entropy of the information source and the mutual information provided by KB. Based on the STU objective, we formulate a non-convex semantic-aware power allocation optimization problem. To solve this problem, we propose a Modified Water-Filling-based Algorithm. Finally, we build an Arduino-based hardware prototype with LEDs and validate the effectiveness of the proposed algorithm through physical experiments. Experimental results clearly demonstrate how our algorithm intelligently balances between bit error rate (BER) and semantic error rate (SER) to maximize the overall semantic performance of the system.

1.5 Dissertation Organization

The rest of this thesis is organized as follows:

Chapter 2 provide a detailed review of the theoretical foundations of SemCom, existing methods for 0-1 mixed optimization, reliability models for distributed consensus, and physical layer technologies for VLC-NOMA, laying the theoretical groundwork for subsequent sections.

Chapter 3 delves into the core issue of this thesis, the trade-off between transmission utility and knowledge consensus cost. We will construct a knowledge obsolescence model and a 0-1 MINLP problem, and provide a detailed account of the design and performance of the MPC-IMCA algorithm.

Chapter 4 is to provide on the content of unified optimization-learning hybrid optimization framework, namely the relaxation-guided reinforcement learning framework. We will provide its mathematical derivation, theoretical proof of HPZ, algorithm pseudocode, and universal simulation validation, serving as a methodological toolkit for subsequent chapters.

Chapter 5 is elaborated an in-depth reliability analysis framework for knowledge consensus protocols. We formulate an optimization problem to determine the optimal quorum and demonstrate the advantages of the reputation-weighted mechanism. This chapter provides the theoretical foundation for the reliability of knowledge consensus discussed in Chapter 3.

Chapter 6 demonstrates the work of semantics-aware resource allocation in VLC-NOMA. As a specific physical layer application for SemCom resource allocation, We detail the definition of STU, the design of the semantic watermarking algorithm, the construction of the hardware platform, and the experimental and simulation results.

Finally, Chapter 7 summarize all contributions of this thesis, reiterate the central theme of knowledge maintenance and 0-1 mixed optimization, and discuss future research directions.

Chapter 2

Background and Literature Review

This research aims to establish a unified, intelligent resource allocation and maintenance framework for emerging SemCom networks. This highly interdisciplinary field spans four critical frontiers in both theoretical foundations and practical challenges: 1. SemComtheory and its core reliance on KB; 2. 0-1 MINLP methods for addressing network resource scheduling problems; 3. The reliability of consensus protocols for maintaining distributed system consistency; 4. Physical layer communication technologies such as VLC and NOMA. This chapter aims to provide a comprehensive and in-depth review of the current state, mainstream approaches, and limitations across these four domains.

2.1 SemCom: From Theory to Knowledge Management

The theory of SemCom provides the foundation for this thesis, but its practical implementation hinges on a core component: the shared KB. This section reviews the literature surrounding SemCom, moving from its theoretical emergence to the critical, challenge of managing the KB lifecycle, which is a challenge that motivates the core conflicts addressed in this work.

2.1.1 The Emergence of SemCom and Its Core Dependency

In recent years, breakthroughs in DL, particularly Deep Joint Source-Channel Coding (DeepJSCC), have emerged as the mainstream paradigm for realizing this vision [24, 25]. Studies demonstrate that in low signal-to-noise ratio (SNR) regions, DeepJSCC significantly outperforms traditional methods in tasks like image reconstruction and text transmission, successfully validating its substantial advantages in achieving end-to-end task performance [26].

However, the ultra-high efficiency of SemCom hinges entirely on a critical prerequisite: a shared KB between the sender and receiver [10]. In the DeepJSCC paradigm, this KB typically refers to a pre-trained and parameter-aligned deep neural network (i.e., encoder and decoder). The sender utilizes the KB (encoder) to compress high-dimensional source data (e.g., image

[24], text [27], and speech [28]) into low-dimensional, task-critical semantic features; the receiver then employs the identical KB (decoder) to reconstruct the task information .

2.1.2 The Challenge of Knowledge Management: Obsolescence and Maintenance

The transition from wireless resource management to mathematical optimization is driven by the discrete nature of semantic protocols. In our specific SemCom scenario, the transmitter must make a binary choice at each time slot: semantic transmission or knowledge update. This binary decision, coupled with continuous power and bandwidth allocation, naturally formulates a 0-1 MINLP problem. Despite its transformative potential, the vast majority of existing research relies on a fragile, implicit assumption: that the shared KB is static, perfect, and perpetually synchronized [29]. In the non-static real world, however, this assumption is frequently invalidated. In dynamic environments, KBs must evolve with the environment, or they will rapidly become stale [8]. Knowledge staleness leads to a low SemCom efficiency. The outdated KB lacks the necessary concepts to describe the new environment.

2.2 A 0-1 Mixed-Integer Optimization Method for Knowledge Update

As revealed in Section 2.1, the transmission and consensus tradeoff is a complex resource scheduling problem. In fact, the vast majority of resource management problems in wireless networks, such as user association, task offloading, channel selection, and power allocation, mathematically can be modeled as a specific and challenging class of problems: 0-1 MINLP.

This type of problem simultaneously involves binary decisions where $x \in \{0, 1\}$ (e.g., whether to associate with a base station) and continuous variables where $y \in \mathbb{R}$ (e.g., how much power to allocate). Due to their non-convexity and combinatorial nature, 0-1 MINLP problems are NP-hard. Existing solution methodologies can be broadly categorized into three types, but each type has its fundamental limitations, forming a methodological gap:

2.2.1 Traditional optimization methods

Traditional operations research offers various white-box methods for finding global optimal solutions, such as:

- B&B: Solves integer programming by intelligently pruning the search tree [30].
- Cutting Plane Method: Approximates the feasible region by adding linear constraints.
- Dynamic Programming (DP): Suitable for problems with optimal substructure.

- Limitations: While theoretically complete, these methods suffer from the dimension catastrophe. Their computational complexity (in worst-case scenarios) grows exponentially with the number of binary variables. As the problem (3.19), for large-scale problems common in wireless networks (e.g., hundreds of users), the memory and computational time required by B&B are prohibitive.

2.2.2 Heuristic and Meta-Heuristic Algorithms

To strike a balance between optimality and complexity, heuristic algorithms are widely employed, such as:

- Greedy algorithms and local search [31].
- Meta-heuristic algorithms: such as GA [32], simulated annealing (SA), and particle swarm optimization (PSO).

While meta-heuristic algorithms like GA or Particle Swarm Optimization can solve non-convex problems, they typically require thousands of iterations to converge and lack performance guarantees. They are highly sensitive to parameters (such as termination conditions and neighborhood size) and are prone to getting stuck in local optima. In mission-critical networks, this unpredictability and instability are unacceptable.

2.2.3 ML and Reinforcement Learning Methods

In recent years, with the rise of AI, ML methods, particularly DRL, have been regarded as model-free black-box solvers. DRL agents (e.g., DDPG, PPO) learn a strategy (i.e., resource allocation scheme) that maximizes long-term rewards through trial-and-error interactions with the environment [33]. Limitations: Pure DRL methods are inefficient for 0-1 MINLP problems. Their fundamental flaw lies in blind exploration. In a combinatorial action space of size M^N (e.g., N users associated with M base stations), DRL's trial-and-error approach resembles searching for a needle in a haystack. This results in extremely low sample efficiency, excessively long training times, and a black-box nature that makes learned policies difficult to interpret or guarantee.

2.3 Distributed Consensus and Reliability Models

As described in Section 2.1, KB maintenance relies on a reliable knowledge consensus protocol [34]. This section reviews the reliability model of consensus protocols to reveal the gap.

2.3.1 Classic Deterministic Fault Tolerance Models (BFT/CFT)

The foundational work in distributed consensus is Paxos (for crash fault tolerance, CFT) [35] and the "Byzantine Generals Problem" (for BFT) [36], proposed by Lamport et al. Subsequent Practical BFT (PBFT) [37] enabled the practical implementation of BFT protocols in asynchronous systems. At the core of these classical protocols lies the deterministic fault model. They assume the existence of a precise, bounded number of faulty nodes f in the system (CFT assumes $f < n/2$, BFT assumes $f < n/3$). Provided this assumption holds, the system can provide 100% guarantees of safety and liveness (in synchronous systems).

This worst-case deterministic model is severely incompatible with modern wireless networks. Failures in wireless networks are typically benign and uncertain in nature, such as unstable communication links, temporary hardware glitches, or software bugs, rather than deterministic, premeditated malicious attacks. In this benign failure environment, the assumption of BFT $f < n/3$ appears overly conservative. For instance, to guard against an extremely low-probability event where more than $n/3$ nodes simultaneously experience benign crashes, the system may be forced to configure double or even triple redundant resources, resulting in severe performance and flexibility losses.

2.3.2 Emerging probabilistic models

To overcome the rigidity of BFT, the research community began exploring probabilistic approaches. This effort follows two main lines: Randomized Consensus: Primarily aimed at circumventing the FLP Impossibility Theorem [38] (i.e., the impossibility of achieving deterministic consensus in asynchronous systems). By introducing random sources such as coin flipping, protocols can achieve probabilistic liveness (i.e., a probability of 1 that consensus will eventually be reached) in asynchronous systems. Stochastic Fault Tolerance: It no longer assumes a fixed number of failures f, but instead assumes each node i has an independent failure probability p_i. Recent work [39, 40] has begun analyzing the reliability of Raft and PBFT protocols in wireless networks based on this model. It is crucial to distinguish between randomized consensus (where the protocol itself uses randomness, e.g., gossip algorithms) and stochastic fault tolerance (where the system is designed to withstand probabilistic failures of the environment/channel). This thesis focuses on the latter.

2.4 Physical Layer Applications: Semantic Optimization in VLC and NOMA

One of the ultimate objectives of this thesis is to validate the effectiveness of the aforementioned SemCom framework in real physical systems. VLC serves as an ideal candidate platform.

2.4.1 VLC's Opportunities and Challenges

VLC leverages ubiquitous LEDs to transmit data while providing illumination, offering significant advantages such as vast unlicensed spectrum availability, resistance to electromagnetic interference, and high physical security. However, a critical bottleneck of VLC lies in its limited modulation bandwidth (typically in the MHz range), which severely constrains its data rate.

2.4.2 NOMA: VLC's Bit Layer Enhancement Technology

To serve multiple users within limited bandwidth, NOMA is considered a key technology for VLC. NOMA leverages power domain multiplexing and relies on the inherent, significant user channel gain differences within VLC systems to substantially enhance system throughput through successive interference cancellation (SIC). Currently, academic research on NOMA-VLC is highly active [23], primarily focused on maximizing bit rate. Examples include studies on NOMA-VLC power allocation (PA) strategies [41], integration with intelligent reflecting surfaces (IRS) [42], and MIMO-NOMA systems [43].

Chapter 3

Resource Allocation in SemCom: A Trade-off Between Transmission Utility and the Opportunity Cost of Knowledge Consensus

3.1 Introduction

As the research community conceptualizes the future wireless networks, it confronts a set of challenges that are fundamentally different in nature [5]. The future communication network is defined not merely by transmitting bits faster, but by reliably enabling distributed intelligence. Future applications, are not primarily concerned with the perfect reconstruction of source data. Instead, their objective is the successful execution of a specific task [44]. For instance, an autonomous driving system does not require a bit-perfect 4K video stream from its sensors; it requires a timely and reliable answer to the query: “Is a pedestrian entering the vehicle’s path?"

This transition from a data-oriented to a task-oriented or goal-oriented paradigm exposes the limitations of traditional communication design. In an era of increasing spectrum scarcity and exponential growth in demand for intelligent tasks, consuming precious wireless resources to transport vast quantities of data that may ultimately be discarded as redundant by an AI agent is profoundly inefficient. In response to this challenge, SemCom has emerged as a revolutionary paradigm [8, 45]. Diverging from the Shannon-based approach of compressing data entropy at the source, SemCom aims to extract and transmit the meaning or semantics inherent in the data. The core premise of SemCom is the utilization of a shared, a priori KB between the transmitter and receiver [46]. The KB can encapsulate any form of shared context, such as pre-trained DL models, shared feature extractors, semantic vocabularies, or environment-specific maps. By leveraging this shared KB, an edge sensor can encode high-dimensional source data (e.g., a video feed) into a low-dimensional semantic representation. The edge server then uses its identical KB

to interpret this information, understand its meaning, and reconstruct the task-critical features.

3.1.1 Related Works

The potential efficiency gains are exponential, as the communication bottleneck is no longer the raw data entropy but the rate at which new semantic information is generated. In pioneering works, particularly those leveraging DL for joint source-channel coding [47, 48], SemCom has demonstrated superior task performance (e.g., image classification) at extremely low SNR regimes, far exceeding the capabilities of conventional systems that separate source and channel coding [49]. Early and foundational work in SemCom has been focused on establishing its theoretical framework and feasibility. The efforts in [50] linked semantic information theory to AI-driven tasks, laying the groundwork for the field. Concurrently, substantial research, such as [47–49], has focused on DL applications, particularly DeepJSCC, for transmitting text [47], images [48], and speech. These studies have successfully demonstrated SemCom's superiority over traditional methods in terms of end-to-end task performance versus BER. However, this body of work almost universally assumes a pre-shared, static KB. The primary focus is on the design of the semantic encoder/decoder architecture and its robustness to channel noise, not on the lifecycle management of the knowledge itself.

A more recent line of inquiry has begun to address resource allocation (RA) for SemCom networks. Researchers have explored power and bandwidth allocation to maximize system-wide semantic utility [51, 52], minimize semantic transmission delays [53], or manage interference in multi-user semantic systems [54]. These works are vital for operational SemCom, but they fundamentally optimize the use of a given KB. In these models, the KB is a static assumption, not a dynamic variable to be optimized. The problem of when and how to allocate resources for the costly maintenance of the KB itself remains unaddressed. Consensus is a well-established problem in distributed computing (e.g., Paxos [55], Raft [56]) and has seen growing application in wireless networks, particularly for federated learning (FL) [57], blockchain [58], and decentralized IoT networks [59]. This research is primarily concerned with the reliability, convergence speed, and security of the consensus protocol itself, e.g., how to achieve model parameter agreement over lossy wireless channels. In these contexts, achieving consensus is the primary and sole objective. This literature lacks a framework to quantify the opportunity cost of the resources consumed by the consensus process, especially when those resources are in direct competition with a parallel, high-priority data-plane task like semantic transmission.

MPC is an advanced control strategy adept at managing dynamic systems under uncertainty. It operates by solving a finite-horizon optimization problem at each time step, executing only the first step of the plan, and then re-optimizing at the next step with updated state information [60]. MPC has been widely and successfully applied to wireless resource allocation problems, such as power control [61], scheduling [62], and caching [63]. These studies validate MPC as a robust tool for online optimization. However, its application to the specific, nuanced problem of

semantic knowledge management remains unexplored. Furthermore, standard MPC frameworks are not inherently equipped to solve the unique, non-convex, and mixed-integer optimization structure that this particular problem presents.

3.1.2 Motivation

Despite its transformative potential, the vast majority of existing SemCom research relies on a critical, yet often implicit, assumption: the KB shared between parties is static, perfect, and perpetually synchronized. This assumption, however, is frequently invalidated in dynamic, real-world environments. Consider a fleet of autonomous vehicles employing SemCom for remote situational awareness. The vehicles and a coordinating base station share a KB optimized for an urban street environment (e.g., models of pedestrians, traffic signals). This allows for highly efficient semantic transmission. However, when the fleet transitions to a rural highway environment, the KB becomes instantaneously stale.

This knowledge staleness precipitates a collapse in SemCom efficiency. The old KB lacks the necessary concepts to describe the new environment (e.g., wildlife, highway debris). The system is now faced with a critical dilemma:

- Transmit with Stale KB: The vehicles continue transmitting using the obsolete KB. The semantic encoder fails to match new features, resulting in semantic distortion or error. At the RX, this can lead to task failure (e.g., misclassifying debris as a shadow).
- Revert to Traditional Mode: The encoder recognizes the KB's inadequacy and reverts to sending high-volume raw sensor data to describe the unknown environment, completely nullifying the efficiency gains of SemCom.

Therefore, to maintain high efficiency, the KB itself must be a dynamic entity, capable of evolving with the environment. This realization introduces the central conflict of this work: the process of updating the KB consumes the very same finite wireless resources that are needed for data transmission.

We consider a high-saturation network scenario, where mission-critical data is perpetually queued for transmission. In this context, the total bandwidth, B_{total}, is a fixed, scarce resource. At any given moment, the network controller (e.g., the BS) faces a zero-sum game:

- Option A: Allocate bandwidth for Semantic Transmission. The system uses the current (and potentially stale) KB to transmit task-oriented data. The utility of this action is a function of the allocated bandwidth and the current degree of knowledge staleness. As staleness increases, the semantic utility per unit of bandwidth degrades.
- Option B: Allocate bandwidth for Knowledge Update. The system pauses or curtails data transmission to execute a protocol to synchronize a new, more accurate KB. This action

immediately incurs an Opportunity Cost, as the bandwidth consumed for the update is no longer available for data transmission.

In many realistic architectures, knowledge discovery is distributed (i.e., a UE is the first to detect the new environment). The BS must then coordinate a network-wide update. This process must be robust against errors and secure against malicious attacks (e.g., preventing a single faulty UE from poisoning the network's KB). It requires a reliable synchronization protocol—termed herein as "Knowledge Consensus"—to ensure atomicity and consistency across all relevant nodes.

This creates a complex, dynamic trade-off. Allocating resources to consensus incurs an immediate opportunity cost but yields a future reward (a more accurate KB, leading to higher transmission utility). Conversely, continuing semantic transmission yields immediate utility but risks catastrophic failure as the KB becomes increasingly stale. This resource allocation problem must be solved online, navigating the inherent trade-off between the immediate utility of transmission and the long-term cost of knowledge staleness, all under the stochastic, time-varying nature of wireless channels.

3.1.3 Contributions

To address the aforementioned research gap, this work proposes a novel, online control framework to dynamically manage the trade-off between semantic transmission utility and the cost of knowledge consensus. Our methodology is grounded in a rigorous problem formulation that, for the first time, captures the specific costs and constraints of dynamic KB maintenance. This formulation and its corresponding solution provide a new theoretical and practical tool for the design of next-generation intelligent communication systems.

The main contributions of this work are articulated as follows:

- We are the first to establish a dynamic SemCom system model that explicitly incorporates "Knowledge Staleness," $\mathscr{K}(t)$, as a quantifiable state variable and links it directly to semantic utility degradation. We formally model the KB update procedure as a resource-consuming "Knowledge Consensus" protocol with specific data size and reliability requirements.

- We articulate this dynamic trade-off as a complex 0-1 MINLP. This formulation uniquely incorporates a binary decision variable to model the fixed activation costs associated with initiating the consensus protocol, thereby capturing real-world operational overheads and preventing trivial, unrealistic solutions.

- Recognizing that the formulated MINLP is NP-hard and computationally intractable for real-time control, we propose a low-complexity, deterministic heuristic algorithm, termed IMCA. IMCA is based on the classical economic principle of equalizing marginal costs

and is capable of rapidly finding a high-quality, sub-optimal solution for the fixed-horizon planning sub-problem.

- To operate the planner in a stochastic wireless environment and obviate the need for perfect long-term channel prediction, we embed the IMCA planner within a MPC framework. This online controller re-optimizes at each decision epoch, enabling it to adapt dynamically to real-time channel conditions and make robust allocation decisions.
- We conduct extensive simulations to evaluate the proposed MPC-IMCA framework. We demonstrate its significant superiority in long-term cumulative semantic utility when compared against several baseline strategies, including a greedy transmit policy and a periodic update policy, thereby validating its efficacy in managing dynamic knowledge in SemCom systems.

3.2 System Model and Problem Formulation

This section formally defines the system architecture, the key state variables, the decision space, and the associated cost and utility functions. We conclude by formulating the dynamic resource allocation problem as a stochastic optimal control problem.

3.2.1 KB Staleness and Stochastic Updates

We consider a centralized wireless communication system comprising a single central controller, such as a Base Station (BS) or Access Point (AP), and a set of N associated user equipments (UEs). The system operates over a total available bandwidth B_{total}, which is considered a fixed and scarce resource. The system evolves in a discrete-time framework, indexed by $t \in \{0,1,2,...\}$, where each index represents a decision epoch (e.g., a time slot) of duration T_{epoch}. We assume a high-saturation, mission-critical scenario where UEs always have semantic data (e.g., sensor readings, video features) queued for transmission to the BS (or vice-versa). Consequently, any bandwidth allocated to control-plane tasks (such as knowledge maintenance) directly competes with data-plane transmission, creating a quantifiable opportunity cost. At time $t = 0$, we assume a knowledge update event is triggered. This signifies that a new, more accurate KB, KB_{new}, has been identified (e.g., proposed by a UE and validated by the BS) and is ready for dissemination. This event defines the initial state of our problem.

Fig. 3.1 illustrates the system model of the proposed knowledge-driven semantic communication network. The system consists of a transmitter and a receiver communicating over a dynamic wireless channel. Distinct from traditional architectures, both terminals are equipped with a KB to support semantic feature extraction and recovery. The efficacy of the SemCom system is dictated by the fidelity of its shared KB. We model the system's evolution using a state

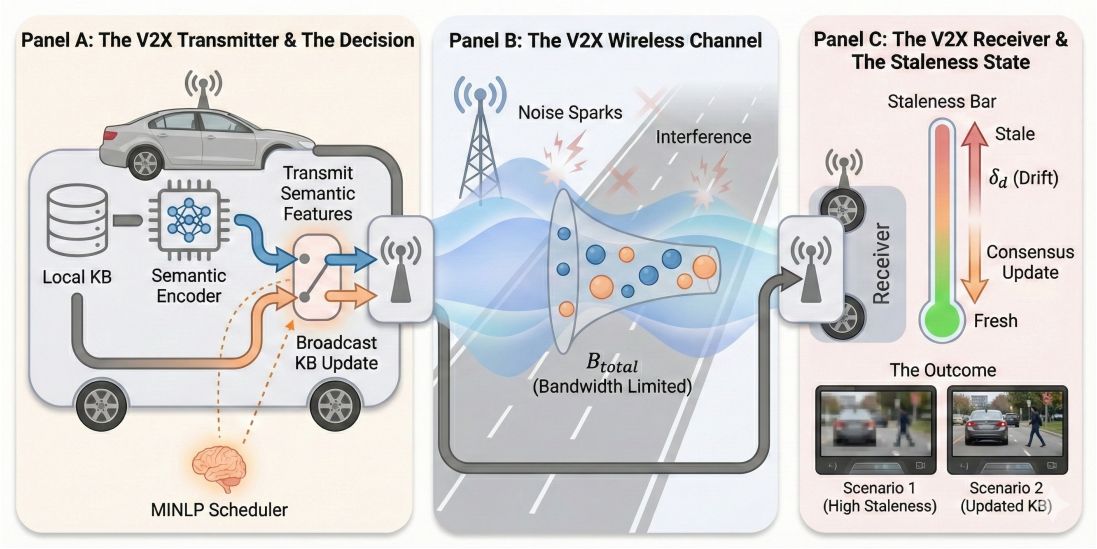


Figure 3.1: The system model of the knowledge-driven semantic communication.

vector $\mathscr{S}_t$. The minimal state required for decision-making at epoch t consists of the Knowledge Staleness, the remaining consensus work, and the current channel realization. We define a continuous state variable $\mathscr{K}(t) \in [0, \mathscr{K}_{max}]$ to quantify the semantic staleness of the currently active KB. $\mathscr{K}(t) = 0$ represents a perfectly accurate and synchronized KB. $\mathscr{K}(t) > 0$ indicates a quantifiable divergence between KB_{old}, leading to a loss in SemCom efficiency. In a dynamic environment, the relevance of any KB decays over time. We model this as a non-decreasing process. While the consensus protocol for KB_{new} is in-progress (i.e., not yet completed and activated), the staleness of KB_{old} continues to increase. We model this as a drift $\delta_d \geq 0$ per epoch, δ_d represents the knowledge obsolescence rate. It is a system-specific parameter derived from the rate at which the environment changes. A higher δ_d implies the current KB loses its validity faster. This captures the average rate at which the environment diverges from the static KB_{old}. The evolution is given by:

$$\mathscr{K}(t+1) = \min(\mathscr{K}_{max}, \mathscr{K}(t) + \delta_d). \tag{3.1}$$

Upon completion of the consensus protocol at a future time T_c, the new KB_{new} is activated, and the staleness state is reset to a target value $\mathscr{K}_{target}$, typically $\mathscr{K}_{target} = 0$. Specifically, the execution of the consensus protocol is the counter-measure in the system that resets $\mathscr{X}(t)$ to zero, thereby restoring the semantic fidelity.

The dissemination of KB_{new} requires a finite amount of work. We define D_{KB} as the total workload (e.g., in reliable bits) required to successfully synchronize KB_{new} with the necessary quorum of UEs, ensuring network-wide consistency. D_{KB} is a known constant at $t = 0$. We define a state variable $D_r(t) \geq 0$ to track the consensus workload remaining at the beginning of epoch t. Wireless channels are stochastic. We let $h(t)$ represent the channel gain (or a vector of channel gains) during epoch t. We assume $h(t)$ is constant within an epoch but varies between

epochs according to a stationary ergodic process. The BS is assumed to have knowledge of $h(t)$ at the start of epoch t (e.g., via channel sounding). Thus, the system state at the beginning of epoch t is

$$\mathscr{S}_t = (\mathscr{K}(t), D_r(t), h(t)). \tag{3.2}$$

The consensus activation decision is a binary decision variable denoted as $x(t) \in \{0,1\}$. The consensus module is active in epoch t:

$$x(t) = 1, \tag{3.3}$$

or the consensus module is inactive in epoch t:

$$x(t) = 0. \tag{3.4}$$

This action is necessary to perform any consensus work and is associated with a fixed start-up cost.

The fraction of bandwidth allocated to the consensus protocol is a continuous decision variable denoted as $b(t)$, if knowledge consensus is activated:

$$b(t) \in [0,1]. \tag{3.5}$$

Thus, the bandwidth allocated to consensus is

$$B_{con}(t) = x(t)b(t)B_{total}, \quad \forall t, \tag{3.6}$$

and the bandwidth allocated to communication is

$$B_{tx}(t) = B_{total} - B_{con}(t), \quad \forall t. \tag{3.7}$$

Thus, at the beginning of each epoch t, the BS observes the state $\mathscr{S}_t$ and executes an action (decision) a_t. This action determines the resource allocation for the upcoming epoch. The action a_t is a tuple

$$a_t = \big(x(t), b(t)\big). \tag{3.8}$$

3.2.2 Semantic Transmission Utility Model

The immediate reward in epoch t is the STU achieved using the transmission bandwidth $B_{tx}(t)$ and the current KB staleness $\mathscr{K}(t)$. We model the loss of semantic efficiency due to knowledge staleness $\mathscr{K}(t)$ using a mapping $\eta : [0, \mathscr{K}_{max}] \to [0,1]$. We adopt an exponential decay model,

which is common for modeling efficiency loss:

$$\eta(\mathscr{K}) = e^{-\alpha_K \mathscr{K}}, \tag{3.9}$$

where $\alpha_K > 0$ is a system sensitivity parameter that dictates how rapidly staleness degrades performance. This function is strictly decreasing, with $\eta(0) = 1$. The exponential decay model is widely adopted in age of information literature and Information Theoretic reliability analysis to model the value of freshness [64].

Given a transmission bandwidth allocation B_{tx} and staleness $\mathscr{K}$ over an epoch T_{epoch}, the achievable STU in one epoch is modeled based on the Shannon capacity formula, scaled by the semantic efficiency factor:

$$U_{tx}(B_{tx}, \mathscr{K}, h) = \eta(\mathscr{K}) \cdot T_{epoch} \cdot B_{tx} \cdot \log_2\left(1 + \frac{P_{tx} h}{B_{tx} N_0}\right), \tag{3.10}$$

where h is the channel gain, P_{tx} is the transmission power, and N_0 is the noise power spectral density. This model captures the core premise that as staleness $\mathscr{K}$ increases, the STU-per-Hz, U_{tx}/B_{tx}, rapidly decrease.

3.2.3 Knowledge Consensus Model

We model the consensus process as a project to reliably transmit the D_{KB} block, which competes for bandwidth B_{con}. We define $C_{eff}(B_{con}, h_{con})$ as the effective work (in reliable bits) completed by the consensus protocol in one epoch, given bandwidth B_{con} and its corresponding channel h_{con}. We abstract the complex, multi-round protocol into an effective throughput model that accounts for both transmission rate and reliability.

Allocating bandwidth B_{con} does not guarantee success, it only buys a probability of successful progress. The effective progress in one epoch is the probability-weighted rate:

$$C_{eff}(B_{con}, h_{con}) = p_s(B_{con}, h_{con}) \cdot T_{epoch} \cdot B_{con} \cdot \log_2(1 + \mathrm{SINR}_{con}), \tag{3.11}$$

where $SINR_{con}$ is the signal-to-interference-plus-noise ratio on the consensus channel, $p_s(\cdot)$ is the probability that the transmitted bits are successfully received and acknowledged by the required quorum. The term $SINR_{con}$ denotes the Signal-to-Interference-plus-Noise Ratio. We assume an orthogonal frequency division scheme where the total bandwidth B_{total} is split into B_{tx} and B_{con} without spectral overlap. Therefore, the transmission does not cause intra-system interference. This probability is an increasing, concave function of B_{con}, reflecting the reliability guarantees of the consensus protocol:

$$p_s(B_{con}, h_{con}) = 1 - e^{-\beta_B B_{con}}. \tag{3.12}$$

where $\beta_B > 0$ is a parameter scaling the reliability of the consensus channel. This form reflects that allocating more bandwidth not only increases the potential rate but also enhances protocol reliability.

Based on the models above, we can now write the formal state transition functions. Given a state $\mathscr{S}_t = (\mathscr{K}(t), D_r(t), h(t))$ and an action $a_t = (x(t), b(t))$, the state at $t+1$ evolves as:

$$\mathscr{K}(t+1) = \begin{cases} \mathscr{K}_t, & \text{if } D_r(t+1) \leq 0; \\ \min(\mathscr{K}_{max}, \mathscr{K}(t) + \delta_d), & \text{if } D_r(t+1) > 0; \end{cases} \tag{3.13}$$

and

$$D_r(t+1) = D_r(t) - C_{eff}(B_{con}(t), h(t)). \tag{3.14}$$

We assume the transmission and consensus channels $h(t)$ are sufficiently correlated or that $h(t)$ represents the relevant channel state for both.

3.2.4 The 0-1 Mixed-Integer Optimization Problem

The goal of the system is to create a schedule over a future horizon T that completes the D_{KB} update while minimizing the total opportunity cost incurred during this transition. Maximum potential utility, $R_{max}(t)$, is the utility that would have been achieved if all resources were dedicated to transmission:

$$R_{max}(t) = U_{tx}\Big(B_{total}, \mathscr{K}(t), h(t)\Big). \tag{3.15}$$

Immediate transmission utility, $R(t)$), is the utility actually achieved with action a_t:

$$R(t) = U_{tx}\Big(B_{tx}(t), \mathscr{K}(t), h(t)\Big). \tag{3.16}$$

The opportunity cost ($J_{opp}(t)$) in an epoch t is the utility lost due to allocating resources to consensus:

$$J_{opp}(t) = R_{max}(t) - R(t). \tag{3.17}$$

Fixed activation cost, denoted as $J_{fixed}(t)$, is the cost associated with the binary decision $x(t)$. It represents the start-up overhead (e.g., computation, signaling) of activating the consensus protocol for one epoch:

$$J_{fixed}(t) = \lambda \cdot K_{fixed} \cdot x(t), \tag{3.18}$$

where K_{fixed} is a fixed cost value, and λ is a weighting factor that translates this activation cost into the same units as the utility cost.

At t_0, given $\mathscr{K}_0$ and D_{KB}, find the plan $\{x(t), b(t)\}$ and horizon T that minimizes total cost. Thus, the problem can be formulated as:

$$\min_{T,\{x(t),b(t)\}_{t=1}^{T}} \quad J = \sum_{t=1}^{T} \Big[J_{opp}(t) + \lambda \cdot K_{fixed} \cdot x(t) \Big], \tag{3.19}$$

$$\text{s.t.} \quad \sum_{t=1}^{T} C_{eff}(B_{con}(t), h_t) \geq D_{\text{KB}}, \tag{3.19a}$$

$$\mathscr{K}(t+1) = \mathscr{K}(t) + \delta_d, \quad \forall t, \tag{3.19b}$$

$$x(t) \in \{0,1\}, \quad \forall t, \tag{3.19c}$$

$$0 \leq b(t) \leq 1, \quad \forall t, \tag{3.19d}$$

$$T \in \mathbb{Z}^{+}, \tag{3.19e}$$

The constraint (3.19a) is consensus completion constraint that the total accumulated reliable progress must equal the KB size. The constraint (3.19b) is staleness evolution constraint that staleness evolves during the update, increasing the opportunity cost. This formulation is a deterministic, event-driven optimization problem. It is non-convex and NP-hard, making it extremely difficult to solve with traditional solvers.

3.3 Solution

The problem (3.19) formulated in Section II is a 0-1 Mixed-Integer Non-Linear Program (MINLP) with a variable time horizon T. Solving this problem to global optimality is NP-hard and computationally infeasible for a real-time system controller. Therefore, we develop a low-complexity, deterministic heuristic algorithm designed to find a high-quality, sub-optimal solution rapidly. This algorithm runs at the moment a new knowledge event $(D_{\text{KB}}, \mathscr{K}_0)$ occurs, to generate a concrete, step-by-step execution plan. The core idea is to transform the problem from find the best plan T to find the best plan for a given deadline T, and then iterate over T to find the best deadline.

3.3.1 Fixed-Horizon Resource Allocation

The MPC controller operates in a receding-horizon fashion. At the beginning of each decision epoch t, the controller executes a four-step process. The controller observes the current, true system state $\mathscr{S}_t = (\mathscr{K}(t), D_r(t), h(t))$. The controller predicts the sequence of future channel gains $\{\hat{h}_\tau\}_{\tau=t}^{t+T_H-1}$ over a finite planning horizon T_H

$$\hat{h}_\tau = h(t), \quad \forall \tau \in [t, t+T_H-1]. \tag{3.20}$$

This assumes the channel statistics remain constant over the short planning horizon.

We first simplify the problem by fixing the total completion deadline T. The problem (P1) now becomes a Fixed-Horizon 0-1 MINLP:

$$\min_{\{x(t),b(t)\}_{t=1}^{T}} \quad J(T) = \sum_{t=1}^{T} \left[J_{opp}(t) + \lambda K_{fixed} \cdot x(t) \right] \tag{3.21}$$

$$\text{s.t.} \quad \sum_{t=1}^{T} C_{eff}(x(t)b(t)B_{total}, h_t) \geq D_{\text{KB}} \tag{3.21a}$$

$$\mathscr{K}(t+1) = \mathscr{K}(t) + \delta_d, \quad \forall t, \tag{3.21b}$$

$$x(t) \in \{0,1\}, \quad \forall t, \tag{3.21c}$$

$$0 \leq b(t) \leq 1, \quad \forall t, \tag{3.21d}$$

Problem (3.21) is still a complex MINLP, but its variable set is now fixed. It is a non-linear resource allocation problem that how to distribute the total work D_{KB} across T available time epochs to minimize total cost. The cost of working in epoch t ($J_{opp}(t)$) increases as $\mathscr{K}(t)$ increases. The efficiency of working in epoch t ($C_{eff}(t)$) varies with the channel h_t. Activating work ($x(t) = 1$) incurs a fixed cost K_{fixed}.

Algorithm 1 Iterative Marginal Cost Allocation for Fixed T

Input: $T, D_{\text{KB}}, \mathscr{K}_0, \{h_t\}_{t=1}^{T}, \delta_d, \lambda, K_{fixed}$
Initialize Plan: $x(t) \leftarrow 0, b(t) \leftarrow 0, C(t) \leftarrow 0$ **for** $t = 1 \ldots T$
Initialize State: $\mathscr{K}(t) \leftarrow \mathscr{K}_0 + t \cdot \delta_d$
$D_r \leftarrow D_{\text{KB}}$
$\Delta D \leftarrow D_{\text{KB}}/M$
while $D_r > 0$ **do**
    $t^* \leftarrow \arg\min_{t \in \{1 \ldots T\}} MC(t)$
    if $x(t^*) == 0$ **then**
        $x(t^*) \leftarrow 1$
    end if
    $C(t^*) \leftarrow C(t^*) + \Delta D$
    $b(t^*) \leftarrow C_{eff}^{-1}(C(t^*), h_{t^*})/B_{total}$
    $D_r \leftarrow D_r - \Delta D$
end while
$J_{total} \leftarrow 0$
for $t = 1 \rightarrow T$ **do**
    $B_{con}(t) \leftarrow x(t) \cdot b(t) \cdot B_{total}$
    $B_{tx}(t) \leftarrow B_{total} - B_{con}(t)$
    $J_{opp}(t) \leftarrow R_{max}(t) - R(t)$
    $J_{total} \leftarrow J_{total} + J_{opp}(t) + \lambda K_{fixed} \cdot x(t)$
end for
Return: J_{total}, $\{x(t), b(t)\}_{t=1}^{T}$

3.3.2 Iterative Marginal Cost Allocation

We propose a heuristic algorithm IMCA, to solve problem (3.21). This is a greedy algorithm based on the classic optimization principle of equalizing marginal costs. The algorithm iteratively builds the consensus plan by adding small chunks of progress (ΔD) at a time. At each step, it adds the chunk to the epoch that has the lowest marginal cost to do so.

We must define the cost of adding a small amount of consensus progress, ΔD, to an epoch t that currently has a plan $(x(t), b(t))$. Let $J_t(x,b)$ be the immediate cost in epoch t:

$$J_t(x,b) = J_{opp}(x,b) + \lambda K_{fixed} \cdot x. \tag{3.22}$$

Let $C_t(x,b)$ be the progress made in epoch t:

$$C_t(x,b) = C_{eff}(xbB_{total}, h_t). \tag{3.23}$$

The marginal cost $MC(t)$ is the cost-per-bit of adding the next chunk ΔD to epoch t. This is the derivative of cost with respect to progress:

$$MC(t) = \frac{\partial J_t(x,b)}{\partial C_t(x,b)}. \tag{3.24}$$

This function captures the 0-1 mixed nature:

- If $x(t) = 0$ (starting work): The cost of adding the first chunk ΔD is very high, as it includes the full fixed cost K_{fixed} plus the new opportunity cost J_{opp}.

$$\begin{aligned} MC_{\text{start}}(t) &\approx \frac{J_t(1, b_{\min}) - J_t(0,0)}{C_t(1, b_{\min}) - 0} \\ &\approx \frac{J_{opp}(1, b_{\min}) + \lambda K_{fixed}}{C_{eff}(b_{\min} B_{total}, h_t)}. \end{aligned} \tag{3.25}$$

- If $x(t) = 1$ (continuing work): The cost of adding ΔD is just the incremental increase in opportunity cost (since K_{fixed} is already paid).

$$MC_{\text{cont}}(t) \approx \frac{\partial J_{opp}(t)}{\partial b} \Big/ \frac{\partial C_{eff}(t)}{\partial b}. \tag{3.26}$$

The IMCA (Algorithm 1) solves the fixed-horizon problem (3.21). It iteratively allocates progress D_{KB} chunk by chunk, always picking the epoch t^* with the lowest current marginal cost $MC(t^*)$.

3.3.3 Overall Controller Planning Algorithm

The IMCA algorithm finds the best plan for a given T. The final step is to find the optimal T. The controller does this by running IMCA over a reasonable range of possible deadlines T, from $T = 1$ (a very aggressive plan) to a maximum allowed deadline T_{max} (a very patient plan).The overall algorithm (Algorithm 2) is a simple deadline sweep.

The complexity of the IMCA algorithm (Alg. 1) is dominated by the while loop, which runs M times. In each iteration, it finds the minimum of T values, taking $O(T)$ time. Thus, the complexity of IMCA is $O(M \cdot T)$. The HSP (Alg. 2) calls IMCA for $T = 1...T_H$. The total complexity is $\sum_{T=1}^{T_H} O(M \cdot T) = O(M \cdot T_H^2)$. This $O(M \cdot T_H^2)$ polynomial complexity is the total computational load per decision epoch for the MPC controller. Given that M (chunk precision) and T_H (planning horizon) are design parameters, this low-order polynomial complexity is highly suitable for real-time implementation in a network controller, in stark contrast to the exponential complexity required by any optimal solver for the original MINLP.

Algorithm 2 Overall Planning Algorithm

Input: $D_{\text{KB}}, \mathscr{K}_0, T_{max}, \mathscr{H}, \delta_d, \lambda, K_{fixed}$
Initialize: $J^* \leftarrow \infty$, $\text{Plan}^* \leftarrow \emptyset$
$\{h_t\}_{t=1}^{T_{max}} \leftarrow \mathscr{H}.\text{predict}(T_{max})$
for $T = 1 \rightarrow T_{max}$ **do**
    $[J_T, \text{Plan}_T] \leftarrow$
    $\text{IMCA}(T, D_{\text{KB}}, \mathscr{K}_0, \{h_t\}_{t=1}^{T}, \delta_d, \lambda, K_{fixed})$
    if $J_T < J^*$ **then**
        $J^* \leftarrow J_T$
        $\text{Plan}^* \leftarrow \text{Plan}_T$
    end if
end for
Return: The optimal-cost plan, Plan^*

This deadline sweep approach is a traditional, deterministic optimization heuristic. It avoids the complexities of RL training and provides a concrete, implementable algorithm for the system controller that directly addresses the 0-1 mixed nature of the problem. Since semantic decoding relies on the consistency between the transmitter's and receiver's KBs, any environmental change leads to a staleness. Knowledge consensus is the specific mechanism defined to synchronize these KBs and eliminate mismatch. Furthermore, since this staleness accumulates over time, a decision to delay consensus now results in higher penalties later. This temporal dependency motivates the use of model predictive control, which, unlike greedy algorithms, optimizes the resource trade-off over a future horizon to minimize long-term semantic errors.

3.4 Simulation Results and Analysis

In this section, we conduct extensive simulations to validate the effectiveness of our proposed MPC-IMCA framework. We evaluate its performance in terms of long-term cumulative semantic utility and compare it against two widely recognized baseline:

- Greedy Strategy: This policy maximizes immediate, short-term utility by allocating all available bandwidth (B_{total}) to semantic transmission. It never invokes the knowledge consensus protocol, thus representing a "transmit-only" approach where knowledge staleness $\mathscr{K}(t)$ grows unmitigated.
- Periodic Updates: This strategy employs a static, non-adaptive policy that interrupts semantic transmission at fixed time intervals to execute the knowledge consensus protocol, regardless of the current channel state or staleness level.

To evaluate the long-term performance, we define the cumulative semantic utility as the primary performance metric defined as the summation of the effective semantic throughput. It is important to clarify the relationship between this metric and the optimization problem formulated. Although the problem is formulated as a minimization of the system cost, this is mathematically equivalent to the maximization of semantic utility. We choose to present the results in terms of 'Utility' rather than 'Cost' because it provides a more intuitive visualization of the system's performance gain—specifically, a higher curve directly corresponds to higher semantic utility. We simulate the wireless system, operating over a total bandwidth B_{total} in discrete T_{epoch} time slots. The channel gain $h(t)$ is modeled as a stationary ergodic process. For our proposed controller, the MPC planning horizon is set to $T_H = 30$ and the IMCA chunk precision to $M = 20$. We first analyze the dynamic evolution of cumulative utility over time, and then investigate the robustness of each strategy by varying two critical environmental factors: the Knowledge stale drift rate (δ_d) and the KB update workload (D_{KB}).

Table 3.1: Simulation Parameter Settings

Parameter	Description	Value
B_{total}	Total system bandwidth	20 MHz
T_{epoch}	Decision epoch duration	100 ms
P_{total}	Transmission & Consensus power	1.0 W
α_K	Staleness sensitivity factor	0.1

Fig. 3.2 illustrates the cumulative semantic utility as a function of time, comparing the proposed MPC-IMCA framework against two baseline strategies: a Greedy strategy and a Periodic updates policy. The MPC-IMCA (Proposed) policy consistently achieves the highest cumulative utility, demonstrating a robust, monotonically increasing performance that surpasses 35 units by $t = 60s$. Conversely, the Greedy strategy initially performs well but exhibits significant performance saturation after $t = 40s$, plateauing at approximately 35 units. This saturation

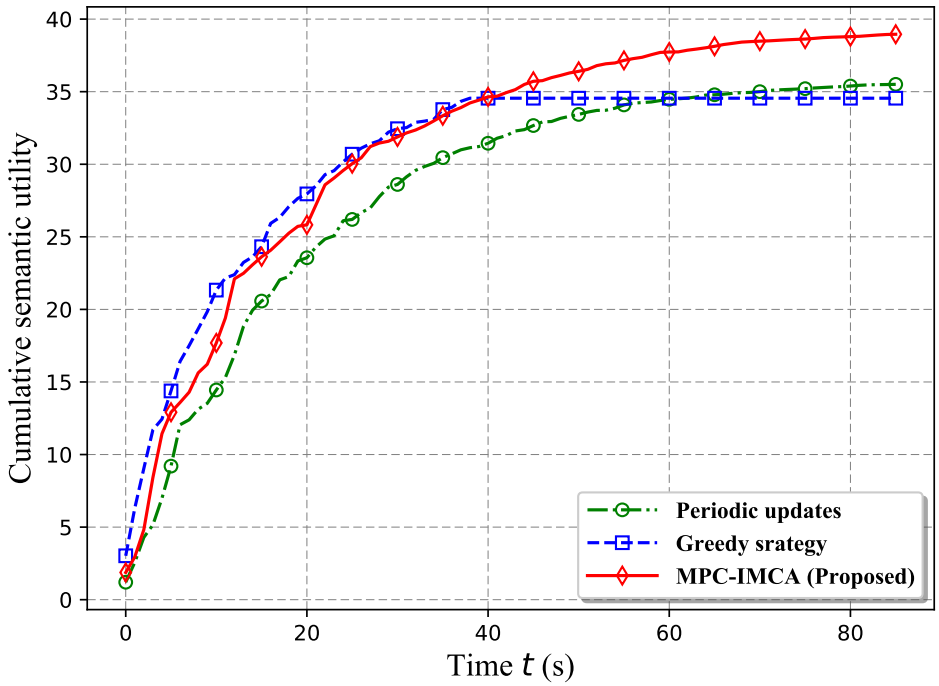


Figure 3.2: A comparison of the evolution of cumulative semantic utility over time.

is attributed to the escalating semantic utility loss as knowledge staleness grows unmitigated. The Periodic updates strategy demonstrates the slowest initial utility accumulation, indicative of the immediate opportunity cost incurred during update cycles, though it eventually converges to a similar performance level as the Greedy approach. This result validates the efficacy of the MPC-IMCA's online, adaptive decision-making, which optimally balances the immediate utility of transmission with the long-term cost of knowledge staleness, thereby outperforming the static, non-adaptive baseline heuristics over the observed time horizon.

Fig. 3.3 presents a comparative analysis of cumulative semantic utility as a function of the knowledge stale drift rate, δ_d, over a low-to-moderate range. The results demonstrate that as the environmental dynamism (i.e., δ_d) increases, the performance of all strategies degrades. However, the proposed MPC-IMCA framework exhibits superior robustness, maintaining the highest utility across the entire spectrum. At near-static conditions ($\delta_d \approx 0.05$), MPC-IMCA's performance correctly converges with the Greedy strategy, both achieving over 80 units of utility. As δ_d escalates, the utility of the Greedy strategy deteriorates sharply, becoming the worst-performing policy for $\delta_d > 0.38$. This highlights the deficiency of a static 'no-update' policy in a dynamic environment. The MPC-IMCA policy's performance degrades far more gracefully, underscoring its adaptive capability to counteract faster knowledge decay by optimizing its update schedule.

Fig. 3.4 investigates the performance robustness of the three policies against an increasing D_{KB}, which represents the total resource cost required to complete one knowledge consensus. The Greedy strategy's utility is, by definition, invariant to D_{KB} and provides a constant baseline

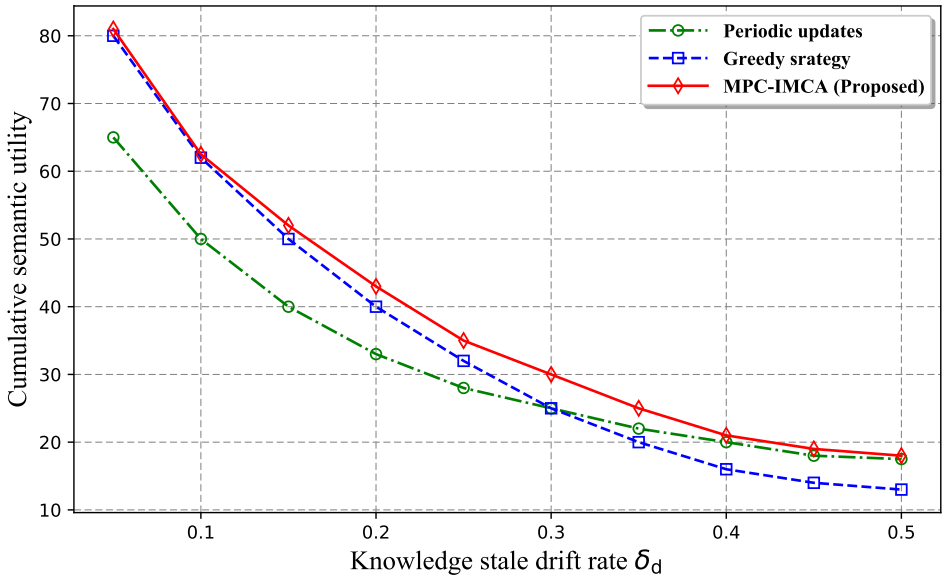


Figure 3.3: The impact of knowledge stale drift rate (δ_d) on total cumulative semantic utility.

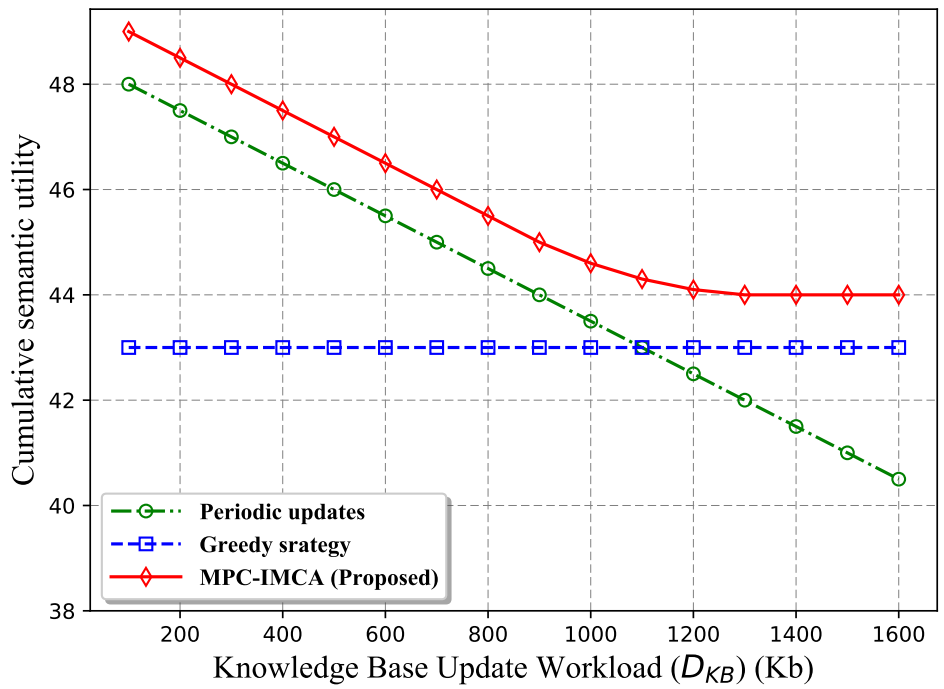


Figure 3.4: The impact of KB update workload (D_{KB}) on total cumulative semantic utility.

at approximately 43 units, as it never incurs an update cost. Conversely, the Periodic updates policy exhibits a significant, near-linear degradation in performance. This is because a larger D_{KB} forces the system to spend more time epochs in a paused update state, thereby accumulating a proportionally larger opportunity cost and falling below the Greedy baseline for $D_{\mathrm{KB}} > 1150$ Kb. The proposed MPC-IMCA framework demonstrates the most robust performance, achieving the highest utility across all tested D_{KB} values. A critical finding is observed for $D_{\mathrm{KB}} \geq 1300$ Kb, where the MPC-IMCA policy's utility curve flattens at 44 units. This demonstrates that the online controller has intelligently identified a cost threshold beyond which the total projected cost of performing the update exceeds its cumulative benefit. At this point, the controller adaptively defaults to a no-update policy, and its performance correctly becomes invariant to any further increase in D_{KB}.

3.5 Conclusions

This work addressed the critical yet under-explored problem of dynamic KB management in SemCom systems. We formulated the inherent trade-off between immediate transmission utility and the opportunity cost of KB maintenance as a 0-1 MINLP. To solve this, we proposed a novel and low-complexity MPC-IMCA online control framework. Extensive simulations validated our approach, demonstrating that the MPC-IMCA controller significantly outperforms static greedy and periodic update strategies. Crucially, our framework demonstrated intelligent adaptation: it robustly manages the update schedule under varying environmental dynamics (δ_d) and update costs (D_{KB}), and correctly defaults to a no-update policy when the cost of consensus is identified as exceeding its benefit. This work provides a practical and effective framework for managing dynamic knowledge in next-generation intelligent communication networks.

Chapter 4

A Learning-based Optimization Framework for 0–1 Mixed Problem in Knowledge-Driven Semantic Communication

4.1 Introduction

As identified in Chapter 3, the resource allocation problem in Knowledge-Driven SemCom networks fundamentally manifests as a 0-1 MINLP problem. Specifically, the network must make discrete decisions (e.g., selecting which user performs knowledge consensus) coupled with continuous resource allocation (e.g., power/bandwidth). While Chapter 3 proposed a specific heuristic for a simplified model, scaling SemCom networks to multi-user, multi-cell scenarios introduces exponential complexity. Standard solvers (like Branch-and-Bound) are too slow for wireless coherence times, and standard RL agents struggle to explore the vast combinatorial action space efficiently. Therefore, the objective of this chapter is to construct a general-purpose, high-efficiency solver engine specifically tailored for the SemCom resource allocation problems formulated in this thesis. We propose a relaxation-guided reinforcement learning framework that leverages the mathematical structure of the problem to accelerate the intelligent agent's learning process.

From the mathematical perspective, solving 0-1 mixed optimization problems is challenging. The challenges stem from the binary variables, the coupling of binary and continuous variables, and often, non-convex objective functions and/or constraints. A variety of optimization-based methods such as linear programming relaxation [65], branch and cut [66], and dynamic programming [67] have been used to solve these problems. These methods can provide the exactly optimal solution for small-scale problems, but they cannot be directly applied to dealing with general large-scale problems due to non-convexity and the problem's combinatorial nature. To

overcome this challenge, wireless networks may rely on heuristic algorithms and ML methods to generate approximate solutions [68]. For example, some DRL approaches now use large reasoning models to enhance performance [69]. Others focus on developing novel, diffusion-based reward shaping schemes [70]. In parallel, distributed frameworks like federated learning are also gaining traction. Researchers are designing long-term contribution incentives in these systems [71]. They are also building robust learning frameworks that use verifiable perturbations [72]. While these approaches provide reasonable solutions for large-scale problems, they cannot guarantee optimality or convergence [73].

4.1.1 Related Works

Prior works [12–14, 16, 19, 74–85] on wireless resource management leverage optimization techniques that can be classified into three categories: traditional optimization-based algorithms [12], heuristic algorithms [16], and more recently ML-based algorithms [74]. In general, traditional optimization methods, such as B&B, cutting planes, and dynamic programming, can guarantee an optimal solution. However, these methods are not suitable for large-scale problems due to their exponential complexity. The authors in [13] integrate Balas's method with B&B search strategies, including depth-first, breadth-first, and best-first search, to reduce computational demands for optimizing the binary integer programming problems. However, memory requirement and computing complexity remain a challenge for large-scale problems because the B&B methods need to generate a list of nodes for evaluation. The authors in [75] and [76] use a linear relaxation on binary variables, and obtain fractional solutions by solving the relaxed problem. A direct rounding method is used in [75] to recover the binary variables. The authors in [76] take the elements of the largest fractional solution of each row to 1 and the rest to 0, because there is a special constraint that the binary variables of each row sum to 1. The authors in [77] use a penalty factor in a suboptimal algorithm to enforce binary constraints on subcarrier allocation in a 0-1 mixed programming framework. This penalty factor penalizes deviations from binary values (0 or 1), ensuring that allocation decisions converge to binary values during the optimization process. In addition, the Hungarian method is also used in [78] to address combinatorial assignment problems in a polynomial time. However, the Hungarian method cannot effectively handle 0-1 mixed problems due to the coupling of binary and continuous variables.

Meanwhile, the works in [19, 79–84] adopt ML solutions for solving 0-1 mixed optimization problems in various wireless scenarios. In general, ML techniques are used in two ways to solve these problems: using neural networks to map inputs (problem parameters) and outputs (solutions), or employing reinforcement learning (RL) to explore solution space. In [79], the authors employ an RL-based curriculum learning framework, which incrementally identifies hierarchical belief structures and refines semantic event descriptions to optimize task execution efficiency, communication costs, and belief utilization. The authors in [80] use an ML solution by introducing a federated echo state network to predict the locations and orientations of vir-

tual reality (VR) users in a distributed manner, optimizing the user-base station association to minimize breaks in presence events in wireless networks. The work of [81] uses a neural combinatorial deep reinforcement learning (NCRL) framework to optimize strategies in wireless networks. Their approach combines convex optimization with a state representation based on long short-term memory (LSTM). The goal is to jointly optimize the trajectory and scheduling for UAVs. The works in [82] and [83] introduce DL-based frameworks for optimization in wireless networks. In [82], binary offloading decisions are separated from resource allocation tasks, using a DNN-based module for the binary decisions and an optimization-based module for resource allocation. Similarly, the work in [83] proposes to approximate solutions for non-convex constrained optimization problems, using binarization and a primal-dual training method. The authors in [84] integrate imitation learning with the B&B algorithm to optimize pruning policies, achieving near-optimal performance. In [19], the authors employ RL in a game-theoretic framework to optimize energy efficiency through iterative updates of user association and resource scheduling strategies. In summary, DNN-based methods like those in [19, 79–84] are useful for generating fast approximate solutions with low computational intensity. However, the quality of the DNN output depends on the robustness of the dataset and the effectiveness of the training process. RL-based methods can search the solution space based on historical data, however, they can be inefficient for problems with large solution spaces, particularly in the context of large-scale wireless network.

Furthermore, the works in [14, 16, 85] use heuristic methods to obtain acceptable solutions with significantly reduced computational complexity, despite these methods may lack guarantees of convergence and optimality. The authors in [85] conduct a survey of the applications of GAs in wireless networks. This paper covers how GAs have been utilized to address problems characterized by large, complex search spaces, particularly the 0-1 optimization problems in wireless networks. The authors in [16] develop a heuristic algorithm that combines a greedy approach for constructing initial feasible solutions and a local search method for exploring neighboring solutions to optimize binary variables. The authors in [14] present a fast heuristic algorithm for channel assignment in wireless networks, which utilizes weighted maximal independent sets within a conflict graph to manage cumulative interference, enhancing computational speed and optimality. Overall, heuristic algorithms such as those in [14, 16, 85] face a tradeoff between solution optimality and computational complexity, i.e., they target for fast problem-solving with compromises on optimality. However, these algorithms are typically tailored to specific problem structures, and their performance is highly sensitive to parameters (such as termination criteria or neighborhood size), and it often requires extensive experiments to find the optimal configurations. Additionally, the lack of performance guarantees results in inconsistent outcomes across different instances of the same problem, making these approaches less reliable for applications with high precision requirements.

4.1.2 Contributions

The main contribution of this work is to overcome the limitations of prior works by developing a unified approach for solving 0-1 mixed optimization problems in wireless networks. While pure RL methods explore the solution space based on trial-and-error, which is inefficient in the vast solution spaces of large-scale wireless networks. The core novelty of our work is the creation of a framework that transforms this process into an informed search. We achieve this by creating a unique synergy between convex optimization and RL, where the former provides a prior to guide the exploration of the latter. We first model the process of solving binary variables as an Markov decision process (MDP). Then, we relax binary variables to reformulate the original problem into a convex framework, where a feasible relaxed solution is derived based on convex optimization theory. Using the relaxed solution as prior information, RL is exploited to determine the suboptimal solution of the binary variables. When the RL algorithm reaches a terminal state, a binary solution is determined, and the continuous variables can be easily optimized using traditional optimization method. In each episode, the objective value serves as the reward for RL, i.e., it is used to update the search policy. This process is repeated until the termination condition for the whole problem is met. This proposed approach can enhance the efficiency of searching binary variables space as well as the likelihood of converging to a suboptimal solution. In summary, our key contributions include:

- Considering the problem scalability and time complexity requirement in wireless networks, we integrate convex optimization with RL to tackle the binary decision of 0–1 mixed problems. We solve a relaxed sub-problem before binary decision making, and use the relaxed solution as prior information to guide RL search policy.
- We theoretically prove that the neighbourhood of the relaxed solution has a higher probability of containing suboptimal solutions. Thus, relaxed solutions can provide prior information for RL to guide its searching policy effectively.
- We discuss the extensions of our approach to handle non-convex objective functions and non-convex constraints. These extensions broaden the applicability of our approach to complex real-world wireless networking scenarios.
- We conduct simulations to validate the proposed learning-based optimization approach for wireless problems with convex and non-convex objective functions. The simulations demonstrate that for large-scale wireless problems, our proposed approach is more effective in exploring better solutions than traditional optimization methods and pure RL.

4.2 A General 0-1 Mixed Optimization Problem for SemCom

In a typical 0-1 mixed optimization problem in wireless networks, binary variables are used to

capture discrete decision variables, such as user association, channel assignment, or task allocation; while continuous variables are introduced to manage wireless resources, such as power, bandwidth, or time. The objective is generally to optimize network performance or resource utilization, subject to constraints on the quality-of-service (QoS) requirements as well as resource limitations. In this section, we formulate a general 0-1 mixed optimization problem, and provide several common examples from wireless networks including bandwidth optimization, power allocation, and task offloading.

4.2.1 Problem Description and Formulation

Consider a scenario with N users and M base stations (BSs) in a wireless network in which a general 0-1 mixed optimization problem for network resource management can be formulated as follows:

$$\max_{x,y} \quad f(x,y) \tag{4.1}$$

$$\text{s.t.} \quad g_l(x,y) \leq 0, \quad \forall l \in \mathscr{L}; \tag{4.1a}$$

$$\sum_{j=1}^{M} x_{ij} = 1, \quad \forall i \in \mathscr{N}; \tag{4.1b}$$

$$x_{ij} \in \{0,1\}, \quad \forall i \in \mathscr{N}, \forall j \in \mathscr{M}. \tag{4.1c}$$

Binary Variables: $x \in \{0,1\}^{N\times M}$ is a matrix of binary decision variables, where each $x_{ij} \in \{0,1\}$ indicates a yes/no, on/off, or included/excluded decision.[1] For example, x_{ij} might represent whether or not to establish a connection between user i and BS j, where $x_{ij} = 1$ if user i is connected to BS j, and $x_{ij} = 0$ otherwise.

Continuous Variables: $y \in \mathbb{R}_0^{+N\times M}$ is a matrix that consists of continuous variables. It could represent the amount of resources allocated to links or nodes, such as bandwidth or power. For example, y_{ij} can be the transmit power at BS j to user i, or the bandwidth allocated from BS j to user i. These variables usually take on positive real values, and subject to resource budgets.

Objective Function: The objective function $f(x,y) : \left(\{0,1\}^{N\times M}, \mathbb{R}_0^{+N\times M}\right) \to \mathbb{R}$ is a wireless network performance metric that depends on both the binary variables x and continuous variables y. This function can capture the networks goal of maximizing the overall network throughput, minimizing the total power consumption, or optimizing any other relevant network performance metrics. It should be noted that multi-objective optimization is out of the scope of our work.

Constraints $g_l(x,y) \leq 0$*:* A set of constraints is imposed by the network in order to ensure that the minimum QoS requirements are met within the given resource budget. For example, the sum of bandwidth allocated by BS j to its associated users should not exceed its capacity, i.e., $\sum_i^N x_{ij} y_{ij} \leq B_j$, if we use x_{ij} and y_{ij} to represent cell association and bandwidth allocation,

[1] Our approach can still work under multiple sets of binary variables $x \in \{0,1\}^{N\times M\times K}$ or continuous variables $y \in \mathbb{R}_0^{+N\times M\times K}$. For example, if we consider a joint optimization of bandwidth and power allocation with continuous variables, our approach can still work.

respectively. Meanwhile, each user might have its own requirements on service quality, in terms of throughput, latency, data rate, etc.

Exclusivity Constraint $\sum_{j=1}^{M} x_{ij} = 1$*:* The exclusivity constraint is commonly needed in wireless networking optimization problems. It ensures that exactly one of the options or resources is selected from a set of possible choices for each instance i. This constraint is required typically for BS association, caching, channel allocation, time slot allocation, etc. For example, for user association, this constraint ensures that each user is uniquely associated to exactly one BS to access networks.

Problem (4.1) can be viewed as a general resource management problem for a wireless network with M BSs and N users. This problem is clearly formulated as a 0-1 mixed optimization problem. The goal is to optimize the objective function $f(x,y)$ subject to L constrains $g_l(x,y) \leq 0$ and an exclusivity constraint $\sum_{j=1}^{M} x_{ij} = 1$. Note that, this framework can be extended to problems with higher-dimensional variables, such as $x \in \{0,1\}^{N\times M\times K}$. This generalization is possible because the problem can still be framed as a sequence of N decisions. The fundamental principle of using a relaxed solution to inform the RL policy is independent of the variable dimension.

4.2.2 Challenges of Solving Problem (4.1) in SemCom

The commonly formulated problem (4.1), in wireless networks, is generally intractable. The main difficulty lies in the binary variables, which make the objective function and constraints non-differentiable, and, thus, we cannot use a gradient-based approach to solve problem (4.1). In addition, the large-scale nature of wireless networks, characterized by dense devices and base stations, increases the size of the problem and leads to an exponential explosion in solution space. This exponential computational complexity makes it infeasible to evaluate every possible solution of the problem [74]. The complexity of problem (4.1) also arises from the coupling of binary variables x and continuous variables y, particularly the cross term xy. These cross terms are non-convex and commonly appear in wireless networking optimization problems. This interdependence means that optimizing x while fixing y (or vice versa) may not lead to an optimal solution, as the optimality of one set of variables directly affect the other. Furthermore, the diverse QoS requirements complicate the constraints, leading to the non-convexity of the problem caused by feasible region. The non-convexity of objective function and constraints makes it challenging for optimization to avoid local optima traps and ensure that the solutions are as close to the global optimum as possible.

4.3 A Learning-based Optimization Approach for Solving 0-1 Mixed Optimization Problems

Our approach combines optimization theory with RL, using relaxed solutions to provide information for RL to efficiently deal with binary variables. In this section, we consider the case in which $f(x,y)$ and $g_l(x,y)_{l\in\mathscr{L}}$ are convex functions[2], and the non-convexity of optimization problem is caused by the binary variables x. We begin with this convexity assumption to first establish our core methodology in a theoretically tractable setting. This allows for a clear presentation of how relaxed solutions can guide the RL search policy. However, we acknowledge that many practical problems feature non-convexity beyond the binary variables. For example, objective functions for throughput maximization often contain non-convex terms coupling user association and resource allocation. Recognizing this challenge, we dedicate Section IV to extending our framework to realistic non-convex scenarios, where we employ techniques like convex envelopes and constraint transformation to adapt our approach.

In our proposed approach, the binary variables are first relaxed to continuous variables, then we can easily find the optimal relaxed solution using convex optimization methods due to the convexity of functions $f(x,y)$ and $g_l(x,y)$. When an episode is completed, the binary solution is then obtained through RL. Given the binary solution, an optimization-based method processes the continuous variables and obtains the objective function value. This value is then fed back into RL to update the policy for obtaining binary solutions. Repeat until convergence or the termination condition is met.

4.3.1 Transformation to Sequential Decision Problems

The process of optimizing the binary variables in problem (4.1) can be equivalently represented as an MDP. By doing so, the optimal policy of the MDP would correspond to the optimal solution of the optimization problem. The components of the transformed MDP framework are described next.

State Space: The state $s_i = \{0,1\}^{N\times M}$ represents the binary variable matrix x, where the values of first i rows have already been determined. Construct a 0-matrix $s_0 = \{0\}^{N\times M}$ with the same dimension of the binary variables as the *initial state*. For a problem with a variable size of $N \times M$, N decisions are needed to determine which element of each row of x is equal to 1, and each decision has M choices. Therefore, there are a total of M^N *terminal states* s_N, i.e., the size of possible solution space, and there are $\sum_{i=1}^{N-1} M^i$ *intermediate states*.

Action Space: There are M possible actions $\{a_i^1 \cdots a_i^M\}$ for each intermediate state s_i representing an incomplete solution, where $a_i^k \in \{0,1\}^M$. The current state is extended by updating the next decision row as $s_i[i,:] = a_i$.

[2]We use the term "convex" to refer to both convex and concave functions. By convention, we use "convex" to encompass both types of functions, allowing us to streamline the discussion and apply convex analysis techniques uniformly.

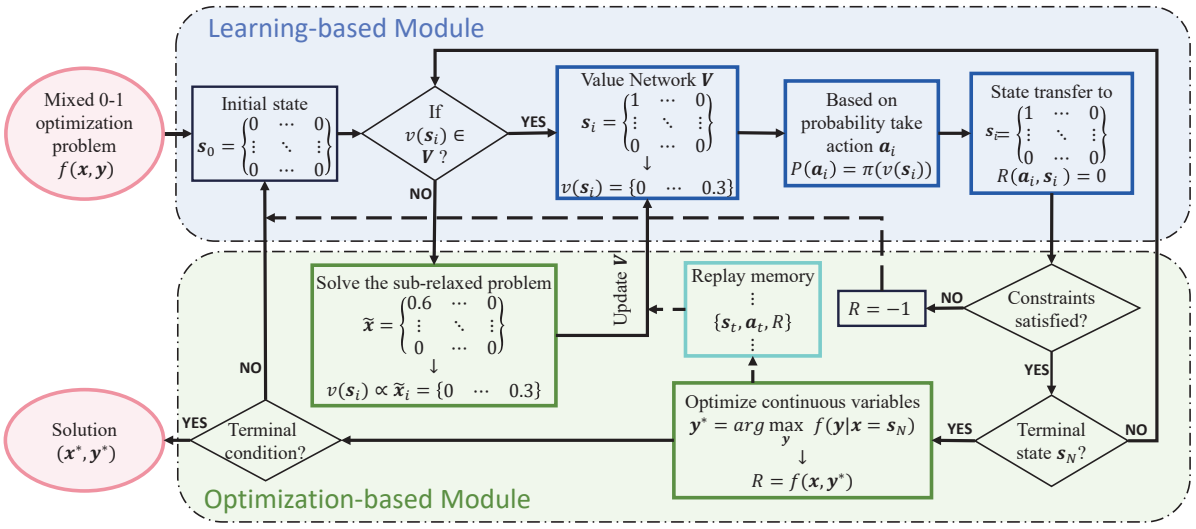


Figure 4.1: Flowchart of the proposed approach.

State Transition: The state transition is determined by the action we take. Given a intermediate state s_i, the system transitions to state s_{i+1}, as described by:

$$P[s_{i+1}|s_i] = \pi(a_i), \tag{4.2}$$

where $\pi(a_i)$ represents the probability that the search strategy takes action a_i.

Reward Function: The reward function can be defined as $R(s_i, a_i, s_{i+1}) = 0$ for all transitions until a solution is reached. When a complete solution is achieved, $R(s_i, a_i, s_{i+1})$ can be the objective function value of the 0-1 optimization problem.

$$R(s_i, a_i, s_{i+1}) = \begin{cases} 0, & i < N; \\ f(x = s_N, y'), & i = N; \end{cases} \tag{4.3}$$

where $y' = \arg\max_y f(y|x = s_N)$.

Given these mappings, it is evident that any terminal state of the MDP corresponds to a feasible solution to the binary part of a 0-1 mixed optimization problem. The process of obtain binary solution is shown as follow:

$$s_0 \xrightarrow{a_0} s_1 \xrightarrow{a_1} s_2 \cdots \xrightarrow{a_{N-1}} s_N = x.$$

Due to the large state space and delayed rewards, RL is employed to solve the MDP. Once *terminal state* s_N is reached, determining the suboptimal allocation of continuous variables becomes straightforward, i.e., $y' = \arg\max_y f(y|x = s_N)$, s.t. constraints (1a). However, exploring all possibilities is not feasible due to the huge state space. In order to improve the efficiency of exploration, relaxed solutions are utilized as prior information to guide the RL exploration.

4.3.2 Relaxed Solutions as Priors for RL

Given that $f(x,y)$ and $g_l(x,y), \forall l \in \mathscr{L}$, in problem (4.1) are convex functions, the problem becomes a convex optimization problem after relaxing $x_{ij} \in \{0,1\}$ to $\widetilde{x}_{ij} \in [0,1]$.

$$\max_{\widetilde{x},y} \quad f(\widetilde{x},y) \tag{4.4}$$

$$\text{s.t.} \quad g_l(\widetilde{x},y) \leq 0, \quad \forall l \in \mathscr{L}; \tag{4.4a}$$

$$\sum_{j=1}^{M} \widetilde{x}_{ij} = 1, \quad \forall i \in \mathscr{N}; \tag{4.4b}$$

$$\widetilde{x}_{ij} \in [0,1], \quad \forall i \in \mathscr{N}, \forall j \in \mathscr{M}. \tag{4.4c}$$

The relaxed solutions to problem (4.4) can be easily obtained using convex optimization methods. Then, the relaxed solutions are used to guide RL in prioritizing the exploration of regions with a higher likelihood of containing better solutions to the original problem. To formalize the concept of such regions in the solution space, we define the following term:

Definition 1. *A high potential zone (HPZ), is a subset of the possible solution space $X_d \subseteq X$, has higher likelihood of containing suboptimal solutions compared to the feasible space:*

$$\mathbb{E}[f(x|y)|x \in X_d] \geq \mathbb{E}[f(x,y)|x \in X]. \tag{4.5}$$

First, we present the following lemma [86], known as Cantelli's inequality, which provides a probabilistic bound on random variables.

Lemma 1. *(from [86]) If Z is a random variable with mean μ and variance σ^2, then $\forall a > 0$,*

$$P[Z \leq \mu - a] \leq \frac{\sigma^2}{\sigma^2 + a^2}. \tag{4.6}$$

Then, we introduce the following proposition, to characterize the relationship between the relaxed solution and the HPZ in the context of a 0-1 mixed optimization problem.

Proposition 1. *For a 0-1 mixed optimization problem*

$$\max_{x \in X} f(x,y), X = \{0,1\}^{N \times M},$$

if the relaxed problem

$$\max_{\widetilde{x} \in \widetilde{X}} f(\widetilde{x},y), \widetilde{X} = [0,1]^{N \times M}$$

is convex, the neighborhood of the relaxed optimal solution $U(\widetilde{x}^*,d)$ is the HPZ of original problem, where $d = \mathbb{E}[||x - \widetilde{x}^*||, x \in U]$.

Proof: Assume there are two subsets $\mathscr{O}_1(\widetilde{x}^*, d_1)$ and $\mathscr{O}_2(\widetilde{x}^*, d_2)$, where $d_1 \leq d_2$, each contains at least 1 feasible binary solution. Let $\widetilde{\mu}_1$ and $\widetilde{\sigma}_1$ represent the mean and standard deviation of $f(\widetilde{x}, y) | \widetilde{x} \in \mathscr{O}_1(\widetilde{x}^*, d_1)$, and let $\widetilde{\mu}_2$ and $\widetilde{\sigma}_2$ represent the mean and standard deviation of $f(\widetilde{x}, y) | \widetilde{x} \in \mathscr{O}_2(\widetilde{x}^*, d_2)$.

It is obvious that $\widetilde{\mu}_1 \geq \widetilde{\mu}_2$ and $\widetilde{\sigma}_1 \leq \widetilde{\sigma}_2$, thus

$$\widetilde{\mu}_1 - k\widetilde{\sigma}_1 \geq \widetilde{\mu}_2 - k\widetilde{\sigma}_2, \quad k \geq 0. \tag{4.7}$$

According to Lemma 1, for any distribution random variable Z with μ and σ, let $a = k\sigma$, then we have

$$\begin{aligned} P[Z \geq \mu - k\sigma] &\geq 1 - P[Z \leq \mu - k\sigma] \\ &\geq \frac{k^2}{1+k^2}. \end{aligned} \tag{4.8}$$

According to (4.8), there is

$$P[f(\widetilde{x}, y) \geq \widetilde{\mu}_1 - k_1\widetilde{\sigma}_1 | \widetilde{x} \in \mathscr{O}_1] \geq k_1^2/1 + k_1^2, \tag{4.9}$$

$$P[f(\widetilde{x}, y) \geq \widetilde{\mu}_2 - k_2\widetilde{\sigma}_2 | \widetilde{x} \in \mathscr{O}_2] \geq k_2^2/1 + k_2^2. \tag{4.10}$$

For $\mathscr{O}_1$, consider independent functions $f_1, f_2, ..., f_n$ with the same relaxation value distribution $\widetilde{\mu}_1$ and $\widetilde{\sigma}_1$, denote radom variables $F_1^{(1)}, F_1^{(2)}, \ldots, F_1^{(n)}$ as the objective values from these functions, with $\mu_1^{(1)}, \mu_1^{(2)}, \ldots, \mu_1^{(n)}$ and $\sigma_1^{(1)}, \sigma_1^{(2)}, \ldots, \sigma_1^{(n)}$. Here is

$$P\left[\mu_1^{(i)} \geq (\widetilde{\mu}_1 - k\widetilde{\sigma}_1)\right] \geq k^2/1 + k^2. \tag{4.11}$$

When the sample problems capacity n is large enough:

$$\lim_{n\to\infty} P\left[\left|\overline{\mu_1^n} - \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(\mu_1^{(i)})\right| > \varepsilon\right] = 0, \tag{4.12}$$

where $\overline{\mu_1^n} = \frac{1}{n}\left(\mu_1^{(1)} + \mu_1^{(2)} + \cdots + \mu_1^{(n)}\right)$ and $\varepsilon > 0$.
Thus, the mean expectation of discrete objective values in $\mathscr{O}_1$ has lower confidence bound:

$$\lim_{n\to\infty} \overline{\mu_1^n} \xrightarrow{p} \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\mu_1^{(i)}\right] \geq \frac{(\widetilde{\mu}_1 - k\widetilde{\sigma}_1)k^2}{1+k^2}. \tag{4.13}$$

The same with $\mathscr{O}_2$:

$$\lim_{n\to\infty} \overline{\mu_2^n} \xrightarrow{p} \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\mu_2^{(i)}\right] \geq \frac{(\widetilde{\mu}_2 - k\widetilde{\sigma}_2)k^2}{1+k^2}. \tag{4.14}$$

There exist $k_1^* \geq 0$ and $k_2^* \geq 0$ to let these two lower bounds achieve their maximum values

respectively, and with (4.7) there is:

$$\frac{(\widetilde{\mu}_1 - k_1^* \widetilde{\sigma}_1) k_1^{*2}}{1 + k_1^{*2}} \geq \frac{(\widetilde{\mu}_1 - k_2^* \widetilde{\sigma}_1) k_2^{*2}}{1 + k_2^{*2}} \geq \frac{(\widetilde{\mu}_2 - k_2^* \widetilde{\sigma}_2) k_2^{*2}}{1 + k_2^{*2}}. \tag{4.15}$$

Which means $f(x,y)|x \in \mathscr{O}_1$ has a better lower confidence bound of potential solution than $f(x,y)|x \in \mathscr{O}_2$.
Therefore, the neighborhood of relaxed optimal solution $U(\widetilde{x}^*, d)$ is the HPZ for a 0-1 mixed optimization problem. □

According to the proof of Proposition 1, we have:

$$\mathbb{E}\big[f(x,y)|x \in U(\widetilde{x}^*, d_i)\big] \geq \mathbb{E}\big[f(x,y)|x \in U(\widetilde{x}^*, d_{i+1})\big], \tag{4.16}$$

where $U(\widetilde{x}^*, d_i) \subseteq U(\widetilde{x}^*, d_{i+1})$. Thus, it is justified to utilize the relaxed solutions as prior information for RL to guide the design of the search policy. This search policy prioritizes exploring potential solutions within the HPZ. In an HPZ, our proposed approach provides a probabilistic lower bound on solution quality. Specifically, the approach guarantees a lower confidence bound on the expected objective value of the solutions. An HPZ with a better lower confidence bound is given more opportunity to be explored. Specifically, for two subsets $U(\widetilde{x}^*, d_i)$ and $U(\widetilde{x}^*, d_{i+1})$ the search policy assigns higher exploration probability density to the HPZ with a better confidence bound, as expressed by the following equation:

$$\pi\left[x|x \in U(\widetilde{x}^*, d_i)\right] \geq \pi\left[x|x \in U(\widetilde{x}^*, d_{i+1})\right]. \tag{4.17}$$

This ensures that the search process focuses on regions with a higher likelihood of containing suboptimal solutions, thereby improving the efficiency of RL exploration.

The practical implementation of this guided search policy is achieved by using the specific values from the relaxed solution to initialize the state-value function for the RL agent. Specifically, at any decision step, the algorithm solves a relaxed sub-problem to obtain a vector of continuous values, corresponding to each possible action. These values are used to calculate the action-selection probabilities in our exploration strategy. This method creates a soft probabilistic neighborhood by imposing a probability distribution over the action space, rather than defining hard boundaries. It effectively biases the search towards the promising regions indicated by the HPZ while still allowing for the exploration of less likely paths.

4.3.3 Exploration Strategy in RL

Given the relaxed solutions, we use them as priors to guide the RL agent in exploring the HPZs in the solution space. In each iteration of the RL process, when a terminal state s_N is reached,

we compute the value of the optimization objective, and use this value as reward:

$$R = \max_{y} f(y|x = s_N). \tag{4.18}$$

This reward is then propagated back through the previous states to update the search policy, ensuring that the agent learns which areas of the solution space yield better results.

From (4.2), we can simplify the transition probability as:

$$P(s_{i+1}|s_i) = P(a_i|s_i). \tag{4.19}$$

Thus, in our RL process, according to (4.3) and (4.19), the action-value function can be written as:

$$\begin{aligned} q(s_i, a_i) &= R_{s_i}^{a_i} + \gamma \sum_{s_{i+1}} P\left[s_{i+1}|s_i, a_i\right] v(s_{i+1}) \\ &= \gamma v(s_{i+1}), \end{aligned} \tag{4.20}$$

where γ is discount factor and $v(s_i)$ is the value of state s_i. Given this, the policy for selecting actions will be given by:

$$\pi(a_i|s_i) = P(a_i|s_i) = \frac{e^{v(s_{i+1})}}{\sum_{s_{i+1}} e^{v(s_{i+1})}}. \tag{4.21}$$

As the agent interacts within the solution space, its policy is updated based on the objective values it receives. Next, we state the following proposition to establish the convergence of the policy.

Proposition 2. *Suppose there exist feasible solutions in the feasible domain and the objective value is bounded. Then, as the number of iterations k tends to infinity, the objective value converges, i.e.,*

$$\lim_{k\to\infty} f(x_{k+1}, y'_{k+1}) - f(x_k, y'_k) \leq \varepsilon, \forall \varepsilon \geq 0.$$

Proof: Let π_{k+1} be the updated policy after an iteration, and π_k represent the previous policy. For states that do not satisfy the constraints, the probability of selecting actions that lead to these states decreases after the policy update:

$$P_{\pi_{k+1}}\left[a_i|s_i, s_{i+1} \in S_T\right] \leq P_{\pi_k}\left[a_i|s_i, s_{i+1} \in S_T\right], \tag{4.22}$$

where S_T is the set of states which are not satisfied with constraints, and $V(s_i|s_i \in S_T) < 0$. For states where the constraints are satisfied, the probability of selecting suboptimal actions increases:

$$P_{\pi_{k+1}}\left[a_i^*|s_i, s_{i+1} \notin S_T\right] \geq P_{\pi_k}\left[a_i^*|s_i, s_{i+1} \notin S_T\right], \tag{4.23}$$

where $a_i^* = \arg\max_{a_i} q(s_i, a_i)$.

The value of state s_i under policy π is:

$$V_\pi(s) = \sum_{a_i} \pi(a_i|s_i)\big[R(s_{i+1}) + \gamma V_\pi(s_{i+1})\big]. \tag{4.24}$$

From (4.19), it can be rewrite as:

$$\begin{aligned} V_\pi(s_i) = &\sum_{a_i} P_\pi\big[a_i|s_i, s_{i+1} \in S_T\big]\big[R(s_i) + \gamma V_\pi(s_{i+1})\big] \\ &+ \sum_{a_i} P_\pi\big[a_i|s_i, s_{i+1} \notin S_T\big]\big[R(s_i) + \gamma V_\pi(s_{i+1})\big]. \end{aligned} \tag{4.25}$$

Substitute (4.22), and (4.23) into (4.25), we have:

$$V_{\pi_{k+1}}(s) \geq V_{\pi_k}(s), \tag{4.26}$$

which means the value function is monotonically increasing. Therefore, the objective value is non-decreasing as the number of iterations increases:

$$f(x_{k+1}, y'_{k+1}) \geq f(x_k, y'_k). \tag{4.27}$$

Since the objective value is bounded, then we have:

$$\lim_{k\to\infty} f(x_{k+1}, y'_{k+1}) - f(x_k, y'_k) \leq \varepsilon, \forall \varepsilon \geq 0. \tag{4.28}$$

□

Proposition 2 indicates that by updating the searching policy based on the rewards obtained from the objective values, the agent will eventually converge. The exploration strategy presented in this section allows the RL agent to efficiently explore the HPZs in the solution space. The framework of the proposed approach is illustrated in Fig. 4.1 and is summarized in Algorithm 3. If an action at any step leads to a state that violates a problem constraint, the episode terminates prematurely, and a penalty reward of $R = -1$ is assigned. This teaches the agent to avoid such paths. For all valid, non-terminal steps within an ongoing episode, the immediate reward is $R = 0$. This ensures the agent focuses on the final outcome rather than the path length.

Algorithm 3 RL-based Optimization Algorithm for 0-1 Mixed Problem

Initialize states value $V(s)$; initialize replay buffer D with capacity d; and let $m = 0$.
Repeat(for each episode):
  Initialize $s \leftarrow \{0\}^{N\times M}$, $i = 0$.
  Repeat(for each step of an episode):
    If $s \notin \mathrm{dom}(V(s))$:
      Solve relaxed sub-problem:
        $\widetilde{x} = \arg\max_{\widetilde{x}\in\widetilde{X}} f(y,\widetilde{x}|s)$.
      Initialize $V(s) = \widetilde{x}$.
    State For a given state s, generate an action $a \leftarrow \pi_v(a|s)$,
    using policy derived form $V(s)$.
    Update $s \leftarrow s+a$,$i = i+1$.
  Until: s is a terminal state or constraints cannot be satisfied.
  Reward

$$R = \begin{cases} -1, & s \text{ is not a terminal state;} \\ \max_y f(y|x=s), & s \text{ is a terminal state.} \end{cases}$$

  Update replay buffer D and states value $V(s_i) \leftarrow \{V(s_i), R\}_{i\in\mathscr{N}}$.
  $m = m+1$.
Until: converges within a prescribed accuracy or a maximum number of iterations has reached.

The complexity of the approach can be analyzed by considering both the computation of HPZs and the iterative RL process. First, obtaining the HPZs involves solving a relaxed optimization problem, which is a convex optimization problem. The complexity of solving this convex relaxation, assuming the use of standard convex optimization algorithms, is typically $O\big((N\times M)^3\big)$ [87]. Then, we define the number of states as $|S|$, the number of actions as $|A|$, the complexity of a single RL policy update is $O(|S|\cdot|A|)$. Over E iterations, the overall complexity becomes $O\big(T\cdot(N\times M)^3 + T\cdot|S|\cdot|A|\big)$. Our approach offers a potential improvement by using relaxed solutions to focus exploration, which can reduce the search space compared to blind exploration. By guiding the agent toward HPZs, we expect a faster convergence to suboptimal solutions. This makes our approach computationally more efficient in cases where the solution space is large, compared to traditional RL methods that explore the entire solution space more uniformly.

4.4 Extensions to Non-convex Scenarios

In wireless network communication scenarios, there are many cases in which the objective function and/or constraints are non-convex. For example, optimizing the throughput via power control in a multi-user environment often leads to a non-convex objective function because of the interference between transmitters. the relationship between transmit power and interference is non-linear. Additionally, constraints related to QoS requirements can often be non-convex, making the optimization problems challenging. In this section, we discuss how to apply our proposed optimization approach to the problems with non-convex objective function and/or constraints.

4.4.1 Non-Convex Objective Function Problems

Optimization problems such as bandwidth allocation, power control, energy efficiency maximization, and multi-carrier resource allocation typically have non-convex objective functions. This is mainly due to the presence of cross terms or fractional terms. Let $\hat{f}(x,y)$ be a non-convex objective function. The proposed solution in Section III cannot be directly applied since it is challenging to obtain the optimal relaxed solutions considering the presence of multiple local optimum and/or saddle points of $\hat{f}(\widetilde{x},y)$. To extend our proposed solution to this case, one straightforward way is using the convex envelope $\text{conv}(\hat{f})(\widetilde{x},y)$ to approximate $\hat{f}(\widetilde{x},y)$. If the convex envelope is difficult to obtain, another efficient way is estimating $\hat{f}(x,y)$ with numerical estimation methods. The double conjugate is commonly employed to transform a non-convex function into a convex approximation. Formally, the double conjugate $\hat{f}^{**}$ of function $\hat{f}$, is defined as:

$$\hat{f}^{**}(z) = \sup_{z^*} \{\langle z^*, z\rangle - \hat{f}^*(z^*)\}, \tag{4.29}$$

where $\hat{f}^*(z^*) = \sup_z \{\langle z^*, z\rangle - \hat{f}(z)\}$. We can reverse it to upper convex envelope for maximization problem

$$\hat{f}_u^{**}(z) = \inf_{z^*} \{-\langle z^*, z\rangle - \hat{f}_u^*(z^*)\}, \tag{4.30}$$

where $\hat{f}_u^*(z^*) = \inf_z \{-\langle z^*, z\rangle - \hat{f}(z)\}$. $\hat{f}_u^{**}$ is the least upper semi-continuous upper convex function that overestimates.

Lemma 2 reveals that $\hat{f}^{**}$ is the convex envelope of $\hat{f}$, providing the closest convex approximation. Then, Proposition 3 demonstrates that the double conjugate does not move the global optimum of $\hat{f}$ while being convex and more tractable for optimization.

Lemma 2. *(Fenchel–Moreau–Rockafellar Theorem [88]) For every function f, whenever f admits a continuous affine minorant, there is*

$$f^{**} = conv(f), \tag{4.31}$$

where $conv(f) := \sup\{g : g \text{ is convex, semi-continuous, and } g \leq f\}$.

The same holds true for f_u^{**}.

Proposition 3. *Assume function* f *admits a continuous affine minorant and has a unique global optimum, then*

$$\max f_u^{**} = f_u^{**}(\arg\max f) = \max f. \tag{4.32}$$

Proof: Assume

$$(\arg\max f_u^{**}, \max f_u^{**}) \neq (\arg\max f, \max f), \tag{4.33}$$

then,

$$f_u^{**}(\arg\max f) > \max f, \tag{4.34}$$

or

$$\max f_u^{**} > \max f. \tag{4.35}$$

It is easy to find a convex function $g \geq f$,

$$\max g = g(\arg\max f) = \max f. \tag{4.36}$$

Then, there exists a convex function

$$g^* = \inf\{g, f_u^{**}\}. \tag{4.37}$$

From (4.34-4.37), we have $g^* \leq f_u^{**}$, which is contradiction to Lemma 2.
Therefore, f_u^{**} and f have the same global optimum point. □

Proposition 3 indicates that the optimization landscape of $\hat{f}^{**}$ retains the global optimal solutions of $\hat{f}$, making it a faithful convex proxy. Thus, solving the convex problem given by $\hat{f}^{**}$ can provide prior information for the RL. Integrating convex estimation method, our approach could be extended to problems with non-convex objective function. As shown in Algorithm 4, an estimation method could be integrated into the steps of obtaining relaxed solutions to provide prior information for RL.

Algorithm 4 RL-based Optimization Algorithm for 0-1 Mixed Problem|Non Convex Objective Function

Same with Algorithm 3 steps 1-5.
    If $\exists \text{conv}(\hat{f})$:
        Solve relaxed sub-problem:
        $\widetilde{x} = \arg\max_{\widetilde{x}} \text{conv}(\hat{f})(y, \widetilde{x}|s)$.
    Else:
        Estimated $\widetilde{x}$ by numerical estimation methods.
    Initialize $V(s) = \widetilde{x}$.
Same with Algorithm 3 steps 9-11.
    Reward

$$R = \begin{cases} -1, & s \text{ is not completed;} \\ \max_y \text{conv}(\hat{f})(y|x=s), & s \text{ is completed.} \end{cases}$$

Same with Algorithm 3 steps 12-14.

4.4.2 Non-Convex Constraints Problems

Besides the non-convexity in the objective function discussed in Section IV. A, let us discuss how to extend the proposed solution to problems with non-convex constraints. Let $\hat{g}_l(x,y)_{l\in\mathscr{L}}$ be the non-convex constraints, which typically lead to a feasible region non-convex. The non-convexity of feasible region makes it harder to search for the global optimum. Three potential ways could be introduced to extend our proposed solution to this case: convex hull [89], Lagrangian dual method [90], and/or arc consistency judgment [91].

The convex hull approach transforms the non-convex feasible region into a convex hull using relaxation techniques, enabling efficient convex optimization [89]. The convex hull retains optimality but compromises feasibility. The Lagrangian dual method handles coupling non-convex constraints by decomposing them into independent subproblems, which are easier to handle independently [90]. The arc consistency method reduces the search space by enforcing constraints between variables. It works by iteratively removing inconsistent values from the variables domains [91]. This process simplifies the problem and helps manage non-convexity.

By applying these three methods individually or in combination, our approach can be extended to problems with non-convex constraints. As shown in Algorithm 5, the constraints, which are difficult to approximate convexly, could be ignored during solving the relaxed problem, making feasible region convex. Then, consider arc consistency in RL when initial state values, setting the values of the states do not satisfy arc consistency to -1. This negative assignment serves as a strong penalty for actions that violate arc consistency. The mechanism is directly linked to our policy selection rule in (4.21), where the probability of selecting an action is proportional to the value of the subsequent state. During the policy calculation, the negative

value ensures that the probability of selecting the invalid action will tend to zero. As a result, the RL agent is steered away from exploring this infeasible path, which prunes this branch from the search space.

Algorithm 5 RL-based Optimization Algorithm for 0-1 Mixed Problem|Arc Consistency Judgment

Same with Algorithm 3 steps 1-4.
    If $s \notin \mathrm{dom}(V(s))$:
        Solve relaxed sub-problem:
            $\widetilde{x} = \arg\max_{\widetilde{x}\in\widetilde{X}} \mathrm{conv}(f)(y,\widetilde{x}|s)$.
        For $j \in \mathscr{M}$:
            If a_i^j does not satisfy arc consistency:
                Let $\widetilde{x}_{ij} = -1$.
        Initialize $V(s) = \widetilde{x}$. State Same with Algorithm 3 steps 8-14.

4.5 Simulation Results and Analysis

In this section, we evaluate the performance of our proposed optimization approach through numerical simulations. Traditional RL approach and B&B approach are selected as benchmarks.

- B&B approach: This approach keeps the same process to obtain the relaxed solution first, then employing depth-first heuristic exploration to determine binary variables. This branching process continues until the binary part of solution obtained from all branches are binary.

- Traditional RL approach: RL is employed to explore the binary part of solutions. The initial values of the states are averaged due to the lack of prior information, which differs from the proposed approach. Arc consistency judgment is used to improve searching efficiency. After obtaining binary solution, the continuous variables are determined using convex optimization.

For ease of comparison, the *normalized objective value* representing the ratio of objective value to the upper bound of the relaxation problem is utilized. Note that this upper bound is not the optimal objective value. Thus it is meaningless to compare this normalized value between different figures (i.e., with different settings), but meaningful to compare it and reveal the performance gain within one figure (i.e., with the same settings). Before presenting the results, we define the key hyperparameter $\alpha \in [0,1]$ as the learning rate. This parameter controls how much the final reward from an episode updates the state-value function, managing the trade-off between faster learning and more stable updates. The simulation codes are built with Python 3.8 and conducted on an Intel Core i7 CPU with 16GB of RAM.

4.5.1 Simulation Results for Convex Relaxed Problems

First, we conduct the simulations for problems with a convex objective function. Consider the below problem as a common use case. For example, this problem can formulate a joint user-association and edge-computing resource allocation problem, as follows:

$$\max_{x,y} \quad \sum_{i}^{N}\sum_{j}^{M}\left[\log(1+s_{ij}x_{ij})+d_{ij}y_{ij}\right] \tag{4.38}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} y_{ij} \leq D_j, \quad \forall j \in \mathscr{M}; \tag{4.38a}$$

$$\sum_{i=1}^{N} c_{ij}x_{ij} \leq C_j, \quad \forall j \in \mathscr{M}; \tag{4.38b}$$

$$\sum_{j=1}^{M} x_{ij} = 1, \quad \forall i \in \mathscr{N}; \tag{4.38c}$$

$$x_{ij} \in \{0,1\}, \quad \forall i \in \mathscr{N}, \forall j \in \mathscr{M}; \tag{4.38d}$$

where x_{ij} is the association-decision variable, y_{ij} is the computing-resource variable, s_{ij} and d_{ij} represent the transmission and computation utility weights, respectively; c_{ij} is the amount of communication resources required, and D_j and C_j represent, respectively,the computation and communication resource budgets at server j.

Figs. 4.2 and 4.3 compare the normalized objective value over the time under two small scales for the three algorithms. From Fig. 4.2 with a problem size $N = 8$, $M = 4$, it can be observed that the objective value of all three approaches increases rapidly at the beginning. For this small-scale problem, the proposed approach and the traditional RL approach yield very similar results. Although the B&B approach shows a lower initial increase, its normalized objective value is slightly higher than the other two approaches. As shown in Fig. 4.3 with a problem size of $N = 20$, $M = 5$, it becomes clear that the proposed approach outperforms the two benchmarks in terms of achieved objective value. The proposed approach reduces the convergence time by 30% compared to the B&B approach, and shows a 5% improvement in the normalized objective value compared to the RL approach, and a 15% improvement compared to the B&B approach.

We then examine the normalized objective value for different value of N (typically N can be treated as the number of users in a network) in problem (4.38), as shown in Fig. 4.4. As B&B approach cannot be deployed in large-scale problems, this simulation still focuses on small scale problem, but the next figure will examine large scale problems without B&B approach. From this figure, we observe that the proposed approach always maintains the highest normalized objective value that is also stable. In contrast, the RL approach experiences a slight decrease in normalized objective value as the problem size increases, but it remains relatively stable overall. The B&B approach demonstrates a degradation as the problem size increases, due to the high computational complexity.

Fig. 4.5 presents the trend of normalized objective value under large scale problems. This figure compares the normalized objective value of the proposed approach under two different learning rate ($\alpha = 0.5$ and $\alpha = 0.9$) with that of the RL approach. It is evident that the proposed

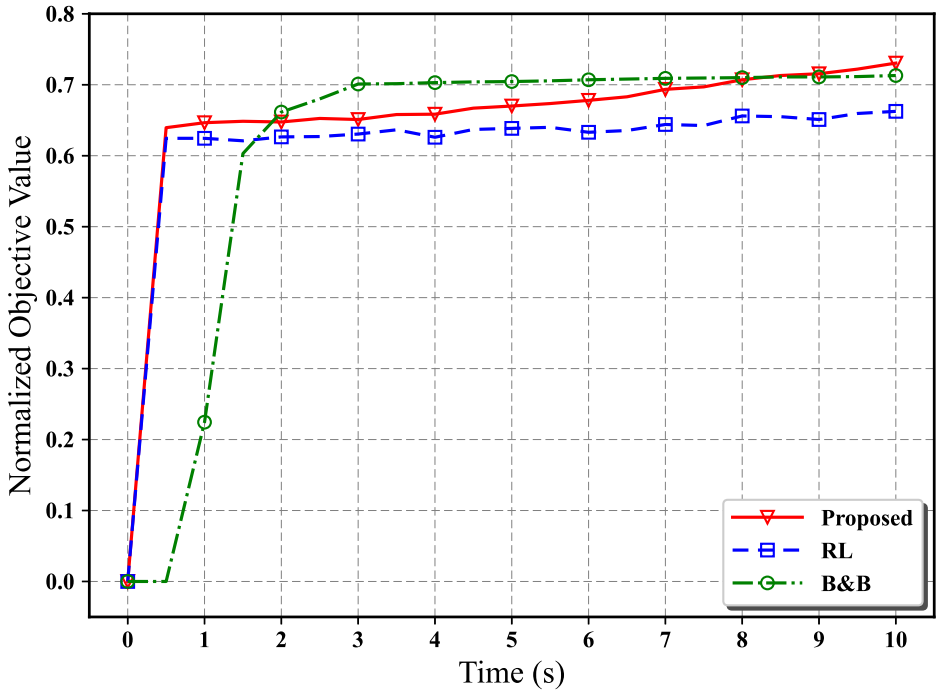


Figure 4.2: Running time vs. normalized objective value ($N = 8, M = 4$).

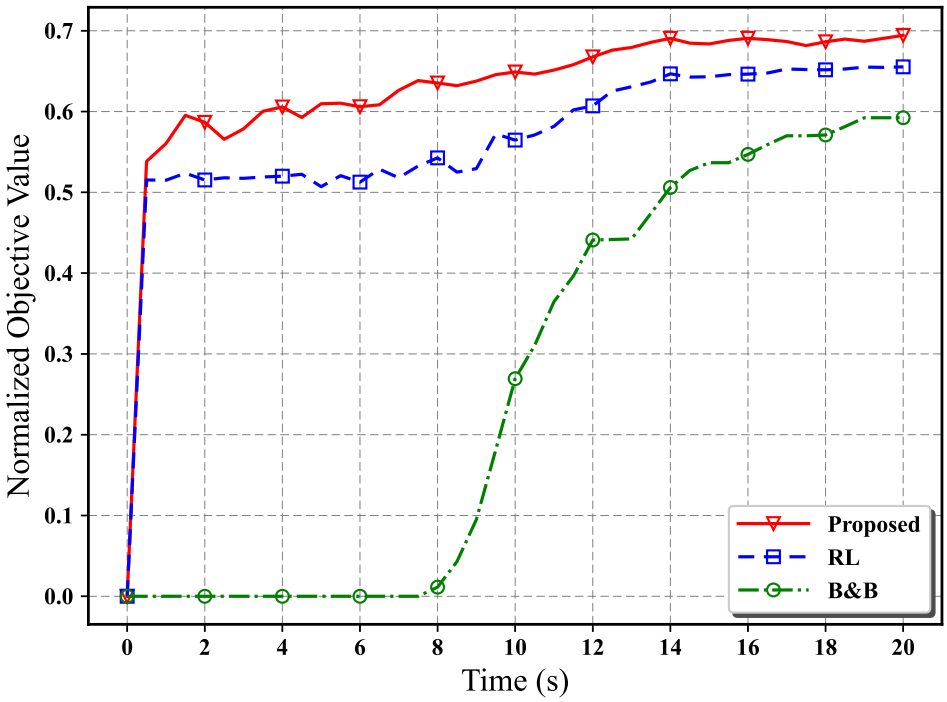


Figure 4.3: Running time vs. normalized objective value ($N = 20, M = 5$).

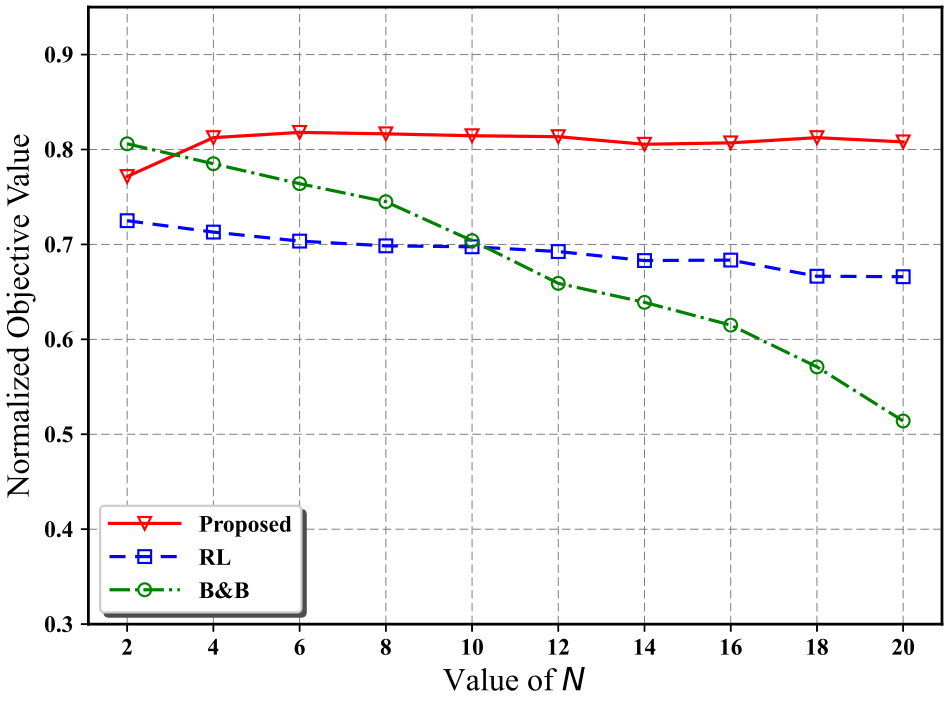


Figure 4.4: Comparison of normalized objective value for different values of N under $M = 5$.

Table 4.1: Simulation Parameters

Parameters	**Values**
Number of BSs (M)	10
Bandwidth limitation of BSs (C)	20 MHz
Density of BSs	16 /km^2
Transmit power of UDs (P_r)	20 dBm
Noise power (P_n)	-114 dBm [92]
Path loss model ($L(d)$)	$L(d) = 34 + 40\log_{10}(d)$

approach, particularly when $\alpha = 0.5$, outperforms the RL approach by 20% when $N \geq 70$. Fig. 4.6 compares the running time across different problem sizes. The running time of the proposed approach is slightly higher than that of the RL approach. This reveals that the proposed approach achieves higher objective value with a minor compromise on computation time. The extra computing complexity mainly lies in solving relaxed subproblems. Note that computing entities in practical networks should be more powerful, which can further mitigate the extra running time.

4.5.2 Simulation Results for Non-Convex Relaxed Problems

Then, we conduct the simulations for non-convex relaxed problems. Consider the below problem as a common use case. For example, this problem can formulate a joint user association and bandwidth allocation, using a logarithmic utility function as the objective function, as follows:

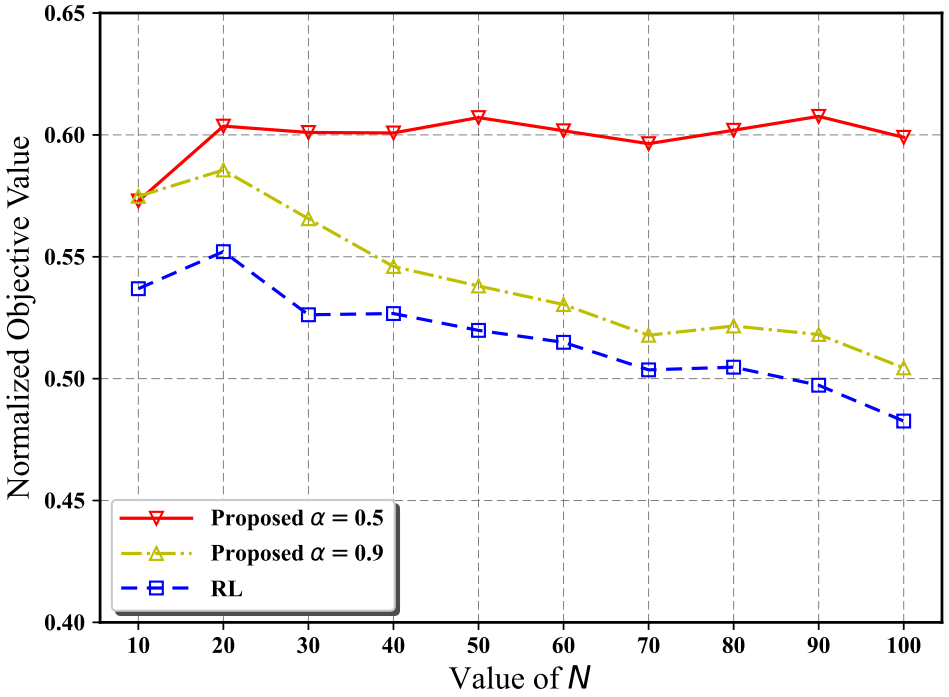


Figure 4.5: Comparison of normalized objective value for different values of N under $M = 10$.

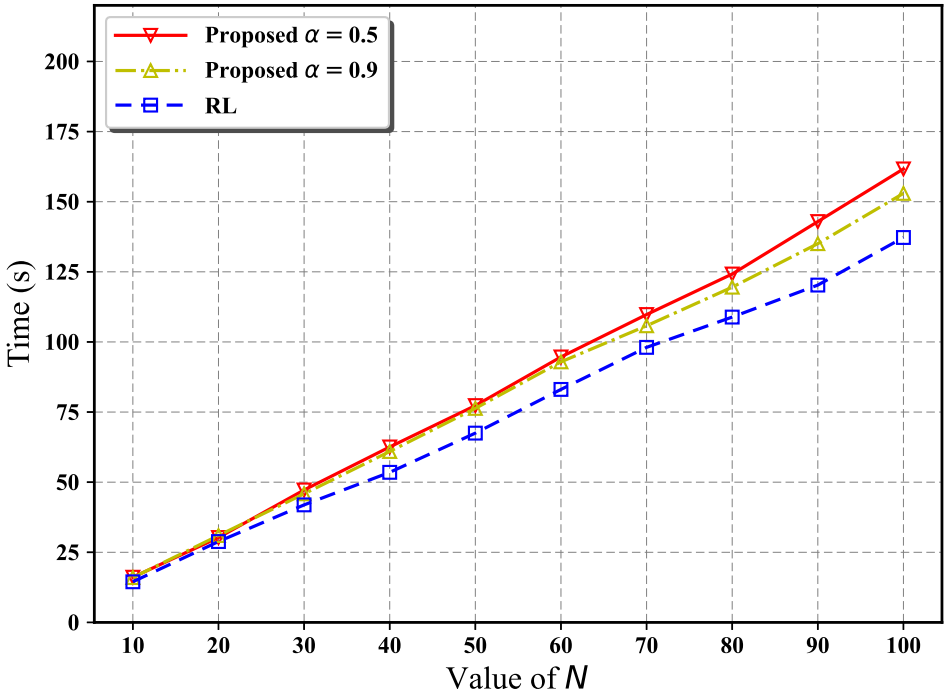


Figure 4.6: Comparison of running time for different values of N under $M = 10$.

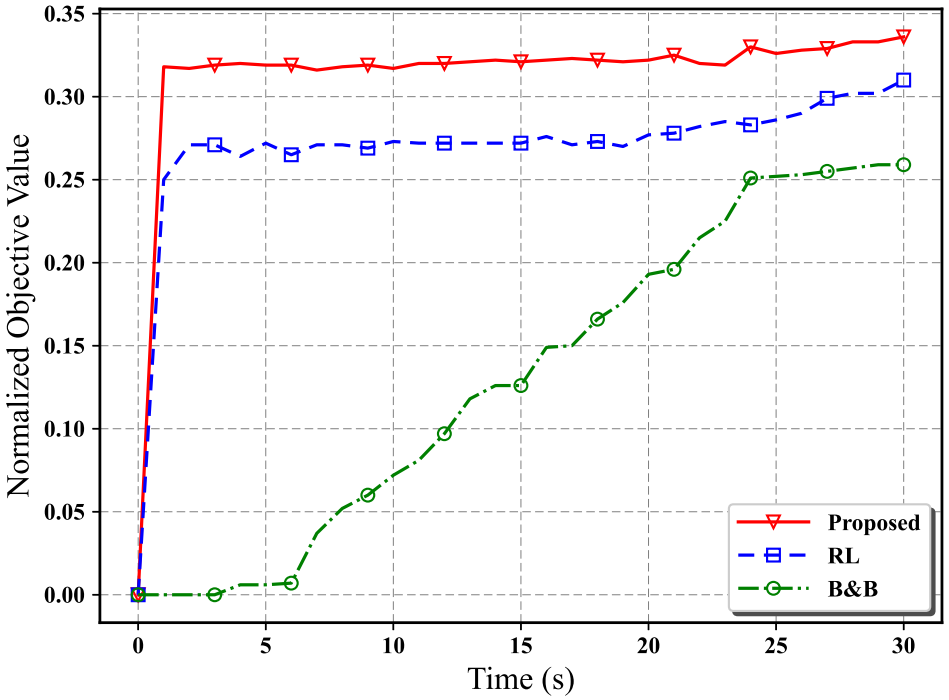


Figure 4.7: Running time vs. normalized objective value ($N = 10, M = 5$).

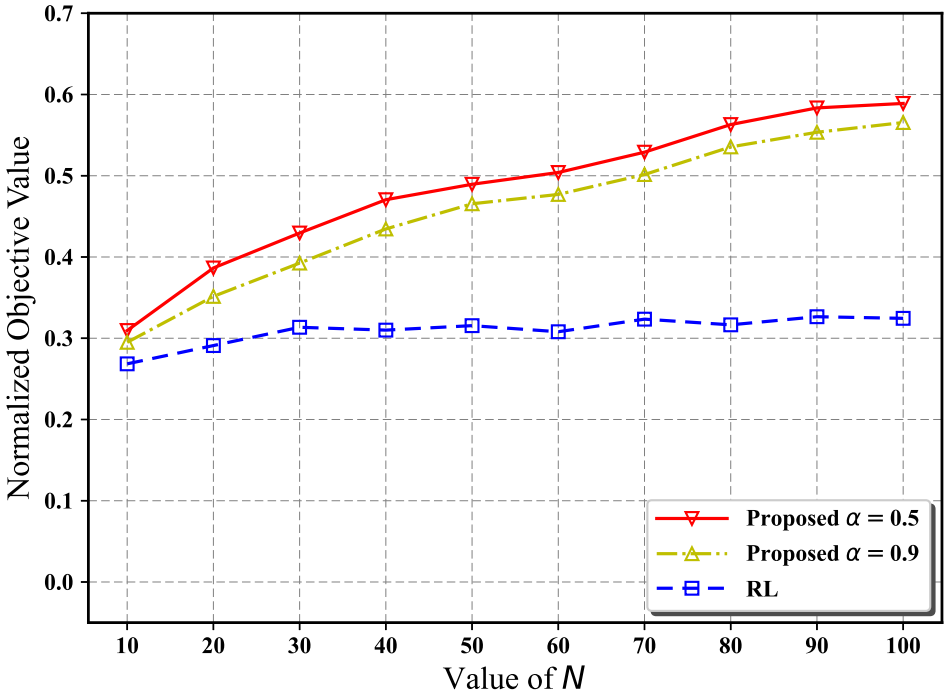


Figure 4.8: Comparison of normalized objective value for different values of N under $M = 10$ and $Q = 0.1$.

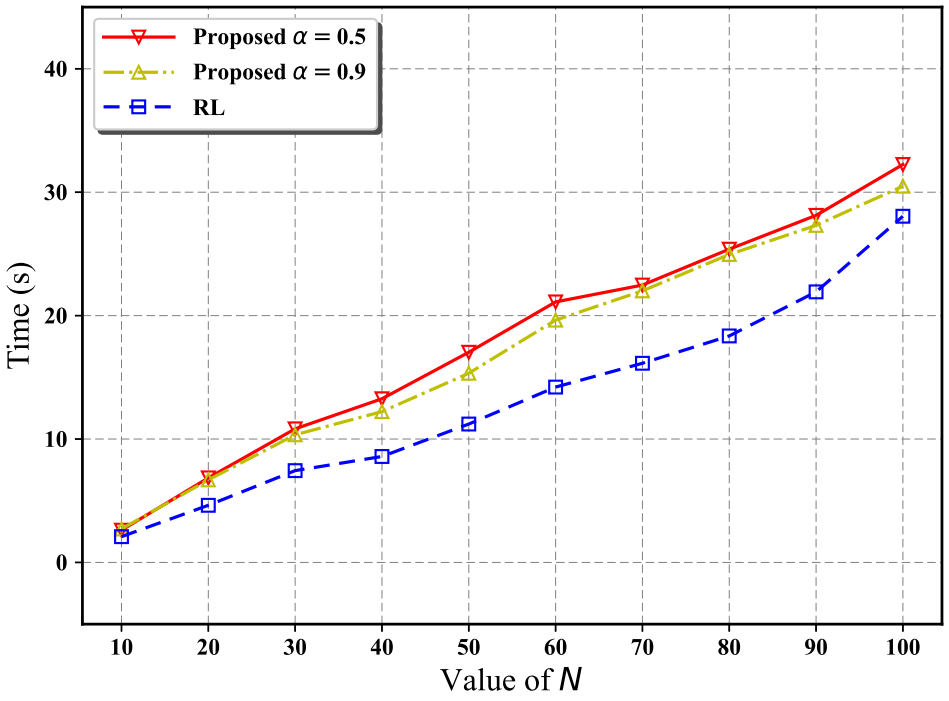


Figure 4.9: Comparison of running time for different values of N under $M = 10$ and $Q = 0.1$.

$$\max_{x,y} \sum_{i}^{N} \sum_{j}^{M} \log \left[x_{ij} y_{ij} \log_2(1+\gamma_{ij}) \right] \tag{4.39}$$

$$\text{s.t.} \quad \sum_{j=1}^{M} x_{ij} y_{ij} \log_2(1+\gamma_{ij}) \leq Q, \forall i \in \mathcal{N}; \tag{4.39a}$$

$$\sum_{i=1}^{N} y_{ij} \leq C, \quad \forall j \in \mathcal{M}; \tag{4.39b}$$

$$\sum_{j=1}^{M} x_{ij} = 1, \quad \forall i \in \mathcal{N}; \tag{4.39c}$$

$$x_{ij} \in \{0,1\}, \quad \forall i \in \mathcal{N}, \forall j \in \mathcal{M}; \tag{4.39d}$$

where x_{ij} is the association-decision variable, y_{ij} is the bandwidth resource variable. To ensure a generic and unbiased evaluation, we model a standard wireless network topology. In the simulations, BSs are placed on a uniform grid within a square area, and UDs are randomly distributed throughout the area. This setup avoids any performance bias that might result from a fixed, structured topology and is representative of typical urban or suburban deployments. The specific parameters for path loss, transmit power, and resource constraints are then applied to this randomly generated topology and are detailed in Table I.

We start with a small size problem with $M = 5$ and $N = 10$, and evaluate the normalized objective value in problem (4.39) over time. As shown in Fig. 4.7, we see that the proposed approach always achieves the highest and most stable objective value compared to the two benchmarks. Moreover, the proposed approach shows a fast convergence speed compared with B&B.

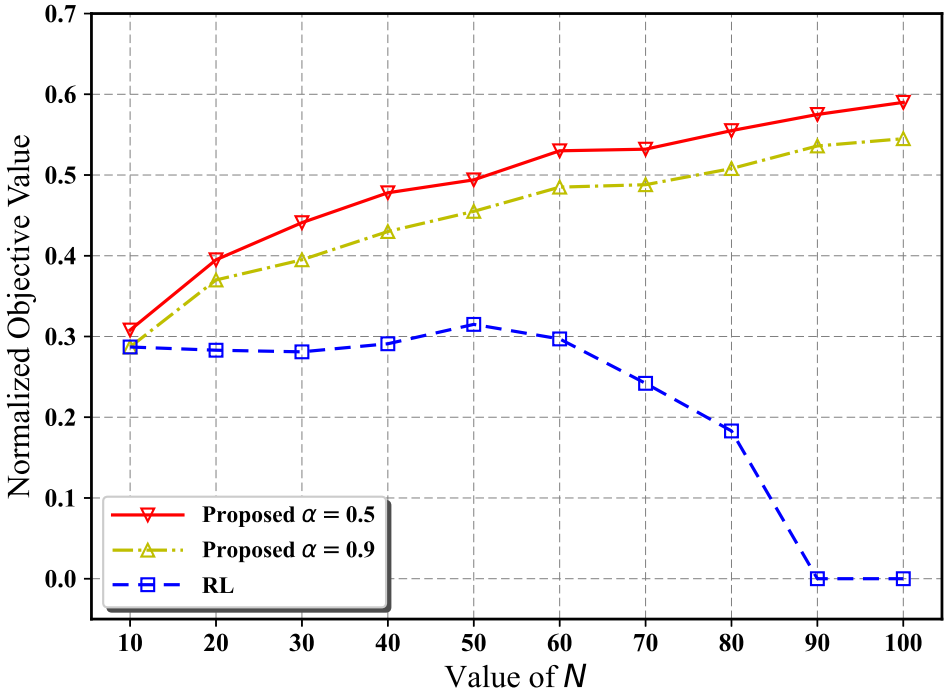


Figure 4.10: Comparison of normalized objective value for different values of N under $M = 10$ and $Q = 0.5$.

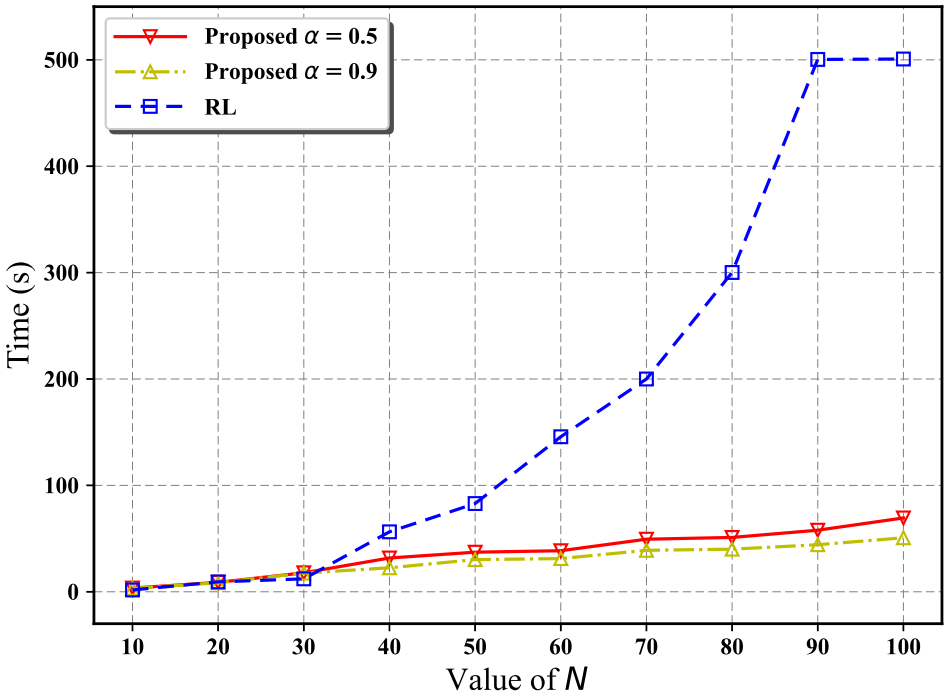


Figure 4.11: Comparison of running time for different values of N under $M = 10$ and $Q = 0.5$.

Moreover, we find that although the RL approach achieves a similar convergence speed, the objective value obtained is lower than our approach. The B&B approach performs the worst due to inefficiencies in branching.

Next, we evaluate the performance for large size problems, where B&B cannot be used because of complexity. Figs. 4.8 and 4.9 show the normalized objective value and running time respectively for different sizes of problem (4.39), with $Q = 0.1$ in (22a). In Fig. 4.8, the proposed approach significantly outperforms the RL approach. When the problem size increases, the normalized objective value of the RL approach remains constant, while the proposed approach shows a consistent upward trend, particularly when $\alpha = 0.5$. This is because RL tends to get trapped in local optima while exploring feasible solutions, whereas the proposed approach efficiently explores the solution space using relaxed solutions as prior information. From Fig. 9, it is observed that the running time of the proposed approach is slightly higher than that of the RL approach. Although the running time of the proposed approach increases by about 30% compared to RL when $N \geq 60$, the computational overhead is acceptable considering the 50% improvement in the objective value.

Figs. 4.10 and 4.11 show the normalized objective value and running time respectively for different sizes of problem (4.39), with $Q = 0.5$ in the constraint (22a). In Fig. 4.10, the proposed approach maintains a high normalized objective value, whereas the RL approach exhibits a significant performance decline. Fig. 4.11 shows that, as the problem size increases, the time needed to execute the RL process increases significantly, far exceeding that of the proposed approach. The normalized objective value of the RL approach decreases to zero as N approaches 90, and the running time reaches 500 seconds, which is set as the unacceptable time threshold. This indicates that RL approach suffers from exploration inefficiency when the constraints become tight and feasible solutions become sparse in the possible solution space. In contrast, the proposed approach maintains a stable objective value, with running time remains controlled.

4.6 Conclusions

In this work, we have proposed an optimization approach that integrates convex optimization with RL in the binary decision-making process to address 0-1 mixed problems. Specifically, we have solved a relaxed sub-problem before making decisions and use the relaxed solution as prior information to guide RL search policy. We have theoretically proven that the neighborhood of the relaxed solution has a higher probability of containing the suboptimal solutions. Hence, relaxed solutions can provide prior information to RL for effective guidance of its searching policy. Furthermore, we have discussed how to extend our approach to deal with non-convex objective functions and non-convex constraints, thus generalizing the proposed approach to complex real-world networking scenarios. Simulation results validated the superiority of our approach under various scenarios, showing a significant improvement in objective value with a minor compromise in computing complexity. In general, this work proposed a unified approach combining RL with optimization theory to efficiently solve a series of 0-1 mixed problems in the field of networking. Our approach bridges the existing gap in methodologies for solving complicated

networking problems in future emerging applications. Finally, it is crucial to reiterate the role of this algorithmic framework within the broader context of Knowledge-Driven Semantic Communication Networks. As formulated in Chapter 3, the optimal scheduling of knowledge consensus involves solving complex NP-hard 0-1 MINLP problems. The relaxation-guided RL framework proposed in this chapter serves as the core computational engine for this thesis.

Chapter 5

Reliability Analysis and Optimization of Knowledge Consensus in Wireless Semantic Networks

5.1 Introduction

In Chapter 3, we established that SemCom relies on Knowledge Consensus to maintain KB freshness. In Chapter 4, we developed a generic solver to schedule these consensus updates. However, both chapters treated the consensus protocol as a black box with a fixed success probability. In this chapter, we address the fundamental question: Given the noisy, error-prone nature of wireless channels, how can we design a consensus mechanism that is both reliable and robust? Unlike wired networks where links are reliable, wireless SemCom networks face probabilistic packet erasures. Therefore, this chapter proposes an availability-robustness analysis framework specifically for wireless knowledge consensus, linking the protocol's physical layer reliability directly to the upper-layer semantic accuracy.

Achieving reliable agreement among a set of participants is a foundational problem in distributed computing, especially when some participants may fail or behave unpredictably [93]. Distributed consensus protocols provide the fundamental building blocks for ensuring such consistency and enabling coordinated operations in decentralized systems. The conventional protocols are used for ensuring reliability in controlled environments, such as enterprise databases within data centers, where node behavior and network conditions were largely stable. Today, the application of consensus system is rapidly expanding into far more uncertain and dynamic scenarios, such as vehicular networks, edge computing infrastructures, and permissioned blockchains. This shift poses challenges to classic protocols, where most existing consensus protocols are designed and analyzed under deterministic assumptions [94]. In such models, each node is either entirely correct or fails in a clearly defined and bounded manner. Protocols such as CFT and BFT rely on fixed fault thresholds, with CFT tolerating up to $f_{\mathrm{CFT}} < \frac{n}{2}$ crash faults [95]

and BFT tolerating up to $f_{\text{BFT}} < \frac{n}{3}$ Byzantine faults [36, 37]. Both models assume a worst-case but well-characterized fault model to ensure safety [96, 97]. However, real-world systems often operate under far more uncertain conditions, particularly in emerging application scenarios such as vehicular networks [98], edge computing infrastructures [99], and permissioned blockchains [100], node failures are frequently caused by benign but unpredictable factors, such as temporary hardware glitches, software bugs, or unstable communication links [101]. In such situations, relying solely on deterministic models can render the system theoretical guarantees of correctness invalid, potentially leading to unpredictable behavior or forcing designers into over-conservative configurations that limit performance and flexibility [102].

5.1.1 Related Work

To address the challenges posed by uncertainty, probabilistic approaches in distributed consensus have emerged [103], offering more realistic and flexible solutions. These researches broadly developed along two distinct lines: efforts to circumvent the Fischer-Lynch-Paterson (FLP) impossibility theorem [104–106], and a more recent focus on developing stochastic fault tolerance models [20, 39, 107, 108]. The FLP theorem drove probabilistic consensus research, which demonstrates that deterministic consensus is impossible in asynchronous systems with even a single crash failure [104]. To circumvent this fundamental limitation, foundational work on randomized Byzantine agreement emerged. These protocols achieve consensus with high probability by leveraging techniques like “coin-flipping" or other random choices within the protocol’s logic. This breaks symmetry and ensures termination or agreement in environments where deterministic solutions would stall, thereby overcoming the limitations of determinism in asynchronous settings [105, 106]. This line of research primarily focuses on relaxing the strict liveness guarantee [97] to a probabilistic liveness guarantee [105] while often maintaining strict safety.

On the other hand, a distinct but equally crucial line of research focuses on adapting consensus to more realistic operational conditions by explicitly embracing stochastic fault tolerance, which proposes changing the fault tolerance assumption from a deterministic limit on the number of faulty nodes to a probabilistic model of node reliability, where reliability is evaluated as the probability of non-faulty behavior. This paradigm shift allows for a more nuanced understanding of system robustness in practical, unpredictable scenarios, moving beyond the binary “faulty or non-faulty" and “fixed fault count" assumptions [109]. Recent surveys highlight this increasing focus on stochastic fault tolerance in distributed systems, encompassing a wider range of probabilistic failure scenarios, from message loss to transient errors, which aligns with the practical challenges faced by modern distributed systems [100]. For instance, [39] was a pioneering work that proposed a distributed consensus model based on Raft, notably adopting a probabilistic failure model to evaluate consensus system reliability. The work of [107] analytically derived the consensus reliability for both PBFT and Raft in wireless autonomous systems

by modeling probabilistic node and communication link failures. Similarly, [108] delved into a dynamic fault model for PBFT, also embracing a node probabilistic failure perspective and leveraging Bernoulli models to derive the consensus success probability. Consistent with this, [101] further investigated the reliability of probabilistic consensus within the wireless communication context, operating under the same node failure assumptions. Going beyond a uniform reliability assumption, [56] pushed the boundaries by relaxing the assumption of homogeneous node reliability, proposing that individual node reliability could follow an independent distribution rather than being uniformly deterministic, which more accurately reflects diverse real-world deployments.

5.1.2 Motivation

In distributed consensus protocols, a quorum represents the minimum number of nodes that must agree on a decision for it to be accepted [110]. In this context, a valid request is an input that conforms to all predefined rules and protocols (e.g., a correctly signed financial transaction), while a non-valid request violates these criteria (e.g., a malformed data packet or a fraudulent transaction). A primary limitation of existing approaches is that defining system reliability via a single performance metric, the overall probability of successful consensus, which is insufficient. In these approaches, the quorum size is selected simply to ensure the success probability meets a predefined threshold [111, 112]. A single metric fails to distinguish between two different types of errors regarding these requests: incorrectly rejecting a valid request and incorrectly accepting a non-valid request. A single success probability simply averages these two distinct risks together. This monolithic view cannot be used to optimize for the nuanced requirements of specific applications. For example, a financial transaction system must prioritize correctness—strongly rejecting all non-valid or fraudulent transactions, even if it means slowing down valid ones. In contrast, a high-throughput IoT data network might prioritize progress—ensuring most valid sensor readings are accepted quickly, even at the risk of occasionally accepting a malformed data packet. This motivates the need for flexible quorum models that account for probabilistic fault behavior.

Another challenge comes from the assumptions made about node behavior. Many existing models assume that the fault probabilities of nodes are independent. While in practice, this is not always true. Malicious nodes may act in coordination, or collude, to disrupt the consensus process. By exploiting weaknesses in quorum selection, such as small group sizes or predictable selection rules, they can strategically influence the outcome to their advantage. It is important to build models that can reflect such strategic behavior and to develop quorum designs that are robust even in the presence of active adversaries.

5.1.3 Contributions

To address the limitations outlined above, this work proposes a quorum optimization framework based on probabilistic modeling of nodes. To capture system reliability under uncertainty and adversarial threats, we introduce two complementary performance metrics. Availability is defined as the ratio of accepted valid requests to the total number of valid requests, reflecting the system's responsiveness. Robustness is defined as the ratio of rejected non-valid requests to the total number of non-valid requests, representing the system's ability to maintain correctness in the face of errors or attacks. These metrics jointly characterize the practical reliability of the system, offering more granular insight than a single success probability. Then, our framework is designed to explicitly manage the trade-off between two key performance: availability and robustness. In this framework, we assume that the honest nodes may fail independently due to benign causes, while the malicious nodes are allowed to collude and strategically disrupt the consensus process.

Building on these metrics, we develop a targeted quorum optimization framework that explicitly quantifies the trade-off between availability and robustness. Our model formulates this as a constrained optimization problem and derives closed-form bounds for both performance dimensions under different quorum sizes. Furthermore, to address the impact of collusive malicious behavior, we introduce a worst-case analysis framework where adversaries can manipulate quorum formation by selectively engaging in malicious actions. This analysis reveals the conditions under which the system can still maintain reliability guarantees, and provides insights into how quorum parameters should be adjusted to defend against such strategic threats.

The main contributions of this work are summarized as follows.

- We propose a dual-metric reliability model based on availability and robustness to provide a comprehensive and practical assessment of consensus system performance under uncertain nodes behavior.
- We develop a quorum optimization framework that considers the trade-off between availability and robustness. The framework is formulated as a constrained optimization problem, and we derive closed-form expressions to guide the selection of quorum sizes in different reliability scenarios.
- We introduce an extended fault model that includes both independent benign failures and collusive malicious behavior. Malicious nodes are allowed to act strategically and collaborate to disrupt consensus, offering a more realistic reflection of potential threats in practical systems.
- We establish worst-case performance guarantees by analyzing the minimum achievable levels of availability and robustness under adversarial strategies. This analysis reveals

the conditions under which reliability can still be maintained and provides insights for designing robust quorum configurations.

5.2 Probabilistic Node Behavior and Consensus Reliability

In this section, we first define the model for probabilistic nodes behavior and then introduce the dual performance metrics, availability and robustness, used to evaluate consensus system performance.

5.2.1 Node Modeling

Consider a consensus system with N nodes denoted as $\mathscr{N} = \{1, 2, \ldots, N\}$. Let $\mathscr{V} = \{v_1, v_2, \ldots, v_N\}$ represent the states of N nodes in the consensus network, where each $v_i \in \{0, 1\}$. Here, $v_i = 0$ indicates that the i-th node is faulty, while $v_i = 1$ denotes that the i-th node is non-faulty. Let $\mathscr{P}_f = \{p_1, p_2, \ldots, p_N\}$ denote the set of fault probabilities for the N nodes, where p_i represents the probability that the i-th node is faulty:

$$P(v_i = 0) = p_i. \tag{5.1}$$

Let f denote the number of malicious nodes in the system. To distinguish between malicious nodes and honest nodes, let $\mathscr{V}^M = \{v_1^M, \ldots, v_f^M\}$ represent the states of malicious nodes, and $\mathscr{V}^H = \{v_1^H, \ldots, v_{N-f}^H\}$ represent the states of honest nodes. The marginal failure probabilities of malicious nodes are denoted as $\mathscr{P}_f^M = \{p_1^M, \ldots, p_f^M\}$, and failure probabilities of honest nodes are represented by $\mathscr{P}_f^H = \{p_1^H, \ldots, p_{N-f}^H\}$. The malicious nodes, modeled as rational agents, maximize their utility functions based on predefined incentive structures [113].

In the system, note that distinguishing between honest and malicious nodes is a non-trivial task. In practice, we cannot directly observe the type of a node (whether it is honest or malicious) based solely on its behavior $\mathscr{V} = \mathscr{V}^M \cup \mathscr{V}^H$. Both types of nodes might fail in similar ways. Consequently, the only observable metric we have for each node is its failure probability $\mathscr{P}_f = \mathscr{P}_f^M \cup \mathscr{P}_f^H$. Thus, the system analysis must rely on failure probabilities $\mathscr{P}_f$ as the primary observable characteristic. The failure probability of each node is an empirical measure, and the system must tolerate this uncertainty when making decisions about consensus.

5.2.2 Consensus Performance Metrics

In the context of a consensus system, a valid request is an input that conforms to all predefined rules and protocols, while a non-valid request violates one or more of these criteria. To systematically analyze the performance of a consensus system, we can frame its results in four distinct

Table 5.1: The Four Quadrants of Consensus Outcomes

	Accept	**Reject**
Valid Request	True Positive: A valid request is correctly accepted.	False Negative: A valid request is incorrectly rejected.
Non-Valid Request	False Positive: A non-valid request is incorrectly accepted.	True Negative: A non-valid request is correctly rejected.

outcomes, which are illustrated in the framework presented in Table 5.1. The four quadrants are defined as follows:

- True Positive (TP): A valid request is correctly accepted, which represents a successful operation.
- False Negative (FN): A valid request is incorrectly rejected, which is a failure that the system fails to make progress.
- False Positive (FP): A non-valid request is incorrectly accepted, which is a failure that may lead to a corrupted.
- True Negative (TN): A non-valid request is correctly rejected, successful protecting the system from error inputs.

To evaluate the effectiveness of a consensus system within this probabilistic framework, we define our dual performance metrics based on the two successful outcomes:
Availability, denoted as η_A, defined as the ratio of the number of passed valid requests to the total number of valid requests:

$$\eta_A = \frac{\text{The Number of Accepted Valid Requests}}{\text{The Number of Total Valid Requests}}. \tag{5.2}$$

A high η_A indicates that the system maintains good responsiveness despite potential node failures.
Robustness, denoted as η_R, defined as the ratio of rejected non-valid requests to the total number of non-valid requests:

$$\eta_R = \frac{\text{The Number of Rejected Non-valid Requests}}{\text{The Number of Total Non-valid Requests}}. \tag{5.3}$$

A high η_R implies strong fault resilience and robustness against erroneous or potentially harmful inputs.

Based on these definitions, our dual metrics directly correspond to the successful outcomes in Table 5.1: Availability (η_A) is the TP rate, and robustness (η_R) is the TN rate. The corresponding failure modes, FN and FP rate, are the mathematical complements of availability ($1-\eta_A$) and robustness ($1-\eta_R$). Therefore, our analysis focuses on optimizing these two primary measures of successful system operation.

5.3 Analysis of Crash Failure

In this section, we provide an analysis of crash failures, which are a type of benign fault in distributed systems. While these faults are not malicious, their occurrence directly impacts the system's ability to form a quorum. We will now use probabilistic framework to quantify how crash failures affect system availability and robustness.

5.3.1 Availability-Robustness Trade-off under Quorum Size

The quorum size Q plays a critical role in the performance of consensus systems, particularly in balancing the goals of availability and robustness. In probabilistic settings, where each node independently fails with a certain probability, the choice of Q determines the number of consistent node responses required to reach agreement. A smaller quorum size generally enhances availability by allowing the system to accept valid requests even when several nodes are faulty or unresponsive. However, this may come at the cost of robustness, as a smaller quorum increases the risk that a majority of responding nodes are faulty and may incorrectly accept non-valid requests.

Conversely, increasing the quorum size improves robustness by requiring a higher level of agreement among nodes, reducing the likelihood that non-valid requests are accepted. However, this makes the system more vulnerable to benign faults, as fewer valid requests may reach the stricter acceptance threshold, thereby reducing availability. This trade-off highlights the need to carefully calibrate Q based on the expected fault distribution and desired reliability guarantees. In the following analysis, we examine how different values of Q influence η_A and η_R, and we explore optimal quorum configurations for maximizing overall system effectiveness.

Let variable $S = \sum_1^N v_i$ denote the total number of non-faulty (i.e., correctly functioning) nodes in the system. From N nodes, the number of combinations of selecting S nodes is $\binom{N}{S}$, and the set of these combinations is represented as:

$$\mathscr{C}(N,S) = \{\{i_1, i_2, \ldots, i_S\} \mid i_j \in \mathscr{N}\}. \tag{5.4}$$

Then, given the fault probabilities $\mathscr{P}_f$, the availability under a quorum size Q can be calculated

as:

$$\begin{aligned}\eta_A &= P(S \geq Q) = \sum_{k=Q}^{N} P(S=k) \\ &= \sum_{k=Q}^{N} \sum_{A \in \mathscr{C}(N,k)} \prod_{i \in A} (1-p_i) \prod_{j \notin A} p_j.\end{aligned} \tag{5.5}$$

Similarly, the robustness can be calculated as:

$$\begin{aligned}\eta_R &= P(N-S < Q) = P(S \geq N-Q) \\ &= \sum_{k=N-Q}^{N} P(S=k) \\ &= \sum_{k=N-Q}^{N} \sum_{A \in \mathscr{C}(N,k)} \prod_{i \in A} (1-p_i) \prod_{j \notin A} p_j.\end{aligned} \tag{5.6}$$

From Equations (5.5) and (5.6), it follows that a smaller quorum size Q tends to increase the availability metric η_A while decreasing the robustness metric η_R. Therefore, Q should be optimized to balance the trade-off between availability and robustness, depending on the specific reliability requirements of the system.

$$\max_{Q} \quad \eta(Q) = \alpha_A \eta_A(Q) + \beta_R \eta_R(Q), \tag{5.7}$$

$$\text{s.t.} \quad Q \in \mathscr{N}, \tag{5.7a}$$

$$\eta_A(Q) \geq T_{\mathscr{L}}, \tag{5.7b}$$

$$\eta_R(Q) \geq T_{\mathscr{S}}, \tag{5.7c}$$

where $\alpha_A, \beta_R \geq 0$ are user-defined weights reflecting the importance of availability and robustness, respectively; $T_{\mathscr{L}}$ and $T_{\mathscr{S}}$ represent minimum acceptable thresholds for availability and robustness, respectively.

5.3.2 Exact and Approximate Calculation Methods

Since $\mathscr{V}$ can be regarded as a set of independent Bernoulli variables, where $v_i \sim \text{Bernoulli}(1-p_i)$, the sum variable S follows a Poisson-Binomial distribution:

$$S \sim \text{Poisson-Binomial}(1-\mathscr{P}_f). \tag{5.8}$$

The probability mass function (PMF) of S, denoted as F_S, can be efficiently computed using a dynamic programming approach. The initial condition is given by:

$$F_S^{(0)} = \{f_0^{(0)} = 1, \ldots, f_k^{(0)} = 0, \ldots, f_N^{(0)} = 0\}. \tag{5.9}$$

The recurrence relations for the PMF are:

$$\begin{aligned} f_0^{(i)} &= p_i f_0^{(i-1)}, \\ f_k^{(i)} &= p_i f_k^{(i-1)} + (1-p_i) f_{k-1}^{(i-1)}, \quad \forall k > 0, \end{aligned} \tag{5.10}$$

where $f_k^{(i)} = P(\sum_{j=1}^{i} v_j = k)$ denotes the probability that the partial sum of the first i Bernoulli trials equals k. This recursive formulation enables the exact computation of the full PMF $F_S = \{f_k^{(N)}\}_{k=0}^{N}$.

Alternatively, when N is large enough, the central limit theorem (CLT) can be used to approximate the distribution of the sum S. According to the CLT, the distribution of S approaches a normal distribution:

$$S \sim \mathcal{N}(\mu, \sigma^2), \tag{5.11}$$

where $\mu = \sum_{i=1}^{N}(1-p_i)$, $\sigma^2 = \sum_{i=1}^{N} p_i(1-p_i)$. Therefore, the probability of approximation can be expressed as:

$$P(S \geq k) \simeq 1 - \Phi\left(\frac{k-\mu}{\sigma}\right), \tag{5.12}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution.

5.4 Analysis of Malicious Behavior

In this section, we extend our model to incorporate malicious tolerance, considering the presence of collusive adversarial nodes. We quantified the lower bounds of availability and robustness that can be guaranteed in the presence of a bounded number of malicious nodes.

5.4.1 Malicious Nodes Behavior

To prevent consensus divergence caused by malicious node attacks, Q should satisfy the following condition [114]:

$$2Q - N > f, \tag{5.13}$$

which means that there is at least one common honest node between two quorums [114]. It should be noted that we assume the honest nodes do not send conflict messages when they fail. Therefore, satisfying condition (5.13) ensures the consensus system will not experience divergence.

To ensure the system's availability and prevent malicious nodes from colluding to block valid requests, q should satisfy the following condition [114]:

$$Q \leq N - f, \tag{5.14}$$

which means that the quorum value should not exceed the number of honest nodes.

Unlike traditional Byzantine fault tolerance (BFT), due to the probabilistic nature of node behavior, even when conditions (5.13) and (5.14) are satisfied, the consensus outcome of the system remains uncertain. In traditional BFT, if a Quorum Q is honest, consensus is guaranteed. However, in our wireless model, satisfying the Quorum condition (selecting enough nodes) is necessary but not sufficient. Even if a valid Quorum is formed, the packet error rate of the wireless links implies that vote messages may be lost during transmission. Thus, the final consensus outcome remains a probabilistic event governed by the joint channel states, leading to the uncertainty described above. Therefore, we focus on two performance metrics, η_A and η_R, to evaluate the system performance with malicious nodes.

Let $v^M \in \{0,1\}^f$ denote the action vector of the f malicious nodes, where v_i^M indicates whether ith malicious node engages in malicious behavior ($v_i^M = 0$) or not ($v_i^M = 1$) during a given consensus round. The ℓ_0-norm $\|v^M\|_0$ thus represents the number of malicious nodes that choose not to engage in malicious behavior in this consensus round. Let $\mathscr{V}_{\mathscr{F}}$ represents the action vector space and $\pi_f\left(v^M\right)_{v^M \in \mathscr{V}_{\mathscr{F}}}$ represent the strategy of malicious nodes, which defines the probability of selecting each possible joint action vector v^M. Thus, the sum of probabilities for all possible joint actions in the sample space must be 1:

$$\sum_{v^M \in \mathscr{V}_{\mathscr{F}}} \pi_f(v^M) = 1. \tag{5.15}$$

Let variable $S^H = \sum_1^{N-f} v_i^H$ denote the total number of honest nodes that are functioning correctly. From $N-f$ honest nodes, the number of combinations of selecting S^H nodes is $\binom{N-f}{S^H}$, and the set of these combinations is represented as:

$$\mathscr{C}(N-f, S^H) = \{\{i_1, i_2, \ldots, i_{S^H}\} \mid i_j \in \mathscr{N}^H\}. \tag{5.16}$$

Then,

$$P(S^H = k) = \sum_{A \in \mathscr{C}(N-f,k)} \prod_{i \in A} (1 - p_i^H) \prod_{j \notin A} p_j^H. \tag{5.17}$$

For malicious nodes, the objective is to disrupt the consensus process as much as possible. Specifically, for a given quorum size Q, their strategy aims to maximize the probability that a

valid request fails to reach the quorum, thereby preventing the system from accepting it:

$$\begin{aligned}
\text{P2}: \max_{\pi_f(v^M)} \quad & P\Big(\|v^M\|_0 + S^H < Q\Big) && (5.18)\\
&= \sum_{v^M \in \mathscr{V}_{\mathscr{F}}} \pi_f(v^M)\, P\big(\|v^M\|_0 + S^H < Q\big), \\
\text{s.t.} \quad & \sum_{v^M \in \mathscr{V}_{\mathscr{F}}} v_i^M \pi_f(v^M) = p_i^M, \quad \forall i \in \mathscr{F}, && (5.18a)\\
& \sum_{v^M \in \mathscr{V}_{\mathscr{F}}} \pi_f(v^M) = 1, && (5.18b)\\
& 0 \le \pi_f(v^M) \le 1, \quad \forall v^M \in \mathscr{V}_{\mathscr{F}}, && (5.18c)
\end{aligned}$$

where constraint (5.18a) ensures that the action policy of each malicious node is consistent with its marginal fault probability and (5.18c) ensures that $\pi_f(v^M)$ is a valid probability value.

While solving P2 can yield the optimal strategy for malicious nodes, we are not concerned with the exact optimal strategy adopted by malicious nodes. Our focus is on the worst-case system performance when malicious nodes act optimally to disrupt the consensus. Specifically, we aim to characterize a lower bound on the availability metric η_A under the most adversarial conditions. Given the quorum size Q and the number of malicious nodes f, the minimum achievable η_A, serves as the availability lower bound, can be obtained by solving the following minimization problem:

$$\begin{aligned}
\text{P3}: \min_{\bar{p}_f, \bar{f}} \quad & \eta_A(\bar{p}_f, \bar{f} | Q) = \bar{p}_f P(S^H + (f - \bar{f}) \ge Q) \\
& \qquad + (1 - \bar{p}_f) P(S^H + f \ge Q), && (5.19)\\
\text{s.t.} \quad & \bar{p}_f \bar{f} = \sum_{i=1}^{f} p_i^M, && (5.19a)\\
& \min\{p_i^M\}_{i=1}^{f} \le \bar{p}_f \le 1, && (5.19b)\\
& 0 < \bar{f} \le f, && (5.19c)
\end{aligned}$$

where $\bar{p}_f$ denotes the frequency of collusive malicious actions, and $\bar{f}$ represents the average number of active malicious nodes involved in such collusive behavior.

When $N - f$ is large enough, the CLT can be used to approximate the distribution of the sum S^H. According to CLT, the distribution of S^H tends to be normal:

$$S^H \sim \mathscr{N}(\mu^H, \sigma^{H^2}), \tag{5.20}$$

where $\mu^H = \sum_{i=1}^{N-f}(1 - p_i^H)$, $\sigma^{H^2} = \sum_{i=1}^{N-f} p_i^H (1 - p_i^H)$. Therefore, based on the approximation

of the probability, the objective function of P3 can be reformulated as:

$$\begin{aligned}\eta_A(\bar{p}_f, \bar{f}|Q) \simeq 1 - \bar{p}_f \Phi\left(\frac{(Q+\bar{f}-f)-\mu^H}{\sigma^H}\right) \\ - (1-\bar{p}_f)\Phi\left(\frac{(Q-f)-\mu^H}{\sigma^H}\right).\end{aligned} \tag{5.21}$$

Similarly, given the quorum size Q and the number of malicious nodes f, the minimum achievable η_R, serves as the robustness lower bound, can be obtained by solving the following minimization problem:

$$\begin{aligned}\text{P4}: \min_{\bar{p}_f, \bar{f}} \quad & \eta_R(\bar{p}_f, \bar{f}|Q) = \bar{p}_f P(N - S^H - f + \bar{f} < Q) \\ & + (1-\bar{p}_f)P(N - S^H - f < Q), \qquad (5.22) \\ \text{s.t.} \quad & \text{(5.19a), (5.19b), (5.19c).} \qquad (5.23)\end{aligned}$$

The objective function of P4, under probabilistic approximation, can be expressed as:

$$\begin{aligned}\eta_R(\bar{p}_f, \bar{f}|Q) \simeq 1 - \bar{p}_f \Phi\left(\frac{(N-f+\bar{f}-Q)-\mu^H}{\sigma^H}\right) \\ - (1-\bar{p}_f)\Phi\left(\frac{(N-f-Q)-\mu^H}{\sigma^H}\right).\end{aligned} \tag{5.24}$$

Let $\underline{\eta_A}$ and $\underline{\eta_R}$ represent the availability and robustness lower bounds, obtained by solving P3 and P4, respectively. The lower bounds of availability and robustness provide quantitative guarantees of the system reliability in the presence of both benign faults and malicious behavior. By leveraging probabilistic approximations, we express these objectives in a tractable form, enabling analytical evaluation and joint optimization.

5.5 Problem Formulation and Proposed Solution

In the previous sections, we established a probabilistic framework and derived the worst-case lower bounds for availability and robustness under adversarial strategies. Building on this analysis, this section now formally presents our solution framework. First, we formulate a multi-objective optimization problem to guide the selection of the optimal quorum size Q that balances the availability-robustness trade-off. Then, we extend this framework to a reputation-weighted quorum model, demonstrating how to enhance system reliability by incorporating node trustworthiness.

5.5.1 Multi-Objective Optimization of Q

For a given probabilistic consensus system, given a fixed number of malicious nodes tolerance f, our goal is to select a quorum size $Q \in \mathscr{N}$ that jointly maximizes the system's reliability, measured in terms of the worst-case availability and robustness guarantees under adversarial behavior. We formulate the problem as a multi-objective optimization:

$$\text{P5}: \max_{Q} \quad \left[\min_{\bar{p}_f, \bar{f}} \eta_A\left(\bar{p}_f, \bar{f} | Q\right), \min_{\bar{p}_f, \bar{f}} \eta_R\left(\bar{p}_f, \bar{f} | Q\right) \right] \tag{5.25}$$

$$s.t. \quad 2Q - N > f, \tag{5.25a}$$

$$Q \leq N - f, \tag{5.25b}$$

This formulation captures the fundamental trade-off between robustness and availability. To solve it, we compute the Pareto frontier over the range of feasible quorum sizes. Each point on this frontier corresponds to a quorum size that yields a non-dominated trade-off between availability and robustness, allowing system designers to select a configuration aligned with specific application priorities.

Proposition 4. *Consider a consensus system with a given set of node fault probabilities $\mathscr{P}_f = \{p_1, p_2, \ldots, p_N\}$ and f malicious nodes, if the quorum size $Q_1 < Q_2$, then the corresponding performance bounds satisfy*

$$\underline{\eta_A}(Q_1) > \underline{\eta_A}(Q_2), \tag{5.26}$$

$$\underline{\eta_R}(Q_1) < \underline{\eta_R}(Q_2). \tag{5.27}$$

Proof: From (5.18) and (5.19), the lower bound of availability for given Q_1 is

$$\begin{aligned} \underline{\eta_A}(Q_1) &= \eta_A(\pi_1^* | Q_1) \\ &= 1 - \sum_{v^M \in \{0,1\}^f} \pi_1^*(v^M)\, P\big(\|v^M\|_0 + S^H < Q_1\big), \end{aligned} \tag{5.28}$$

where π_1^* is the best policy of malicious nodes to break the consensus when quorum size is Q_1.

Given $Q_2 > Q_1$, there is

$$P\big(\|v^M\|_0 + S^H < Q_1\big) < P\big(\|v^M\|_0 + S^H < Q_2\big). \tag{5.29}$$

Then, for quorum size Q_2, when malicious nodes take policy π_1^*, there is

$$\eta_A(\pi_1^* | Q_2) < \eta_A(\pi_1^* | Q_1). \tag{5.30}$$

Thus,

$$\underline{\eta_A}(Q_2) = \eta_A(\pi_2^*|Q_2) \leq \eta_A(\pi_1^*|Q_2) < \underline{\eta_A}(Q_1). \tag{5.31}$$

The similar with $\underline{\eta_R}$, when $Q_2 > Q_1$, we can have

$$\underline{\eta_R}(Q_1) < \underline{\eta_R}(Q_2). \tag{5.32}$$

Therefore, the $\underline{\eta_A}$ is strictly monotonically decreasing with respect to Q, and the $\underline{\eta_R}$ is strictly monotonically increasing with respect to Q. □

Intuitively, from proposition 1, a larger quorum sets a higher threshold for consensus, requiring more nodes to agree and respond. This makes it more difficult for an adversary to create conflicting decisions, thus increasing robustness η_R. However, this same high threshold makes it easier for an adversary to stall the system simply by ensuring not enough nodes respond, thereby lowering the system's ability to make progress η_A. Consequently, the selection of Q is not a matter of simple optimization but a critical balancing act, and this proposition proves that this trade-off is monotonic and inherent to the system's design.

Proposition 5. *Given a multi-objective optimization problem of the form*

$$\max_{Q} \quad \left[\underline{\eta_A}(Q), \underline{\eta_R}(Q)\right] \tag{5.33}$$

defined over a non-empty feasible set $\mathcal{Q}$. Then, every feasible value of the quorum size Q lies on the Pareto front.

Proof: We use proof by contradiction.

Assume, that the proposition is false. This means there exists at least one feasible solution $Q' \in \mathcal{Q}$ is not Pareto optimal. Then, Q' must be dominated by at least one other feasible solution $Q^* \in \mathcal{Q}$, where $Q' \neq Q^*$.

For Q^* to dominate Q', the following conditions must be met simultaneously:

$$\underline{\eta_A}(Q^*) \geq \underline{\eta_A}(Q'), \tag{5.34}$$

$$\underline{\eta_R}(Q^*) \geq \underline{\eta_R}(Q'). \tag{5.35}$$

Case 1: Assume $Q' < Q^*$, then, from (5.26), we have

$$\underline{\eta_A}(Q^*) < \underline{\eta_A}(Q'). \tag{5.36}$$

This directly contradicts condition (5.34), thus, this case is impossible.

Case 2: Assume $Q' > Q^*$, then, from (5.27), we have

$$\underline{\eta_R}(Q^*) < \underline{\eta_R}(Q'). \tag{5.37}$$

This directly contradicts condition (5.35), thus, this case is impossible.

Both possible cases lead to a contradiction. This contradicts the initial assumption that there exists Q' is not Pareto optimal. Therefore, any feasible solution $Q \in \mathcal{Q}$ is Pareto optimal. □

Proposition 2 reveals that the entire feasible set of Q constitutes the Pareto front. This is because Q acts as the monotonic control for the trade-off between availability and robustness, where any gain in one metric leads to a loss in the other. Therefore, the engineering challenge shifts from finding a dominant solution to selecting a value of Q that best aligns with the system's specific priorities and risk tolerance.

5.5.2 Extension to Reputation Weighted Quorum

In this section, we extend the model by incorporating node reputation into the quorum formulation, to enhance the reliability of the probabilistic consensus system. The reputation of a node is an indicator of its trustworthiness and is used to assign a weight to its response in the consensus process. We denote the reputation of node i as r_i, which is assumed to be negatively correlated with its failure probability, such that $r_i = (1 - p_i)$.[1] A reputation score of $r_i = 0$ indicates a completely unreliable node, while $r_i = 1$ represents a fully trustworthy node. When messages are received for quorum formation, we apply a reputation-weighted aggregation to account for the reliability of each source node. With this approach, the quorum is no longer defined by a simple count of nodes but rather by the aggregated reputation weights of the participating nodes. This allows the system to place greater trust in highly reliable nodes.

Let variable $S_r = \sum_1^N v_i r_i$ denotes the sum of reputations of nodes, and let Q_r denotes the reputation-weighted quorum size. Then, for a reputation quorum Q_r, the availability can be calculated as:

$$\eta_A = P(S_r \geq Q_r) = \sum_{S_r=Q_r}^{N_r} P(S_r) \tag{5.38}$$

where $N_r = \sum_{i=1}^N r_i$. The robustness can be calculated as:

$$\eta_R = P(S_r \geq N_r - Q_r) = \sum_{S_r=N_r-Q_r}^{N_r} P(S_r) \tag{5.39}$$

Assume there are f malicious nodes, and the optimal strategy for malicious nodes is π^* under given Q_r. Let $\bar{p_f^*}$ denotes the frequency of collusive malicious behaviors, and $\bar{f_r^*}$ represents the expected total reputation of malicious nodes that are actively engaged in a collusive attack. Then,

[1] In this work, for simplicity, we model each node's reputation as being negatively correlated with its failure probability, allowing the system to weigh messages based on source trustworthiness during consensus.

the lower bound of availability can be approximated as:

$$\begin{aligned}\underline{\eta_A} \simeq 1 - \bar{p_f^*}\Phi\left(\frac{(Q_r + \bar{f_r^*} - f_r) - \mu_r^H}{\sigma_r^H}\right) \\ - (1 - \bar{p_f^*})\Phi\left(\frac{(Q_r - f_r) - \mu_r^H}{\sigma_r^H}\right),\end{aligned} \tag{5.40}$$

where

$$f_r = \sum_{i=1}^{f} r_i^M, \tag{5.41}$$

$$\mu_r^H = \sum_{i=1}^{N-f} (1 - p_i) r_i^H, \tag{5.42}$$

and

$$\sigma_r^{H\,2} = \sum_{i=1}^{N-f} p_i^H (1 - p_i^H) r_i^{H\,2}. \tag{5.43}$$

Similar with (5.24), the lower bound of robustness of the reputation-weighted quorum system can be approximated as:

$$\begin{aligned}\underline{\eta_R} \simeq 1 - \bar{p_f^*}\Phi\left(\frac{(N_r - f_r + \bar{f_r^*} - Q_r) - \mu_r^H}{\sigma_r^H}\right) \\ - (1 - \bar{p_f^*})\Phi\left(\frac{(N_r - f_r - Q_r) - \mu_r^H}{\sigma_r^H}\right).\end{aligned} \tag{5.44}$$

These derived lower bounds for availability (5.40) and robustness (5.44) provide an analytical framework for the reputation-weighted system under worst-case adversarial conditions. This enables us to explore the trade-off between these two metrics, just as we did for the standard model.

5.6 Simulations and Discussions

To validate the theoretical formulations and probabilistic approximations presented in earlier sections, this part conducts a comprehensive set of simulations to evaluate the trade-offs between robustness and availability under varying quorum configurations, malicious node behaviors, and system fault tolerance levels. The results are reported with both linear and logarithmic visualizations to highlight distinct operational regimes, including near-optimal consensus efficiency and extreme reliability boundaries. Particular attention is paid to the role of quorum size, the strategy intensity of malicious nodes, and the upper bound on Byzantine participation, all of which significantly affect system resilience and responsiveness. The Pareto frontier visualizations clearly illustrate the inherent tension between robustness and availability, while the empirical alignment of simulations with dynamic and norm-based analytical models further substantiates the

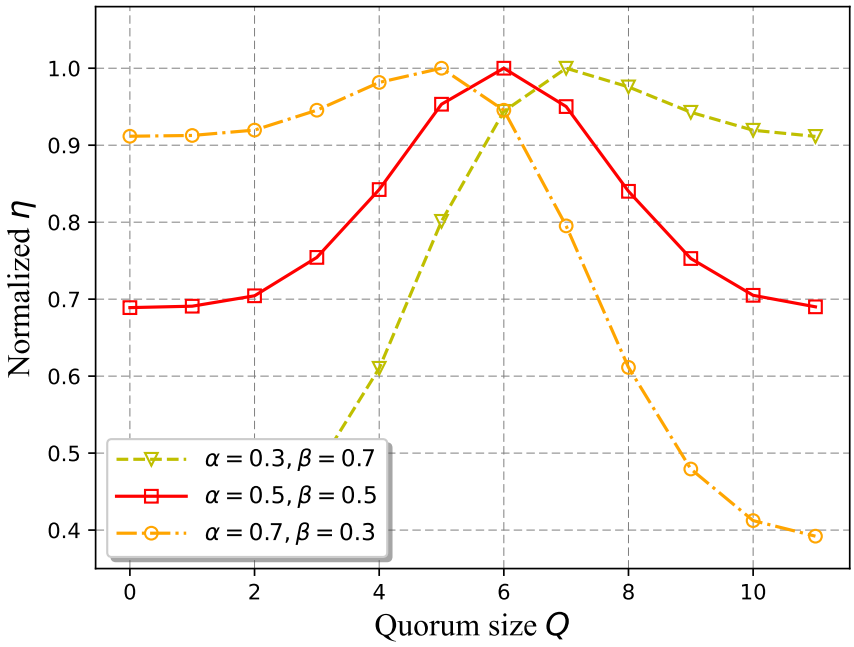
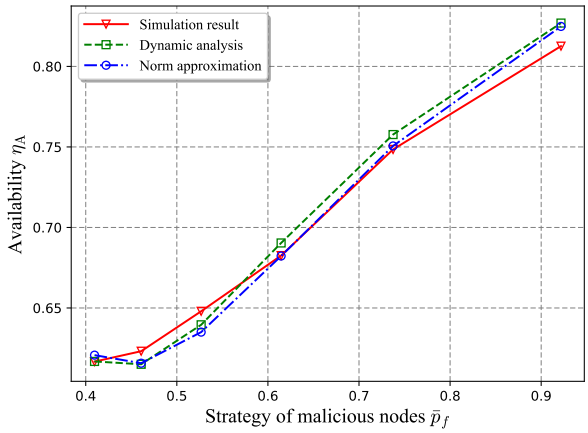


Figure 5.1: Normalized performance versus quorum size.

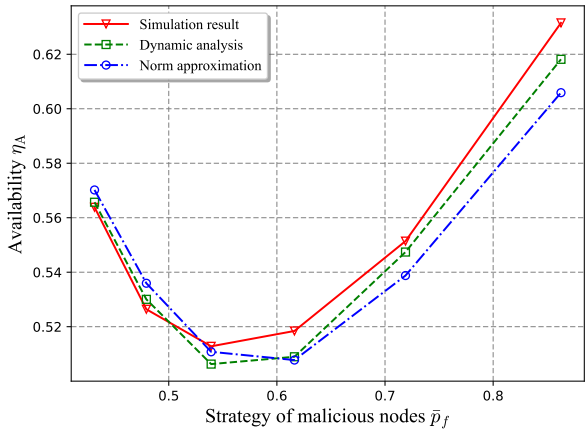
accuracy and applicability of our proposed optimization framework.

Fig. 5.1 illustrates the relationship between the quorum size Q and the normalized consensus performance metric η, from problem (5.7), under varying parameter configurations of α_A and β_R, which jointly characterize the probabilistic model of node behavior. Across all configurations, η exhibits a non-monotonic trend with respect to Q, indicating the existence of an optimal quorum size that balances the trade-off between consensus availability and robustness. For instance, when $\alpha_A = 0.5$ and $\beta_R = 0.5$, η achieves its peak at a moderate quorum size, suggesting a well-balanced system condition. In contrast, more skewed parameter combinations lead to sharper peaks or more rapid declines, reflecting increased sensitivity to quorum selection under imbalanced network conditions. These results highlight the necessity of carefully calibrating quorum sizes according to system-specific probabilistic characteristics in order to maximize consensus performance.

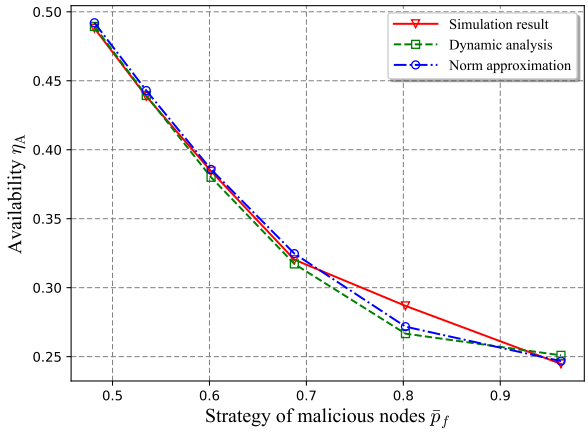
Fig. 5.2 illustrates the impact of varying malicious node strategies $\bar{p}_f$ on the system availability metric η_A, under three different quorum configurations. Each sub-figure compares results obtained via simulation, dynamic probabilistic analysis, and a norm-based analytical approximation. The close agreement among the three curves across all settings demonstrates the reliability of the proposed analytical models in capturing adversarial influence under diverse conditions. The Fig. 5.2(a) corresponds to a relatively lenient quorum configuration, where the required quorum size is low. As $\bar{p}_f$ increases, representing increasingly aggressive behavior from malicious nodes, the availability metric steadily improves. This counterintuitive trend arises because, in low-quorum systems, benign nodes still dominate decision-making unless malicious nodes



(a) Quorum size: $Q = 22$.



(b) Quorum size: $Q = 24$.



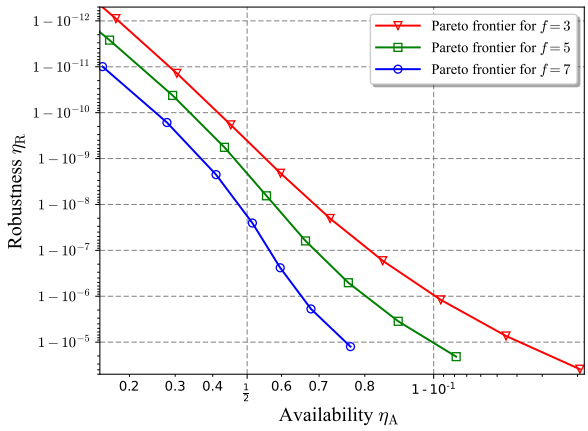
(c) Quorum size: $Q = 26$

Figure 5.2: A comparison of availability (η_A) under varying malicious node strategies ($\bar{p}_f$) across simulation, dynamic analysis, and norm-based approximation, $N = 35$, $f = 8$. The theoretically valid range for the quorum size Q is $[22, 27]$, derived from (5.13) and (5.14) conditions. We therefore analyze three representative configurations.

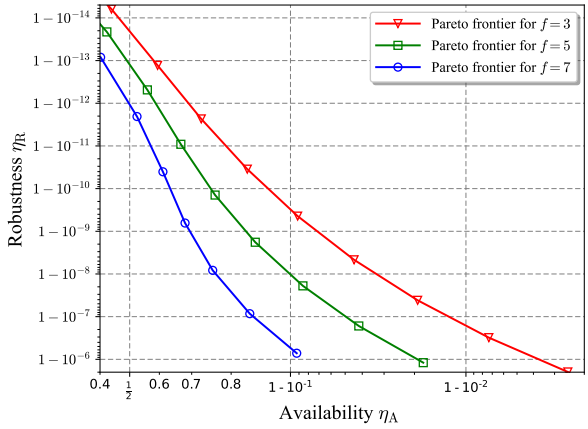
choose a collective action strategy. The Fig. 5.2(b) reflects an intermediate quorum threshold. Here, η_A initially decreases as malicious strategies tends to be more high-frequency dispersed actions, reaching a minimum at moderate $\bar{p}_f$. Beyond this point, availability begins to recover. This U-shaped behavior suggests that moderate malicious strategies are most effective at delaying consensus. When $\bar{p}_f$ is low, the actions of malicious nodes are overly concentrated. While this increases the success rate of attack in a single consensus process, the frequency of such attacks is low. Conversely, when $\bar{p}_f$ is high, their actions become too dispersed. Although the attack frequency increases, the success rate per attempt decreases due to fewer malicious nodes participating in a single consensus round. The Fig. 5.2(c) corresponds to a strict quorum regime, where the required threshold for consensus is high. In this setting, the system is more sensitive to malicious disruption, and the availability metric η_A decreases monotonically with increasing $\bar{p}_f$. This reveals that the optimal adversarial strategy for disrupting consensus is adaptive rather than fixed, varying with system parameters. Consequently, it is essential to model this worst-case behavior to establish the guaranteed lower bounds for system availability and robustness.

Fig. 5.3 illustrates the Pareto frontiers that characterize the trade-off between robustness (η_R) and availability (η_A) under three different tolerances for malicious nodes: $f = 3$, $f = 5$, and $f = 7$. Each curve represents the optimal achievable pairs of (η_A, η_R) for a given malicious nodes threshold, where increasing one metric leads to a degradation in the other. The observed trend confirms that as the quorum size Q is varied, one can trace out a frontier of feasible operating points that balance robustness and availability. As shown, increasing the tolerated number of malicious nodes results in a leftward and downward shift of the frontier. Specifically, for higher f, achieving a given level of availability entails a significantly larger sacrifice in robustness, and vice versa. The nearly linear behavior of each curve in log–log space suggests a power-law relationship between robustness degradation and availability gain. This implies that improvements in consensus responsiveness, even by a small factor, may require exponentially greater risk exposure when fault tolerance demands are high. The visualization thus serves as a design reference, enabling system architects to quantify how aggressively the quorum parameters can be tuned without breaching critical robustness thresholds.

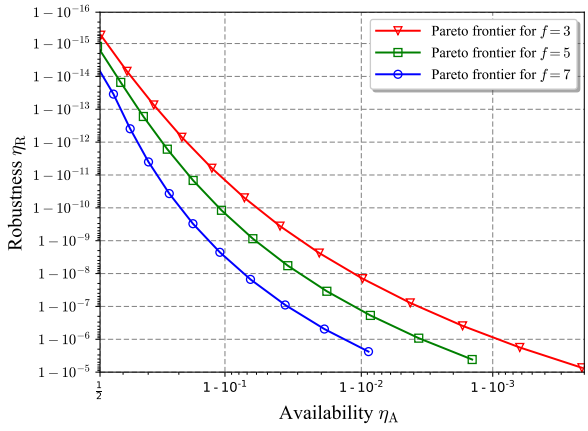
Fig. 5.3(a) and 5.3(b) illustrate the trade-offs between robustness (η_R) and availability (η_A) in systems with expected failure values of 0.2 and 0.1, respectively. This emphasizes the importance of selecting quorum parameters that do not merely balance average-case performance but are aligned with extreme-case reliability constraints. It offers a more precise and practical view of the design trade-offs in robustness-critical systems. Fig. 5.3(c) illustrates the performance of the reputation-weighted quorum mechanism by depicting the Pareto frontier between system availability η_A and robustness η_R under varying numbers of malicious nodes. From the fig. 5.3(c), a clear and inherent trade-off between availability and robustness is observable across all configurations. An improvement in robustness can only be achieved at the expense of availability, and vice versa. Comparing Fig. 5.3(c) with Fig. 5.3(a) reveals that reputation-



(a) A standard consensus system with a mean node fault probability of 0.2.



(b) A standard consensus system with a mean node fault probability of 0.1.



(c) Reputation-weighted quorum system with a mean node fault probability of 0.2.

Figure 5.3: Pareto frontiers illustrating the availability-robustness trade-off under different system parameters and quorum mechanisms. The subplots show the optimal performance for varying numbers of malicious nodes (f=3,5,7).

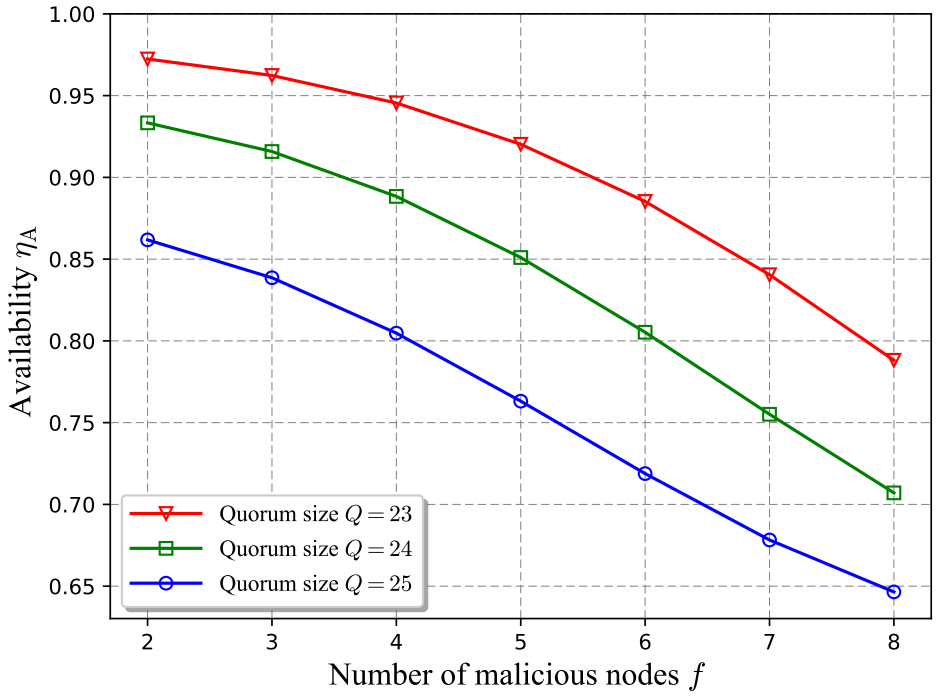


Figure 5.4: Normalized performance versus quorum size.

weighted mechanism delivers significant performance gains. While the Pareto frontier curves of the two figures share similar shapes, their robustness metrics η_R differ in magnitude. Specifically, under identical conditions of availability η_A and number of failed nodes f, the robustness in Fig. 5.3(c) surpasses that in Fig. 5.3(a) by several orders of magnitude. This comparative result demonstrates that reputation-weighted quorum mechanism can significantly enhance the system's robustness against malicious behaviors without compromising system availability.

Fig. 5.4 illustrates the impact of the number of malicious nodes, f, on the system availability, η_A, under three distinct quorum size configurations ($Q = 23, 24, 25$). Two primary trends are evident.First, for any given quorum size, the availability η_A demonstrates a monotonic decrease as the number of malicious nodes f increases. This is expected, as a higher f implies a greater number of nodes that may fail or act adversarially, thus reducing the probability that a sufficient number of non-faulty nodes will approve a valid request. Second, the figure clearly shows an inverse relationship between the quorum size Q and availability. At any fixed value of f, a larger quorum size (e.g., $Q = 25$) results in a significantly lower η_A compared to a smaller quorum size (e.g., $Q = 23$). This is because a stricter quorum threshold imposes a higher burden for consensus, making the system less likely to achieve the required number of responses to accept a valid request.

Fig. 5.5 depicts the system robustness, η_R, as a function of the number of malicious nodes f, with Q as a parameter. As with availability, robustness degrades as the number of malicious nodes f grows. This signifies that as adversarial pressure increases, the system's ability to correctly reject non-valid requests diminishes. However, in direct contrast to the availability

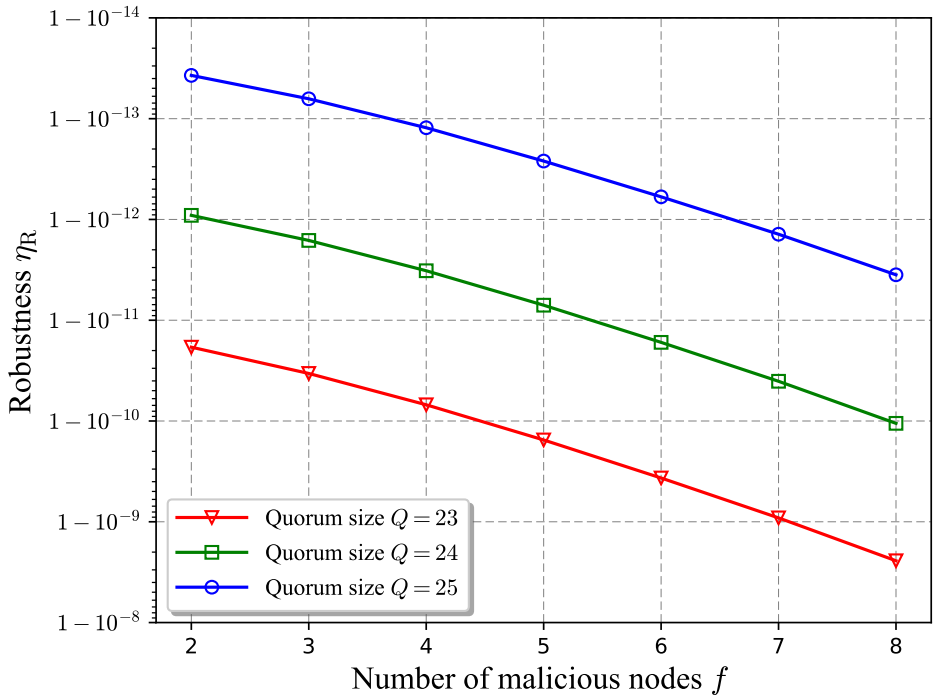


Figure 5.5: Normalized performance versus quorum size.

results, a larger quorum size Q significantly enhances robustness. The $Q = 25$ configuration consistently yields the best robustness, followed by $Q = 24$ and $Q = 23$. A stricter quorum (higher Q) makes it exponentially more difficult for malicious nodes to collude and force the acceptance of a non-valid request.

5.7 Conclusions

This work proposed a probabilistic framework for analyzing consensus availability and robustness under both benign and malicious faults. By modeling each node as a Bernoulli agent with an individual fault probability, we derived closed-form expressions for availability and robustness, revealing their trade-off as a function of quorum size. We formulated two distinct optimization problems: The first models the adversary's behavior to determine the optimal strategy for maximizing the probability of consensus failure, and the second addresses the system designer's challenge of selecting the optimal quorum size to best balance availability and robustness against this worst-case threat. Our analysis captures worst-case scenarios using nested min–max formulations, and simulations confirm the accuracy of the derived bounds. The results offer practical guidelines for quorum sizing in systems like IoT and permissioned blockchains, showing that high reliability can be maintained even in the presence of faults and adversaries. Future work will consider correlated failures, network delays, and adaptive mechanisms for real-time quorum adjustment.

As established in Chapter 3, the decoding accuracy of a SemCom system is strictly dependent on the freshness of the shared KB. The consensus protocol analyzed in this chapter acts as the synchronization engine for the KB. Therefore, the Reliability, related to $p_s(\cdot)$ in equation (3.11), optimized in this chapter is not merely a network-layer metric; it serves as the physical guarantee for semantic fidelity. A failure in the consensus protocol directly translates to a failure in KB synchronization, which in turn leads to high semantic errors at the receiver. By ensuring an optimal trade-off between availability and robustness, the proposed framework minimizes the probability of Knowledge Staleness, thereby enabling the semantic encoder and decoder to operate efficiently even in error-prone wireless environments.

Chapter 6

Case Study: Design and Demonstration of a Visible Light Based SemCom System

6.1 Introduction

In response to the continuous growth of wireless data traffic, VLC has attracted much attention as a complementary wireless technology [9]. By utilizing ubiquitous LED to perform the dual functions of illumination and data transmission, VLC can effectively alleviate the problem of spectrum resource limitation [115]. VLC offers a number of advantages, including vast unlicensed spectrum access, inherent physical security due to limited optical penetration, immunity to electromagnetic interference, and energy efficiency by leveraging existing lighting infrastructure [116]. However, despite these advantages, VLC faces a critical bottleneck in practical deployment: the modulation bandwidth of commercial LEDs is very limited, which severely constrains their data rate and system performance. Therefore, the development of innovative communication techniques that break through the bandwidth limitation is crucial for VLC.

SemCom is a highly promising approach to transcending this bandwidth limitation [5]. Unlike traditional communication that pursues bit-level accuracy in transmission, SemCom focuses on efficiently conveying the meaning or significance of information [7]. This approach can considerably reduce the amount of data required for transmission, thereby improving communication efficiency while maintaining high semantic fidelity, which is particularly beneficial in bandwidth-constrained VLC systems [117]. Therefore, applying SemCom to VLC to build a VL-SemCom system is an effective way to break through the bandwidth bottleneck. In order to support multiple users within a single VLC illumination area, NOMA is a highly applicable technique [118]. NOMA employs power-domain multiplexing to simultaneously serve multiple users on the same time-frequency resource, effectively exploiting the significant differences in channel conditions inherent in the VLC system in order to maximize spectral efficiency [119]. The fusion of SemCom with NOMA promises heightened transmission accuracy and spectral efficiency, as SemCom inherently reduces redundancy and interference among NOMA users,

simplifying the decoding complexity at receivers [120]. Nevertheless, how to build such a VL-SemCom system and develop optimal power allocation strategies with the consideration of the resource budget and quality-of-service (QoS) constraints is a complicated yet important issue.

6.1.1 Related Works

As an emerging field, recent research in SemCom [17, 121–124] has explored a variety of ML models for semantic information extraction and recovery, and preliminarily investigated its advantages in resource-constrained wireless networks. The work of [125] proposed a semantically enhanced two-user uplink NOMA framework, which analyzes the rate region of semantic and bit communication and proposes opportunistic communication strategies under fading channels to strike a balance between performance and interference. The authors in [6] investigated how to balance the bandwidth allocation for data sensing and data transmission in SemCom, by constructing an optimization problem and solving it using the projected gradient method to maximize the semantic spectral efficiency of the system. The authors of [126] explored a semantic-assisted hybrid NOMA system where a far user can switch between SemCom and bit communication (BitCom) to transmit simultaneously with a near user. In [127], the authors introduced a framework that enables the coexistence of multiple semantic and bit users by employing a hybrid NOMA and OMA approach for improved resource management, with the goal of maximizing the system's overall semantic spectrum efficiency.

In the field of VLC, researchers have proposed various techniques to enhance its performance. To solve the multi-user access problem, NOMA is identified as a viable technique for overcoming inherent LED bandwidth limitations. Existing research primarily investigates fixed power allocation (FPA) and gain-ratio-power-allocation (GRPA) strategies for maximizing bit-rates or enhancing fairness under VLC-specific constraints. The work of [120] explored the application of NOMA in VLC downlink networks to improve the data rate and a new power allocation strategy, while the effect of transceiver angle on the performance is investigated. The authors of [128] explored the use of IRS to improve the reliability of NOMA-based VLC systems by jointly optimizing NOMA and IRS parameters, demonstrating significant reductions in bit-error-rate, especially in challenging channel conditions like blockage. The authors in [129] proposed a new method called NGDPA (Normalized Gain Difference Power Allocation) for multiple-input-multiple-output (MIMO)-based NOMA-VLC systems in order to solve the power allocation efficiency and complexity problems. In [23], the authors proposed a smart demodulator based on convolutional neural network (CNN) to solve the problems of signal distortion and difficulty in obtaining channel information in NOMA-VLC system, which combines the two steps of signal distortion compensation and data recovery into one and demonstrates that the system performance can be significantly improved without the need of accurate channel information. However, these studies typically neglect the semantic dimension of transmitted information. Although SemCom and VLC system are actively researched for next-generation

wireless systems, there is still a lack of building a VL-SemCom system and solving the critical problem of resource allocation for multiple semantic users.

While the aforementioned studies have made progress in applying SemCom to NOMA systems, they predominantly focus on conventional radio frequency (RF) networks. For instance, the work of [126] investigates a semantic-assisted hybrid NOMA system where a user can dynamically switch between semantic and bit communication modes. Similarly, the work of [127] proposes a hybrid NOMA/OMA framework to support a heterogeneous mix of semantic and bit users. Our work distinguishes itself from these studies in several fundamental ways. Firstly, we tailor our system design specifically for the VLC environment, introducing a set of physical layer constraints related to illumination and limited LED bandwidth. Secondly, we focus on optimizing power allocation for a system with multiple semantic users, a different challenge from managing hybrid communication modes or heterogeneous user types. Most importantly, unlike the simulation-based validation, a key contribution of our work is the design and implementation of a hardware demonstration platform, which validates the practical feasibility of our proposed VL-SemCom system in a real-world setting.

6.1.2 Motivations and Challenges

The integration of SemCom in VLC systems, leverages the advantages of both technologies, thereby overcoming individual technological limitations and enabling novel applications aligned with future intelligent wireless systems. Specifically, SemCom significantly reduces the required transmitted data, alleviating the primary VLC bottleneck of limited LED modulation bandwidth. Coupled with NOMA, this facilitates efficient multiplexing of semantic streams, enhancing overall bandwidth utilization. The SemCom framework combined with the inherent indoor nature of VLC enables complex multi-user application scenarios. In environments such as smart homes, smart offices, and healthcare facilities, this combination can leverage existing lighting systems to deliver personalized semantic information securely and efficiently. SemCom inherently aligns with future task-oriented communication systems, efficiently conveying relevant information. NOMA and VLC together enable localized communication zones for collaborative tasks, optimizing semantic information distribution while simultaneously maintaining illumination requirements [130]. Transitioning theoretical advancements into practical demonstrations highlights the tangible value of integrating these technologies, motivating this study's extension from previous theoretical explorations into realistic VLC deployments [131].

While the benefits of VL-SemCom systems are expected, designing effective power allocation strategies for this integration still have significant challenges. First, achieving the synchronization required for NOMA across multiple user devices is non-trivial, and the system should remain robust to channel variations caused by shadowing or ambient light fluctuations, which may make statically optimized power allocation schemes ineffective. Second, translating semantic QoS requirements into practical NOMA power levels within VLC's stringent optical

limits and LED nonlinearities poses a substantial challenge, complicating the management of semantic interference among users. Third, extending or redefining semantic throughput metrics suitable for VL-SemCom contexts, accounting for illumination constraints, remains essential for effective power allocation optimization. And last, power allocation becomes a non-convex optimization task, aiming to simultaneously maximize semantic throughput, ensure user fairness, adhere to power budgets, satisfy semantic QoS, and meet illumination constraints. Efficiently finding globally optimal solutions remains challenging.

6.1.3 Contributions and Organization

In this chapter, we investigate building a VL-SemCom system that achieves an optimal power allocation scheme. A preliminary hardware platform is established where a VL-SemCom system is implemented using an Arduino-based LED transmitter and a BPW21 photodiode receiver for pulse-width-modulation (PWM) signal transmission and semantic-aware decoding. Moreover, an optimized power allocation scheme is proposed for this system to maximize throughput and consider fairness at the semantic level. The main contributions are summarized as follows:

- We first develop a comprehensive mathematical model for a VL-SemCom system. The model systematically integrates the features of semantic information, the key constraints of the VLC physical layer, and uses NOMA as the multiplexing technique for multi-user access. In addition, this work defines the semantic transmission rate, thus providing new theoretical perspectives on SemCom systems.

- We next formulate an optimization problem for power allocation considering successive interference cancellation (SIC), QoS requirements, and resource constraints. The objective is to maximize the system throughput in semantic unit (STU) subject to user fairness, total power budget, and VLC-specific operational constraints. Then, we propose a modified water-filling-based algorithm and provide a theoretical proof of its optimality.

- Finally, to bridge the gap between theory and practice, we build and test a preliminary hardware demonstration platform. This platform not only implements the proposed VL-SemCom system but also validates the effectiveness of our power allocation strategy over real VLC links. Both demonstration and simulation results reveal the superiority of the VL-SemCom system in terms of STU and fairness, compared with traditional benchmarks.

6.2 System Model

We consider a downlink multi-user SemCom system operating over a VLC channel, as illustrated in Fig. 6.1. The system consists of one access point (AP), embodied by an LED array, serving N users multiplexed in power domain by means of NOMA. The communication process is assisted

by KB. The AP leverages a shared KB for semantic encoding, while each user has a personal KB for semantic decoding. At the receiver, the channel for each user k is characterized by a channel gain h_k, where $k \in \{1,\ldots,N\}$. After photo detection, a SIC decoder is employed to mitigate interference and detect the user's own signal from the composite stream. Subsequently, the user's semantic decoder recovers the interpreted information. Without loss of generality, we assume that the channel state information (CSI) for all users is perfectly known at the AP to perform NOMA power allocation. For indoor VLC scenarios considered here, as the channels are typically dominated by a stable line-of-sight (LoS) path and change very slowly for static or low-mobility users, allowing for highly accurate CSI estimation [132].

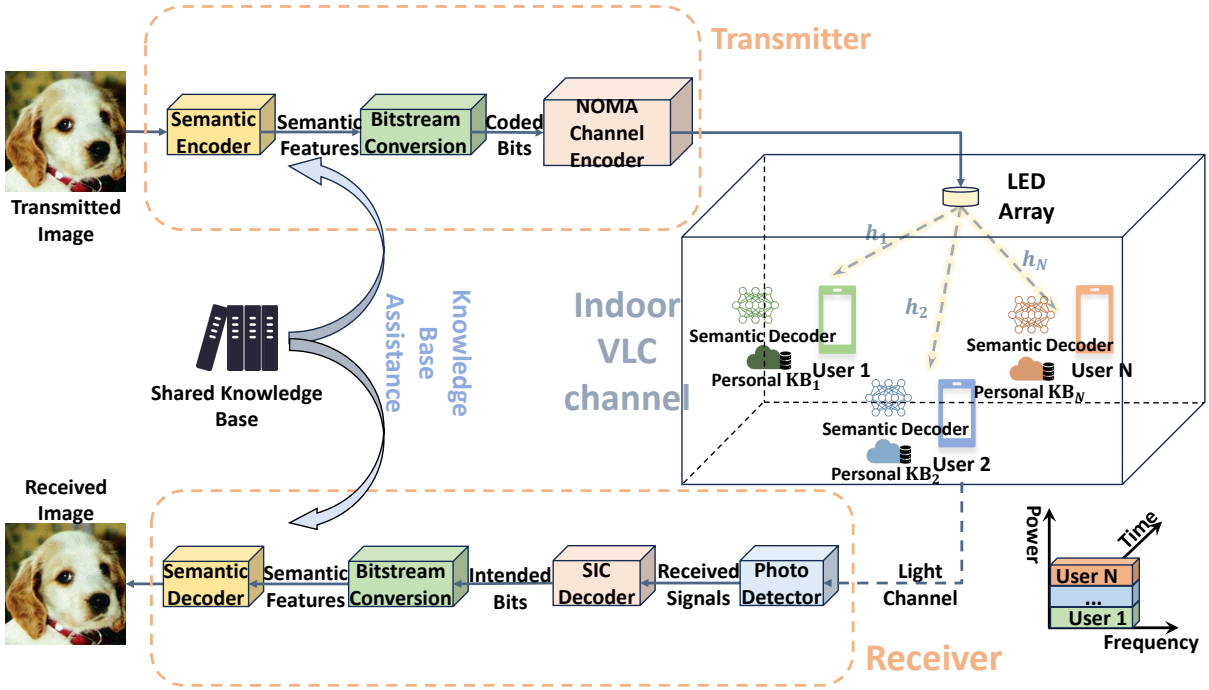


Figure 6.1: A Knowledge-Assisted Semantic Communication Framework for Multi-User NOMA-VLC Networks.

6.2.1 VLC Channel Model

The LoS channel gain from the LED to the k-th user's photo-detector (PD) is given by [133]:

$$h_k = \begin{cases} \mathscr{V}_{k,\mathrm{LED}}^{\mathrm{LoS}} \cos^m(\phi_{k,\mathrm{LED}}) \cos(\psi_{k,\mathrm{LED}}), 0 \le \psi_{k,\mathrm{LED}} \le \Psi_c, \\ 0, \qquad\qquad\qquad\qquad \text{otherwise}, \end{cases} \tag{6.1}$$

where $\phi_{k,\mathrm{LED}}$ is the angle of irradiance with respect to the axis normal to the LED plane; $\psi_{k,\mathrm{LED}}$ is the angle of incidence with respect to the axis normal to the receiver plane, Ψ_c denotes the

field-of-view (FoV) of the receiver, m is the Lambertian order [133]:

$$m = -\frac{1}{\log_2(\cos\Phi_{1/2})}, \tag{6.2}$$

$\Phi_{1/2}$ is the LED half-intensity angle, and

$$\mathscr{V}_{k,\mathrm{LED}}^{\mathrm{LoS}} = \frac{A(m+1)}{2\pi d_{k,\mathrm{LED}}^2}\mathscr{G}_f\mathscr{G}_c, \tag{6.3}$$

where $d_{k,\mathrm{LED}}$ denotes the Euclidean distance from the LED to the k-th user; A is the physical area of the PD; $\mathscr{G}_f$ is the gain of the optical filter; $\mathscr{G}_c$ is the gain of the optical concentrator [120]:

$$\mathscr{G}_c = \begin{cases} \dfrac{\varsigma^2}{\sin^2\Psi_c}, & 0 \le \psi_{k,\mathrm{LED}} \le \Psi_c \\ 0, & \text{otherwise}, \end{cases}, \tag{6.4}$$

where ς stands for the refractive index. Furthermore, from analytical geometry, the irradiance and incidence angles can be calculated as [120]:

$$\cos(\phi_{k,\mathrm{LED}}) = \frac{\mathbf{d}_{k,\mathrm{LED}} \cdot \mathbf{n}_{\mathrm{LED}}}{\|\mathbf{d}_{k,\mathrm{LED}}\|}, \tag{6.5a}$$

$$\cos(\psi_{k,\mathrm{LED}}) = \frac{-\mathbf{d}_{k,\mathrm{LED}} \cdot \mathbf{n}_{\mathrm{PD}}}{\|\mathbf{d}_{k,\mathrm{LED}}\|}, \tag{6.5b}$$

where $\mathbf{n}_{\mathrm{LED}}$ and $\mathbf{n}_{\mathrm{PD}}$ are the normal vectors at the LED and the receiver planes, respectively, and the symbols $\cdot$ and $\|\cdot\|$ denote the inner product and the Euclidean norm operators, respectively.

6.2.2 NOMA Link Modeling

According to the principle of NOMA, the AP allocates a higher power level to the user with lower channel gain and vice versa. Thus, we assume that the users $U_1, U_2, ..., U_N$ are sorted in an ascending order according to their channel gains, i.e., $h_1 \le h_2 \le ... \le h_N$ [120]. Using NOMA, the AP transmits the real and non-negative signals $s_1, s_2, ..., s_N$ with power $P_1, P_2, ..., P_N$, where s_i conveys information intended for U_i. To this effect, the N transmitted signals are superimposed in the power domain as follows [120]:

$$x = \sum_{i=1}^{N} P_i s_i, \tag{6.6}$$

where $\sum_{i=1}^{N} P_i \le P_{\mathrm{MAX}}$ with P_{MAX} being the total transmit power.

The received signal at user U_k is expressed as [120]:

$$y_k = h_k \sum_{i=1}^{N} P_i s_i + n_k, \tag{6.7}$$

where n_k denotes zero-mean additive white Gaussian noise (AWGN) with variance σ_k^2. It is widely assumed that the multi-user interference at user U_k can be eliminated by means of SIC. Based on this, in order to decode its own signal, U_k needs to successfully decode and subtract the signals of all other users with lower decoding order, i.e., $s_1, s_2, ..., s_{k-1}$. Meanwhile, the residual interference from $s_{k+1}, ..., s_N$ is considered negligible and treated as noise [120].

In order to facilitate SIC decoding, the AP should allocate higher transmission power to users with poor channel conditions, i.e., $P_1 > P_2 > ... > P_N$. In the SIC processing, when U_k decodes and removes the signal s_j, the signal-to-interference-plus-noise ratio (SINR) of s_j for this decoding step must be high enough [134].

$$SINR_{U_k}^{s_j} = \frac{P_j h_k}{\sum_{i=j+1}^{N} P_i h_k + \sigma_k^2} \geq \gamma_k, \forall j < k, \forall j, k \in \mathcal{N}, \tag{6.8}$$

where γ_k represent the minimum SINR requirement of $s_j, \forall j < k$ at U_k for successful SIC.

6.2.3 Semantic Channel Capacity Modelling

The classic Shannon capacity expresses the limit of data rate for a channel (in the unit of bits per second, bit/s). For the k-th channel, the Shannon capacity assuming perfect SIC is [120]:

$$C_k^b = B \log_2 \left(1 + \frac{P_k h_k}{\sum_{i=k+1}^{N} P_i h_k + \sigma_k^2} \right), \tag{6.9}$$

where B denotes the allocated bandwidth. It is noted that the Shannon capacity formula cannot accurately capture the semantic transmit rate in SemCom system, due to the distinct features of the semantic channel.

Definition 2. ***Semantic unit (sut):*** *Let a semantic source $\mathscr{Z}_k$ be a set of discrete and mutually exclusive semantic entities $\mathscr{Z}_k = \{z_1, z_2, \ldots, z_N\}$, where each entity z_i represents a unique, indivisible element of meaning (e.g., a word, an object). Define the sut as the fundamental unit to quantify the amount of semantic information, such that each distinct semantic entity $z_i \in \mathscr{Z}_k$ is defined to contain one sut.* [1]

The relationship between a sut and the bits needed to represent it is determined by the semantic entropy of the source. For a source $\mathscr{Z}_k$ with discrete semantic entities z_i, the conventional

[1] In this work, we assume the semantic entities have the same importance. This work could be extended to scenarios that the sut value of entities may different depend on their importance.

Shannon entropy is calculated as [135]:

$$H(\mathscr{Z}_k) = -\sum_{z_i \in \mathscr{Z}_k} p(z_i) \log p(z_i). \tag{6.10}$$

where $p(z_i)$ is the probability of occurrence of entity z_i. This value represents the average number of bits required per sut without any contextual knowledge. Then, to quantify the semantic throughput, we define the sut rate of U_k without KB as:

$$C_k^s = \frac{C_k^b}{H(\mathscr{Z}_k)}. \tag{6.11}$$

In a SemCom system, we model KB_k of $\mathscr{Z}_k$ as conditional probability distributions containing numerous semantic entities along with the dependencies between these entities. These dependencies are quantified through the mutual information, $I(z_i, z_j)$, between neighboring entities, which measures how much the information of one entity reduces the uncertainty of another, i.e.,

$$\begin{aligned} I_{KB_k} &= \frac{1}{2} \sum_{z_i \in \mathscr{Z}_k} \sum_{z_j \in \mathscr{Z}_k} I(z_i, z_j) \\ &= \frac{1}{2} \sum_{z_i \in \mathscr{Z}_k} \sum_{z_j \in \mathscr{Z}_k} p(z_i, z_j) \log \frac{p(z_i, z_j)}{p(z_i) p(z_j)}, \end{aligned} \tag{6.12}$$

where I_{KB_k} is the semantic-relevant mutual information depending on the KB_k. Thus, knowing the information of one entity can decrease the ambiguity and uncertainty of related entities. Based on these dependencies, a SemCom system can transmit information more efficiently. For instance, when transmitting one entity, the system can infer information about other related entities based on the mutual information stored in the KB_k [136]. Assume the transmitter and receiver have the identical KB_k which perfectly contains probability distributions of source $\mathscr{Z}_k$. Then, the semantic information entropy of the source $\mathscr{Z}_k$ with KB_k can be expressed as

$$H^s(\mathscr{Z}_k) = H(\mathscr{Z}_k) - I_{KB_k}. \tag{6.13}$$

In practice, semantic coding is subject to generalization error, which depends on the inference capability of the models used at the source and destination [8]. This error is linked to the amount of data (V_k) available for model training. The error follows an empirically validated power-law form [8]:

$$\varepsilon_k = \alpha V_k^{\beta} + \gamma, \tag{6.14}$$

where α is a constant related to the model architecture, β is the learning rate exponent, and γ is the irreducible error. Therefore, the semantic channel capacity (in the unit of suts per second,

sut/s) for the k-th user with a given bandwidth B can be written as

$$S_k(C_k^b) = \frac{C_k^b}{H^s(\mathscr{Z}_k)} \frac{1}{1+\varepsilon_k}. \tag{6.15}$$

6.3 Power Allocation in VL-SemCom

In this section, we formulate an optimization problem that aims to maximize the STU based on NOMA power allocation, and we investigate how to allocate the limited power resources in the NOMA-based SemCom system.

6.3.1 Problem Formulation

Given the channel conditions and the semantic coding models, we are able to adjust the NOMA power allocation so as to optimize the corresponding achievable sut rate. First, we define *STU* as the objective function, which measures the overall sut rates obtained by all users in the network, expressed as

$$\begin{aligned} STU &= \sum_{k=1}^{N} S_k(C_k^b) \\ &= \sum_{k=1}^{N} S_k \left(B \log_2 \left(1 + \frac{P_k h_k}{\sum_{i=k+1}^{N} P_i h_k + \sigma_k^2} \right) \right). \end{aligned} \tag{6.16}$$

Then, based on the system model, the power allocation optimization problem can be formulated as

$$P1: \quad \max_{P} STU, \tag{6.17}$$

$$s.t. \quad \sum_{i=1}^{N} P_i \leq P_{\mathrm{MAX}}; \tag{6.17a}$$

$$\frac{P_j h_k}{\sum_{i=j+1}^{N} P_i h_k + \sigma_k^2} \geq \gamma_k, \quad \forall j < k, \forall j,k \in \mathscr{N}; \tag{6.17b}$$

$$S_i(C_i^b) \geq C_{min}, \quad \forall i \in \mathscr{N}; \tag{6.17c}$$

where (6.17a) is the total transmit power constraint, (6.17b) is the SINR constraint for successful SIC, and (6.17c) is the minimum sut rate required for each user to avoid unfair power allocation.

6.3.2 Proposed Solution

In order to analyze the concavity of the optimization problem, we calculate the partial derivative of the objective function with respect to P_k. Considering P_k should introduce interference to U_j, $\forall j < k$, the partial derivative of the objective function with respect to P_k can be expressed as

$$\frac{\partial STU}{\partial P_k} = \frac{\partial S_k}{\partial P_k} + \sum_{j=1}^{k-1} \frac{\partial S_j}{\partial P_k}. \tag{6.18}$$

First, let

$$Q_k = \frac{1}{\sum_{i=k+1}^{N} P_i + \sigma_k / h_k}, \tag{6.19}$$

then

$$\frac{\partial S_k}{\partial P_k} = \frac{\partial S_k}{\partial C_k^b} \frac{\partial C_k^b}{\partial P_k} = \frac{1}{H^s(\mathscr{Z}_k)} \frac{1}{1+\varepsilon_k} \frac{B}{\ln 2} \frac{Q_k}{1+Q_k P_k}. \tag{6.20}$$

The partial derivative of interference at U_j can be written as

$$\frac{\partial S_j}{\partial P_k} = \frac{\partial S_j}{\partial C_j^b} \frac{\partial C_j^b}{\partial P_k} = \frac{1}{H^s(\mathscr{Z}_j)} \frac{1}{1+\varepsilon_j} \frac{\partial C_j^b}{\partial P_k}, \tag{6.21}$$

$$\frac{\partial C_j^b}{\partial P_k} = \frac{\partial C_j^b}{\partial Q_j} \frac{\partial Q_j}{\partial P_k} = \frac{B}{\ln 2} \frac{P_j}{1+Q_j P_j} \frac{\partial Q_j}{\partial P_k}, \tag{6.22}$$

$$\frac{\partial Q_j}{\partial P_k} = \frac{-1}{(\sum_{i=j+1}^{N} P_i + \sigma_j / h_j)^2}. \tag{6.23}$$

Substituting (6.22), (6.23) into (6.21), we have

$$\frac{\partial S_j}{\partial P_k} = \frac{1}{H^s(\mathscr{Z}_j)} \frac{1}{1+\varepsilon_j} \frac{B}{\ln 2} \frac{-Q_j^2 P_j}{1+Q_j P_j}. \tag{6.24}$$

From (6.18), it is found that the objective function is non-convex, which makes solving the optimization problem not straightforward. The term $\frac{\partial S_k}{\partial P_k}$ dominates the partial derivative $\frac{\partial STU}{\partial P_k}$ under the condition $\sum_{j=1}^{k-1} \frac{Q_j^2 P_j}{H^s(\mathscr{Z}_j)} < \frac{Q_k P_k}{H^s(\mathscr{Z}_k)}$. Consequently, the objective function STU is concave with respect to P_k, meaning $\frac{\partial^2 STU}{\partial P_k^2} \leq 0$. Conversely, when $\sum_{j=1}^{k-1} \frac{Q_j^2 P_j}{H^s(\mathscr{Z}_j)} > \frac{Q_k P_k}{H^s(\mathscr{Z}_k)}$, $\sum_{j=1}^{k-1} \frac{\partial S_j}{\partial P_k}$ plays a dominant role in $\frac{\partial STU}{\partial P_k}$, which means $\frac{\partial^2 STU}{\partial P_k^2} \geq 0$, as shown in Fig.6.2.

Proposition 6. *For an optimization problem with the objective function* $\max_x \sum_{i=1}^{N} f_i(x_i)$ *and constraint* $\sum_{i=1}^{N} x_i \leq T$*, if* $f_i''(x_i) \geq 0, \forall i \in \mathscr{N}$*, the optimal solution is always achieved at a vertex.*

Proof: Take a cross-section (fix other variables) of the original objective function with re-

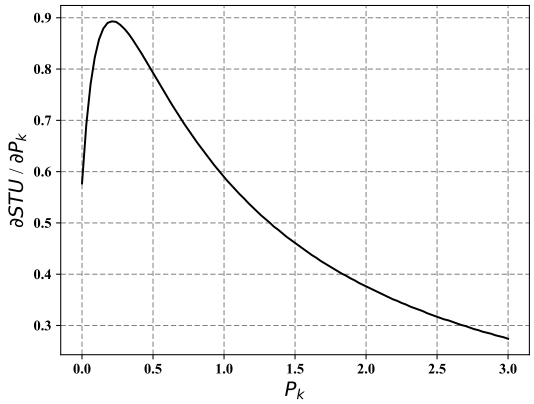


Figure 6.2: Example of $\partial STU/\partial P_k$ vs. P_k.

spect to x_i and x_j, we have

$$\begin{aligned} &\max_{x_i,x_j} f_{ij} = f_i(x_i) + f_j(x_j) \\ &s.t. \quad x_i + x_j \leq T_{ij}, \\ &\qquad x_i \geq 0, \quad x_j \geq 0. \end{aligned}$$

Then, it can be rewritten as

$$\begin{aligned} &\max_{x_i} f_{ij} = f_i(x_i) + f_j(T_{ij} - x_i) \\ &s.t. \quad 0 \leq x_i \leq T_{ij}. \end{aligned} \tag{6.25}$$

Due to $f_i''(x_i) \geq 0, \forall i \in \mathcal{N}$, we have

$$\frac{\partial^2 f_j(T_{ij} - x_i)}{\partial x_i^2} = \frac{\partial^2 f_j(T_{ij} - x_i)}{\partial (T_{ij} - x_i)^2} \left(\frac{\partial (T_{ij} - x_i)}{\partial x_i}\right)^2 \geq 0. \tag{6.26}$$

Then,

$$\frac{\partial^2 f_{ij}}{\partial x_i^2} = \frac{\partial^2 f_i}{\partial x_i^2} + \frac{\partial^2 (f_j(T_{ij} - x_i))}{\partial x_i^2} \geq 0. \tag{6.27}$$

Thus, the optimal solution of the sub-problem is $x_i = 0$ or $x_i = T_{ij}$. Therefore, the optimal solution of the original problem is achieved at a vertex. ■

To optimize power allocation in the proposed VL-SemCom framework, we develop a water filling-based iterative algorithm, as summarized in Algorithm 6. The algorithm aims to maximize the *STU* under a total power constraint. Initially, the transmit power for each user P_i is set to the minimum feasible value constrained by the sub-conditions (6.17b) and (6.17c), ensuring the initialization respects system-level constraints.

The optimization proceeds in two stages, depending on the concavity of the *STU* function. In the first stage, where the second-order derivative $\frac{\partial^2 STU}{\partial P_k^2} \geq 0$ holds, the user with the steepest ascent (i.e., largest gradient ∇_i) is selected to incrementally increase its power P_k. Meanwhile, other users receive proportional updates based on their respective constraints. If the sum power

Algorithm 6 Water Filling-based Optimization Algorithm for Power Allocation

Initialize $p_i^0 = max\{min(6.17b), min(6.17c)\}$; and let $m = 0$;
While $\partial^2 STU/\partial P_k^2 \geq 0$:
    $\nabla_i = \frac{\partial STU}{\partial p_i}$;
    $k = \arg\max_i \nabla[i]$;
    Repeat:
        $p_k^{m+1} = p_k^m + \Delta_k$;
        $\Delta_i = max\{min(6.17b), min(6.17c)\}, \forall i < k$;
        $p_i^{m+1} = p_i^m + \Delta_i$;
        If $\sum_{i=1}^N P_i > P_{\text{MAX}}$:
            reset $p_i^{m+1} = p_i^m$;
            $\Delta_k = \Delta_k - \zeta$;
    Until: $\sum_{i=1}^N P_i \leq P_{\text{MAX}}$.
    $m = m + 1$;
While $\partial^2 STU/\partial P_k^2 \leq 0$:
    Set $\mathscr{L} = \lambda$;
    $k = N$;
    Repeat:
        While $STU'(P_k) > \mathscr{L}$:
            $p_k^{m+1} += \Delta_k$;
        $k = k - 1$
    Until: $k = 0$.
    If $\sum_{i=1}^N P_i > P_{\text{MAX}}$:
        reset $p_i^{m+1} = p_i^m$;
        $\mathscr{L} = \mathscr{L} + \Delta_\lambda$
    else: $\mathscr{L} = \mathscr{L} - \Delta_\lambda$
    $m = m + 1$;
Until: Termination condition reached or $\sum_{i=1}^N P_i = P_{\text{MAX}}$.

exceeds the maximum available power P_{MAX} , the updates are rolled back and the step size Δ_k is adaptively reduced to maintain feasibility.

In the second stage, triggered when $\frac{\partial^2 STU}{\partial P_k^2} \leq 0$, a Lagrangian-based refinement is applied. The algorithm sets a dual threshold $\mathscr{L} = \lambda$ and backtracks from the last user, gradually increasing P_k as long as the marginal utility $STU'(P_k)$ remains above $\mathscr{L}$. The dual variable λ is updated in each iteration to satisfy the power constraint. This two-phase strategy ensures convergence toward a locally optimal power allocation that balances individual semantic utility gains with overall resource constraints.

Therefore, in $\frac{\partial^2 STU}{\partial P_k^2} \geq 0$ domain, allocate power based on Proposition 1, and in $\frac{\partial^2 STU}{\partial P_k^2} \leq 0$ domain, allocate power using water filling-based algorithm depend on equation (6.18), as shown in Algorithm 6. Since allocating power to user U_k introduces interference to all users U_j with a lower decoding order ($j < k$), the water-filling process is performed starting with user U_N. Considering the domain item is $\frac{\partial STU}{\partial P_k}$, allocate to U_N first could reduce the impact of subsequent allocations on prior allocations, although allocate power to U_j will affect the water level of

$U_k, \forall k > j$.

6.3.3 Complexity Analysis

To evaluate the practicality of our proposed power allocation scheme, we analyze its computational complexity and compare it with the benchmark FPA and GRPA strategies. The complexity of our proposed Algorithm 1 is primarily determined by its iterative structure. The algorithm seeks an optimal solution by progressively refining the power allocation over multiple iterations. Let I_{iter} be the total number of iterations required for the algorithm to converge. In each iteration, the core operations involve calculating the gradients of the STU function and updating the power variables for all N users. Consequently, the computational complexity of our algorithm is on the order of $O(I_{\text{iter}} \cdot N)$.

In contrast, the FPA and GRPA benchmarks are non-iterative and computationally much simpler. The FPA scheme allocates power according to fixed, predefined ratios. The GRPA scheme allocates power based on ratios calculated directly from the users' channel gains2. For both methods, the power assigned to each of the N users can be determined in a single computational pass once the total power budget is known. Therefore, their complexity is linear with the number of users, i.e., $O(N)$.

It is important to note that this increased computational load is handled by the AP, which is mains-powered and equipped with capable processors, making the complexity of our iterative algorithm manageable. Furthermore, the indoor VLC is often characterized by static or low-mobility users, leading to channel conditions that change slowly. Thus, the power allocation vector does not need to be re-calculated at a high frequency, which further reduces the average computational demand on the AP. Therefore, while our algorithm is more computationally intensive than FPA and GRPA, it achieves a superior solution under these practical considerations.

6.4 Hardware Demonstration platform of VL-SemCom

To validate the feasibility of the proposed VLC framework in a practical environment, we implement a VL-SemCom demo using off-the-shelf components and open-source tools. The system consists of Arduino-based transmitter and receiver, configured to support pulse-level visible light modulation and intensity decoding.

6.4.1 Transmitter Design

The transmitter hardware employs an Arduino UNO, which supports multiple 8-bit PWM channels operating at 490–980 Hz depending on the timer configuration. By reconfiguring the timer

registers, the PWM frequency of the Arduino UNO can be configured up to a theoretical maximum of 62.5 kHz; in our implementation, a frequency of 10 kHz is adopted to balance transmission speed and response limitations. A single high-brightness white LED is driven via a transistor-based current amplifier circuit, where the PWM output pin modulates light intensity in accordance with the encoded composite signal. The Arduino generates PWM waveforms by modulating the aggregate NOMA-encoded signal, where multiple users' semantic symbols are weighted and summed prior to transmission.

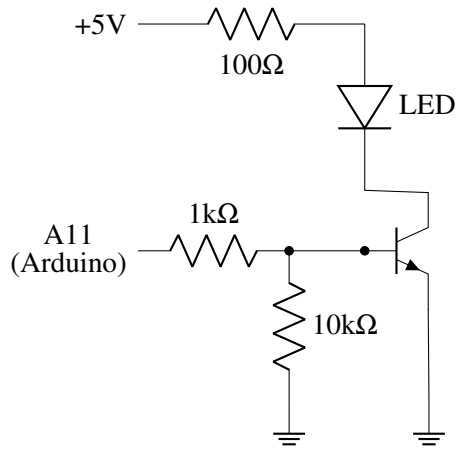


Figure 6.3: Schematic of the VLC transmitter circuit.

At the software level, source data is first semantically encoded on a PC using lightweight encoders such as principal component analysis (PCA). The encoded semantic vectors are then partitioned per user and prioritized based on their task utility contribution. A NOMA power allocation is applied to assign higher weights to users with lower channel quality. Let $\alpha_i \in (0,1)$ represent the corresponding power weight of user i determined offline. The superimposed transmit signal is generated as:

$$x = P_{\text{MAX}} \sum_{i=1}^{N} \alpha_i s_i, \tag{6.28}$$

where $\sum_i^N \alpha_i \leq 1$. The superimposed signal is quantized into discrete intensity levels and transmitted as a sequence of PWM signals. Each frame includes a short synchronization header followed by the superimposed semantic payload. The Arduino microcontroller stores the PWM duty cycle sequence in SRAM and transmits each symbol frame at fixed intervals. Although the UNO has limited memory and clock speed, this architecture supports low-rate VLC transmission suitable for demonstrating the feasibility of SemCom with NOMA in constrained platforms.

6.4.2 Receiver Design

The schematic of the receiver front-end is shown in Fig. 6.4. A BPW21 photodiode is reverse-biased by a 5V supply and connected in series with a 100kΩ resistor. The photodiode converts the incident optical signal into a photocurrent, which is translated into a voltage signal via the

load resistor. This voltage reflects the light intensity modulated by the transmitter's PWM output. To smooth out high-frequency fluctuations and suppress noise, the voltage signal is further filtered by a passive RC low-pass filter composed of a 1kΩ resistor and a shunt capacitor. The filtered analog signal is then sampled by the Arduino analog input (A0), which employs a 10-bit ADC with a reference voltage of 5V.

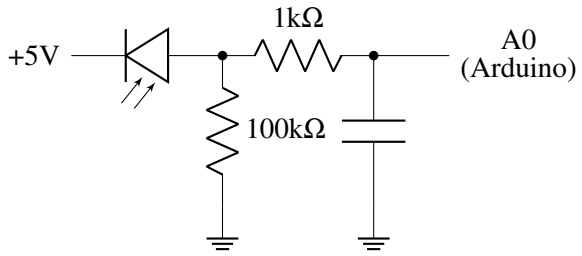


Figure 6.4: Schematic of the VLC receiver circuit.

Once digitized, the analog samples are processed to estimate the duty cycle of each PWM period. To separate individual user signals, the receiver applies a SIC process. The symbol with the highest power level is decoded first, based on the known power allocation pattern. Its estimated contribution is subtracted from the aggregate signal, allowing the next lower-power user's symbol to be extracted. The decoded symbols are then mapped back to semantic vectors, which subsequently serve as input for a pre-trained diffusion-based reconstructor to generate the final output.

6.4.3 Semantic Transmission over VLC with Diffuser-Based Reconstruction

To evaluate the performance of semantic image transmission over VLC, we conduct a proof-of-concept experiment where RGB images of size $224 \times 224 \times 3$ are first compressed and then transmitted through a VLC channel. At the transmitter, to balance data reduction with feature preservation for our proof-of-concept demo, we employed image resizing as an efficient compression method. Each $224 \times 224 \times 3$ image was downsampled to a compact representation of $75 \times 75 \times 3$ using the LANCZOS resampling algorithm (via Image.Resampling.LANCZOS from the Python PIL library). This high-quality interpolation method is known for preserving sharp details and effectively reduced the transmitted data size by 88.8%.

The compressed images are then modulated using intensity modulation and transmitted through a VLC link implemented with LED and photodiode components. Transmission occurs in a controlled indoor setting to ensure LoS conditions. The receiving end captures the optical signal and reconstructs the compressed image using analog-to-digital conversion followed by signal decoding. For image restoration, we utilize a diffusion-based generative model (Diffuser) as the semantic decoder. After receiving the $75 \times 75 \times 3$ image, the Diffuser performs a progressive denoising process to recover the original image resolution and structure.

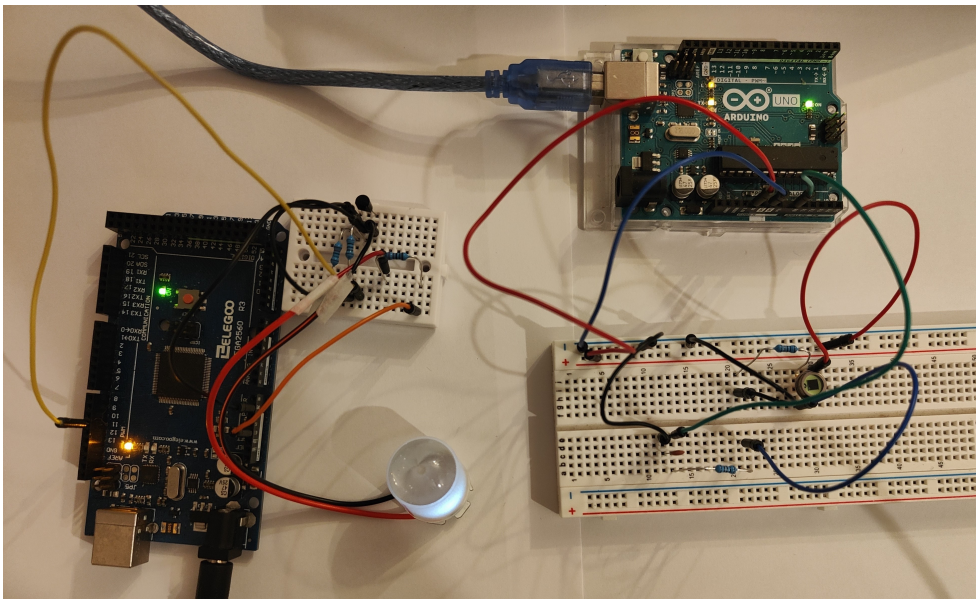


Figure 6.5: The hardware demonstration platform for algorithm validation.

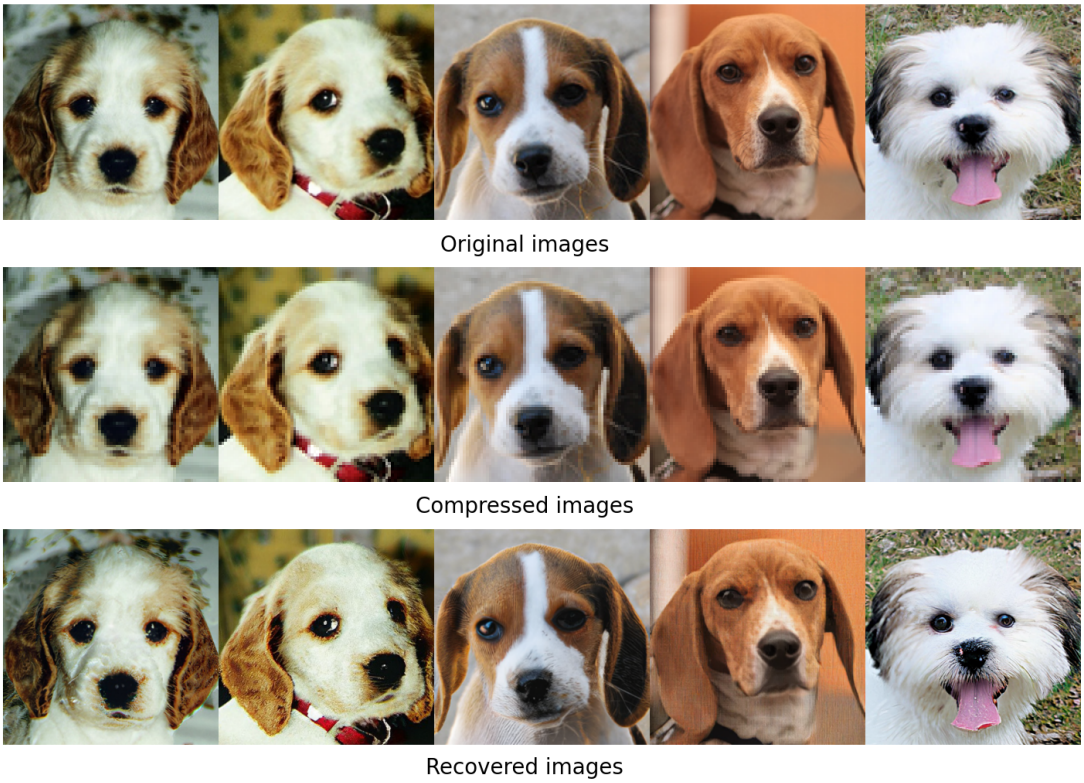


Figure 6.6: Visual results of semantic image transmission using the proposed framework.

The model used was pre-trained on the large-scale ImageNet (ILSVRC 2012) dataset. For this experiment, we used the pre-trained model directly for inference without any task-specific fine-tuning, demonstrating the framework's ability to leverage general-purpose generative models for semantic reconstruction.

The visual results of semantic image transmission under the VL-SemCom system are illustrated in Fig. 6.6. The top row presents the original high-resolution RGB images ($224 \times 224 \times 3$) used as reference. In the middle row, the semantically compressed images ($75 \times 75 \times 3$) are shown. These images are generated by the transmitter-side semantic encoder prior to VLC transmission and clearly exhibit noticeable degradation in fine-grained details due to aggressive dimensionality reduction. The bottom row displays the reconstructed images at the receiver, recovered from the compressed inputs using the diffusion-based generative model. Compared to the intermediate compressed images, the recovered images demonstrate substantial improvement in perceptual quality, including restoration of edge sharpness, texture continuity, and structure. In particular, the eyes, nose, and fur contours of the dog subjects are more accurately restructured by the diffusion model, closely resembling the original images. This confirms that the proposed semantic compression, when combined with diffusion-based decoding, is capable of preserving and reconstructing critical semantic content even under aggressive bit rate reduction.

6.5 Results and Discussions

In this section, we will comprehensively evaluate the performance of the proposed NOMA-SemCom VLC power allocation algorithm, which will be developed in two dimensions: hardware demonstration and numerical simulations. In the demonstration subsection, we show the test results of the built hardware demonstration platform to verify the feasibility of the proposed framework in real physical links. In the numerical simulation subsection, we present the parameter settings and then compare the performance of our proposed algorithm with two benchmark strategies to quantitatively analyze its superiority. For performance benchmarking, we compare our algorithm against the fixed power allocation (FPA) and GRPA strategies. The FPA strategy allocates power based on a predetermined ratio between users (i.e., $P_i = \alpha_p P_{i+1}$). In GRPA strategy, power is allocated according to channel conditions, i.e., $P_i = (\frac{h_1}{h_i})^{\beta}_{p} P_{i-1}$ [137]. In our simulations, we choose $\alpha_p = 0.4$ and $\beta_p = 1$.

6.5.1 Demo Results of the VL-SemCom System

This section presents an experimental validation of our proposed NOMA-SemCom power allocation algorithm on a hardware demonstration platform. The setup comprises two users, denoted as U_1 and U_2. To establish distinct channel conditions characteristic of NOMA systems, U_2 is positioned as the near user, located 0.5 meters closer to the transmitter than the far user, U_1.

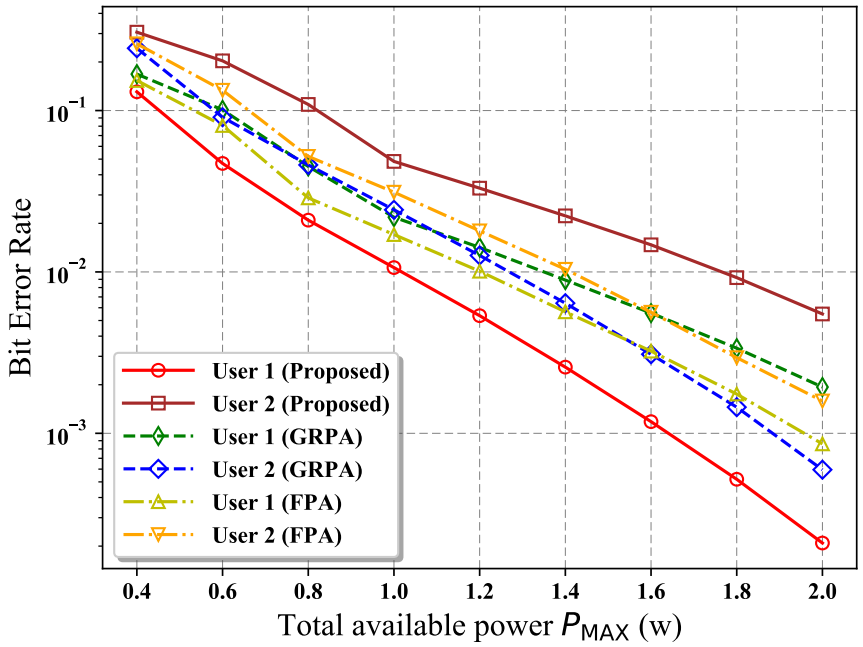


Figure 6.7: Bit Error Rate vs. Total available power P_{MAX}.

Critically, U_1 is set as a traditional communication user, while U_2 operates as a SemCom user. The simulation results analyze the system's performance under various total power levels.

Fig. 6.7 illustrates the performance curves of the bit-error-rate (BER) of the two users with respect to the total available power P_{MAX}, comparing the proposed power allocation algorithm with the FPA and GRPA strategies. As shown, the BER decreases monotonically with increasing P_{MAX}, since higher transmission power enhances the signal-to-noise ratio. Crucially, it is observed that for any given total power P_{MAX}, our proposed power allocation algorithm yields a higher BER for U_2, while achieving a lower BER for U_1. This is an intended result of the semantic-aware trade-off. The goal of traditional communication is to maximize bit-level accuracy, thus, its resource allocation is oriented towards BER reduction. In contrast, our algorithm reallocates power from U_2 to U_1 because U_2 possesses stronger semantic recovery capability than U_1. This action increases U_2's BER while decreasing U_1's BER, as seen in the plot. As a result, our semantic framework makes this bit-level sacrifice to maximize the overall semantic throughput.

In contrast to Fig. 6.7, Fig. 6.8 presents the corresponding sut error rate (SER) performance under the same system parameters and strategies. As shown, the SER for both strategies decreases as P_{MAX} increases, indicating that higher transmission power also enhances the reliability of semantic information conveyance. In detail, these strategies are very different in terms of user performance. Under the FPA and GRPA strategies, a clear performance gap exists where U_1 consistently experiences a higher SER than U_2. In contrast, our proposed algorithm, while making U_2 much larger than U_1 in terms of BER, brings the two users close to the same level

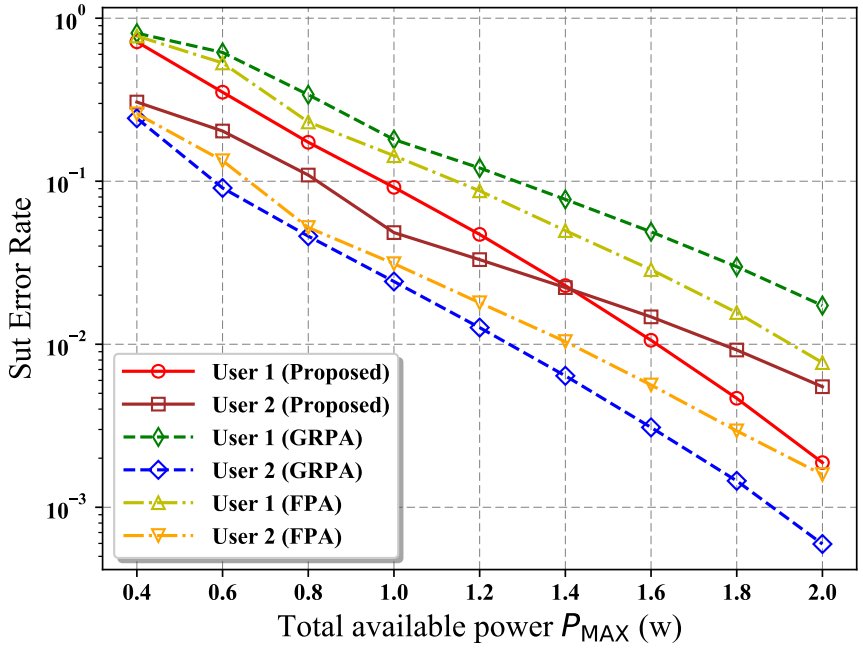


Figure 6.8: Sut Error Rate vs. Total available power P_{MAX}.

in terms of SER. Since U_2 has strong semantic recovery and can tolerate lower signal fidelity, our proposed algorithm reallocates the energy-saving power from U_2 to U_1. This ensures that the overall semantic integrity of the system is maximized, rather than being constrained by the bit-level requirements of individual users. This shows that, unlike conventional strategies, our algorithm is able to evaluate the semantic level, thus providing a fairer power allocation for the system.

Fig. 6.9 shows the comparison of the effect of receiving and recovering pictures from different users when different power allocation strategies are used. In this case, our proposed power allocation algorithm is used for the upper pictures (a), (b), (c) and the FPA strategy is used for the lower pictures (d), (e), (f). In our proposed algorithm, the algorithm evaluates that trans-

Table 6.1: Comparison of Image Quality Metrics

	U_2 received	U_2 recovered	U_1 received
PSNR			
Proposed	21.37 dB	24.66 dB	31.68 dB
FPA	24.42 dB	26.54 dB	28.01 dB
SSIM			
Proposed	0.3658	0.6255	0.8528
FPA	0.5203	0.6871	0.6955
LPIPS			
Proposed	0.4460	0.1432	0.0588
FPA	0.2834	0.1115	0.1753

(a) U_2 received (proposed) (b) U_2 recovered (proposed) (c) U_1 received (proposed)

(d) U_2 received (FPA) (e) U_2 recovered (FPA) (f) U_1 received (FPA)

Figure 6.9: Image transmission and recovery results for users under different allocation strategies.

mitting information for the semantic user U_2 yields a huge semantic gain, i.e., the core image meaning can be conveyed with very little power resources. Therefore, the algorithm reduces the power allocated to U_2, so (a) shows that U_2 received the initial signal with noise. The saved power resources are then reallocated to U_1, which does not benefit from SemCom. This enables U_1 to receive high-fidelity images as shown in (c), with a solid guarantee of quality of service. Meanwhile, (b) shows that U_2 is still able to reconstruct clear images from low-power signals due to its strong semantic recovery capability. In contrast, the FPA strategy has a single decision dimension, which is unable to sense and utilize any semantic gain, based only on the physical channel or using fixed allocation. This strategy leads to its inability to perform efficient resource allocation, resulting in insufficient guarantees for U_1 and wasted resources for U_2. In addition, we evaluated the images using three standard metrics: Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Learned Perceptual Image Patch Similarity (LPIPS). While PSNR and SSIM measure pixel-level fidelity, LPIPS is a deep feature-based metric that aligns closely with human perception, making it an excellent indicator of semantic fidelity. The results are presented in TABLE I. Therefore, our proposed algorithm is able to consider both physical channel resources and semantic gains to achieve rational power allocation.

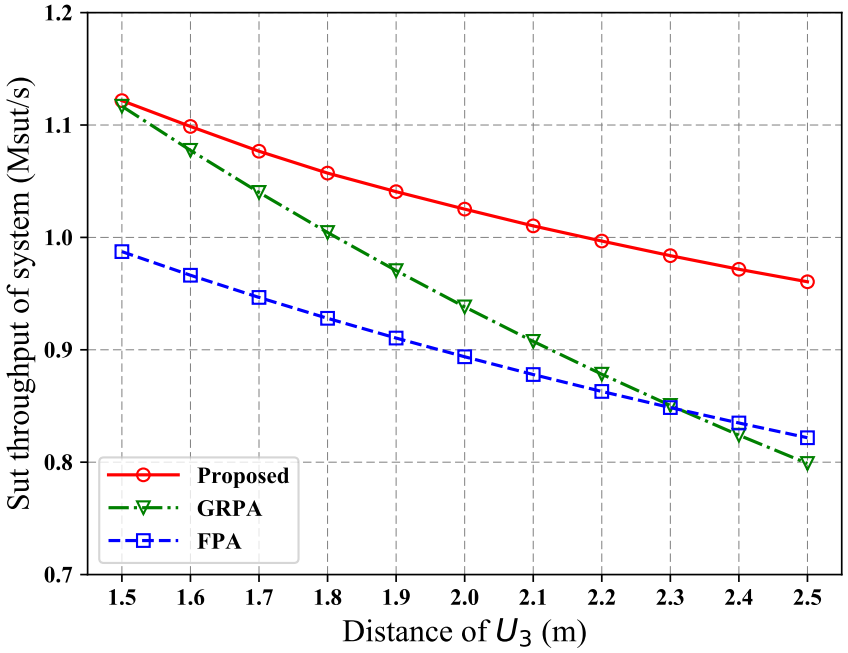


Figure 6.10: Sut throughput vs. Distance of U_3.

6.5.2 Performance Evaluation of Power Allocation Algorithm

We now evaluate the performance of our proposed algorithm through numerical simulations. Unless otherwise specified, the simulation parameters are set as follows. We consider a NOMA system consisting of three users, with a total bandwidth of 1 MHz. The total transmit power of a NOMA link is set to 1 W, and the noise power is -85 dBm. The main parameters of the simulation settings are listed in TABLE II.

Table 6.2: Simulation Parameters

Parameters	**Values**
Number of NOMA users	3
Bandwidth (B)	1 MHz
Transmit power (P_{MAX})	1 W
Noise power (P_n)	-85 dBm
LED half-intensity angle ($\Phi_{1/2}$)	60°
Physical area of the PD (A)	7.5mm^2

Next, we compare the sut throughput of the system under the different distance from U_3 to the LED. Note that while the distance between U_3 and the AP changes, the positions of other two users remain unchanged, and the distance between U_1 and the AP, and U_2 and the AP are set to 4m and 3m, respectively. As shown in Fig. 6.10, our proposed algorithm consistently outperforms GRPA and FPA for all distances. As the distance increases, although the performance of all strategies decreases, the performance gap between the proposed algorithm and the

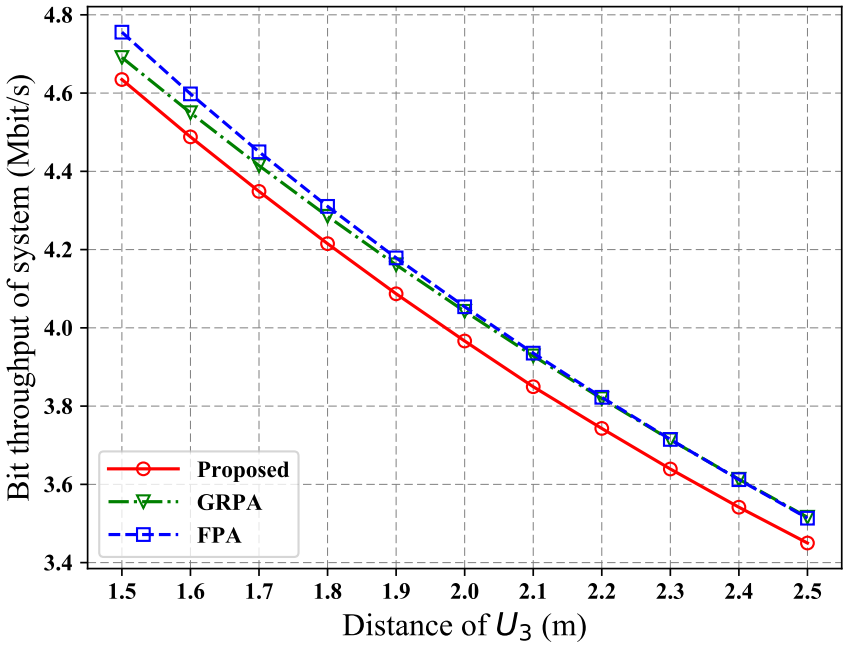


Figure 6.11: Bit throughput vs. Distance of U_3.

other strategies remains stable and even tends to widen, showing its advantages under different channel conditions. The proposed algorithm considers not only the physical channel but also the semantic gain when allocating power, resulting in its high semantic throughput. The FPA strategy exhibits the lowest sut throughput almost all distances. We note that the FPA and GRPA curves crossover at 2.3 m, which occurs because both FPA and GRPA are heuristic strategies. While GRPA is channel-aware, its allocation rule is sub-optimal. In the specific region, the FPA happens to yield a power distribution coincidentally more effective for semantic transmission than GRPA. This crossover between the two benchmarks demonstrates their unpredictable nature and highlights the need for our adaptive optimization-based algorithm.

Besides sut rate, Fig. 6.11 examines the system bit throughput under different distances from U_3. Contrast to the performance of sut rate, the FPA strategy shows the better performance, with the curve almost coinciding with the GRPA strategy. Our proposed algorithm shows slightly lower bit throughput than FPA and GRPA strategies. This is because our algorithm considers not only the physical channel conditions but also the semantic gain during power allocation. As a result, our algorithm trades some bit-level throughput for a gain in the overall semantic throughput.

We then analyze how the fairness of sut rate of the system varies with the total available power P_{MAX}. In this simulation, we use Jain's Fairness Index to measure the fairness of power allocation, which is calculated as $(\sum_{k=1}^{N} S_k)^2 / N \sum_{k=1}^{N} S_k^2$ [138]. From Fig. 6.12, our proposed algorithm presents the highest sut rate fairness over the entire power range, and its fairness index shows an upward trend as the total power increases. This shows our algorithms can fairly

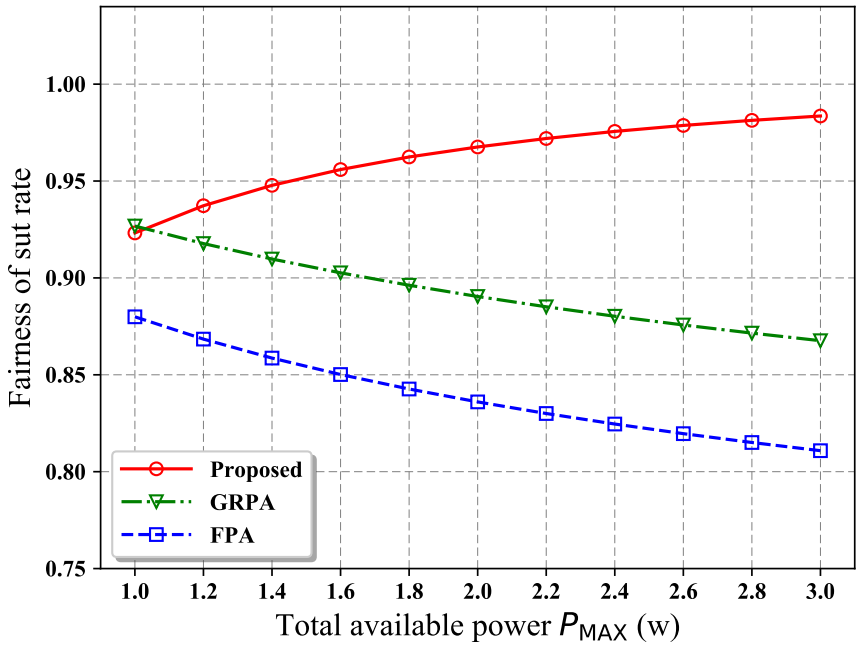


Figure 6.12: Fairness of sut rate vs. Total available power P_{MAX}.

benefit all users with the added power resources, so that the system performs more balanced when resources are abundant. In contrast, the fairness of both GRPA strategy and FPA strategy decreases with increasing power. This illustrates that these traditional algorithms tend to disproportionately allocate more resources to users with good channel conditions when the power budget becomes large, thus increasing the unfairness among users.

6.6 Conclusions

In this chapter, we have developed and analyzed a novel framework for a multi-user VL-SemCom system, leveraging NOMA to enhance spectral efficiency. Our objective was to maximize the system's semantic throughput while ensuring user fairness. To this end, we first developed a comprehensive mathematical model that uniquely integrates semantic transmission features with the physical layer characteristics of VLC. Building upon this model, we formulated a power allocation optimization problem aimed at maximizing the total system throughput. To solve this non-convex problem, we proposed a modified water-filling algorithm with a proof of its optimality. Validation via hardware demonstration and simulation confirmed our system significantly outperforms traditional benchmarks in semantic throughput and user fairness. Future work can focus on incorporating user mobility, exploring advanced semantic extraction techniques, and scaling the hardware prototype for larger deployments.

Chapter 7

Conclusions and Future Trends

This thesis has systematically investigated the complex domain of intelligent resource allocation for SemCom networks, focusing on the critical, yet often overlooked, challenge of KB maintenance. The research was motivated by a core insight: in dynamic environments, the ultra-high efficiency of SemCom is fundamentally tethered to the freshness of its shared KB. The process required to update this KB, termed knowledge consensus, creates a novel and fundamental resource conflict, as it competes directly with ongoing semantic data transmission for the same limited wireless resources. This central conflict, along with the practical imperatives of deploying SemCom in real-world systems, gives rise to a tripartite challenge:

- the need for a new optimization methodology to solve the resultant NP-hard 0-1 mixed problems;
- the necessity of a new reliability framework to ensure the consensus protocols themselves are reliable;
- the demand for new physical layer optimization metrics and hardware validation.

This thesis has constructed a unified framework that provides a cohesive set of solutions to these interconnected challenges.

7.1 Conclusions

First, we formally modeled and solved the core resource conflict between knowledge obsolescence and knowledge consensus cost. This work is the first, to our knowledge, to mathematically quantify knowledge staleness ($\mathscr{K}(t)$) as a dynamic state variable. We formulated the dynamic trade-off between immediate semantic transmission utility (which degrades with staleness) and the opportunity cost of performing consensus as a 0-1 MINLP. To solve this online, we designed a low-complexity MPC framework. The MPC's internal planner utilizes IMCA, which efficiently approximates the 0-1 MINLP subproblem by iteratively equalizing marginal

costs. Simulations confirmed that our MPC-IMCA controller significantly outperforms static baselines, such as “Greedy Transmission" and “Periodic Updates," in long-term cumulative semantic utility.

Second, this thesis developed a unified optimization-learning framework to solve the general 0-1 MINLP optimization problems that are endemic to wireless networking. We identified a key methodological gap where traditional optimization methods like B&B suffer from exponential complexity, while pure RL methods are inefficient due to blind exploration in vast combinatorial action spaces. Our proposed relaxation-guided RL framework bridges this gap. By first solving the continuous relaxation of the problem, we obtain a relaxed solution that is theoretically proven to define a HPZ containing high-quality integer solutions. This solution is then used as a potent prior to guide the RL agent's search. This synergy transforms the learning process from blind trial-and-error to an efficient, guided search. Simulation results validated this approach, demonstrating a 20% improvement in objective value over pure RL and a 30% reduction in convergence time compared to B&B in large-scale scenarios.

Third, we established a novel reliability framework for the consensus protocols used in KB maintenance, based on an availability-robustness trade-off. We argued that the traditional deterministic BFT/CFT models are overly conservative for the probabilistic failure modes of wireless networks. Furthermore, we identified that emerging probabilistic models, which rely on a single success probability, are insufficient as they fail to distinguish between two fundamentally different failure types: False Negatives (losing Availability) and False Positives (losing Robustness). Our dual-metric model explicitly quantifies this trade-off, enabling the optimal design of the quorum size by solving a constrained optimization problem. Moreover, we extended this framework to a reputation-weighted quorum mechanism, demonstrating through simulation that this extension can enhance system robustness by several orders of magnitude.

Finally, we designed, optimized, and demonstrated a multi-user SemCom system over a physical VLC-NOMA platform. This work addresses the critical physical layer bandwidth bottleneck of VLC by applying SemCom. We rigorously defined a new optimization objective, STU, based on semantic entropy and KB mutual information, moving beyond traditional bit rate metrics. We formulated the non-convex, semantic-aware power allocation problem and proposed a modified water-filling-based algorithm to solve it. The validation, performed on a custom-built Arduino-based hardware platform, confirmed the efficacy of our solution. The experimental results clearly showed our algorithm intelligently trades higher BER for a significantly lower SER, thereby maximizing overall system semantic integrity and demonstrating a clear performance gain over traditional approaches.

In synthesis, these four contributions provide a cohesive and comprehensive framework for intelligent resource management in semantic networks, bridging the gap from foundational optimization methodology and reliability theory to the core problem of knowledge maintenance and, ultimately, to physical layer implementation and validation.

7.2 Future Trends

The conclusions of this thesis open several promising avenues for future research, extending the foundations laid herein toward more complex, distributed, and secure systems.

7.2.1 Synergistic Integration of Optimization Frameworks

This thesis optimized communication and computation resources. A more general direction for 6G is the integration of Sensing. particularly in the VLC scenarios explored in Chapter 6. Future research can explore how semantic features extracted for communication can be simultaneously reused for sensing tasks (e.g., localization or motion detection), optimizing the trade-off between semantic fidelity and sensing accuracy in a unified resource framework.

7.2.2 Distributed Knowledge Maintenance and Multi-Agent Consensus

This thesis largely modeled KB maintenance from a centralized perspective and analyzed the reliability of a single consensus instance. However, future networks, such as V2X or UAV swarms, are inherently distributed. Future research should extend our framework to multi-agent scenarios. This introduces new challenges, such as: (1) Distributed Scheduling, moving from a centralized MPC to a decentralized MARL policy for managing the transmission-consensus trade-off; (2) KB Conflict Resolution, developing protocols for scenarios where multiple agents discover conflicting new knowledge simultaneously; and (3) Correlated Failures, an explicit future work. In V2X networks, geographically co-located vehicles may experience correlated channel fading or failures, violating the independence assumption in our current reliability model and requiring new, location-aware quorum design.

7.2.3 Refinement of Semantic Metrics and Staleness Models

Our work provided foundational, quantifiable models for STU and knowledge staleness ($\mathscr{K}(t)$), both of which can be significantly refined. The STU model assumes all semantic units (sut) are of equal importance. A crucial extension is to incorporate task-importance or contextual value, where semantic units like pedestrian or warning signal are assigned a much higher value than “tree" or “cloud." This would lead to a value-weighted STU objective, enabling more critical resource allocation. Similarly, the staleness model uses a deterministic drift, could be advanced to a stochastic process where staleness evolves based on the rate of novel environmental events, offering a more realistic depiction of dynamic worlds.

While this thesis focused on synchronizing specific, task-oriented KBs, the recent surge in Large Language Models and Generative AI offers a more general paradigm. Future work will investigate Generative Semantic Communication, where the KB is replaced by a pre-trained Foundation Model. The challenge shifts from KB synchronization to prompt engineering and

parameter efficient fine-tuning to align the generative models of the sender and receiver with minimal communication overhead.

Current work relies on empirical utility functions (e.g., semantic similarity). A more general theoretical challenge is to establish the fundamental limits of semantic compression. Future work aims to leverage Graph Theory or Category Theory to mathematically model the structure of knowledge, providing a rigorous information-theoretic bound for the Knowledge-Driven SemCom networks proposed in this thesis.

7.2.4 Hardware Prototyping and Semantic Security

This thesis laid the groundwork for both physical layer implementation and reliability analysis. The logical progression is to merge these threads, focusing on robust hardware and security. The VL-SemCom prototype successfully validated the proposed algorithms on an Arduino platform. Future work should scale this prototype to high-performance hardware (e.g., FPGAs, GPUs, or NPUs) capable of executing the complex DNNs of DeepJSCC in real-time. This would create a true end-to-end, high-data-rate VL-SemCom system. This advancement, however, highlights the critical security vulnerability of SemCom: KB Poisoning. An attacker could attempt to inject a malicious or corrupted KB update. Future work must apply the robustness metric (η_R) to the KB consensus protocol itself, developing secure mechanisms, such as the proposed reputation-weighted quorum, to defend the KB of the entire semantic network.

Bibliography

[1] Wanting Yang, Hongyang Du, Zi Qin Liew, Wei Yang Bryan Lim, Zehui Xiong, Dusit Niyato, Xuefen Chi, Xuemin Shen, and Chunyan Miao. Semantic communications for future internet: Fundamentals, applications, and challenges. *IEEE Communications Surveys & Tutorials*, 25(1):213–250, 2022.

[2] Peiwen Jiang, Chao-Kai Wen, Shi Jin, and Geoffrey Ye Li. Wireless semantic communications for video conferencing. *IEEE Journal on Selected Areas in Communications*, 41(1):230–244, 2022.

[3] Sridhar Iyer, Rajashri Khanai, Dattaprasad Torse, Rahul Jashvantbhai Pandya, Khaled M Rabie, Krishna Pai, Wali Ullah Khan, and Zubair Fadlullah. A survey on semantic communications for intelligent wireless networks. *Wireless Personal Communications*, 129(1):569–611, 2023.

[4] Huiqiang Xie, Zhijin Qin, and Geoffrey Ye Li. Task-oriented multi-user semantic communications for VQA. *IEEE Wireless Communications Letters*, 11(3):553–557, 2021.

[5] Xuewen Luo, Hsiao-Hwa Chen, and Qing Guo. Semantic Communications: Overview, Open Issues, and Future Research Directions. *IEEE Wireless Communications*, 29(1):210–219, 2022.

[6] Shuheng Hua, Yao Sun, Kairong Ma, Swash Rafiq, Wallizada Mohibullah, Zhaohui Yang, and Muhammad Imran. Optimizing Spectral Efficiency through Bandwidth Management in Semantic Communication Systems. In *2024 IEEE International Conference on Communications Workshops (ICC)*, pages 1635–1640. IEEE, 2024.

[7] Yao Sun, Lan Zhang, Linke Guo, Jian Li, Dusit Niyato, and Yuguang Fang. S-RAN: Semantic-Aware Radio Access Networks. *arXiv preprint arXiv:2407.11161*, 2024.

[8] Shuheng Hua, Yao Sun, Kairong Ma, Lei Feng, Mingkai Chen, Zhaohui Yang, and Muhammad Ali Imran. Bandwidth Management in Semantic Communications: A Tradeoff Between Data Sensing and Transmission. *IEEE Transactions on Vehicular Technology*, pages 1–13, 2025.

[9] Mohanad Obeed, Anas M. Salhab, Mohamed-Slim Alouini, and Salam A. Zummo. On Optimizing VLC Networks for Downlink Multi-User Transmission: A Survey. *IEEE Communications Surveys & Tutorials*, 21(3):2947–2976, 2019.

[10] Huiqiang Xie, Zhijin Qin, Geoffrey Ye Li, and Biing-Hwang Juang. Deep learning enabled semantic communication systems. *IEEE transactions on signal processing*, 69:2663–2675, 2021.

[11] VS Mikhalevich, AM Gupal, and VI Norkin. Methods of nonconvex optimization. *arXiv preprint arXiv:2406.10406*, 2024.

[12] Xiaojun Lin, Ness B Shroff, and Rayadurgam Srikant. A tutorial on cross-layer optimization in wireless networks. *IEEE Journal on Selected areas in Communications*, 24(8):1452–1463, 2006.

[13] J. Glover, V. Quan, and S. Zolfaghari. Some new perspectives for solving 0–1 integer programming problems using balas method. *Computational Management Science*, 18(2):177–193, June 2021.

[14] Aizaz U Chaudhry, John W Chinneck, and Roshdy HM Hafez. Fast heuristics for the frequency channel assignment problem in multi-hop wireless networks. *European Journal of Operational Research*, 251(3):771–782, 2016.

[15] Josip Lorincz, Antonio Capone, and Dinko Begusic. Heuristic algorithms for optimization of energy consumption in wireless access networks. *KSII Transactions on Internet and Information Systems (TIIS)*, 5(4):626–648, 2011.

[16] Josip Lorincz, Antonio Capone, and Dinko Begusic. Heuristic algorithms for optimization of energy consumption in wireless access networks. *KSII Transactions on Internet and Information Systems (TIIS)*, 5(4):626–648, 2011.

[17] Kun Lu, Rongpeng Li, Xianfu Chen, Zhifeng Zhao, and Honggang Zhang. Reinforcement Learning-powered Semantic Communication via Semantic Similarity. *arXiv preprint arXiv:2108.12121*, 2021.

[18] Yawen Chen, Yu Liu, Ming Zeng, Umber Saleem, Zhaoming Lu, Xiangming Wen, Depeng Jin, Zhu Han, Tao Jiang, and Yong Li. Reinforcement learning meets wireless networks: A layering perspective. *IEEE Internet of Things Journal*, 8(1):85–111, 2021.

[19] Ying Wang, Xiangming Dai, Jason Min Wang, and Brahim Bensaou. A reinforcement learning approach to energy efficiency and qos in 5g wireless networks. *IEEE Journal on Selected Areas in Communications*, 37(6):1413–1423, 2019.

[20] Yixuan Fan, Huanyu Wu, Zhongxu Dong, Zongyao Li, and Lei Zhang. From human to machine networks: A framework for integrating consensus approaches. submitted.

[21] Yang Xiao, Ning Zhang, Jin Li, Wenjing Lou, and Y Thomas Hou. Distributed consensus protocols and algorithms. *Blockchain for Distributed Systems Security*, 25:40, 2019.

[22] Dongpu Cao, Xiao Wang, Lingxi Li, Chen Lv, Xiaoxiang Na, Yang Xing, Xuan Li, Ying Li, Yuanyuan Chen, and Fei-Yue Wang. Future directions of intelligent vehicles: Potentials, possibilities, and perspectives. *IEEE Transactions on Intelligent Vehicles*, 7(1):7–10, 2022.

[23] Bangjiang Lin, Qiwei Lai, Zabih Ghassemlooy, and Xuan Tang. A Machine Learning Based Signal Demodulator in NOMA-VLC. *Journal of Lightwave Technology*, 39(10):3081–3087, 2021.

[24] Eirina Bourtsoulatze, David Burth Kurka, and Deniz Gündüz. Deep joint source-channel coding for wireless image transmission. *IEEE Transactions on Cognitive Communications and Networking*, 5(3):567–579, 2019.

[25] Harsh Tataria, Mansoor Shafi, Andreas F. Molisch, Mischa Dohler, Henrik Sjöland, and Fredrik Tufvesson. 6g wireless systems: Vision, requirements, challenges, insights, and opportunities. *Proceedings of the IEEE*, 109(7):1166–1199, 2021.

[26] Warren Weaver. Recent contributions to the mathematical theory of communication. *ETC: a review of general semantics*, pages 261–281, 1953.

[27] Xiang Peng, Zhijin Qin, Danlan Huang, Xiaoming Tao, Jianhua Lu, Guangyi Liu, and Chengkang Pan. A robust deep learning enabled semantic communication system for text. In *GLOBECOM 2022-2022 IEEE Global Communications Conference*, pages 2704–2709. IEEE, 2022.

[28] Zhenzi Weng and Zhijin Qin. Semantic communication systems for speech transmission. *IEEE Journal on Selected Areas in Communications*, 39(8):2434–2444, 2021.

[29] Peng Yi, Yang Cao, Xin Kang, and Ying-Chang Liang. Deep learning-empowered semantic communication systems with a shared knowledge base. *IEEE Transactions on Wireless Communications*, 23(6):6174–6187, 2023.

[30] Chen Peng and Bolin Liao. Heavy-head sampling for fast imitation learning of machine learning based combinatorial auction solver. *Neural Processing Letters*, 55(1):631–644, 2023.

[31] Josip Lorincz, Massimo Bogarelli, Antonio Capone, and Dinko Begušić. Heuristic approach for optimized energy savings in wireless access networks. In *SoftCOM 2010, 18th*

International Conference on Software, Telecommunications and Computer Networks, pages 60–65. IEEE, 2010.

[32] Omid Bozorg-Haddad, Mohammad Solgi, and Hugo A Loáiciga. *Meta-heuristic and evolutionary algorithms for engineering optimization*. John Wiley & Sons, 2017.

[33] Ning Yang, Shuo Chen, Haijun Zhang, and Randall Berry. Beyond the edge: An advanced exploration of reinforcement learning for mobile edge computing, its applications, and future research trajectories. *IEEE Communications Surveys & Tutorials*, 27(1):546–594, 2024.

[34] Jinke Ren, Zezhong Zhang, Jie Xu, Guanying Chen, Yaping Sun, Ping Zhang, and Shuguang Cui. Knowledge base enabled semantic communication: A generative perspective. *IEEE Wireless Communications*, 31(4):14–22, 2024.

[35] Shengyun Liu, Paolo Viotti, Christian Cachin, Vivien Quéma, and Marko Vukolić. {XFT}: Practical fault tolerance beyond crashes. In *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, pages 485–500, 2016.

[36] Leslie Lamport, Robert Shostak, and Marshall Pease. The byzantine generals problem. In *Concurrency: the works of leslie lamport*, pages 203–226. 2019.

[37] Miguel Castro, Barbara Liskov, et al. Practical byzantine fault tolerance. In *OsDI*, volume 99, pages 173–186, 1999.

[38] Benjamin Bisping, Paul-David Brodmann, Tim Jungnickel, Christina Rickmann, Henning Seidler, Anke Stüber, Arno Wilhelm-Weidner, Kirstin Peters, and Uwe Nestmann. A constructive proof for flp. *Archive of Formal Proofs*, 20, 2016.

[39] Dachao Yu, Wenyu Li, Hao Xu, and Lei Zhang. Low reliable and low latency communications for mission critical distributed industrial internet of things. *IEEE Communications Letters*, 25(1):313–317, 2021.

[40] Hao Xu, Yixuan Fan, Wenyu Li, and Lei Zhang. Wireless distributed consensus for connected autonomous systems. *IEEE Internet of Things Journal*, 10(9):7786–7799, 2023.

[41] Mohanad Obeed, Hayssam Dahrouj, Anas M Salhab, Salam A Zummo, and Mohamed-Slim Alouini. User pairing, link selection, and power allocation for cooperative noma hybrid vlc/rf systems. *IEEE Transactions on Wireless Communications*, 20(3):1785–1800, 2020.

[42] Aymen Khaleel and Ertugrul Basar. A novel noma solution with ris partitioning. *IEEE Journal of Selected Topics in Signal Processing*, 16(1):70–81, 2021.

[43] Yang Liu, Gaofeng Pan, Hongtao Zhang, and Mei Song. On the capacity comparison between mimo-noma and mimo-oma. *IEEE Access*, 4:2123–2129, 2016.

[44] Huiqiang Xie, Zhijin Qin, Xiaoming Tao, and Khaled B Letaief. Task-oriented multi-user semantic communications. *IEEE Journal on Selected Areas in Communications*, 40(9):2584–2597, 2022.

[45] Yulin Shao, Qi Cao, and Deniz Gündüz. A theory of semantic communication. *IEEE Transactions on Mobile Computing*, 23(12):12211–12228, 2024.

[46] Jie Bao, Prithwish Basu, Mike Dean, Craig Partridge, Ananthram Swami, Will Leland, and James A Hendler. Towards a theory of semantic communication. In *2011 IEEE Network Science Workshop*, pages 110–117. IEEE, 2011.

[47] Shengshi Yao, Kai Niu, Sixian Wang, and Jincheng Dai. Semantic coding for text transmission: An iterative design. *IEEE Transactions on Cognitive Communications and Networking*, 8(4):1594–1603, 2022.

[48] Jialong Xu, Tze-Yang Tung, Bo Ai, Wei Chen, Yuxuan Sun, and Deniz Gündüz. Deep joint source-channel coding for semantic communications. *IEEE Communications Magazine*, 61(11):42–48, 2023.

[49] Jianhao Huang, Kai Yuan, Chuan Huang, and Kaibin Huang. D 2-jscc: Digital deep joint source-channel coding for semantic communications. *IEEE Journal on Selected Areas in Communications*, 2025.

[50] Huiqiang Xie, Zhijin Qin, Xiaoming Tao, and Zhu Han. Toward intelligent communications: Large model empowered semantic communications. *IEEE Communications Magazine*, 2024.

[51] Lei Yan, Zhijin Qin, Rui Zhang, Yongzhao Li, and Geoffrey Ye Li. Resource allocation for text semantic communications. *IEEE Wireless Communications Letters*, 11(7):1394–1398, 2022.

[52] Kairong Ma, Hanaa Abumarshoud, Shuheng Hua, Muhammad Imran, and Yao Sun. Power allocation for throughput maximization in noma-based semantic communication system. In *ICC 2025-IEEE International Conference on Communications*, pages 4288–4293. IEEE, 2025.

[53] Wenjing Zhang, Yining Wang, Mingzhe Chen, Tao Luo, and Dusit Niyato. Optimization of image transmission in cooperative semantic communication networks. *IEEE Transactions on Wireless Communications*, 23(2):861–873, 2023.

[54] Tilahun M Getu, Walid Saad, Georges Kaddoum, and Mehdi Bennis. Performance limits of a deep learning-enabled text semantic communication under interference. *IEEE Transactions on Wireless Communications*, 23(8):10213–10228, 2024.

[55] Heidi Howard, Aleksey Charapko, and Richard Mortier. Fast flexible paxos: Relaxing quorum intersection for fast paxos. In *Proceedings of the 22nd International Conference on Distributed Computing and Networking*, pages 186–190, 2021.

[56] Yuetai Li, Yixuan Fan, Lei Zhang, and Jon Crowcroft. Raft consensus reliability in wireless networks: Probabilistic analysis. *IEEE Internet of Things Journal*, 10(14):12839–12853, 2023.

[57] Yuntao Wang, Haixia Peng, Zhou Su, Tom H Luan, Abderrahim Benslimane, and Yuan Wu. A platform-free proof of federated learning consensus mechanism for sustainable blockchains. *IEEE Journal on Selected Areas in Communications*, 40(12):3305–3324, 2022.

[58] Kotteswaran Venkatesan and Syarifah Bahiyah Rahayu. Blockchain security enhancement: an approach towards hybrid consensus algorithms and machine learning techniques. *Scientific Reports*, 14(1):1149, 2024.

[59] Jose Alfredo Alvarez Aldana, Stephane Maag, and Fatiha Zaidi. A formal consensus-based distributed monitoring approach for mobile iot networks. *Internet of Things*, 13:100352, 2021.

[60] Max Schwenzer, Muzaffer Ay, Thomas Bergs, and Dirk Abel. Review on model predictive control: An engineering perspective. *The International Journal of Advanced Manufacturing Technology*, 117(5):1327–1349, 2021.

[61] Hoa Minh Nguyen, Jose Luis Rueda Torres, Aleksandra Lekić, and Hoan V Pham. Mpc based centralized voltage and reactive power control for active distribution networks. *IEEE Transactions on Energy Conversion*, 36(2):1537–1547, 2021.

[62] Alessandro Bozzi, Simone Graffione, Roberto Sacile, and Enrico Zero. Dynamic mpc-based scheduling in a smart manufacturing system problem. *IEEE Access*, 11:141987–141996, 2023.

[63] Sara A. Elsayed, Sherin Abdelhamid, and Hossam S. Hassanein. Predictive proactive caching in vanets for social networking. *IEEE Transactions on Vehicular Technology*, 71(5):5298–5313, 2022.

[64] Roy D Yates, Yin Sun, D Richard Brown, Sanjit K Kaul, Eytan Modiano, and Sennur Ulukus. Age of information: An introduction and survey. *IEEE Journal on Selected Areas in Communications*, 39(5):1183–1210, 2021.

[65] Zhi-Quan Luo and Wei Yu. An introduction to convex optimization for communications and signal processing. *IEEE Journal on selected areas in communications*, 24(8):1426–1438, 2006.

[66] Arindam Kumar Das, Robert J Marks, Mohamed El-Sharkawi, Payman Arabshahi, and Andrew Gray. Minimum power broadcast trees for wireless networks: integer programming formulations. In *Proc. of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE INFOCOM)*, volume 2, pages 1001–1010, 2003.

[67] Jason L Williams, John W Fisher, and Alan S Willsky. Approximate dynamic programming for communication-constrained sensor network management. *IEEE Transactions on Signal Processing*, 55(8):4300–4311, 2007.

[68] Michel Gendreau, Jean-Yves Potvin, et al. *Handbook of metaheuristics*, volume 2. Springer, 2010.

[69] Feiran You and Hongyang Du. Reacritic: Large reasoning transformer-based drl critic-model scaling for heterogeneous networks. *arXiv preprint arXiv:2505.10992*, 2025.

[70] Feiran You, Hongyang Du, Xiangwang Hou, Yong Ren, and Kaibin Huang. Dress: Diffusion reasoning-based reward shaping scheme for intelligent networks. *arXiv preprint arXiv:2503.07433*, 2025.

[71] Xiangyun Tang, Luyao Peng, Yu Weng, Meng Shen, Liehuang Zhu, and Robert H. Deng. Enforcing differential privacy in federated learning via long-term contribution incentives. *IEEE Transactions on Information Forensics and Security*, 20:3102–3115, 2025.

[72] Xiangyun Tang, Meng Shen, Qi Li, Liehuang Zhu, Tengfei Xue, and Qiang Qu. Pile: Robust privacy-preserving federated learning via verifiable perturbations. *IEEE Transactions on Dependable and Secure Computing*, 20(6):5005–5023, 2023.

[73] Lucy Ellen Lwakatare, Aiswarya Raj, Ivica Crnkovic, Jan Bosch, and Helena Holmström Olsson. Large-scale machine learning systems in real-world industrial settings: A review of challenges and solutions. *Information and software technology*, 127:106368, 2020.

[74] Yoshua Bengio, Andrea Lodi, and Antoine Prouvost. Machine learning for combinatorial optimization: a methodological tour d'horizon. *European Journal of Operational Research*, 290(2):405–421, 2021.

[75] Suzhi Bi and Ying Jun Zhang. Computation rate maximization for wireless powered mobile-edge computing with binary computation offloading. *IEEE Transactions on Wireless Communications*, 17(6):4177–4190, 2018.

[76] Thinh Quang Dinh, Jianhua Tang, Quang Duy La, and Tony QS Quek. Offloading in mobile edge computing: Task allocation and computational frequency scaling. *IEEE Transactions on Communications*, 65(8):3571–3584, 2017.

[77] Yan Sun, Derrick Wing Kwan Ng, Zhiguo Ding, and Robert Schober. Optimal joint power and subcarrier allocation for full-duplex multicarrier non-orthogonal multiple access systems. *IEEE Transactions on Communications*, 65(3):1077–1091, 2017.

[78] Ahmad Ali Khan, Raviraj S. Adve, and Wei Yu. Optimizing downlink resource allocation in multiuser mimo networks via fractional programming and the hungarian algorithm. *IEEE Transactions on Wireless Communications*, 19(8):5162–5175, 2020.

[79] Mohammad Karimzadeh Farshbafan, Walid Saad, and Merouane Debbah. Curriculum learning for goal-oriented semantic communications with a common language. *IEEE Transactions on Communications*, 71(3):1430–1446, 2023.

[80] Mingzhe Chen, Omid Semiari, Walid Saad, Xuanlin Liu, and Changchuan Yin. Federated echo state learning for minimizing breaks in presence in wireless virtual reality networks. *IEEE Transactions on Wireless Communications*, 19(1):177–191, 2019.

[81] Aidin Ferdowsi, Mohamed A Abd-Elmagid, Walid Saad, and Harpreet S Dhillon. Neural combinatorial deep reinforcement learning for age-optimal joint trajectory and scheduling design in uav-assisted networks. *IEEE Journal on Selected Areas in Communications*, 39(5):1250–1265, 2021.

[82] Xian Li, Liang Huang, Hui Wang, Suzhi Bi, and Ying-Jun Angela Zhang. An integrated optimization-learning framework for online combinatorial computation offloading in mec networks. *IEEE Wireless Communications*, 29(1):170–177, 2022.

[83] Hoon Lee, Sang Hyun Lee, and Tony Q. S. Quek. Deep learning for distributed optimization: Applications to wireless resource management. *IEEE Journal on Selected Areas in Communications*, 37(10):2251–2266, 2019.

[84] Yifei Shen, Yuanming Shi, Jun Zhang, and Khaled B Letaief. Lorm: Learning to optimize for resource management in wireless networks with few training samples. *IEEE Transactions on Wireless Communications*, 19(1):665–679, 2019.

[85] Usama Mehboob, Junaid Qadir, Salman Ali, and Athanasios Vasilakos. Genetic algorithms in wireless networking: techniques, applications, and issues. *Soft Computing*, 20:2467–2501, 2016.

[86] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*, chapter 2, pages 46–47. Oxford University Press, 02 2013.

[87] Stephen Boyd. *Convex optimization*, chapter 10, pages 553–555. Cambridge UP, 2004.

[88] Rafael Correa, Abderrahim Hantoute, and Marco A López. *Fundamentals of convex analysis and optimization*, chapter 3, pages 63–69. Springer, 2023.

[89] Thomas Lachand-Robert and Édouard Oudet. Minimizing within convex bodies using a convex hull method. *SIAM Journal on Optimization*, 16(2):368–379, 2005.

[90] Natashia Boland, Jeffrey Christiansen, Brian Dandurand, Andrew Eberhard, and Fabricio Oliveira. A parallelizable augmented lagrangian method applied to large-scale nonconvex-constrained optimization problems. *Mathematical Programming*, 175:503–536, 2019.

[91] Christian Bessiere. Arc-consistency and arc-consistency again. *Artificial intelligence*, 65(1):179–190, 1994.

[92] Le Xia, Yao Sun, Dusit Niyato, Daquan Feng, Lei Feng, and Muhammad Ali Imran. xURLLC-Aware Service Provisioning in Vehicular Networks: A Semantic Communication Perspective. *IEEE Transactions on Wireless Communications*, 23(5):4475–4488, 2024.

[93] Shancang Li, George Oikonomou, Theo Tryfonas, Thomas M Chen, and Li Da Xu. A distributed consensus algorithm for decision making in service-oriented internet of things. *IEEE Transactions on Industrial Informatics*, 10(2):1461–1468, 2014.

[94] Yang Xiao, Ning Zhang, Wenjing Lou, and Y Thomas Hou. A survey of distributed consensus protocols for blockchain networks. *IEEE communications surveys & tutorials*, 22(2):1432–1465, 2020.

[95] Leslie Lamport. The part-time parliament. In *Concurrency: the works of Leslie Lamport*, pages 277–317. 2019.

[96] Leslie Lamport. Proving the correctness of multiprocess programs. *IEEE transactions on software engineering*, (2):125–143, 1977.

[97] Xinru Han, Jianming Zhan, Zeshui Xu, and Luis Martínez. Trust risk test-based group consensus with probabilistic linguistic preference relations under social networks. *IEEE Transactions on Fuzzy Systems*, 32(6):3508–3520, 2024.

[98] Mario Gerla and Leonard Kleinrock. Vehicular networks and the future of the mobile internet. *Computer Networks*, 55(2):457–469, 2011.

[99] Priti Kumari and Parmeet Kaur. A survey of fault tolerance in cloud computing. *Journal of King Saud University-Computer and Information Sciences*, 33(10):1159–1176, 2021.

[100] Lei Zhang, Hao Xu, Oluwakayode Onireti, Muhammad Ali Imran, and Bin Cao. How much communication resource is needed to run a wireless blockchain network? *IEEE network*, 36(1):128–135, 2021.

[101] Haoxiang Luo, Xiangyue Yang, Hongfang Yu, Gang Sun, Bo Lei, and Mohsen Guizani. Performance analysis and comparison of nonideal wireless pbft and raft consensus networks in 6g communications. *IEEE Internet of Things Journal*, 11(6):9752–9765, 2024.

[102] Christian Cachin and Marko Vukolić. Blockchain consensus protocols in the wild. *arXiv preprint arXiv:1707.01873*, 2017.

[103] Michael Ben-Or. Another advantage of free choice (extended abstract) completely asynchronous agreement protocols. In *Proceedings of the second annual ACM symposium on Principles of distributed computing*, pages 27–30, 1983.

[104] Michael J Fischer, Nancy A Lynch, and Michael S Paterson. Impossibility of distributed consensus with one faulty process. *Journal of the ACM (JACM)*, 32(2):374–382, 1985.

[105] Michael Ben-Or. Another advantage of free choice (extended abstract) completely asynchronous agreement protocols. In *Proceedings of the second annual ACM symposium on Principles of distributed computing*, pages 27–30, 1983.

[106] Paul Feldman and Silvio Micali. Optimal algorithms for byzantine agreement. In *Proceedings of the twentieth annual ACM symposium on Theory of computing*, pages 148–161, 1988.

[107] Hao Xu, Yixuan Fan, Wenyu Li, and Lei Zhang. Wireless distributed consensus for connected autonomous systems. *IEEE Internet of Things Journal*, 10(9):7786–7799, 2023.

[108] Martin Nischwitz, Marko Esche, and Florian Tschorsch. Bernoulli meets pbft: Modeling bft protocols in the presence of dynamic failures. In *2021 16th Conference on Computer Science and Intelligence Systems (FedCSIS)*, pages 291–300, 2021.

[109] Yixuan Fan, Lei Zhang, Ruiyu Wang, and Muhammad Ali Imran. Insight into voting in daos: Conceptual analysis and a proposal for evaluation framework. *IEEE Network*, 38(3):92–99, 2024.

[110] Moni Naor and Avishai Wool. The load, capacity, and availability of quorum systems. *SIAM Journal on Computing*, 27(2):423–447, 1998.

[111] Xin Yao and Cho-Li Wang. Probabilistic consistency guarantee in partial quorum-based data store. *IEEE Transactions on Parallel and Distributed Systems*, 31(8):1815–1827, 2020.

[112] Byeng D Youn and Kyung K Choi. Selecting probabilistic approaches for reliability-based design optimization. *AIAA journal*, 42(1):124–131, 2004.

[113] Ittai Abraham, Danny Dolev, Rica Gonen, and Joe Halpern. Distributed computing meets game theory: robust mechanisms for rational secret sharing and multiparty computation. In *Proceedings of the twenty-fifth annual ACM symposium on Principles of distributed computing*, pages 53–62, 2006.

[114] Dahlia Malkhi and Michael Reiter. Byzantine quorum systems. *Distributed computing*, 11(4):203–213, 1998.

[115] Seba Adnan Yosef and Lwaa Faisal Abdulameer. A Systematic Review on Non-Orthogonal Multiple Access (NOMA) Based on Visible Light Communication for Intelligent Transportation Systems. *Tikrit Journal of Engineering Sciences*, 31(1):278–290, 2024.

[116] Harald Haas, Elham Sarbazi, Hanaa Marshoud, and John Fakidis. *Optical Fiber Telecommunications VII*, chapter 11, pages 443–493. Academic Press, 2020.

[117] Qi Sun, Shuangfeng Han, Chin-Lin I, and Zhengang Pan. On the Ergodic Capacity of MIMO NOMA Systems. *IEEE Wireless Communications Letters*, 4(4):405–408, 2015.

[118] Hanaa Marshoud, Paschalis C. Sofotasios, Sami Muhaidat, George K. Karagiannidis, and Bayan S. Sharif. On the Performance of Visible Light Communication Systems With Non-Orthogonal Multiple Access. *IEEE Transactions on Wireless Communications*, 16(10):6350–6364, 2017.

[119] Xidong Mu and Yuanwei Liu. Exploiting Semantic Communication for Non-Orthogonal Multiple Access. *IEEE Journal on Selected Areas in Communications*, 41(8):2563–2576, 2023.

[120] Hanaa Marshoud, Vasileios M. Kapinas, George K. Karagiannidis, and Sami Muhaidat. Non-Orthogonal Multiple Access for Visible Light Communications. *IEEE Photonics Technology Letters*, 28(1):51–54, 2016.

[121] Hongwei Zhang, Shuo Shao, Meixia Tao, Xiaoyan Bi, and Khaled B Letaief. Deep Learning-enabled Semantic Communication Systems with Task-unaware Transmitter and Dynamic Data. *IEEE Journal on Selected Areas in Communications*, 41(1):170–185, 2022.

[122] Ke Yang, Sixian Wang, Jincheng Dai, Kailin Tan, Kai Niu, and Ping Zhang. WITT: A Wireless Image Transmission Transformer for Semantic Communications. In *2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5, 2023.

[123] Shugang Liu, Zhan Peng, Qiangguo Yu, and Linan Duan. A Novel Image Semantic Communication Method via Dynamic Decision Generation Network and Generative Adversarial Network. *Scientific Reports*, 14(1):19636, 2024.

[124] Nivine Guler and Zied Ben Hazem. Semantic Communication-based Convolutional Neural Network for Enhanced Image Classification. *Franklin Open*, 9:100192, 2024.

[125] Xidong Mu and Yuanwei Liu. Exploiting Semantic Communication for Non-Orthogonal Multiple Access. *IEEE Journal on Selected Areas in Communications*, 41(8):2563–2576, 2023.

[126] Wenqian Tang, Ronghua Liu, Qiang Li, Zian Meng, Ashish Pandharipande, and Xiaohu Ge. Joint Optimization on Mode Switching and Power Adaptation in Semantic-assisted Hybrid NOMA. *IEEE Open Journal of the Communications Society*, 2024.

[127] Ishtiaque Ahmed and Leila Musavian. Hybrid NOMA Assisted Heterogeneous Semantic and Bit Users Communication. *arXiv preprint arXiv:2505.03379*, 2025.

[128] Hanaa Abumarshoud, Bassant Selim, Mallik Tatipamula, and Harald Haas. Intelligent Reflecting Surfaces for Enhanced NOMA-based Visible Light Communications. In *2022 IEEE International Conference on Communications (ICC)*, pages 571–576, 2022.

[129] Chen Chen, Wen-De Zhong, Helin Yang, and Pengfei Du. On the Performance of MIMO-NOMA-Based Visible Light Communication Systems. *IEEE Photonics Technology Letters*, 30(4):307–310, 2018.

[130] Jingjing Cui et al. Optimal User Scheduling and Power Allocation for Millimeter Wave NOMA Systems. *IEEE Transactions on Wireless Communications*, 17(3):1502–1517, 2018.

[131] Weizhi Li, Haotai Liang, Chen Dong, Xiaodong Xu, Ping Zhang, and Kaijun Liu. Non-Orthogonal Multiple Access Enhanced Multi-user Semantic Communication. *IEEE Transactions on Cognitive Communications and Networking*, 2023.

[132] Qiong Zhao, Jiacheng Fan, Bangjiang Lin, and Chen Chen. A Quantized CSI Acquisition Strategy Based on Position Information Feedback for Indoor VLC-Based IoT Systems. *IEEE Internet of Things Journal*, 12(13):25371–25380, 2025.

[133] D. Wu, Z. Ghassemlooy, H. Le Minh, S. Rajbhandari, M. A. Khalighi, and X. Tang. Optimisation of Lambertian Order for Indoor Non-directed Optical Wireless Communication. In *2012 IEEE International Conference on Communications in China Workshops (ICCC)*, pages 43–48, 2012.

[134] Kenichi Higuchi and Anass Benjebbour. Non-orthogonal multiple access (NOMA) with successive interference cancellation for future radio access. *IEICE Transactions on Communications*, 98(3):403–414, 2015.

[135] Roman F. Nalewajski. *Perspectives in Electronic Structure Theory*, chapter 8, pages 371–395. Springer Berlin Heidelberg, 2012.

[136] Shuheng Hua, Yao Sun, Kairong Ma, Dusit Niyato, and Muhammad Ali Imran. A Mathematical Framework of Semantic Communication based on Category Theory. *arXiv preprint arXiv:2504.11334*, 2025.

[137] Hanaa Marshoud et al. Optical Non-Orthogonal Multiple Access for Visible Light Communication. *IEEE Wireless Communications*, 25(2):82–88, 2018.

[138] Raj Jain, Arjan Durresi, and Gojko Babic. Throughput Fairness Index: An Explanation. *ATM Forum contribution*, 99(45):1–13, 1999.