RECONFIGURABLE INTELLIGENT SURFACE AIDED 3D LOCALIZATION USING REINFORCEMENT LEARNING

LOCALISATION 3D ASSISTÉE PAR SURFACE INTELLIGENTE RECONFIGURABLE UTILISANT L'APPRENTISSAGE PAR RENFORCEMENT

A Thesis Submitted to the Division of Graduate Studies
of the Royal Military College of Canada
by

Horobetz J.W.

Lieutenant-Commander

In partial Fulfillment of the Requirements for the Degree of
Master of Applied Science in Electrical Engineering

April, 2026

Acknowledgements

Firstly, I would like to thank my wife for supporting me through this endeavour, especially as we forge our way through our new world of parenthood. I would also like to thank my son Zachary for bringing me happiness whenever I felt overwhelmed; his joy brought me back to Earth.

I would also like to thank Dr. Mostafa Hefnawi for providing invaluable insight into my research and all the technical staff at RMC for supporting me in the lab.

Finally, I would like to thank Mike Turpin and Colin Taylor at Major Surface Combatants (MSC) 9 for making this possible and sponsoring my studies, and I also gratefully acknowledge the support of the Directorate of Technical Airworthiness and Engineering Support (DTAES-6), which funded this research.

Abstract

Global Navigation Satellite Systems (GNSS) provide highly accurate positioning, navigation, and timing. Still, their performance degrades in contested, urban, and indoor environments due to denial and non-line-of-sight conditions. To address this challenge, Reconfigurable Intelligence Surfaces (RIS) offer a potential low-cost, energy-efficient solution for GNSS-denied and non-line-of-sight-assisted localization. RISs consist of passive reflective elements that manipulate the signal propagation path by altering the phase of an incident wave, thereby improving network coverage. These features make RISs a compelling technology for enhancing wireless systems in both civilian and defense applications. For military aircraft and Unmanned Aerial Vehicles (UAVs), RIS-based electromagnetic large apertures (ELAs) can be integrated into vehicle surfaces to significantly enhance communication reliability, target sensing, and overall situational awareness.

This work proposes using Reinforcement Learning (RL) to optimize RIS-aided three-dimensional localization via a power maximization approach. Traditional optimization and estimation approaches for RIS-aided localization suffer from high computational complexity and degrade performance in complex, changing environments. RL offers adaptability by enabling an agent to learn about the environment through interaction, reducing reliance on assumptions required for rigid traditional approaches.

This research will first compare traditional algorithms, such as Beam Sweep Grid Search (BSGS), with RL-aided RIS localization in a simulated environment. Building on these results, real-world vehicle trajectory data in a GNSS-denied environment will be used to validate the RL approach. Finally, the developed methodology will be tested in the RMC's RIS laboratory using millimeter-wave (mmWave) TMYTEK RISs.

The results of this research demonstrate the feasibility and advantages of configuring the RIS through Reinforcement Learning for accurate, low complexity, and adaptive localization in nosy, multipath and dynamic channels. The proposed RL based framework enables efficient RIS configurations using received power measurements alone, eliminating the need for complex channel state information estimation. Through simulation, real-world trajectory evaluation, and experimental validation with mmWave RIS hardware, the approach is shown to achieve reliable localization performance in GNSS-denied conditions. These results highlight the potential of RL-enabled RIS systems to support robust and secure localization in challenging environments.

Resumé

Les systèmes mondiaux de navigation par satellite (GNSS) fournissent un positionnement, une navigation et une synchronisation extrêmement précis. Cependant, leurs performances se dégradent dans les environnements contestés, urbains et intérieurs, en raison du déni et de l'absence de ligne de visée. Pour relever ce défi, les surfaces intelligentes reconfigurables (RIS) offrent une solution potentiellement peu coûteuse et économe en énergie pour la localisation assistée sans GNSS ni ligne de visée. Les RISs sont constituées d'éléments réfléchissants passifs qui manipulent le trajet de propagation d'un signal en modifiant la phase d'une onde incidente, ce qui permet d'améliorer la couverture du réseau. Ces caractéristiques font des RIS une technologie prometteuse pour améliorer les systèmes sans fil dans les applications civiles et militaires. Pour les avions militaires et les véhicules aériens sans pilote (UAV), les grandes ouvertures électromagnétiques (ELA) basées sur les RIS peuvent être intégrées à la surface de ces véhicules afin d'améliorer considérablement la fiabilité des communications, la détection des cibles et la connaissance globale de la situation.

Ce travail propose d'utiliser l'apprentissage par renforcement (RL) pour optimiser la localisation tridimensionnelle assistée par RIS à l'aide d'une approche de maximisation de la puissance. Les approches traditionnelles d'optimisation et d'estimation pour la localisation assistée par RIS souffrent d'une grande complexité computationnelle et présentent des performances dégradées dans des environnements complexes et changeants. L'apprentissage par renforcement offre une adaptabilité en permettant à un agent d'apprendre à connaître l'environnement par interaction, réduisant ainsi la dépendance aux hypothèses requises par les approches traditionnelles rigides.

Cette recherche comparera d'abord les algorithmes traditionnels, tels que la recherche par balayage de grille (BSGS), à la localisation assistée par RIS avec apprentissage par renforcement dans un environnement simulé. Sur la base de ces résultats, des données réelles sur la trajectoire des véhicules dans un environnement sans GNSS seront utilisées pour valider l'approche RL. Enfin, la méthodologie développée sera testée dans le laboratoire RIS du CMR à l'aide d'un RIS TMYTEK à ondes millimétriques (mmWave).

Les résultats de cette recherche démontrent la faisabilité et les avantages de la configuration du RIS par apprentissage par renforcement pour une localisation précise, peu complexe et adaptative dans des canaux bruyants, à trajets multiples et dynamiques. Le cadre proposé, basé sur l'apprentissage par renforcement, permet des configurations efficaces du RIS en utilisant uniquement les mesures de puissance reçue, éliminant ainsi le besoin d'une estimation complexe de l'état du canal. Grâce à des simulations, à l'évaluation de trajectoires en conditions réelles et à une validation expérimentale avec du matériel RIS mmWave, cette approche s'avère capable d'atteindre des performances de localisation fiables dans des conditions où le GNSS est indisponible. Ces résultats soulignent le potentiel des systèmes RIS basés sur l'apprentissage par renforcement pour assurer une localisation robuste et sécurisée dans des environnements difficiles.

Contents

List of Tables

List of Figures

Acronyms

Acronym	Meaning
5G	Fifth Generation
6G	Sixth Generation
AI	Artificial Intelligence
AoA	Angle of Arrival
AoD	Angle of Departure
AP	Access Point
AWGN	Additive White Gaussian Noise
B5G	Beyond 5G
BS	Base Station
BSGS	Beam Sweep Grid Search
CDF	Cumulative Distribution Function
CoA	Curvature of Arrival
CSI	Channel State Information
DDPG	Deep Deterministic Policy Gradient
DNN	Deep Neural Network
DL	Down-Link
DoA	Direction of Arrival
DRL	Deep Reinforcement Learning
ECEF	Earth-Centered, Earth Fixed
EKF	Extended Kalman Filter
ELA	Electromagnetic Large Aperture
FFT	Fast Fourier Transform
GNSS	Global Navigation Satellite Systems
HPBW	Half-Power Beamwidth
IID	Independent and Identically Distributed
IMU	Inertial Measurement Unit
INS	Inertial Navigation System
LAS	Lobe Angular Spread
LMS	Least Mean Squares
LOS	Line-of-sight
MAC	Multiple Access Channel
MDP	Markov Decision Process
MIMO	Multiple Input Multiple Output
MISO	Multiple Input Single Output
MLE	Maximum Likelihood Estimation
ML	Machine Learning
mm wave	Millimetre wave
MOP	Measure of Performance
MU-MIMO	Multi-User Multiple Input Multiple Output
NLOS	Non-line-of-sight
OFDM	Orthogonal Frequency Division Multiplexing

PBP	Posson Bipolar Process
PDF	Probability Density Function
PNT	Positioning Navigation Timing
P-CRLB	Posterior Cramer-Rao Lower Bound
PIN	Positive Intrinsic Negative
PLL	Phase Lock Loop
PNA	PNA-X Network Analyzer
PPP	Poisson Point Process
PPO	Proximal Policy Optimization
PRS	Position Reference Signal
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
RF	Radio Frequency
RIS	Reconfigurable Intelligent Surface
R-IMU	Reduced-IMU
RISS	Reduced Inertial Sensor System
RL	Reinforcement Learning
IMU	Inertia Measurement Unit
RSS	Received Signal Strength
RTT	Round Trip Time
SARSA	State-Action-Reward-State-Action
SDR	Software Defined Radio
SIMO	Single Input Multiple Output
SISO	Single Input Single Output
SNR	Signal to Noise Ratio
TDOA	Time Difference of Arrival
TOA	Time of Arrival
UAV	Unmanned Aerial Vehicle
UE	User Equipment
URA	Uniform Rectangular Array
USRP	Universal Software Radio Peripheral
V2V	Vehicle-to-Vehicle
VNA	Vector Network Analyzer

1 Introduction

1.1 Background

Global Navigation Satellite Systems (GNSS) has become ubiquitous in modern society through its provision of high-precision Positioning, Navigation, and Timing (PNT) information. However, GNSS is also prone to outages and noise errors. In contested environments where GNSS use is denied, high-precision positioning information becomes challenging. In addition, in urban areas, where high-precision positioning is crucial, GNSSs suffer from non-line-of-sight (NLOS) conditions. These geographic features create a GNSS-denied environment. Reduced Inertia Measurement Units (R-IMU) can be used as a complement, but they suffer from stochastic and systematic errors that compound over time. Even an integrated navigation system that combines GNSS and R-IMU to balance their trade-offs will still drift in GNSS-denied environments due to the inherent limitations of R-IMU [1]. To produce a high-accuracy navigational solution in a GNSS-denied environment, novel solutions must be implemented to assist in the positioning of military vehicles, aircraft, and UAVs. One promising approach is the use of Reconfigurable Intelligent Surfaces (RIS) aided localization.

Artificial Intelligence (AI) has experienced rapid adoption in recent years, offering numerous use cases for optimizing non-convex problem sets. Deep Reinforcement Learning (DRL) allows adaptability to changing environments through thorough training. Applying DRL to RIS-aided localization offers a promising opportunity to develop novel approaches for rapid search and localization of User Equipment (UE) with RIS. By incorporating positional and velocity information into DRL algorithms, intelligent, adaptive searches of the environment can be performed, enabling localization of the UE.

1.2 Problem Statement

Literature has outlined various applications of RIS in communication systems and localization. Current RIS-aided localization approaches face significant limitations that must be overcome to achieve reliable performance. Existing optimization methods deliver inaccurate position estimates in complex environments due to their sensitivity to environmental conditions. Hybrid algorithms that attempt to broach this problem incur a high computational cost, incurring more integration time in real-world applications[2], [3].

1.3 Motivation

The motivation for this thesis is to provide a means of achieving high-accuracy positional estimation in GNSS-denied, urban, and indoor environments where GNSS positioning is not practical. Reliable positioning in these environments is crucial for the safe and effective operations of civilian and military ground and air vehicles.

Emerging communication technologies, such as Reconfigurable Intelligent Surfaces, provide a new opportunity to leverage it ability to not only enhance the communication network but to be a new reliable and accurate localization method. To address the limitations of traditional localization algorithms with RIS, RL has been proposed as an alternative [3]. The adaptive nature of RL algorithms has the potential to improve positioning estimation accuracy in more complex, changing environments.

By integrating RIS with an adaptive Reinforcement Learning algorithm, it becomes possible to improve localization robustness and accuracy in complex environments.

1.4 Objectives

This research aims to develop and validate a Reinforcement Learning-based framework for Reconfigurable Intelligent Surface-aided three-dimensional localization for positioning in GNSS-denied environments. The proposed research focuses on improving localization accuracy, adaptability and RIS access efficiency.

The objectives of this thesis are:

- Develop a RIS configuration framework for RIS-aided three-dimensional localization without signal modulation.
- Design a hierarchical reinforcement learning architecture to enable efficient beam search and localization.
- Analyze the impact of RIS configuration on localization performance.
- Evaluate the proposed RIS-aided localization framework with a single and dual RIS configuration using real-world GNSS-denied vehicle trajectory data.
- Experimentally validate the proposed approach using RIS hardware, assessing localization performance in both two-dimensional and three-dimensional scenarios, including single and dual RIS configurations.

By achieving these objectives, this research aims to demonstrate that reinforcement learning provides an effective, adaptive mechanism for RIS-aided 3D localization

in GNSS-denied environments, making them suitable for deployment in GNSS denied environments.

1.5 Thesis Statement

Reinforcement Learning provides an effective, adaptive mechanism for RIS-aided 3D localization in GNSS-denied environments, where it outperforms traditional localization algorithms in accuracy, robustness and RIS access efficiency.

1.6 Methodology

The following activities will be conducted to realize RL RIS-aided localization. First, a literature review will be conducted, focusing on RIS-aided localization with DRL.

Second, after reviewing the current research on RIS localization and DRL, a methodology for applying RL in RIS-aided localization will be proposed. The proposed approach will be developed in a simulated environment, where the opportunities and challenges associated with DRL-aided RIS localization can be analysed. After confirming the validity of the proposed DRL RIS-aided localization, the algorithm will be implemented in a simulated environment that incorporates real-world vehicle trajectory data. The trajectory represents a GNSS-denied environment and requires RIS-aided localization for positioning along the vehicle path.

Lastly, the algorithm will be tested experimentally. The positional accuracy and time to localize and track will be used as measures of performance (MOP).

1.7 Thesis Organization

This thesis is organized into 6 chapters. Chapter 2 provides a literature review of RIS-aided localization. Fundamentals of RIS and RL are covered, along with a summary of current research on RIS-aided localization using reinforcement learning.

Chapter 3 describes the approach to RIS-aided localization with Beam Sweep Grid Search (BSGS) and the developed RL algorithm. The quantization of the phases of the RIS's elements is discussed and the impact on localization. Finally, a simplified environment is presented with an implementation of the RIS-aided localization with RL and BSGS approaches.

Chapter 4 presents the implementation and the results of the RIS-aided localization for positioning using a real-world trajectory.

Chapter 5 presents the laboratory setup and measurement results of RIS-aided localization.

Finally, Chapter 6 presents the findings and conclusions of this thesis and discusses future work opportunities regarding RIS-aided localization integrated with RL.

2 Literature Survey

The following section provides the background for the proposed thesis topic. First, RISs will be discussed with their proposed opportunities and challenges. Research on the use of RIS-aided localization will be presented, followed by an overview of RL and its application in RIS configuration and RIS-aided localization.

2.1 Reconfigurable Intelligent Surfaces

Reflective Intelligent Surfaces are planar devices that are composed of *M*-by-*N* meta-surface elements that work in conjunction to control the reflected propagation of an incident wave. Research into meta-surfaces has been conducted over decades to understand their properties associated with surface materials and the frequencies used [4]; however, fundamentally, they all work by changing the impedance of the meta-surface's elements. Through the control of an element's impedance, the RIS reflects the signal by changing the meta-surfaces' reflection coefficient [5]. By changing the reflective coefficient of an element's meta-surface, the incident signal's propagation path can be manipulated so that its amplitude can be changed between 0 and 1, and the incident angle can be changed between negative π and π. Thus, RIS effectively beamforms the incident signal by controlling the reflective coefficients of the RIS's elements [5]. RIS does not add any power to the incident wave but only attenuates it and shifts its phase to manipulate the propagation of the reflected wave. A RIS can be treated as a Uniform Rectangular Array (URA) while eliminating the RF chains of an URA transceiver and the associated power consumption, hardware complexity, and cost. A typical RIS is illustrated in Figure 2.1 and consists of three layers, including a controller. The outermost layer is the meta-surface composed of a dielectric substrate that can be tuned to change its reflective coefficient. The middle layer of the RIS is a copper plate on which the meta-surface is printed, and is used to minimize leakage from the meta-surface during reflections, and the inner-most layer is a control circuit board that connects all the RIS's elements together and to the controller [6].

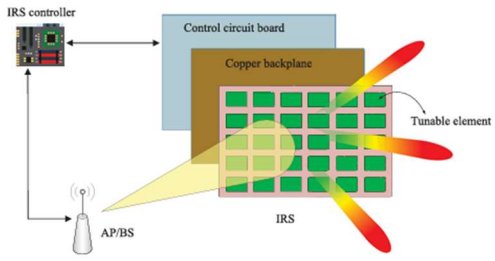


Figure 2.1: Construction of a typical RIS [6]

RISs can be considered semi-passive or passive, with semi-passive RISs containing active components used for sensing the communication channel. With a semi-passive RIS, the RIS assists in calculating the Channel State Information (CSI), while with the fully passive RIS, which does not contain any sensing ability, it is completely controlled through the network for its manipulation [6]. RIS are not co-located with the transmitter or receiver in a communication channel but deployed along its propagation path. Having a controllable device in the communication channel offers opportunities that would otherwise not be possible.

RISs must be rapidly tunable to adapt to a dynamic channel. The RIS must be able to receive control signals, or decide on its own how it must be configured to best suit the network and adapt to the environment [6]. In a semi-passive RIS, sensors measure the channel. With a fully passive RIS, the RIS receives configuration control signals via the back channel. In both cases, it is advantageous if the RIS can rapidly and continuously change the reflection coefficient of individual elements. However, recent manifestations of RIS do not reflect the aspirational desire of RIS. To change the reflection coefficient of a meta-surface element, its impedance is adjusted using an electronic device, such as a Field-Effect Transistor or a Positive-Intrinsic-Negative (PIN) diode [6]. These electronics enable rapid changes in the impedance of the meta-surfaces, but practical challenges must be overcome before the idealized RIS can be realized. Firstly, current RIS implementations have only been implemented with discrete changes in amplitude or phase. It has been shown that meta-surfaces exhibit nonlinearity in their operation. When the impedance is changed, the real component changes the amplitude, while the imaginary component changes the phase. That is to say, the resistance changes the amplitude, and the inductance or capacitance changes the phase [5]. However, in practical implementations, the non-linearity of the meta-surfaces causes changes to both amplitude and phase. Consequently, recent implementations have only seen limited changeability with the RIS, where the amplitude is only changed to zero or one, and the phase is changed to negative π or positive π [6].

What's more, mutual coupling has been identified in the operation of RIS's elements, where large numbers of elements have been constructed for RIS due to

their close spacing. When the elements are not adequately spaced, coupling between the control electronics can occur, resulting in coarse adjustability of the RIS [6].

RISs are expected to mitigate some of the challenges associated with Fifth-Generation (5G) and Sixth-Generation (6G) telecommunication [6]. Due to the frequencies employed in 5G and 6G, which are expected to operate in the millimeter-wave (mm-wave) and terahertz ranges [6], a line-of-sight (LOS) link is required between the Base Station (BS) and the User Equipment (UE). To provide the reliability, latency, and data rates required by 5G and beyond, service providers would have to deploy more BSs, Access Points (APs), and relays. With a super-dense network of BSs being prohibitively expensive in terms of upfront costs and maintenance, and with the potential for significant intercell interference due to small cell size, the use of low-cost RIS is expected to enhance network coverage and reliability [6] and has even seen recent adoption with real-world developments [7] in Europe.

With coordination across the overall telecommunication network, BS and RIS can achieve increased reliability of communication channels. If the UE is in a BS-shadowed region, the network can coordinate RIS operation to maintain seamless connectivity. [8]. In addition, network architecture coordination can improve overall channel quality. If BSs, RISs, and UEs understand the channel state and their environments, they can improve the communication channel by providing multipath signals to the UE via multiple RISs. The UE could then coherently combine the multi-path signal and improve the channel rank. Through this means, the power requirements for communication of the UE with the BS would also be reduced [8].

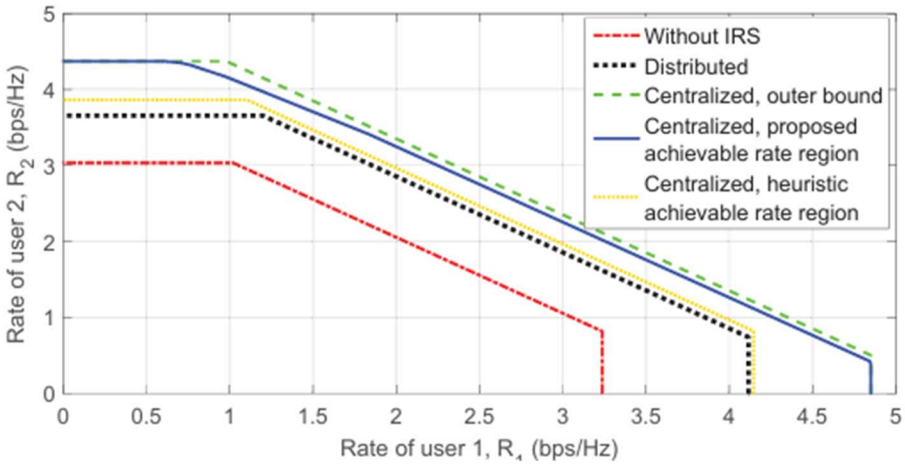


Figure 2.2: Enhanced Network Quality with RIS Employment [10]

As depicted in Figure 2.2, by properly deploying RIS in the network architecture, the cell capacity can be improved in a Multiple Access Channel (MAC) [10]. It is also shown that through a well-defined RIS deployment strategy, the power consumption for the BS and UE was reduced [10].

Because RIS can attenuate and reflect signals in a direction away from a specific location, it can be used to increase the spectral efficiency of the network. Massive MIMO (mMIMO) systems provide a higher spectral efficiency than traditional MIMO systems, but at the expense of reduced coverage area [11]. With mMIMO, as the distance between the UE and BS increases, its advantages diminish until the distance is large enough to meet the keyhole condition. In mMIMO, keyholing is the lower bound on achievable capacity due to antenna separation distances [12]. However, the use of a reconfigurable intelligent surface (RIS) enables re-beamforming of mMIMO signals, refocusing the propagated energy to extend the effective beam coverage while improving spectral efficiency by limiting signal dispersion beyond the target user equipment (UE). An RIS can also block a signal to a location. If the network detects intercell interference where edge UEs experience degraded signal quality, it could employ RISs to attenuate signals from one cell to another [8]. Not only do RISs increase a network's capacity and reliability, but they can also increase the security of the network by having precise spatial control of the RIS beam.

2.1.1 Reflective Intelligent Surface Aided Channel Model

Research in [13] indicates that, with passive RIS, the RIS dimensions must be extremely large to have an impact on the channel. This is because larger RISs can reflect more of the incident wave than smaller ones. Because of this, RISs need to be significantly larger than those of active transmit and receive arrays [14]. Despite the opportunities they offer, larger RISs face significant challenges due to their greater size. Primarily, as the size of the RIS increases, the far-field distances also increase, and thus near-field modelling needs to be investigated.

The reactive near-field is the region of the electro-magnetic field that is closest to the point source where the amplitude variations are significant, and the wavefront curvature needs to be accounted for. Beyond the reactive near-field lies the radiative near-field, where the amplitude variations stabilize, but the wavefront curvature must still be accounted for. This region occurs when the distance from the source is greater than the source's wavelength, λ. This region, commonly called the near-field, is the focus of investigation for RIS. Outside the near-field is the far-field, where the wavefront curvature is negligible [14]. The transition point between the near-field and far field is called the Rayleigh distance and is denoted by d_{FF} [14]

$$d_{FF} = \frac{2D^2}{\lambda} \tag{2.1}$$

D is the largest antenna dimension, and λ is the signal wavelength. Within the near field, there is a sub-field called the quadratic near-field, or Fresnel region, where the phase variations across the wavefront can be modelled with a quadratic function. The distance to where the quadratic near-field begins is denoted as d_{qNF}[14]:

$$d_{qNF} = \sqrt{\frac{2D^3}{3\sqrt{3}\lambda}} \approx 0.62\sqrt{\frac{D^3}{\lambda}} \tag{2.2}$$

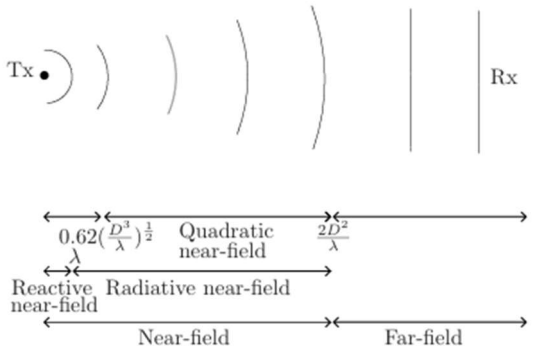


Figure 2.3: Electromagnetic field illustration [14]

In [15], it was stated that the RIS has a near field range that is dependent on both the BS RIS distance r_1, and the RIS-UE distance r_2:

$$\frac{r_1 r_2}{r_1 + r_2} < \frac{2D^2}{\lambda} \tag{2.3}$$

Understanding the electromagnetic regions for RIS is crucial for developing the appropriate RIS configuration. In their research, [14] provides a channel model of the electromagnetic regions for different RIS configurations at 5 GHz and 28 GHz, as illustrated in Figure 2.4.

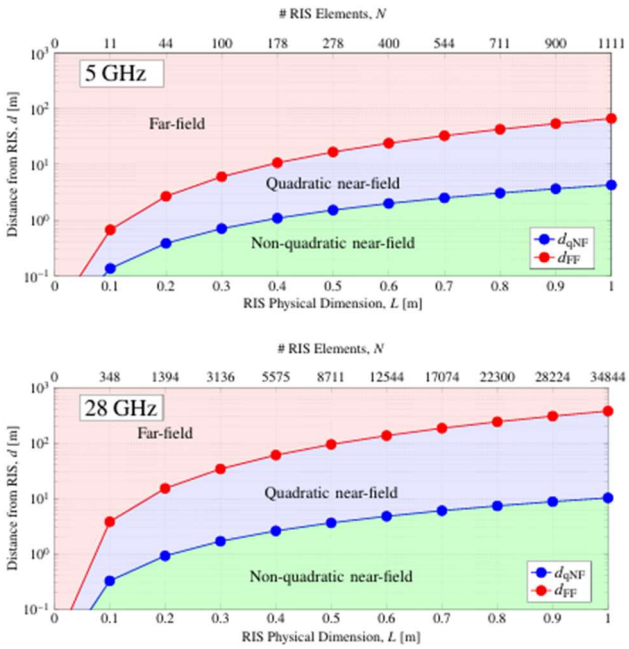


Figure 2.4: Electromagnetic regions of RIS configurations[14]

Single-reflection models have been developed for both the near-field and far-field for a single RIS along the propagation path between the BS and UE. There has also been a study on cascade channels, where multiple RIS are in a single communication channel.

For the general case, the free-space path loss of the cascaded BS–RIS–UE channel, applicable to both near-field and far-field operation, can be expressed as [16]:

$$PL = \frac{64\pi^3}{G_t G_r G d_x d_y \lambda^2} \frac{1}{\left| \sum_{m=1}^{M} \sum_{n=1}^{N} \frac{\Gamma_{n,m} \sqrt{F_{n,m}}}{r^t_{n,m} r^r_{n,m}} e^{\frac{-j2\pi(r^t_{n,m} + r^r_{n,m})}{\lambda}} \right|^2} \tag{2.4}$$

Where G_t and G_r are the transmit and receive antenna gains, with G being the scattering gain of a single RIS element, which can be expressed by [16]:

$$G = \frac{4\pi d_x d_y}{\lambda^2} \tag{2.5}$$

d_x and d_y represent the width and length of each RIS element with N by M total elements. The distances from the transmitter and receiver from each RIS element are $r^t_{n,m}$and $r^r_{n,m}$, respectively. $F_{n,m}$ is the combined power radiation pattern of the transmitter, receiver, and RIS. Γ is the reflection coefficient of each RIS element and λ is the wavelength of the signal.

Two different models have been developed for when the RIS is in the near-field. The first situation is when the RIS reflection coefficients are configured to broadcast in a desired direction. That is, the RIS reflects the incident wave in the general direction of the UE. An empirical model has been developed for this situation [16] :

$$PL_{NF} = \frac{1}{G_t G_r}\left[\frac{4\pi d_x d_y}{\lambda A}\right]^2 \tag{2.6}$$

Where A is the maximum gain of the RIS elements. In the second situation, where the optimal RIS reflection pattern is configured to focus the incident wave at the receiver, the near-field pathloss can be modelled as [16] :

$$PL = \frac{64\pi^3}{G_t G_r G d_x d_y \lambda^2 A^2 \left|\sum_{m=1}^{M}\sum_{n=1}^{N}\frac{\sqrt{F_{n,m}}}{r^t_{n,m} r^r_{n,m}}\right|^2} \tag{2.7}$$

Beyond the Rayleigh distance, the wavefront can be approximated by a plane wave, as the far-field model applies. If the RIS is configured so that the reflection wave is focused on the UE, the pathloss channel is:

$$PL = \frac{64\ {}^3(d_x d_y)^2}{G_t G_r G N M^2 d_x d_y \lambda^2 F(\theta_t \varphi_t) F(\theta_r \phi_r) A^2} \tag{2.8}$$

where $F(\theta_t \varphi_t)$ denotes the normalized radiation pattern of each RIS element toward the transmitter, where θ and ϕ represent the elevation and azimuth angles measured from the center of the RIS to the transmitter. Similarly, $F(\theta_r \phi_r)$ denotes the normalized radiation pattern toward the receiver. In this far-field model, A is assumed to be the same for all RIS elements [16].

2.1.2 Configuration of the Reflective Intelligent Surface

As indicated in the far-field and near-field channel models, the optimal configuration of the reflection coefficients is crucial for the models to be effectively used. In a known system model where the locations of the BS, UE, and RIS are known, [17] has developed a joint iterative optimization strategy that configures the RIS to the optimal reflection pattern given the system model.

While BS and RIS can be assumed to be stationary and fixed over time, the UE can be moving and at any location visible to the RIS and the BS. With current construction of RISs only having discrete phase changes, significant research has been conducted in finding the optimal reflection pattern for a RIS, given the unknown channel and discrete phase changes in the RIS' elements. The optimization of the RIS's beampattern with only discrete phase changes is subject to two constraints.

The first constraint is the unit-modulus constraint caused by the passive phase shift of the RIS [18]. To deal with the unit-modulus constraint, different methods have been developed, including semi-definite relaxation (SDR-SDP) and manifold optimization (Manopt) [19].

The second constraint is the discrete phase configuration. While it would be convenient to round the continuous reflection pattern algorithm to the nearest quantization point for a discrete RIS, [19] indicates that this can cause significant performance degradation, and when the phase bits are small, such as 1 bit, rounding may cause significantly poor performance. Much like with finding the optimal reflection pattern with continuous phase options, finding the optimal discrete reflection pattern for the system model is resource intensive and is a non-convex optimization problem with many optimization algorithms converging to local optima [19]. A Divide and Sort (DaS) algorithm proposed by [19] is able to produce a guaranteed global optimum, achieving the optimal reflection pattern while being not as computationally intensive as traditional approaches to this problem, such as extensive search.

Through their research, [19] also showed that increasing the phase bits to 4 bits enables discrete phase changes to achieve the same performance as a continuous-phase change RIS. It has also been outlined in [20] that the increase in the resolution of the phase is more significant than the increase in the resolution of the amplitude for the RIS reflection coefficients, thus most RIS only focus on the changing of the phase.

2.2 Reconfigurable Intelligent Surfaces Aided Localization

When RIS are treated as extra anchors within a network the use of both time and angle based techniques can be used to localize a UE in a network Figure 2.5 outlines the general localization scenarios for Single Input Single Output (SISO), Multiple Input Single Output (MISO), Single Input Multiple Output (SIMO), and Multiple Input Multiple Output (MIMO), which can be used with RIS. For the different scenarios the RIS can localize the UE by estimating the angles and range of the UE. Additional anchors can be added to the scenario by using a directional BS or by adding more RIS.

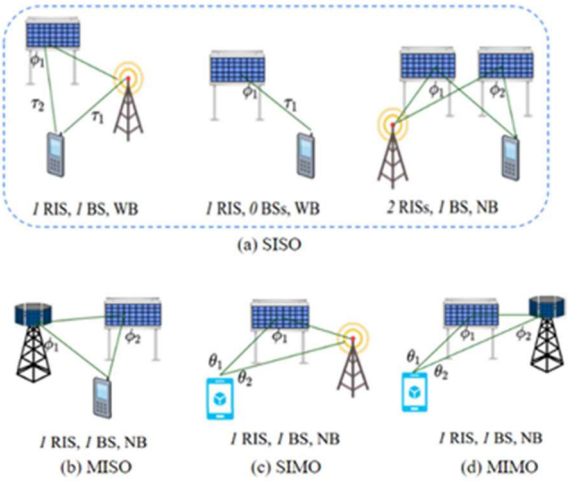


Figure 2.5: Localization scenarios [16]

The location of the UE can be estimated based on a variety of different measurements, depending on the scenario being utilized. The known azimuth and elevation angles of the RIS-reflected signal can be used to assist in localizing the UE. In 5G and beyond, Positioning Reference Signals (PRS) are used to calculate the Time of Arrival (TOA) of the PRS being sent from the BS to the UE. The time of arrival is achieved by first correlating the received signal from the LOS BS-UE channel, $g[n]$, with the known PRS signal, $x_{prs}[n]$, in the discrete time domain. The cross correlation between the received signal and the PRS is given by [21]:

$$R_{gx}[k] = \sum_n g[n] x^*_{prs}[n-k] \tag{2.9}$$

Where $(.)^*$ denotes the complex conjugate. The propagation delay is estimated by cross-correlating the received signal with the known PRS sequence. The delay corresponds to the lag at which the magnitude of the cross-correlation is maximized. Dividing this lag index by the sampling frequency f_s yields the delay estimate in seconds [21]:

$$\tau_{delay} = \frac{\arg\max_{\tau k} R_{gx}[k]}{f_s} \tag{2.10}$$

The delay is equivalent to the TOA. It is then used to calculate the range from the BS and UE by multiplying it by the speed of light [22]. By passing the PRS through the RIS, and since the range from the BS to the RIS is fixed, the range from the RIS to the UE can also be calculated, providing an accurate position estimation [16]. The accuracy of UE localization can be increased by using an additional RIS or by employing a MIMO BS, which enables the addition of BS bearing and elevation measurements to the estimated UE position.

The UE position is determined through trilateration of the ranges from the BS and the RIS. By using the bearing angle between the RIS and UE, any bearing ambiguity encountered by the trilateration can be resolved, even if bearing accuracy is low, as illustrated in Figure 2.6.

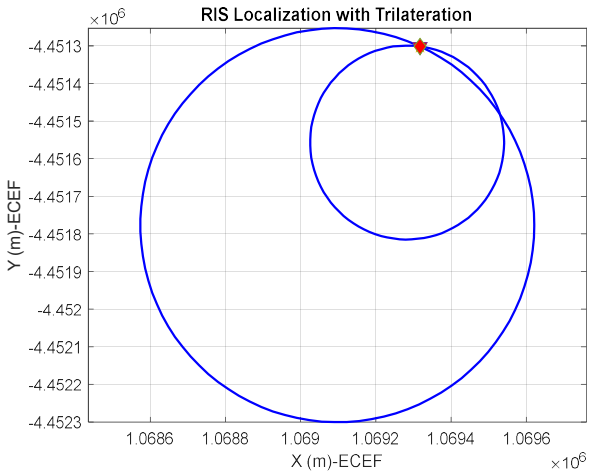


Figure 2.6: RIS Localization with Trilateration in the Earth-Centered, Earth Fixed (ECEF) Frame

In Figure 2.6, ranges are used in conjunction with the BS and RIS locations to determine the intersection points of the circles. Using a vector line, the range between the RIS and UE, along the RIS-UE bearing and centered at the RIS location, is assumed to be the closest intersection point, taken to be the UE location. This approach for localization is also used in [23], as illustrated in Figure 2.7, but in their implementation, they use multiple RIS with a single BS.

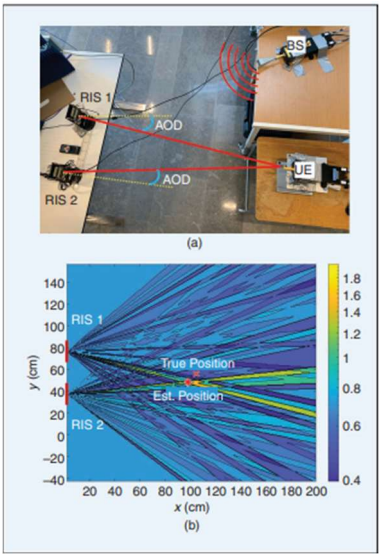


Figure 2.7: RIS Localization with Trilateration [23]

With PRS, high-accuracy, synchronized timing signals are required to achieve positional accuracy. With TOA there is uncertainty of start times of the timing signal, which causes time offsets and inaccuracy in positional calculations. Thus, the use of TOA is often limited to GNSS, which uses high-precision synchronized clock systems [8]. One method that is used to overcome the synchronization problem is to use the difference in the ToA measurements, referred as Time Difference of Arrival (TDoA). In TDoA to determine the position of a UE the difference in time arriving from different anchors is used, rather than the absolute time from one anchor. As depicted in Figure 2.8, with the use of TDoA the intersection of either hyperboles for 2D or hyperbolic surfaces for 3D localization is performed. With the TDoA localization scheme there is no requirement to detect signal transmission time as only the difference in arrival time is required. The synchronization of the clocks is required at the network anchors, reducing the clock accuracy constraints of the user. Using of TDoA and Round Trip Time (RTT) the 5G NR noise measurement vector can be modeled with errors ranging from approximately 15 centimetres to 2.5 metres [24].

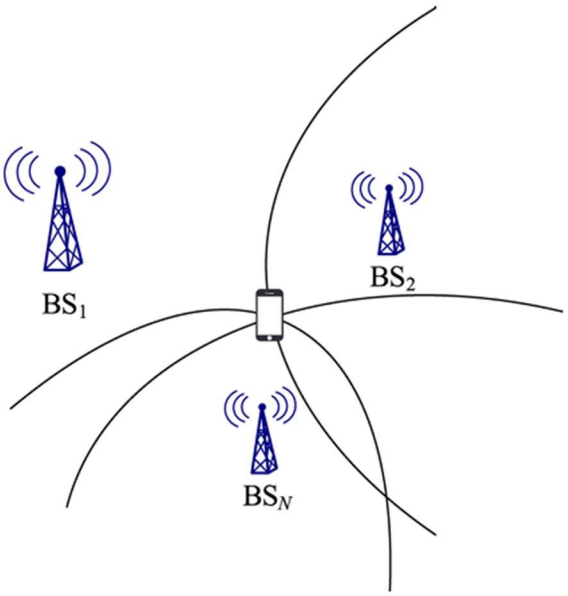


Figure 2.8: TDoA based localization [20]

Another approach to positional estimation is through power maximization. Power maximization provides a low-complexity channel estimation approach that allows the extraction of geometric information from the channel [20]. In their work, [25], [26], [27], [28] the RIS-based beamforming is used to maximize the received power at the UE. Traditional approaches that use power maximization in conjunction with RSS use a propagation model with a close-in distance with a known received power (P_0) at a reference point d_0 and channel state coefficient α [20]:

$$Pr(d) = P_0 - 10\alpha log(\frac{d}{d_0}) \tag{2.11}$$

In the general case, Angle of Arrival (AoA) and Angle of Departure (AoD) information can be used to configure the RIS phase profile so that the reflected beam is steered toward the UE, thereby enhancing the received power and supporting localization. As [28] outlines, the RIS can sweep and scan the working area of the RIS, which is defined by the beamforming limits of the RIS, and the highest received signal power from the RIS will determine the AoD from the UE to the RIS. Assuming the UE is located at the center of the beam, the UE's location can be estimated. Using the AoD, the 2-D position can be calculated as [28] and illustrated in Figure 2.9:

$$\hat{\boldsymbol{p}} = \left[x_{RIS} + \hat{d}\cos\varphi_{max}, y_{RIS} + \hat{d}\sin\varphi_{max}\right] \tag{2.12}$$

where x_{RIS} and y_{RIS} are the RIS 2-D coordinates and $\hat{d}$ is the distance between the UE and RIS.

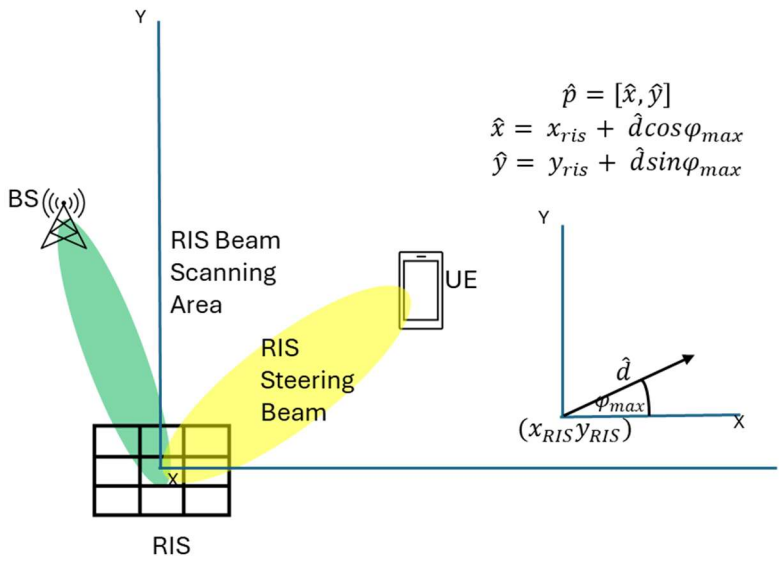


Figure 2.9: AoD-Based Localization Estimation

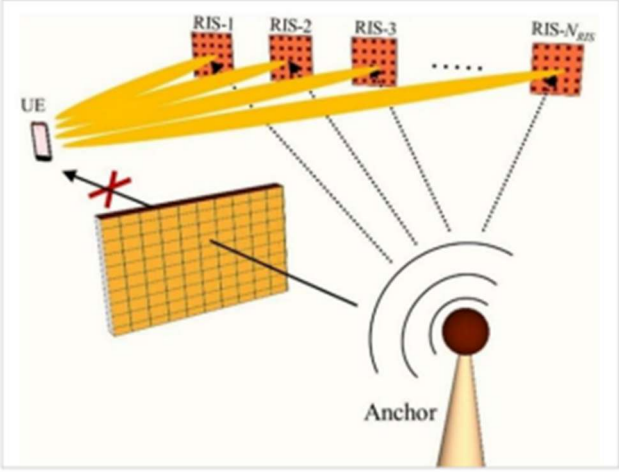


Figure 2.10: Multiple RIS localization model [28]

Using this method with multiple RIS, as depicted in Figure 2.10, resulted in higher positional accuracy [28]. However, when beam sweeping is carried out across multiple RISs, signaling overhead is increased and can lead to higher latency. As [25] states, using multiple localization methods, such as TOA and AoA estimation, increases overall positional accuracy while reducing computational overhead.

In [29], two or three RISs are used in their estimation of the UE location by extracting the geometric information from the reflection coefficient's solution set, which is due to the unique reflection coefficient for unique AoA, AoD and ranges the position of the UE could be estimated. In their work with multiple RISs, they can have effective localization in noisy channels. They note in their work that the strategic placement of the RIS is crucial for accurate localization, while adding more RIS to the network can achieve a similar effect. They also note that larger RIS assist in accurate localizationWhile [28] and [29] localize the UE at short ranges using the beamwidth of the reflected beam, [30] uses both the TDoA and AoA to localize an emitter at distances greater than 1 km. By leveraging the low beamwidth of the RIS pattern and multiple RIS, they perform sequential scanning in the desired environment and localize by triangulation, as depicted in Figure 2.11.

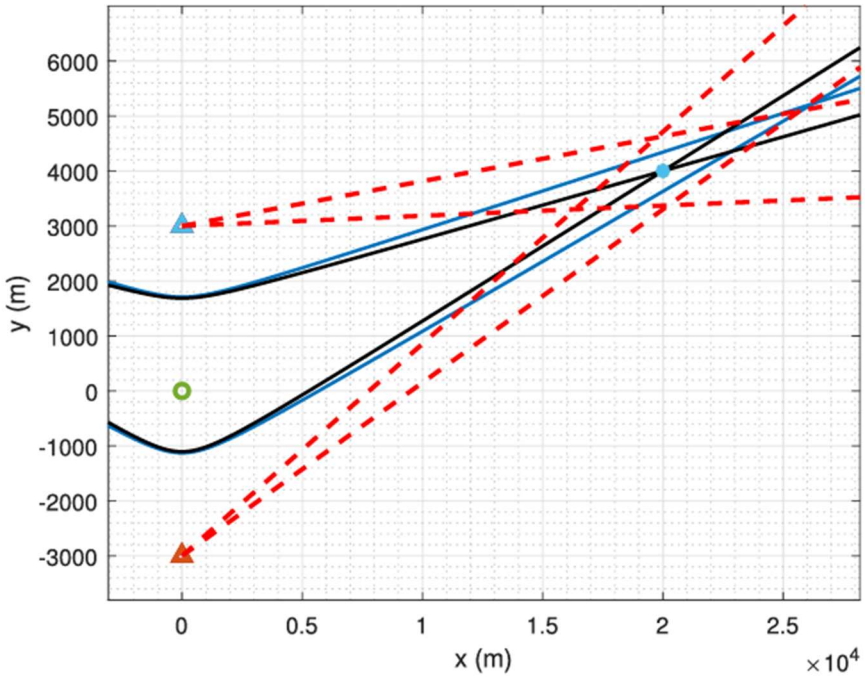


Figure 2.11: AOA and TDoA RIS-aided localization [30]

When the UE is in the near field of a RIS, the wavefront is no longer planar but spherical. By having a curved wavefront, the phase variation across the array depends on both direction and distance. Therefore, in addition to AoA or AoD, the signal also contains range-related information that can be exploited for localization. As shown in [31] a joint estimation of bearing and range can be extracted without time based measurements. This contrasts with far-field models, where only angular information can be extracted, and range must be inferred from other means, such as ToA or RSS. The principle of extracting the AoA or AoD and the range from the

curved wavefront is known as the Curvature of Arrival (CoA). By exploiting the curved wavefront, the CoA enables joint estimation of the UE's range and angle.

Although [31] considers large antenna arrays (LAAs), the same concept can be applied to CoA RIS-aided localization in the Fresnel region of the near field. In this region, the RIS reflection pattern is jointly determined by the UE angle and range. Consequently, the combination of AoA, AoD, and range yields a unique sparse reflection-coefficient solution, enabling UE localization using a single RIS.

To address the localization of a moving UE in RIS-aided localization, [24] investigated the Extended Kalman filter. In their work [24], they recognize the imprecision of TDoA with RTT when using only the BS and UE. They propose using an Extended Kalman filter with a RIS to achieve precision measurements, employing two RIS modes to measure the RTT and TDoA. Their two proposed modes of operation are reflective, which is the normal mode that all RIS use, and absorption, which does not reflect any of the signal. In their assessment, they note that the RIS provides a highly localized reflected signal in space. To reduce the noise vector of the TDoA the use of a tracker to correlate the measurements temporally is proposed with a Kalman filter. In their work [24], using real-world trajectory data, they demonstrate comparable performance to that of multicell RTT and TDoA localization. The advantage they propose for their work over multicell RTT and TDoA localization is that their method could be cheaper to deploy with RIS, due to reduced BS deployments.

To simplify the initialization of localization, so as not to do a full environment search, in their work on RIS localization, [32] assumes prior knowledge of the user's location before RIS localization. By initializing from a known position using RIS localization, higher positioning accuracy can be achieved.

RIS localization is prone to errors. When using PRS for localization, an accurate TOA calculation requires a non-corrupted PRS signal to determine the delay. When using autocorrelation to estimate the TOA, multipath fading can significantly corrupt the TOA estimate [21]. A noisy channel with uncontrolled multipath will degrade localization accuracy by interfering with LOS and RIS signals [23]. Clock drift between the BS and UE can significantly degrade localization accuracy. When clock drift causes the BS and UE to be out of sync, calculating the TOA becomes difficult. However, this can be overcome by using multiple BS signals to the UE with Time Difference of Arrival (TDOA) [21]. In addition, depending on the size of the RIS, the directivity of the propagated signal may be limited. To achieve high TOA accuracy, a high time sampling rate must also be used. The highest achievable positional accuracy with a given sample frequency is the speed of light divided by that frequency. Therefore, the use of the PRS in localization requires complex hardware and algorithms to ensure a high degree of accuracy in localization.

Beam sweeping is also susceptible to errors. The primary contributors to errors in power-based localization are Additive White Gaussian Noise (AWGN), multipath channels, and the finite limits of the beam-sweeping process.

When the noise floor is elevated by AWGN and the received power is comparable, accurate positional estimation becomes challenging. In such conditions, the received signal becomes difficult to distinguish from background noise, leading to uncertainty in the power measurement used for localization. Consequently, localization accuracy degrades, and positional estimation errors are introduced.

To minimize the effects of AWGN on RIS-aided localization, the transmitted power must be sufficiently high so that the received power exceeds the noise floor, thereby improving the SNR and enabling more reliable positional estimation.

Since RIS is envisioned for use in urban environments, multipath channels need to be modelled. A multipath channel poses a significant challenge to RIS-aided localization in urban and indoor environments, as it comprises multiple overlapping propagation paths [33]. In the context of the power maximization localization approach, this creates a distributed received power over the adjacent cells of the search grid with dead zones and hot spots throughout the grid. Research such as [33] indicates that multipath channel effects cause angle and range bins to spread into adjacent cells, potentially leading to location ambiguity.

Finally, another source of error in measured power-based localization is the limitations of the beam-sweeping codebook. The beams generated during codebook-based sweeping are only approximations of the desired patterns, leading to spatial leakage, broader lobes, and reduced angular accuracy. These imperfections cause the energy associated with a given direction to spread into adjacent angular bins, which, in turn, degrades angle and range discrimination in localization and increases the positioning error bound relative to ideal beam forms. Such beam approximation errors have been shown to degrade localization performance when realistic RIS hardware and codebook limitations are considered [34].

As detailed by [6] and [13], with fewer elements in RIS, their beamforming gain is low, and they are unable to achieve high directivity. Consequently, accurate elevation and bearing angles are not expected, resulting in low accuracy with power maximization localization. Therefore, to achieve high position accuracy, a physically large RIS with thousands of elements is required. In their work on power optimization in RIS localization, [25] outlines the power requirements of an RIS as a function of its size. They demonstrated with a sufficiently large RIS and enough power for the BS-RIS-UE channel distance, they were able achieve centimetre-level position accuracy. An idealized RIS can beamform from $[-\pi,\pi]$; however physical and manufacturing constraints limit its field of view to $[-\pi/2,\pi/2]$ or less. If localization is not achievable outside the RIS's field of view, an inaccurate position will be returned.

Additionally, the network's architecture is closely related to localization accuracy with RIS [25]. By properly configuring the network's topography, including the Base Stations, RIS, and their relationships to the UE, positioning accuracy can be ensured. As outlined in [10], the deployment of RIS throughout the network can significantly reduce outage probability and increase cell capacity. With a well-implemented RIS deployment strategy, overall coverage is improved by reducing dead zones. In RIS localization, there must be a LOS between the RIS and the UE. Therefore, without a well-organized network topology, localization accuracy cannot be guaranteed due to inadequate coverage. Supporting this, [20] outlines that increasing the number of RIS in a network is shown to increase the positional accuracy but consideration to the configuration and position of the RIS in that network must be undertaken.

RIS localization is suited for urban jungle scenarios where GNSS is unreliable. It is envisioned that RIS will be deployed throughout densely populated urban areas to meet the coverage and reliability needs of 5G and beyond [6]. Since they are already deployed for communication purposes, they can also be leveraged to assist in localization and positioning when GNSS positioning is unreliable.

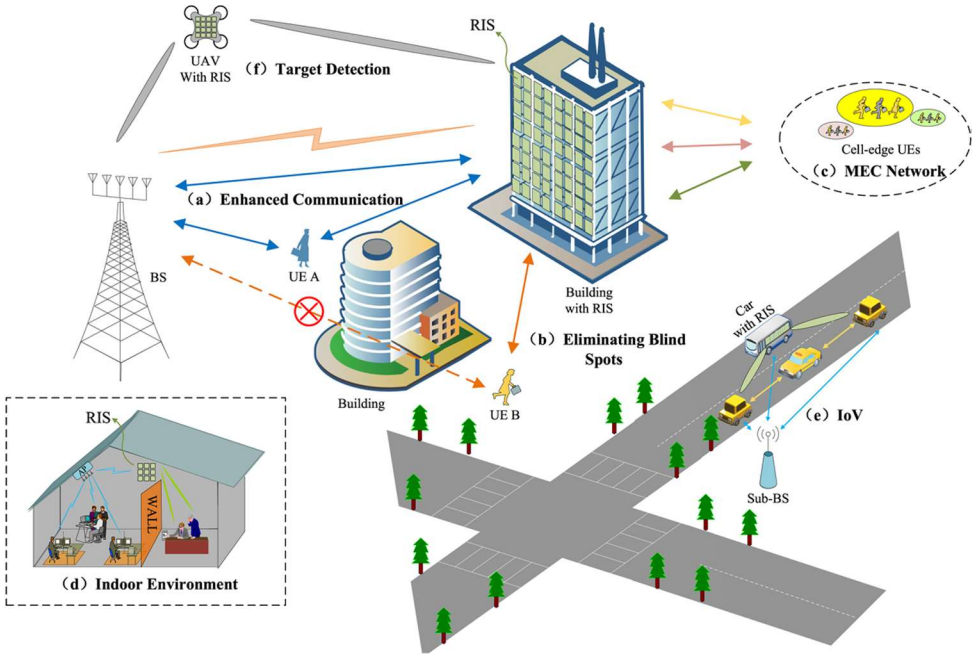


Figure 2.12: RIS uses in urban scenarios [20]

Different use cases of RIS-aided localization have been pursued in research and are illustrated in Figure 2.12. Firstly, in Figure 2.12(f), the use of Unmanned Aerial Vehicles (UAVs) has been proposed in RIS-aided localization. Both [35] and [36] have proposed using UAVs as RIS relays, acting as additional mobile anchors within the network. In their architecture, they use the PRS from both the direct and

RIS paths. Through their work, they demonstrate that their scheme can double positional accuracy compared to one that does not use a RIS deployment. The advantage of the use of a UAV in their network architecture is that the network can dynamically adjust to suit the needs of the current moment, or the UAV can be prepositioned to adjust to future predicted demand. Unlike a fixed architecture that requires significant planning and compromises for RIS placement, a mobile RIS can adjust its placement to maximize LOS probabilities and reduce outages, thereby increasing localization accuracy by strategically locating the RIS at the optimal location. However, as [36] outlines, scheduling the UAV for localization would increase network scheduling overhead, particularly for multiple users.

In their work [35], the authors propose a RIS-mounted UAV to assist in the localization of ground vehicles. They consider UAV trajectory planning, RIS phase shift optimization for communication and localization. To tackle UAV trajectory planning, they use Deep Reinforcement Learning, and for their phase-shift optimization, they use Block Coordinate Descent. Through their developed framework, they demonstrate that they are able to effectively improve the communication and localization qualities in urban blind areas [35].

Another application of RIS-aided localization is Vehicle-to-Vehicle (V2V) communication, as depicted in Figure 2.12(e). To support high-speed mmWave communication between vehicles, which suffers from LoS limitation, [37] proposes the use of RISs to assist in both localization and enhancement of the communication channel, as depicted in Figure 2.13.

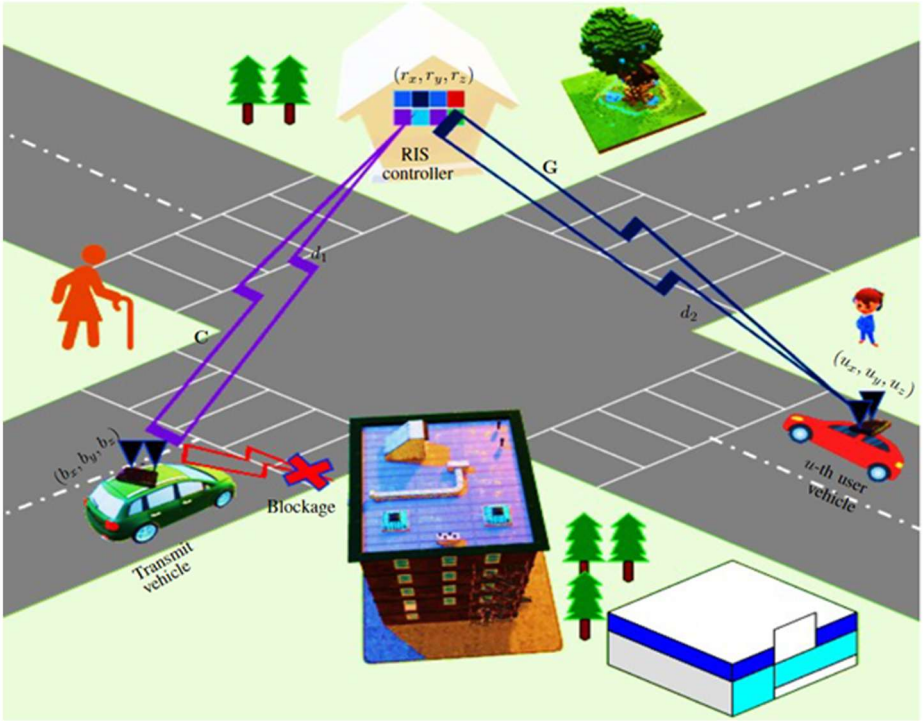


Figure 2.13: RIS-assisted V2V communication and localization model [37]

Through their work, [37] were able to significantly improve spectral efficiency in the mmWave V2V system and were able to reduce the effects of localization errors with their proposed passive beamforming method.

2.3 Reinforcement Learning

Artificial Intelligence aims to replicate the human brain and produce results that resemble human intelligence. Machine Learning (ML) is a subfield of AI that focuses on developing software that learns and improves with experience [38]. There are three main categories of ML: Supervised Learning, Unsupervised Learning, and Reinforcement Learning.

Supervised learning uses labelled datasets that humans or other methods have categorized. Unsupervised learning aims to identify inherent patterns in data without human intervention. Reinforcement Learning (RL) is a type of ML that seeks to achieve a specific goal or objective by having an agent interact with the environment through exploration and exploitation [38]. By exploring the environment, the agent receives feedback on its actions, either being rewarded or punished, and gains knowledge about its environment and how to navigate it best to maximize its reward [38]. Figure 2.14 illustrates the RL problem. The agent needs to balance its current knowledge and how to exploit it with exploration and learning.

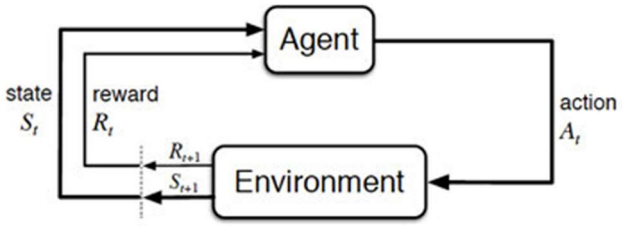


Figure 2.14: The RL Problem [38]

Crucial to RL is the ability to define the goal that the agent is trying to achieve with a Markov Decision Process (MDP). The MDP provides the roadmap of the optimal actions for each state and environment [38]. With an MDP, the roadmap is only based on the current state. It does not use past states or future states. The agent will make an optimal decision about its next steps based on its current state, then re-evaluate its current state.

To understand RL with regard to an MDP, certain terms will need to be first defined, and then further explained in terms of RL applied to RIS configuration.

The RL environment is the world in which the RL agent interacts. It is the set of rules that the agent must follow, the possible states the agent can be in, and the

rewards or punishments the agent receives for performing actions to achieve a different state [38].

In RL, states refer to the places, positions, or configurations of the environment. When an agent transitions from one state to another, it is called a visit. When an agent moves to a state that ends the process, this state is called a terminal state. When multiple visits occur, such as from the initial state to a terminal state, then this is known as an episode [38].

In the environment, the agent is allowed to take specific actions to achieve a new state. Anything the agent is allowed to do is called an action.

To provide feedback to the agent based on its performed actions in the environment, a reward or punishment is given. The amount of reward or punishment given to the agent is defined as policy. The policy dictates how the agent should act based on its current state. In addition, with the reward, the use of the Unity function as defined in (2.17) [38] balances exploration vs exploitation, in that depending on the discount factor, γ, current rewards may be favored over past rewards or may be treated equally, resulting in more or less exploration.

$$U_k = Rs_0 + \gamma \quad Rs_1 + \gamma^2 Rs_2 + \ldots + \gamma^n Rs_n \tag{2.13}$$

with R being the reward, s_n being the nth state, and γ being the discount factor.

In RL, the agent seeks to maximize its cumulative reward by transitioning to desirable states and avoiding undesirable ones, thereby maximizing its long-term reward.

There are many different algorithms employed with RL. These algorithms can be classified into two categories: model-based and model-free algorithms. With model-based RL algorithms, a model of the environment and how the agent transitions within it is created, allowing the algorithm to predict how the agent will respond to it. With a model-free algorithm, no model of the environment is learned or created; instead, the algorithm learns from observations made while interacting with the environment [38]. Model-free algorithms have experienced widespread adoption in recent years, mainly due to the increasing availability of computational power and datasets.

One widely used RL algorithm is the Deep Q-Network [38], which uses neural networks to solve complex state spaces that may be non-linear or non-convex. DQN is a model-free algorithm that seeks to maximize future rewards, using a Q-function to estimate the expected reward for each action. Traditional Q-learning algorithms use Q tables to perform this; however, in a complex state space, NNs are used instead [38]. Figure 2.15 details the DQN framework, where a neural network approximates the Q-function to map system states to optimal actions, enabling the

agent to learn decision making policies through interacting with the environment. DQN and related RL algorithms are value-based, aiming to maximize their Q-values. Rather than maximizing its Q value, other algorithms learn the states of an environment and the actions it can take in each state. The agent then seeks to maximize its reward by taking the best actions in each state. This type of algorithm is called a policy-based agent.

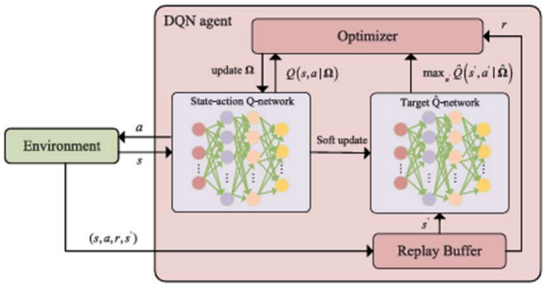


Figure 2.15: DQN Framework [39]

2.4 Reinforcement Learning in RIS-Aided Localization

Although RL offers significant opportunities for solving complex problems, considerable challenges remain. For proper training of the RL models, large amounts of data are required. If the datasets are not large enough, the RL algorithms lack sufficient space to explore and learn about the environment. Even when datasets are available, if the environment remains complex and partially observable, RL algorithms may never be able to fully model the environment and converge to a solution. Therefore, RL must be applied appropriately to a solvable problem with enough datasets.

Significant research has investigated the applicability of AI to RIS, mainly focusing on optimizing the reflection pattern to the communication channel. However, there has been relatively little research investigating the use of RL in RIS localization. Therefore, this section will outline the main research contributions to RIS-aided localization using RL.

In their work, [3] provides a novel integration of DRL with RIS for DoA estimation. Utilizing DRL with a 2-bit RIS, they demonstrated an average RMSE of less than 0.5 degrees after 32 channel samples. This shows significant improvement in DoA estimation accuracy for different scenarios.

Similar to [6] and [13], [3] also indicates that estimation accuracy is significantly reduced with a smaller RIS that has fewer elements. In their research, the DRL algorithm estimated the DoA by localizing via sequential sampling. After 32-channel samples for the DRL agent using a 2-bit RIS with 128 and 64 elements, it

achieved an RMSE of approximately 0.25 degrees for the DoA, demonstrating a high predicted angle accuracy with limited samples.

While they did not specifically focus on RIS, [40] used DRL with a DQN architecture to construct a grid of possible beam directions around a Beyond 5G (B5G) base station. Using their DQN algorithm, the BS beam steers toward the UE, achieving a 2-degree accuracy in both AoA and AoD, despite the UE's movement and blockages between the UE and BS. In their research, they use a hybrid approach to reduce overall training time for the RL algorithm. First, they train a DNN as a codebook for the BS phase shifts, then use an RL agent to optimize beam selection and fine-tune the phase shifts. Through this approach, they achieve faster convergence during training and higher accuracy in deployment compared to independent DNN and RL implementations.

Reinforcement Learning offers a promising opportunity for RIS-aided localization by adding an adaptable algorithm that can conform to a complex environment. If the RL agent is sufficiently trained, it can quickly localize the UE by configuring the RIS. The outlined literature suggests that further research into its application should be pursued, given the opportunities it presents.

2.5 Summary

As discussed throughout the literature, RIS-aided localization has shown strong potential, particularly in controlled or semi-controlled deployments where the geometry between the transmitter, RIS, and UE can be exploited.

Many works assume idealized propagation models or continuous-phase RIS elements, which simplify the analysis but become unrealistic when hardware constraints, such as 1-bit or a few-bit phase quantization, are introduced. This quantization not only introduces strong ambiguity lobes but also complicates the localization process when estimating the UE’s position.

Several approaches attempt to overcome these limitations by either reducing the complexity of the search space, such as restricting angle grids, coarse-to-fine search, or probabilistic region prediction, or by incorporating physical constraints directly into the optimization. While others rely on additional information, such as multiple RIS configurations or prior knowledge of likely UE regions, to improve robustness.

Overall, the literature suggests that RIS-aided localization is feasible, but achieving high accuracy requires effective management of measurements, to ensure the measurement overhead is reduced. Recent works have also shown that reinforcement learning can be used to adaptively select RIS configurations or measurement locations, reducing the search space while still converging toward the

correct position. The insights gained here motivate the development of an approach that not only produces accurate location estimates using minimal measurements but also integrates ambiguity handling directly into the localization algorithm.

3 Proposed RIS-Aided Localization Approach and Implementation

This chapter provides a full overview of the proposed RIS-aided localization approach and implementation. Sections 3.1 through 3.3 provides a comprehensive overview of the system model, including single- and dual-RIS configurations, and of power maximization with beam sweeping for localization. It also provides an overview of how BSGS is employed within RIS-aided localization and how BSGS will be used in our implementation. Section 3.4 describes the RL algorithm architecture, its employment, and its validation for use within our models. A key limitation of current RISs is the quantization of the phase shifts of their elements. In Section 3.5 how a RIS is configured is presented with its potential effects of localization. Finally, Section 3.6 presents a simplified RIS-aided localization environment for the employment of the presented system model.

3.1 System Model

Utilizing the SISO high-level system models from Figure 2.5 for both a single and dual RIS, a general system model for a RIS-aided communication channel is depicted in Figure 3.1. In this model, the BS is equipped with a single antenna, with the RIS equipped with *N* by *M* elements and inter-element spacing is $\Delta d = \lambda/2$. The RIS is $Q = 2^b$ configurable, such that each reflecting element can switch its phase among Q discrete values while the amplitude remains constant. In the system model, the LOS path, H_{bu}, is either weak or blocked by an obstruction. The channels from the BS to the RIS are denoted as $\boldsymbol{H}_{br}$ *and* $\boldsymbol{H}_{br}$ and the channel from the RIS to the UE is denoted as $\boldsymbol{H}_{r1u}$ *and* $\boldsymbol{H}_{r2u}$.

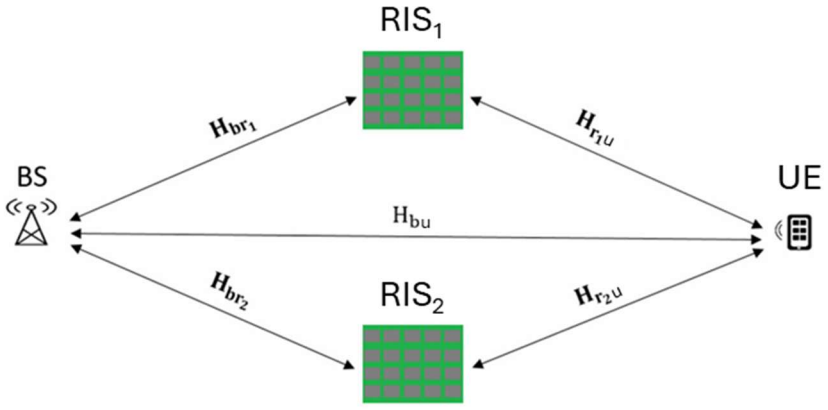


Figure 3.1: RIS aided communication channel

The geometric channel coefficients of these links are [29], [41]:

$$\boldsymbol{H}_{bu} = \rho_{bu} e^{-j2\pi f \tau_{bu}} \tag{3.1}$$

$$\boldsymbol{H}_{br1} = \rho_{br1} e^{-j2\pi f \tau_{br1}} a_{RIS}(\Psi_{br1}) \tag{3.2}$$

$$\boldsymbol{H}_{br2} = \rho_{br2} e^{-j2\pi f \ _{br2}} a_{RIS}(\Psi_{br2i}) \tag{3.3}$$

$$\boldsymbol{H}_{r1u} = \rho_{r1m} e^{-j2\pi f \tau_{r1u}} a_{RIS}(\Psi_{r1u}) \tag{3.4}$$

$$\boldsymbol{H}_{r2u} = \rho_{r2u} e^{-j2\pi f \tau_{r2u}} a_{RIS}(\Psi_{r2u}) \tag{3.5}$$

where f is the carrier frequency and $j = \sqrt{-1}$. The subindex b denotes the BS, r_i denotes the i-th RIS, and $(.)^{\mathrm{T}}$ represents the transpose operation. The channels from the BS to the $i-$th RIS are denoted as $\boldsymbol{H}_{br_i}$, and the channel from the $i-$th RIS to the UE is denoted as $\boldsymbol{H}_{r_i m}$. The free-space path loss is ρ_{il}. The AoA and AoD are expressed as, $\Psi_{bri} = \Psi_{riu} = (\theta, \phi)$ in the far-field and $\Psi_{bri} = \Psi_{riu} = (\theta, \phi, \mathrm{r})$ in the near field with θ, ϕ as the azimuth and elevation angles, respectively, and r as the distance from the user to the first RIS element. Thus, the array response vector for a uniform planar RIS with inter-element spacing d_x and d_y and element indices $m = 0, \ldots, M-1, n = 0, \ldots, N-1$ can be expressed as follows.

In the far-field, the incident wave can be expressed as a planar wave front and thus the array response vector only depends on the angles. Therefore, the far-field array response vector from the $i-$th node to $l-$th node is given by [42]:

$$a_{RIS}(\theta_{il}, \phi_{il}) = \begin{bmatrix} 1 \\ e^{\frac{j2\pi f}{c}(d_x sin\theta_{il} cos\phi_{il} + d_y sin\theta_{il} sin\phi_{il})} \\ \vdots \\ , e^{\frac{j2\pi f}{c}((M-1)d_x sin\theta_{il} cos\phi_{il} + (N-1)d_y sin\theta_{il} sin\phi_{il})} \end{bmatrix} \tag{3.6}$$

The same array response applies to both the BS-RIS and RIS-UE links. In the near field a planar wavefront cannot be assumed, and a spherical wavefront must be accounted for. Following [43] the near-field array response vector is expressed as:

$$a_{RIS}(r_{il}, \theta_{il}, \phi_{il}) = \left[1, \ldots, e^{\frac{-j2\pi f}{c}\Delta r_n^{RIS}}\right]^T \tag{3.7}$$

where in the near field within the Fresnel region, the distance difference between the reference RIS element and UE can be expressed as [43]:

$$\Delta r_n^{RIS} = -\boldsymbol{k}^T \boldsymbol{p}_n + \frac{\|\boldsymbol{p}_n\|^2 - (\boldsymbol{k}^T \boldsymbol{p}_n)^2}{2r} \tag{3.8}$$

$$\mathbf{k} = [cos\theta cos\varphi, cos\theta sin\phi, sin\theta]^T \tag{3.9}$$

$$\mathbf{p_n} = [x_n y_n z_n]^T \tag{3.10}$$

Given equations (3.6-3.10) the AoD and AoA are explicitly expressed in the formulation of the channel coefficients.

The free-space path loss, ρ_{il}, and phase shift due to the delay, τ_{il} are obtained as follows [29]:

$$\rho_{il} = \|i - l\|_2^{-\mu/2} = \Delta R_{il}^{-\mu/2}, \quad \tau_{il} = \frac{\|i-l\|_2}{c} \tag{3.11}$$

$$\Delta R, i, l \in \{b, r_1, r_2, m\}$$

$\Delta R, i, l$ is the distance between nodes i and l, and μ is the path loss exponent for the given channel. With the transmitter sending a known signal s, the received signal from the three paths is [29]:

$$y = (\boldsymbol{H}_{bm} + \boldsymbol{H}_{br1}\boldsymbol{\Omega}_1\boldsymbol{H}_{r1m} + \boldsymbol{H}_{br2}\boldsymbol{\Omega}_1\boldsymbol{H}_{r2m})s + n \tag{3.12}$$

where $\boldsymbol{\Omega}_i$ are the phase shift matrices of the *i*-th RIS and *n* is a complex additive white Gaussian noise with variance σ^2.

The phase shift matrix $\boldsymbol{\Omega}_i$ can be expressed as:

$$\boldsymbol{\Omega}_i = diag(\beta_1 e^{j\theta_1}, \ldots, \beta_N e^{j\theta_N}) \tag{3.13}$$

where β_n is the reflection coefficient of the $n-$th antenna element. It is assumed that:

$$\beta_n = e^{j\theta_n}, \theta_N = \frac{2\pi q_n}{Q}, \; q_n \in \{0, 1, \ldots, Q-1\}, \; Q = 2^b \tag{3.14}$$

where b denotes the number of quantization bits per RIS element**,** i.e., each reflecting element is controlled by a b-bit phase shifter and can realize $Q = 2^b$discrete phase states uniformly distributed over $[0,2\pi)$ while maintaining a constant amplitude.

In this model, the CSI is known between the BS and RIS but not between the RIS and the UE. This is because the BS and RIS are fixed and assumed to be known from calculations and empirical validation. The users are assumed to be low mobility and that the channel remains static during each coherence block and is independent between blocks [39]. This work focuses on the use of the CoA for localization. Therefore, it is assumed all environments are operating in the near

field. This approach was taken to align with one of the expected operating environments where UEs will be operating, which is operating near large aperture RIS in indoor and urban environments. This restrictive approach that only relies on CoA, and not incorporating a ToA into the system model is to align work in the simulated environment and the laboratory. Due to limitations during the time of this thesis, no modulation of the transmitted 28GHz signal was possible. Therefore, only power maximization is used.

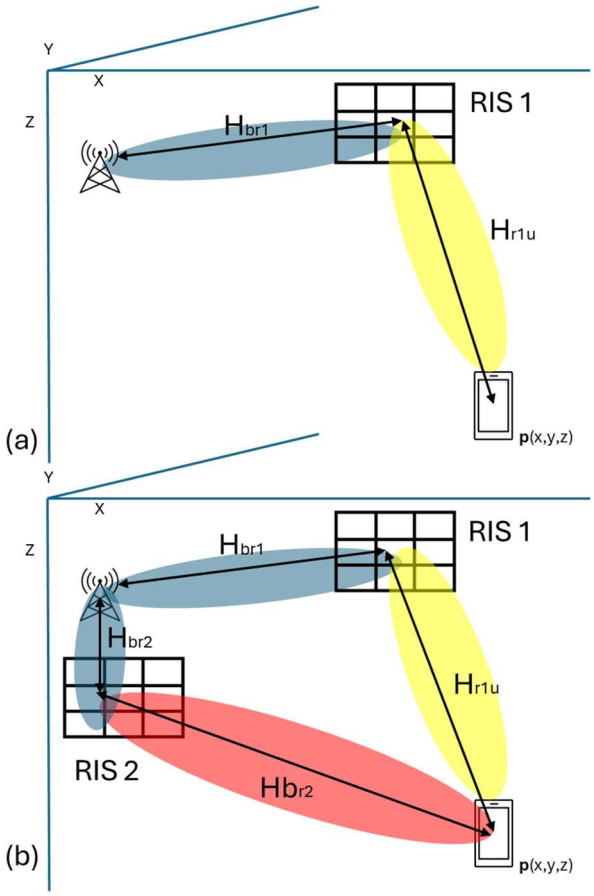


Figure 3.2: (a) Single RIS and (b) Dual RIS system models

The localization model, illustrated in Figure 3.2, consists of a BS, either one or two RISs, and a mobile UE. The coordinate system used is Cartesian, with the ECEF coordinate system for real-world trajectory data and conversion to geodetic coordinates when required.

3.2 Beam-Sweep Grid Search Localization

In our implementation, we focused on beam-sweep grid search localization in the near field as this approach relies purely on direct measurements, without computing predicted signals or evaluating statistical likelihoods, which is required in algorithms such as Maximum Likelihood Estimation (MLE). Beam sweeping significantly reduces per-grid-point computation, which is important for real-time RIS control.

In Beam-Sweep Grid Search (BSGS) localization in the near field, the environment is constructed to generate a global power map for 3D localization of the UE. Each cell of the global power map represents the received power corresponding to a hypothetical UE location. The origin of the global power map is the central grid point, designated based on the last known position. The current position is computed based on the maximum received power, expressed as:

$$\boldsymbol{p}_n = arg \max_{(x,y,z)} P_r(x, y, z) \tag{3.15}$$

When determining the maximum power, the CoA enables distinctive power measurements at each grid points in the near field. This compares to the far field scenario where only angles would be able to be distinguished, and range would be ambiguous along the angle vector. To perform the BSGS, the RIS beamforming strategy follows the angular-range-dependent models defined in equations 3.16-3.20 for the position $\boldsymbol{p}_n$. The previously estimated UE location, $\boldsymbol{p}_{ref} = (x_{ref}, y_{ref}, z_{ref})$, is used to compute the reference angles and range denoted as $\boldsymbol{p}_{ref} = (\theta_{ref}, \varphi_{ref}, r_{ref})$. This reference location serves as the central grid location in the grid search region of size G^2 with $\Delta_\theta \Delta_\varphi \Delta_r$resolutions.

Discrete steps are performed in the grid search for a 2D and 3D search defined as:

$$\theta_i = \theta_{ref} + \left(i - \frac{G+1}{2}\right)\Delta_\theta,\ i = 1,2,\ldots,G \tag{3.16}$$

$$\varphi_j = \varphi_{ref} + \left(j - \frac{G+1}{2}\right)\Delta_\varphi,\ j = 1,2,\ldots,G \tag{3.17}$$

$$r_k = r_{ref} + \left(k - \frac{G+1}{2}\right)\Delta_z,\ k = 1,2,\ldots,G \tag{3.18}$$

For a given (θ_i, φ_i) and fixed height z_{ref}, the corresponding range is obtained by geometric conversion from the polar to the Cartesian system. The 2D angular search is performed by maximizing the received power over the grid while fixing the range at r_{ref},

$$(\theta_n, \phi_n) = arg \max_{(i,j)} P_r(\theta_i, \varphi_j, z_{ref}) \tag{3.19}$$

followed by a 1D height search along the estimated angular direction,

$$(r_n) = \underset{(k)}{arg\max} P_r(\theta_n, \varphi_n, z_k) \tag{3.20}$$

The angular position is then converted to the Cartesian coordinate system for finding the new estimated UE location $\boldsymbol{p}_n$. The complete grid search procedure is outlined in Algorithm 1.

Algorithm 1 RIS beam sweeping grid search strategy

Input: Previous UE location $\boldsymbol{p}_{ref}$, Grid size G, and resolution Δ

Output: Estimate UE location $\boldsymbol{p}_n$

Initialize $P_{max} \leftarrow -\infty$

1: **for** i = 1 to G **do**

2: $\theta_i = \theta_{ref} + (i - (G+1)/2)\Delta_\theta$

3: **for** j = 1 to G **do**

4: $\varphi_j = \varphi_{ref} + (j - (G+1)/2)\Delta_\varphi$

5: Configure RIS toward $(\theta_i, \varphi_j, z_{ref})$

6: Measure $P_r(\theta_i, \varphi_j, z_{ref})$

7: **if** $P_r > P_{max}$ **then**

8: $P_{max} \leftarrow P_r$

9: $(\theta_n, \varphi_n) \leftarrow (\theta_i, \varphi_j)$

10: **end if**

11: **end for**

12: **end for**

13: for k = 1 to G **do**

14: $z_k = z_{ref} + (k - (G+1)/2)\Delta_z$

15: Configure RIS toward $(\theta_n, \varphi_n, z_k)$

16: Measure $P_r(\theta_n, \varphi_n, z_k)$

17: **if** $P_r > P_{max}$ **then**

18: $P_{max} \leftarrow P_r$

19: $z_n \leftarrow z_k$

20: **end if**

21: **end for**

22: Return $(\theta_n, \varphi_n, z_n)$

In the near-field (NF), the RIS can focus towards a *point*, sweeping (θ, ϕ) and then sweeping z effectively localizing a 3D point. Once the beam sweep returns $(\theta_n, \phi_n)$and the z-sweep returns z_n, the range can be computed along the ray by intersecting the ray with the plane $z = z_n$.

In the far-field (FF), beam sweeping gives a direction but not the range. With only one RIS and BS pair and only RSS, the FF generally cannot uniquely determine distance along the ray. Many points along the same direction can produce similar maxima. So, for FF, you need a ToA-based range estimate, multi-RIS or multi-BS triangulation.

3.3 Reinforcement Learning Developed Architecture

The approach for integrating an RL algorithm into RIS-aided localization combines a supervised regression model with a management RL agent to achieve fast, accurate current location predictions, as detailed in Figure 3.3 and Figure 3.4. The objective of the RL agent is to predict the next UE position based on its previous location and velocity, minimizing the number of localization measurements required.

The RL algorithm that was developed receives the previous known position and velocity. It uses that information in a DNN trained on kinematics to predict a search area for where the UE may be. The agent then performs a grid search of the area, where each grid location corresponds to a beam sweep location of the RIS. The RL agent does not directly control the RIS; instead, it provides the RIS with sweep parameters, which then performs beamforming. The agent provides the grid locations, which are used to calculate the RIS's azimuth, elevation, and range.

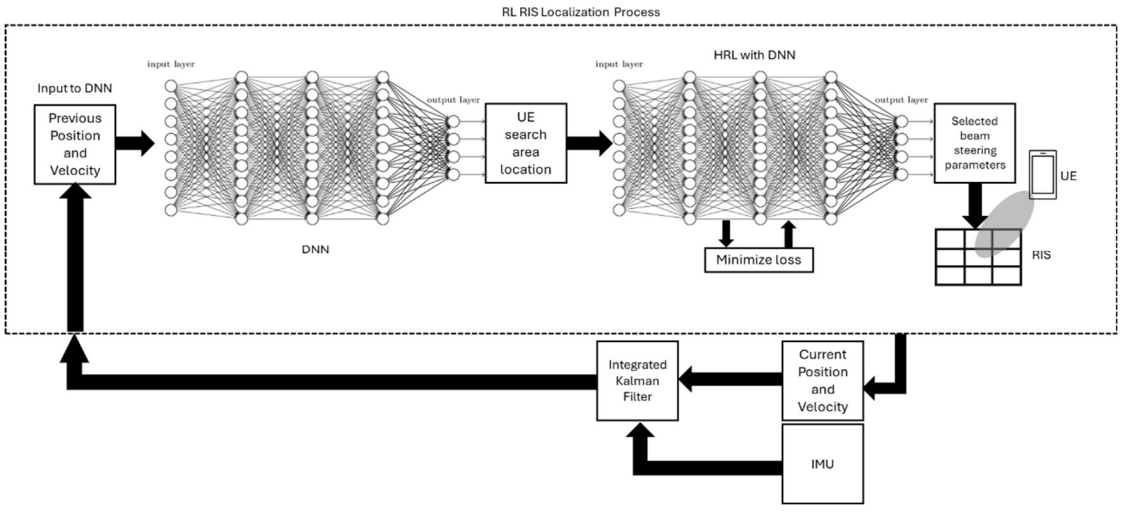


Figure 3.3: RIS-aided localization with RL and Kalman filter

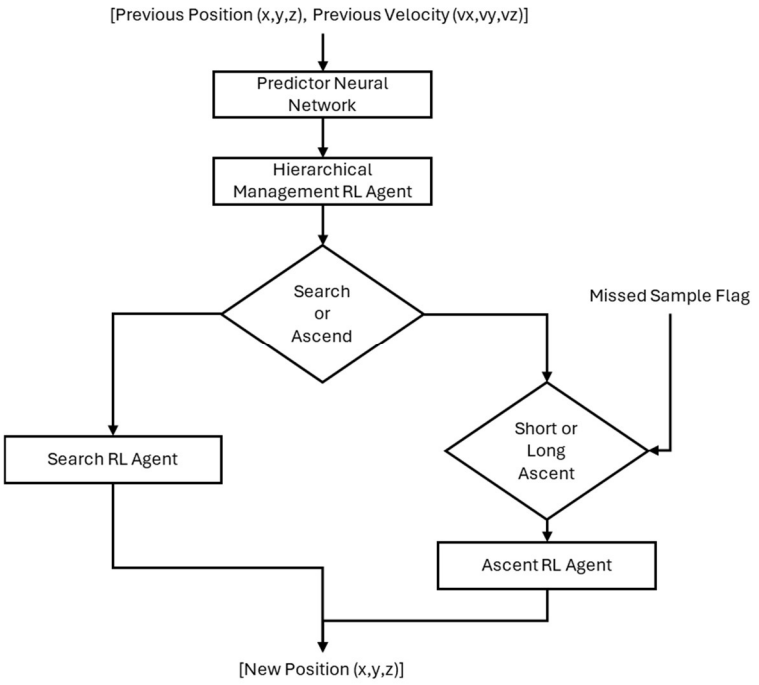


Figure 3.4: RL agent architecture flow chart

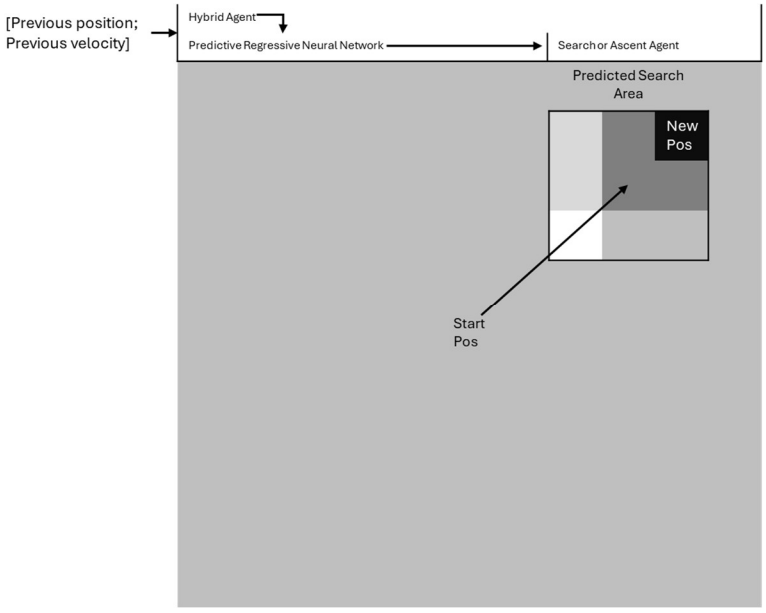


Figure 3.5: Graphical representation of the RL actions

With RIS-aided localization, it is anticipated that due to the LoS requirements, there may be instances of blockages. There will be instances where a sample will not be achievable, and thus the RL agent will have to adapt. Therefore, training on predicting 2 samples ahead was incorporated into the agent's environment and training. Due to limitations in the used dataset, only one missing sample will be

predicted at a time. However, with a higher-sampling-frequency dataset, further predictions could be trained in future work.

Table 3.1: Hierarchical RL Agent breakdown

Category	Hybrid UE Prediction Agent Description
Environment	A 31 × 31 2-D spatial grid representing discrete UE positions. Each episode evaluates whether the agent can correctly predict the next UE location based on a single velocity input.
State Representation	A continuous 4 × 1 vector: [vx; vy; dx; dy], where (vx, vy) is the UE velocity and (dx, dy) is the regression model's coarse predicted offset.
Observation Space	A continuous 4 × 1 vector normalized to [-1, 1], ensuring stable learning and reducing scaling differences between velocity and predicted offset.
Action Space	Continuous 2 × 1 offset**:** [dx_residual; dy_residual], representing a small correction added on top of the regression prediction.
Reward Function	Reward = 1 − (distance / maxDistance) the agent receives higher reward the closer its corrected prediction is to the true next UE position.
Episode Termination	A single-step episode
Reset Policy	UE starts at the grid center.
Agent Type	DDPG, chosen because the residual offset action is continuous.
Function Approximator	Actor: 2 hidden layers (64 neurons each), ReLU activations, outputs residual correction. Critic: Separate state and action processing paths merged to estimate Q-values for continuous actions.
Exploration Strategy	Standard DDPG exploration: deterministic policy with noise added during training to explore different residual corrections.
Discount Factor	$\gamma = 0.99$
Training Episodes	2000, because the task is single-step and the regression model already provides a strong baseline.
Max Steps per Episode	1, to predict next position and evaluate
Stopping Criterion	Training stops if average reward ≥ 0.95, small prediction error.
Training Output	A trained hybrid agent consisting of the regression model and DDPG agent. The RL agent outputs a correction that improves the regression prediction.
Validation	Tested on multiple UE velocity samples to compute the mean prediction error between predicted and true next positions.

This hybrid UE prediction agent combines a supervised regression model with a reinforcement learning residual-correction strategy. The regression model provides a coarse estimate of the UE's next position based on the previous known velocity, while the DDPG agent learns to refine that prediction by applying small continuous adjustments that minimize error. Each episode is only one step long, so the agent learns directly how its correction affects prediction accuracy.

Table 3.2: Search and ascent RL agent breakdown

Category	**Ascent Agent (Peak-Finding)**	**Search Agent (Grid Exploration + Max-Finding)**
Environment	A 2-D grid world (9 × 9 or 5 × 5) where the goal is to reach a single peak cell.	A (3 × 3) grid world where the agent must explore all cells and ultimately move to the max value cell.
State Representation	A finite state index representing the combination of agent position and peak location. Total states = 6561 (for 9 × 9 grid).	A continuous [row; col] representation of the agent's location, optionally extended with a visited cell mask and a max-value tracker.
Observation Space	All discrete agent/peak position pairs.	Agent's continuous position.
Action Space	8 discrete movement actions.	8 discrete movement actions, with optional stay action for fine control.
Reward Function	Sparse reward: +1 for reaching the peak; −0.01 otherwise.	Dense reward structure: strong incentives for exploration and discovering high value cells, first-visit bonus, with penalties for revisits, invalid moves, and unnecessary steps.
Episode Termination	Episode ends when the agent reaches the peak or hits the step limit (30 steps).	Episode ends when the agent visits all cells and ends on the max value cell, or when the step limit is reached (15 steps).
Reset Policy	Agent always starts at the center. Peak positions are randomized each episode to ensure generalization.	Agent always starts at the center [2,2]; grid values and the max-value location is fixed.
Agent Type	Tabular SARSA.	Deep Q-Network (DQN).
Function Approximator	A tabular Q-table mapping each state–action pair to a Q-value.	A neural network with two 16-neuron ReLU layers to approximate Q-values over continuous states.
Exploration Strategy	ε-greedy with initial ε = 1.0 and slow decay (0.001), supporting broad exploration due to large initial state space.	ε-greedy with ε = 1.0, ε_min = 0.05, and decay = 0.9975, allowing exploration early and stable exploitation in later training.
Discount Factor	γ = 0.95	γ = 0.95
Training Episodes	6000	1700
Max Steps per Episode	30	15
Stopping Criterion	Training stops when the 30-episode moving average reward exceeds 0.98.	Training stops when the 10 episode moving average reward exceeds 88.
Training Output	A learned Q-table for all possible agent/ peak configurations.	A trained DQN policy capable of efficiently exploring the grid and navigating to the max value cell.
Validation	Multiple rollout trajectories plotted to verify consistent peak seeking without exploration noise.	Rollouts plotted to confirm complete exploration and correct navigation toward the max value cell.

The ascent agent and search agent are built for two different behaviors. The ascent agent is a simple peak-seeking controller that operates in a fully discrete state space and uses tabular SARSA, so its policy is easy to interpret and depends only on on-policy updates from sparse rewards. In contrast, the search agent performs full grid scanning and max-value detection. It uses a neural network to approximate the action–value function. This requires denser feedback and a more extensive model to generalize across states.

3.3.1 Reinforcement Learning Agent Validation

To ensure the proper functioning of the developed RL agents, they must be validated in a simulated environment and tested across multiple trials. The ascent and search agents were tested across 100 trials, and their results are summarized below.

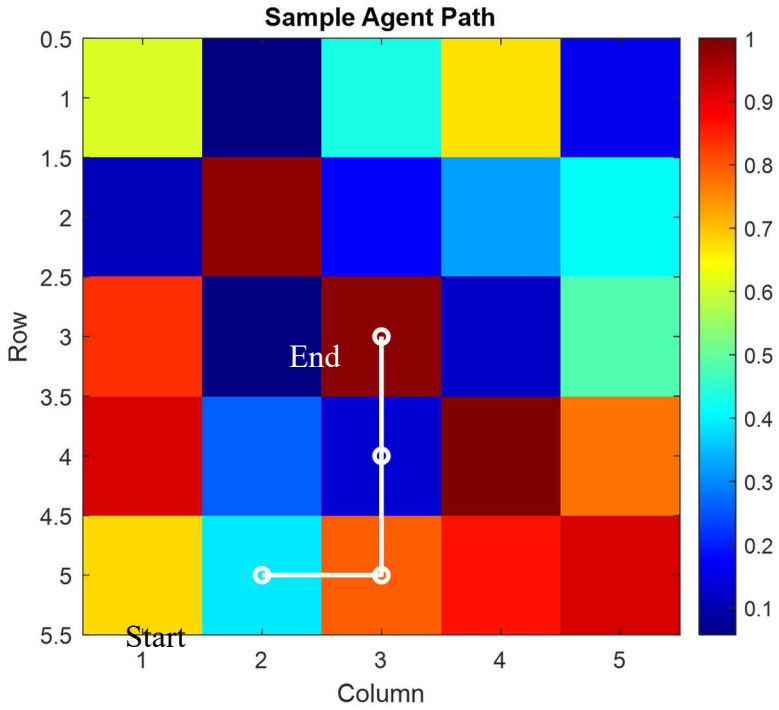


Figure 3.6: Ascent RL path

Table 3.3: Ascent agent validation trials

Total Trials	Success Rate	Average Steps
100	96.00%	4.43

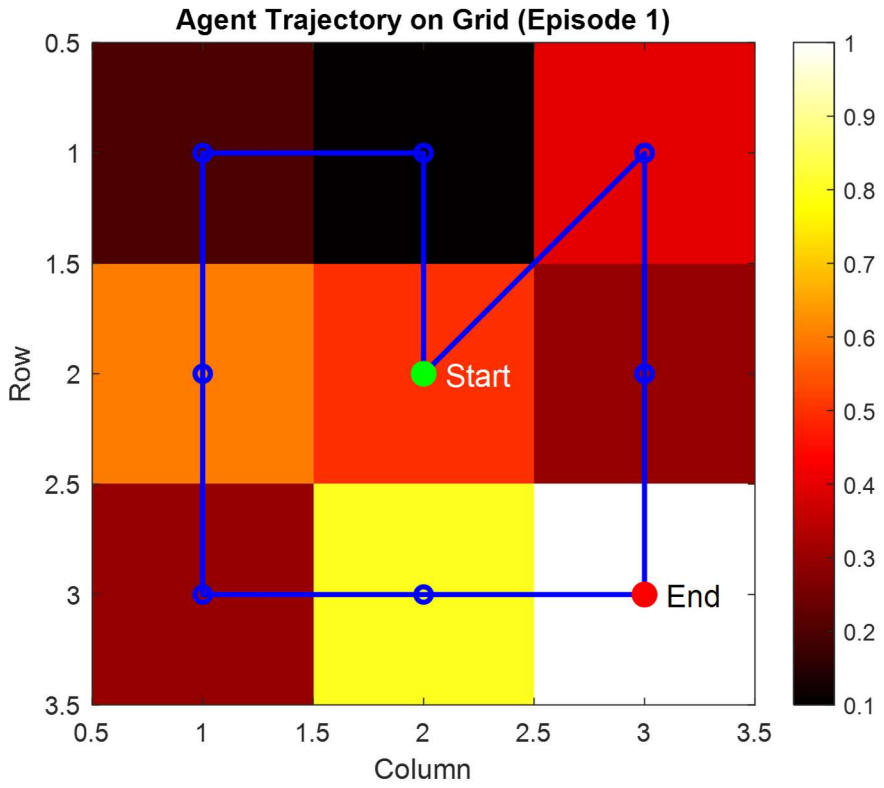


Figure 3.7: Search RL path

Table 3.4: Search agent validation trials

Total Trials	Success Rate	Average Steps
100	100%	11.75

Through the validation of the ascent and search agents it can be seen that the ascent agent is able to find the maximum in less steps but is not always able to find the maximum, with only 96% success rate. Conversely, the search agent is able to find the maximum 100% of the time but takes more steps and thus will incur more use time with the RIS.

3.4 RIS Configuration

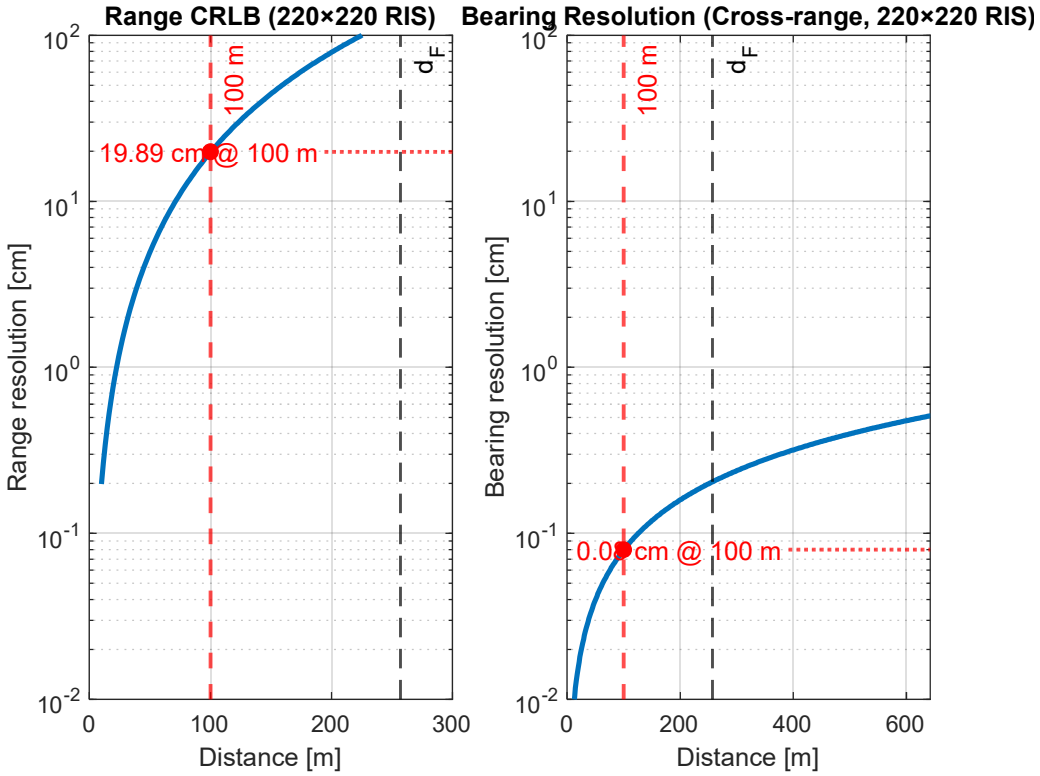
Configuring the RIS for the channel is a significant challenge, and many approaches have been proposed. In this work, the channel modeling framework described in [16] is adopted, and the RIS reflection patterns are obtained using a codebook-based beam-sweeping approach.

This work assumes that the UE is operating in the near field of the RIS. In their work [44], [45] state that the array response, the reflection pattern of the RIS, is directly related to not only the azimuth and elevation angles, but also the range of the UE from the RIS. Therefore, 3D localization is possible with strictly one RIS and a UE operating in the near field [44]. This work assumes that the reflection pattern is generated and encoded with angles and range information. Since the reflection pattern encodes the azimuth, elevation, and range information of the UE, the theoretical lower bound on localization accuracy must be determined, and this can be done with the Cramér-Rao lower bound. As previously stated, an RIS can be treated as a UPA or URA. The posterior Cramér-Rao lower bound (P-CRLB) for a URA in the near field of a UE is well defined in [31]. Using the derived P-CRLB, the theoretical accuracy estimates for the defined system can be determined. Using both a 220-by-220 for the simulated environment and a 32-by-32 for the laboratory, the RIS can be modelled and bearing and range resolutions calculated. The RIS is modelled as a URA at 100m for the 220-by-220 RIS, and 1m for the 32-by-32 RIS. The bearing and range resolutions are calculated using equations 48 and 49 from [31]. The resulting model is presented in Figure 3.8.

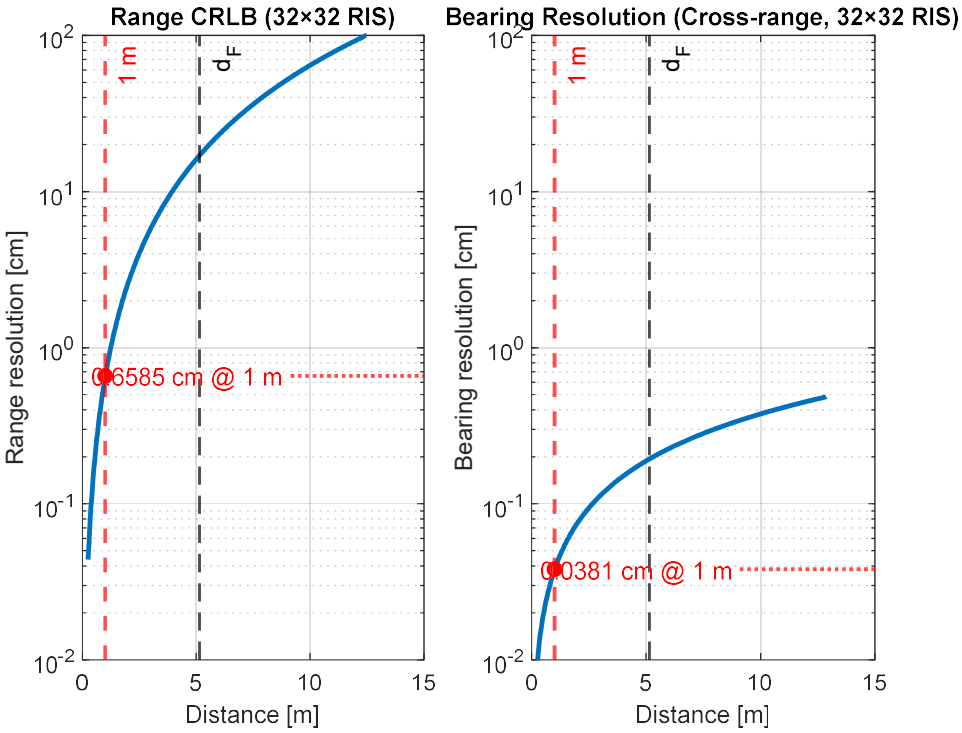
The derived range and bearing resolutions of the 220-by-220 RIS is approximately 20cm and sub-1 cm, respectively. With the 32-by-32 RIS the accuracy at 1m is sub-centimetre, illustrating that centimetre-level accuracy is possible with the outlined system model.

Ideal RIS would use continuous amplitude and phase changes to configure the RIS to suite the channel. However, as detailed in Section 2.1.2, current RIS do not have this flexibility and thus suffer significant restrictions in their ability to change their amplitude and phase. As stated in [20], the ability to change the phase provides the most advantage to beamform; thus, only phase is quantized in current RIS with only limited quantization levels. Only having 1 or 2 bits severely limits the ability of the RIS to change its phase and further analysis needs to be conducted to determine the effects on localization accuracy.

Quantization of the phase has a variety of effects on the beam and consequently may affect the effectiveness of RIS-aided localization. The chief effects of phase quantization has on RIS beamforming are: beamforming gain, main beam beamwidth (HPBW), grating lobe suppression, and beam pointing errors [46].



(a)



(b)

Figure 3.8: P-CRLB range and bearing resolution: (a) 220-by-220 RIS at 100m, (b) 32-by-32 RIS at 1m

The derived range and bearing resolutions of the 220-by-220 RIS is approximately 20cm and sub-1 cm, respectively. With the 32-by-32 RIS the accuracy at 1m is sub-centimetre, illustrating that centimetre-level accuracy is possible with the outlined system model.

Ideal RIS would use continuous amplitude and phase changes to configure the RIS to suite the channel. However, as detailed in Section 2.1.2, current RIS do not have this flexibility and thus suffer significant restrictions in their ability to change their amplitude and phase. As stated in [20], the ability to change the phase provides the most advantage to beamform; thus, only phase is quantized in current RIS with only limited quantization levels. Only having 1 or 2 bits severely limits the ability of the RIS to change its phase and further analysis needs to be conducted to determine the effects on localization accuracy.

Quantization of the phase has a variety of effects on the beam and consequently may affect the effectiveness of RIS-aided localization. The chief effects of phase quantization has on RIS beamforming are: beamforming gain, main beam beamwidth (HPBW), grating lobe suppression, and beam pointing errors [46].

3.4.1 Phase quantization effects on beamforming gain

In their analysis of the effect of RIS phase quantization on the received power in a BS-RIS-UE channel, [46] found a minimal effect, as depicted in Figure 3.9.

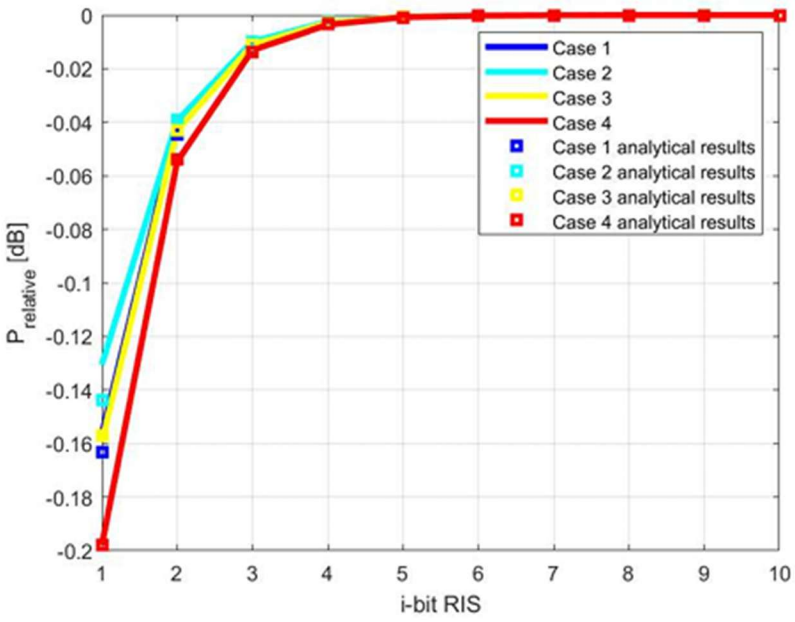
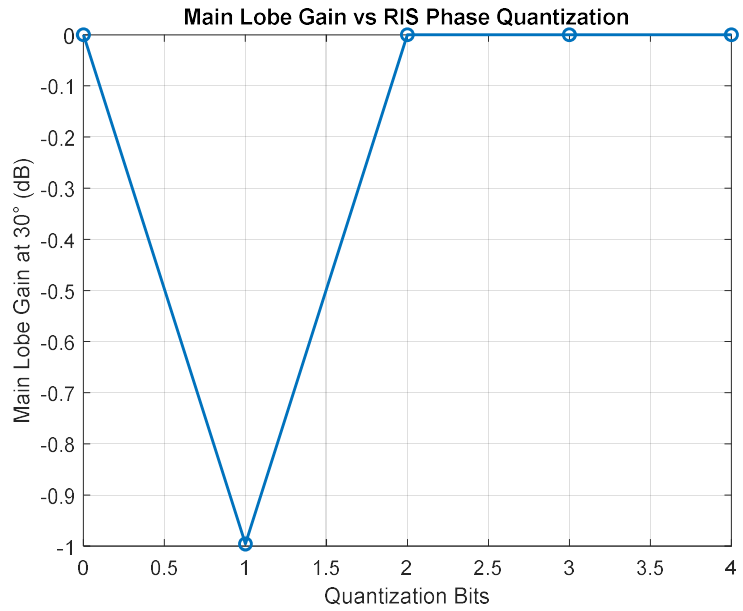
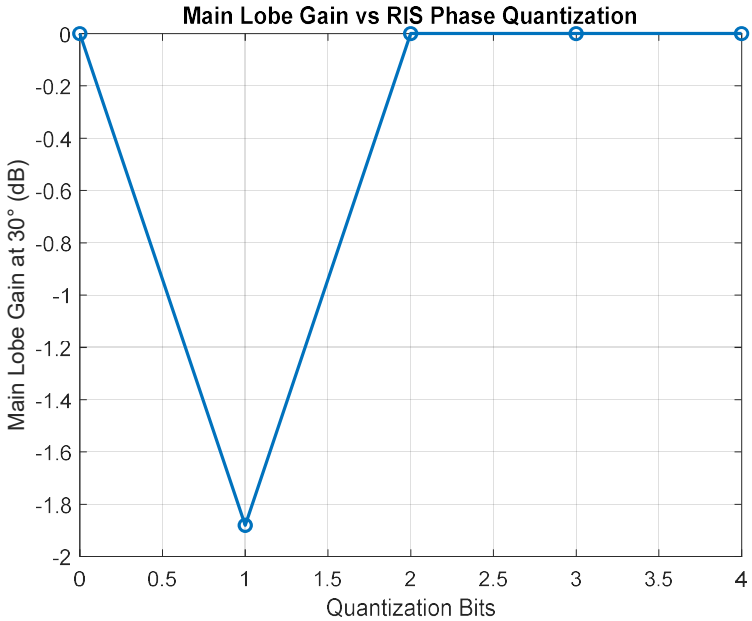


Figure 3.9: Relative received power effects due to quantization [46]

In their results, they show that even with a 1-bit quantization of the phase, the relative difference between the received power is less than 0.2 dB, and with 4-bit quantization, less than 0.1 power decrement can be achieved. Applying the same principle to our simulations, we are able to reproduce similar results as depicted in Figure 3.10 and Table 3.5 using an RIS with 220 by 220 elements. In our results, a continuous phase shift is classified as 0-bit quantization as a baseline comparison.



(a)



(b)

Figure 3.10: Effects of RIS phase quantization on received power for (a) 220-by-220 RIS and (b) 32-by-32 RIS

Table 3.5: Effects of RIS phase quantization on received power

Bits	**1**	**2**	**3**	**4**
Relative Gain (dB) 220-by-220	-1	0	0	0
Relative Gain (dB) 32-by-32	-1.8	0	0	0

In our results, we also see that when quantizing the phase of the RIS, we only see approximately -2 dB change in the main lobe gain with a continuous and 1-bit phase change, with minimal change for more bits. Therefore, it can be concluded that phase quantization will not significantly affect the received power.

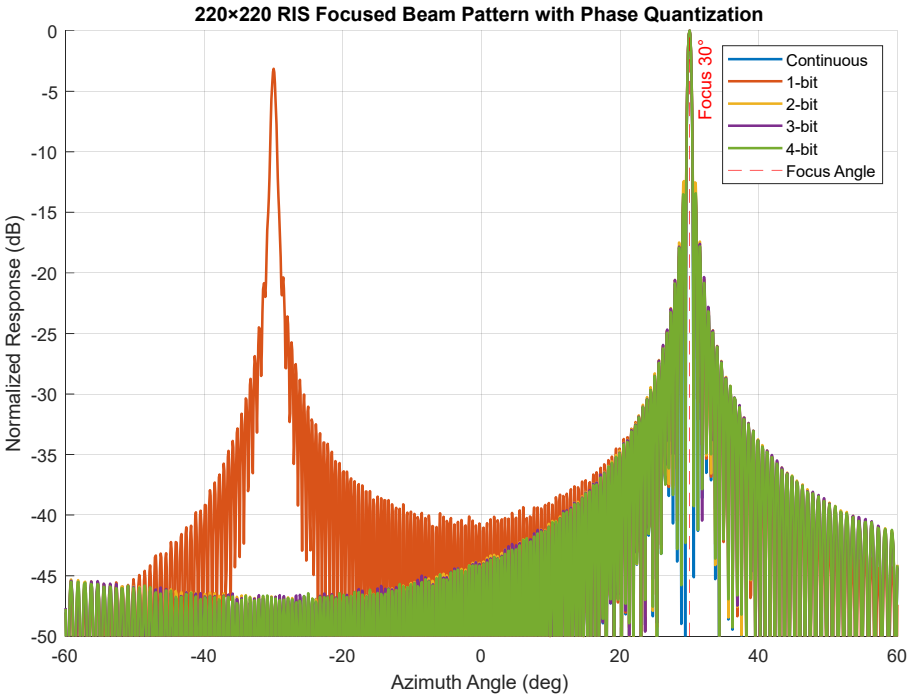
3.4.2 Main beamwidth (HPBW)

In their analysis of the effects of phase quantization on the HPBW of the main beam of a RIS, [46] were able to show that when comparing a 1-bit quantization to a continuous phase, the HPBW were comparable as depicted in Table 3.6.

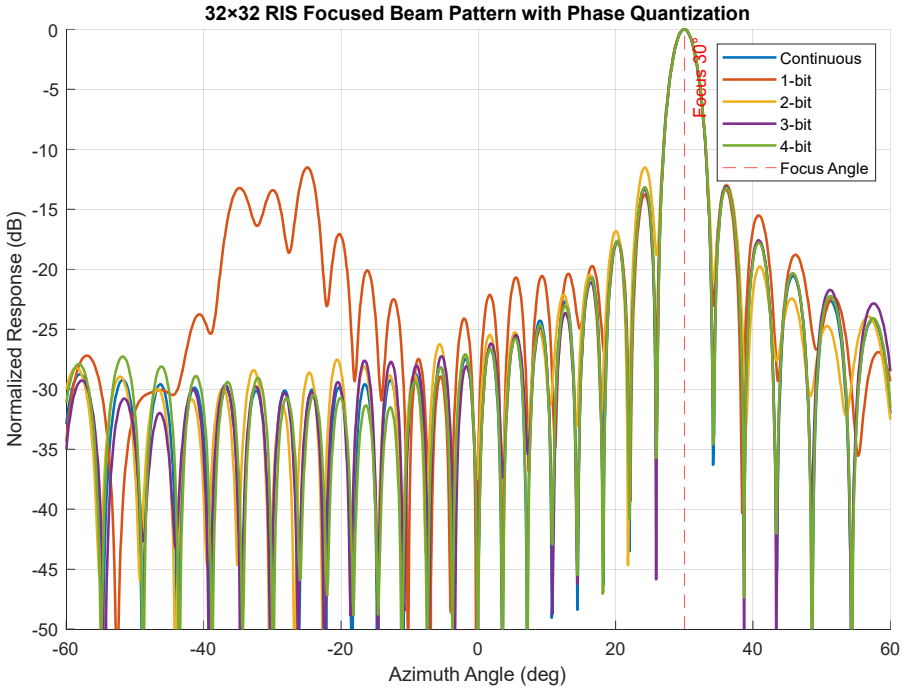
Table 3.6: Effects of RIS phase quantization on beamwidth[46]

	Scenario 1	**Scenario 2**	**Scenario 3**	**Scenario 4**
Continuous	5.6°	4.8°	5.6°	9.8°
1 Bit	5.6°	4.9°	5.6°	9.8°

Applying the same methodology to a RIS with 220-by-220 elements and a RIS with 32-by-32 elements yielded similar results, as depicted in Figure 3.11 and Table 3.7.



(a)



(b)

Figure 3.11: Effects of RIS quantization on beamwidth

Table 3.7: Effects of RIS phase quantization on beamwidth

Bits	Continuous	1	2	3	4
3 dB Beamwidth at 110m (degrees) for 220-by-220 RIS	0.6	0.6	0.6	0.6	0.6
3 dB Beamwidth at 1m (degrees) for 32-by-32 RIS	3.8	3.8	3.8	3.8	3.8

To verify our simulated results, it can be compared to the provided TMYTEK test results for their RCS gain, provided in Figure 3.12. Our results align closely with those provided for the 32-by-32 RIS with 1 bit quantization.

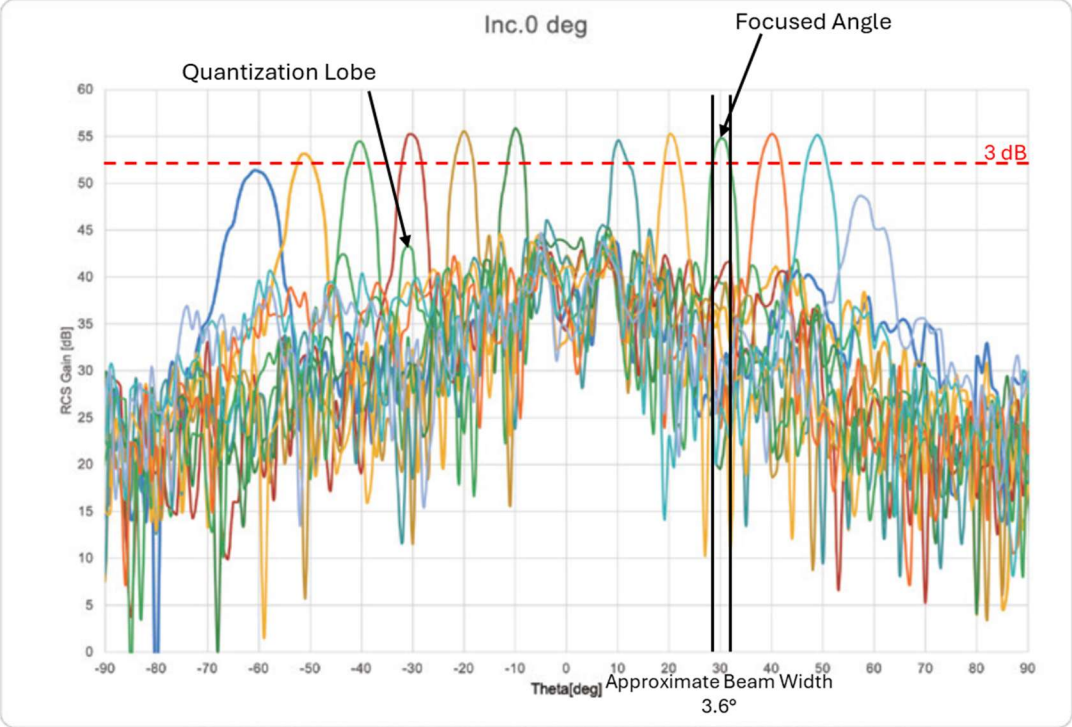


Figure 3.12: TMYTEK Test results for RCS Gain [47] with approximate 3dB Beamwidth overlay

3.4.3 Grating lobe suppression

A major limitation of RIS with phase quantization is that of suppressing grating lobes, especially with single-bit quantization. This is due to the periodicity in phase rounding errors with low bit quantization [48]. The effects are particularly evident

in single-bit quantization. With single-bit phase quantization, the quantization lobe can be of similar magnitude to that of the main lobe [48], potentially having a significant effect on RIS-aided localization.

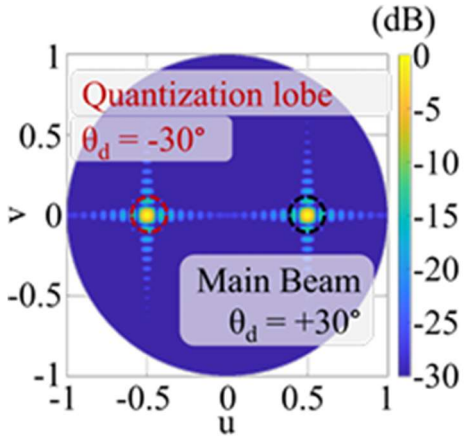
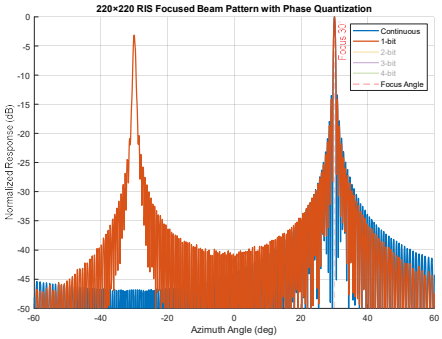


Figure 3.13: Effects on grating lobe suppression with 1-bit quantization [48]

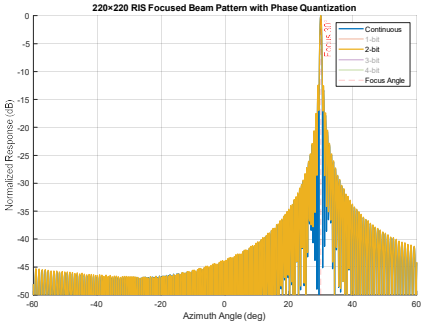
Table 3.8: Quantization effects on grating lobe suppression

Bits	**1**	**2**	**3**	**4**
SL relative suppression to ML (dB) for 220^2 RIS	-12	-30	-30	-30
SL relative suppression to ML (dB) for 32^2 RIS	-3	-47	-47	-47

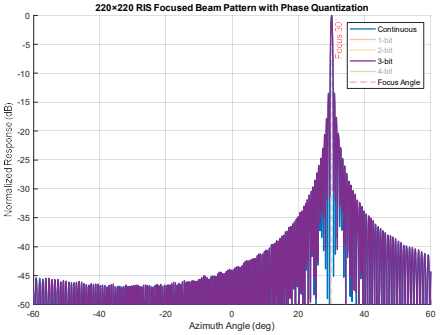
Comparing the simulated quantization effects on the quantization lobe to that provided by TMYTEK's test's results demonstrate a strong alignment, achieving the same quantization suppression for a 32-by-32 RIS with 1-bit quantization, as illustrated in Figure 3.15.



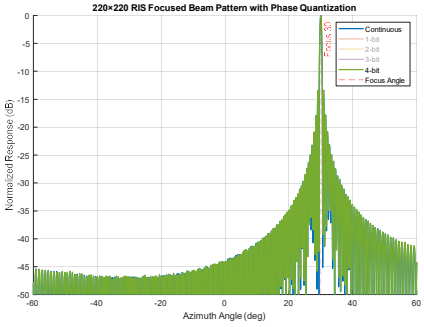
(a) 1-bit vs continuous 220^2 RIS



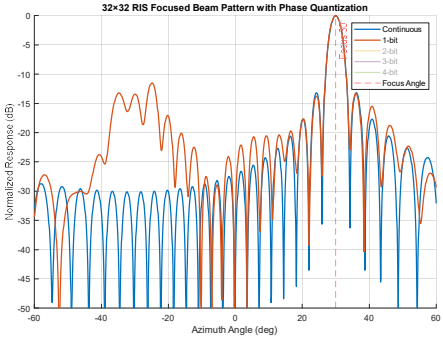
(b) 2-bit vs continuous 220^2 RIS



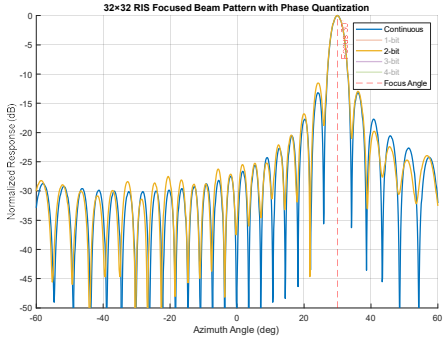
(c) 3-bit vs continuous 220^2 RIS



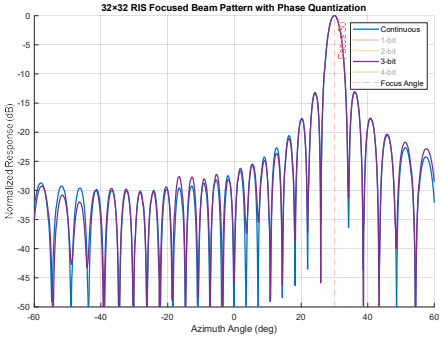
(d) 4-bit vs continuous 220^2 RIS



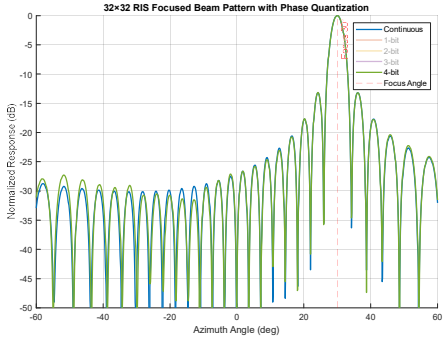
(e) 1-bit vs continuous 32^2 RIS



(f) 2-bit vs continuous 32^2 RIS



(g) 3-bit vs continuous 32^2 RIS



(h) 4-bit vs continuous 32^2 RIS

Figure 3.14: Quantization effects on grating lobe suppression

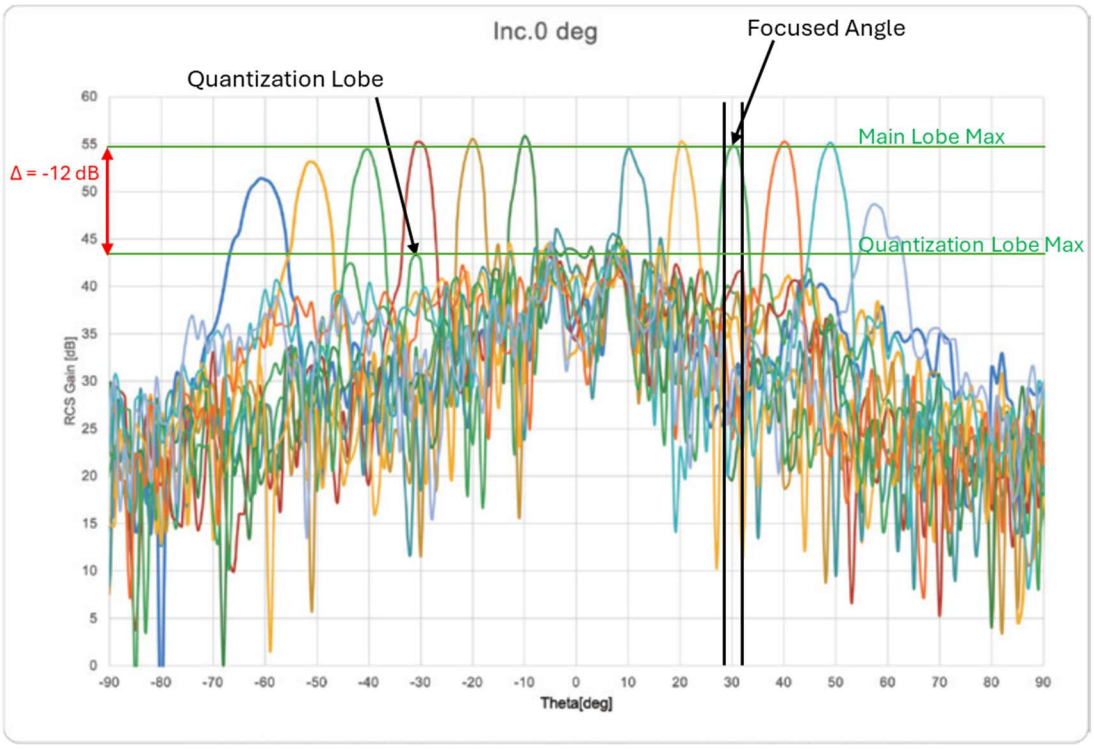


Figure 3.15: TMYTEK Test results for RCS Gain [47] with approximate quantization lobe suppression overlay

This is further supported by the work conducted by [49], where they analyze the effects of spatial and phase quantization. As illustrated in Figure 3.16.

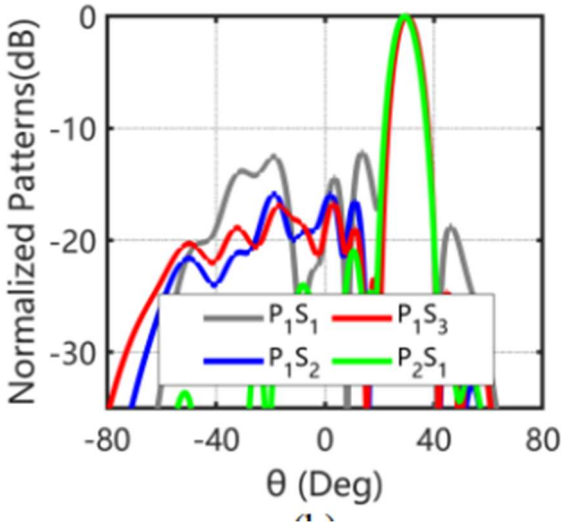


Figure 3.16: Phase quantization effects on grating lobe suppression [49]

When the quantization of the phase is increased from 1-bit (P_1S_1) to 2-bits (P_2S_1) the grating lobe suppression is significantly increased. They demonstrated that the largest increase in grating lobe suppression came from the increase of the phase quantization levels and closely aligns with the non-quantized phase.

Extending the issue of 1-bit phase quantization to RIS-aided localization, there can be a position ambiguity due to the similar magnitudes of the main lobe and the quantization lobe. The quantization lobe makes a strong ambiguity at the mirrored angle, as does the main lobe. Thus, the actual position P_1 and the mirrored position P_2 can be calculated as:

$$P_1 = (rcos\theta, rsin\theta), P_2 = (rcos\theta, -rsin\theta) \tag{3.21}$$

The ambiguity distance can be calculated the Euclidean distance:

$$d_{amb} = \|P_1 - P_2\| \tag{3.22}$$

Therefore, where the highest ambiguity to positional localization is to occur can be calculated. Figure 3.17 illustrates the positional ambiguity due to 1-bit quantization for the 220-by-220 RIS used in our simulated environment.

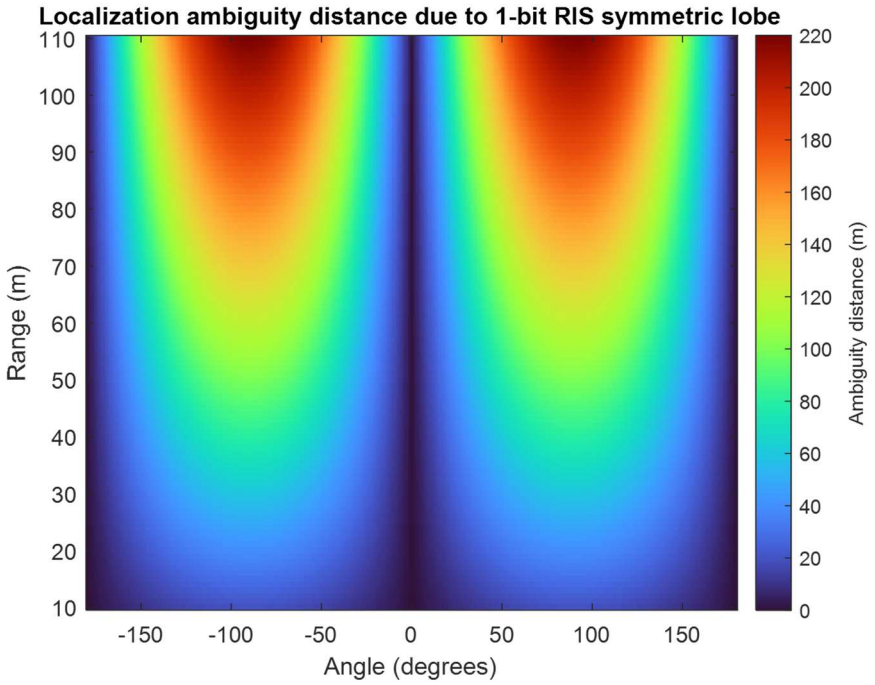


Figure 3.17: Localization ambiguity caused by 1-bit quantization grating lobe

Furthermore, using physical constraints of the environment, when the ambiguity is greater than 20m apart, the other lobe can be ignored because of the large separation distance. Only focusing on when the ambiguity is less than 20m Figure 3.18 is produced.

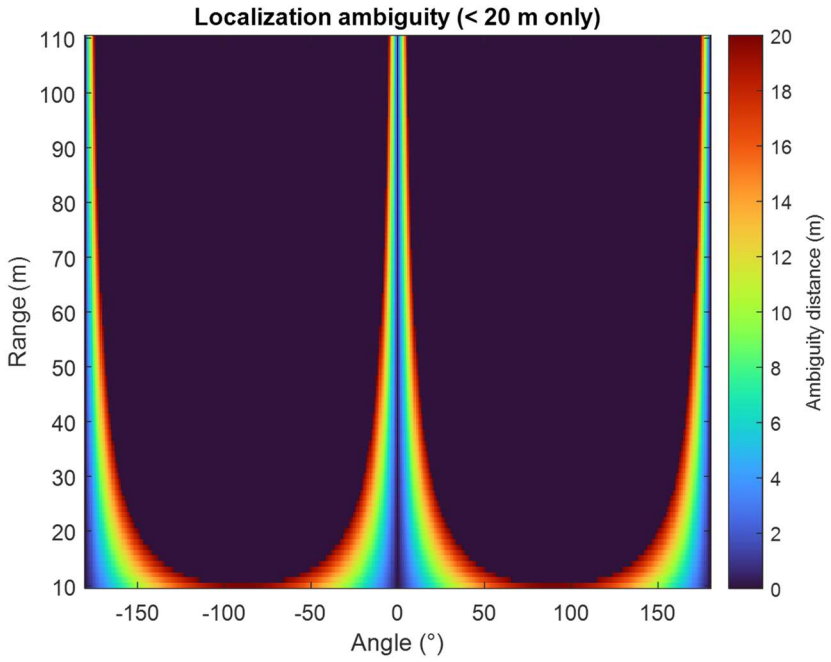


Figure 3.18: Localization ambiguity caused by 1-bit quantization grating lobe (constrained)

3.4.4 Beam pointing errors

In mmWave communication and localization, high-resolution beam steering is essential with the use of RIS to ensure alignment. Extensive research has investigated the effects of phase quantization on array pointing error. As noted by [17], RIS can be represented as URAs; therefore, research into URA pointing error provides insight into its effect on RIS. In their work on beam pointing accuracy of phased arrays for satellite communication [50], the RMS error in beam pointing in a multi-element array can be expressed as:

$$\delta\Phi_{RMS} = \frac{\Delta p}{k\Delta cco\ (s_0)N^{3/2}} \tag{3.23}$$

Where k = phase constant ($\frac{2\pi}{\lambda}$), Δc = element spacing, N = number of elements on steering axis and s_0 = steering angle.

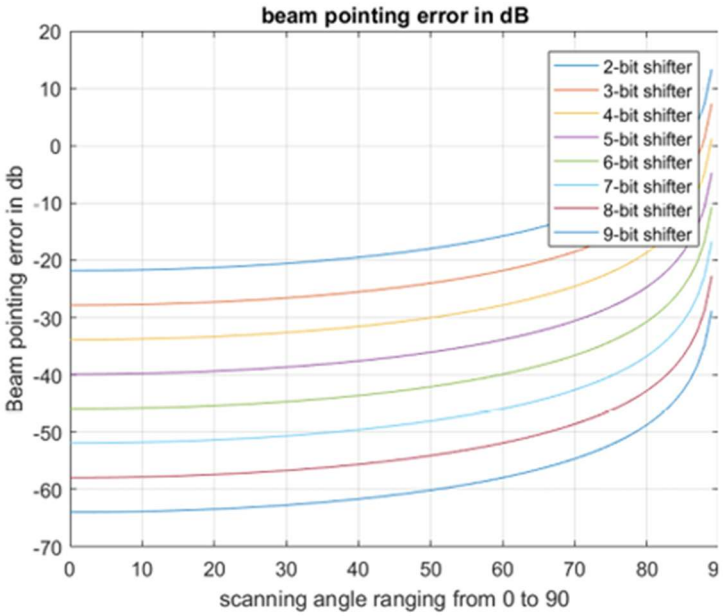
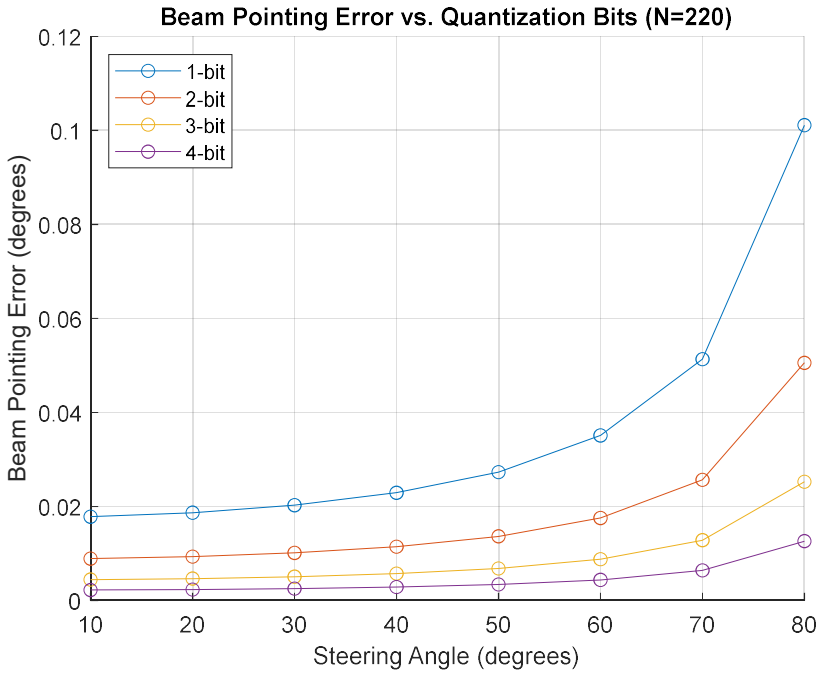
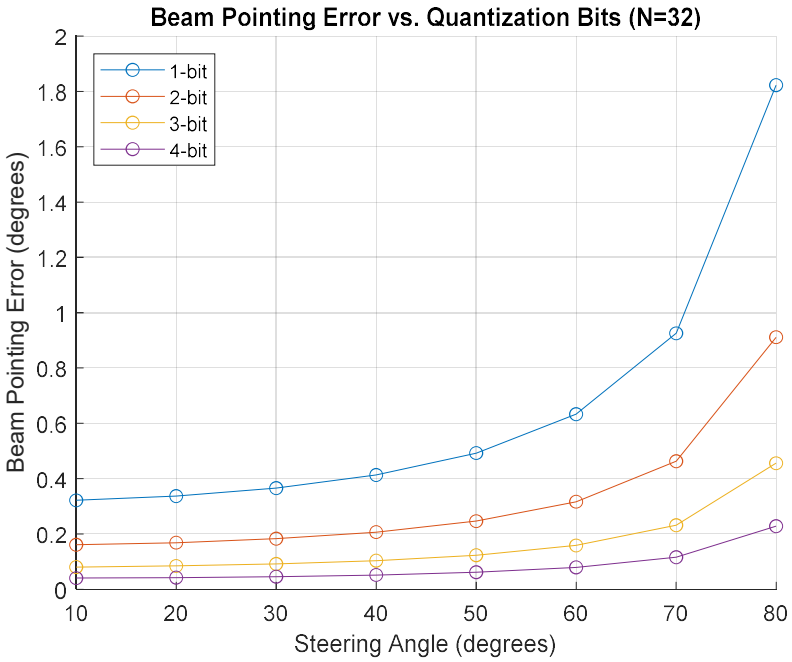


Figure 3.19: The variation of beam pointing error in dB with varying number of bits and the scanning angle [50]

Using Figure 3.18, we can calculate the effects of phase quantization on our 220-by-220 and 32-by-32 RIS which is depicted in Figure 3.20 and Table 3.9.



(a)



(b)

Figure 3.20: Phase quantization effects on beam pointing error for (a) 220-by-220 and (b) 32-by-32 RIS

Table 3.9: Phase quantization effects on beam pointing error

Bits	**1**	**2**	**3**	**4**
Positional error at 110m (m) at 70° offset for 220^2 RIS	0.098	0.049	0.025	0.012
Positional error at 1m (m) at 70° offset for 32^2 RIS	0.016	0.008	0.004	0.002

Recognizing the current limitations of RIS in configurability, [51] limits the phase changeability to 3 bits (8 levels) and, due to the nonlinearity in the amplitude-phase, only the phase is configurable. To optimize the RIS configuration for maximum channel throughput, a greedy algorithm was used to calculate the reflection coefficients for the RIS elements. Due to this limitation, the computational requirements for configuring the RIS were reduced. Despite the limitations of the RIS with the element's phase only having eight states, through their approach, they were able to show that as long as the optimal solution for the configuration of the RIS elements was within the bounds of their limitations, they could achieve near the optimal configuration that a continuous configurable RIS would be able to achieve. Similarly, through our analysis of phase quantization, we can show that if a 3-bit quantized phase is used minimal difference between it and a continuous phase can be achieved.

3.5 Simulated Environment

This section will cover the creation of a simplified RIS-aided localization environment. The purpose of the simplified environment was to ensure the validity of the localization approach and to test aspects of the RL agents. Furthermore, work in the simplified environment allowed experimentation on the limits of the localization approach, which were implemented in further phases of this work.

The power grid was modelled both uplink and downlink scenarios. In the uplink model the UE was represented as an isotropic point source radiating uniformly based off a geometric near field model that uses a propagated spherical wavefront. To estimate the UE's location, the RIS sequentially steers its beam toward each discrete grid position. At each candidate position, the BS' received power is computed based off the free-space propagation of the UE's emissions and the RIS' beamforming gain at that location.

In the downlink model, an analogous procedure is employed, where the RIS directs its beam towards each grid point, and the UE records the corresponding received power level. The downlink model is also based off a geometric near field model that uses a propagated transmitted spherical wavefront. A consideration in the downlink model is the selection of the grid resolution relative to the beam's 3-dB beamwidth to minimize positional ambiguity.

A simplified RIS-aided localization environment was created using these approaches. The RIS and Access Point (AP) are fixed with the new UE position being user defined. The previous location is randomly chosen in the vicinity of the user defined position with movement only possible within one unit of the previously known position.

```
Enter a value for Y in meters for UE position between (3,9):4
Enter a value for X in meters for UE position (3,9):5
Enter a value for Z in meters for UE position (1,2):2
```

Figure 3.21: User input for UE location in the simplified environment

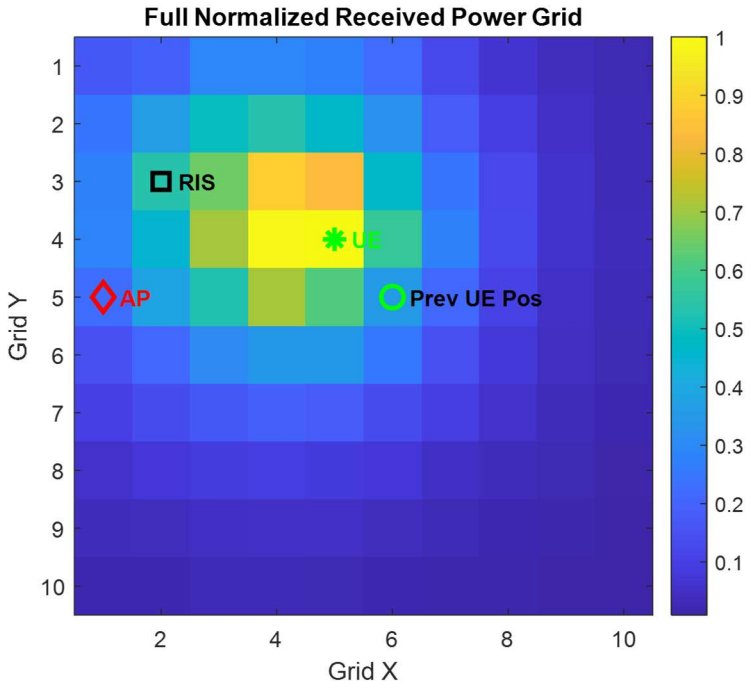


Figure 3.22: Single RIS-aided localization power grid

With a single RIS for localization, there is ambiguity between grid points (4,4) and (5,4), where the received power is comparable; however, the RL agent was able to localize to the correct location As shown in Figure 3.23.

Following the same principles as [30], a map of the area of interest was created by beam shaping the RIS to provide coverage of where the highest probability of the UE may be located.

To perform the sequential area scan to localize the UE as detailed in [30] the time to perform the scan is [30]:

$$t_{scan} = \frac{\theta_s t_{int}}{2} N \tag{3.24}$$

Where t_{int} is the integration time for a signal received from a source and θ_s is the total coverage area of interest.

However, by increasing the size of N the time to scan the environment increases. Thus, for scanning a fast-moving target having more elements in a RIS reduces the ability of the RIS to track the target due to the time to scan the area of interest. But if the area of interest can be reduced with the use of RL and smart scanning then this issue can be mitigated.

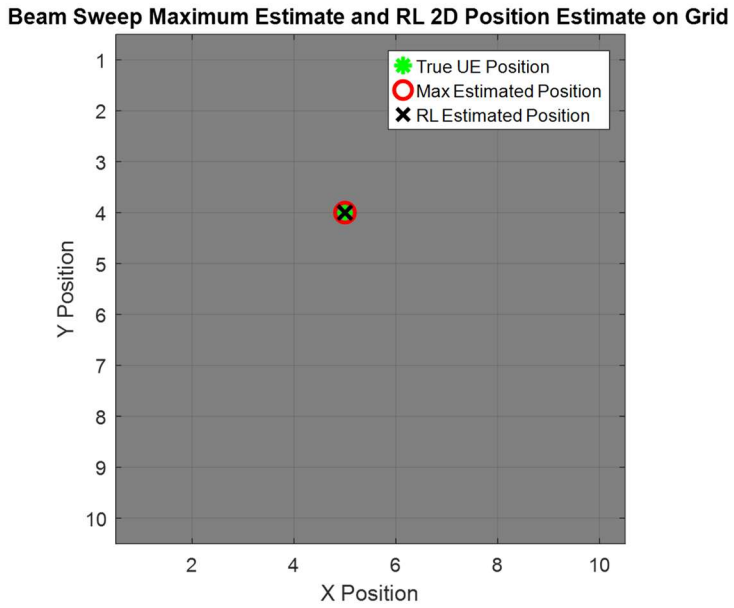


Figure 3.23: 2D RIS-aided localization with a single RIS

Another environment was created in Figure 3.25 and Figure 3.26 to determine the performance of a dual RIS-aided localization. When there is ambiguity in the UE location with Figure 3.22, a dual-RIS configuration resolves it.

```
Enter a value for Y in meters for UE position between (3,9):7
Enter a value for X in meters for UE position (3,9):8
Enter a value for Z in meters for UE position (1,2):1
```

Figure 3.24: User input for UE location in the simplified environment

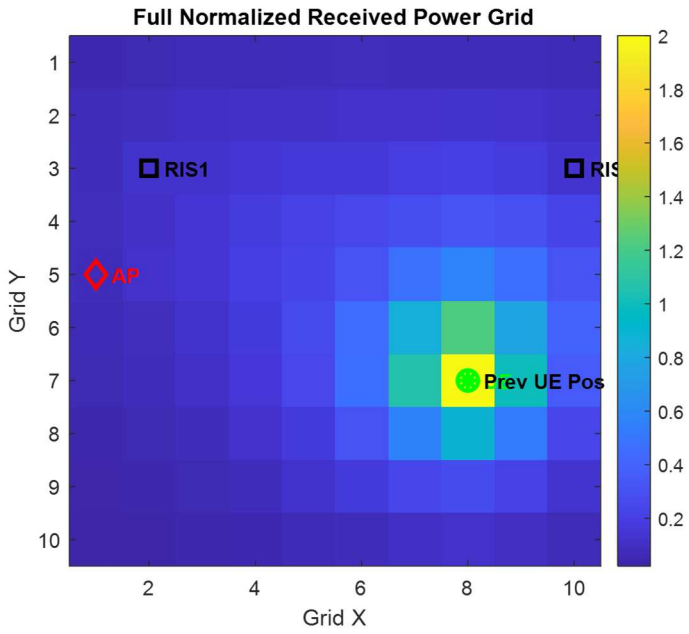


Figure 3.25: Dual RIS-aided localization power grid

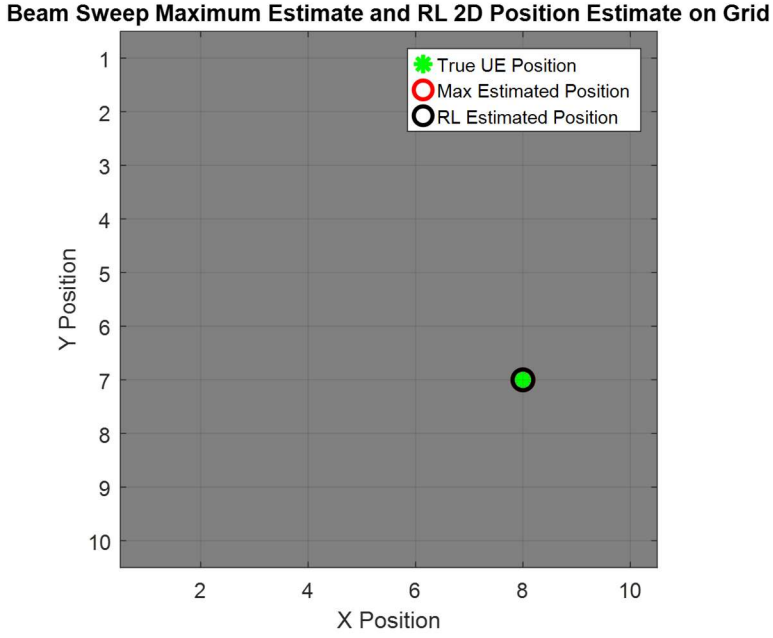


Figure 3.26: 2D RIS-aided localization with two RIS

The simulation is modelled assuming that the UE is within the near field to allow the CoA to be exploited. The RIS' reflection pattern, which is encoded with geometric information based off the range and angles between the UE and RIS allows the 3D position to be determined. If far-field was assumed the 3D position

can be determined based off triangulation; however, due to the near field having a curved wavefront the CoA can be used with the RIS reflection pattern to determine the 3D position. The CoA with a 3D grid search was implemented in the simplified environment, as illustrated in Figure 3.27, and was able to determine the 3D location, validating our approach.

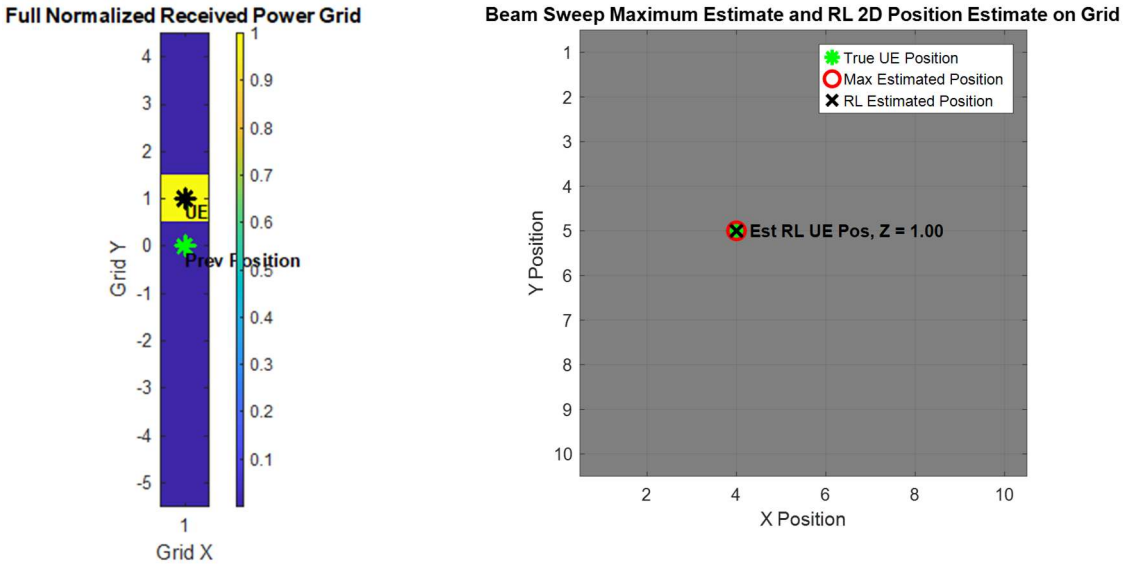


Figure 3.27: Single RIS-aided 3D localization power grid

The objective of the simplified environment was to test the RIS-aided localization approach and RL algorithm. The main aspects that were analyzed in the simplified environment were: effects of RIS size on localization, BS-RIS-UE geometry, and RL performance.

3.5.1 RIS Size Effects on Localization Accuracy

Research [6], [13] has investigated the requirements of a RIS to achieve a narrow beamwidth. Using the centre point of the beamwidth as the position of the UE the accuracy of the localization can be calculated based on the size of the RIS.

Using a RIS of 220 by 220 elements, which corresponds to a size of 1.18m by 1.18m, the positional accuracy can be assumed to be the beam spot size. Figure 3.28 illustrates the beam spot size at broadside. For simplicity, the analysis assumes equal beam spot dimensions along the x- and y-axes, such that, Wx=Wy. With the

size of the RIS being 220 by 220 at a range of 110m, the beam spot size is 1 m by 1 m.

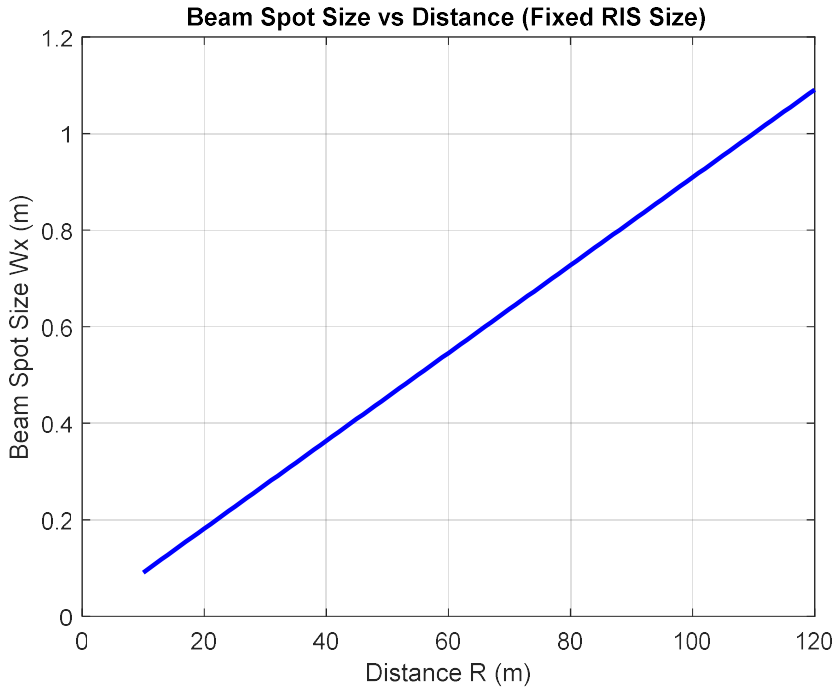


Figure 3.28: Beam spot size vs distance for a 220 by 220 element RIS

In addition, as detailed [30] the AoA of the emitter is estimated to have a precision no worse than the half power beamwidth (HPBW) of the RIS. For a RIS with N^2 with element spacing of half wavelength the HPBW of the RIS plane in azimuth and elevation are:

$$\overline{\Theta_{az}} = \overline{\Theta_{el}} \cong \frac{2}{N} \tag{3.25}$$

So given our size of N = 220: $\overline{\Theta_{az}} = \overline{\Theta_{el}} \cong \frac{2}{220} = 0.0091rad$

And at 110m the beamwidth footprint can be calculated as: $d\theta = 0.0091 * 110 \cong 1m$

3.5.2 BS-RIS-UE Geometry

To maintain at least 1 m in accuracy with a maximum distance between the RIS and the UE of 110m the size of the RIS has to be N=M=220 elements. It was also found that the elevation angle separation between the RIS and UE had to be less than 25 degrees to maintain the beam spot in the accuracy bounds of 1m.

Assessing the impact of multiple RIS in the BS-RIS-UE geometry it was found when there is ambiguity in the position of the UE it was found that the use of multiple RIS reduced the ambiguity. It did this by providing a more defined maximum power location, as can be seen comparing the single RIS configuration of Figure 3.22 to the dual RIS configuration of Figure 3.25.

3.5.3 RL Performance

Analysis compared the localization accuracy of the RL agent and BSGS. In this simplified environment, it was found that there was no significant difference between the two. This is not surprising because the environment was very simple in geometry, size, and channel quality; nonetheless the simulation showed that the RL agent can perform localization comparable to that of BSGS.

4 RIS-Aided Localization with a Real-World Trajectory

4.1 Introduction

The knowledge gained from the previous phases were applied to a real-world trajectory captured from a vehicle that drove around Kingston, Ontario. A GNSS denied environment was simulated and RIS-aided localization was employed for positioning. Significant work was required to first create a network topology that had enough RIS to provide coverage to the UE throughout the GNSS outage. The GNSS denied trajectory and RIS placements are illustrated in Figure 4.1.

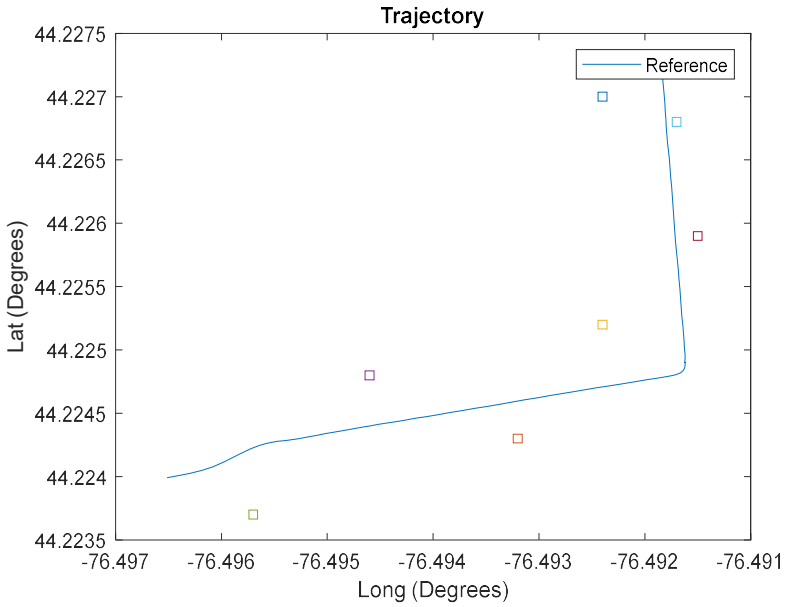


Figure 4.1: GNSS denied environment vehicle trajectory with RIS placed along route (squares)

Leveraging the advantages of GNSS single-point position and RIS localization, along with the Inertia Measurement Unit (IMU), which requires no external sensor information, a loosely coupled integration balances the challenges of GNSS and RIS localization, which are prone to noise, and of the IMU, which is prone to compounded errors and bias. The integrated navigational solution is achieved through a Kalman filter, as depicted in Figure 4.2.

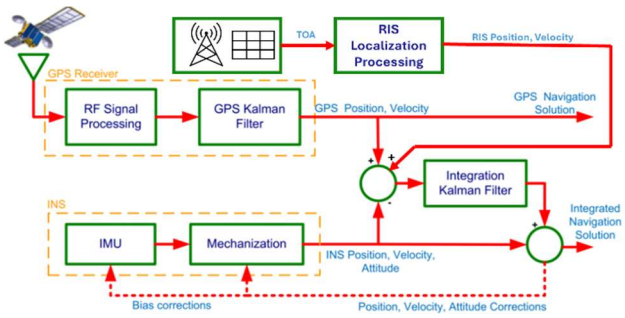


Figure 4.2: Loosely coupled integration of R-IMU mechanization GNSS and RIS Localization [52]. [Used with permission]

Kalman filters for an integrated navigational solution optimally estimate the errors state of a system caused by errors introduced into the measured signal [53]. Figure 4.3 outlines the general concept of a Kalman filter with Figure 4.4 detailing the flow chart of a Kalman filter.

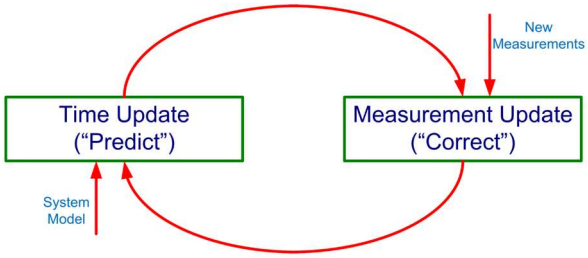


Figure 4.3: General concept of Kalman Filter [52]. [Used with permission]

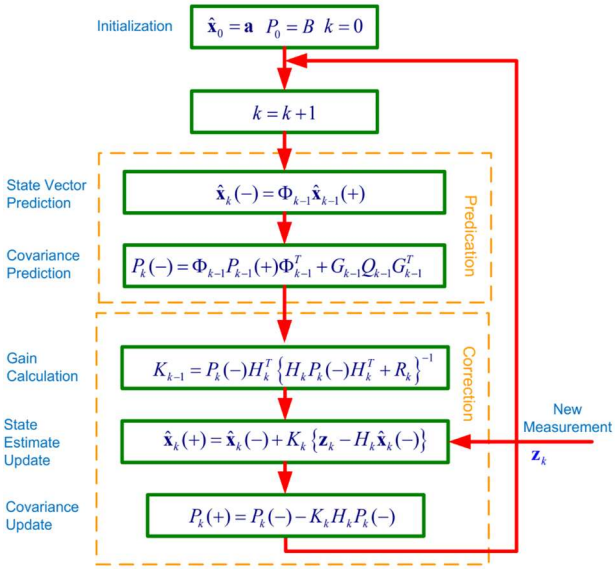


Figure 4.4: Flow chart of a Kalman filter [52]. [Used with permission]

A more detailed description of a Kalman filter can be found in [54], but by following Figure 4.4 [55] the integration of navigational sources through a Kalman filter can be realised.

As depicted in Figure 4.2, for the integration of the navigational sources, first the Reduced-IMU (R-IMU) mechanization occurs, and the outputs are compared to the GNSS single-point positioning and RIS localization, when either are available, and the errors are calculated. The errors are fed into the integrated Kalman filter, where it relies on the dynamic R-IMU model and stochastic sensor error model. The Kalman filter provides the estimate of the difference in the navigational solution, which then is applied to the R-IMU mechanization output to correct for the errors. This implementation uses a Loosely Coupled Kalman filter based off of [53] to provide feedback to the mechanization process.

4.2 RIS-Aided Localization for Positioning with real-world Trajectory Simulation

To verify the RIS localization approach and RL algorithms developed, an environment with real-world trajectory was employed. Using the simulated environment the RIS-aided localization with RL and BSGS were compared during a GNSS denied environment.

Using the Kalman filter to integrate the GNSS, RIS and R-IMU navigation sources, an integrated navigational solution was computed and compared to reference trajectories. An outage of the GNSS signal was simulated for at two different instances for two minutes each. For the first outage, the RIS localization provided new measurements to the model for positioning, while during the second the model only relied on the INS. The following section is broken down as follows:

- An outline of the implementation of the integrated navigation solution; and
- A discussion on the results of the integrated navigational system.

A road trajectory from around Kingston, Ontario was used to produce results for this implementation. The trajectory consisted of a combination of urban and highway driving with a total time of approximately forty minutes. In our implementation the integrated GNSS RIS R-IMU Kalman filter results were compared to a reference trajectory.

During the first GNSS outage RIS localization was used to provide positioning and velocity to the integrating Kalman filter. Figure 4.5 details the reference trajectory with the placement of BS and RIS.

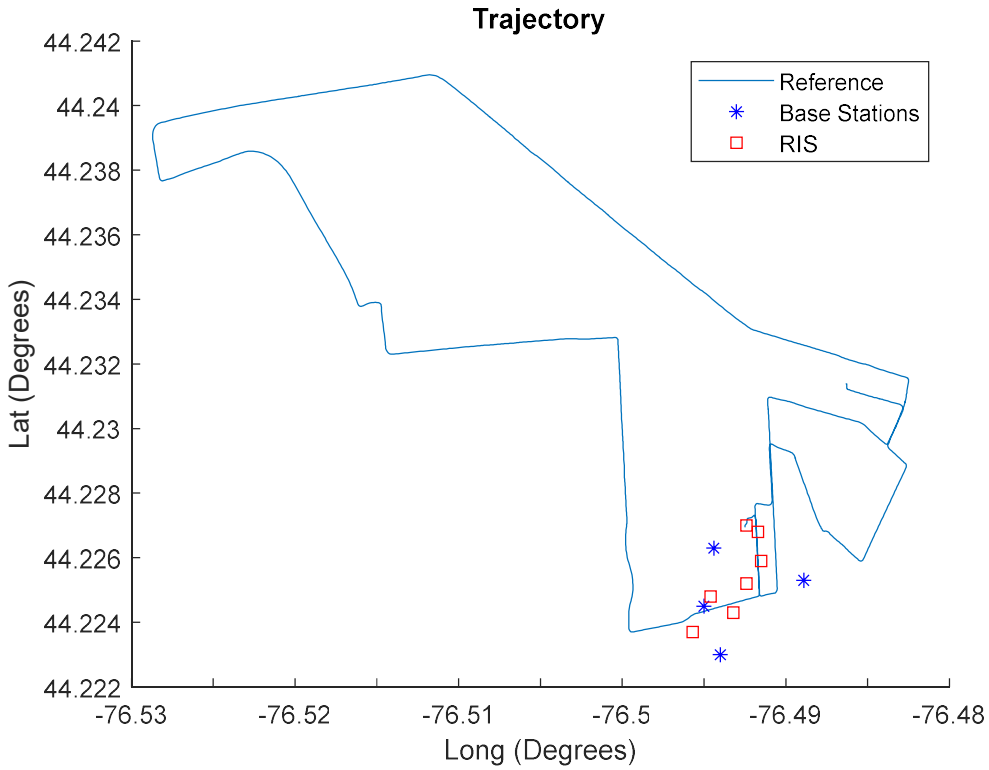


Figure 4.5: Complete Reference Trajectory with Base Station and RIS Locations

To simulate the operation of a UE moving along the trajectory with the UE-RIS-BS channel, the reference trajectory was used as the location of the UE at the sampling times and AWGN channels with a free space model were created. Initial simulations did not use any additive Gaussian noise and perfect Channel State Information (CSI) was assumed. A maximum channel distance of 5 km was also assumed with a RIS UE maximum distance of 110m. A signal operating at $f_c = 28$ GHz was used. The transmitted downlink signal had a peak power of $P_{peak} = 20$ W with 0 gain while the transmitted uplink signal had a peak power of $P_{peak} = 200$ mW with 0 gain during RIS-aided localization. The RIS were constructed with $N = M = 220$ with half wavelength element height and length. As a baseline continuous phase is used for the RIS configuration, unless otherwise stated.

A SISO localization scenario was implemented for the RIS aided localization as illustrated in Figure 3.2 for the given trajectory at the first GNSS outage.

4.2.1 RL vs BSGS Localization for Positioning

As discussed in [54], for an accurate navigational solution to be produced with a Kalman filter it has to be tuned.

GNSS and RIS localization produce high accuracy in their positioning information however, both are prone to outages and non-accumulating errors, so generally their measurements can be trusted. When there are outages of GNSS, the Kalman filter will have to rely on RIS localization for positioning and if that is not available, then on the R-IMU mechanization process and its model. To add more weight to the measurement provided by the GNSS and RIS sources the Kalman filter was given weight to the GNSS and RIS over the R-IMU, thus tuning the Kalman filter.

Outages were simulated at 4, and 10 minutes into the trajectory. During GNSS outages the Kalman filter relies on RIS aided localization, if available, otherwise it relies on its state estimation, which introduces drifts caused by the inherent errors within the R-IMU mechanization process.

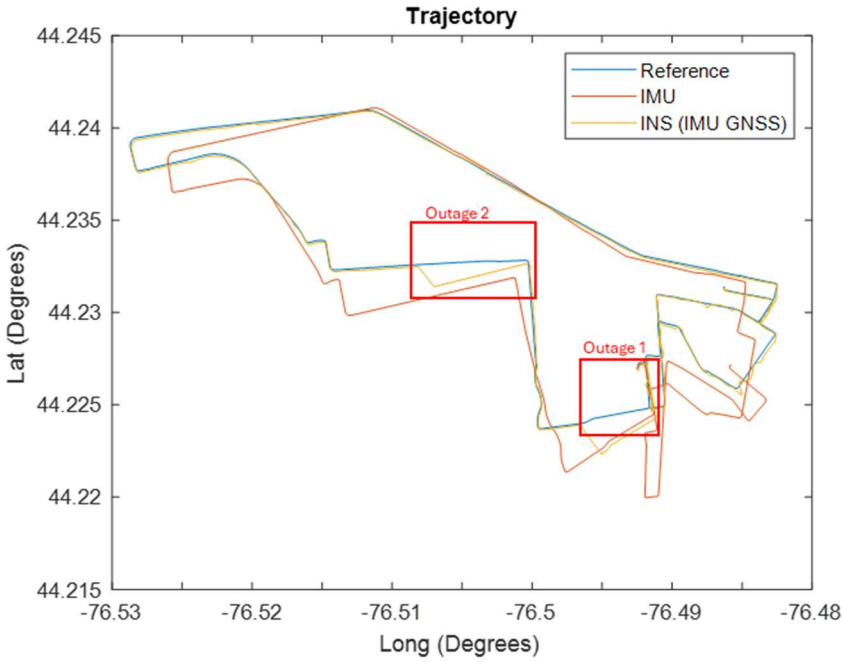


Figure 4.6: IMU, GNSS, reference and Integrated Navigational Trajectories with outages without RIS-aided Positioning

As outlined in Figure 4.6, during the outage at 4 minutes there are errors in the trajectory of the vehicle caused by the drift in the mechanization process. In the integrated navigational solution, when there is an outage, the navigational solution resembles the R-IMU trajectory from the start of the outage to the end. That is, the

integrated solution fixes the R-IMU to GNSS trajectory, but when the GNSS signal is lost the R-IMU takes over. Once the GNSS is returned, the trajectory returns to resemble the reference signal. Figure 4.7 and Figure 4.8 illustrate the trajectories with RIS positioning implemented during the first GNSS outage.

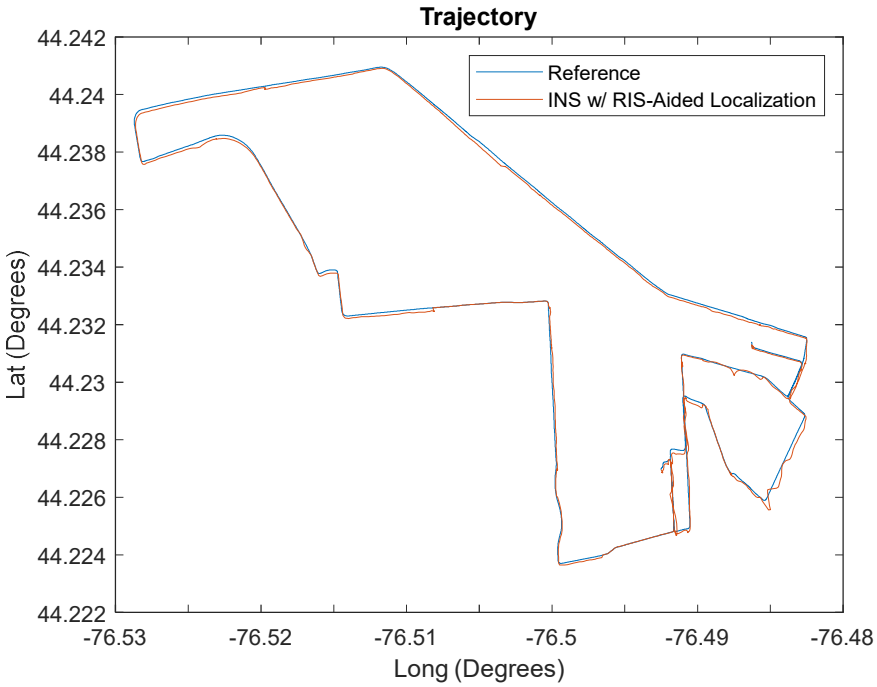


Figure 4.7: Reference and Integrated Navigational Trajectories with outages with RL RIS Positioning

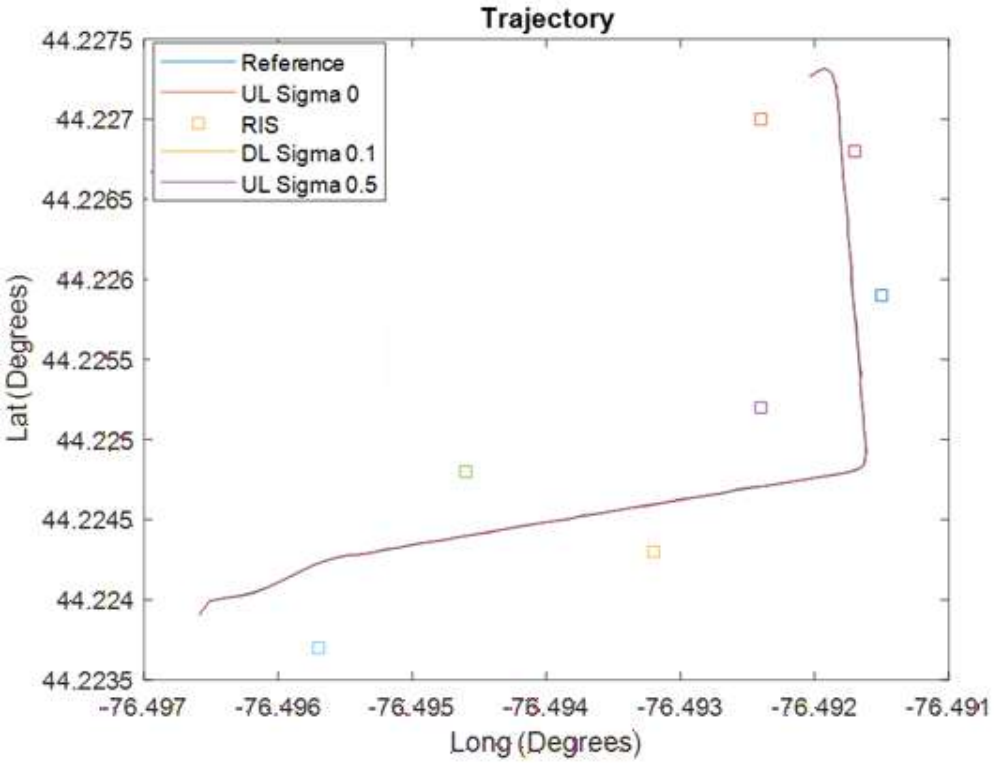


Figure 4.8: IMU, GNSS, reference and Integrated Navigational Trajectories with outages with RIS Positioning

With the implemented integrated navigational solution, the following results were produced for position and velocity, with a focus on the first outage, where RIS-aided positioning was implemented with σ being zero.

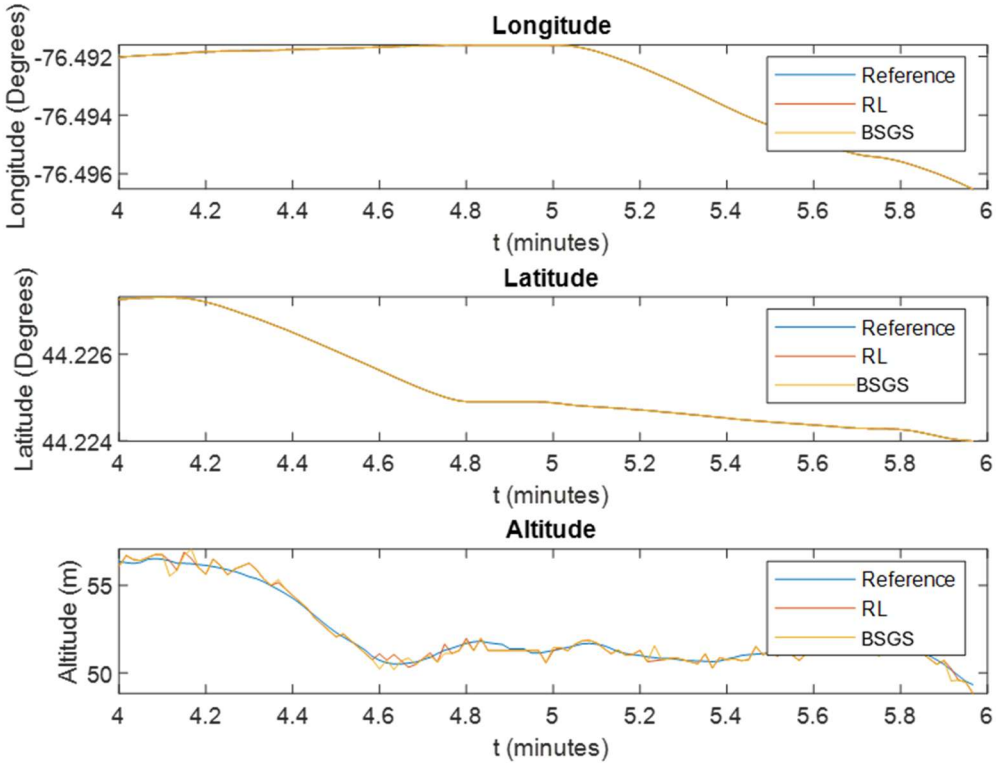


Figure 4.9: Positioning Throughout First Outage

Figure 4.9 details the comparison between the longitude, latitude, and altitude of the integrated navigational solution with RIS positioning compared to the reference signal. During the outage, where RIS localization is utilized, a highly accurate position is maintained, with only small variations in the altitude. It is evident that the results for RL and BSGS are comparable in a non-nosy channel. This is consistent with the expected result as they both localize using the same process its just BSGS does a broader search compared to RL.

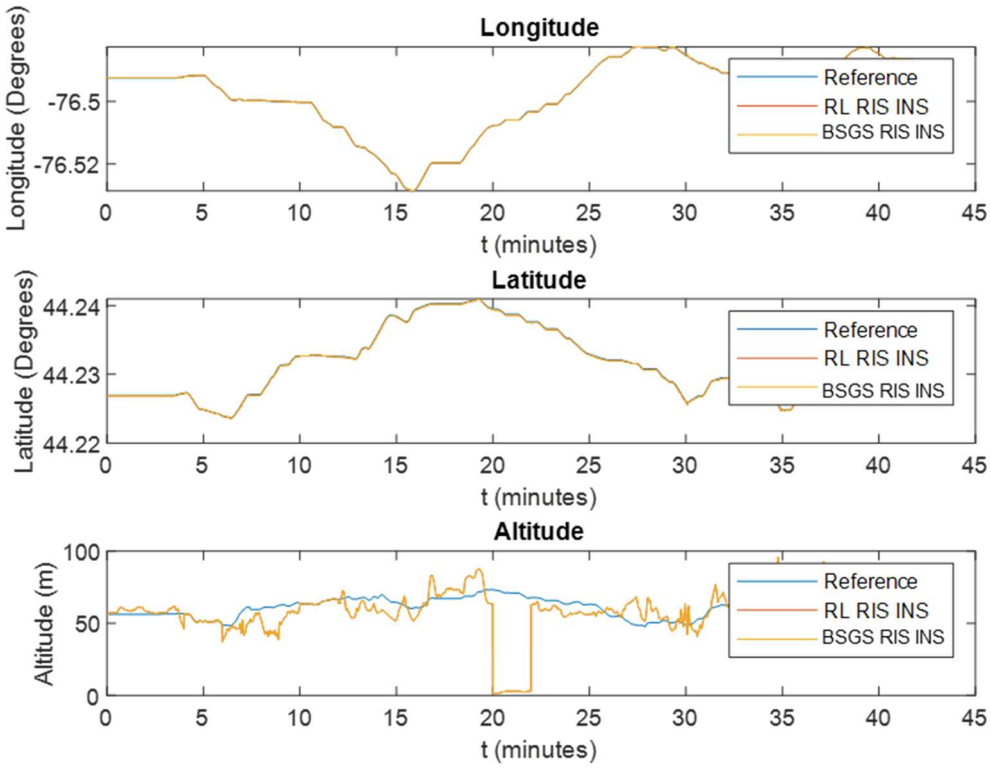


Figure 4.10: Positioning Throughout Full Trajectory

As can be seen in Figure 4.10 over the course of the complete trajectory during the last GNSS outage, where RIS positioning is not used, there is a drift introduced by the R-IMU mechanization process. Due to this implementation of the R-IMU mechanization not being designed for altitude [56], when the mechanization takes over during outages without RIS-aided positioning the altitude diverges but returns to the reference trajectory altitude once the GNSS signal is returned.

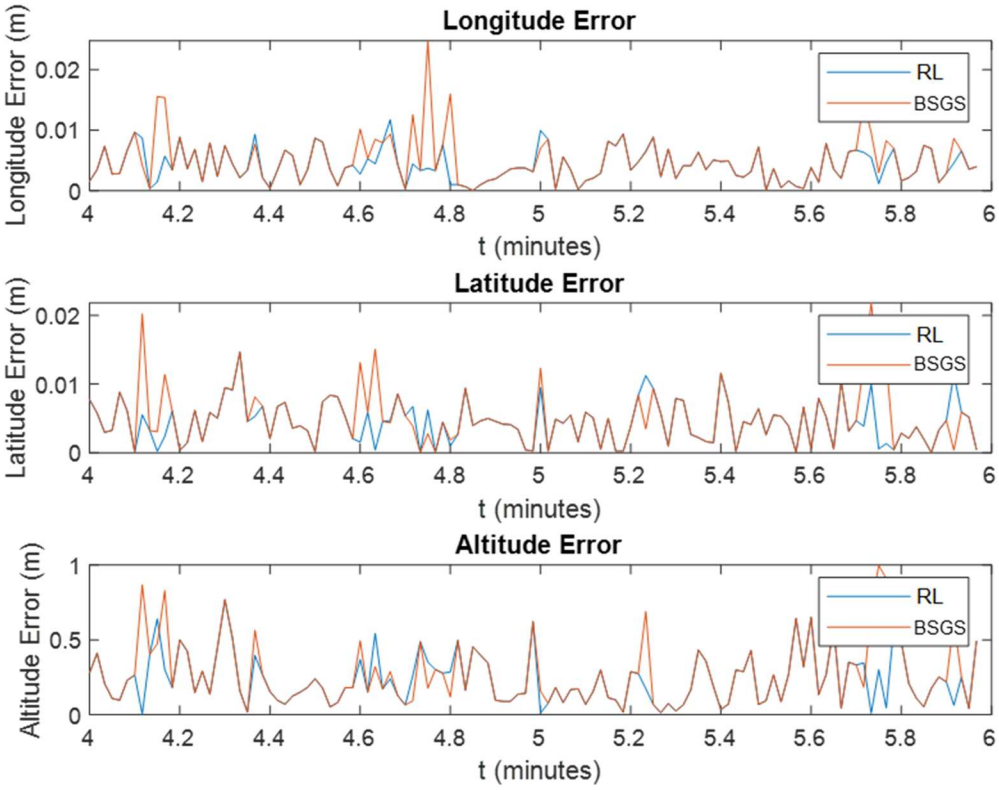


Figure 4.11: Positioning Errors Throughout First Outage

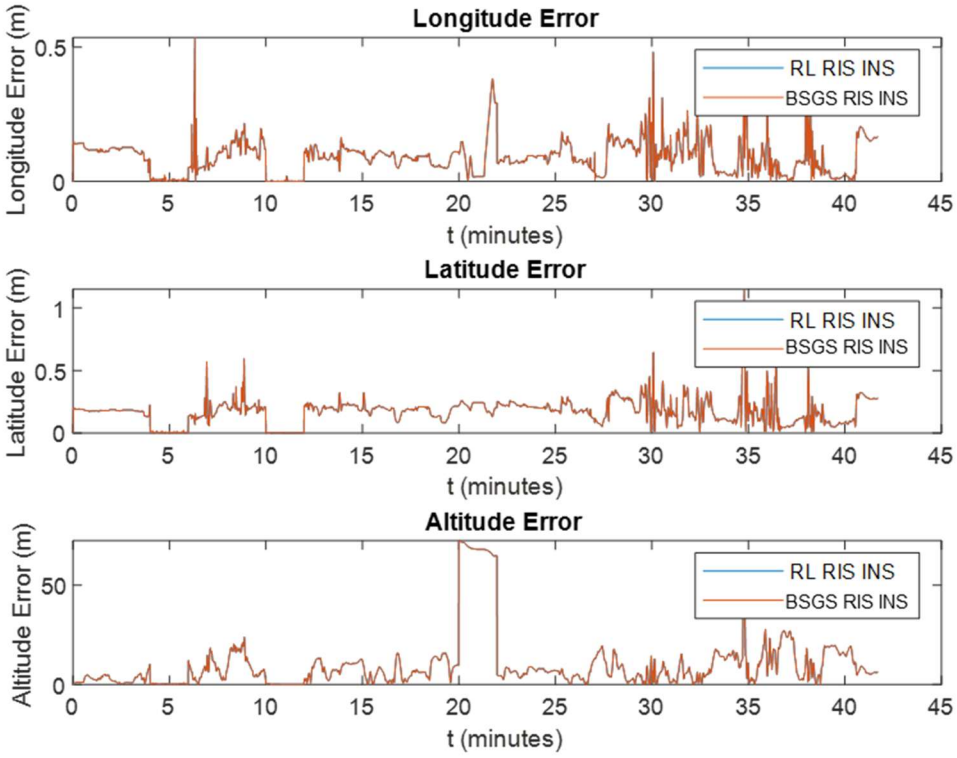


Figure 4.12: Positioning Errors Throughout Full Trajectory

Table 4.1 details the RMSE errors of the implemented trajectories for latitude, longitude and altitude.

Table 4.1: Positioning errors for RIS-aided positioning for the first outage

RMSE	RL RIS INS	BSGS RIS INS
Longitude	0.005m σ 0.003m	0.006m σ 0.004m
Latitude	0.005m σ 0.003m	0.007m σ 0.003m
Altitude	0.29m σ 0.17m	0.34m σ 0.21m

The RIS positioning provided a very accurate latitude and longitude and altitude positioning. Comparing the RMSE of the RL to the BSGS they are comparable with very little difference in their results. Again, this is as expected, as with BSGS and RL when the localization is performed and the maximum power is determined there should only be one position evident in the power map. There should not be any ambiguity in the estimated UE position.

Similar to the position calculations, the north, east, and up velocities were calculated

throughout the trajectory and are illustrated in Figure 4.13 and Figure 4.14. Like the results for positioning, Figure 4.13 shows a high accuracy in velocity estimation with comparable results for BSGS and RL.

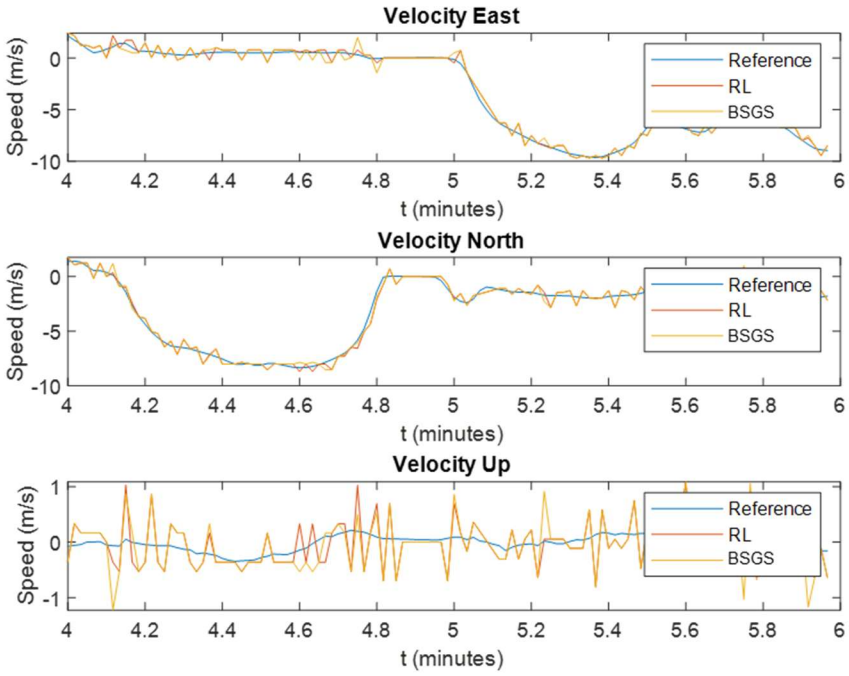


Figure 4.13: Velocities Throughout First Outage

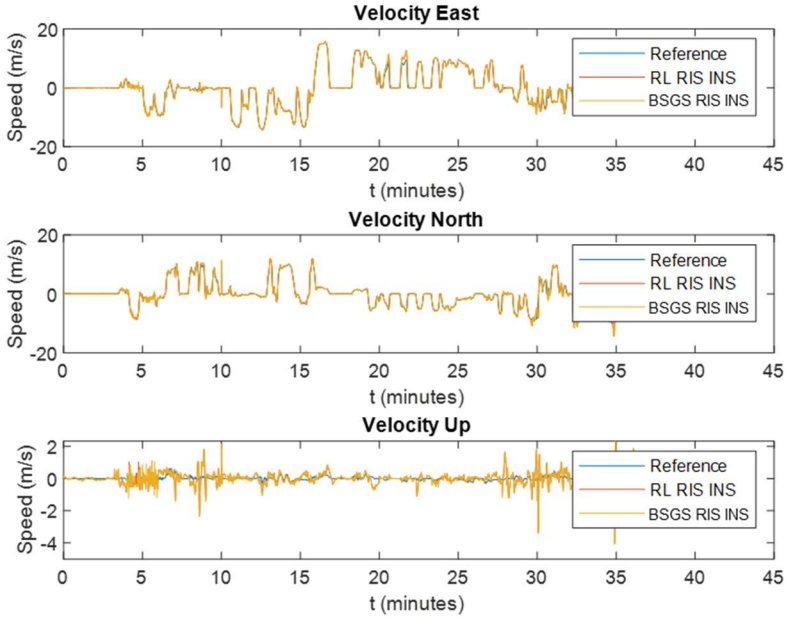


Figure 4.14: Velocities Throughout Full Trajectory

Table 4.2 details the RMSE errors of the implemented trajectories for velocities east, north, and up. It shows that the results of BSGS and RL are comparable during the first outage.

Table 4.2: Velocity errors for RIS-aided positioning for the first outage

RMSE	RL RIS INS	BSGS RIS INS
Ve	0.47m/s σ 0.47m/s	0.51m/s σ 0.51m/s
Vn	0.46m/s σ 0.45m/s	0.50m/s σ 0.50m/s
Vu	0.39m/s σ 0.39m/s	0.43m/s σ 0.43m/s

By providing an accurate positioning source during GNSS outages, the RIS positioning enhances the overall accuracy of the velocity calculations as depicted in Figure 4.16 and Table 4.2. While the R-IMU and R-IMU INS drift with GNSS outages, the accuracy of the velocities is maintained when RIS positioning is used.

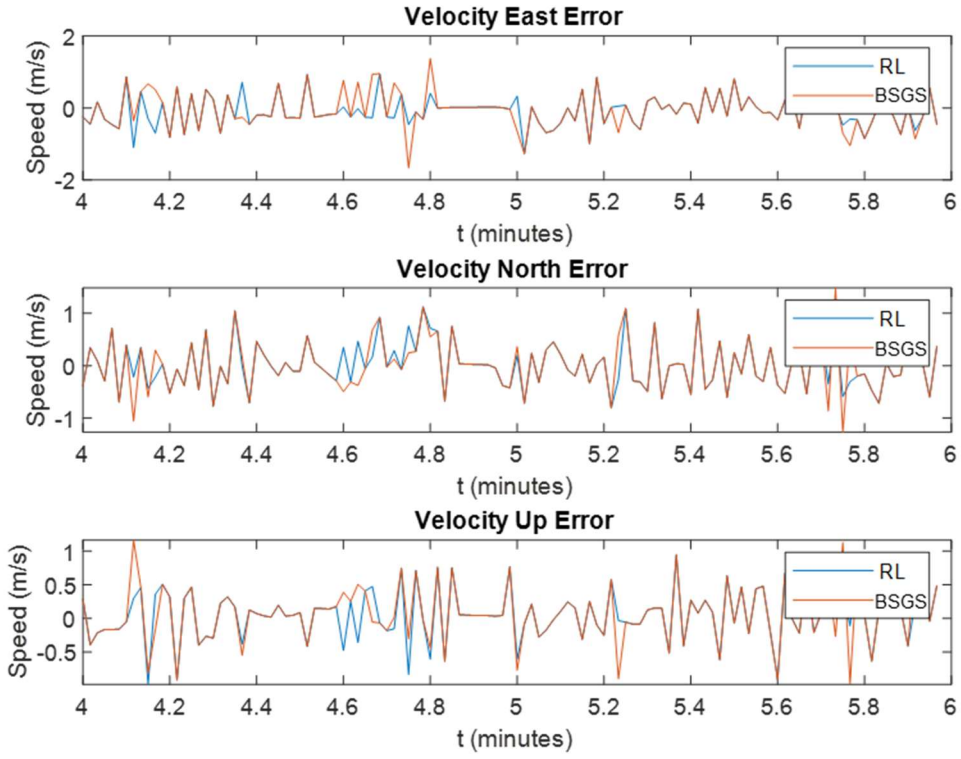


Figure 4.15: Velocity Errors Throughout First Outage

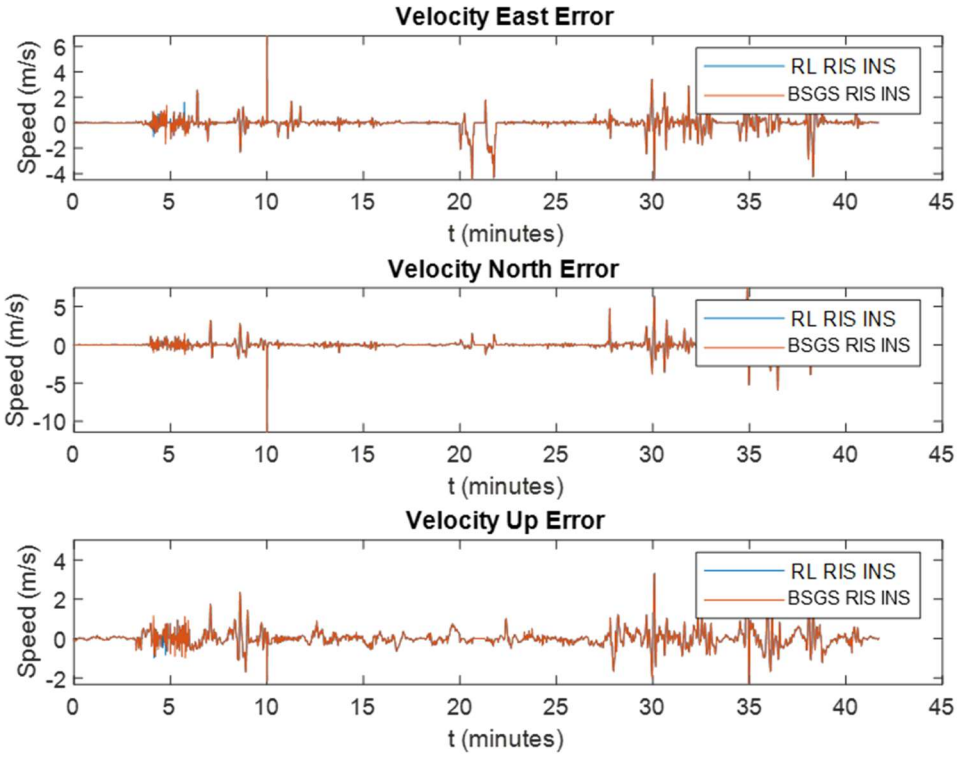


Figure 4.16: Velocity Errors Throughout Full Trajectory

Table 4.3: Quantization impact on integrated trajectory

RMSE	$\sigma^2 = 0$ 4 Bits	$\sigma^2 = 0$ 2 Bits	$\sigma^2 = 0$ 1 Bits
Long (m)	0.005 σ 0.003	0.005 σ 0.003	0.005 σ 0.003
Lat (m)	0.005 σ 0.003	0.005 σ 0.003	0.005 σ 0.003
Alt (m)	0.30 σ 0.18	0.30 σ 0.18	0.30 σ 0.18
Ve (m/s)	1.16 σ 0.15	1.16 σ 0.15	1.16 σ 1.15
Vn (m/s)	1.84 σ 1.08	1.84 σ 1.08	1.84 σ 1.08
Vu (m/s)	0.12 σ 0.12	0.12 σ 0.12	0.12 σ 0.12

Table 4.4: Quantization impact on RIS positioning trajectory

RMSE (m)	$\sigma^2 = 0$ 4 Bits	$\sigma^2 = 0$ 2 Bits	$\sigma^2 = 0$ 1 Bits
RL UL	0.5075	0.5075	0.5075
RL DL	0.6967	0.6967	0.6967
BSGS UL	0.6004	0.5985	0.5827
BSGS DL	0.7795	0.7471	0.8107

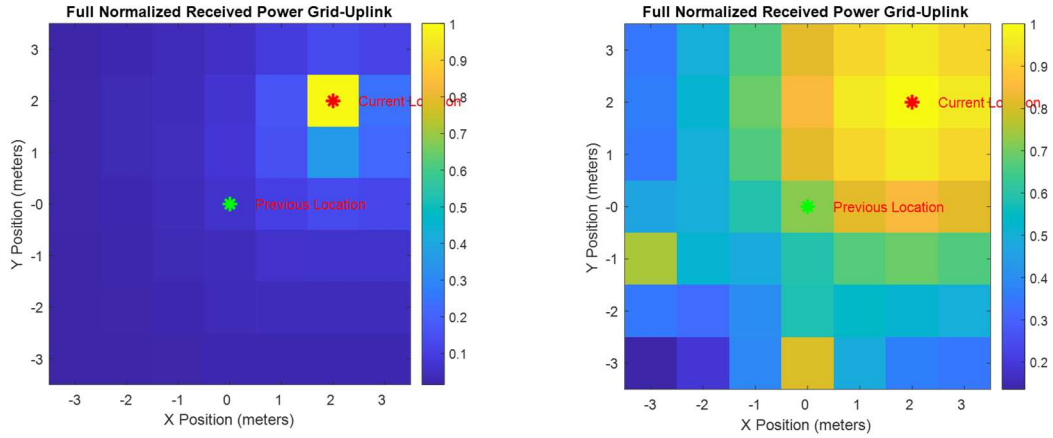
Simulation scenarios were created to verify our methods conducted in the quantitative analysis of the effects of RIS phase quantization on localization. From Table 4.3 and Table 4.4 it can be seen that the phase quantization of the elements of the RIS has negligible impact on the overall positioning performance. For the integrated navigation solution, in Table 4.3 there was no change in the estimated position and the resulting trajectory remained consistent across the different quantization levels.

This is further reinforced when the raw positioning is analyzed in Table 4.4. The RL-based approach yields the same RMSE values across all quantization levels; however, the BSGS-based approach sees minor fluctuations across quantization levels. The difference in performance between the RL-based and BSGS-based positioning approaches can be contributed to the effects of the quantization lobe and the ambiguity it introduces. While the RL-based positioning approach was able to overcome the ambiguity, when present, due to its focused localization grid search, the BSGS-based localization approach was not. Overall, the results suggest that RIS phase quantization has minimal influence on positioning accuracy. However, the advantage of the RL approach is now evident as it is able to overcome the quantization lobe ambiguity compared to the BSGS approach.

RIS positioning is prone to noise in the channel. To simulate a noisy channel, Gaussian noise was introduced to the channel with a variance of 0.01, 0.1, and 0.5, which saw significant errors for the uplink channel as depicted in Table 4.5.

The configuration tested for multipath was a Gaussian spatial response from the UE position. This simplified model is supported by literature such as [57], which provides a measured mmWave outdoor environment. In their measured environment there is typically 2.5 lobes with the Lobe Angular Spread (LAS) of approximately 40 degrees with the RMS of the LAS approximately 7.8 degrees, meaning that even in a multipath environment the spread of the main beam is still relatively minimized. Based off of [57] a multipath mask was created and applied to our simulated environment. Figure 4.17 illustrate the mask application to the environment. Using the developed multipath

environment in our real-world trajectory the effects of the multipath channel were determined and provided in Table 4.5.



(a) Non-multipath Applied Channel (b) Multipath Applied Channel

Figure 4.17: Multipath Mask Channel

Throughout the trajectory there is an expectation that there will be missed samples. There will be instances that there may be a blockage and a sample is not able to be achieved. As the RL algorithm was trained to overcome this, it was implemented in our real-world trajectory with a 10% of a missed sample. The impact of all the effects on the first outage trajectory is detailed in Table 4.5.

Table 4.5: Error vectors and impact on trajectory

RMSE	$\sigma^2 = 0$ Up-Link	$\sigma^2 = 0$ Missed Sample Up-Link	$\sigma^2 = 0.01$ w/ 2 RIS Up-Link	$\sigma^2 = 0.1$ Down-Link	$\sigma^2 = 0.5$ Down-Link	Multi-path Up-Link/ Down-Link
Longitude (m)	0.005 σ 0.003	0.007 σ 0.005	0.005 σ 0.003	0.005 σ 0.003	0.005 σ 0.003	0.276 σ 0.274
Latitude (m)	0.005 σ 0.003	0.034 σ 0.028	0.006 σ 0.003	0.006 σ 0.003	0.006 σ 0.003	0.310 σ 0.310
Altitude (m)	0.29 σ 0.17	2.055 σ 1.696	0.29 σ 0.17	0.29 σ 0.17	0.29 σ 0.17	0.300σ 0.301
Velocity East (m/s)	1.16 σ 0.12	1.16 σ 0.15	1.16 σ 1.15	1.16 σ 1.15	1.16 σ 1.15	1.16 σ 1.15
Velocity North (m/s)	1.84 σ 0.14	1.84 σ 0.14	0.89 σ 1.81	1.84 σ 1.08	1.84 σ 1.08	1.84 σ 1.08
Velocity Up(m/s)	0.12 σ 0.01	0.12 σ 0.12	1.84 σ 1.08	0.12 σ 0.12	0.12σ 0.12	0.12 σ 0.12

The results in Table 4.5 show that RIS-based positioning is largely insensitive to AWGN across all tested noise variances, with only minor degradation observed in position and velocity RMSE. This is primarily because of the use of the downlink channel being used. With the use of the down-link channel, which has a higher transmitted power, the received signal is above the noise floor. Missed samples introduce moderate performance loss, most notably in the latitude and altitude components. This is shown in Figure 4.18, where there are deviations in the trajectory, but the overall trajectory remains close to the reference trajectory. In contrast, multipath has a pronounced impact on positioning accuracy, causing substantial increases in position RMSE, particularly in the longitudinal and latitude directions. These results indicate that while noise and sampling imperfections are well tolerated, multipath propagation remains the dominant limiting factor for accurate RIS-based positioning.

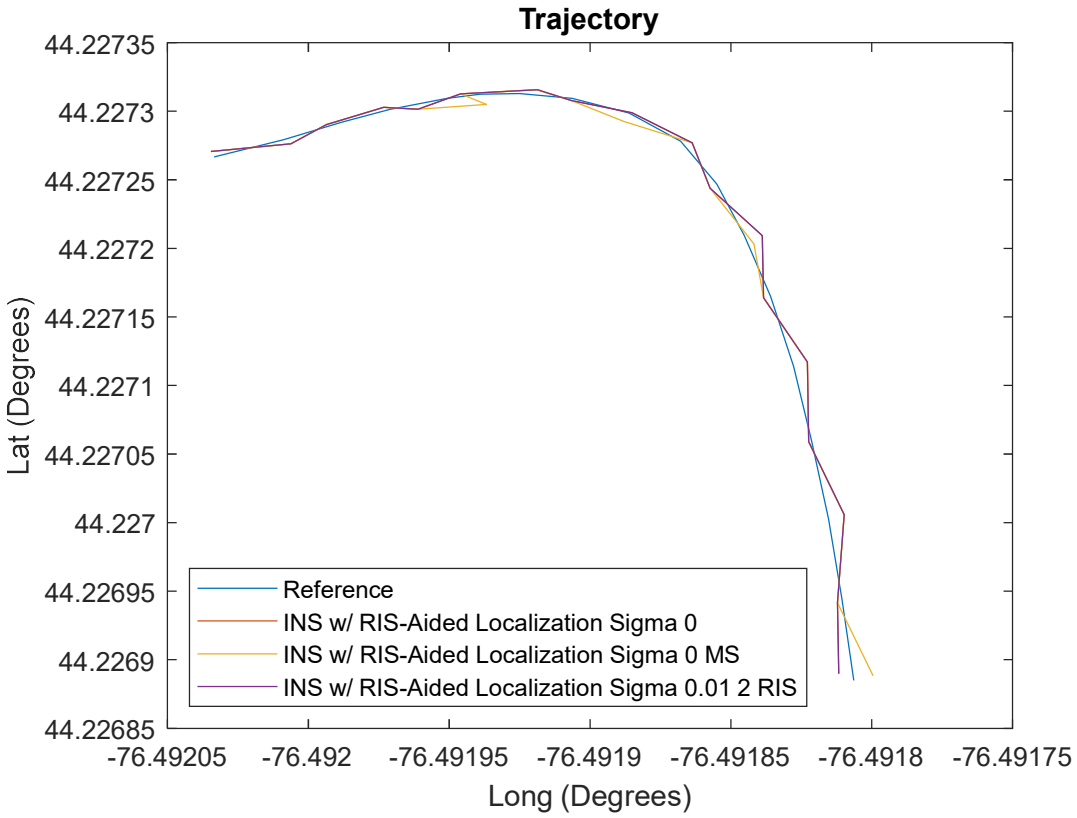


Figure 4.18: First section of GNSS Outage with RIS-Aided Positioning (MS=Missed Samples)

For each run of the trajectory the best method for positioning was used to achieve the RMSE. A breakdown of the differences between BSGS RL and UL and DL can be seen in Table 4.6. Table 4.6 provides the RMSE of RIS-aided positioning only. It does not include the RMSE of the integrated solution. Only how the raw RIS-aided positioning compares to the reference trajectory.

Table 4.6: RMSE breakdown of DL, UL, BSGS and RL

Positioning Approach	RMSE (m) σ 0	RMSE (m) w/ 1 random missed sample σ 0	RMSE (m) σ 0.01 2 RIS UL	RMSE (m) σ 0.1	RMSE (m) σ 0.5	RMSE (m) w/ 1 random missed sample and σ .01	RMSE (m) Multipath High reflections
RL Uplink	0.5103	2.8587	6.4648	NA	NA	NA	0.512
RL Downlink	0.7520	2.9745	0.6975	0.6975	0.6975	0.9815	0.686
BSGS Uplink	0.6203	2.8419	8.0890	NA	NA	NA	2.934
BSGS Downlink	5.3694	6.2598	0.7803	0.7803	0.7919	1.0118	2.438

During the simulation runs depicted in Table 4.6, what was seen with a multipath channel was that there was no significant difference between the BSGS and RL. However, this is true only in a weak multipath channel. If the main beam were to be attenuated or there was constructive interference with the reflections, there may be instances where a reflection has a higher received power than the main beam. Therefore, to test this, the multipath channel was changed so that ten percent of the time, one of the reflections would have a higher received power than the main channel.

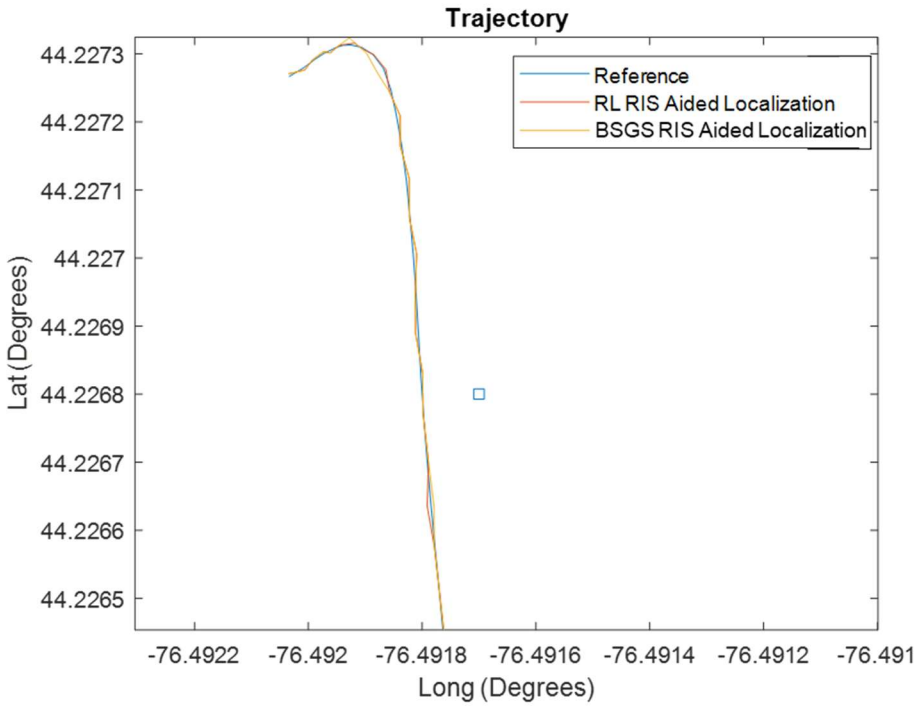


Figure 4.19: RL vs BSGS trajectory with σ = 0.

A comparison of the positional accuracy of BSGS and RL, as presented in Table 4.6, indicates that both approaches achieve comparable performance in non-complex channel conditions. Specifically, when the received signal remains above the noise floor and multipath effects are absent, BSGS and RL produce nearly identical localization results.

When missed samples are introduced, the RL agent demonstrates a slight performance improvement over BSGS due to its ability to adapt to dynamic conditions. However, this improvement is not pronounced for two primary reasons. First, only a single missed sample could occur at any given time, limiting deviation from the reference trajectory and allowing BSGS to recover effectively. Second, missed samples occurred only 10% of the time, resulting in a relatively small number of deviations throughout the outage period and reducing the likelihood of significant accumulated error. It is expected that increasing either the frequency or severity of missed samples would amplify the performance advantages of the RL approach. This could not be fully explored in the simulation due to the limited number of samples in the simulated environment.

In contrast, under multipath conditions, RL demonstrated a measurable improvement over BSGS, particularly when a dominate reflection is present. This performance gain is attributed to the RL agent's ability to conduct a more refined and adaptive search of the environment compared to the fixed structure of BSGS.

Overall, these results indicate that while BSGS and RL perform similarly in ideal conditions, the RL based approach provides superior performance in dynamic and complex channel environments.

Although BSGS and RL demonstrate comparable positional accuracy in an ideal channel, their RIS access efficiency differs significantly. In the BSGS-based localization system, the number of RIS accesses grows directly with the size of the search grid. For the first outage that lasts two minutes, there is a significant number of required measurements. In 2D, the full grid must be scanned, leading to 21,488 RIS accesses. Estimating the height component adds another 1,460 measurements, bringing the total to 22,948 RIS accesses for full 3D localization. This illustrates the fundamental limitation of BSGS in this context. The BSGS does not adapt its measurement strategy and instead evaluates every point in the search space, regardless of whether those measurements contribute meaningful information.

In contrast, the RL approach dramatically reduces the measurement burden by learning where to sample instead of exhaustively scanning. When no samples are missed, the RL agent requires only 5,880 RIS accesses to complete 2D localization. Extending to 3D keeps the access count at 6,237 measurements for the full position estimate. This is more than an order of magnitude fewer measurements than BSGS,

yet it still converges to the correct location because the agent learns to focus only on the regions of the grid that contribute to improving the UE position.

Even when the environment becomes less reliable such as when 10% of the RIS samples are randomly missed the RL agent still maintains a low measurement footprint. Under this condition, the 2D access count increases only to 6,264, and the 3D access count increases to 6,621. The increase is modest because the policy compensates for uncertainty by gathering a few additional samples, but it still completely avoids the full-grid behavior seen in BSGS. This demonstrates not only robustness to missing data but also the fundamental efficiency of the RL approach.

In uncertain environments, such as the high reflection multipath channels tested the RL agent remains robust to misleading measurements and avoids being dominated by false reflections. In contrast, the BSGS algorithm takes a holistic approach and evaluates all measurements uniformly and may incorrectly assess strong reflections as valid position estimates.

Overall, this comparison shows that the RL-based localization framework provides the same localization capability in an ideal channel, and better performance in complex channels both with drastically fewer RIS accesses. BSGS relies full grid evaluation, while RL learns an adaptive, information driven measurement pattern.

5 Laboratory Setup and Experimental Validation of RIS-Aided Localization

5.1 Introduction

Employing the system model from Figure 3.2 (a) in the laboratory as illustrated in Figure 5.1 the RIS-aided localization approach was tested. With a simulated BS located at the origin the RIS was placed one meter directly in front of the BS. The UE was then placed at various locations around the BS and the ability of our approach was used to localize the UE. With our setup the BS was used as the origin [0,0,0] in our 3D environment with the RIS located 1 meter away on the y-axis at the same elevation as the BS [0,1,0]. A Keysight (Agilent) N5244A PNA-X Network Analyzer (VNA) was used for the BS and UE, acting as the transmitter and receiver. The VNA had a power output of 0 dBm at 28 GHz and Aaronia PowerLOG 40400 horn antennas were connected to the VNA with Accu-Test R95-0004-36 cables.

Figure 5.1: Single RIS laboratory setup

5.2 Initial Localization and Bias Correction

To assess the localization accuracy a resolution of 0.05 and 0.2 meters for azimuth and elevation, respectively, was selected. The UE was initially placed at a known reference position of the localization of [0.4,0,0], corresponding to 40 cm away from the BS at the same elevation. The resulting localization output is depicted in Figure 5.2.

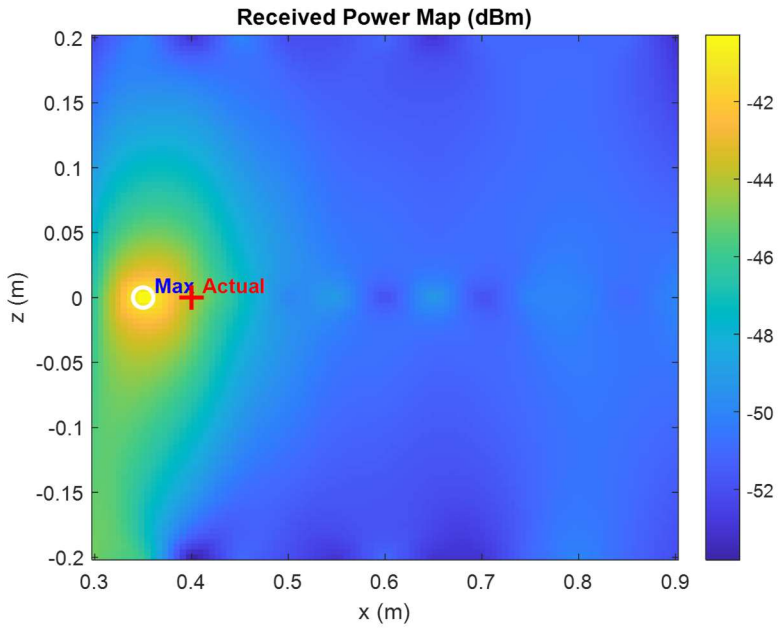


Figure 5.2: Laboratory localization at [0.4,0,0]

As observed in Figure 5.2, the UE position exhibits a bias towards the BS due to misalignment or misplacement. Consequently, the power maximum is not at the actual position. To compensate for this systematic error, a bias correction factor is introduced. To find the bias correction factor multiple measurements are performed at known locations and the error is averaged in the x, y, and z axis. This bias correction can be considered a bias correction initialization and needs to be performed whenever a new setup is established, specifically with the BS and RIS. Despite significant effort being taken to align the BS, and RIS errors in geometry are unavoidable. For the initial setup it was found that the bias was [-.05,0,0].

5.3 2D Localization Performance

The simulated environment aligns closely with the laboratory results as illustrated in Figure 5.3. In both cases, a sharp received power maximum is observed at the UE location with corresponding gradient drop off.

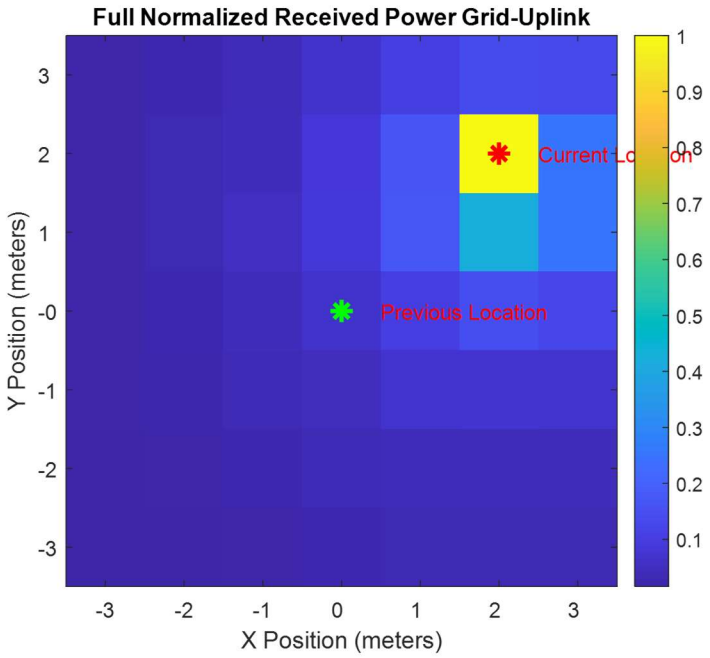
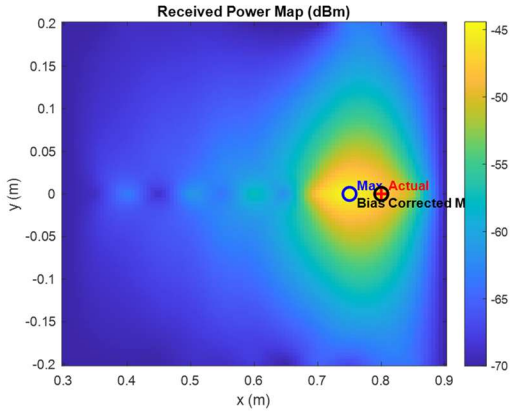


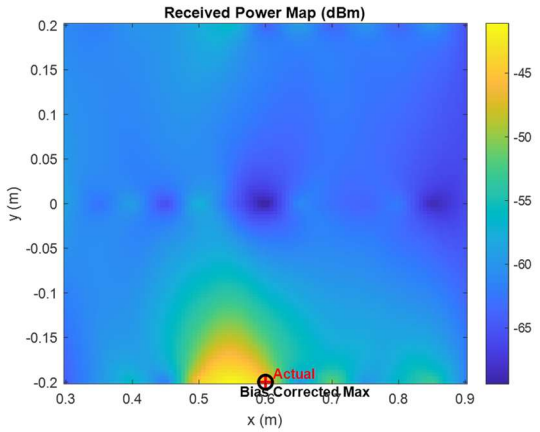
Figure 5.3: Sample output from simulated environment

Laboratory localization results for multiple UE positions are summarized in Figure 5.4, including varying horizontal and vertical positions.

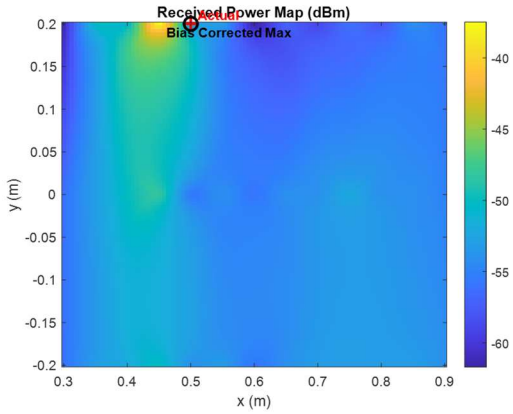
UE Location: [0.8,0,0] (a)



UE Location: [0.6,0,-0.2] (b)



UE Location: [0.5,0,0.2] (c)



UE Location: [0.5,0,0] (d)

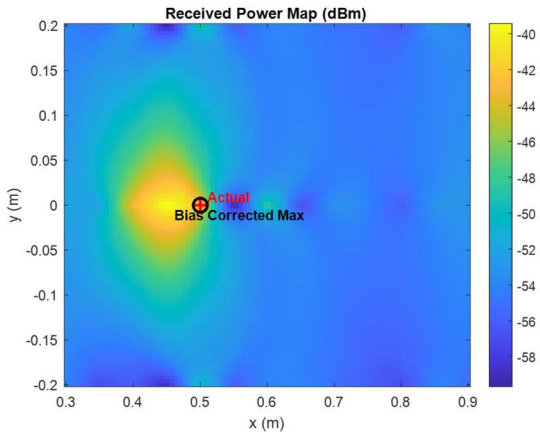
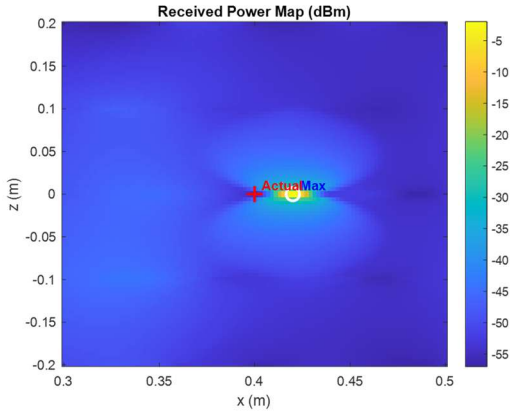


Figure 5.4: Laboratory output from various UE locations

Further experiments were conducted to evaluate the resolution limits of the localization setup. First, the azimuth resolution was examined using a resolution of 0.01m and 0.1 for elevation. The UE was positioned at [0.4,0,0] with bias correction factor applied.

UE Location: [0.4,0,0]



Bias correction: [-0.02,0,0]

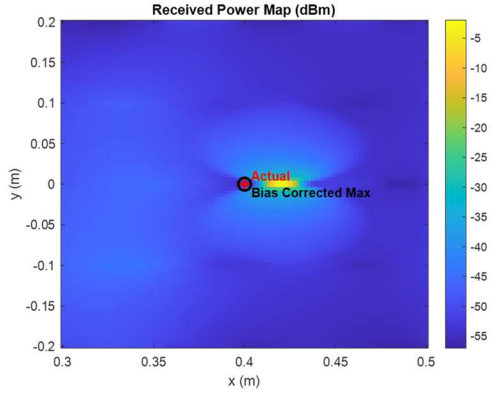


Figure 5.5: Azimuth resolution limit testing

As shown in Figure 5.5 a high precision in the azimuth is achievable with localization errors on the order of 2 cm. The azimuth resolution is constrained by hardware limitations where increments of only 1 degree is possible.

Elevation resolution was evaluated separately, as shown in Figure 5.6.

UE Location: [0.4,0,0]

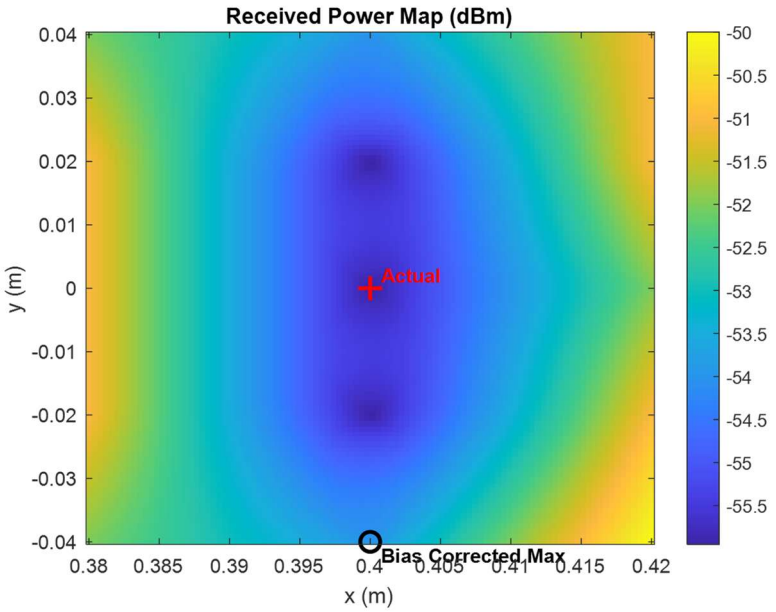
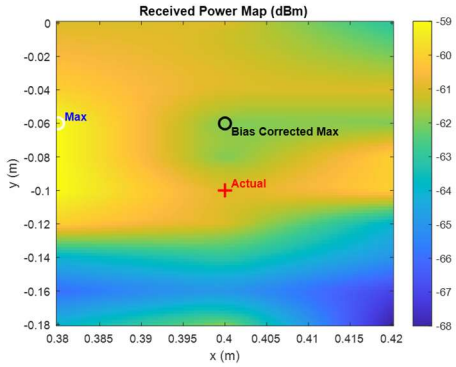


Figure 5.6: Elevation resolution limit testing

The elevation estimation exhibits reduced accuracy, with localization errors of approximately 6 cm. It is suspected that a lower elevation resolution is a result of the Aaronia PowerLOG 40400 horn antennas having a smaller width dimension than the length. Based off these observations, a coarse elevation resolution of 10 cm and a fine azimuth resolution of 2 cm were adopted for subsequent experiments.

Using the established resolution limits, additional localization measurements were performed, as illustrated in Figure 5.7.

UE Location: [0.4,0,-0.1]



UE Location: [0.4,0,-0.1]

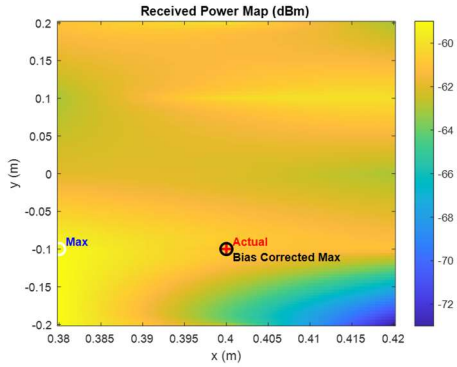


Figure 5.7: Laboratory localization with resolution limits

The 2D RMSE of the RIS-aided localization approach was computed across multiple measurements. Since the bias factor can cause significant error in positional error it is important to account for it in the RMSE. Therefore, Table 5.1 established the uncorrected and corrected 2D RMSE of this RIS-aided localization.

Table 5.1: 2D RIS-Aided localization RMSE

2D RMSE without bias correction	**2D RMSE with bias correction**
0.04m σ0.2mm	0.02m σ0.3mm

5.4 3D Localization Performance

To evaluate our 3D localization approach, a beam sweeping experiment was conducted. Bias correction was accounted for in the initialization of the experimentation. A resulting 3D localization output is shown in Figure 5.8.

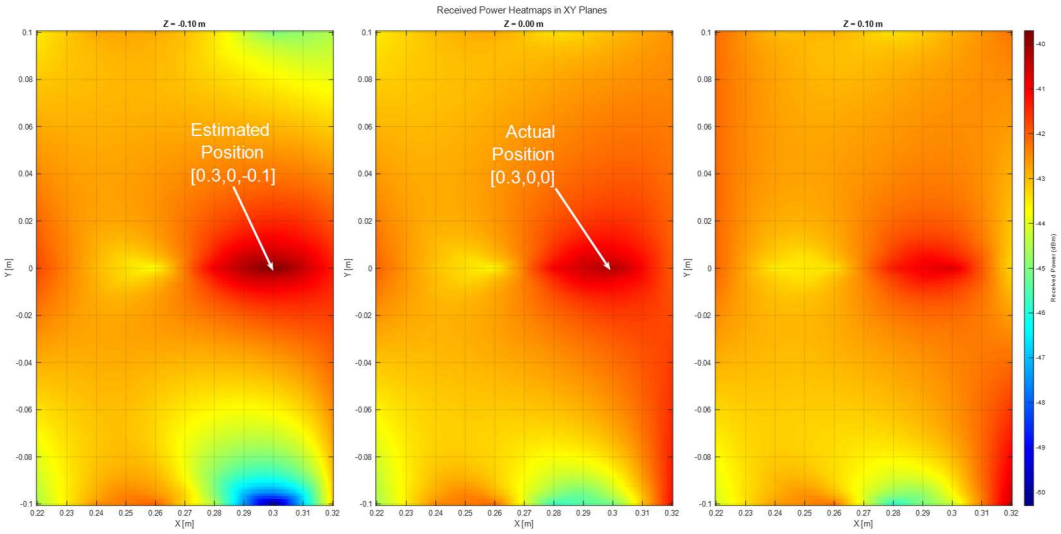


Figure 5.8: 3D RIS-aided localization using beam sweeping

Using a single RIS, a 3D RMSE of 0.10m σ0.01mm was achieved, demonstrating that beam sweeping enables accurate 3D localization.

5.5 Near Field Operation and Curvature of Arrival

Despite the high localization accuracy achieved using the beam sweeping approach, it is important to assess whether the experimental setup operated in the near field or far field. Using (2.3) with D being the diagonal of the RIS, the near field and far

field boundaries were analyzed. As illustrated in Figure 5.9 the far-field is never reached.

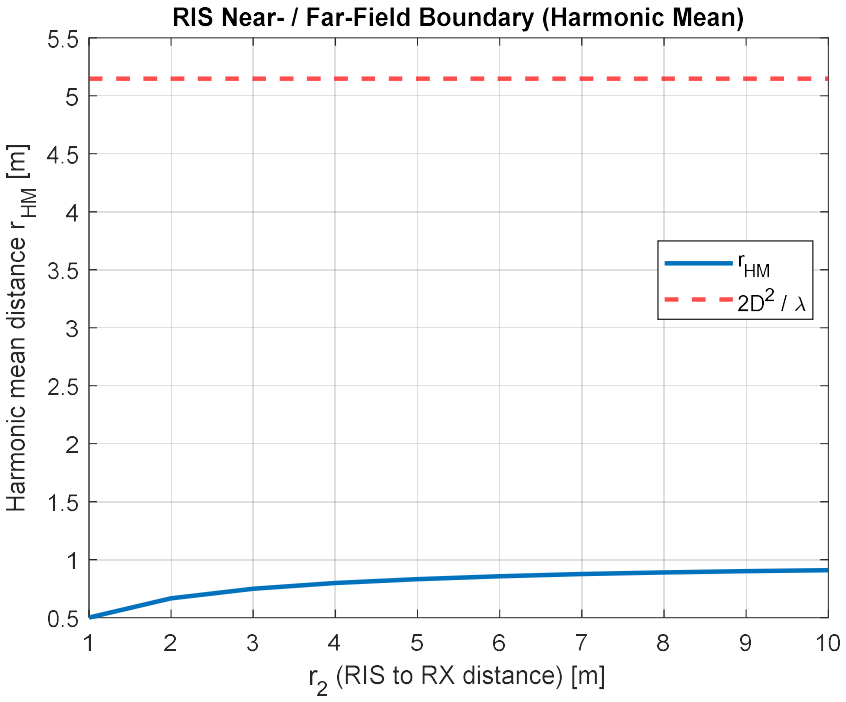


Figure 5.9: Laboratory Near Field Analysis

Consequently, the use of the CoA can be exploited to estimate the range between the RIS and the UE. Three-dimensional RIS-aided localization can be achieved without explicit beam sweeping.

By only changing the range variable with a stationary UE and using the established codebook sweeping approach for the range between the RIS and the UE it can be shown that there is a change in the received power at the UE, as illustrated in Figure 5.10.

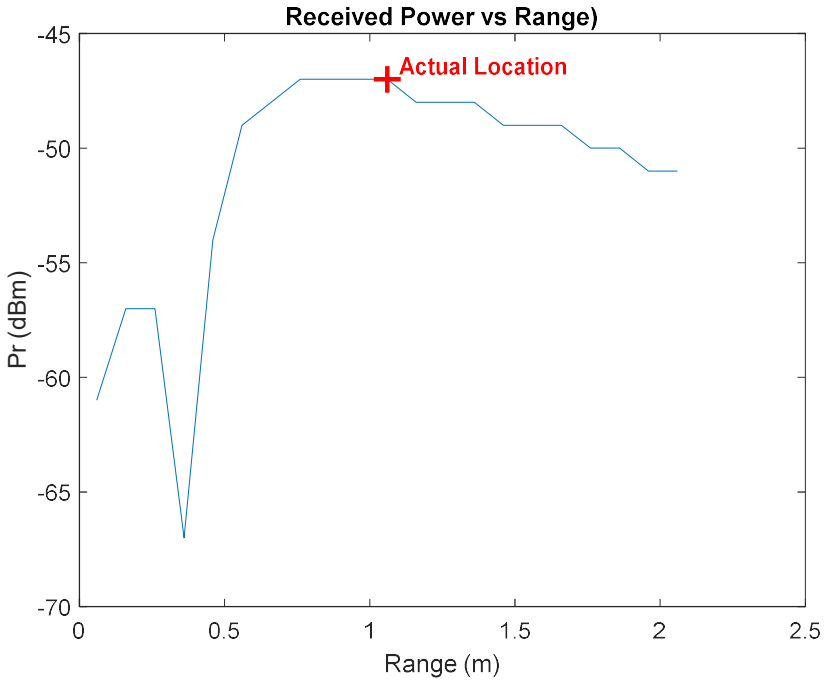


Figure 5.10: Received power at UE with RIS range sweeping

Figure 5.10 shows the received power as the RIS range variable is swept across where the UE is located. The maximum power is achieved at 1 m, where the UE is located.

By combining the 2D localization approach with the CoA based range approximation, 3D RIS-aided localization can be achieved. Experimental results yield a 3D RMSE of 0.12m σ14mm. A comparison between the beam sweeping approach and CoA based localization is provided in Table 5.2.

Table 5.2: RIS Localization approaches comparison

Beam Sweeping 3D RMSE	**CoA 3D RMSE**
0.10m σ 0.01mm	0.12m σ14mm

While the beam sweeping based localization provided a more accurate position estimate the number of accesses to the RIS is more, compared to the CoA-based localization. Using the laboratory runs performed, it was found that the CoA-based localization only used the RIS approximately 81% of the total access of that of the beam sweeping localization. Therefore, there is trade-off of accuracy of localization, and access time to the RIS.

5.6 Dual-RIS Localization Evaluation

To evaluate the effectiveness of a dual RIS configuration, a new experimental setup conforming to Figure 3.2 (b) was implemented. With a separation of 1.5 m between the RIS and UE/BS, localization accuracy degrades at larger distances due to reduced angular and range sensitivity. To assess the effects on each RIS, experiments were first conducted with only one RIS at a time; the results are illustrated in Figure 5.11.

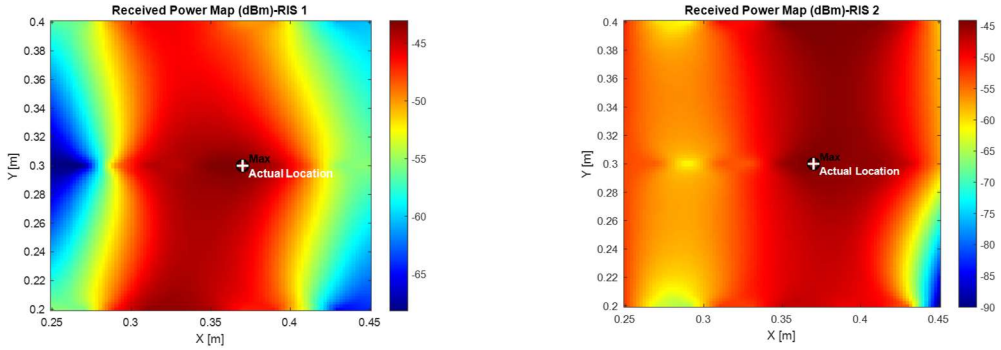


Figure 5.11: Single RIS Localization at a separation distance of 1.5m

The same experimental setup was then employed with the dual-RIS, with the results in Figure 5.12.

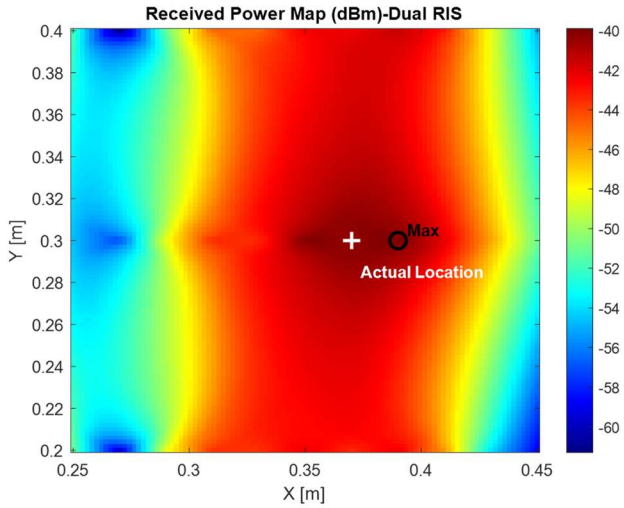


Figure 5.12: Dual RIS Localization at a separation distance of 1.5m.

An additional test of the dual-RIS configuration was performed with a separation distance of 1 m. When both RISs are used, a significantly sharper received peak

power is observed at the UE location, as shown in Figure 5.12. Specifically, in Figure 5.12, the maximum received power increases from -43 dBm to -39.8 dBm, resulting in a 3 dBm improvement. Another effect of using dual RIS is a smoother power map. While the maximum power is more distinct in the dual setup, there is more stability in the power map with deeper nulls, and the power stays within 15 dBm across the grid search.

The 2 RIS setup acts also acts effectively as an enlarged RIS with 64 by 32 elements. This is evident in the experimentation in Figure 5.12, where increased spatial fidelity is observed along the horizontal axis compared to the vertical axis.

Using the beam-sweeping approach with the dual RIS configuration, a positional RMSE of 0.1m (σ = 0.08 mm) was achieved. Furthermore, since the proposed experimental localization operates in the near field, the CoA-based localization with two RISs can be exploited. A positional RMSE of 0.29m σ0.04m was achieved. The higher RMSE of the CoA approach is primarily due to errors in the range estimate, which affect the calculated z-position. In the dual RIS configuration, the overall positional RMSE is less than the single RIS configuration. It is suspected that the reason for this was because of the more complex setup of the equipment and resulting configurations of the RIS. In the setup, there could be more misalignments and misplacements of equipment compounding any errors and cascading over to the generation of the reflection patterns for each RIS.

5.7 BSGS vs RL Localization

With the fundamental approach to localization validated through experimentation, the BSGS and RL algorithms can be applied. In the employment of the RL and BSGS algorithms, they are applied to the laboratory results, rather than during laboratory experimentation. That is, they are applied offline rather than real-time. This approach was used to simplify the laboratory setup and ease of troubleshooting. The use of the RL requires a previously known velocity. To provide a velocity to the algorithm, the known position is used as a baseline velocity, and a random offset is applied; thereby, simulating a previously know velocity. Applying both the single and dual RIS 3D results, the resulting 3D localization with BSGS and RL algorithms for the beam sweeping and CoA are:

Table 5.3: 3D CoA RIS-Aided localization BSGS and RL RMSE

BSGS 3D RMSE w/ CoA Localization	**RL 3D RMSE w/ CoA Localization**
0.17m σ18mm	0.04m σ0.9mm

Table 5.4: 3D Beam Sweeping RIS-Aided localization BSGS and RL RMSE

BSGS 3D RMSE w/ Beam Sweeping Localization	**RL 3D RMSE w/ Beam Sweeping Localization**
0.10m σ2.6mm	0.08m σ1.7mm

Table 5.5: 3D RIS-Aided localization BSGS and RL RMSE with all laboratory runs

All Lab Runs 3D RMSE w/ BSGS-Based Localization	**All Lab Runs 3D RMSE w/ RL-Based Localization**
0.15m σ11mm	0.06m σ1.7mm

The laboratory results demonstrate a consistent advantage of the RL-based localization over the conventional BSGS approach, particularly for CoA localization, as shown in Table 5.3. Much like what is observed in Chapter 4, when introduced to complex environment RL outperforms BSGS. Given the more complex environment of the laboratory compared to the initial simulation environments generated in Chapter 4, the RL agent can adapt to the environment compared to BSGS.

With beam-sweep localization, outlined in Table 5.4 it is shown that the performance between BSGS and RL is much more similar. This can be accounted for by the more limited search area of beam sweeping, particularly in the z-axis. When beam sweeping is used, only a limited number of steps along the z-axis are taken, compared to the CoA, which ranges in and out with lower resolution and more samples. Thus, with the CoA-based approach, the search area is larger and more uncertain, leading to differences in results between BSGS and RL-based localization with the CoA and beam sweeping approaches.

Considering all laboratory runs collectively, illustrated in Table 5.5, the overall trend remains consistent. The RL-based localization yields a better localization estimate and is more consistent across runs. This suggests that RL-based localization is more resilient to environmental inconsistencies and errors than BSGS-based localization.

When comparing the BSGS results to those of the RL, the positional accuracy is not the only factor to consider. The number of RIS accesses is a crucial metric because of the integration time required for each localization. The more time it takes to localize a UE, the more RIS time is taken away from other users. Therefore, when analysing the number of RIS accesses using the laboratory samples, with the post-processed employment of BSGS and RL algorithms, it was found that for CoA localization, the RL algorithm used the RIS approximately 32% of the time

compared to the BSGS algorithm. In beam-sweep localization, the RL algorithm used the RIS approximately 20% of the time, compared to the BSGS algorithm.

5.8 Overall Results Summary

As shown in the results of the simulated environment, the proposed RL-based RIS-aided localization for positioning achieved more robust performance than BSGS. In higher-noise and multipath environments, the RL RIS-aided localization maintained stable, low-error bounds. In our simulation, we showed that the RL-based approach produced an average RMSE of 2.4m in the uplink, better than that of the BSGS in a multipath channel, thereby demonstrating the viability of the RL-based approach for RIS-aided localization for positioning in complex channels. The RIS access were also significantly reduced with the use of the RL based localization. In our simulation, over the course of the two minutes where RIS-aided localization was used, the RL RIS-aided localization accessed the RIS 20% of the number of BSGS RIS accesses. Significantly reducing a single user's time required with the RIS for localization.

Extending our RIS-aided localization to a controlled laboratory with a physical RIS, the proposed localization was validated, and BSGS and RL results were compared. Demonstrating the validity of the proposed RL RIS-aided localization again, the use of RL achieved an RMSE that was 0.13m better than that of BSGS for CoA-based localization and 0.02m for beam-sweeping-based localization. This is because it was able to have a more refined search area compared to BSGS, reducing the potential for ambiguity in positional estimation. Further validating our simulated results for RIS access efficiency in our laboratory results the RL RIS-aided localization accessed the RIS 32% of the number of accesses for BSGS with CoA based localization and with beam sweeping the RL RIS-aided localization accessed the RIS 20% of the amount compared to the BSGS RIS-aided localization.

It can be concluded that RL provides an effective, adaptive mechanism for RIS-aided 3D localization, where it outperforms traditional BSGS in accuracy, and robustness in complex channels and RIS access efficiency.

6 Conclusion

In closing, this chapter will recap the key material from earlier sections, emphasize the main contributions of the thesis, evaluate the thesis hypothesis and its outcomes, and outline potential directions for future research.

6.1 Summary

In Chapter 1, the subject of RIS-aided localization was introduced, with RL as a potential adaptive control mechanism for determining a UE's positional information. A problem statement, a motivation for the work, and a thesis statement were also presented.

In Chapter 2, the topics of RIS, RIS-aided localization, and AI were discussed. Their individual and collective opportunities and challenges were outlined. These included lower deployment costs for RIS compared to BS, the ability to provide higher network reliability and coverage, and, finally, dual-purpose use for enhanced telecommunications and localization. Their challenges of low phase quantization, network architecture, and non-convex optimization were discussed, along with how AI can be used to overcome some of the challenges. The challenges associated with AI use were also discussed.

Chapter 3 presented the proposed system model and the developed RL architecture and BSGS algorithm. A simplified environment was created for the fundamental testing of our RIS-aided localization approach and RL architecture.

Chapter 4 implemented the proposed system model in a second environment that used real-world trajectory data to validate our approach for RIS-aided localization for positioning. The results from the simulated real-world trajectory were discussed, highlighting the advantages of the proposed RL architecture over the BSGS algorithm.

In Chapter 5, the RIS-aided localization was extended to a controlled laboratory setting with a physical RIS, our approach was validated with our testing of a single and dual RIS setup.

6.2 Contributions

This thesis makes several contributions to the field of RIS-aided localization, focusing on the use of reinforcement learning for three-dimensional positioning in GNSS-denied environments.

First, a reinforcement learning based RIS configuration framework for 3D localization is developed and demonstrated. Unlike traditional RIS-aided localization approaches that rely on pre-defined channel models (MLE) or exhaustive scanning (BSGS), the proposed method enables the RIS to configure its beam pattern solely based on the received power. This significantly reduces reliance on accurate channel-state information and environmental assumptions.

Second, a hierarchical RL architecture is proposed to address the search problem inherent in RIS beam sweeping. By breaking down the localization task into a fine search stage with kinematic prediction, the approach enables efficient localization.

Third, the impact of practical RIS limitations on localization performance is investigated. In particular, the effects of phase quantization, grating lobe formation, beamwidth, and beam pointing errors are analyzed in the context of localization. These results provide insight into how hardware constraints directly affect localization accuracy and highlight scenarios in which traditional power-based localization methods degrade.

Fourth, the proposed RL-based RIS localization framework is validated using real-world vehicle trajectory data in a GNSS-denied scenario. By integrating RIS-based position updates into a loosely coupled Kalman filter, the results demonstrate improved positional stability during GNSS outages when compared to traditional methods.

Finally, the feasibility of RL-configured RIS localization is demonstrated experimentally using millimeter-wave RIS hardware in a laboratory environment. The laboratory results confirm that the proposed approach can achieve consistent 2D and 3D localization performance under realistic measurement noise, resolution limits, and hardware constraints. This experimental validation represents a key step toward practical deployment of RL-assisted RIS localization systems.

Collectively, these contributions demonstrate that reinforcement learning provides a viable and flexible alternative to conventional RIS-aided localization techniques, enabling accurate and adaptable positioning in complex and GNSS-denied environments.

6.3 Discussion

In Chapter 1, the following thesis statement was identified:

“Reinforcement Learning provides an effective, adaptive mechanism for RIS-aided 3D localization in GNSS-denied environments, where it outperforms traditional localization algorithms in accuracy, robustness and RIS access efficiency.”

To verify the hypothesis that reinforcement learning can effectively configure an RIS for accurate localization, the methodologies described in Chapter 3 were developed to support both traditional and reinforcement learning-based localization approaches. Specifically, a power maximization framework using beam sweeping was implemented, with Beam Sweep Grid Search as a baseline, while a reinforcement learning-based approach was introduced to reduce the search space and adaptively configure the RIS.

The analysis of phase quantization effects identified key limitations of power-based localization approaches. Strong grating lobes and beam ambiguity introduced by low-bit phase resolution were shown to degrade localization accuracy when left unaddressed. The RL-based framework inherently mitigates some of these effects by learning search strategies that exploit spatial consistency and prior motion information, rather than using a holistic power measurement estimation, which BSGS relies on.

In Chapter 4 and 5, the performance of the proposed RL-based localization framework was evaluated using simulated environments, real-world vehicle trajectory data in GNSS-denied scenarios, and laboratory measurements using mmWave RIS hardware. Localization accuracy, robustness during GNSS outages, and sensitivity to noise, multipath, and missed samples were used as the primary measures of performance. Successful validation was defined as the ability of the RL-based approach to achieve localization accuracy comparable to or better than traditional methods while reducing RIS accesses.

The results demonstrate that the reinforcement learning based RIS localization approach consistently achieves accurate position estimates in both 2D and 3D scenarios. In simulated real-world trajectory experiments, the RL-assisted RIS localization maintained positional stability during GNSS outages and reduced errors compared to the R-IMU-only solution. Compared to BSGS-based localization, the RL approach demonstrated improved robustness in the presence of multipath effects, and missed measurements, while requiring fewer scans of the search space.

Laboratory experiments further validate the feasibility of the proposed approach under practical hardware constraints. Despite limited angular resolution, phase quantization, and measurement noise inherent to the experimental setup, the RL-based localization framework produced consistent localization results across multiple UE positions. These results confirm that reinforcement learning can successfully adapt RIS configurations using received-power feedback alone, without relying on channel-state information.

Based on the results presented in Chapters 4 and 5, the thesis statement has been validated. It can be concluded that reinforcement learning provides an effective and adaptable mechanism for RIS-aided localization, enabling accurate 3D positioning

in GNSS-denied environments under realistic system constraints. The proposed approach demonstrates clear advantages over traditional methods in terms of adaptability, robustness, and practical deployability, supporting the use of RL-configured RIS as a viable solution for next-generation localization systems.

6.4 Future Work

In their work on near-field localization with Large Antenna Arrays (LAA) [31], they developed the state models for the Extended Kalman Filter and Particle Filter and evaluated their performance. Further work with our approach should build on this work and extend the work already performed by [31], and apply it to real-world trajectories and laboratory evaluations.

While this work developed a DRL algorithm for UE localization, it was designed as a framework that can be expanded, and further learning can occur. During algorithm development, agents of different sizes were generated. That is, the hierarchical agent may learn to use varying-sized search grids. This could be employed with a confidence score of the previously known velocity. The acceleration or standard deviation of previously known velocities can be used to determine the agent's search grid size. This would decrease the overall RIS accesses. Also, the DRL algorithm could be trained to develop a state model rather than relying on the loosely coupled Kalman filter currently employed.

Furthermore, in this work, only the maximum received power is used to localize. As illustrated, when the transmitted power is low, such as from the UE in the uplink scenario, the ability to localize the UE becomes challenging due to the higher noise floor. However, using a modulation technique with the PRS, such as OFDM, that is robust to AWGN can yield higher accuracy in a noisy environment.

Bibliography

[1] A. Noureldin, T. B. Karamat, and J. Georgy, *Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013. doi: 10.1007/978-3-642-30466-8.

[2] S. Keşir, M. Y. Yağan, I. Hökelek, A. E. Pusane, and A. Görçin, "Rapid CNN-Assisted Iterative RIS Element Configuration," in *2023 International Symposium on Networks, Computers and Communications (ISNCC)*, Doha, Qatar: IEEE, Oct. 2023, pp. 1–6. doi: 10.1109/ISNCC58260.2023.10323787.

[3] A. Paul, M. Katwe, K. Singh, A. Kaushik, G. C. Alexandropoulos, and C.-P. Li, "RIS-Empowered 3D DoA Estimation of Multiple Aerial Targets via Deep Reinforcement Learning," in *2025 IEEE Wireless Communications and Networking Conference (WCNC)*, Milan, Italy: IEEE, Mar. 2025, pp. 1–6. doi: 10.1109/WCNC61545.2025.10978492.

[4] E. Björnson, H. Wymeersch, B. Matthiesen, P. Popovski, L. Sanguinetti, and E. de Carvalho, "Reconfigurable Intelligent Surfaces: A Signal Processing Perspective With Wireless Applications," *IEEE Signal Process. Mag.*, vol. 39, no. 2, pp. 135–158, Mar. 2022, doi: 10.1109/MSP.2021.3130549.

[5] E. Björnson and Ö. T. Demir, *Introduction to Multiple Antenna Communications and Reconfigurable Surfaces*. Now Publishers, 2024. doi: 10.1561/9781638283157.

[6] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, "Intelligent Reflecting Surface-Aided Wireless Communications: A Tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3313–3351, May 2021, doi: 10.1109/TCOMM.2021.3051897.

[7] "Turkcell and ZTE conduct Europe's first dynamic RIS trial in a 5G Advanced mmWave network." Accessed: Apr. 10, 2026. [Online]. Available: https://www.zte.com.cn/content/zte-site/www-zte-com-cn/global/about/news/turkcell-and-zte-conduct-europes-first-dynamic-ris-trial-in-a-5g-advanced-mmwave-network

[8] H. Jiao, H. Liu, and Z. Wang, "Reconfigurable Intelligent Surfaces aided Wireless Communication: Key Technologies and Challenges," in *2022 International Wireless Communications and Mobile Computing (IWCMC)*, Dubrovnik, Croatia: IEEE, May 2022, pp. 1364–1368. doi: 10.1109/IWCMC55113.2022.9824117.

[9] M. Toka *et al.*, "RIS-Empowered LEO Satellite Networks for 6G: Promising Usage Scenarios and Future Directions," Feb. 12, 2024, *arXiv*: arXiv:2402.07381. Accessed: Nov. 19, 2024. [Online]. Available: http://arxiv.org/abs/2402.07381

[10] S. Zhang and R. Zhang, "Intelligent Reflecting Surface Aided Multiple Access: Capacity Region and Deployment Strategy," in *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Atlanta, GA, USA: IEEE, May 2020, pp. 1–5. doi: 10.1109/SPAWC48557.2020.9153884.

[11] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An Overview of Massive MIMO: Benefits and Challenges," *IEEE J. Sel. Top. Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014, doi: 10.1109/JSTSP.2014.2317671.

[12] R. T. Schwarz, T. Delamotte, K.-U. Storek, and A. Knopp, "MIMO Applications for Multibeam Satellites," *IEEE Trans. Broadcast.*, vol. 65, no. 4, pp. 664–681, Dec. 2019, doi: 10.1109/TBC.2019.2898150.

[13] W. Tang *et al.*, “Wireless Communications With Reconfigurable Intelligent Surface: Path Loss Modeling and Experimental Measurement,” *IEEE Trans. Wirel. Commun.*, vol. 20, no. 1, pp. 421–439, Jan. 2021, doi: 10.1109/TWC.2020.3024887.
[14] M. Delbari, G. C. Alexandropoulos, R. Schober, and V. Jamali, “Far- versus Near-Field RIS Modeling and Beam Design,” Jan. 30, 2024, *arXiv*: arXiv:2401.08237. doi: 10.48550/arXiv.2401.08237.
[15] M. Cui, Z. Wu, Y. Lu, X. Wei, and L. Dai, “Near-Field MIMO Communications for 6G: Fundamentals, Challenges, Potentials, and Future Directions,” *IEEE Commun. Mag.*, vol. 61, no. 1, pp. 40–46, Jan. 2023, doi: 10.1109/MCOM.004.2200136.
[16] “ETSI GR RIS 003 V1.1.1. ‘Reconfigurable Intelligent Surfaces (RIS); Communication Models, Channel Models, Channel Estimation and Evaluation Methodology.’”
[17] T. Zhou and K. Xu, “Multi-Intelligent Reflecting Surface-Aided Wireless Network With Achievable Rate Maximization,” in *2020 International Conference on Wireless Communications and Signal Processing (WCSP)*, Nanjing, China: IEEE, Oct. 2020, pp. 552–557. doi: 10.1109/WCSP49889.2020.9299714.
[18] Q. Wu and R. Zhang, “Intelligent Reflecting Surface Enhanced Wireless Network via Joint Active and Passive Beamforming,” *IEEE Trans. Wirel. Commun.*, vol. 18, no. 11, pp. 5394–5409, Nov. 2019, doi: 10.1109/TWC.2019.2936025.
[19] R. Xiong, X. Dong, T. Mi, K. Wan, and R. C. Qiu, “Optimal Discrete Beamforming of RIS-Aided Wireless Communications: An Inner Product Maximization Approach,” in *2024 IEEE Wireless Communications and Networking Conference (WCNC)*, Dubai, United Arab Emirates: IEEE, Apr. 2024, pp. 1–6. doi: 10.1109/WCNC57260.2024.10571003.
[20] T. Ma, Y. Xiao, X. Lei, L. Zhang, Y. Niu, and G. K. Karagiannidis, “Reconfigurable Intelligent Surface-Assisted Localization: Technologies, Challenges, and the Road Ahead,” *IEEE Open J. Commun. Soc.*, vol. 4, pp. 1430–1451, 2023, doi: 10.1109/OJCOMS.2023.3292052.
[21] Z. Zhang and S. Kang, “Time of arrival estimation based on clustering for positioning in OFDM system,” *IET Commun.*, vol. 14, no. 15, pp. 2584–2591, Sep. 2020, doi: 10.1049/iet-com.2019.0943.
[22] “NR Positioning Using PRS.” Accessed: Mar. 14, 2025. [Online]. Available: https://www.mathworks.com/help/5g/ug/nr-prs-positioning.html
[23] K. Keykhosravi *et al.*, “Leveraging RIS-Enabled Smart Signal Propagation for Solving Infeasible Localization Problems: Scenarios, Key Research Directions, and Open Challenges,” *IEEE Veh. Technol. Mag.*, vol. 18, no. 2, pp. 20–28, Jun. 2023, doi: 10.1109/MVT.2023.3237004.
[24] M. Ammous and S. Valaee, “Positioning and Tracking Using Reconfigurable Intelligent Surfaces and Extended Kalman Filter,” in *2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, Helsinki, Finland: IEEE, Jun. 2022, pp. 1–6. doi: 10.1109/VTC2022-Spring54318.2022.9860999.
[25] Z. Feng, B. Wang, Y. Zhao, M. Luan, and F. Hu, “Power optimization for target localization with reconfigurable intelligent surfaces,” *Signal Process.*, vol. 189, p. 108252, Dec. 2021, doi: 10.1016/j.sigpro.2021.108252.
[26] A. Fascista, M. F. Keskin, A. Coluccia, H. Wymeersch, and G. Seco-Granados, “RIS-Aided Joint Localization and Synchronization With a Single-Antenna Receiver: Beamforming Design and Low-Complexity Estimation,” *IEEE J. Sel. Top. Signal Process.*, vol. 16, no. 5, pp. 1141–1156, Aug. 2022, doi: 10.1109/JSTSP.2022.3177925.

[27] M. Rahal *et al.*, “Performance of RIS-aided near-field localization under beams approximation from real hardware characterization,” *EURASIP J. Wirel. Commun. Netw.*, vol. 2023, no. 1, p. 86, Aug. 2023, doi: 10.1186/s13638-023-02294-9.
[28] A. Alhafid, Y. Ali, and S. Younis, “Far-Field Localization for RIS Empowered Wireless Systems Leveraging Beamforming,” *Jordan J. Electr. Eng.*, vol. 10, no. 3, p. 1, 2024, doi: 10.5455/jjee.204-1701118386.
[29] A. Nasri, A. H. A. Bafghi, and M. Nasiri-Kenari, “Wireless Localization in the Presence of Intelligent Reflecting Surface,” *IEEE Wirel. Commun. Lett.*, vol. 11, no. 7, pp. 1315–1319, Jul. 2022, doi: 10.1109/LWC.2022.3165427.
[30] A. E. Asl, S. M. Karbasi, and H. Behroozi, “Reconfigurable Intelligent Surface-Aided Emitter Localization,” *IEEE Sens. J.*, vol. 22, no. 22, pp. 21866–21876, Nov. 2022, doi: 10.1109/JSEN.2022.3209983.
[31] A. Guerra, F. Guidi, D. Dardari, and P. Djuric, “Near-Field Tracking With Large Antenna Arrays: Fundamental Limits and Practical Algorithms,” *IEEE Trans. Signal Process.*, vol. 69, pp. 5723–5738, 2021, doi: 10.1109/TSP.2021.3101696.
[32] M. Haghshenas, P. Ramezani, M. Magarini, and E. Björnson, “Parametric Channel Estimation With Short Pilots in RIS-Assisted Near- and Far-Field Communications,” *IEEE Trans. Wirel. Commun.*, vol. 23, no. 8, pp. 10366–10382, Aug. 2024, doi: 10.1109/TWC.2024.3371715.
[33] Y. Li and W. Yu, “Localization in Multipath Environments via Active Sensing With Reconfigurable Intelligent Surfaces,” *IEEE Commun. Lett.*, vol. 28, no. 9, pp. 2061–2065, Sep. 2024, doi: 10.1109/LCOMM.2024.3421370.
[34] M. Rahal *et al.*, “Performance of RIS-aided near-field localization under beams approximation from real hardware characterization,” *EURASIP J. Wirel. Commun. Netw.*, vol. 2023, no. 1, p. 86, Aug. 2023, doi: 10.1186/s13638-023-02294-9.
[35] J. Luo, T. Liang, C. Chen, and T. Zhang, “A UAV mounted RIS aided communication and localization integration system for ground vehicles,” *IEEE ICC*, 2022.
[36] J. Zhang, Z. Zheng, Z. Fei, and Z. Chang, “RIS-Assisted Multi-User Localization in UAV-Enabled mmWave Wireless Networks”.
[37] A. L. Sarker, W. Son, and D. S. Han, “RIS-Assisted Hybrid Beamforming and Connected User Vehicle Localization for Millimeter Wave MIMO Systems”.
[38] M. Naeem, S. T. H. Rizvi, and A. Coronato, “A Gentle Introduction to Reinforcement Learning and its Application in Different Fields,” *IEEE Access*, vol. 8, pp. 209320–209344, 2020, doi: 10.1109/ACCESS.2020.3038605.
[39] P. Chen, X. Li, M. Matthaiou, and S. Jin, “DRL-Based RIS Phase Shift Design for OFDM Communication Systems,” *IEEE Wirel. Commun. Lett.*, vol. 12, no. 4, pp. 733–737, Apr. 2023, doi: 10.1109/LWC.2023.3242449.
[40] G. Eappen, J. Cosmas, S. T, R. A, R. Nilavalan, and J. Thomas, “Deep learning integrated reinforcement learning for adaptive beamforming in B5G networks,” *IET Commun.*, vol. 16, no. 20, pp. 2454–2466, Dec. 2022, doi: 10.1049/cmu2.12501.
[41] P. Ramezani, M. A. Girnyk, and E. Björnson, “Dual-Polarized Reconfigurable Intelligent Surface-Assisted Broad Beamforming,” *IEEE Commun. Lett.*, vol. 27, no. 11, pp. 3073–3077, Nov. 2023, doi: 10.1109/LCOMM.2023.3317439.
[42] B. Zhang, W. Liu, and C. Wang, “Two-dimensional Antenna Array Design for Directional Modulation”.

[43] W.-X. Long *et al.*, “Channel Estimation for 6G Near-Field Wireless Communications: A Comprehensive Survey,” Jul. 31, 2025, *arXiv*: arXiv:2507.23526. doi: 10.48550/arXiv.2507.23526.
[44] Y. Pan, C. Pan, S. Jin, and J. Wang, “Localization in the Near Field of a RIS-Assisted mmWave/subTHz System,” in *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, Rio de Janeiro, Brazil: IEEE, Dec. 2022, pp. 3905–3910. doi: 10.1109/GLOBECOM48099.2022.10001107.
[45] Z. Abu-Shaban, K. Keykhosravi, M. F. Keskin, G. C. Alexandropoulos, G. Seco-Granados, and H. Wymeersch, “Near-field Localization with a Reconfigurable Intelligent Surface Acting as Lens,” in *ICC 2021 - IEEE International Conference on Communications*, Montreal, QC, Canada: IEEE, Jun. 2021, pp. 1–6. doi: 10.1109/ICC42927.2021.9500663.
[46] A. Haskou and H. Khaleghi, “On the Effect of RIS Phase Quantization on Communications System Performances,” in *2023 International Wireless Communications and Mobile Computing (IWCMC)*, Marrakesh, Morocco: IEEE, Jun. 2023, pp. 1406–1411. doi: 10.1109/IWCMC58020.2023.10182737.
[47] “TMYTEK_UM_Dynamic RIS_GL_v1.0.0.”
[48] A. S. Shekhawat, B. G. Kashyap, R. W. Raldiris Torres, F. Shan, and G. C. Trichopoulos, “A Millimeter-Wave Single-Bit Reconfigurable Intelligent Surface With High-Resolution Beam-Steering and Suppressed Quantization Lobe,” *IEEE Open J. Antennas Propag.*, vol. 6, no. 1, pp. 311–325, Feb. 2025, doi: 10.1109/OJAP.2024.3506453.
[49] X. Zong, B. Zhang, F. Yang, S. Xu, and M. Li, “Suppressing RIS Sidelobe via Spatial Quantization and Phase Quantization Approach,” in *2025 IEEE MTT-S International Wireless Symposium (IWS)*, Xi’an, China: IEEE, May 2025, pp. 1–3. doi: 10.1109/IWS65943.2025.11177719.
[50] K. C. S. Kavya, S. K. Kotamraju, B. N. Kumar, S. Singh, and A. Sidda, “BEAM POINTING ACCURACY OF PHASED ARRAYS FOR SATELLITE COMMUNICATION,” . *Vol.*, no. 10, 2005.
[51] M. S. Abouamer and P. Mitran, “Practical Configuration of a Discrete RIS With Phase-Dependent Amplitude Response,” *IEEE Commun. Lett.*, vol. 28, no. 11, pp. 2678–2682, Nov. 2024, doi: 10.1109/LCOMM.2024.3468872.
[52] A. Noureldin, “Chapter 4 Integrated Navigation System,” RMC, Feb. 10, 2025.
[53] T. Karamat, M. Atia, and A. Noureldin, “An Enhanced Error Model for EKF-Based Tightly-Coupled Integration of GPS and Land Vehicle’s Motion Sensors,” *Sensors*, vol. 15, no. 9, pp. 24269–24296, Sep. 2015, doi: 10.3390/s150924269.
[54] Lc. J. Horobetz, “Single-Point Global Positioning System Positioning”.
[55] Aboelmagd Noureldin, *EEE523 Chapter 4 Video 5*, (Nov. 19, 2021). Accessed: Feb. 26, 2025. [Online Video]. Available: https://www.youtube.com/watch?v=a1zZdzDth6Q
[56] U. Iqbal, A. F. Okou, and A. Noureldin, “An integrated reduced inertial sensor system — RISS / GPS for land vehicle,” in *2008 IEEE/ION Position, Location and Navigation Symposium*, Monterey, CA, USA: IEEE, 2008, pp. 1014–1021. doi: 10.1109/PLANS.2008.4570075.
[57] T. S. Rappaport, R. W. H. Jr, R. C. Daniels, and J. N. Murdock, “Millimeter Wave Wireless Communications”.