

REINFORCEMENT LEARNING AND BEAMFORMING DESIGN FOR MULTI-USER MULTI-TARGET UAV-ENABLED ISAC SYSTEM

BY

ABDALLA MOHAMED ABDALLA SALIH

A Thesis Presented to the
COLLEGE OF GRADUATE AND INTERDISCIPLINARY STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

ELECTRICAL ENGINEERING

DEC 2024

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN 31261, SAUDI ARABIA

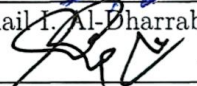
DEANSHIP OF GRADUATE STUDIES

This thesis, written by **ABDALLA MOHAMED ABDALLA SALIH** under the direction of his thesis adviser and approved by his thesis committee, has been presented to and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE IN ELECTRICAL ENGINEERING**.

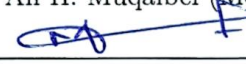
Thesis Committee



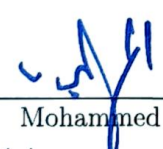
Dr. Suhail I. Al-Dharrab (Adviser)



Dr. Ali H. Muqaibel (Member)



Dr. Mudassir M. Ali (Member)



Dr. Mohammed M. Al-Muhaini
Department Chairman



Dr. Suliman S. Al-Homidan
Dean of Graduate Studies

Date



©Abdalla Salih
2024

To Allah, the Most Merciful, for His endless blessings, guidance, and strength that have carried me through every challenge. To my father, whose guidance, encouragement, and steadfast faith in me have always been a driving force of inspiration. To my mother, whose love, sacrifices, and prayers have shaped the person I am today. This work is a tribute to your unconditional love, encouragement, and inspiration.

ACKNOWLEDGMENTS

I would like to express my heartfelt gratitude to my esteemed supervisor and mentor, Dr. Suhail Al-Dharrab, for his invaluable guidance, unwavering patience, and steadfast support throughout my research journey. Dr. Al-Dharrab's expertise, insightful feedback, and thoughtful mentorship have been invaluable in shaping the direction and scope of my thesis, leaving an indelible mark on my academic and professional growth. I sincerely thank the members of my thesis committee for their insightful feedback and meaningful contributions, which have significantly enhanced the quality of my research. My profound gratitude goes to my father, my mother, and my friends and colleagues, whose unwavering belief in my abilities has been a constant source of inspiration and motivation throughout this journey. Finally, I am deeply thankful to KFUPM for accepting me, and enabling me to pursue my graduate studies and successfully complete my thesis. Their commitment to advancing education and fostering future prospects has been truly invaluable, and I am immensely grateful for their support.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	v
LIST OF TABLES	ix
LIST OF FIGURES	x
LIST OF SYMBOLS	xiii
LIST OF ABBREVIATIONS	xvi
ABSTRACT (ENGLISH)	xix
ABSTRACT (ARABIC)	xxi
CHAPTER 1 INTRODUCTION	1
1.1 Integrated Sensing and Communications	2
1.2 Airborne ISAC Systems	3
1.3 Reinforcement Learning for ISAC Systems	4
1.4 Motivation	7
1.5 Main Contributions	8
1.6 Thesis Structure	9
CHAPTER 2 LITERATURE REVIEW	10
2.1 Joint Sensing and Communications	10
2.2 UAV-Enabled ISAC Systems	12
2.3 UAV Trajectory Optimization	13

2.4	Beamforming Designs	16
 CHAPTER 3 SPECTRAL EFFICIENCY AND HYBRID BEAM-		
FORMING FOR MULTI-USER MULTI-TARGET ISAC SYSTEM		19
3.1	System Model	21
3.2	Channel Model	24
3.3	HBF and DFRC Signal Design	28
3.3.1	Hybrid Beamforming	29
3.3.2	DFRC Received Signal	31
3.4	Communication Performance Analysis	32
3.4.1	Spectral Efficiency	32
3.4.2	Information Outage Probability	33
3.5	Sensing Performance Analysis	35
3.5.1	Target Detection	35
3.5.2	Target Localization	36
3.5.3	Target Tracking	38
3.6	Numerical Simulations	42
3.7	Summary	52
 CHAPTER 4 GENERALIZED GEOMETRIC MODEL FOR PROB-		
ABILISTIC LOS IN IRREGULAR URBAN AREAS		53
4.1	Background	54
4.2	Geometric LoS Probability	57
4.3	Numerical Results	61
4.4	Summary	67
 CHAPTER 5 REINFORCEMENT LEARNING IN MULTI-USER		
MULTI-TARGET UAV-ENABLED ISAC SYSTEM		68
5.1	System Model and Problem Formulation	70
5.1.1	System Model	70
5.1.2	Problem Formulation	72

5.2	Methodology	75
5.2.1	Proximal Policy Optimization Algorithm	76
5.2.2	Bayesian Optimization Method	79
5.2.3	Nonstationary UAV and Target Tracking	82
5.3	RL Training and Numerical Results	84
5.4	Summary	92
CHAPTER 6 CONCLUSIONS AND FUTURE DIRECTIONS		93
6.1	Conclusions	93
6.2	Future Directions	95
APPENDIX A		96
A.1	Spectral Efficiency Derivation	96
A.2	Achievable Rate Derivation	98
A.3	Information-Outage Probability Derivation	99
A.4	Cramér-Rao Lower Bound for Target Localization	101
A.5	Jacobian Matrix for Nonlinear Measurement Model	107
REFERENCES		109
VITAE		121

LIST OF TABLES

Table 2.1:	Summary of recent contributions on UAV-enabled ISAC systems.	18
Table 3.1:	Numerical simulation parameters	42
Table 4.1:	Parameters of the International Telecommunication Union (ITU)-R P.1410 Model for Selected Environments [1]	56
Table 5.1:	Training and Numerical simulation parameters	84
Table 5.2:	Hyperparameters Optimization Range	85
Table 5.3:	Optimized Hyperparameters using Bayesian Optimization . .	85

LIST OF FIGURES

Figure 1.1: Integrated Sensing and Communication (ISAC) system example using single base station for both functionalities. . . .	1
Figure 1.2: Reinforcement Learning (RL) agent and environment interaction	5
Figure 3.1: UAV-assisted ISAC system with hybrid beamforming.	22
Figure 3.2: Hybrid beamforming structure	22
Figure 3.3: One realization of the 3D random scene with locations of UAV, users, scatterers, targets, and users in extreme scenarios.	43
Figure 3.4: Angle estimation at the UAV.	44
Figure 3.5: Spectral efficiency comparison for a single-user under LoS and NLoS conditions with various number scatterers and increasing N_{RF}	44
Figure 3.6: Information-outage probability for single-user with code-words lengths: 6, 15, 30, 60, and 120.	45
Figure 3.7: Spectral efficiency for best and worst user location with antenna elements $N_U = 8$ and 4.	46
Figure 3.8: Average spectral efficiency for various UAV heights/altitudes with $N_V = 32$ and 16.	47
Figure 3.9: Average spectral efficiency for different UAV heights under various environments.	47
Figure 3.10: Average spectral efficiency as number of users increases in various environments.	48

Figure 3.11: Detection probability of one target as the number of communication tasks increases.	49
Figure 3.12: CRLB for distance, velocity, and AoA as the number of communication tasks increases.	50
Figure 3.13: CRLB for distance, velocity, and AoA as the number of sensing pulses increases.	51
Figure 3.14: Illustration of UAV, users, and moving targets and their tracking using EKF.	51
Figure 3.15: CRLB for distance, velocity, and AoA calculated for the tracking process.	52
Figure 4.1: An environment that contain a typical grid layout (R1 and R2) and additional street (R3).	58
Figure 4.2: Comparison between different models in urban environment, $M = \frac{\pi}{4}$	62
Figure 4.3: Comparison between different built-up environment (urban, suburban and dense urban, $M = 2\pi$	63
Figure 4.4: Illustration of multi-legged intersection	64
Figure 4.5: Effect of number of legs in an intersection region on P_{LoS}^* , $M = 2\pi$	65
Figure 4.6: Illustration of an environment with 6 regions in (a) equal distance apart, and (b) non equal distance apart	65
Figure 4.7: Distance between intersection legs effect on P_{LoS}^* , $M = 2\pi$. .	66
Figure 5.1: Illustration of UAV-enabled ISAC system with a trajectory. .	70
Figure 5.2: Average channel gain at an altitude of 10m (black circle: user location).	74
Figure 5.3: Iterative execution of the Bayesian optimization method. . .	86
Figure 5.4: Training performance in terms of cumulative reward.	87
Figure 5.5: Training performance in terms of number of steps.	87

Figure 5.6: A simulation of the trained UAV trajectory, users locations,
and target locations and tracking. 88

Figure 5.7: Trained UAV performance in terms of P_{LoS} and its average. . 89

Figure 5.8: Trained UAV performance in terms of channel gain and its
average. 90

Figure 5.9: Average spectral efficiency with UAV undertaking its planned
trajectory, with and without sensing. 91

LIST OF SYMBOLS

$\boldsymbol{\alpha}[\tau]$	Agent action at time slot τ
$\mathbf{F}_{\text{RF}}$	Analog beamformer
φ_p	Angle of arrival of the p th path for a user
β_t	Average channel power gain for the t th target
ρ	Average number of buildings per unit area
θ_t	Azimuth angle of received echo of the t target
$f(\boldsymbol{\rho})$	Bayesian optimization objective function
$\mathbf{H}_i$	Channel matrix for the i th user DL signal
$\mathbf{H}_t$	Channel matrix for the t th target echo signal
β_{ip}	Average channel power gain for the i th user through the p th path
$\mathbf{F}_{\text{BB}}$	Digital beamformer
d_{ki}	Distance from the k th scatterer to the i th user
d_{ui}	Distance from the UAV to the i th user
d_{uk}	Distance from the UAV to the k th scatterer
d_{ut}	Distance from the UAV to the t th target
ϕ_i	Elevation angle of the UAV for the i th user

ϕ_t	Elevation angle of the UAV for the t th target
$\mathbf{u}_i$	Location of the i th user in 3D space
$\mathbf{k}_k$	Location of the k th scatterer in 3D space
$\mathbf{t}_t$	Location of the t th target in 3D space
$\mathbf{v}$	Location of UAV in 3D space
L_{actor}	Loss function for the actor
L_{critic}	Loss function for the critic
P_T	Maximum transmit power
N_{RF}	Number of activated RF chains
L	Number of communication users
N_P	Number of propagation paths from the UAV to the i th user
N_{RF}^c	Number of RF chains activated for communication
N_{RF}^s	Number of RF chains activated for sensing
$N_{\text{RF}_{\text{max}}}$	Maximum number of available RF chains at the UAV
K	Number of scattering objects
T	Number of targets
$\mathcal{A}$	PPO agent action space
$R[\tau]$	PPO agent reward at time slot τ
$\pi(\psi)$	PPO policy function
$V(\kappa)$	PPO value function
P_{LoS}^*	Probability of LoS for the Intersection region

P_{LoS_i}	Probability of LoS for the i th user
P_{LoS}^r	Probability of LoS for the r th region
ε	Ratio of land covered by buildings
ϑ	Rayleigh parameter for building height variability
S_r	Region r street width
$\mathcal{S}$	RL state space
Θ	Set of angles of propagation paths for users and targets
$\boldsymbol{\rho}$	Set of hyperparameters
$\mathcal{K}$	Set of scatterer locations
$\mathcal{T}$	Set of target locations
$\mathcal{U}$	Set of user locations
$\mathbf{B}(\boldsymbol{\varphi}_i)$	Steering matrix for the i th user
$\mathbf{A}(\Theta)$	Steering matrix for the UAV
$\mathbf{b}(\boldsymbol{\varphi}_p)$	Steering vector for the p th AoA at the user
Δ_T	Time slot duration
N_V	Total number of UAV antenna elements
N_U	Total number of UE antenna elements
$\mathbf{F}_V$	Zero-forcing beamformer
$\mathbf{W}_i$	Zero-forcing combiner at the i th user

LIST OF ABBREVIATIONS

3GPP	Third-Generation Partnership Project
6G	Sixth Generation
AoA	Angle of Arrival
AoD	Angle of Departure
AOI	Age of Information
AWGN	Additive White Gaussian Noise
CAR	Communication-assisted Radar
CDF	Cumulative Distribution Function
CRLB	Cramér-Rao Lower Bound
CRSS	Communication and Radar Spectrum Sharing
CSI	Channel State Information
DDPG	Deep Deterministic Policy Gradient
DFRC	Dual-Function Radar and Communication
DL	Downlink Communication
DP	Downlink Pilot Signal
DRL	Deep Reinforcement Learning

EI	Expected Improvement Acquisition Function
EKF	Extended Kalman Filter
GP	Gaussian Process
HBF	Hybrid Beamforming
ISAC	Integrated Sensing and Communication
ITU	International Telecommunication Union
JRC	Joint Radar and Communication
LFM	Linear Frequency Modulation
LoS	Line of Sight
MDP	Markov Decision Process
MI	Mutual Information
MIMO	Multiple-Input Multiple-Output
ML	Machine Learning
MMSE	Minimul Mean Squere Error
mmWave	Millimeter Wave
MUSIC	MUltiple SIgnal Classification
NLoS	Non-Line of Sight
PDF	Probability Distribution Function
PPO	Proximal Policy Optimization
RAC	Radar-assisted Communications
RCC	Radar and Communication Coexistence

RL	Reinforcement Learning
RT	Ray Tracing
SAC	Soft Actor-Critic
UAV	Unmanned Aerial Vehicle
ULA	Uniform Linear Array
UP	Uplink Pilot Signal
URA	Uniform Rectangular Array

THESIS ABSTRACT

NAME: Abdalla Mohamed Abdalla Salih

TITLE OF STUDY: Reinforcement Learning and Beamforming Design for
Multi-User Multi-Target UAV-Enabled ISAC System

MAJOR FIELD: Electrical Engineering

DATE OF DEGREE: Dec 2024

Unmanned Aerial Vehicles (UAVs) are becoming crucial in future Integrated Sensing and Communication (ISAC) systems, leveraging their unique capabilities in complex environments. The performance of multi-user communications while sensing multi-target in UAV-enabled ISAC systems relies significantly on the UAV position, beamforming design, flight trajectory, and accurate radio propagation channel model for the scene/environment. In this thesis, we investigate the system-level performance for UAV-enabled ISAC systems in terms of main communications and sensing metrics. This includes spectral efficiency, information outage probability, probability of detecting target(s), localization and tracking of target(s), and their associated trade-offs. We further propose a low-complexity hybrid beamforming (HBF) design enabling multi-user aerial communications and sensing of non-stationary targets simultaneously. Due to

the unique nature of air-to-ground channels for UAVs in urban areas, we propose a generalized geometrical model to characterize the probability of line-of-sight (LoS) for regular and irregular built-up environments. Moreover, we investigate and propose a deep reinforcement learning (DRL)-based model using the proximal policy optimization (PPO) algorithm to plan UAV trajectory and enhance the overall system performance. Hyperparameters for the DRL-based model are obtained using Bayesian optimization method, and we finally provide a comparative analysis of the UAV trajectory and overall system performance under diverse conditions.

ملخص الرسالة

الاسم: عبدالله محمد عبدالله صالح

عنوان الدراسة: تصميمات التعلم المعزز و تشكيل حزم الموجات الراديوية لنظام الإتصالات و الإستشعار المتكامل المدعوم بطائرات

بدون طيار للمستخدم و الاهداف المتعددة

التخصص: الهندسة الكهربائية

تاريخ الدرجة العلمية: ديسمبر 2024

من المتوقع أن تلعب الطائرات بدون طيار دوراً مهماً في أنظمة الاستشعار والاتصالات المتكاملة المستقبلية. ومع ذلك، فإن أداء الاتصالات متعددة المستخدمين مع القيام بمهام الاستشعار متعددة الأهداف في الطائرات المدعومة بدون طيار يعتمد بشكل كبير على موقع الطائرة، وتصميم تشكيل حزم موجات الراديو، ومسار الطيران، وكذلك على نموذج دقيق لقناة الانتشار الراديوي للمشهد/البيئة. في هذه الأطروحة، نقترح نموذجاً غير معقد لنظام الاستشعار والاتصالات المتكامل مدعوم بالطائرات بدون طيار، وندرس أدائه الاتصالي من حيث الكفاءة الطيفية واحتمال انقطاع المعلومات، وأدائه في مجال الاستشعار من حيث كشف الأهداف وتحديد مواقعها وتعبئها، بالإضافة إلى الموازنة بين الاستشعار والاتصالات في ظل قيود الطاقة المحدودة للطائرة. يعتمد النموذج المقترح على نموذج تشكيل حزم موجات الراديو الهجين وتقنية التصفير القسري للحزم، مما يمكن الطائرة بدون طيار من التواصل مع عدة مستخدمين بالتوازي، مع القدرة على الاستشعار المتزامن للأهداف ذات الاهتمام. نناقش نماذج احتمال وجود خط نظر مباشر بين المرسل والمستلم ونقترح نموذجاً هندسياً معمماً يأخذ بعين الاعتبار زوايا الارتفاع والأفق بين المستخدم والطائرة، بالإضافة إلى عدم انتظام البيئة المحيطة بالمستخدم. كما نقترح نموذجاً للتعلم العميق المُعزَّز بالاعتماد على تحسين السياسة المقاربة لتخطيط مسار الطائرة، مع الأخذ في الاعتبار معدل فقدان قوة الإشارة و نموذج احتمالية وجود خط نظر مباشر المقترح. بالإضافة إلى ذلك، نقوم بتحسين عوامل النموذج باستخدام طريقة التحسين البايزي. وأخيراً، نحلل سلوك النموذج المُدرَّب في بيئات عشوائية، بضمان عدم تأثر أداء الاستشعار سلباً بحركة الطائرة.

CHAPTER 1

INTRODUCTION

The immense potential of Integrated Sensing and Communication (ISAC) to shape the next generation of wireless communication has sparked significant interest in the industry, particularly in the context of sixth-generation (6G) networks [2]. By merging sensing and communication functionalities, ISAC is anticipated to play a pivotal role in diverse applications, such as autonomous systems, while enhancing communication speeds and reducing latency [3].

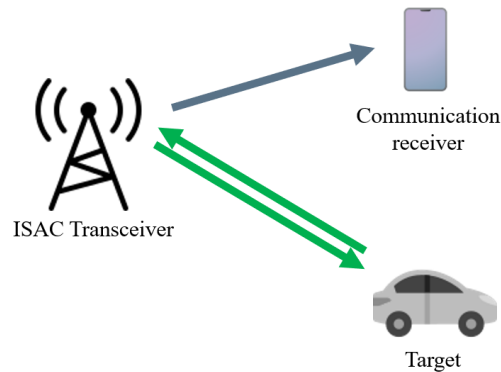


Figure 1.1: ISAC system example using single base station for both functionalities.

The technology is expected to leverage advanced millimeter-wave (mmWave) frequencies (30 GHz to 300 GHz) and Terahertz frequencies (300 GHz to 3 THz) to

achieve extremely high data rates and minimal latency while ensuring sensing accuracy [4]. Furthermore, ISAC is poised to deliver precise and reliable sensing capabilities. With advancements and increased rate of Unmanned Aerial Vehicles (UAVs) manufacturing, UAVs have been proposed as alternatives to range-diminished and overloaded base stations in scenarios such as football matches, concerts, and emergencies [5]. Consequently, UAVs are now attracting significant attention for their potential to assist ISAC systems. Although UAVs offer remarkable mobility compared to location-fixed base stations, designing their trajectories poses challenges due to the demand of their tasks and the limited power resources provided by in the form of batteries. As such, the joint design of ISAC systems and UAV trajectories has garnered the interest of academic and scientific researchers.

1.1 Integrated Sensing and Communications

ISAC is a fully integrated system that incorporates cross-temporal, cross-spatial, and cross-spectrum sharing of resources, providing mutual benefits for both communication and sensing functionalities [6]. Although the term Integrated Sensing and Communication is widely adopted in recent literature, it is also known by other names and acronyms, such as Joint Radar and Communication (JRC), Radar-assisted Communications (RAC), and Communication-assisted Radar (CAR).

Since sensing can be regarded as communication with non-cooperative users, entire communication networks can also function as sensing networks without disrupting their communication functionalities. This dual functionality has become increasingly

feasible due to the existing communication infrastructure. Such networks offer additional benefits, including synchronization, sensory data transportation, and waveform sharing, while enabling advanced sensing capabilities such as high-accuracy localization, activity recognition, and passive object detection and tracking [6]. Similarly, sensing can enhance communication functionalities by providing physical information that aids in improving communication systems. Examples of such information include precise physical locations for accurate beamforming, blockage detection and prediction, and Channel State Information (CSI) [6].

Beyond the informational, spatial, and modular sharing benefits, ISAC also offers spectrum-sharing advantages. By utilizing the same bands in some applications, ISAC has the potential to free up large amounts of spectrum for other uses. Additionally, ISAC’s temporal sharing capability allows it to operate both communication and sensing functionalities in real-time, eliminating the need to coordinate operational windows for these systems.

1.2 Airborne ISAC Systems

The evolution of airborne ISAC systems has been primarily driven by the need for more efficient, reliable, and versatile technologies. Traditionally, airborne platforms such as aircraft, drones, and balloons utilized separate systems for sensing and communication. These systems operated independently, often resulting in inefficiencies in data processing, transmission, and decision-making, as well as limitations in their spatial operations. Due to their capability for mobile deployment, airborne platforms

inherently possess the ability to establish direct Line of Sight (LoS) with ground users or sensing targets. By integrating sensing and communication functionalities into airborne platforms, capabilities such as real-time sensing and data transmission, enhanced situational awareness, and improved operational efficiency in highly dynamic become achievable.

Recently, a wide variety of applications for airborne ISAC systems have emerged, including air traffic management, environmental monitoring, disaster rapid response, infrastructure inspection, and telecommunications, to name a few. However, these benefits come with significant challenges. The integration and miniaturization of such systems are highly complex, and issues such as resource allocation and availability, security, and reliability present additional challenges that demand advanced solutions and focused attention.

1.3 Reinforcement Learning for ISAC Systems

ISAC systems aim to maximize resources utilization by enabling joint sensing and communication, thereby addressing the demands of emerging applications such as autonomous systems, smart cities, and advanced radar systems. However, the inherent complexity of ISAC systems, characterized by high-dimensional data, dynamic environments, and stringent performance requirements, poses significant design and optimization challenges [3]. Machine Learning (ML), is a powerful tool for solving complex, high-dimensional optimization problems in wireless systems. By leveraging large-scale data, ML techniques can model intricate relationships, predict system

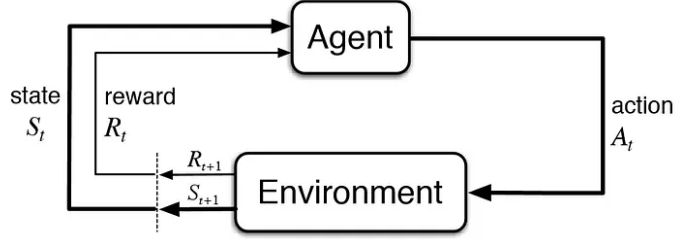


Figure 1.2: RL agent and environment interaction

behaviors, and optimize resource allocation strategies. In the context of ISAC, ML is increasingly being utilized to enhance functionalities such as JRC waveform design, beamforming, resource management, interference management, feature extraction from sensed data, and blockage and hand-off prediction, to name a few. These data-driven approaches provide a level of adaptability and efficiency that traditional rule-based methods often fail to achieve [7].

Among ML paradigms, RL, in particular Deep Reinforcement Learning (DRL), stands out as a promising framework for ISAC systems due to its ability to model decision-making in dynamic and uncertain environments.

RL is a machine learning technique that allows an agent to learn within an interactive environment by trial and error, leveraging feedback from its actions and experiences. In the context of ISAC systems, the agent can represent either the ISAC system itself or an enabling entity, while the environment encompasses data collected regarding the operation of the ISAC system, such as communication metrics, sensing metrics, or physical metrics. Figure 1.2 illustrates a typical interaction between the agent and the RL environment. The agent observes the environment's state S_t , based on which it performs an action A_t . This action alters the environment in some way, resulting in a reward R_t , which indicates the effectiveness of the action taken in in-

fluencing the environment. RL algorithms are broadly classified into model-free and model-based approaches [8]. Model-free algorithms enable the agent to learn directly through interactions with the environment, without requiring a model of the environment's dynamics. In contrast, model-based algorithms utilize a learned or predefined model of the environment to allow the agent to predict future states and rewards, enabling planning and decision-making based on these predictions.

The ability to learn from interacting with the environment and acting upon this interaction is particularly relevant for ISAC systems, where real-time adaptation to operational and environmental changes is critical for maintaining sensing accuracy and communication performance [9], [10]. By leveraging RL, ISAC systems can dynamically optimize parameters such as resource allocation, trajectory planning (specially in UAV-enabled scenarios), and resource sharing between sensing and communication tasks to balance trade-offs [11].

1.4 Motivation

Despite significant advancements in the study of UAV-enabled ISAC systems, there remains a critical gap in the practical implementation of ISAC beamforming and channel characterization models that address several key challenges. These challenges include the limited resources available at the UAV, the interference caused to users by transmitting sensing signals, the complexities of operating in built-up environments, and the multi-dimensional problem of tracking a moving target of interest while the UAV itself is mobile. To address these gaps, the main objectives of the research in this study are summarized as follows:

- Investigate the system-level performance in UAV-enabled ISAC systems and explore effective methods for multi-user communications and multi-target sensing under constrained conditions.
- Develop a generalized channel model that accounts for environmental and resource constraints specific to UAV-enabled ISAC systems.
- Design and optimize ISAC trajectory and plan strategies under similar environmental and resource constraints.

1.5 Main Contributions

In this thesis, we propose a comprehensive framework for UAV-enabled ISAC systems that explicitly considers the power constraints inherent to UAV operations. The framework includes an HBF design tailored to enable simultaneous communication and sensing. We develop a general geometric-based model to characterize the probability of LoS between the UAV and ground-based targets or users in complex built-up environments. This model accounts for the spatial distribution and density of buildings in an urban environment, to accurately predict the LoS probability for various UAV altitudes and positions. Lastly, to address the challenges posed by the randomized environments, we introduce a DRL-based model for optimizing UAV trajectories. The proposed model leverages the developed geometric model for LoS probability to guide trajectory planning and decision-making. By incorporating environmental constraints and real-time feedback, the DRL model learns to navigate the UAV efficiently, ensuring the optimality of communication performance while actively sensing moving targets. Therefore, the primary contributions of this thesis can be summarized as follows:

- A design for UAV-enabled ISAC system that performs sensing and communications simultaneously by employing a hybrid beamforming model.
- Analysis of the proposed model's spectral efficiency and information outage probability, and target's detection, localization and tracking.
- A generalized geometric-based model to characterize the probability of LoS in

regular and irregular landscapes.

- A DRL-based UAV trajectory optimization model with emphasis on communication performance while performing sensing operations.
- hyperparameters for the DRL-based model using the Bayesian optimization method.

1.6 Thesis Structure

The thesis is structured as follows.

- **Chapter 2:** discusses previous publications on integrated sensing and communications systems, UAV-enabled ISAC system, UAV-enabled ISAC trajectory optimizations, and ISAC Beamforming designs.
- **Chapter 3:** presents a low-complexity UAV-enabled ISAC framework that enables simultaneous sensing and communication.
- **Chapter 4:** discusses the characterization of built-up urban environments, and proposes a geometrical model that characterizes the probability of line of sight (LoS).
- **Chapter 5:** proposes a proximal policy optimization model to optimize the trajectory of a UAV-enabled ISAC system.
- **Chapter 6:** discusses final remarks and conclusions, and presents future research direction.

CHAPTER 2

LITERATURE REVIEW

In this chapter, we review and present the state-of-the-art advancements in ISAC systems, with a focus on UAV integration, trajectory optimization, and beamforming designs. These developments highlight the growing importance of ML techniques, particularly RL, in addressing the complex challenges of ISAC systems.

2.1 Joint Sensing and Communications

By leveraging advanced sensing and communication capabilities, ISAC offers transformative opportunities across a range of industries and applications [12], [13]. These capabilities are enabled by techniques such as Communication and Radar Spectrum Sharing (CRSS), spatial and temporal coexistence, and real-time operation. Over the past decade, research has focused on exploring innovative approaches to enhance the design of ISAC systems. One research direction emphasized the independent enhancement of radar and communication functionalities, leading to Radar and Communication Coexistence (RCC) designs that aim to establish an interface between

these functionalities to mitigate spectrum interference [13], [14], [15], [16], [17]. However, coexistence designs have been shown to be inefficient due to the complexity of developing control layers or techniques to manage the sharing of information between radar and communication systems [13]. An alternative research direction has focused on dual-functionality designs that enable simultaneous remote sensing and communication. These designs can be categorized into radar-centric, communication-centric, or joint approaches. Dual-Function Radar and Communication (DFRC) systems have gained considerable attention for their real-time operation, efficient spectrum sharing, and temporal and spatial processing capabilities [13].

Such designs have been widely investigated with a focus of optimizing radar performance under communication constraints. For instance, the authors in [18] proposed a JRC scheme that maximizes the Mutual Information (MI) of a radar target by optimizing radar transmission power under constraints of covertness and information rate. In [19], a matched filter bank technique was introduced to enhance radar performance in systems utilizing DFRC signals. Li et al. in [20] developed a range-space-time response method to estimate range, angle, and Doppler profiles from a 3D radar signal model employing a radar system with frequency diverse array Multiple-Input Multiple-Output (MIMO). Additionally, in [21], the Lagrange multiplier method combined with a cyclic optimization algorithm was used to design a radar-centric system for joint carrier and power allocation.

2.2 UAV-Enabled ISAC Systems

Many of the existing ISAC networks provide both communication and sensing functionalities from a fixed location or a ground-based mobile platform, such as vehicle-to-everything (V2X) or vehicle-to-vehicle (V2V) systems. However, these systems are severely affected by the status of LoS conditions [22]. To mitigate this limitation, researchers are looking into UAVs due to their flexibility and attractive features in positioning themselves in favorable locations, thereby ensuring LoS communications and sensing [23].

Despite these advantages, UAVs face significant challenges stemming from their limited size, weight, and power availability [24]. Taking LoS conditions into account, Oriakumhi et al. in [25] designed a UAV-assisted ISAC system for multi-user downlink communication in the Millimeter Wave (mmWave) band. This system integrates environmental sensing with user tracking while simultaneously providing communication services. By utilizing sensed information, the system predicted potential blockages and identifies alternative communication paths to maximize the future communication sum-rate. Their findings highlighted the significant influence of UAV deployment and trajectory design on effectively managing the inherent sensing and communication trade-offs.

In scenarios where blockages hinder accurate sensing and communication, UAVs have also been proposed as relays. For instance, the authors in [26] introduced a power allocation scheme designed to optimize achievable sum rate in a decode-and-forward relay network utilizing ISAC, while simultaneously ensuring minimum positioning ac-

curacy for both UAVs and users. Similarly, a multi-objective optimization problem was formulated in [27] to maximize the energy efficiency of a UAV-assisted ISAC system and minimize radar MI. This was achieved through the joint optimization of user scheduling, transmission power, and UAV trajectory.

In addition to these advancements, researchers have explored algorithms to enhance resource management in single or multi-UAV scenarios. For example, the authors in [28] proposed an approach utilizing matching theory and fractional programming to efficiently manage coupling between ground users and multiple UAV-enabled ISAC systems. Furthermore, the study in [4] focused on optimizing sensing channel estimation by introducing a Hybrid Beamforming (HBF) architecture, pilot waveform design, and compressive sensing framework, all while addressing hardware constraints of a single UAV. To address the dynamic nature of UAV operations, researchers proposed an improved Unscented Kalman Filter (UKF) beam tracking scheme to predict the angles of arrival and departure AoA and AoD, accounting for UAV mobility [29]. This approach significantly enhances tracking accuracy and adaptability in dynamic environments, further contributing to the effectiveness of UAV-assisted ISAC systems.

2.3 UAV Trajectory Optimization

Researchers have a growing interest in improving the performance of UAV-assisted ISAC systems. These efforts span a wide range of joint objectives, including trajectory optimization, resource allocation, Age of Information (AOI) management, and computation offloading in edge computing environments. One of the notable con-

tributions is by Meng et al. [30], who proposed an integrated periodic sensing and communication (IPSAC) design for UAVs to communicate with a cluster of single-antenna users while sensing a ground target. Their approach focused on optimizing communication by providing an optimum trajectory without location constraints. On the other hand, Jing et al. [31] introduced a UAV trajectory arrangement for ISAC systems, formulating an optimization problem aimed at concurrently enhancing the downlink communication rate and ensuring precise target localization. To optimize the Cramér-Rao Lower Bound (CRLB), they devised a multi-phase strategy for designing trajectories and proposed an iterative algorithm in order to mitigate the challenges posed by non-convex constraints and objectives.

Considering the dynamic complexities and harsh environments UAVs operate in, researchers have also explored trajectory models that incorporate specific environmental factors. For instance, in [32], a trajectory planning method was proposed that accounts for AOI and UAV geometry as an anchor node. In cases where accurate sensing information about users is unavailable, UAVs have been utilized to gather such information by approaching the users' vicinity. Addressing this challenge, the authors in [33] reformulated the trajectory optimization problem into a catenary shape, proposing a continuous-form trajectory solution derived from second-order equations considering force analysis and static equilibrium to optimize network performance. Sequential quadratic programming has also been employed to find sub-optimal solutions for UAV trajectory optimization in multi-UAV, multi-user scenarios [28]. In [34], Meng et al. enhanced the achievable data rate in a UAV-based ISAC system by jointly optimizing user association, sensing time allocation, beamforming, and UAV trajec-

tory. Khalili et al. in [35] proposed a joint resource allocation and trajectory design for a multi-user multi-target UAV-enabled ISAC system, optimizing UAV trajectory, velocity, beamforming, and the sensing power and time. They minimized the consumed power while maintaining quality of service for communications and successful sensing of targets.

In addition to statistical and deterministic trajectory models, DRL is considered to be a powerful tool for addressing trajectory optimization in UAV-assisted ISAC systems [36]. RL has been applied extensively in multi-UAV, multi-user scenarios. For example, in [37], an online convex-assisted RL method was proposed to jointly optimize UAV trajectories, UAV-user associations, and link power control for solar-cell-equipped UAVs. Similarly, in [38], an enhanced RL model was introduced to optimize beam gain by adjusting the trajectory of UAVs in ISAC settings. Several RL algorithms have been studied for trajectory optimization, with policy optimization methods such as Deep Deterministic Policy Gradient (DDPG), Soft Actor-Critic (SAC), and Proximal Policy Optimization (PPO) being the most prominent. The authors in [10] employed DDPG to optimize the trajectory planning of UAV-enabled ISAC systems, while in [18], centralized and decentralized multi-agent SAC algorithms were explored for joint resource allocation and trajectory optimization. To detect multiple targets and transmit data to central UAV in collaborative UAV setting, the authors in [39] used DDPG to maximize the average utility function and minimize the completion time. Lastly, PPO was utilized in multi-agent setting in [40], [41] to address the non-convexity of jointly optimizing UAV trajectories, beamforming, and UAV-user associations. Several of the proposed models have not considered the

channel effect of the communication channel apart from the beam gain, the resultant trajectory on the sensing performance, and the nature of the surrounding environment.

2.4 Beamforming Designs

Beamforming is a critical component in UAV-assisted ISAC systems, as it ensures efficient operation under the dual constraints of sensing and communication. Researchers proposed a UAV-assisted design employing orthogonal time-frequency space (OTFS) modulation with sub-array hybrid beamforming [42]. Using the Minimum Mean Square Error (MMSE) beamforming technique, their design effectively mitigates performance degradation caused by high-speed movements or ground vehicle dynamics. Lyu et al. [43] explored UAV-enabled ISAC in the quasi-stationary and fully mobile scenarios. Their objective was to enhance the total throughput for users, represented as a weighted sum-rate, through jointly optimizing UAV positioning and beamforming strategies for both information transmission and sensing. Their results demonstrated significant improvements in balancing communication and sensing requirements under dynamic conditions. Authors in [44] proposed a UAV-focused, sensing-assisted communication approach for ISAC systems, utilizing a frame structure and extended Kalman filter (EKF)-based prediction. They also used temporal-assisted beamforming enhanced by SCA and iterative rank minimization (IRM) methods to improve directional beamforming, communication efficiency, and object state estimation. For multi-UAV systems, fractional programming (FP) has been employed to reformulate the original non-convex beamforming subproblem into a more tractable form. The

authors in [28] alternately optimized auxiliary variables and beamforming matrices, achieving a local-optimal solution that ensures a balance between sensing and communication performance.

Recent research has also incorporated RL to further enhance beamforming designs. For example in [45], an RL-based antenna division strategy dynamically allocated antennas to non-overlapping sub-arrays in automotive ISAC systems. Convex relaxation was used to optimize the configuration, resulting in improved sensing accuracy and constrained communication performance. Similarly, the authors in [46] utilized a DRL framework with a Wolpertinger policy [47] to optimize both antenna selection and the phases of quantized double-phase shifters. Further extending DRL applications, a beamforming design in [48] employed a Deep Q-Network (DQN) algorithm. This framework modeled beamforming and power allocation as actions within a Markov Decision Process (MDP), with a reward function designed to maximize the achievable sum-rate while meeting CRLB constraints for sensing accuracy. These advancements underscore the potential of DRL in addressing complex beamforming challenges in UAV-assisted ISAC systems.

The following table summarizes the recent literature regarding UAV-assisted or UAV-enabled ISAC, as well as their trajectory optimizations.

Table 2.1: Summary of recent contributions on UAV-enabled ISAC systems.

Ref.	HBF	Height variation	Channel Model	Multi-User	Interference
[4]	✓	-	Simulated Rician fading only	✓	-
[10]	-	-	Analytical large scale fading, probabilistic LoS conditions, and shadowing effects	✓	-
[18], [27], [40]	-	-	Analytical large scale fading only	✓	✓
[25]	-	-	Small scale fading and Doppler effect.	✓	✓
[28]	-	-	Analytical large scale fading, probabilistic LoS conditions, and shadowing effects	✓	✓
[29], [39]	-	✓	Analytical large scale fading only	-	-
[30], [31]	-	-	Analytical large scale fading only	-	-
[33]	-	-	No channel model	✓	-
[38]	-	✓	Analytical large scale fading and simulated Rician fading	✓	✓
[42]	✓	-	Analytical small scale fading and Doppler effect	✓	✓
[43]	-	-	Simulated Rician fading only	✓	✓
[41]	-	✓	Analytical large scale fading only	✓	-
[34], [35]	-	-	Analytical large scale fading only	✓	-
[44]	-	✓	Analytical large scale fading only	✓	✓
Proposed	✓	✓	Analytical large scale fading, probabilistic LoS conditions, and shadowing effects	✓	✓

CHAPTER 3

SPECTRAL EFFICIENCY AND HYBRID BEAMFORMING FOR MULTI-USER MULTI-TARGET ISAC SYSTEM

Beamforming in UAV-enabled ISAC systems is challenging due to both power and hardware constraints, which increases the complexity of 3D beamforming for a high-precision alignment in both azimuth and elevation. Moreover, interference management becomes critical as UAVs have to minimize interference with ground user(s) and/or other UAVs while ensuring efficient sensing and communications. The aforementioned literature indicates a noticeable gap in the hybrid aspect of beamforming design and the limited consideration to interference in the designed models. While beamforming has been extensively studied and utilized in ISAC systems, there is a

for comprehensive approaches to address the challenges posed by interference and incorporate hybrid techniques. Therefore, we propose a low complexity UAV-enabled ISAC system that enables simultaneous communication and sensing while taking into consideration interference, LoS, and Non-Line of Sight (NLoS), environmental conditions¹.

In our considered system, we employ a HBF design that utilizes zero forcing technique to enable the UAV to detect, locate and track targets of interest, while communicating with multiple ground users simultaneously. Our main contributions can be summarized in the following:

- Proposing a low-complexity hybrid beamforming (HBF) design that enables the UAV to simultaneously communicate with multiple users and sense multiple moving targets.
- Validating our formulated and derived expressions and bounds for the sensing and communication system-level performance under several multi-user multi-target scenarios.

The rest of the chapter is organized as follows. We first describe the ISAC system and the components in which we conduct our study in section 3.1. In section 3.2, we describe the DFRC channel. In section 3.3, we give a detailed description about the DFRC and HBF design. We evaluate the system communication performance using spectral efficiency analysis and information outage probability in section 3.4. In section 3.5, we evaluate sensing performance in terms of target detection probability,

¹The work in this chapter has been submitted for publication and it is now under review (2nd round) [49].

localization through CRLB, and tracking by using Extended Kalman Filter (EKF). Finally, we perform numerical simulations of results and provide a summary in sections 3.6 and 3.7, respectively.

3.1 System Model

In this section, we describe the airborne ISAC system model with its surroundings consisting of a UAV, multiple users, scatterers, and targets as shown in Figure 3.1. In such a three-dimensional environment, we have a single UAV, L communication users, T targets that are of interest for UAV to sense, and K scattering objects. The UAV location is denoted by $\mathbf{v} = [x_v, y_v, z_v]^T \in \mathbb{R}^{3 \times 1}$, and the i th user is located at $\mathbf{u}_i = [x_i, y_i, z_i]^T$ among a set of users at $\mathcal{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_L\}$. Each target in set of targets $\mathcal{T} = \{\mathbf{t}_1, \dots, \mathbf{t}_T\}$ is located in $\mathbf{t}_t = [\hat{x}_t, \hat{y}_t, \hat{z}_t]^T \in \mathbb{R}^{3 \times 1}$, while each scatterers in set of scatterers $\mathcal{K} = \{\mathbf{k}_1, \dots, \mathbf{k}_K\}$ is located at $\mathbf{k}_k = [x'_k, y'_k, z'_k]^T \in \mathbb{R}^{3 \times 1}$. Both the UAV and individual users are equipped with Uniform Linear Array (ULA) or Uniform Rectangular Array (URA) with a total number of element N_V , while the individual users are equipped with ULA with N_U elements. The UAV is assumed to have a fully connected HBF structure [50] with RF chains, N_{RF} , to handle sensing and communication as shown in Figure 3.2, such that $1 \leq N_{\text{RF}} \leq N_{\text{RF}_{\max}} \leq N_V$, where $N_{\text{RF}_{\max}}$ is the maximum available RF chains. HBF is a beamforming technique that combines analog and digital beamforming to optimize the transmission and reception of signals in multi-antenna systems. Sensing and communication signals are initially processed using the digital beamformer, $\mathbf{F}_{BB}$, where amplitude weighting is applied to optimize

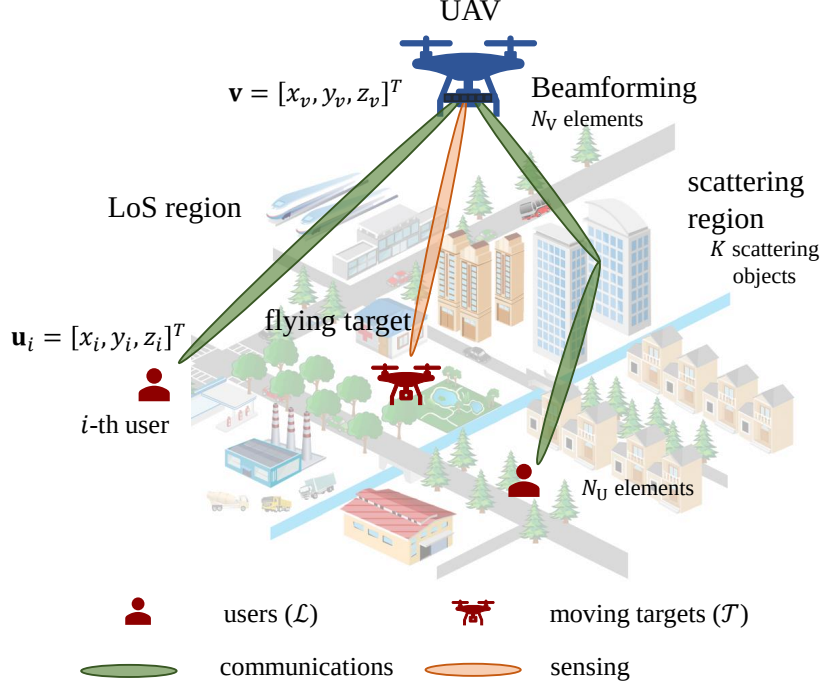


Figure 3.1: UAV-assisted ISAC system with hybrid beamforming.

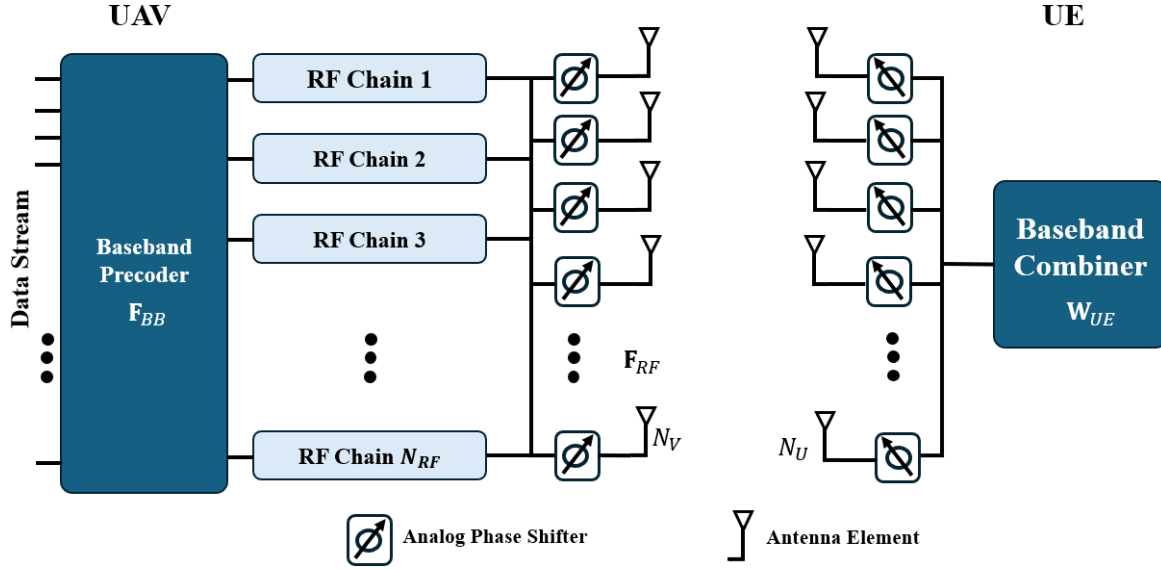


Figure 3.2: Hybrid beamforming structure

data streams and sensory signals. Subsequently, the signals pass through the analog beamformer, $\mathbf{F}_{RF}$, where phase shifters control the directionality across the N_V antenna elements, focusing the transmission on desired directions. Unlike fully digital

beamforming, where each antenna element is connected to a dedicated RF chain for independent signal processing and beamforming, HBF employs a limited number of RF chains, which are fully connected to the N_V antenna elements, by splitting the beamforming process between the digital and analog domains, allowing for greater flexibility and control. This two-stage process enables effective spatial multiplexing and directional signal enhancement while minimizing hardware complexity and energy consumption, which is particularly advantageous for systems with hardware and energy constraints such as UAV-enabled ISAC systems. The number of active RF chains N_{RF} in our model will be affected by the sensing activity and the number of users and their propagation paths, i.e, $N_{\text{RF}} = N_{\text{RF}}^c + N_{\text{RF}}^s$, where N_{RF}^c and N_{RF}^s are the active number of RF chains for communication and sensing respectively. Each user is associated with the UAV and communicating with it only, therefore we assume each user to have a single RF chain as shown in Figure 3.2.

We focus in our communication analysis on the Downlink Communication (DL) and the induced interference where the UAV is actively sensing for targets; however, there is no communications among the users. In our considered scenario, the UAV sends an omnidirectional Downlink Pilot Signal (DP) that reaches targets and users either directly, or reflected through scatterers. Using such received signals, the user can estimate the Angle of Arrival (AoA) at an array of antenna elements with a Multiple Signal Classification (MUSIC) algorithm [51].

An analog beamforming precoder is formulated by the user based on the estimated AoA where the user sends an Uplink Pilot Signal (UP). The UAV then senses the incoming directions of UP signals from the users and the DP signal reflected by

sensing targets using MUSIC algorithm. By decoding the received signals, the UAV can distinguish between the users and targets of interest and proceed to formulate the hybrid beamformer. By doing so, the UAV utilizes the sensing functionality to assist the communication functionality. It is important to note that angle estimation techniques such as MUSIC algorithm yield suboptimal results while the UAV and/or user(s)/target(s) are non-stationary, which could be mitigated and further refined by using techniques such as EKF [44]. For the sake of simplicity, we assume accurate angle estimations.

3.2 Channel Model

We consider in the UAV-assisted ISAC system a probabilistic path loss for the channel between the UAV and users in the sense that it depends on the environment attributes. For mathematical tractability, we assume that both propagation delay and Doppler effects are compensated at the receivers. With this assumption, the channel power gain for the i th user can be determined by the free-space path loss [52]

$$L_{\text{FS}}(d_{\mathbf{v}i}) = 20 \log_{10} \left(\frac{4\pi f_c d_0}{c} \right) + 20 \log_{10} (d_{\mathbf{v}i}), \quad (3.1)$$

where f_c is the carrier frequency, c is the speed of light, d_0 is a reference distance of 1m, and $d_{ui} = \|\mathbf{v} - \mathbf{u}_i\|$ is the distance from the UAV to the intended i th user². The overall path loss experienced is probabilistic where the probability of LoS P_{LoS_i} for the

²Several stochastic channel models, like WINNER II and COST2100, use channel measurements and ray tracing techniques to more accurately represent the channel, however they are beyond the scope of this work.

i th user is [53]

$$P_{\text{LoS}i} = \frac{1}{1 + a \exp(-b(\phi_i + a))}, \quad (3.2)$$

where a and b are the environment-dependant constants [53], and ϕ_i is the elevation angle. We can obtain the average path loss induced on the p th signal propagation path depending on the LoS conditions from the UAV to the i th user using the proposed path loss expression as follows

$$L_{ip}(d) = \begin{cases} P_{\text{LoS}i} L_{\text{FS}}(d_{\mathbf{v}i}) + (1 - P_{\text{LoS}i})(L_{\text{FS}}(d_{\mathbf{v}i}) + L_s), & \text{LoS/Deflection} \\ L_{\text{FS}}(d_{\mathbf{v}k} + d_{ki}), & \text{Reflection} \end{cases}. \quad (3.3)$$

The distances $d_{uk} = \|\mathbf{v} - \mathbf{k}_k\|$ and $d_{ki} = \|\mathbf{k}_k - \mathbf{u}_i\|$ are from the UAV to k th scatterer, and from k th scatterer to i th user, respectively, with the assumption of LoS between them, and L_s is the additional shadowing loss. The average channel power gain β_{ip} for the i th user through the p th propagation path is

$$\beta_{ip} = \frac{1}{\sqrt{L_{ip}(d)}}. \quad (3.4)$$

Similarly, under the assumption of line of sight to target t , the channel gain β_t is

$$\beta_t = \frac{1}{\sqrt{L_{\text{FS}}(d_{\mathbf{v}t})}}, \quad (3.5)$$

where $d_{ut} = \|\mathbf{v} - \mathbf{t}_t\|$ is the distance from the UAV to target t . For the i th user considering multipath propagation from the UAV, we can represent the average channel

power gain vector β_i as

$$\beta_i = [\beta_{i1}, \dots, \beta_{ip}, \dots, \beta_{iN_P}], \quad (3.6)$$

N_P is the number of propagation paths from the UAV to the i th user. Each of those paths has an RF chain dedicated to it, where several pre-transmission operation (such as pre-equalization) could be performed. The UAV is equipped with N_V antenna elements, which has steering capability towards different direction with a steering vector of

- UAV equipped with a ULA antenna

$$\mathbf{a}(\theta_p) = [1, e^{j\frac{2\pi}{\lambda}d_f \sin \theta_p}, \dots, e^{j\frac{2\pi}{\lambda}d_f(N_V-1)\sin \theta_p}]^T \in \mathbb{C}^{N_V \times 1} \quad (3.7)$$

- UAV equipped with a URA antenna

$$a_{n,m}(\theta_p, \phi_p) = \exp \left[j \frac{2\pi}{\lambda} (d(n-1) \sin \theta_p \sin \phi_p - d(m-1) \cos \theta_p \sin \phi_p) \right] \quad (3.8)$$

where n and m indicate the (n, m) th antenna element, knowing that $N_{V_n} \times$

$$N_{V_m} = N_V$$

$$\mathbf{A}(\theta_p, \phi_p) = \begin{bmatrix} a_{1,1}(\theta_p, \phi_p) & a_{1,2}(\theta_p, \phi_p) & \dots & a_{1,m}(\theta_p, \phi_p) \\ a_{2,1}(\theta_p, \phi_p) & a_{2,2}(\theta_p, \phi_p) & \dots & a_{2,m}(\theta_p, \phi_p) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1}(\theta_p, \phi_p) & a_{n,2}(\theta_p, \phi_p) & \dots & a_{n,m}(\theta_p, \phi_p) \end{bmatrix} \quad (3.9)$$

$$\mathbf{a}(\theta_p, \phi_p) = \text{vec}(\mathbf{A}(\theta_p, \phi_p)) \in \mathbb{C}^{N_V \times 1} \quad (3.10)$$

is the j propagation path, ϕ_p is the j th elevation Angle of Departure (AoD), and $d_f = \lambda/2$ is the spacing between the UAV antenna elements where λ is the wavelength. It is important to note that both ULA and URA will give similar numerical results assuming both arrays have the same number of elements. Unless otherwise stated, we will assume the UAV is using ULA antenna. On the other hand, the steering vector $\mathbf{b}(\varphi_p)$ at the end user side is

$$\mathbf{b}(\varphi_p) = [1, e^{j\frac{2\pi}{\lambda}d_f \sin \varphi_p}, \dots, e^{j\frac{2\pi}{\lambda}d_f (N_U - 1) \sin \varphi_p}]^T \in \mathbb{C}^{N_U \times 1} \quad (3.11)$$

where φ_p is p th AoA for the user. In case of multiple propagation paths, the steering matrices for the UAV $\mathbf{A}(\Theta)$ and the i th user $\mathbf{B}(\varphi_i)$ based on (3.7) and (3.11), respectively, are

$$\mathbf{A}(\Theta) = [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_{N_{RF}})] \in \mathbb{C}^{N_V \times N_{RF}} \quad (3.12)$$

$$\mathbf{B}(\varphi_i) = [\mathbf{b}(\varphi_1), \dots, \mathbf{b}(\varphi_{N_P})] \in \mathbb{C}^{N_U \times N_P} \quad (3.13)$$

where $\Theta = \{\theta_1, \dots, \theta_{N_{\text{RF}}}\}$ containing the angles of the propagation paths for all users and targets (i.e. N_{RF}), which themselves can be directly or through known scatterers, which depend on LoS conditions. While $\boldsymbol{\varphi} = \{\varphi_1, \dots, \varphi_{N_P}\}$ contains angles of N_P propagation paths for the i th user. Using (3.4), (3.12) and (3.13), we can formulate and represent the channel matrix $\mathbf{H}_i$ for the i th user as

$$\mathbf{H}_i = \mathbf{B}(\boldsymbol{\varphi}_i) \text{diag}(\boldsymbol{\beta}_i) \mathbf{A}(\Theta_i)^T \in \mathbb{C}^{N_U \times N_V}, \quad (3.14)$$

and, therefore, the received DL signal $\mathbf{y}_i$ at the i th user through N_P paths is

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{x}_i + \mathbf{n}_i \in \mathbb{C}^{N_U \times 1} \quad (3.15)$$

where $\mathbf{x}_i$ is the transmitted signal with a certain energy per bit intended for the i th user, and $\mathbf{n}_i$ is the Additive White Gaussian Noise (AWGN), $\mathbf{n}_i \sim \mathcal{N}(0, \sigma_n^2 \mathbf{I}_{N_U})$, where $\mathbf{I}_{N_U}$ is the identity matrix of size N_U . In the next section, we derive the hybrid beamforming taking into consideration the received DFRC signal.

3.3 HBF and DFRC Signal Design

For sensing and communication functionalities, we design a zero forcing HBF and the DFRC signal. This design integrates analog and digital beamforming techniques to minimize the $N_{\text{RF}_{\text{max}}}$ and maximize the received power at the target user. In zero forcing beamformer, most of the power is directed to the intended receiver while nulls

are obtained in other directions³.

3.3.1 Hybrid Beamforming

First, we form a generalized steering matrix for both sensing and communications functionalities at UAV using (3.7), i.e.

$$\mathbf{A}(\Theta) = [\mathbf{A}(\Theta_1), \dots, \mathbf{A}(\Theta_L), \mathbf{A}(\Theta_T)] \in \mathbb{C}^{N_V \times N_{\text{RF}}} \quad (3.16)$$

where $\mathbf{A}(\Theta_1)$ through $\mathbf{A}(\Theta_L)$ are the steering matrix towards communication users, and $\mathbf{A}(\Theta_T)$ is the steering matrix for sensing T targets (assuming each with a single propagation path). With this formulation, the analog beamformer $\mathbf{F}_{\text{RF}}$ can be obtained taking the conjugate of $\mathbf{A}(\Theta)$ in (3.16)

$$\mathbf{F}_{\text{RF}} = \mathbf{A}^*(\Theta) \in \mathbb{C}^{N_V \times N_{\text{RF}}}. \quad (3.17)$$

To formulate the digital beamformer, we first introduce a zero-forcing precoder $\mathbf{F}_V$, which can be expressed as

$$\mathbf{F}_V = \tilde{\mathbf{H}}^H (\tilde{\mathbf{H}} \tilde{\mathbf{H}}^H)^{-1} \in \mathbb{C}^{N_V \times N_P} \quad (3.18)$$

where $\tilde{\mathbf{H}} = \mathbf{A}^T(\Theta_i)$ and $(\cdot)^H$ is the Hermitian transpose. The complex digital beamformer $\mathbf{F}_{\text{BB}}$ is a square matrix of order N_{RF} can be found by minimizing the objective

³We assume perfect channel state information for the zero-forcing beamformer to nullify transmission in directions other than intended user(s) as in [13], [35], [54].

function[13]

$$\begin{aligned} & \arg \min_{\mathbf{F}_{\text{BB}}} \|\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}} - [\mathbf{F}_{\text{V}}, \mathbf{0}]\|_F^2 \\ & \text{s.t. } \mathbf{F}_{\text{BB}}\mathbf{F}_{\text{BB}}^H = \frac{P_T}{N_{\text{RF}}N_{\text{V}}}\mathbf{I}_{N_{\text{RF}}}, \end{aligned} \quad (3.19)$$

where $\|\cdot\|_F$ is the Frobenius norm. The closed-form solution of (3.19) is given by

$$\mathbf{F}_{\text{BB}} = \sqrt{\frac{P_T}{N_{\text{RF}}N_{\text{V}}}}\tilde{\mathbf{U}}\tilde{\mathbf{V}}^H \quad (3.20)$$

where P_T is the maximum transmit power, and $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{V}}^H$ are the singular value decomposition (SVD)⁴ of $\mathbf{F}_{\text{RF}}^H[\mathbf{F}_{\text{V}}, \mathbf{0}]$, where $\mathbf{0}$ is a zero matrix of dimension $N_{\text{V}} \times (N_{\text{RF}} - N_P)$, which will minimize the interference for i th user [13]. The size of resultant matrix depends on the number of beams transmitted. We can rewrite $\mathbf{F}_{\text{BB}}$ as

$$\mathbf{F}_{\text{BB}} = [\mathbf{F}_{\text{BB}_1}, \mathbf{F}_{\text{BB}_2}, \mathbf{F}_{\text{BB}_3}] \in \mathbb{C}^{N_{\text{RF}} \times N_{\text{RF}}} \quad (3.21)$$

where $\mathbf{F}_{\text{BB}_1} \in \mathbb{C}^{N_{\text{RF}} \times N_P}$ is for the intended communication user, $\mathbf{F}_{\text{BB}_2} \in \mathbb{C}^{N_{\text{RF}} \times L-1}$ is for the unintended users in the scene, and $\mathbf{F}_{\text{BB}_3} \in \mathbb{C}^{N_{\text{RF}} \times T}$ is for sensing targets. When receiving the DL signal, the i th user will combine the signal to be successfully decoded with a zero forcing combiner $\mathbf{W}_i$ using $\mathbf{B}(\varphi_i)$ in (3.13) as

$$\mathbf{W}_i = (\mathbf{B}^H(\varphi_i)\mathbf{B}(\varphi_i))^{-1}\mathbf{B}^H(\varphi_i) \in \mathbb{C}^{N_P \times N_{\text{U}}}. \quad (3.22)$$

⁴The typical computation of SVD is in order of $\mathcal{O}(mn^2)$ for a matrix of size m by n , therefore the computational complexity of computing $\mathbf{F}_{\text{BB}}$ is in the order of $\mathcal{O}(N_{\text{V}}N_{\text{RF}}^2)$.

3.3.2 DFRC Received Signal

Let $\mathbf{x} = [\mathbf{x}_i, \mathbf{x}_l, \mathbf{x}_t]^T$ denotes the DFRC signal sent by the UAV, which consists of vectors of orthogonal Linear Frequency Modulation (LFM) signals intended for communication and sensing, and assumed to be independent with zero mean and covariance matrix $\mathbb{E}[\mathbf{x}\mathbf{x}^H] = \mathbf{I}_{L+T}$. Here, $\mathbf{x}_i \in \mathbb{C}^{N_P \times 1}$ contains useful communication information intended for the i th user, while $\mathbf{x}_l \in \mathbb{C}^{(N_{\text{RF}}^c - N_P) \times 1}$ contains useful communication information however its intended for neighboring users. $\mathbf{x}_t \in \mathbb{C}^{T \times 1}$ contains no useful information since it is dedicated to only sense targets, however, it uses the same waveform as the communication signals.

Using this signal, the UAV can reliably sense targets and communicate with multiple users. Therefore, the received signal at the i th user is

$$\mathbf{y}_i = \mathbf{W}_i \mathbf{H}_i \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_1} \mathbf{x}_i + \mathbf{W}_i \mathbf{H}_i \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_2} \mathbf{x}_l + \rho \mathbf{W}_i \mathbf{H}_i \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_3} \mathbf{x}_t + \mathbf{W}_i \mathbf{n}_i \quad (3.23)$$

where $\rho = 1$ when sensing is active and 0 otherwise, the desired signal is the first term in (3.23), the second term is the interfering communication signal intended for the neighboring users, however, it is received by the i th user. The third term is the interfering sensing signal, and $\mathbf{n}_i$ is the AWGN.

3.4 Communication Performance Analysis

3.4.1 Spectral Efficiency

To evaluate the spectral efficiency for the i th user, we identify the desired part of $\mathbf{y}_i$ in (3.23) and remaining terms will be the interference and noise. The information for the i th user, $\mathbf{x}_i$, is conditioned on the knowledge of the channel matrix $\mathbf{H}_i$. Therefore, spectral efficiency for the i th user is given by

$$SE_i = \log_2 \det \left(\mathbf{I}_{N_P} + \mathbf{R}_i' \mathbf{R}_{in}'^{-1} \right) \quad (3.24)$$

Proof: See the Appendix A.1.

Here, $\mathbf{R}_i$ and $\mathbf{R}_{in}$ correspond to the matrix of the intended i th user and the matrix of the interference, respectively, which are given by

$$\mathbf{R}_i = \mathbf{W}_i \mathbf{H}_i \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_1} \mathbf{F}_{\text{BB}_1}^H \mathbf{F}_{\text{RF}}^H \mathbf{H}_i^H \mathbf{W}_i^H \quad (3.25)$$

$$\mathbf{R}_{in} = \mathbf{R}_l + \mathbf{R}_t + \sigma^2 \mathbf{W}_i \mathbf{W}_i^H \quad (3.26)$$

where

$$\mathbf{R}_l = \mathbf{W}_i \mathbf{H}_i \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_2} \mathbf{F}_{\text{BB}_2}^H \mathbf{F}_{\text{RF}}^H \mathbf{H}_i^H \mathbf{W}_i^H \quad (3.27)$$

$$\mathbf{R}_t = \mathbf{W}_i \mathbf{H}_i \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_3} \mathbf{F}_{\text{BB}_3}^H \mathbf{F}_{\text{RF}}^H \mathbf{H}_i^H \mathbf{W}_i^H. \quad (3.28)$$

Since we are using a zero-forcing technique, the interference term in (3.24) will have

less effect on the spectral efficiency compared to the noise. To calculate it for different users, we prioritize that users steering vector in (3.12), which means we reorder the columns to have the steering vector of that user with high priority first then redo the steps till we get the bandwidth efficiency.

Using $\mathbf{R}_i$ and $\mathbf{R}_{in}$, the theoretical reliable communication with infinite codeword for the i th user is possible at any point under the curve found from

$$\frac{R_i}{B} < \log_2 \det \left[\mathbf{I}_{N_P} + \frac{R_i}{B} \gamma_0 \mathbf{R}_i' \left(\frac{R_i}{B} \gamma_0 (\mathbf{R}_l' + \mathbf{R}_t') + \mathbf{W}_i \mathbf{W}_i^H \right)^{-1} \right] \quad (3.29)$$

where γ_0 is the signal to noise ratio per bit, and R/B is the spectral efficiency.

Proof: See the Appendix A.2.

3.4.2 Information Outage Probability

In this subsection, we consider the realistic constraints of finite codeword lengths. Conventionally, the spectral efficiency (or ergodic capacity) analysis assumes infinite codeword length with tractable mathematical derivations [55]. However, in practical communication systems, the codeword length is inherently finite. To bridge this gap between theory and practice, we consider the approximated information-outage probability as a metric to assess the efficacy and reliability of the proposed ISAC system with HBF. We incorporate the approximation of information-outage probability under the Gaussian channel to evaluate the system performance.

Let n denotes the length of the finite codeword, and $\mathbf{x} \sim \mathcal{N}(0, \mathbf{R}_i)$ is a vector of independent and identically distributed (i.i.d.) normally distributed random variables

(r.v.'s). Therefore, the MI rate between the sample vector, $\mathbf{x} \in \mathbb{C}^N$, and received sample vector, $\mathbf{y} \in \mathbb{C}^N$, is

$$\begin{aligned} i(\mathbf{y}; \mathbf{x}) &= \frac{1}{n} \log \frac{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})}{p_{\mathbf{Y}}(\mathbf{y})} \\ &= SE_i + W_n \end{aligned} \quad (3.30)$$

where W_n follows Laplace distribution with zero mean and variance of $\sigma_W^2 = \det(2\mathbf{R}_i(\mathbf{R}_i + \mathbf{R}_{in})^{-1})$ in case of $n = 1$. For the case of $n > 1$, W_n in (3.30) is a Bessel- K r.v. with a zero mean, variance of σ_W^2 , and the distribution depends on the modified Bessel function of second kind with order $n - \frac{1}{2}$. As the codeword length increases, i.e., $n \rightarrow \infty$, the MI tend to be approximated by Gaussian using the central limit theorem [55]. In this context, it is safe to assume that $i(\mathbf{y}; \mathbf{x})$ is approximately Gaussian with mean SE_i and variance σ_W^2/n . Consequently, the information-outage probability P_{out} with finite codeword length can be formulated as follows

$$P_{out} = P[i(\mathbf{y}; \mathbf{x}) \leq R_e] = Q\left(\frac{SE_i - R_e}{\sqrt{\sigma_W^2/n}}\right) \quad (3.31)$$

where $R_e = R_c \log(2)$ is the equivalent code rate in nats per symbol, code rate $R_c = k/n$, k is information bits in a message, and n is number of symbols in a codeword.

Proof: See the Appendix A.3.

3.5 Sensing Performance Analysis

3.5.1 Target Detection

In resource-constraint sensing platforms such as UAVs, sensing in omni-directional manner can lead to wasting resources, which will degrade the performance of other functionalities. Therefore, utilizing initial estimated AoA can help narrow the focus of sensing, and thus, ensuring adequate power reservation. By gaining initial information of the target echo AoA, the UAV can direct a beam towards the general direction of the target.

Assuming the target is present, the reflected echo $\mathbf{y}_t$ received by the UAV is

$$\mathbf{y}_t = \mathbf{w}_t \mathbf{H}_t \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_3} \mathbf{x}_t + \mathbf{w}_t \mathbf{n} \quad (3.32)$$

where $\mathbf{H}_t = \beta_t \mathbf{a}(\theta_t) \mathbf{a}^H(\theta_t) e^{j2\pi f_D l}$, $\mathbf{w}_t = \mathbf{a}^H(\theta_t)$ is the analog combiner at the UAV for target t beam, f_D is the Doppler frequency, $\mathbf{n}$ is the complex-valued AWGN noise with zero mean and covariance $\sigma^2 \mathbf{I}_{N_V}$. We formulate the probability of detection as the binary hypothesis problem, where hypothesis $\mathcal{H}_0$ for no target detection, and hypothesis $\mathcal{H}_1$ for target detection. With the decision threshold δ is set by the operator to achieve a desired probability of false alarm [56]

$$\mathcal{H}_0 : \|\mathbf{y}_t\|^2 < \delta \quad (3.33)$$

$$\mathcal{H}_1 : \|\mathbf{y}_t\|^2 \geq \delta \quad (3.34)$$

As such, under $\mathcal{H}_0$ [57]

$$\|\mathbf{y}_t\|^2 \sim \frac{\mathbf{w}_t \mathbf{w}_t^H \sigma^2}{2} \chi_2^2, \quad (3.35)$$

while under $\mathcal{H}_1$

$$\|\mathbf{y}_t\|^2 \sim \left(\frac{\mathbf{w}_t \mathbf{w}_t^H \sigma^2}{2} + \frac{\|\mathbf{w}_t \mathbf{H}_t \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_3}\|^2}{2} \right) \chi_2^2, \quad (3.36)$$

where χ_2^2 is the central chi-squared distribution with two degrees of freedom [58].

Assuming a constant probability of false alarm P_{FA} , δ can be determined by

$$\delta = \frac{\mathbf{w}_t \mathbf{w}_t^H \sigma^2}{2} F_{\chi_2^2}^{-1}(1 - P_{FA}) \quad (3.37)$$

where $F_{\chi_2^2}^{-1}(\cdot)$ is the inverse Cumulative Distribution Function (CDF) of the chi-square random variable. Consequently, we can formulate the probability of detection as

$$P_D = 1 - F_{\chi_2^2} \left(\frac{2\delta}{\mathbf{w}_t \mathbf{w}_t^H \sigma^2} / \left(1 + \frac{\|\mathbf{w}_t \mathbf{H}_t \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_3}\|^2}{\mathbf{w}_t \mathbf{w}_t^H \sigma^2} \right) \right) \quad (3.38)$$

where $F_{\chi_2^2}(\cdot)$ is the CDF of the chi-square random variable.

3.5.2 Target Localization

Localization is generally defined as parameter estimation of an object in terms of location, range, velocity or direction. While mean squared error (MSE) is a straightforward metric for assessing localization accuracy by comparing true parameters with their estimates, it is often challenging to explicitly characterize. Instead, the CRLB is frequently utilized as it establishes a theoretical lower limit on the variance achievable

by any unbiased estimator. As such we present the CRLB for estimating the range, velocity, and direction to the t th target as follows:

- **CRLB for distance $d_{\mathbf{v}t}$**

$$\begin{aligned} CRLB_{d_{\mathbf{v}t}} &= \frac{1}{E \left[\frac{\partial}{\partial d_{\mathbf{v}t}} \ln \mathbb{P}(\mathbf{y}_t[\tau]; [\theta_t, \phi_t, v_t, d_{\mathbf{v}t}]) \right]^2} \\ &= \frac{1}{\left[-\frac{Q \frac{\partial C}{\partial d_{\mathbf{v}t}} (U-2C)}{2C^2} \right]^2} = \frac{4C^4}{\left[Q \frac{\partial C}{\partial d_{\mathbf{v}t}} (U-2C) \right]^2} \end{aligned} \quad (3.39)$$

- **CRLB for velocity v_t**

$$\begin{aligned} CRLB_{v_t} &= \frac{1}{E \left[\frac{\partial}{\partial v_t} \ln \mathbb{P}(\mathbf{y}_t[\tau]; [\theta_t, \phi_t, v_t, d_{\mathbf{v}t}]) \right]^2} \\ &= \frac{1}{\left[-\frac{Q \frac{\partial C}{\partial v_t} (U-2C)}{2C^2} \right]^2} = \frac{4C^4}{\left[Q \frac{\partial C}{\partial v_t} (U-2C) \right]^2} \end{aligned} \quad (3.40)$$

- **CRLB for azimuth angle θ_t**

$$\begin{aligned} CRLB_{\theta_t} &= \frac{1}{E \left[\frac{\partial}{\partial \theta_t} \ln \mathbb{P}(\mathbf{y}_t[\tau]; [\theta_t, \phi_t, v_t, d_{\mathbf{v}t}]) \right]^2} \\ &= \frac{1}{\left[-\frac{Q \frac{\partial C}{\partial \theta_t} (U-2C)}{2C^2} \right]^2} = \frac{4C^4}{\left[Q \frac{\partial C}{\partial \theta_t} (U-2C) \right]^2} \end{aligned} \quad (3.41)$$

- In case of using URA is used, **CRLB for elevation angle** ϕ_t

$$\begin{aligned}
CRLB_\phi &= \frac{1}{E \left[\frac{\partial}{\partial \phi_t} \ln \mathbb{P}(\mathbf{y}_t[\tau]; [\theta_t, \phi_t, v_t, d_{\mathbf{v}t}]) \right]^2} \\
&= \frac{1}{\left[-\frac{Q \frac{\partial C}{\partial \phi_t} (U - 2C)}{2C^2} \right]^2} = \frac{4C^4}{\left[Q \frac{\partial C}{\partial \phi_t} (U - 2C) \right]^2} \quad (3.42)
\end{aligned}$$

where Q is the number of sensing pulses echoed by the t th target, C is the variance of the received echo, and U is the variance of sensing signal, which are given by

$$C = \|\mathbf{w}_t \mathbf{H}_t \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_3}\|^2 + \mathbf{w}_t \mathbf{w}_t^H \sigma^2, \quad (3.43)$$

$$U = \|\mathbf{w}_t \mathbf{H}_t \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_3}\|^2. \quad (3.44)$$

Proof: See the Appendix A.4.

3.5.3 Target Tracking

The third sensing task, tracking, is generally considered to be an extension of localization. While localization involves determining the position of a static or moving target, tracking focuses on continuously estimating the position and velocity of a moving target over time. In this context, tracking is particularly challenging due to the dynamic nature of the target, the presence of noise in measurements, and potential uncertainties in the system model. We employ the EKF, to estimate the state of a moving target effectively. For simplicity, we consider the tracked target is moving at a constant speed. The problem of tracking a moving target becomes significantly

more challenging when the UAV itself is in motion. This complexity arises due to the need for simultaneous estimation of the target's state and the dynamic adjustments required to account for the UAV's changing position. The relative motion of the UAV and the sensed target introduces additional uncertainties and noise into the system, complicating the process of accurately predicting the target's trajectory. To simplify the analysis at this stage, we assume that the UAV is hovering in a fixed position, thereby eliminating the complications introduced by UAV movement. The more complex scenario of tracking a moving target while the UAV is also in motion will be addressed in detail in Chapter 5.

As such, the motion state vector of the t -th target is expressed at time slot τ as

$$\boldsymbol{\xi}_{t,\tau} = [\mathbf{t}_{t,\tau}, \dot{\mathbf{t}}_{t,\tau}]^T, \quad (3.45)$$

where $\mathbf{t}_{t,\tau}$ and $\dot{\mathbf{t}}_{t,\tau} = [\dot{x}_{t,\tau}, \dot{y}_{t,\tau}, \dot{z}_{t,\tau}]^T$ are the position and velocity vectors of the t -th target. As such we can model the state evolution as follows

$$\boldsymbol{\xi}_{t,\tau} = \mathbf{T}\boldsymbol{\xi}_{t,\tau-1} + \tilde{\mathbf{n}}_{t,\tau}, \quad (3.46)$$

where

$$\mathbf{T} = \begin{bmatrix} 1 & \Delta_T \\ 0 & 1 \end{bmatrix} \otimes \mathbf{I}_3, \quad (3.47)$$

is the transition matrix, Δ_T is the time slot duration. $\tilde{\mathbf{n}}_{t,\tau}$ is the process noise that follows the Gaussian distribution with zero mean and $\mathbf{Q}$ covariance and noise intensity

of σ_p^2

$$\mathbf{Q} = \begin{bmatrix} \frac{\Delta_T^3}{3} & \frac{\Delta_T^2}{2} \\ \frac{\Delta_T^2}{2} & \Delta_T \end{bmatrix} \otimes \sigma_p^2 \mathbf{I}_3. \quad (3.48)$$

The nonlinear measurement observed from the received echo can be expressed as follows

$$\mathbf{z}_{t,\tau} = h(\boldsymbol{\xi}_{t,\tau}) + \check{\mathbf{n}}_{t,\tau}, \quad (3.49)$$

such that $h(\cdot)$ is the non-linear operation for transforming Cartesian to Spherical coordinates

$$h(\boldsymbol{\xi}_{t,\tau}) = \begin{cases} d_{ut,\tau} = \|\mathbf{v} - \mathbf{t}_{t,\tau}\| \\ v_{t,\tau} = (\hat{x}_{t,\tau}\dot{\hat{x}}_{t,\tau} + \hat{y}_{t,\tau}\dot{\hat{y}}_{t,\tau} + \hat{z}_{t,\tau}\dot{\hat{z}}_{t,\tau})/d_{ut,\tau} \\ \theta_{t,\tau} = \arctan(\hat{y}_{t,\tau}/\hat{x}_{t,\tau}) \\ \phi_{t,\tau} = \arctan\left(\hat{z}_{t,\tau}/\sqrt{\hat{x}_{t,\tau}^2 + \hat{y}_{t,\tau}^2}\right) \end{cases}, \quad (3.50)$$

where $\check{\mathbf{n}}_{t,\tau}$ is the measurement noise that is assumed to be AWGN with zero mean and variance of

$$\boldsymbol{\sigma}_m^2 = \text{diag}(\sigma_d^2, \sigma_v^2, \sigma_\theta^2, \sigma_\phi^2). \quad (3.51)$$

We assume that $\sigma_d^2, \sigma_v^2, \sigma_\theta^2$, and σ_ϕ^2 to be the CRLB for distance, range, azimuth and elevation angles respectively. The extended Kalman filter for tracking the t -th target begins with the prediction phase, which consists of two key components:

- State prediction $\boldsymbol{\xi}_t$, which we denote as $\boldsymbol{\xi}_{t,\tau}^-$

$$\boldsymbol{\xi}_{t,\tau}^- = \mathbf{T}\boldsymbol{\xi}_{t,\tau-1}. \quad (3.52)$$

- Covariance prediction $\mathbf{P}_t$, which we denote as $\mathbf{P}_{t,\tau}^-$

$$\mathbf{P}_{t,\tau}^- = \mathbf{T}\mathbf{P}_{t,\tau-1}\mathbf{T} + \mathbf{Q}. \quad (3.53)$$

After the prediction phase, the EKF proceeds to the update phase, where the state prediction is corrected using incoming measurements $\mathbf{z}_{t,\tau}$:

1. Kalman gain calculation $\mathbf{K}_{t,\tau}$

$$\mathbf{K}_{t,\tau} = \mathbf{P}_{t,\tau}^- \tilde{\mathbf{J}}^T \left(\tilde{\mathbf{J}}\mathbf{P}_{t,\tau}^- \tilde{\mathbf{J}}^T + \sigma_m^2 \right)^{-1}, \quad (3.54)$$

where $\tilde{\mathbf{J}}$ is the Jacobian matrix of $h(\boldsymbol{\xi}_{t,\tau})$

$$\tilde{\mathbf{J}}_\tau = \frac{\partial h(\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} \Big|_{\boldsymbol{\xi}=\boldsymbol{\xi}_{t,\tau}^-}. \quad (3.55)$$

Proof: See the Appendix A.5

2. State update, notated by $\boldsymbol{\xi}_{t,\tau}^+$

$$\boldsymbol{\xi}_{t,\tau}^+ = \boldsymbol{\xi}_{t,\tau}^- + \mathbf{K}_{t,\tau} (\mathbf{z}_{t,\tau} - h(\boldsymbol{\xi}_{t,\tau}^-)). \quad (3.56)$$

3. Covariance update

$$\mathbf{P}_{t,\tau} = \mathbf{P}_{t,\tau}^- - \mathbf{K}_{t,\tau} \tilde{\mathbf{J}} \mathbf{P}_{t,\tau}^- \quad (3.57)$$

It is important to note that the estimate $\boldsymbol{\xi}_{t,\tau}^+$ will follow the normal distribution, i.e., $\boldsymbol{\xi}_{t,\tau}^+ \sim \mathcal{N}(\boldsymbol{\xi}_{t,\tau}, \mathbf{P}_{t,\tau})$.

3.6 Numerical Simulations

In this section, we provide and discuss extensive Monte-Carlo simulations of our system model considering the parameters given in Table 3.1. Unless otherwise stated, the location of users, scatterers and targets are uniformly distributed within the standard ranges in (x, y, z) directions set in the table. We assume the users are located on the floor of the 3D plane, i.e. $z = 0$, while the targets and scatterers can be in the feasible ranges.

Table 3.1: Numerical simulation parameters

Parameter Description	Value
Maximum transmit power (P_T)	20 dBm
Noise power (σ^2)	−80 dBm
Carrier frequency (f_c)	28 GHz
Number of scatterers (K)	4
Targets to be sensed (T)	2
Number of neighboring users (L)	9
Number of ULA elements at UAV (N_V)	32
Number of URA elements at UAV ($[N_{V_n}, N_{V_m}]$)	[8, 8]
Number of ULA elements at user (N_U)	8
Environment Range in x -axis	[−100, 100] m
Environment Range in y -axis	[−100, 100] m
Environment Range in z -axis	[0, 100] m
Shadowing loss (L_s)	10 dB
Code Rate (R_c)	2/3 bits/symbol
Time slot duration (Δ_T)	0.1 s
UAV cruise altitude variations (z -axis)	[45, 55] m
Constants (a, b) for suburban	(4.88, 0.43)
Constants (a, b) for urban	(9.61, 0.158)
Constants (a, b) for dense urban	(12.08, 0.11)

The flying UAV in the scene can be either in the cruising or hovering mode, and a possible realization is depicted in Figure 3.3. In the cruising mode, UAV is assumed

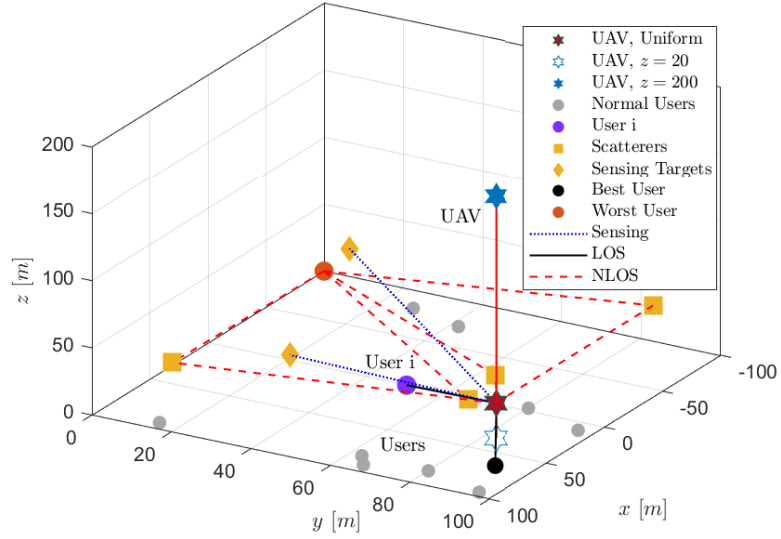


Figure 3.3: One realization of the 3D random scene with locations of UAV, users, scatterers, targets, and users in extreme scenarios.

to have uniformly distributed variations from the exact altitude; while if it is hovering over a fixed location, the UAV can reach to a higher/lower altitudes. We assume an urban environment, unless otherwise stated, and the UAV already has knowledge about the users' locations, target of interest and scatterer in the environment when assessing the system's performance.

In our experiments, we use the MUSIC algorithm in Figure 3.4 to estimate the angle of arrival at the UAV using the UP and reflected DP. In this case, we are assuming direct LoS with all sensing targets and users.

In the single-user scenario, we investigate the effect of active sensing and LoS condition on the quality of communication for a single user. According to (3.20), P_T is divided by the number of RF chains, N_{RF} , which yields number of communication paths, N_P , to the i th user, neighboring users, L , and sensing targets, T . As the UAV is actively sensing for targets of interest, it allocates power and dedicated beams towards

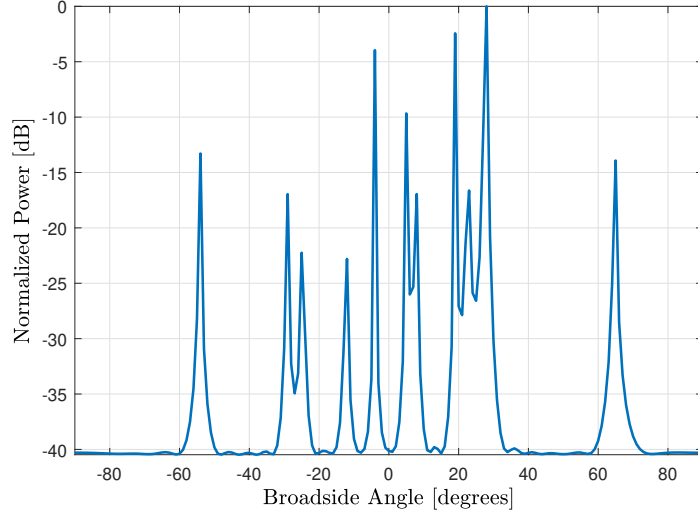


Figure 3.4: Angle estimation at the UAV.

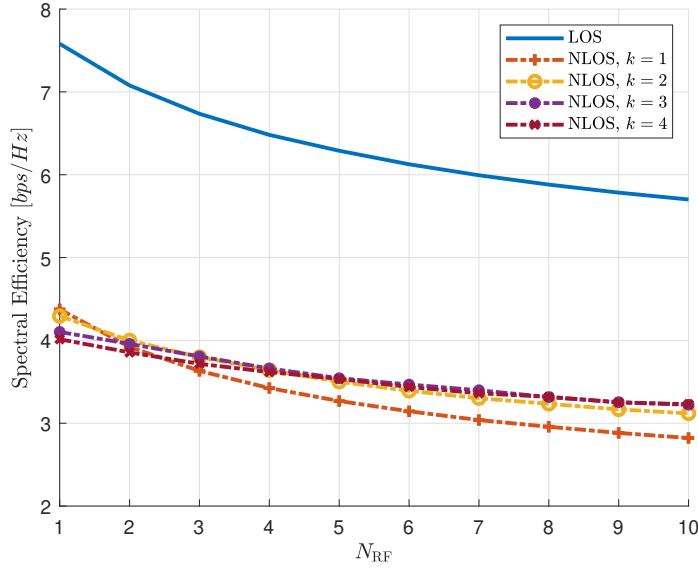


Figure 3.5: Spectral efficiency comparison for a single-user under LoS and NLoS conditions with various number scatterers and increasing N_{RF} .

those targets, which require a dedicated RF chain and N_{RF} increases. If P_T is divided by a larger N_{RF} , the power dedicated for the i th user decreases, which affects the spectral efficiency as clearly shown in Figure 3.5. It also shows the effect of additional N_{RF} on the spectral efficiency for a single-user, where the additional N_{RF} can be used

to communicate with different users or to sense more targets. As N_{RF} increases, the spectral efficiency decreases where the number of known scatterers, i.e. $k = 1, \dots, 4$, increases as well. We observe a slight performance enhancement for $N_{\text{RF}} < 2$ but for any addition/increase in N_{RF} it will cause the dedicated transmission power per path towards a user to be decreasing, thus resulting in a diminishing and deterioration in the spectral efficiency.

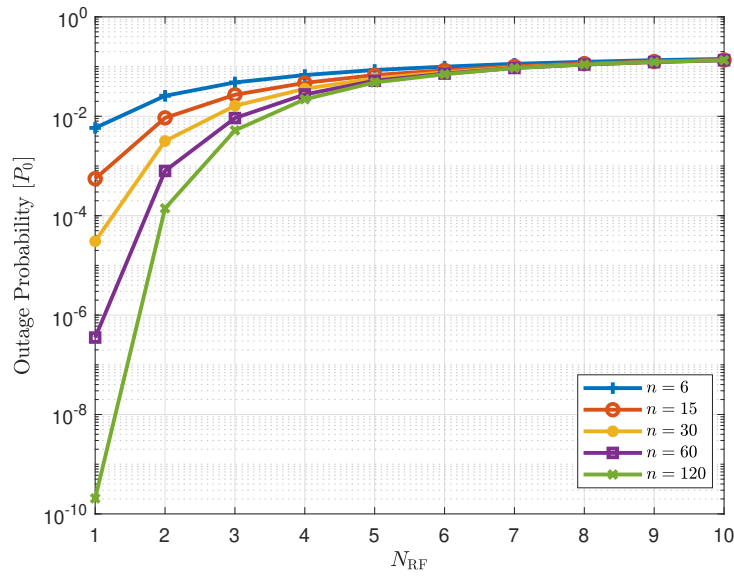


Figure 3.6: Information-outage probability for single-user with codewords lengths: 6, 15, 30, 60, and 120.

In Figure 3.6, the information outage probability is evaluated across multiple codeword lengths for a single user against additional N_{RF} . The results are shown as a function of N_{RF} over multiple codeword lengths. In the case of single-user communications, i.e., $N_{\text{RF}} = 1$, the codeword length can greatly affect the outage probability, which is not as severe if N_{RF} has been increased.

In another scenario, we investigate users' extreme locations with respect to the UAV considering the sensing activity, and antenna elements at users. On one hand, a

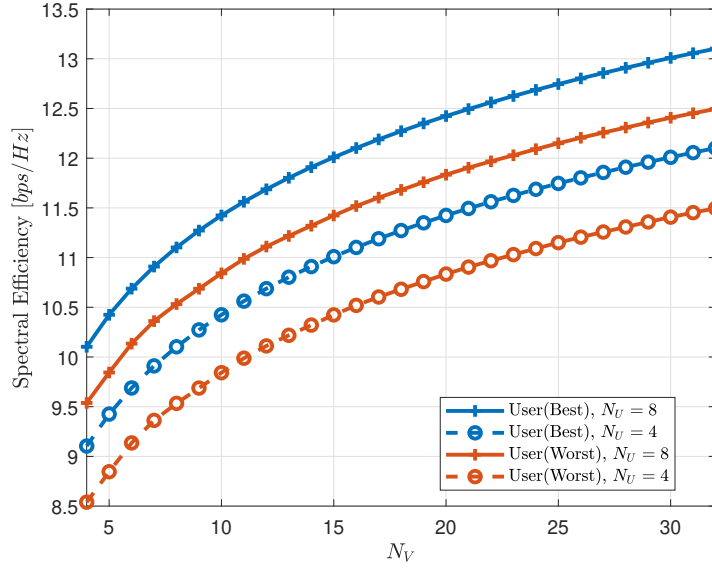


Figure 3.7: Spectral efficiency for best and worst user location with antenna elements $N_U = 8$ and 4.

user is positioned close to the UAV with a LoS condition, and no targets of interest to actively sense. On the other hand, a user is placed in NLoS conditions and far away from UAV, while the UAV actively senses for targets of interest which results in more power consumption. Figure 3.7 compares the spectral efficiency as the UAV antenna elements, N_V , increases. It illustrates the effect of the positioning of the UAV, the power allocation expressed by N_{RF} and N_U on the quality of service at the users.

So far, we have noticed that the UAV position is critical to reduce the over-sharing of power. It is also important to reduce the range between the UAV and the users, which will reduce the large-scale fading impact. In this case, we investigate the effect of UAV height (in the range from 20 to 200 m) on the quality of service while being stationary in the same (x, y) plane. From Figure 3.8, as the UAV initially ascends, the average spectral efficiency improves due to improved probability of LoS. However, this improvement is hindered by the effect of increasing the range/distance to the user

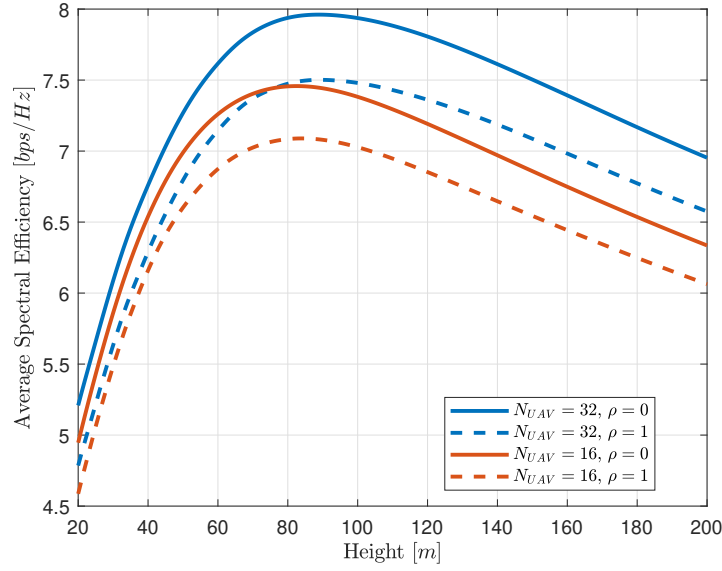


Figure 3.8: Average spectral efficiency for various UAV heights/altitudes with $N_V = 32$ and 16.

which increases the large-scale path loss effects.

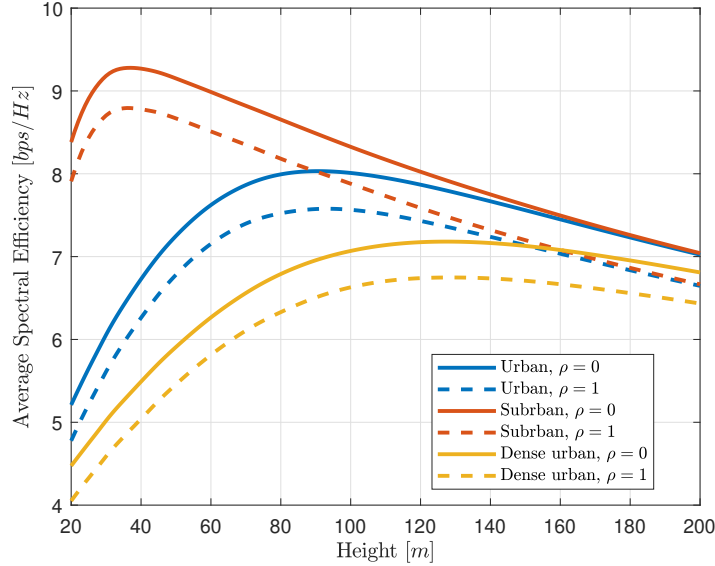


Figure 3.9: Average spectral efficiency for different UAV heights under various environments.

In Figure 3.9, we consider the *suburban*, *urban*, and *dense urban* environments and study the average spectral efficiency. The performance has an optimal UAV altitude

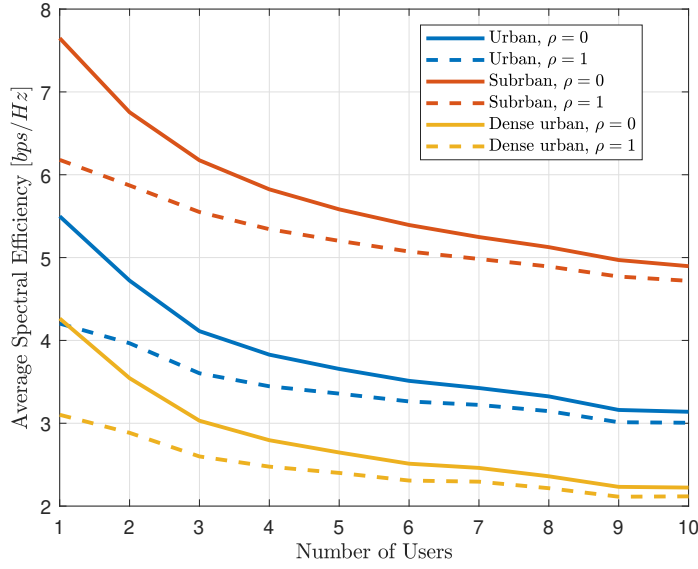


Figure 3.10: Average spectral efficiency as number of users increases in various environments.

at which the average spectral efficiency is maximum. These variations occur due to the changes in the probability of LoS as the UAV ascends, as well as the additional RF chains (by actively sensing targets). For multi-user communications, we emphasis on the effect of increasing RF chains by increasing number of users, and having the UAV sensing for targets under different environments. In Figure 3.10, we show the average spectral efficiency assuming all the users in the vicinity are communicating through one path, i.e., LoS/diffraction conditions, and compare it with activated sensing for T targets of interest. A degradation is observed in average spectral efficiency as the N_{RF} increases, which coincides with our earlier observations in single-user case.

Since the total transmission power is equally shared among all N_{RF} RF chains, any additional activation of RF chains to assist in communication tasks will divert power to the N_{RF}^c RF chains and reduce the power dedicated to the N_{RF}^s RF chains. In Figure 3.11, we assess the detection probability, P_D , of a single target to illustrate

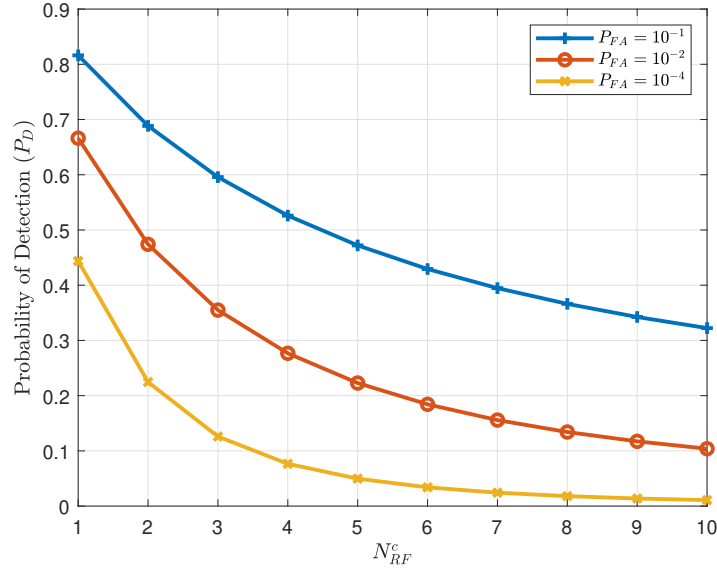


Figure 3.11: Detection probability of one target as the number of communication tasks increases.

the trade-off effect of added communication tasks. We evaluate P_D with constant probabilities of false alarm (P_{FA}), specifically, 10^{-1} , 10^{-2} , and 10^{-4} . As more RF chains are allocated to communication tasks, more power resources are dedicated to communication and less to target sensing. This power distribution can hinder the detection probability, with varying intensity depending on the P_{FA} .

To illustrate the relationship between the localization performance and the number of communication tasks, we compare the CRLB with the number of communication tasks N_{RF}^c in Figure 3.12. In a scenario where the UAV is hovering at $\mathbf{v} = [50, 50, 50]^T$ while the sensed target is positioned at $\mathbf{t} = [50, -70, 35]^T$ moving at $\dot{\mathbf{t}} = [-20, 20, 0]m/s$. The CRLB is shown for three parameters of interest: distance to the target d_{ut} , the velocity of the target v_t , the AoA of the target, namely the azimuth θ_t . The sensing system is configured with $Q = 10$ sensing pulses per time slot. As N_{RF}^c increases, the CRLB values for all three parameters exhibit a gradual

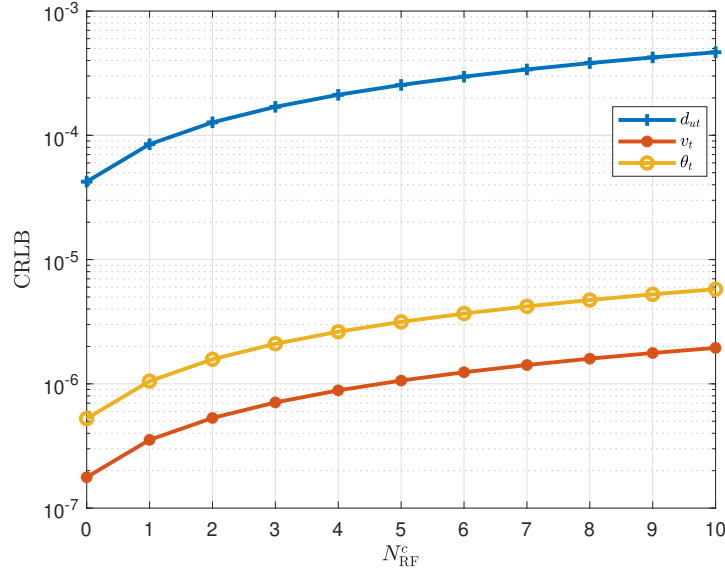


Figure 3.12: CRLB for distance, velocity, and AoA as the number of communication tasks increases.

rise, indicating a decrease in the estimation accuracy. This trend occurs due to the decrease of transmission power dedicated for that target. This observation further illustrates the communication and sensing trade-off in UAV-enabled ISAC systems. Figure 3.13 illustrates the increasing of number of sensing pulses Q effect on the CRLB for distance, velocity and AoA.

Figure 3.14 shows a scenario where the UAV is two moving targets, $\xi_1 = [50, -70, 35, -20, 20, 0]$ and $\xi_2 = [80, 80, 35, 0, -20, 0]$. Under the assumption of L users are actively communicating with the UAV, meaning $N_{\text{RF}}^c = L$ and $N_{\text{RF}}^s = T$, the UAV is equipped with URA, the number of pulses Q is 50, and tracking results are picked every Δ_T . The figure shows the effectiveness of the proposed EKF tracking process at a large number of N_{RF} activated. In Figure 3.15, We show the CRLB for the targets as they move through the environment. By assuming the UAV is equipped with URA antenna, we can readily solve the CRLB for the elevation angles for both

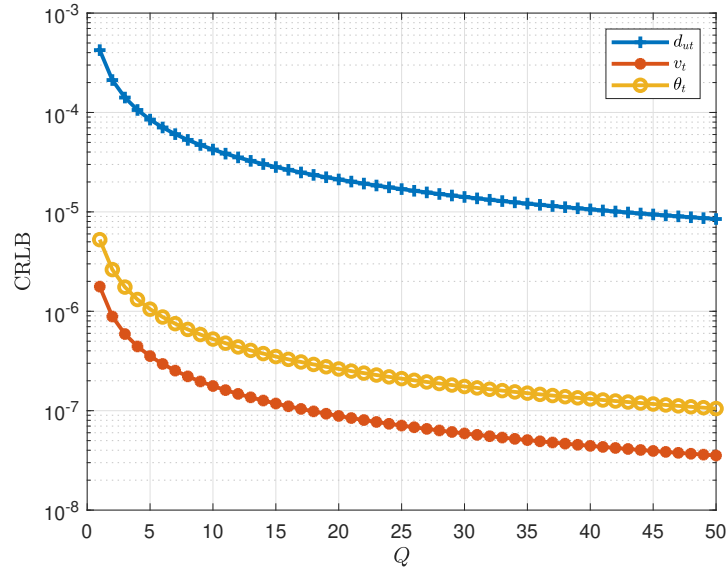


Figure 3.13: CRLB for distance, velocity, and AoA as the number of sensing pulses increases.

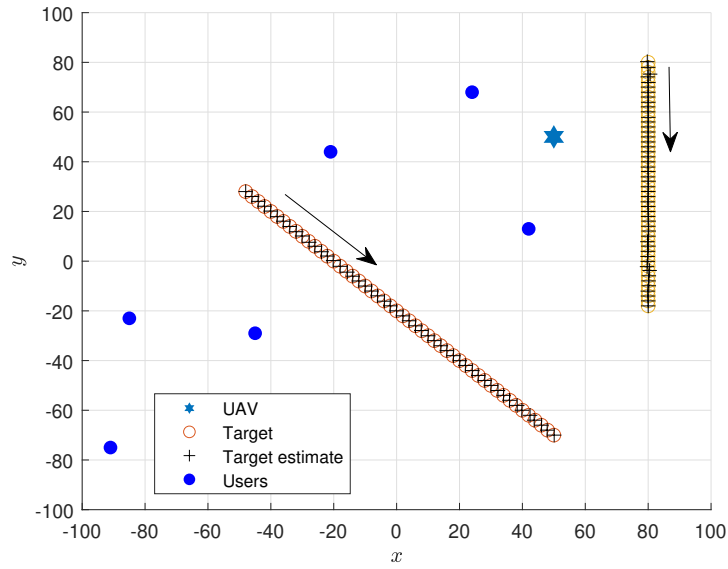


Figure 3.14: Illustration of UAV, users, and moving targets and their tracking using EKF.

targets.

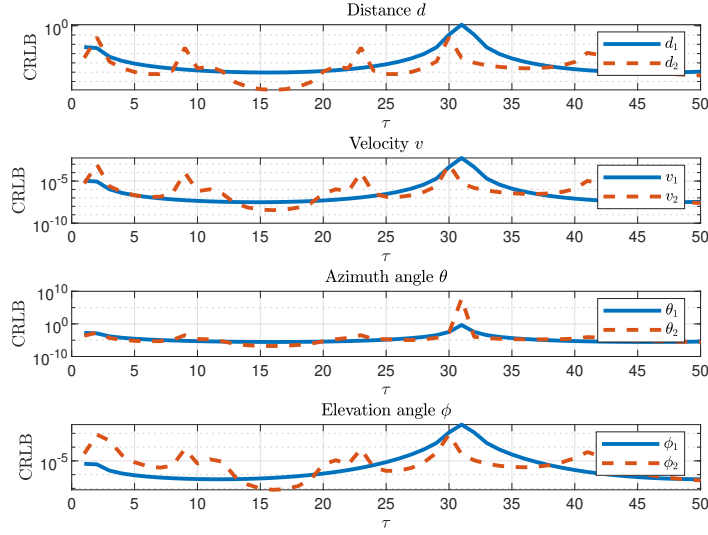


Figure 3.15: CRLB for distance, velocity, and AoA calculated for the tracking process.

3.7 Summary

In this chapter, we proposed a low-complexity hybrid beamforming UAV-enable ISAC system enabling the UAV to simultaneously sense targets of interest and provide multi-user communications. We also proposed an accurate large-scale path loss model for the downlink communications taking into consideration reflections and diffractions. We tested our proposed system against a variety of scenarios and demonstrated the effect of UAV positioning and the interference caused by UAV on users. We have also assessed the spectral efficiency, average spectral efficiency and information outage probability performance considering practical finite-length codewords. Finally, we investigated in depth the effect of the number of RF chains and its impact on power distribution, and thus the quality of service at the user(s) and detection probability. In future work, sensing and communication functionalities can be extending to include cooperative multi-UAV based on federated learning and game theory.

CHAPTER 4

GENERALIZED GEOMETRIC MODEL FOR PROBABILISTIC LOS IN IRREGULAR URBAN AREAS

The accurate modeling of path loss is crucial for optimizing performance in any aerial wireless communication systems where carrier frequency, range, and probability of LoS, among other parameters play a significant role [59]. The distance and carrier frequency are typically predetermined by the application, the LoS probability becomes the primary controllable factor, which is directly influenced by the UAV's location. Several of the existing probability of LoS models are based on the standard Manhattan grid shape of the area in which the probability is calculated, and do not address the deviation caused by the change in azimuth angle or the irregularity of the environment.

This chapter proposes a general geometrical model to calculate the probability of LoS for three dimensional environments. The model considers the shape of the environment based on the International Telecommunication Union (ITU) recommendation P.1410 [60] for characterizing built-up environments, and the characteristics of the non built-up areas in which the ground user is located, as well as the azimuth angle between the UAV and the ground user.

4.1 Background

Several models for estimating LoS probability have been proposed in the literature, broadly categorized into deterministic, empirical, and geometrical approaches [61]. Ray Tracing (RT) is widely employed in deterministic models for its ability to accurately measure the probability of LoS by exhaustively computing the interactions between signals and the environment [62], [63], [64]. While RT models provide high accuracy, they are computationally expensive, making their use in UAV systems impractical as it can significantly hinder overall performance [65].

The Third-Generation Partnership Project (3GPP) model has also been used to propose LoS probability estimations, as demonstrated in [66]. However, empirical models, often based on real-world measurements or simulations, provide an alternative. For instance, in [67], the authors conducted multiple trials in the city of Prague to propose a propagation model. This model incorporated LoS probability and additional shadowing path loss for NLoS conditions as functions of ITU parameters and elevation angle [60]. Similarly, Al-Hourani et al. [68] analytically evaluated LoS probability

using ITU built-up area parameters to optimize coverage for low-altitude platforms (LAPs). Their model demonstrated a direct correlation between elevation angle and path loss. Mohammed et al. [69] extended this approach by presenting a closed-form formula to predict LoS probability for UAV communication as a function of elevation angle. While these empirical models provide accurate representations of LoS probability, they can be time-intensive due to the need for field measurements or long Monte Carlo simulations.

Geometrical models offer a more versatile and adaptive approach by leveraging environmental geometrical data to predict LoS probability. For example, in [70], the authors proposed a stochastic model that considers geographical information and the Fresnel zone to estimate blockage probability in various built-up areas. Another study utilized a stochastic approach to characterize LoS probability and trained a backpropagation neural network (BPNN) using this data [71]. Additionally, in [66], the authors analyzed LoS probability for urban grid scenarios, evaluating base station-to-UAV communication in mmWave frequency bands. This study examined the impact of different building height distributions, including uniform, exponential, and Rayleigh distributions, on LoS probability. Most of the already previously proposed models have measured the LoS probability in specific environments or areas in terms of the elevation angle or the ground distance, they failed to take into consideration the fluctuation that is induced by the change in azimuth angle on the LoS probability or the irregularity of the environment. At the time of writing this thesis, one publication have considered fluctuation induced by change in azimuth angles. Saboor et al. have proposed a model that considers the azimuth angle when calculating the LoS probability [61], however

it and the previously mentioned publication only consider the standard Manhattan grid-like shape. In reality, most of the major cities around the world do not have a regular shapes, furthermore, these areas may have more than two streets intersections and include different streets and buildings widths. Motivated by this, we propose a generalized model of characterizing the LoS probability that takes into consideration the azimuth angle, multiple streets intersections with different characterizations. We consider built-up standard parameters provided in the following table

Table 4.1: Parameters of the ITU-R P.1410 Model for Selected Environments [1]

Environment	ε	ϱ	ϑ
Suburban	0.1	750	8
Urban	0.3	500	15
Dense Urban	0.5	300	20
High-rise Urban	0.5	300	50

According to the ITU-R P.1410 recommendation for propagation prediction methods, these selected environments can be characterized by the following parameters:

- ε : represents the ratio of land covered by buildings.
- ϱ : represents the average number of buildings per unit area (in $1km$).
- ϑ : which is a parameter of the Rayleigh distribution for building height variability.

The rest of this chapter is structured as follows: Section 4.2 presents our proposed model, followed by numerical results in Section 4.3, and a summary of the chapter in Section 4.4.

4.2 Geometric LoS Probability

In this section, we expand upon Saboor et al. proposed model [61], to include the azimuth angle (θ) for each street in the area, and the ability to calculate LoS probability of the area (P_{LOS}) considering the different characteristics of streets, in terms of widths (S), orientation (θ_r). Assume a 3D built-up environment where a user is located along S following the uniform distribution $\mathcal{U}[0, S]$ and communicating with a flying UAV. For a typical grid environment, the street width and buildings width (W) -assuming equidistantly spaced- are

$$W = 1000\sqrt{\frac{\varepsilon}{\varrho}}, \quad (4.1)$$

$$S = \frac{1000}{\sqrt{\varrho}} - W. \quad (4.2)$$

However, it is possible to have streets or buildings that are different in width or angle of alignment within an area. Therefore, we divide such area into distinctive regions where the user could be located, and calculate the LoS probability for each. A region r is defined as an area where a possible ground user is located. Between the user and the UAV is an obstructing building that its height follows Rayleigh distribution where it is probable for the user's communication to be blocked by such building given a certain UAV height (h_{UAV}). Figure 4.1 gives an example for a built-up grid environment divided into regions, $R_1, \dots, R_4$, and an intersection region R_i between them. Each region is aligned with angle θ_r , where it affects the line in which the user is assumed to be located depending on the azimuth angle of the user. For a

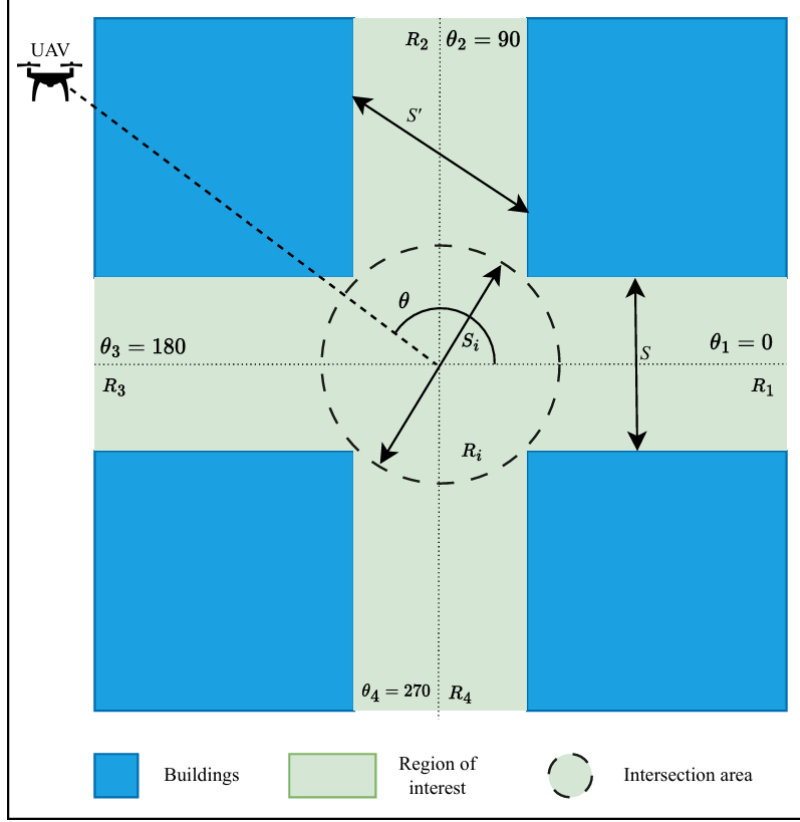


Figure 4.1: An environment that contain a typical grid layout (R1 and R2) and additional street (R3).

certain elevation angle (ϕ), the maximum height of the first obstructing building that will block LoS across the entire r th region is said to be $h_{\max_1}^r(\phi) = S_r \tan(\phi)$, where S_r is the region's width. This is true assuming the line between the UAV and user is perpendicular to the r th region, otherwise however, the region and building widths will differ. Thus $h_{\max_1}^r$ is

$$h_{\max_1}^r = S'_r \tan(\phi), \quad (4.3)$$

where

$$S'_r = \frac{S_r}{|\sin(\theta - \theta_r)|}, \quad (4.4)$$

In a built-up area, it is possible for the user to be obstructed by multiple build-

ings. Unless some information about the environment is provided (i.e., the street and building widths between the UAV and user), the number of possible buildings that can obstruct the LoS communication -assuming equidistantly spaced- is

$$N = \left\lfloor \frac{h_{\text{UAV}}}{(S' + W') \tan(\phi)} \right\rfloor, \quad (4.5)$$

$$W' = \frac{W}{|\sin(\theta - \theta_r)|} \quad (4.6)$$

$\lfloor \cdot \rfloor$ represents the floor function, ensuring that the calculated number of buildings is rounded down to the nearest integer. For the subsequent obstructing buildings, only part of such buildings are able to block the transmission. Therefore, for a certain ϕ , the minimum and maximum heights of the n th building that will block the transmission is

$$h_{\min_n}^r = ((n - 1) \times (S' + W')) \tan(\phi), \quad (4.7)$$

$$h_{\max_n}^r = ((n - 1) \times (S' + W') + S') \tan(\phi). \quad (4.8)$$

The LoS probability for the the first obstructing building in for a user located in the r th region is [61]

$$P_{\text{LoS}_1}^r(\phi, \theta, \theta_r) = 1 - \sqrt{\frac{\pi}{2}} \frac{\vartheta}{h_{\max_1}^r} \times \left(\text{erf} \left(\frac{h_{\max_1}^r}{\sqrt{2}\vartheta} \right) \right), \quad (4.9)$$

while the LoS probability for the n th obstructing building is

$$P_{\text{LoS}_n}^r(\phi, \theta, \theta_r) = 1 - \sqrt{\frac{\pi}{2}} \frac{\vartheta}{h_{\text{max}_1}^r} \times \left(\text{erf}\left(\frac{h_{\text{max}_n}^r}{\sqrt{2}\vartheta}\right) - \text{erf}\left(\frac{h_{\text{min}_n}^r}{\sqrt{2}\vartheta}\right) \right). \quad (4.10)$$

As such, the LoS probability P_{LoS}^r for a user located in the r th region is

$$P_{\text{LoS}}^r(\phi, \theta, \theta_r) = \prod_{n=1}^N P_{\text{LoS}_n}^r(\phi, \theta, \theta_r). \quad (4.11)$$

When multiple regions are oriented at various angles, the shape of their intersection will inevitably vary as these angles change, thus, complicating the calculation of such intersection effective width. Consequently, we assume a circular shaped intersection to generalize the model. To account for the irregularity of the intersection area with varying numbers of legs and their orientations, we approximate the area as a circular region with a radius of $S/2$. With the possibility of a user to be located within an intersection region, it is possible for the transmission to have a full LoS at a certain θ for all ϕ . Therefore, the previously mentioned expression of LoS probability will not correctly represent the that of an intersection region. For a certain θ , the first possible obstructing building is most likely to be adjacent to the normal region which its θ_r is close to θ . In certain instances, when the user's θ is close to or aligned with the orientation a region, the LoS probability can be maximized. As such, in Algorithm 1 we propose a simple algorithm to calculate the LoS probability of such an intersection where $\mathcal{R}$ is a set containing R regions, each with its own orientation angle θ_r .

Algorithm 1 Probability of LoS for Multi-leg Intersection

```
1:  $r \leftarrow 1$ 
2: for  $r \in \mathcal{R}$  do
3:    $a \leftarrow \frac{\theta_{r-1} + \theta_r}{2}$ 
4:    $b \leftarrow \frac{\theta_{r+1} + \theta_r}{2}$ 
5:   if  $a < \theta \leq b$  then
6:     Calculate:  $P_{\text{LoS}}^*(\phi, \theta, \theta_r)$ 
7:   else
8:      $r \leftarrow r + 1$ 
9:   end if
10: end for
```

As such, the LoS probability for the entire area A can be calculated by taking the weighted average of all regions

$$P_{\text{LoS}}(\phi, \theta) = \frac{A^*}{A} P_{\text{LoS}}^*(\phi, \theta, \theta_r) + \frac{1}{A} \sum_{r=1}^R A^r P_{\text{LoS}}^r(\phi, \theta, \theta_r) \quad (4.12)$$

where $P_{\text{LoS}}^*(\phi, \theta, \theta_r)$ is the probability of LoS for the multi-leg intersection region, A^* and A^r are the area of the multi-leg intersection region and the r th region respectively. In order to evaluate P_{LoS} for a specific ϕ , we average over a range of azimuth angles $[0, M]$

$$P_{\text{LoS}}(\phi) = \frac{1}{M} \int_0^M P_{\text{LoS}}(\phi, \theta) d\theta. \quad (4.13)$$

4.3 Numerical Results

In this section, the proposed LoS probability characterization for normal and intersection regions is evaluated for urban, suburban, dense urban, and high-rise built-up environments. In order to ensure fair comparison between Al-Hourani et al. model [68], Saboor et al. model [61], and the proposed model, we evaluate models in grid

based urban environment against elevation angles $[0^\circ, 90^\circ]$ by limiting the integration in (4.13) to $M = \pi/4$ in Figure 4.2, assuming all regions have equal width of S , and a square shaped intersection area of $A^* = S^2$. The figure shows the distinctive difference between geometric based models and analytical models, which mostly relay only on elevation angels. The figure also show the resemblance between the proposed model and Saboor et al. model [61], where the difference is in result is the way the effective region width S'_r is evaluated.

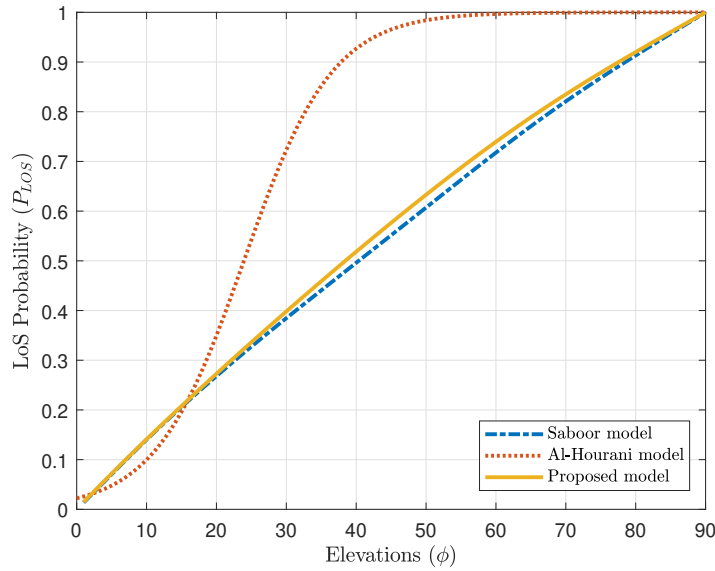


Figure 4.2: Comparison between different models in urban environment, $M = \frac{\pi}{4}$

In Figure 4.3, under the assumption of equal widths for all regions, and a circular intersection region of width $S/2$, we compare different built-up environments where a multi-leg intersection region is surrounded by three regions in irregular shape, which are oriented in 0° , 120° , and 270° degrees. The figure shows the effect of region width on the proposed model when evaluated over the entire 360° .

To clearly show the effect of the number of leg in an intersection on the LoS

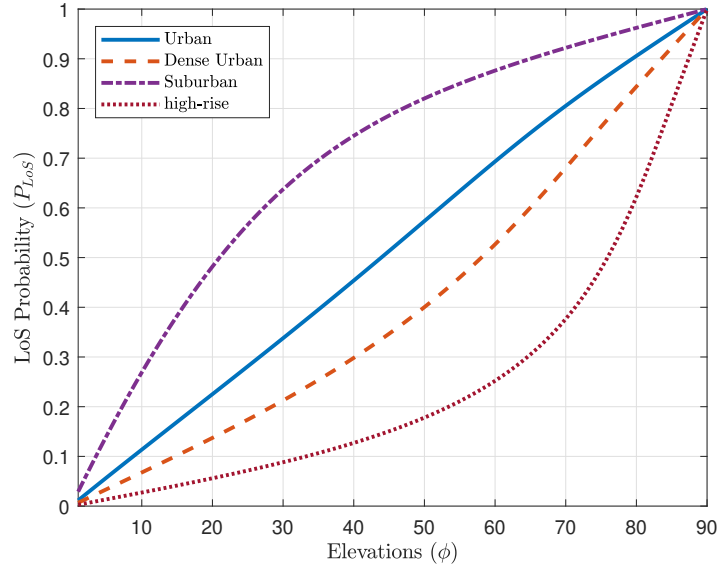


Figure 4.3: Comparison between different built-up environment (urban, suburban and dense urban, $M = 2\pi$

probability, Figure 4.5 compares P_{LoS}^* by increasing the number of intersection legs with the configuration illustrated in Figure 4.4, starting with 2 regions in 90° between them, to 8 legs with 45° between them, and evaluated at $\phi = 45^\circ$. We set the following orientation angles for

- Two regions: $\{0^\circ, 270^\circ\}$
- Three regions: $\{0^\circ, 90^\circ, 270^\circ\}$
- Four regions: $\{0^\circ, 90^\circ, 180^\circ, 270^\circ\}$
- Five regions: $\{0^\circ, 90^\circ, 135^\circ, 180^\circ, 270^\circ\}$
- Six regions: $\{0^\circ, 90^\circ, 135^\circ, 180^\circ, 270^\circ, 315^\circ\}$
- Seven regions: $\{0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 270^\circ, 315^\circ\}$
- Eight regions: $\{0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ, 315^\circ\}$

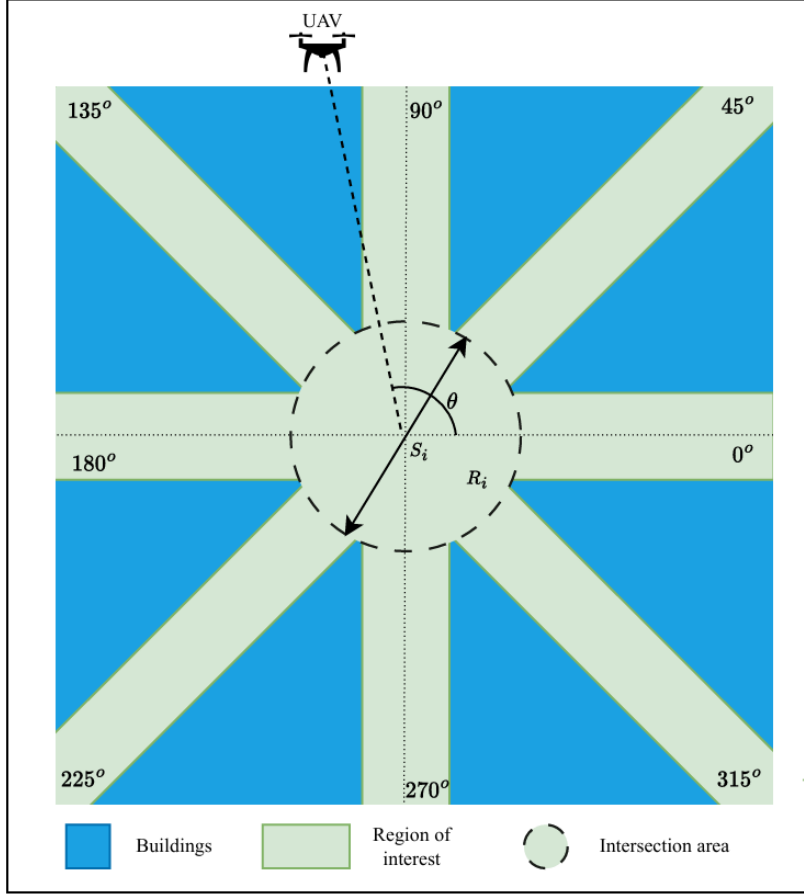


Figure 4.4: Illustration of multi-legged intersection

As the number of legs increases, its less probable for an obstructing building to block LoS communication for a user located within the intersection region. Another insight taken from Figure 4.5, is that increasing the number of legs has higher effect in denser environments than less dense environment. For example, the increment in P_{LoS}^* for 5 legs intersection to 4 legs is 0.018, 0.04, 0.052 and 0.045 for suburban, urban, dense urban and high-rise respectively. In highly dense environments, the potential height of obstructing buildings (ϑ), is generally greater than in less dense areas. This shows variations in LoS probability when azimuth angles change, particularly in dense, high-rise settings. Consequently, it highlights not only the influence of the surrounding

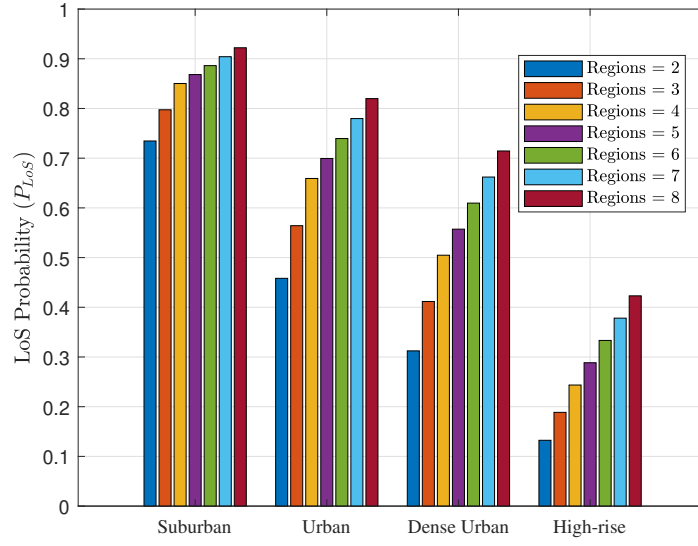


Figure 4.5: Effect of number of legs in an intersection region on P_{LoS}^* , $M = 2\pi$ environment but also the impact of the number of intersection region segments on the LoS probability.

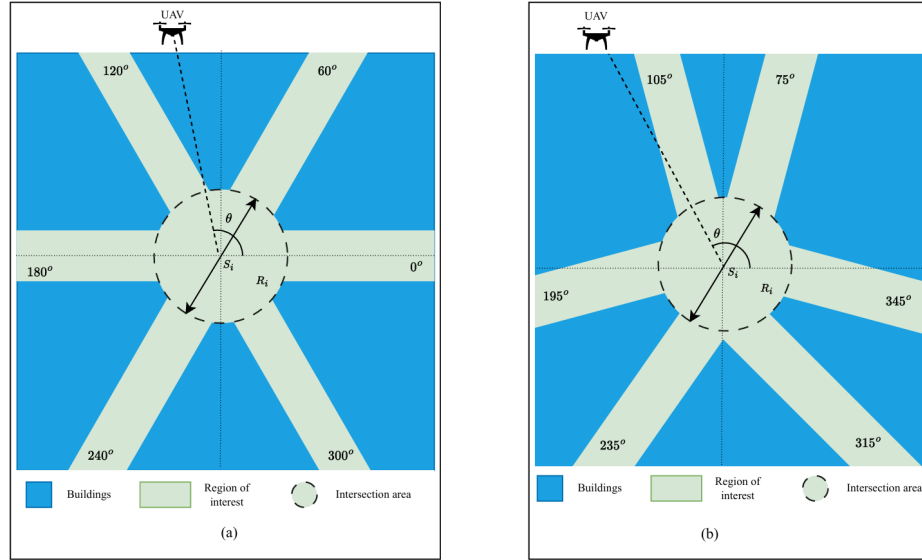


Figure 4.6: Illustration of an environment with 6 regions in (a) equal distance apart, and (b) non equal distance apart

The distance between intersection legs can also influence the LoS probability. Figure 4.7 presents P_{LoS}^* on different built-up environments, comparing cases where six

intersection legs are either equidistant or non-equidistant with configuration as shown in Figure 4.6. The figure demonstrates that P_{LoS}^* of equidistant legs are in general higher than non-equidistant legs, which is attributed to the probability of blockage being greater when azimuth angle vary from 0 to 2π on average.

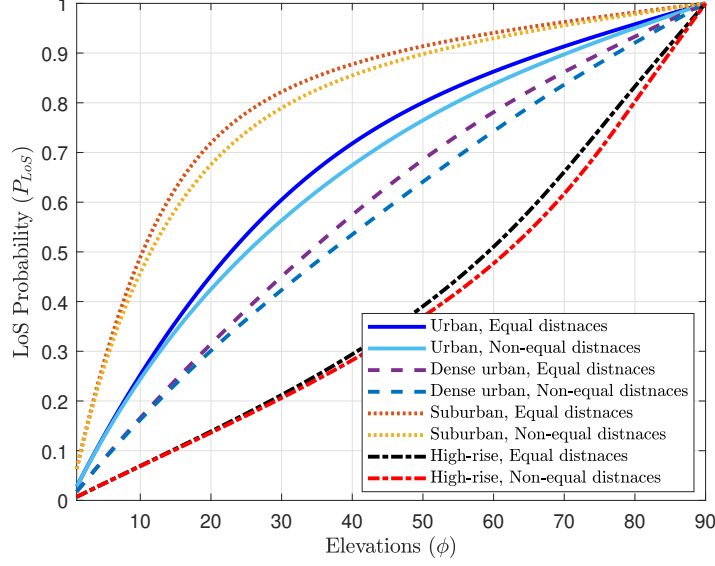


Figure 4.7: Distance between intersection legs effect on P_{LoS}^* , $M = 2\pi$

4.4 Summary

In this chapter, we proposed a geometrical model to characterize the P_{LOS} for regular and irregular built-up environments. The model considers any shape or configuration of regions where a user is located, and captures the impact of different regions widths and orientations, and the azimuth angle. We presented numerical results to show the models effect considering varies regions configurations and their intersection in different environments.

CHAPTER 5

REINFORCEMENT LEARNING IN MULTI-USER MULTI-TARGET UAV-ENABLED ISAC SYSTEM

The performance of UAV-enabled ISAC system, discussed earlier, depends on the UAV trajectory and resource allocation, which have a significant impact on the system. For instance, ensuring optimal positioning not only reduces energy consumption under communication and sensing requirements but also enhances system adaptability in dynamic environments. In this chapter, we propose a low-complexity DRL model to optimize the UAV trajectory. Specifically, we employ the Proximal Policy Optimization (PPO) algorithm, which is well-suited for balancing exploration and exploitation in dynamic environments. The proposed model leverages PPO's algorithm ability to

learn from environmental interactions, enabling the UAV to adapt its trajectory to maximize communication and sensing performance within the constraints. We summarize the main contributions in this chapter as follows.

- Proposing a low-complexity DRL-based model for UAV-enabled ISAC system using the PPO algorithm to plan UAV trajectories and enhance the overall system performance.
- Optimizing the hyperparameters for the DRL-based model using the Bayesian optimization method.
- Providing a comparative analysis of the agent or UAV behavior under varying probabilities of line-of-sight and diverse built-up environment conditions.

In this chapter, we first present the system model and formulate the problem of UAV trajectory optimization to maximize system performance in section 5.1. In Section 5.2, we introduce the PPO-based reinforcement learning model and discuss its role in optimizing the UAV trajectory according to the specified criteria. Furthermore, we discuss the adoption of the Bayesian optimization method for hyperparameter optimization and fine-tuning. Section 5.3 presents the training and numerical results, our observations in the comparative analysis. Finally, we summarize the chapter's findings in Section 5.4.

5.1 System Model and Problem Formulation

5.1.1 System Model

In this subsection, we consider the UAV-enabled ISAC system that operates in the environment shown in Figure 5.1. It consists of a UAV located at time slot τ in $\mathbf{v}[\tau] = [x_v, y_v, z_v]^T \in \mathbb{R}^{3 \times 1}$, $\tau = 0, \dots, N_T - 1$, within the 3D physical environment boundaries, $\mathcal{E}$. The total UAV flight duration is $T_f = N_T \Delta_T$ with a fixed time slot duration of Δ_T . We assume that the environment conditions are unchanged if Δ_T is

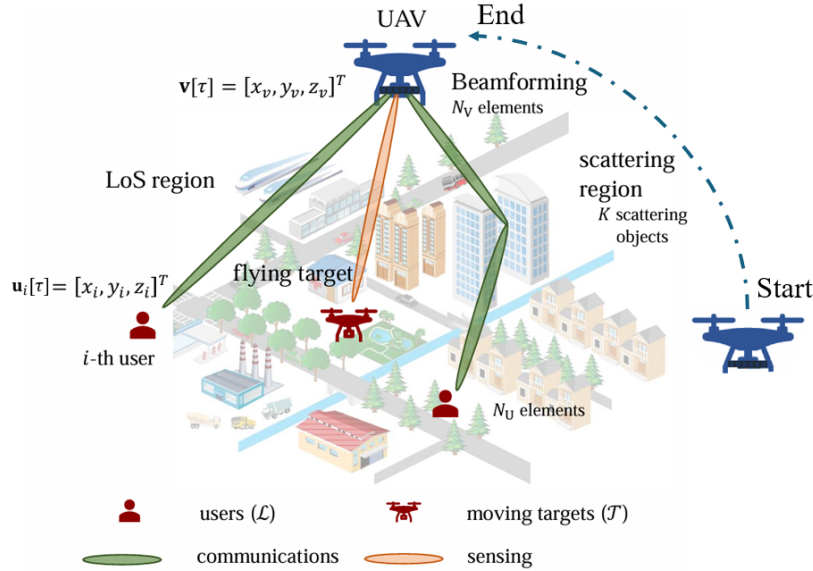


Figure 5.1: Illustration of UAV-enabled ISAC system with a trajectory.

sufficiently small. Similar to Chapter 3, we consider now a dynamic, not necessarily stationary, where the UAV is equipped with $N_{\text{RF}_{\max}}$ RF chains to serve a set of L ground users, and sense a set of T targets of interest denoted by $\mathcal{U} = \{\mathbf{u}_1[\tau], \dots, \mathbf{u}_L[\tau]\}$ and $\mathcal{T} = \{\mathbf{t}_1[\tau], \dots, \mathbf{t}_T[\tau]\}$ respectively. Furthermore, we assume that a set of K possible scattering objects denoted by $\mathcal{K} = \{\mathbf{k}_1, \dots, \mathbf{k}_K\}$ are within the vicinity of the UAV, and can change as the UAV changes its position.

The physical environment, $\mathcal{E}$, is defined by

$$\mathcal{E} = \{[x, y, z]^T \in \mathbb{R}^{3 \times 1} \mid x \in [x_{\min}, x_{\max}], y \in [y_{\min}, y_{\max}], z \in [z_{\min}, z_{\max}]\}. \quad (5.1)$$

where the sets for users, targets and scattering objects are assumed to be within the 3D physical scene, i.e. $\mathcal{U}$, $\mathcal{T}$, and $\mathcal{K}$ are proper subsets of $\mathcal{E}$.

The UAV is equipped with a uniform rectangular array (URA) consisting of N_v elements in total. Such arrangement enables an accurate estimation of AoA in two dimensions. While not all RF chains are simultaneously activated as discussed in Chapter 3, the number of activated RF chains, N_{RF} , depends on the communication and sensing tasks, such that $N_{\text{RF}} \leq N_{\text{RF}}^c + N_{\text{RF}}^s$, where N_{RF}^c and N_{RF}^s denote the RF chains dedicated to communication and sensing respectively. For the feasibility of the study, we assume that $L + T \leq N_{\text{RF}_{\max}}$. Each user is equipped with a ULA of size N_U , and a single RF chain. Furthermore, we assume all users are associated with the UAV and communicating only with it¹. By employing the same hybrid beamforming model proposed in Chapter 3, the received signal at the i th user can be expressed as

$$\begin{aligned} \mathbf{y}_i[\tau] = & \mathbf{W}_i[\tau] \mathbf{H}_i[\tau] \mathbf{F}_{\text{RF}}[\tau] \mathbf{F}_{\text{BB}_1}[\tau] \mathbf{x}_i[\tau] + \mathbf{W}_i[\tau] \mathbf{H}_i[\tau] \mathbf{F}_{\text{RF}}[\tau] \mathbf{F}_{\text{BB}_2}[\tau] \mathbf{x}_l[\tau] \\ & + \mathbf{W}_i[\tau] \mathbf{H}_i[\tau] \mathbf{F}_{\text{RF}}[\tau] \mathbf{F}_{\text{BB}_3}[\tau] \mathbf{x}_t[\tau] + \mathbf{W}_i[\tau] \mathbf{n}_i[\tau] \end{aligned} \quad (5.2)$$

Thus, the average spectral efficiency for the i th user at time slot τ can be formulated

¹We assume that all users are within the range of the UAV, therefore, associated with that UAV. User association problem will be tackled in future studies.

similar to the one in Chapter 3, i.e.

$$SE_i[\tau] = \log_2 \det \left(\mathbf{I}_{N_P} + \frac{1}{N_P} \mathbf{R}_i'[\tau] \mathbf{R}_{in}'^{-1}[\tau] \right). \quad (5.3)$$

5.1.2 Problem Formulation

Our objective in this section is to determine the optimal UAV trajectory that maximizes the average spectral efficiency for ground users subject to constraints in communications, sensing, and physical boundaries. At $\tau = 0$, the UAV is initially assumed to be on the ground, then flies in a trajectory towards a position that maximizes the performance in terms of spectral efficiency. Therefore, the UAV trajectory optimization problem at time slot τ can be written as

$$(P1) \quad \hat{\mathbf{v}} = \arg \max_{\mathbf{v}[\tau] \in \mathcal{E}} \frac{1}{L} \sum_{i=1}^L SE_i[\tau] \quad (5.4a)$$

$$\text{s.t.} \quad N_{\text{RF}}^c + N_{\text{RF}}^s \leq N_{\text{RFmax}}, \quad (5.4b)$$

$$\mathbf{v}[\tau] \in \mathcal{E}, \quad \forall \tau, \quad (5.4c)$$

$$\frac{\|\mathbf{v}[\tau] - \mathbf{v}[\tau - 1]\|}{\Delta_T} \leq \dot{v}_{\max} \quad \forall \tau, \quad (5.4d)$$

where $\dot{v}_{\max}$ is the maximum UAV velocity. Constraint (5.4b) ensures that dedicated RF chains for communications or sensing do not exceed the maximum available RF chains. Moreover, constraint (5.4c) guarantees the UAV's trajectory stays within the physical operational boundaries, $\mathcal{E}$, at all times, and (5.4d) is the UAV flight constraint.

The optimization problem (P1) is typically formulated as MDP, a framework widely used in reinforcement learning to model decision-making problems. Within this framework, the objective is to identify the optimal policy, which establishes a strategy for selecting the most effective action based on the environment's current state at each time slot. The PPO algorithm is employed to identify this optimal policy, leveraging its ability to balance exploration and exploitation effectively. In this framework, the UAV is the RL agent, which interacts with the environment to learn and adapt its trajectory and resource allocation decisions. This enables the UAV to progressively adjust its actions to enhance overall performance, driven by environmental feedback in the form of rewards.

It is worth noting that calculating the average spectral efficiency at each time slot not only hinders the training phase, but may provide the agent with extra information as feedback. Several aspects have already been optimized by the proposed digital beamforming matrix, $\mathbf{F}_{\text{BB}}$, analog combiner, $\mathbf{W}_i$, and the analog beamformer, $\mathbf{F}_{\text{RF}}$, as discussed in Chapter 3. The channel, $\mathbf{H}_i$, is the key component influencing the received signal, and we will consider the channel gain, $\beta_i[\tau]$, which heavily depends on the 3D geometry of the environment, to maximize $SE_i[\tau]$ for the i th user. Thus, we relax and simplify the optimization problem (P1) into

$$(P2) \quad \hat{\mathbf{v}} = \arg \max_{\mathbf{v}[\tau] \in \mathcal{E}} \frac{1}{L} \sum_{i=1}^L \beta_i[\tau], \quad (5.5)$$

where under LoS between the UAV and i th user, the average channel power gain is

$$\beta_i[\tau] = \sqrt{P_{\text{LoS}_i}[\tau]L_{\text{FS}}(d_{\mathbf{v}i}[\tau]) + (1 - P_{\text{LoS}_i}[\tau])(L_{\text{FS}}(d_{\mathbf{v}i}[\tau]) + L_s)^{-1}}, \quad (5.6)$$

where $d_{\mathbf{v}i}[\tau] = \|\mathbf{v}[\tau] - \mathbf{u}_i[\tau]\|$ is the distance from UAV to the i th user, $L_{\text{FS}}(d_{\mathbf{v}i}[\tau])$ is the free-space path loss, and L_s is the additional shadowing loss. Figure 5.2 illustrates

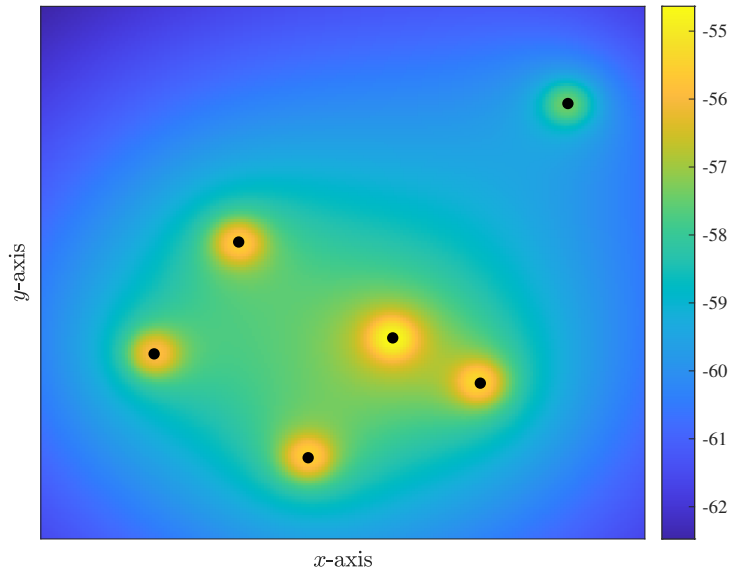


Figure 5.2: Average channel gain at an altitude of 10m (black circle: user location).

the average channel gain in the $x - y$ plane at a fixed altitude of 10m where 6 users are randomly located, the carrier frequency is 28GHz, and probability of LoS using [53] model. We observe local maximum near each user, which indicates the non-convexity of the optimization problem (P2) at lower altitudes. This introduces a challenge for the agent to escape local maxima, and find the optimal location serving all the users simultaneously. We can further split the problem into two sub-problems by considering the probability of LoS, $P_{\text{LoS}_i}[\tau]$, and the distance, $d_{\mathbf{v}i}[\tau]$, which are given

in optimization problems (P3) and (P4), respectively, i.e.

$$(P3) \quad \hat{\mathbf{v}} = \arg \max_{\mathbf{v}[\tau] \in \mathcal{E}} \frac{1}{L} \sum_{i=1}^L P_{\text{LoS}_i}[\tau], \quad (5.7)$$

$$(P4) \quad \hat{\mathbf{v}} = \arg \min_{\mathbf{v}[\tau] \in \mathcal{E}} \frac{1}{L} \sum_{i=1}^L d_{\text{vi}}[\tau]. \quad (5.8)$$

5.2 Methodology

In this section, we propose DRL and use the PPO algorithm to solve the optimization problem (P2) for UAV trajectory optimization. As discussed in Chapter 1, the RL environment consists of three main components, which are described as follows:

- **State space:** The state space, $\mathcal{S} = \{\mathbf{v}, \mathcal{U}\}$, denotes the UAV and all users' locations. Without loss of generality, the users' locations are assumed to be known at the UAV; thus, the state space for an agent at time slot τ will be

$$\mathbf{S}[\tau] = [\mathbf{v}[\tau], \mathbf{u}_1[\tau], \dots, \mathbf{u}_L[\tau]] \in \mathcal{S}, \mathbb{R}^{3 \times (L+1)} \quad (5.9)$$

- **Action space:** The agent takes an action, $\boldsymbol{\alpha}[\tau]$, based on knowing the state $\mathbf{S}[\tau]$ and the current policy π_ψ to move in 3D to the next position. The action space, $\mathcal{A}$, is defined by

$$\mathcal{A} = \{\boldsymbol{\alpha}_a = [x_a, y_a, z_a] \mid x_a, y_a, z_a \in \{-1, 0, 1\}\}, \quad (5.10)$$

which results in a total of 27 possible actions in x , y , and z directions.

- **Reward function:** The reward function, $R[\tau]$, is the accumulated reward based on the action taken by the agent. We propose an adaptive reward function based on (P2), (P3), and (P4), where it considers (P3) and (P4) at lower altitudes until the objective function in (P2) becomes convex at higher altitudes. Then, we consider (P2) to reach an optimal point. We set a condition to change between the rewards based on

$$\frac{1}{L} \sum_{i=1}^L P_{\text{LoS}_i} \geq P_{\text{LoS}_{\text{th}}}. \quad (5.11)$$

As such, we express the reward functions, $R_1[\tau]$ and $R_2[\tau]$, as

$$R_1[\tau] = w_1 \sum_{i=1}^L P_{\text{LoS}_i} + \frac{w_2}{\sum_{i=1}^L d_{vi}}, \quad (5.12)$$

$$R_2[\tau] = w_3 \sum_{i=1}^L \beta_i, \quad (5.13)$$

where w_1 , w_2 , and w_3 are weights to control the influence of the summation of probabilities, P_{LoS_i} , distances, d_{vi} , and average channel gains for the users, β_i 's, respectively.

5.2.1 Proximal Policy Optimization Algorithm

The PPO algorithm is a reinforcement learning approach designed to achieve an optimal balance between efficiency, robustness, and simplicity in policy gradient methods. It operates by alternating between sampling data from the RL environment using the current policy and performing multiple epochs of optimization on a surrogate objective function. According to OpenAI, the inventor of the PPO algorithm, it is a model-

free, on-policy algorithm that directly learns from a collection of experiences [72]. It utilizes two neural networks

- **Actor:** Represented by a policy function $\pi(a|s; \psi)$ where $\pi(\psi)$ is the stochastic policy that determines the probability of the agent taking an action a given a state s .
- **Critic:** Represented by a function $V(s; \kappa)$, where $V(\kappa)$ is the expected cumulative reward for a give state s .

The parameters (weights and biases) for actor and critic neural networks are ψ and κ , respectively. Through learning from experience and interacting with the RL environment, the parameters ψ and κ are modified to maximize a cumulative reward $R[\tau]$. A random number of state-action-reward sequences is selected to train both networks. In each time step, the agent use the generalized advantage estimation (GAE) methods [73] to compute the advantage functions, $\hat{G}[\tau]$, which is given by

$$\hat{G}[\tau] = \sum_{m=\tau}^{N_T} (\gamma\lambda)^{m-\tau} \delta_m, \quad (5.14)$$

$$\delta_m = R[m+1] + \gamma V(\mathbf{S}[m+1]; \kappa) - V(\mathbf{S}[m]; \kappa) \quad (5.15)$$

where γ is the discount factor, and λ denotes the GAE smoothing factor. After each training episode, the algorithm selects a random number of state-action-reward sequences based on the minibatch size (B), update policy $\pi(\psi_{new})$ is found by minimizing the loss function for the actor (L_{actor}) by calculating the average loss for each

entry in the batch

$$L_{\text{actor}} = \frac{1}{B} \sum_{b=1}^B \omega H(\mathbf{S}[b]; \psi) - \min(r(b), c(b, \psi)) \hat{G}(b) \quad (5.16)$$

$$r(b) = \frac{\pi(\boldsymbol{\alpha}[b]|\mathbf{S}[b]; \psi_{\text{new}})}{\pi(\boldsymbol{\alpha}[b]|\mathbf{S}[b]; \psi_{\text{old}})} \quad (5.17)$$

$$c(b; \psi) = \max(\min(r(b), 1 + \epsilon), 1 - \epsilon) \quad (5.18)$$

where ϵ is the clipping factor for the PPO algorithm, which controls the maximum change in policy. ω is a weight for the additional entropy loss which is given by

$$H(\mathbf{S}[b]; \psi) = - \sum_a \pi(\boldsymbol{\alpha}_a|\mathbf{S}[b]; \psi_{\text{new}}) \ln \pi(\boldsymbol{\alpha}_a|\mathbf{S}[b]; \psi_{\text{new}}). \quad (5.19)$$

The loss function for the critic, L_{critic} , is

$$L_{\text{critic}} = \frac{1}{2B} \sum_{b=1}^B \hat{G}^2[b] \quad (5.20)$$

The training process of the PPO algorithm is outlined in Algorithm 2. The training in this section depends on selecting optimal hyperparameters such as discount factor, γ , among other exploration parameters. Their optimization is essential to enhance the model performance which will be discussed in the Bayesian optimization method.

Algorithm 2 PPO algorithm training process

```
1: Randomly establish  $\psi$  and  $\kappa$ .
2: Get policy  $\pi(\psi)$  and  $V(\kappa)$ .
3: for each episode  $e$  do
4:    $\mathbf{S}[0] \leftarrow \{\mathbf{v}[0], \mathcal{U}[\tau]\}$ .
5:   for  $\tau \in T_f$  do
6:     Obtain  $\boldsymbol{\alpha}[\tau]$  based on  $\pi(\psi)$ .
7:      $\mathbf{v}[\tau] \leftarrow \mathbf{v}[\tau - 1] + \boldsymbol{\alpha}[\tau]$ .
8:     if  $\mathbf{v}[\tau] \in \mathcal{E}$  then
9:       if  $\frac{1}{L} \sum_{i=1}^L P_{\text{LoS}_i} \geq P_{\text{LoS}_{\text{th}}}$  then
10:        Calculate reward  $R[\tau]$  based on (5.13).
11:      else
12:        Calculate reward  $R[\tau]$  based on (5.12).
13:      end if
14:    else
15:       $R[\tau] = 0$ .
16:      Terminate episode.
17:    end if
18:     $\mathbf{S}[\tau] \leftarrow \{\mathbf{v}[\tau], \mathcal{U}[\tau]\}$ 
19:    - Compute advantage estimate  $\hat{G}[\tau]$ .
20:    - Add  $\{\mathbf{S}[\tau - 1], \boldsymbol{\alpha}[\tau], R[\tau], \mathbf{S}[\tau], \hat{G}[\tau]\}$  to memory.
21:  end for
22:  - Select mini patch sample of  $\{\mathbf{S}[\tau - 1], \boldsymbol{\alpha}[\tau], R[\tau], \mathbf{S}[\tau], \hat{G}[\tau]\}$  with size  $B$  from memory.
23:  - Update  $\psi$  and  $\kappa$  based on the selected sample using (5.16) and (5.20).
24:  - Repeat sampling and updating  $\psi$  and  $\kappa$  until all data is processed.
25: end for
```

5.2.2 Bayesian Optimization Method

The effective training of DRL models is highly dependent on selecting optimal hyperparameters, including the learning rate, discount factor, and other exploration parameters. Hyperparameter optimization enhances model performance while minimizing training time and computational resources. To achieve this, we propose the Bayesian optimization method [74], [75] to fine-tune the hyperparameters for the proposed DRL-based trajectory optimization model.

Let $f(\boldsymbol{\rho})$ represent the objective function corresponding to the cumulative reward

achieved by PPO algorithm, where $\boldsymbol{\rho}$ is a set of hyperparameters which include

- Minibatch size (B)
- Actor network learning rate (η_a)
- Critic network learning rate (η_c)
- Clipping factor (ϵ)
- Entropy weight (ω_e)
- Number of training epochs,
- Discount factor (γ)

Our objective is to determine hyperparameters $\hat{\boldsymbol{\rho}}$ that maximize the cumulative reward function by the PPO algorithm, i.e.

$$\hat{\boldsymbol{\rho}} = \arg \max_{\boldsymbol{\rho}} f(\boldsymbol{\rho}). \quad (5.21)$$

Bayesian optimization constructs first a surrogate model, $\hat{f}(\boldsymbol{\rho})$, to approximate actual objective function, $f(\boldsymbol{\rho})$. This surrogate model is typically represented by a Gaussian Process (GP), which defines $f(\boldsymbol{\rho})$ as a distribution over functions:

$$\hat{f}(\boldsymbol{\rho}) \sim \mathcal{GP}(\mu(\boldsymbol{\rho}), k(\boldsymbol{\rho}, \boldsymbol{\rho}')), \quad (5.22)$$

where $\mu(\boldsymbol{\rho})$ is the mean function representing the expected value of $f(\boldsymbol{\rho})$, and $k(\boldsymbol{\rho}, \boldsymbol{\rho}')$ is the kernel function representing the similarity between the current hyperparameter

configuration $\boldsymbol{\rho}$ and another set of hyperparameters $\boldsymbol{\rho}'$ [75]. Bayesian optimization uses an acquisition function, $A(\boldsymbol{\rho}, f(\boldsymbol{\rho}))$, which balances exploration (sampling uncertain regions) and exploitation (refining known promising regions). We propose using the Expected Improvement Acquisition Function (EI), which is defined as

$$A(\boldsymbol{\rho}, f(\boldsymbol{\rho})) = \mathbb{E}[\max(0, f(\boldsymbol{\rho}) - f(\boldsymbol{\rho}_{\text{best}}))], \quad (5.23)$$

where $\boldsymbol{\rho}_{\text{best}}$ is the best observed values of $f(\boldsymbol{\rho})$. After evaluating $f(\boldsymbol{\rho})$ for a set of initial hyperparameter configurations with randomly selected values, a Gaussian Process (GP) is fitted to these observations. The GP uses these observations to model the surrogate function $\hat{f}(\boldsymbol{\rho})$, providing both mean predictions $\mu(\boldsymbol{\rho})$ and uncertainties $\sigma(\boldsymbol{\rho})$.

A new hyperparameter configuration $\boldsymbol{\rho}_{\text{new}}$ is then selected by maximizing the acquisition function, i.e.

$$\boldsymbol{\rho}_{\text{new}} = \arg \max_{\boldsymbol{\rho}} A(\boldsymbol{\rho}, \hat{f}(\boldsymbol{\rho})) \quad (5.24)$$

where $A(\boldsymbol{\rho}, \hat{f})$ is the acquisition function, such as the EI function. This step balances exploration, by selecting configurations with high uncertainty, and exploitation, by refining configurations with promising predictions.

Once $\boldsymbol{\rho}_{\text{new}}$ is selected, $f(\boldsymbol{\rho}_{\text{new}})$ is evaluated by training the PPO model with the new configuration, and the observation is added to the dataset. The GP is then updated iteratively refining the surrogate model and acquisition function to guide the search for optimal hyperparameters. With this approach, we can achieve improved PPO performance by systematically balancing exploration of new hyperparameter

configurations and exploitation of promising ones.

5.2.3 Nonstationary UAV and Target Tracking

In section 3.5.3, we described and formulated the EKF process to detect targets by a hovering UAV, i.e. almost stationary. However, our proposed DRL for the optimal UAV trajectory in this section introduces an additional challenge in tracking targets. We will have both the target(s) and UAV to be nonstationary and moving. Such dynamic necessitates a real-time adjustments to the UAV's tracking model. With this in mind, we have adjusted the motion state of the t th moving target as follows

$$\boldsymbol{\xi}_{t,\tau} = \mathbf{T}\boldsymbol{\xi}_{t,\tau-1} - \boldsymbol{\xi}_{v,\tau} + \tilde{\mathbf{n}}_{t,\tau}, \quad (5.25)$$

where $\boldsymbol{\xi}_{v,\tau}$ is the motion state of the UAV at time slot τ , and $\tilde{\mathbf{n}}_{t,\tau}$ is the process noise.

The nonlinear measurement model is expressed as

$$\mathbf{z}_{t,\tau} = h(\mathbf{T}\boldsymbol{\xi}_{t,\tau-1} - \boldsymbol{\xi}_{v,\tau}) + \check{\mathbf{n}}_{t,\tau}, \quad (5.26)$$

where $\check{\mathbf{n}}_{t,\tau}$ is the measurement noise, while $h(\boldsymbol{\xi}_{t,\tau})$ can be expressed as

$$h(\boldsymbol{\xi}_{t,\tau}) = \begin{cases} d_{vt,\tau} = \|\mathbf{v}_\tau - \mathbf{t}_{t,\tau}\| \\ v_{t,\tau} = \frac{((\hat{x}_{t,\tau} - \hat{x}_{v,\tau})(\dot{\hat{x}}_{t,\tau} - \dot{\hat{x}}_{v,\tau}) + (\hat{y}_{t,\tau} - \hat{y}_{v,\tau})(\dot{\hat{y}}_{t,\tau} - \dot{\hat{y}}_{v,\tau}) + (\hat{z}_{t,\tau} - \hat{z}_{v,\tau})(\dot{\hat{z}}_{t,\tau} - \dot{\hat{z}}_{v,\tau}))}{d_{vt,\tau}} \\ \theta_{t,\tau} = \arctan((\hat{y}_{t,\tau} - \hat{y}_{v,\tau})/(\hat{x}_{t,\tau} - \hat{x}_{v,\tau})) \\ \phi_{t,\tau} = \arctan\left((\hat{z}_{t,\tau} - \hat{z}_{v,\tau})/\sqrt{(\hat{x}_{t,\tau} - \hat{x}_{v,\tau})^2 + (\hat{y}_{t,\tau} - \hat{y}_{v,\tau})^2}\right) \end{cases}. \quad (5.27)$$

With (5.25), (5.26) and (5.27), the EKF can track the t th target as explained earlier in the case of hovering UAV.

5.3 RL Training and Numerical Results

In this section, we present numerical simulations of the proposed DRL-based UAV trajectory model and evaluate its performance. The physical environment and simulation parameters are configured as in Table 5.1. At the beginning of each episode, the UAV, users, and targets are randomly located following the uniform distribution. The UAV and users are assumed to be on the ground while the two moving targets have $\xi_1 = [50, -70, 35, -20, 20, 0]$ and $\xi_2 = [80, 80, 35, 0, -20, 0]$, respectively. We consider the probability of LoS, P_{LoS} , based on the model proposed in Chapter 4, for the urban environment ($\varepsilon = 0.3$, $\varrho = 500$, $\vartheta = 15$). We also assume that all users are moving in random directions with random velocities where maximum speed for mobility is set to 1.5 m/s.

Table 5.1: Training and Numerical simulation parameters

Parameter Description	Value
Maximum transmit power (P_T)	20 dBm
Noise power (σ^2)	-80 dBm
Carrier frequency (f_c)	28 GHz
Time slot duration (Δ_T)	0.1ms
Number of time slots (N_T)	300
Number of neighboring users (L)	6
Targets to be sensed (T)	2
URA elements at UAV (N_V)	[8, 4]
ULA elements at user (N_U)	8
$[x_{\min}, x_{\max}]$	$[-100, 100]$ m
$[y_{\min}, y_{\max}]$	$[-100, 100]$ m
$[z_{\min}, z_{\max}]$	$[0, 200]$ m
Shadowing loss (L_s)	40 dB
Actor network hidden layers	[300, 450, 400, 100]
Critic network hidden layers	[400, 550, 500, 200]
Probability of LoS threshold ($P_{\text{LoS}_{\text{th}}}$)	0.5
$[w_1, w_2, w_3]$	[10, 50, 45]

Before training the DRL model, we first optimize the hyperparameters using the proposed Bayesian optimization method, where we set the range for each hyperparameter as given in Table 5.2.

Table 5.2: Hyperparameters Optimization Range

Parameter Description	Range
Minibatch size (B)	$[8, 64]$
Actor network learning rate (η_a)	$[10^{-6}, 10^{-4}]$
Critic network learning rate (η_c)	$[10^{-3}, 10^{-2}]$
Clipping factor (ϵ)	$[0.001, 0.2]$
Entropy weight (ω_e)	$[0.001, 0.1]$
Number of training epochs	$[1, 30]$
Discount factor (γ)	$[0.90, 1.0]$

Generally, using a lower learning rate for the actor compared to the critic allows for stable and incremental policy improvement, leveraging the critic’s faster adaptation to provide reliable feedback. This balance is crucial for efficient training, which is why we have carefully defined the range for the Bayesian optimization method to fine-tune the learning rates accordingly. We executed the optimization method for 30 iterations, yielding the hyperparamters in Table 5.3.

Table 5.3: Optimized Hyperparameters using Bayesian Optimization

B	η_a	η_c	ϵ	ω_e	Epochs	γ
59	8.8333×10^{-6}	0.001512	0.064228	0.085327	20	0.99265

Through the iterative application of the (EI), the hyperparameters that maximize the objective function $f(\boldsymbol{\rho})$ are identified. The resulting optimized values shown in Table 5.3 represent the configuration with the highest $f(\boldsymbol{\rho})$ score, effectively ensuring the best balance between exploration and exploitation during the optimization process.

These values serve as the basis for achieving enhanced model performance. Figure 5.3

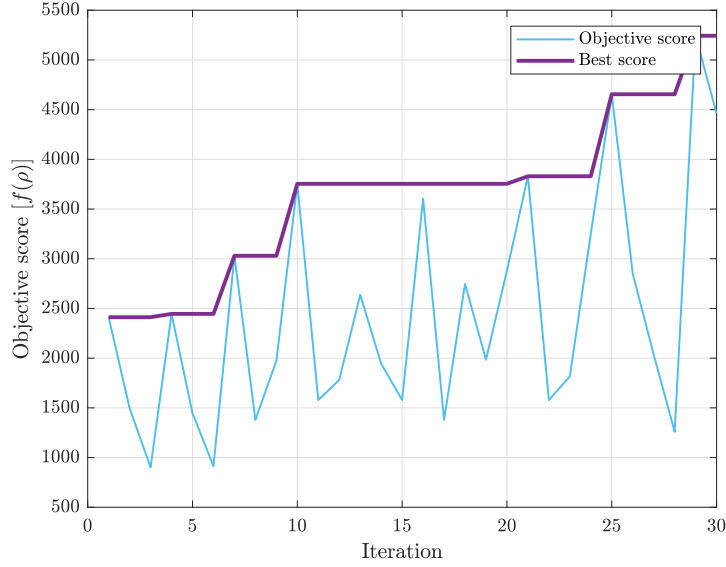


Figure 5.3: Iterative execution of the Bayesian optimization method.

illustrates the progression of scores obtained during each iteration of the Bayesian optimization process, represented by the light blue markers, whereas the best score achieved throughout the entire optimization execution is depicted in violet.

Using the optimized hyperparameter values listed in Table 5.3, we proceed to train the PPO agent, representing the UAV for a total of 11,000 episodes to ensure that the agent learns an effective policy for the given environment and task. Figure 5.4 illustrates the training progression of the UAV trajectory optimization problem using the PPO algorithm described in Algorithm 2. In this figure, the light blue line represents the cumulative reward obtained for each individual episode, reflecting the performance of the agent during that specific episode. Meanwhile, the violet line depicts the moving average of the cumulative reward over episodes, providing a smoothed representation of the overall learning trend and reducing the visual impact

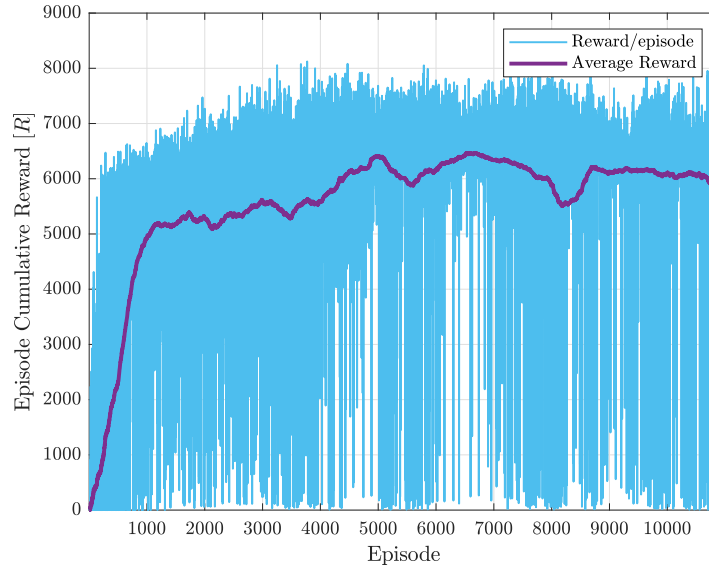


Figure 5.4: Training performance in terms of cumulative reward.

of variability and fluctuations.

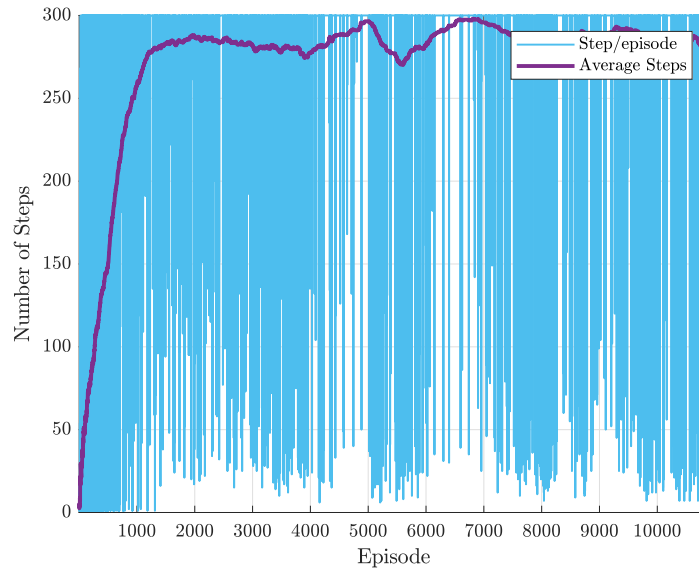


Figure 5.5: Training performance in terms of number of steps.

In DRL, an increase in the number of steps per episode often reflects better decision-making by the agent, as it successfully avoids termination conditions (e.g., out-of-boundary movements) and continues optimizing its trajectory. Consequently,

this improvement translates into higher cumulative rewards, showcasing the agent’s growing competence in navigating the environment. This is evident in Figure 5.5, which shows the steps per episode and the moving average of steps. As the agent progresses through training, episodes with a higher number of steps generally yield higher cumulative rewards. This can be attributed to the agent exploring and interacting with the environment more effectively in these episodes, and not stepping out of the boundaries of the environment $\mathcal{E}$, allowing it to accumulate greater rewards over a longer duration.

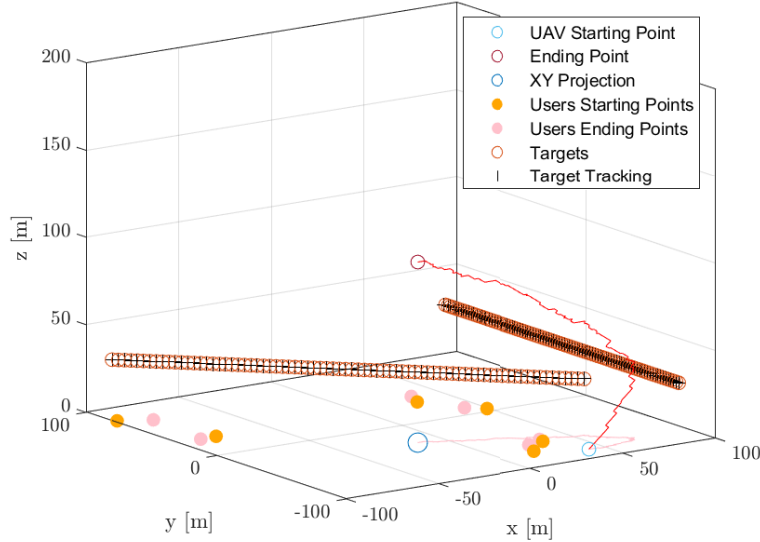


Figure 5.6: A simulation of the trained UAV trajectory, users locations, and target locations and tracking.

Figure 5.6 shows the results of the PPO agent, represented by the UAV, operating within the environment. The trajectory of the UAV is visualized, starting from its initial position (denoted by a blue circle) and moving towards its final position (marked by a dark red circle). The XY projection provides an additional perspective on the UAV’s movement across the environment. The figure also highlights the starting and

ending positions of the users (orange and pink points, respectively) as they move randomly within the environment. The targets are represented by red circles, and the effectiveness of the proposed EKF algorithm is demonstrated by the accuracy in tracking these targets, shown by the red tracking path. One noticeable behavior in the simulation, depicted in Figure 5.6, is the abrupt change in direction near the $x = 50\text{m}$. This behavior can be attributed to two potential reasons. The first, is that the agent likely recognizes that it is near the boundary of the environment, and adjusts its trajectory to remain within the permissible operational area. Second, the agent may have learned through training that such a directional change maximizes its cumulative reward, potentially aligning with its objective.

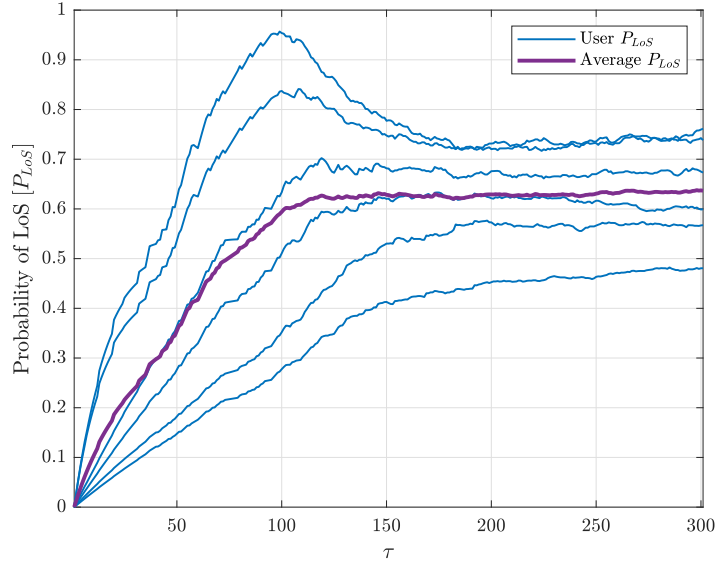


Figure 5.7: Trained UAV performance in terms of P_{LoS} and its average.

Figure 5.7 demonstrates the improvement in the P_{LoS} for individual users and the average P_{LoS} over all users as the UAV navigates its trajectory. The figure highlights how the UAV dynamically adjusts its position to optimize LoS conditions for com-

munication. The individual P_{LoS} trajectories reveal varying levels of improvement depending on the relative position of each user, while the overall average P_{LoS} shows a consistent upward trend, consistent with (5.7), stabilizing as the UAV reaches an optimal trajectory.

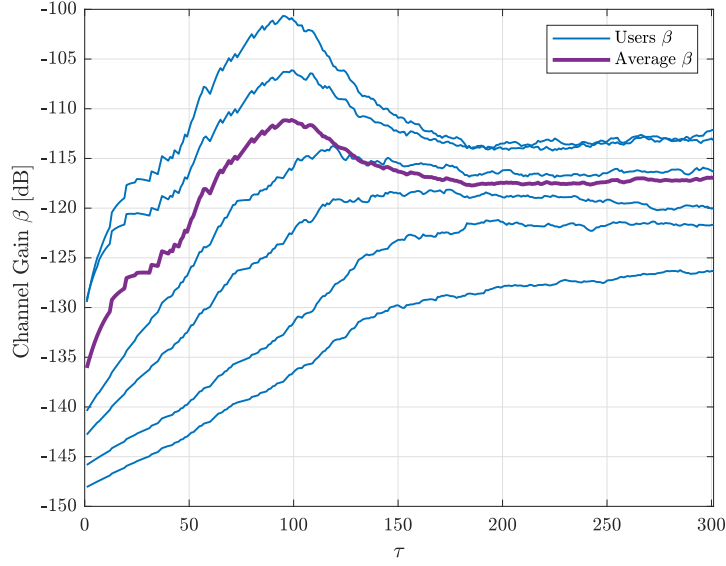


Figure 5.8: Trained UAV performance in terms of channel gain and its average.

Figure 5.8 illustrates the enhancement in channel gains β , for individual users as well as the average β across all users as the UAV progresses along its trajectory. The individual channel gains show variability, reflecting the dynamic adjustments in UAV positioning to optimize conditions for each users (namely the distance and P_{LoS}). The average β exhibits a consistent upward trend, consistent with solving (5.5), indicating the effectiveness of the UAV's trajectory optimization in enhancing communication performance. Comparing the beginning of the scenario to its end, there is a around 19.1 dB improvement in channel gain. The noticeable peaks in individual gains suggest moments when the UAV aligns with favorable conditions for specific users, while the

stabilization of the average gain demonstrates the system’s ability to achieve balanced performance across all users.

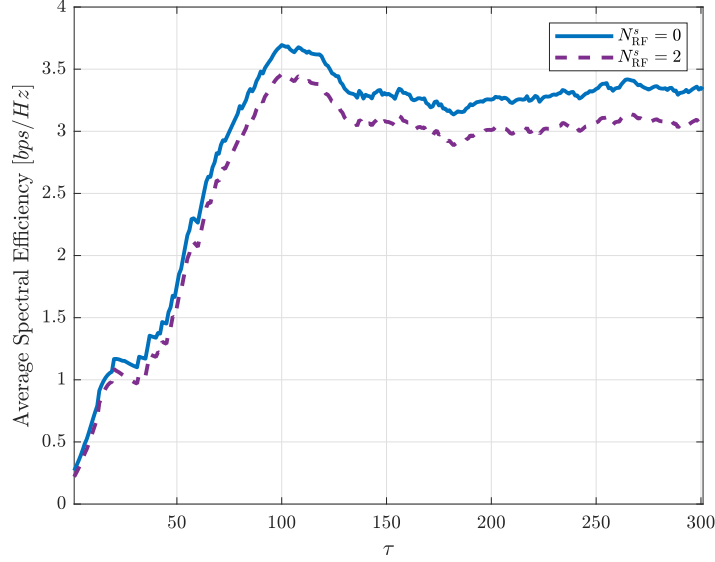


Figure 5.9: Average spectral efficiency with UAV undertaking its planned trajectory, with and without sensing.

Figure 5.9 demonstrates the improvement in average spectral efficiency (ASE) as the UAV progresses along its optimized trajectory. The figure compares the ASE achieved under two scenarios: one with active sensing activity ($N_{\text{RF}} = 2$) and the other without sensing activity ($N_{\text{RF}} = 0$). The results indicate that the spectral efficiency increases significantly during the initial stages of the UAV’s movement as it adjusts its position to optimize communication with the users. Notably, the scenario without sensing activity shows a slightly higher spectral efficiency throughout the trajectory, highlighting the potential communication and sensing trade-off. This trend of ASE, observed with and without active sensing, is consistent with (5.8). It highlights the effectiveness of the proposed DRL model in optimizing the trajectory of a UAV-enabled ISAC system within the defined environment, and demonstrates the model’s

robustness and capability to maximize system performance.

5.4 Summary

In this chapter, we introduced a low-complexity DRL-based framework utilizing the PPO algorithm for UAV-enabled ISAC system trajectory optimization. The proposed model prioritizes the communications' performance while balancing the dual objectives of communication and sensing. To ensure efficient training and enhance model performance, we employed the Bayesian optimization method to fine-tune the hyperparameters of the DRL model. Through a numerical analysis of agent behaviors, we illustrated the effectiveness of the proposed approach in improving communication metrics, such as spectral efficiency and channel gain, while maintaining sensing accuracy. This chapter highlighted the potential of DRL in addressing the complexities of UAV-enabled ISAC systems operating in dynamic and resource-constrained environments.

CHAPTER 6

CONCLUSIONS AND FUTURE DIRECTIONS

6.1 Conclusions

The demand for sensing and communication services is growing with the eventual deployment of 6G wireless networks. UAV-enabled ISAC systems offer enhanced mobility, adaptability, and improved line-of-sight conditions, which suit dynamic and complex environments. This highlights the need for novel designs and approaches, such as hybrid beamforming, reinforcement learning-based trajectory optimization to address the challenges for the next-generation wireless communication systems. This thesis has investigated and proposed DRL-based solutions for multi-user multi-target UAV-enabled ISAC systems under a practical propagation model characterizing irregular geometry in urban areas.

In Chapter 3, we presented the UAV-enabled ISAC system for dynamic and complex environments. We proposed a hybrid beamforming design that enables the UAV

to sense targets of interest and communicate with users while accounting for potential interference caused by integrating sensing tasks. We demonstrated the design’s effectiveness in various environmental structures, such as urban, suburban, and dense urban environments. We investigated spectral efficiency and information outage probability, then evaluated the system’s sensing performance in terms of target detection, localization, and tracking using a hovering UAV.

In Chapter 4, we introduced a novel line-of-sight characterization model based on a geometrical approach, which accounts for the irregularity in terms of widths, lengths, and angles of the spaces where communication users are located. It considered irregular regions, such as intersections or roundabouts, connected to regular regions. We demonstrated the effect of these parameters on the probability of LoS in urban, suburban, dense urban, and high-rise areas. Moreover, we illustrated how the orientation and number of regions where users are located can influence the probability of LoS.

In Chapter 5, we proposed a low-complexity DRL-based model to optimize the trajectory in UAV-enabled ISAC systems for highly dynamic environments. We investigated and proposed reinforcement learning algorithms, specifically, the PPO algorithm, to address the non-convexity of the optimization problem. We used the Bayesian optimization method to determine hyperparameters for the training algorithm, and to ensure robustness. We demonstrated the capability of the proposed model to optimize the UAV’s trajectory policy, and the efficiency under challenging environment conditions.

6.2 Future Directions

There are numerous avenues to extend the thesis work on UAV-enabled ISAC system. One promising direction is the deployment of UAV swarms for sensing and communication across larger areas. While the proposed system demonstrates effectiveness, expanding to a multi-UAV system introduces challenges, e.g. in resource allocation, network architecture, and user association. For resource allocation, machine learning techniques present a promising approach to optimize the distribution of limited resources among the UAVs and users, ensuring efficient and balanced system performance. Regarding network architecture, flying ad hoc networks (FANETs) have emerged to manage inter-UAV networking. In the context of user association, game theory can enable dynamic, decentralized decision-making processes that optimize the interaction between UAVs and users, particularly in competitive or resource-constrained scenarios.

Another promising direction is the integration of neuromorphic receivers and computation. Given the power constraints in UAVs, spiking neural networks (SNNs) paired with neuromorphic hardware offer an energy-efficient and low-latency alternative to traditional machine-learning networks. Neuromorphic processing has not been widely explored for optimizing ISAC performance. SNNs can also be applied in obstacle detection and avoidance in real-time by utilizing sensory information, where UAVs can monitor their surroundings.

APPENDIX A

A.1 Spectral Efficiency Derivation

Based the received signal expression for the i th user in (3.23), we introduce $\mathbf{K}_i$, $\mathbf{K}_l$, and $\mathbf{K}_t$ as follows

$$\mathbf{K}_i = \mathbf{W}_i \mathbf{H}_i \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_1} \quad (\text{A.1})$$

$$\mathbf{K}_l = \mathbf{W}_i \mathbf{H}_i \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_2} \quad (\text{A.2})$$

$$\mathbf{K}_t = \mathbf{W}_i \mathbf{H}_i \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}_3} \quad (\text{A.3})$$

from which we find the covariance matrix, $\mathbf{R}_{y_i}$, which is given by

$$\mathbf{R}_{y_i} = \mathbb{E}[\mathbf{y}_i \mathbf{y}_i^H] = \mathbf{R}_i + \mathbf{R}_{in} \quad (\text{A.4})$$

where $\mathbb{E}[\cdot]$ is the statistical expectation operation, and

$$\mathbf{R}_i = \mathbf{K}_i \mathbb{E}[\mathbf{x}_i \mathbf{x}_i^H] \mathbf{K}_i^H = \mathbf{K}_i \mathbf{C}_{\mathbf{x}_i} \mathbf{K}_i^H, \quad (\text{A.5})$$

$$\mathbf{R}_{in} = \mathbf{R}_l + \mathbf{R}_t + \sigma^2 \mathbf{W}_i \mathbf{W}_i^H \quad (\text{A.6})$$

with $\mathbf{R}_l$ and $\mathbf{R}_t$ in the interference terms are

$$\mathbf{R}_l = \mathbf{K}_l \mathbb{E}[\mathbf{x}_l \mathbf{x}_l^H] \mathbf{K}_l^H = \mathbf{K}_l \mathbf{C}_{\mathbf{x}_l} \mathbf{K}_l^H, \quad (\text{A.7})$$

$$\mathbf{R}_t = \mathbf{K}_t \mathbb{E}[\mathbf{x}_t \mathbf{x}_t^H] \mathbf{K}_t^H = \mathbf{K}_t \mathbf{C}_{\mathbf{x}_t} \mathbf{K}_t^H. \quad (\text{A.8})$$

The channel capacity in bits per complex dimension for the i th user in terms of the conditional mutual information is expressed as

$$\begin{aligned} C_i &= \max_{p(\mathbf{x}_i)} I(\mathbf{x}_i; \mathbf{y}_i | \mathbf{H}_i) \\ &= h(\mathbf{y}_i | \mathbf{H}_i) - h(\mathbf{y}_i | \mathbf{x}_i, \mathbf{H}_i) \end{aligned} \quad (\text{A.9})$$

where $p(\mathbf{x}_i)$ is the distribution of the input signal, in general, following the mathematical derivations in [76] we will have

$$h(\mathbf{y}_i | \mathbf{H}_i) = \log_2 \det \left[\pi e \frac{P_T}{N_P} \mathbf{R}'_{\mathbf{y}_i} \right], \quad (\text{A.10})$$

$$h(\mathbf{y}_i | \mathbf{x}_i, \mathbf{H}_i) = \log_2 \det \left[\pi e P_T \mathbf{R}'_{in} \right] \quad (\text{A.11})$$

where $h(\mathbf{y}_i)$ and $h(\mathbf{y}_i | \mathbf{x}_i)$ are the differential entropy of $\mathbf{y}_i$ and the conditional differential entropy of $\mathbf{y}_i$ given $\mathbf{x}_i$, respectively. Considering a band-limited channel of B Hz, the channel capacity in bps for the i th user substituting (A.10) and (A.11) into (A.9), results in

$$C_i = B \log_2 \det \left[\mathbf{I}_{N_P} + N_P^{-1} \mathbf{R}'_i \mathbf{R}'_{in}{}^{-1} \right] \quad (\text{A.12})$$

and by dividing both sides by B , we obtain the spectral efficiency for the i th user,

SE_i , in (3.24).

A.2 Achievable Rate Derivation

Considering the i th user and an infinite codeword, the channel capacity, C_i , in bits per complex dimension given in (A.9) can be rewritten in terms of the maximum transmit power, P_T , and noise power, σ^2 . The corresponding covariances in (A.5) and (A.6) can be expressed as

$$\mathbf{R}_i = P_T \mathbf{R}_i', \quad (\text{A.13})$$

$$\mathbf{R}_{in} = P_T(\mathbf{R}_l' + \mathbf{R}_t') + \sigma^2 \mathbf{W}_i \mathbf{W}_i^H. \quad (\text{A.14})$$

The differential entropy of $\mathbf{y}_i$ and conditional differential entropy of $\mathbf{y}_i$ given $\mathbf{x}_i$, respectively, are given by

$$h(\mathbf{y}_i | \mathbf{H}_i) = \log_2 \det \left[\pi e \left(P_T \mathbf{R}_i' + P_T(\mathbf{R}_l' + \mathbf{R}_t') + \sigma^2 \mathbf{W}_i \mathbf{W}_i^H \right) \right], \quad (\text{A.15})$$

$$h(\mathbf{y}_i | \mathbf{x}_i, \mathbf{H}_i) = \log_2 \det \left[\pi e \left(P_T(\mathbf{R}_l' + \mathbf{R}_t') + \sigma^2 \mathbf{W}_i \mathbf{W}_i^H \right) \right]. \quad (\text{A.16})$$

Substituting (A.15) and (A.16) into the conditional mutual information expression in (A.9) for a band-limited channel of B Hz will lead to the channel capacity for the i th

user to be

$$\begin{aligned}
C_i &= B [h(\mathbf{y}_i|\mathbf{H}_i) - h(\mathbf{y}_i|\mathbf{x}_i, \mathbf{H}_i)] \\
&= B \log_2 \det \left[\mathbf{I}_{N_P} + \frac{P_T}{\sigma^2} \mathbf{R}'_i \left(\frac{P_T}{\sigma^2} (\mathbf{R}'_l + \mathbf{R}'_t) + \mathbf{W}_i \mathbf{W}_i^H \right)^{-1} \right] \quad (\text{A.17})
\end{aligned}$$

which can be mathematically manipulated by rewriting $\frac{P_T}{\sigma^2}$ in terms of the signal-to-noise ratio per bit, and data rate of i th user, i.e. $\frac{P_T}{\sigma^2} = \frac{\mathcal{E}_b R_i}{N_0 B} = \frac{R_i}{B} \gamma_0$ in (A.17), and the achievable rate will be upper bounded by

$$\frac{R_i}{B} < \log_2 \det \left[\mathbf{I}_{N_P} + \frac{R_i}{B} \gamma_0 \mathbf{R}'_i \left(\frac{R_i}{B} \gamma_0 (\mathbf{R}'_l + \mathbf{R}'_t) + \mathbf{W}_i \mathbf{W}_i^H \right)^{-1} \right] \quad (\text{A.18})$$

where γ_0 is obtained by numerically solving (A.18) with respect to $\frac{R_i}{B}$.

A.3 Information-Outage Probability Derivation

In general, the received signal or output over an AWGN channel can be expressed as $Y = X + N$, where $Y \sim \mathcal{N}_c(0, \mathcal{E}_i + N_0)$ and $X \sim \mathcal{N}_c(0, \mathcal{E}_i)$. Therefore, the mutual information, based on the conditional probability, is expressed as

$$i(y; x) = \log \frac{p_{Y|X}(y|x)}{p_Y(y)}. \quad (\text{A.19})$$

The information is sent through the channel in codewords of length n , and we can

rewrite (A.19) in terms of the MI rate as

$$i(\mathbf{y}; \mathbf{x}) = \frac{1}{n} \log \frac{p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x})}{p_{\mathbf{Y}}(\mathbf{y})} \quad (\text{A.20})$$

for which $\mathbf{y}$ and $\mathbf{x}$ are the output and input sample vectors respectively. Following the mathematical derivations in [55], we can rewrite the MI rate as

$$i(\mathbf{y}; \mathbf{x}) = \log \left(1 + \frac{\mathcal{E}_i}{N_0} \right) + W_n \quad (\text{A.21})$$

where if $n = 1$, the probability distribution of W_n is

$$W_n \sim \mathcal{L} \left(0, \frac{2\mathcal{E}_i}{\mathcal{E}_i + N_0} \right), \quad (\text{A.22})$$

and for a larger n , i.e. the average of n i.i.d Laplacian r.v.'s results in a Bessel- K random variable. For $n \rightarrow \infty$ and using the central limit theorem, we substitute $\mathcal{E}_i$ with $\frac{1}{N_P} \mathbf{R}_i$ from (A.5) and N_0 with $\mathbf{R}_{in}$. We can rewrite (A.20) as

$$i(\mathbf{y}; \mathbf{x}) = \log_2 \det \left(\mathbf{I} + \frac{1}{N_P} \mathbf{R}_i \mathbf{R}_{in}^{-1} \right) + W_n \quad (\text{A.23})$$

resulting in

$$i(\mathbf{y}; \mathbf{x}) \sim \mathcal{N}_c \left(SE_i, \left(\frac{2}{n} \mathbf{R}_i (\mathbf{R}_i + \mathbf{R}_{in})^{-1} \right) \right). \quad (\text{A.24})$$

Hence, the corresponding information outage probability is given by

$$P_{\text{out}} = P[i(\mathbf{y}; \mathbf{x}) \leq R_e] = Q\left(\frac{SE_i - R_e}{\sqrt{\sigma_W^2/n}}\right). \quad (\text{A.25})$$

A.4 Cramér-Rao Lower Bound for Target Localization

Without loss of generality, the received signal echo (3.32) for the t th target is

$$\mathbf{y}_t[\tau] = \mathbf{w}_t[\tau]\mathbf{H}_t[\tau]\mathbf{F}_{\text{RF}}[\tau]\mathbf{F}_{\text{BB}_3}[\tau]x_t[\tau] + \mathbf{w}_t\mathbf{n}[\tau] \quad (\text{A.26})$$

where $\mathbf{H}_t[\tau] = \beta_t \mathbf{a}(\theta_t) \mathbf{a}^H(\theta_t) e^{j2\pi f_D \tau}$, $\mathbf{w}_t[\tau] = \mathbf{a}^H(\theta_t)$ is the analog combiner at the UAV for target t beam, f_D is the Doppler frequency, $\mathbf{n}[\tau]$ is the AWGN noise with zero mean and covariance $\sigma_n^2 \mathbf{I}_{N_V}$. To simplify the derivation, we take advantage of the fact that the zero forcing beamforming negates the power for the other directions, therefore $\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}_3} \rightarrow \sqrt{\frac{P_T}{N_V N_{\text{RF}}}} \mathbf{a}_t^*$. At time slot τ , we can rewrite (A.26) as:

$$\mathbf{y}_t[\tau] = \beta_t[\tau] \sqrt{\frac{P_T}{N_V N_{\text{RF}}}} \mathbf{a}_t^H \mathbf{a}_t \mathbf{a}_t^H \mathbf{a}_t^* e^{j2\pi f_D \tau} x_t[\tau] + \mathbf{a}_t^H n \quad (\text{A.27})$$

From which we can represent $U = \beta_t[\tau] \sqrt{\frac{P_T}{N_V N_{\text{RF}}}} \mathbf{a}_t^H \mathbf{a}_t \mathbf{a}_t^H \mathbf{a}_t^* e^{j2\pi f_D \tau}$ as the amplitude of echo signal. Since $x_t \sim \mathcal{N}(0, 1)$, and $n \sim \mathcal{N}(0, \sigma_n^2)$, then $\mathbf{y}_t \sim \mathcal{N}(0, C)$, where $C = U^2 + \mathbf{a}_t \mathbf{a}_t^H \sigma_n^2$. It is important to note that θ_t, ϕ_t, v_t and d_{vt} are independent from each other, therefore, we can find the CRLB for each estimate individually. The

Probability Distribution Function (PDF) of $\mathbf{y}_t$ is

$$\mathbb{P}(\mathbf{y}_t[\tau]; [\theta_t, \phi_t, v_t, d_{vt}]) = \frac{1}{\sqrt{2\pi C}} \exp \left(-\frac{1}{2C} (\mathbf{y}_t[\tau] - Ux_t[\tau])^2 \right) \quad (\text{A.28})$$

Given Q number of realizations of $\mathbf{y}_t[\tau]$:

$$\mathbb{P}(\mathbf{y}_t[\tau]; [\theta_t, \phi_t, v_t, d_{ut}]) = \frac{1}{(\sqrt{2\pi C})^Q} \exp \left(-\frac{1}{2C} \sum_{q=1}^Q (\mathbf{y}_t[q] - Ux_t[q])^2 \right) \quad (\text{A.29})$$

$$\ln \mathbb{P}(\mathbf{y}_t[\tau]; [\theta_t, \phi_t, v_t, d_{ut}]) = -\frac{Q}{2} \ln 2\pi - \frac{Q}{2} \ln C - \frac{1}{2C} \sum_{q=1}^Q (\mathbf{y}_t[q] - Ux_t[q])^2 \quad (\text{A.30})$$

For θ_t :

$$\frac{\partial}{\partial \theta_t} \ln \mathbb{P}(\mathbf{y}_t[\tau]; [\theta_t, \phi_t, v_t, d_{vt}]) = -\frac{Q}{2C} \frac{\partial C}{\partial \theta_t} - \frac{\partial}{\partial \theta_t} \frac{D}{2C} \quad (\text{A.31})$$

$$D = \sum_{q=1}^Q (\mathbf{y}_t[q] - Ux_t[q])^2 \quad (\text{A.32})$$

$$\frac{\partial D}{\partial \theta_t} = -\sum_{q=1}^Q (\mathbf{y}_t[q] - Ux_t[q]) \frac{\partial U}{\partial \theta_t} x_t[q] \quad (\text{A.33})$$

$$\frac{\partial}{\partial \theta_t} \frac{D}{2C} = \frac{2C \frac{\partial D}{\partial \theta_t} - 2 \frac{\partial C}{\partial \theta_t} D}{(2C)^2} \quad (\text{A.34})$$

$$= \frac{-2C \sum_{q=1}^Q (\mathbf{y}_t[q] - Ux_t[q]) \frac{\partial U}{\partial \theta_t} x_t[q] - 2 \frac{\partial C}{\partial \theta_t} \sum_{q=1}^Q (\mathbf{y}_t[q] - Ux_t[q])^2}{(2C)^2} \quad (\text{A.35})$$

Taking the expectation of (A.34)

$$E \left[\frac{\partial}{\partial \theta_t} \ln \mathbb{P}(\mathbf{y}_t[\tau]; [\theta_t, \phi_t, v_t, d_{\mathbf{vt}}]) \right] = -\mathbb{E} \left[\frac{Q \frac{\partial C}{\partial \theta_t}}{2C} \right] - \mathbb{E} \left[\frac{-2C \sum_{q=1}^Q (\mathbf{y}_t[q] - U x_t[q]) \frac{\partial U}{\partial \theta_t} x_t[q] - 2 \frac{\partial C}{\partial \theta_t} \sum_{q=1}^Q (\mathbf{y}_t[q] - U x_t[q])^2}{(2C)^2} \right]. \quad (\text{A.36})$$

Knowing that

$$\mathbb{E} [\mathbf{y}^T[q] \mathbf{y}[q]] = C \quad (\text{A.37})$$

$$\mathbb{E} [x[q] \mathbf{y}[q]] = U \quad (\text{A.38})$$

$$\mathbb{E} [x^2[q]] = 1 \quad (\text{A.39})$$

$$\mathbb{E} [x[q]] = \mathbb{E} [\mathbf{y}[q]] = 0 \quad (\text{A.40})$$

$$\mathbb{E} [C] = C \quad (\text{A.41})$$

$$\mathbb{E} \left[\frac{\partial C}{\partial \theta_t} \right] = \frac{\partial C}{\partial \theta_t} \quad (\text{A.42})$$

$$\mathbb{E} [U] = U \quad (\text{A.43})$$

$$\mathbb{E} \left[\frac{\partial U}{\partial \theta_t} \right] = \frac{\partial U}{\partial \theta_t} \quad (\text{A.44})$$

we can simplify

$$E \left[\frac{\partial}{\partial \theta_t} \ln \mathbb{P}(\mathbf{y}_t[\tau]; [\theta_t, \phi_t, v_t, d_{\mathbf{vt}}]) \right] = -\frac{Q \frac{\partial C}{\partial \theta_t} (U^2 - 2C)}{2C^2} \quad (\text{A.45})$$

As such, the Cramér-Rao lower bound for θ_t is:

$$CRLB_{\theta_t} = \frac{1}{E \left[\frac{\partial}{\partial \theta_t} \ln \mathbb{P}(\mathbf{y}_t[\tau]; [\theta_t, \phi_t, v_t, d_{vt}]) \right]^2} \quad (\text{A.46})$$

$$= \frac{1}{\left[-\frac{Q \frac{\partial C}{\partial \theta_t} (U^2 - 2C)}{2C^2} \right]^2} = \frac{4C^4}{\left[Q \frac{\partial C}{\partial \theta_t} (U^2 - 2C) \right]^2} \quad (\text{A.47})$$

Similarly, for ϕ_t :

$$CRLB_{\phi} = \frac{1}{E \left[\frac{\partial}{\partial \phi_t} \ln \mathbb{P}(\mathbf{y}_t[\tau]; [\theta_t, \phi_t, v_t, d_{vt}]) \right]^2} \quad (\text{A.48})$$

$$= \frac{1}{\left[-\frac{Q \frac{\partial C}{\partial \phi_t} (U^2 - 2C)}{2C^2} \right]^2} = \frac{4C^4}{\left[Q \frac{\partial C}{\partial \phi_t} (U^2 - 2C) \right]^2} \quad (\text{A.49})$$

For distance d_{ut} :

$$CRLB_{d_{vt}} = \frac{1}{E \left[\frac{\partial}{\partial d_{vt}} \ln \mathbb{P}(\mathbf{y}_t[\tau]; [\theta_t, \phi_t, v_t, d_{vt}]) \right]^2} \quad (\text{A.50})$$

$$= \frac{1}{\left[-\frac{Q \frac{\partial C}{\partial d_{vt}} (U^2 - 2C)}{2C^2} \right]^2} = \frac{4C^4}{\left[Q \frac{\partial C}{\partial d_{vt}} (U^2 - 2C) \right]^2} \quad (\text{A.51})$$

For velocity v_t :

$$CRLB_{v_t} = \frac{1}{E \left[\frac{\partial}{\partial v_t} \ln \mathbb{P}(\mathbf{y}_t[\tau]; [\theta_t, \phi_t, v_t, d_{ut}]) \right]^2} \quad (\text{A.52})$$

$$= \frac{1}{\left[-\frac{Q \frac{\partial C}{\partial v_t} (U^2 - 2C)}{2C^2} \right]^2} = \frac{4C^4}{\left[Q \frac{\partial C}{\partial v_t} (U^2 - 2C) \right]^2} \quad (\text{A.53})$$

where

$$C = U^2 + \mathbf{a}_t^H(\theta_t, \phi_t) \mathbf{a}_t(\theta_t, \phi_t) \sigma_n^2 \quad (\text{A.54})$$

$$\frac{\partial C}{\partial \theta_t} = 2U \frac{\partial U}{\partial \theta_t} + \left(\frac{\partial \mathbf{a}_t^H(\theta_t, \phi_t)}{\partial \theta} \mathbf{a}_t(\theta_t, \phi_t) + \mathbf{a}_t^H(\theta_t, \phi_t) \frac{\partial \mathbf{a}_t(\theta_t, \phi_t)}{\partial \theta_t} \right) \sigma_n^2 \quad (\text{A.55})$$

$$\frac{\partial C}{\partial \phi_t} = 2U \frac{\partial U}{\partial \phi_t} + \left(\frac{\partial \mathbf{a}_t^H(\theta_t, \phi_t)}{\partial \phi_t} \mathbf{a}_t(\theta_t, \phi_t) + \mathbf{a}_t^H(\theta_t, \phi_t) \frac{\partial \mathbf{a}_t(\theta_t, \phi_t)}{\partial \phi_t} \right) \sigma_n^2 \quad (\text{A.56})$$

$$\frac{\partial C}{\partial d_{ut}} = 2U \frac{\partial U}{\partial d_{ut}} \quad (\text{A.57})$$

$$\frac{\partial C}{\partial v_t} = 2U \frac{\partial U}{\partial v_t} \quad (\text{A.58})$$

$$U = \beta_t \sqrt{\frac{P_T}{N_V N_{\text{RF}}}} \mathbf{a}_t^H \mathbf{a}_t \mathbf{a}_t^H \mathbf{a}_t^* e^{j2\pi f_D \tau} \quad (\text{A.59})$$

$$\begin{aligned} \frac{\partial U}{\partial \theta_t} &= \beta_t \sqrt{\frac{P_T}{N_V N_{\text{RF}}}} \frac{\partial \mathbf{a}_t^H}{\partial \theta_t} \mathbf{a}_t \mathbf{a}_t^H \mathbf{a}_t^* e^{j2\pi f_D \tau} \\ &+ \beta_t \sqrt{\frac{P_T}{N_V N_{\text{RF}}}} \mathbf{a}_t^H \frac{\partial \mathbf{a}_t}{\partial \theta_t} \mathbf{a}_t^H \mathbf{a}_t^* e^{j2\pi f_D \tau} \\ &+ \beta_t \sqrt{\frac{P_T}{N_V N_{\text{RF}}}} \mathbf{a}_t^H \mathbf{a}_t \frac{\partial \mathbf{a}_t^H}{\partial \theta_t} \mathbf{a}_t^* e^{j2\pi f_D \tau} \\ &+ \beta_t \sqrt{\frac{P_T}{N_V N_{\text{RF}}}} \mathbf{a}_t^H \mathbf{a}_t \mathbf{a}_t^H \frac{\partial \mathbf{a}_t^*}{\partial \theta_t} e^{j2\pi f_D \tau} \end{aligned} \quad (\text{A.60})$$

$$\begin{aligned} \frac{\partial U}{\partial \phi_t} &= \beta_t \sqrt{\frac{P_T}{N_V N_{\text{RF}}}} \frac{\partial \mathbf{a}_t^H}{\partial \phi_t} \mathbf{a}_t \mathbf{a}_t^H \mathbf{a}_t^* e^{j2\pi f_D \tau} \\ &+ \beta_t \sqrt{\frac{P_T}{N_V N_{\text{RF}}}} \mathbf{a}_t^H \frac{\partial \mathbf{a}_t}{\partial \phi_t} \mathbf{a}_t^H \mathbf{a}_t^* e^{j2\pi f_D \tau} \\ &+ \beta_t \sqrt{\frac{P_T}{N_V N_{\text{RF}}}} \mathbf{a}_t^H \mathbf{a}_t \frac{\partial \mathbf{a}_t^H}{\partial \phi_t} \mathbf{a}_t^* e^{j2\pi f_D \tau} \\ &+ \beta_t \sqrt{\frac{P_T}{N_V N_{\text{RF}}}} \mathbf{a}_t^H \mathbf{a}_t \mathbf{a}_t^H \frac{\partial \mathbf{a}_t^*}{\partial \phi_t} e^{j2\pi f_D \tau} \end{aligned} \quad (\text{A.61})$$

$$\frac{\partial U}{\partial d_{vt}} = \frac{2\sqrt{\beta_0}}{d_{vt}^3} \sqrt{\frac{P_T}{N_V N_{\text{RF}}}} \mathbf{a}_t^H \mathbf{a}_t \mathbf{a}_t^H \mathbf{a}_t^* e^{j2\pi f_D \tau} \quad (\text{A.62})$$

where $\beta_t = \frac{\sqrt{\beta_0}}{d_{ut}^2}$.

$$\frac{\partial U}{\partial v_t} = j2\pi \frac{f_c}{c} \tau \beta_t \sqrt{\frac{P_T}{N_V N_{RF}}} \mathbf{a}_t^H \mathbf{a}_t^H \mathbf{a}_t^* e^{j2\pi f_D \tau} \quad (\text{A.63})$$

where $f_D = \frac{f_c v_k t}{c}$

In case URA antenna of size $N \times M$ is used

$$a_{n,m}(\theta, \phi) = \exp \left[\frac{j2\pi}{\lambda} (d(n-1) \sin \theta \sin \phi - d(m-1) \cos \theta \sin \phi) \right] \quad (\text{A.64})$$

$$\begin{aligned} \frac{\partial a_{n,m}(\theta, \phi)}{\partial \theta} &= \frac{j2\pi}{\lambda} (d(n-1) \cos \theta \sin \phi + d(m-1) \sin \theta \sin \phi) \\ &\times \exp \left[\frac{j2\pi}{\lambda} (d(n-1) \sin \theta \sin \phi - d(m-1) \cos \theta \sin \phi) \right] \end{aligned} \quad (\text{A.65})$$

$$\begin{aligned} \frac{\partial a_{n,m}(\theta, \phi)}{\partial \phi} &= \frac{j2\pi}{\lambda} (d(n-1) \sin \theta \cos \phi - d(m-1) \cos \theta \cos \phi) \\ &\times \exp \left[\frac{j2\pi}{\lambda} (d(n-1) \sin \theta \sin \phi - d(m-1) \cos \theta \sin \phi) \right] \end{aligned} \quad (\text{A.66})$$

$$\frac{\partial \mathbf{A}}{\partial \theta} = \begin{bmatrix} \frac{\partial}{\partial \theta} a_{1,1} & \frac{\partial}{\partial \theta} a_{1,2} & \dots & \frac{\partial}{\partial \theta} a_{1,M} \\ \frac{\partial}{\partial \theta} a_{2,2} & \frac{\partial}{\partial \theta} a_{2,2} & \dots & \frac{\partial}{\partial \theta} a_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial \theta} a_{N,1} & \frac{\partial}{\partial \theta} a_{N,2} & \dots & \frac{\partial}{\partial \theta} a_{N,M} \end{bmatrix} \quad (\text{A.67})$$

$$\frac{\partial \mathbf{A}}{\partial \phi} = \begin{bmatrix} \frac{\partial}{\partial \phi} a_{1,1} & \frac{\partial}{\partial \phi} a_{1,2} & \dots & \frac{\partial}{\partial \phi} a_{1,M} \\ \frac{\partial}{\partial \phi} a_{2,2} & \frac{\partial}{\partial \phi} a_{2,2} & \dots & \frac{\partial}{\partial \phi} a_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial}{\partial \phi} a_{N,1} & \frac{\partial}{\partial \phi} a_{N,2} & \dots & \frac{\partial}{\partial \phi} a_{N,M} \end{bmatrix} \quad (\text{A.68})$$

$$\frac{\partial \mathbf{a}(\theta, \phi)}{\partial \theta} = \text{vec} \left(\frac{\partial \mathbf{A}(\theta, \phi)}{\partial \theta} \right) \quad (\text{A.69})$$

$$\frac{\partial \mathbf{a}(\theta, \phi)}{\partial \phi} = \text{vec} \left(\frac{\partial \mathbf{A}(\theta, \phi)}{\partial \phi} \right) \quad (\text{A.70})$$

In case of ULA antenna of size $N \times 1$ is used

$$a_n(\theta) = \exp \left[\frac{j2\pi}{\lambda} d(n-1) \sin(\theta) \right] \quad (\text{A.71})$$

$$\frac{\partial a_n}{\partial \theta} = \frac{j2\pi}{\lambda} d(n-1) \cos(\theta) \exp \left[\frac{j2\pi}{\lambda} d(n-1) \sin(\theta) \right] \quad (\text{A.72})$$

$$\frac{\partial \mathbf{a}(\theta)}{\partial \theta} = \left[\frac{\partial a_1}{\partial \theta}, \frac{\partial a_2}{\partial \theta}, \dots, \frac{\partial a_N}{\partial \theta} \right]^T \quad (\text{A.73})$$

A.5 Jacobian Matrix for Nonlinear Measurement Model

The Jacobian matrix for $h(\boldsymbol{\xi}_{t,\tau})$ is as follows

$$\tilde{\mathbf{J}}_t = \begin{bmatrix} \frac{\partial d_{\mathbf{v}t}}{\partial x_t} & \frac{\partial d_{\mathbf{v}t}}{\partial y_t} & \frac{\partial d_{\mathbf{v}t}}{\partial z_t} & 0 & 0 & 0 \\ \frac{\partial v_t}{\partial x_t} & \frac{\partial v_t}{\partial y_t} & \frac{\partial v_t}{\partial z_t} & \frac{\partial v_t}{\partial \dot{x}_t} & \frac{\partial v_t}{\partial \dot{y}_t} & \frac{\partial v_t}{\partial \dot{z}_t} \\ \frac{\partial \theta_t}{\partial x_t} & \frac{\partial \theta_t}{\partial y_t} & 0 & 0 & 0 & 0 \\ \frac{\partial \phi_t}{\partial x_t} & \frac{\partial \phi_t}{\partial y_t} & \frac{\partial \phi_t}{\partial z_t} & 0 & 0 & 0 \end{bmatrix}, \quad (\text{A.74})$$

where

$$\frac{\partial d_{\mathbf{v}t}}{\partial x_t} = \frac{x_t}{\left(\sqrt{x_t^2 + y_t^2 + z_t^2}\right)^3} \quad (\text{A.75})$$

$$\frac{\partial d_{\mathbf{v}t}}{\partial y_t} = \frac{y_t}{\left(\sqrt{x_t^2 + y_t^2 + z_t^2}\right)^3} \quad (\text{A.76})$$

$$\frac{\partial d_{\mathbf{v}t}}{\partial z_t} = \frac{z_t}{\left(\sqrt{x_t^2 + y_t^2 + z_t^2}\right)^3} \quad (\text{A.77})$$

$$\frac{\partial v_t}{\partial x_t} = \frac{\dot{x}_t}{\sqrt{x_t^2 + y_t^2 + z_t^2}} - \frac{x_t(x_t\dot{x}_t + y_t\dot{y}_t + z_t\dot{z}_t)}{\left(\sqrt{x_t^2 + y_t^2 + z_t^2}\right)^3} \quad (\text{A.78})$$

$$\frac{\partial v_t}{\partial y_t} = \frac{\dot{y}_t}{\sqrt{x_t^2 + y_t^2 + z_t^2}} - \frac{y_t(x_t\dot{x}_t + y_t\dot{y}_t + z_t\dot{z}_t)}{\left(\sqrt{x_t^2 + y_t^2 + z_t^2}\right)^3} \quad (\text{A.79})$$

$$\frac{\partial v_t}{\partial z_t} = \frac{\dot{z}_t}{\sqrt{x_t^2 + y_t^2 + z_t^2}} - \frac{z_t(x_t\dot{x}_t + y_t\dot{y}_t + z_t\dot{z}_t)}{\left(\sqrt{x_t^2 + y_t^2 + z_t^2}\right)^3} \quad (\text{A.80})$$

$$\frac{\partial v_t}{\partial \dot{x}_t} = \frac{x_t}{\sqrt{x_t^2 + y_t^2 + z_t^2}} \quad (\text{A.81})$$

$$\frac{\partial v_t}{\partial \dot{y}_t} = \frac{y_t}{\sqrt{x_t^2 + y_t^2 + z_t^2}} \quad (\text{A.82})$$

$$\frac{\partial v_t}{\partial \dot{z}_t} = \frac{z_t}{\sqrt{x_t^2 + y_t^2 + z_t^2}} \quad (\text{A.83})$$

$$\frac{\partial \theta_t}{\partial x_t} = -\frac{y_t}{x_t^2 + y_t^2} \quad (\text{A.84})$$

$$\frac{\partial \theta_t}{\partial y_t} = \frac{x_t}{x_t^2 + y_t^2} \quad (\text{A.85})$$

$$\frac{\partial \phi_t}{\partial x_t} = \frac{z_t x_t}{(x_t^2 + y_t^2 + z_t^2)\sqrt{x_t^2 + y_t^2}} \quad (\text{A.86})$$

$$\frac{\partial \phi_t}{\partial y_t} = \frac{z_t y_t}{(x_t^2 + y_t^2 + z_t^2)\sqrt{x_t^2 + y_t^2}} \quad (\text{A.87})$$

$$\frac{\partial \phi_t}{\partial z_t} = \frac{\sqrt{x_t^2 + y_t^2}}{(x_t^2 + y_t^2 + z_t^2)} \quad (\text{A.88})$$

REFERENCES

- [1] J. Holis and P. Pechac, “Elevation dependent shadowing model for mobile communications via high altitude platforms in built-up areas,” *IEEE Transactions on Antennas and Propagation*, vol. 56, no. 4, pp. 1078–1084, 2008.
- [2] K. Meng, Q. Wu, J. Xu, W. Chen, Z. Feng, R. Schober, and A. L. Swindlehurst, “UAV-enabled integrated sensing and communication: Opportunities and challenges,” *IEEE Wireless Communications*, pp. 1–9, 2023.
- [3] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, “Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond,” *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 6, pp. 1728–1767, 2022.
- [4] Z. Wan, Z. Gao, S. Tan, and L. Fang, “Joint channel estimation and radar sensing for UAV networks with mmWave massive MIMO,” in *2022 International Wireless Communications and Mobile Computing (IWCMC)*, 2022, pp. 44–49.
- [5] S. Zhang, Y. Zeng, and R. Zhang, “Cellular-enabled UAV communication: A connectivity-constrained trajectory optimization perspective,” *IEEE Transactions on Communications*, vol. 67, no. 3, pp. 2580–2604, 2019.

- [6] A. Bayesteh, J. He, Y. Chen, P. Zhu, J. Ma, A. Shaban, Z. Yu, Y. Zhang, Z. Zhou, and G. Wang, “Integrated sensing and communication (ISAC)—from concept to practice,” *Communications of Huawei Research*, pp. 4–25, 2022.
- [7] U. Demirhan and A. Alkhateeb, “Integrated sensing and communication for 6G: Ten key machine learning roles,” *IEEE Communications Magazine*, vol. 61, no. 5, pp. 113–119, 2023.
- [8] C. Szepesvári, *Algorithms for reinforcement learning*. Springer nature, 2022.
- [9] N. Wu, R. Jiang, X. Wang, L. Yang, K. Zhang, W. Yi, and A. Nallanathan, “AI-enhanced integrated sensing and communications: Advancements, challenges, and prospects,” *IEEE Communications Magazine*, vol. 62, no. 9, pp. 144–150, 2024.
- [10] H. Lin, Z. Zhang, L. Wei, Z. Zhou, and T. Zheng, “A deep reinforcement learning based UAV trajectory planning method for integrated sensing and communications networks,” in *2023 IEEE 98th Vehicular Technology Conference (VTC2023-Fall)*, 2023, pp. 1–6.
- [11] X. Liu, H. Zhang, K. Sun, K. Long, and G. K. Karagiannidis, “AI-driven integration of sensing and communication in the 6G era,” *IEEE Network*, vol. 38, no. 3, pp. 210–217, 2024.
- [12] W. Saad, M. Bennis, and M. Chen, “A vision of 6G wireless systems: Applications, trends, technologies, and open research problems,” *IEEE Network*, vol. 34, no. 3, pp. 134–142, 2020.

- [13] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo, “Joint radar and communication design: Applications, state-of-the-art, and the road ahead,” *IEEE Transactions on Communications*, vol. 68, no. 6, pp. 3834–3862, 2020.
- [14] L. Han and K. Wu, “Joint wireless communication and radar sensing systems – state of the art and future prospects,” *IET Microwaves, Antennas & Propagation*, vol. 7, no. 11, pp. 876–885, 2013. [Online]. Available: <https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-map.2012.0450>
- [15] C. Shi, F. Wang, M. Sellathurai, J. Zhou, and S. Salous, “Power minimization-based robust OFDM radar waveform design for radar and communication systems in coexistence,” *IEEE Transactions on Signal Processing*, vol. 66, no. 5, pp. 1316–1330, 2018.
- [16] T. Huang, N. Shlezinger, X. Xu, Y. Liu, and Y. C. Eldar, “MAJoRCom: A dual-function radar communication system using index modulation,” *IEEE Transactions on Signal Processing*, vol. 68, pp. 3423–3438, 2020.
- [17] A. R. Chiriyath, B. Paul, G. M. Jacyna, and D. W. Bliss, “Inner bounds on performance of radar and communications co-existence,” *IEEE Transactions on Signal Processing*, vol. 64, no. 2, pp. 464–474, 2016.
- [18] Y. Qin, Z. Zhang, X. Li, W. Huangfu, and H. Zhang, “Deep reinforcement learning based resource allocation and trajectory planning in integrated sensing and communications UAV network,” *IEEE Transactions on Wireless Communications*, vol. 22, no. 11, pp. 8158–8169, 2023.

- [19] S. Şahin and T. Girici, “A novel radar receiver for joint radar and communication systems,” in *2023 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET)*, 2023, pp. 61–66.
- [20] M. Li and W.-Q. Wang, “Performance analysis of joint radar-communication system based on FDA-MIMO radar,” *IEEE Access*, vol. 12, pp. 48 774–48 787, 2024.
- [21] L. Hu, J. Mu, X. Jing, and J. Zhu, “Joint subcarrier and power allocation design based on radar SNR ratio in integrated radar and communication system,” in *2023 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2023, pp. 1–6.
- [22] A. Liu, Z. Huang, M. Li, Y. Wan, W. Li, T. X. Han, C. Liu, R. Du, D. K. P. Tan, J. Lu, Y. Shen, F. Colone, and K. Chetty, “A survey on fundamental limits of integrated sensing and communication,” *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 994–1034, 2022.
- [23] M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, and M. Debbah, “A tutorial on UAVs for wireless networks: Applications, challenges, and open problems,” *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2334–2360, 2019.
- [24] Q. Wu, Y. Zeng, and R. Zhang, “Joint trajectory and communication design for multi-UAV enabled wireless networks,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 2109–2121, 2018.

- [25] I. Orikumhi, J. Bae, and S. Kim, “UAV-assisted integrated sensing and communications for user blockage prediction,” in *2022 13th International Conference on Information and Communication Technology Convergence (ICTC)*, 2022, pp. 471–473.
- [26] H. Kim, M. Hwang, J. Jee, J. Park, and H. Park, “3D state transition modeling and power allocation for UAV-aided ISAC system,” in *2023 IEEE 98th Vehicular Technology Conference (VTC2023-Fall)*, 2023, pp. 1–6.
- [27] Y. Liu, S. Liu, X. Liu, Z. Liu, and T. S. Durrani, “Sensing fairness-based energy efficiency optimization for UAV enabled integrated sensing and communication,” *IEEE Wireless Communications Letters*, vol. 12, no. 10, pp. 1702–1706, 2023.
- [28] R. Zhang, Y. Zhang, R. Tang, H. Zhao, Q. Xiao, and C. Wang, “A joint UAV trajectory, user association, and beamforming design strategy for multi-UAV assisted ISAC systems,” *IEEE Internet of Things Journal*, pp. 1–1, 2024.
- [29] Y. Wu, C. Liu, X. Hu, and M. Peng, “Sensing and beamtracking scheme design for UAV-enabled ISAC systems,” in *2023 International Conference on Wireless Communications and Signal Processing (WCSP)*, 2023, pp. 146–151.
- [30] K. Meng, Q. Wu, S. Ma, W. Chen, and T. Q. S. Quek, “UAV trajectory and beamforming optimization for integrated periodic sensing and communication,” *IEEE Wireless Communications Letters*, vol. 11, no. 6, pp. 1211–1215, 2022.

- [31] X. Jing, F. Liu, C. Masouros, and Y. Zeng, “ISAC from the sky: UAV trajectory design for joint communication and target localization,” *IEEE Transactions on Wireless Communications*, vol. 23, no. 10, pp. 12 857–12 872, 2024.
- [32] Q. Zhu, R. Liu, X. Lv, Q. Meng, and Y. Wang, “AoI-optimal trajectory planning in UAV-assisted ISAC networks,” in *2023 IEEE 23rd International Conference on Communication Technology (ICCT)*, 2023, pp. 428–433.
- [33] Y. Li, X. Li, J. Yang, J. Pan, R. Zhu, and H. Liu, “Continuous UAV trajectory design with uncertain user location in ISAC networks,” in *2024 IEEE Wireless Communications and Networking Conference (WCNC)*, 2024, pp. 1–6.
- [34] K. Meng, Q. Wu, S. Ma, W. Chen, K. Wang, and J. Li, “Throughput maximization for UAV-enabled integrated periodic sensing and communication,” *IEEE Transactions on Wireless Communications*, vol. 22, no. 1, pp. 671–687, 2023.
- [35] A. Khalili, A. Rezaei, D. Xu, F. Dressler, and R. Schober, “Efficient UAV hovering, resource allocation, and trajectory design for ISAC with limited backhaul capacity,” *IEEE Transactions on Wireless Communications*, vol. 23, no. 11, pp. 17 635–17 650, 2024.
- [36] H. Mohammadi, V. Marojevic, and B. Shang, “Analysis of reinforcement learning schemes for trajectory optimization of an aerial radio unit,” in *ICC 2023 - IEEE International Conference on Communications*, 2023, pp. 6423–6428.
- [37] C.-W. Fu, M.-L. Ku, Y.-J. Chen, and T. Q. S. Quek, “UAV trajectory, user association, and power control for multi-UAV-enabled energy-harvesting commu-

- nications: Offline design and online reinforcement learning,” *IEEE Internet of Things Journal*, vol. 11, no. 6, pp. 9781–9800, 2024.
- [38] L. Zhao and B. Li, “Phase design and trajectory optimization of IRS-assisted UAV-isac system,” in *2023 9th International Conference on Computer and Communications (ICCC)*, 2023, pp. 659–663.
- [39] Q. Wang, R. Chai, R. Sun, R. Pu, and Q. Chen, “ISAC-enabled multi-UAV cooperative perception and trajectory optimization,” *IEEE Internet of Things Journal*, vol. 11, no. 24, pp. 40 982–40 995, 2024.
- [40] C. Dai, T. Wu, G. Sun, Y. Zuo, Z. Guo, and F. Xiao, “Joint UAV trajectory and beamforming design for RIS-aided integrated sensing and communication system,” in *IEEE INFOCOM 2024 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, 2024, pp. 01–06.
- [41] Q. Gao, R. Zhong, H. Shin, and Y. Liu, “MARL-based UAV trajectory and beamforming optimization for ISAC system,” *IEEE Internet of Things Journal*, vol. 11, no. 24, pp. 40 492–40 505, 2024.
- [42] J. Zhang and T. Ratnarajah, “Air-ground OTFS in-band-full-duplex integrated sensing and communications systems,” in *ICC 2023 - IEEE International Conference on Communications*, 2023, pp. 2637–2642.
- [43] Z. Lyu, G. Zhu, and J. Xu, “Joint maneuver and beamforming design for UAV-enabled integrated sensing and communication,” *IEEE Transactions on Wireless Communications*, vol. 22, no. 4, pp. 2424–2440, 2023.

- [44] S. Zhou, H. Yang, L. Xiang, and K. Yang, “Temporal-assisted beamforming and trajectory prediction in sensing-enabled UAV communications,” *IEEE Transactions on Communications*, early access, 2024.
- [45] W. Zhai, X. Wang, M. S. Greco, and F. Gini, “Reinforcement learning based integrated sensing and communication for automotive MIMO radar,” in *2023 IEEE Radar Conference (RadarConf23)*, 2023, pp. 1–6.
- [46] L. Xu, S. Sun, Y. D. Zhang, and A. Petropulu, “Joint antenna selection and beamforming in integrated automotive radar sensing-communications with quantized double phase shifters,” in *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp. 1–5.
- [47] G. Dulac-Arnold, R. Evans, H. van Hasselt, P. Sunehag, T. Lillicrap, J. Hunt, T. Mann, T. Weber, T. Degris, and B. Coppin, “Deep reinforcement learning in large discrete action spaces,” 2016. [Online]. Available: <https://arxiv.org/abs/1512.07679>
- [48] Y. Liu, S. Zhang, X. Li, Y. Huang, Y. Fang, and H. Cao, “Deep reinforcement learning-based beamforming design in ISAC-assisted vehicular networks,” in *2024 IEEE Wireless Communications and Networking Conference (WCNC)*, 2024, pp. 1–6.
- [49] A. Salih and S. Al-Dharrab, “Spectral efficiency and hybrid beamforming for multi-user in airborne integrated sensing and communication systems,” *IEEE Open Journal of the Communications Society*, vol. Under 2nd Review, 2024.

- [50] A. F. Molisch, V. V. Ratnam, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda, “Hybrid beamforming for massive MIMO: A survey,” *IEEE Communications Magazine*, vol. 55, no. 9, pp. 134–141, 2017.
- [51] B. Friedlander, “A sensitivity analysis of the MUSIC algorithm,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 10, pp. 1740–1751, 1990.
- [52] G. L. Stüber, *Principles of mobile communication*. Springer, 2001, vol. 2.
- [53] A. Al-Hourani, S. Kandeepan, and S. Lardner, “Optimal LAP altitude for maximum coverage,” *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 569–572, 2014.
- [54] R. Chai, X. Cui, R. Sun, D. Zhao, and Q. Chen, “Precoding and trajectory design for UAV-assisted integrated communication and sensing systems,” *IEEE Transactions on Vehicular Technology*, vol. 73, no. 9, pp. 13 151–13 163, 2024.
- [55] D. Buckingham and M. C. Valenti, “The information-outage probability of finite-length codes over AWGN channels,” in *2008 42nd Annual Conference on Information Sciences and Systems*, 2008, pp. 390–395.
- [56] S. Kay, *Fundamentals of Statistical Signal Processing: Detection theory*, ser. Fundamentals of Statistical Si. Prentice-Hall PTR, 1998.
- [57] E. Fishler, A. Haimovich, R. Blum, L. Cimini, D. Chizhik, and R. Valenzuela, “Spatial diversity in radars—models and detection performance,” *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 823–838, 2006.

- [58] F. Dong, F. Liu, Y. Cui, W. Wang, K. Han, and Z. Wang, “Sensing as a service in 6g perceptive networks: A unified framework for ISAC resource allocation,” *IEEE Transactions on Wireless Communications*, vol. 22, no. 5, pp. 3522–3536, 2023.
- [59] A. A. Khuwaja, Y. Chen, N. Zhao, M.-S. Alouini, and P. Dobbins, “A survey of channel modeling for UAV communications,” *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2804–2821, 2018.
- [60] P. Series, “Propagation data and prediction methods required for the design of terrestrial broadband radio access systems operating in a frequency range from 3 to 60 GHz,” *Recommendation ITU-R*, pp. 1410–1415, 2013.
- [61] A. Saboor, E. Vinogradov, Z. Cui, A. Al-Hourani, and S. Pollin, “A geometry-based modelling approach for the line-of-sight probability in UAV communications,” *IEEE Open Journal of the Communications Society*, vol. 5, pp. 364–378, 2024.
- [62] P. Koivumäki, G. Steinböck, and K. Haneda, “Impacts of point cloud modeling on the accuracy of ray-based multipath propagation simulations,” *IEEE Transactions on Antennas and Propagation*, vol. 69, no. 8, pp. 4737–4747, 2021.
- [63] Z. Zhang, J. Ryu, S. Subramanian, and A. Sampath, “Coverage and channel characteristics of millimeter wave band using ray tracing,” in *2015 IEEE International Conference on Communications (ICC)*, 2015, pp. 1380–1385.

- [64] P. Koivumäki, A. F. Molisch, and K. Haneda, “Line-of-sight probability in cluttered urban microcells: Analyses using poisson point process and point cloud,” *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 3, pp. 2161–2173, 2022.
- [65] M. Lecci, P. Testolina, M. Polese, M. Giordani, and M. Zorzi, “Accuracy versus complexity for mmWave ray-tracing: A full stack perspective,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 12, pp. 7826–7841, 2021.
- [66] M. Gapeyenko, D. Moltchanov, S. Andreev, and R. W. Heath, “Line-of-sight probability for mmWave-based UAV communications in 3D urban grid deployments,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 10, pp. 6566–6579, 2021.
- [67] J. Holis and P. Pechac, “Elevation dependent shadowing model for mobile communications via high altitude platforms in built-up areas,” *IEEE Transactions on Antennas and Propagation*, vol. 56, no. 4, pp. 1078–1084, 2008.
- [68] A. Al-Hourani, S. Kandeepan, and S. Lardner, “Optimal LAP altitude for maximum coverage,” *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 569–572, 2014.
- [69] I. Mohammed, I. B. Collings, and S. V. Hanly, “Line of sight probability prediction for UAV communication,” in *2021 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2021, pp. 1–6.

- [70] M. Pang, Q. Zhu, C.-X. Wang, Z. Lin, J. Liu, C. Lv, and Z. Li, “Geometry-based stochastic probability models for the LoS and NLoS paths of A2G channels under urban scenarios,” *IEEE Internet of Things Journal*, vol. 10, no. 3, pp. 2360–2372, 2023.
- [71] Q. Zhu, F. Bai, M. Pang, J. Li, W. Zhong, X. Chen, and K. Mao, “Geometry-based stochastic line-of-sight probability model for A2G channels under urban scenarios,” *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 7, pp. 5784–5794, 2022.
- [72] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal policy optimization algorithms,” 2017. [Online]. Available: <https://arxiv.org/abs/1707.06347>
- [73] J. Schulman, P. Moritz, S. Levine, M. Jordan, and P. Abbeel, “High-dimensional continuous control using generalized advantage estimation,” 2018. [Online]. Available: <https://arxiv.org/abs/1506.02438>
- [74] M. A. Gelbart, J. Snoek, and R. P. Adams, “Bayesian optimization with unknown constraints,” 2014. [Online]. Available: <https://arxiv.org/abs/1403.5607>
- [75] P. I. Frazier, “A tutorial on bayesian optimization,” 2018. [Online]. Available: <https://arxiv.org/abs/1807.02811>
- [76] A. Goldsmith, *Wireless Communications*. Cambridge university press, 2005.

VITAE

- Name: Abdalla Mohamed Abdalla Salih
- Nationality: Sudanese
- Date of Birth: 12/06/1997
- Email: *amas6176@gmail.com*
- Permanent Address: Muhammad Ibn Ahmad Al Rabae Street, 4652, 8457, Al Buhayrat District, Makkah, Saudi Arabia, 24211.
- Education: Dual Bachelor of science in Communication Engineering, University of Wales, Cardiff, United Kingdom & Modern University for Technology and Information, Cairo, Egypt, 2020
- Work Experience:
 - Administrative assistant, Interdisciplinary Research Center: Communication and Sensing Systems, King Fahd University for Petroleum and Minerals, August 2023 - June 2024.

- Teaching assistant, Electrical Engineering Department, King Fahd University for Petroleum and Minerals, September 2024 - December 2024.
- Publications:
 - A. Salih and S. Al-Dharrab, “Spectral efficiency and hybrid beamforming for multi-user in airborne integrated sensing and communication systems,” *IEEE Open Journal of the Communications Society*, Under 2nd Review, 2024.
 - A General Geometric Model for Probability of Line of Sight for UAV Communication, *Under preparation*.
 - UAV-enabled ISAC trajectory planning using PPO-based Deep Reinforcement Learning, *Under preparation*.