UCLA
UCLA Electronic Theses and Dissertations

Title
Design and Electromagnetic Characterization of Circularly-Polarized Shorted-Annular-Ring GPS Antenna for Autonomous Rail Vehicles Positioning

Permalink
https://escholarship.org/uc/item/96r7p1sc

Author
Huang, Tianjian

Publication Date
2022

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA

Los Angeles

Design and Electromagnetic Characterization of Circularly-Polarized Shorted-Annular-Ring GPS Antenna for Autonomous Rail Vehicles Positioning

A thesis submitted in partial satisfaction
of the requirements for the degree
Master of Science in Electrical & Computer Engineering

by

Tianjian Huang

2022

© Copyright by
Tianjian Huang
2022

ABSTRACT OF THE THESIS

Design and Electromagnetic Characterization
of Circularly-Polarized Shorted-Annular-Ring GPS Antenna
for Autonomous Rail Vehicles Positioning

by

Tianjian Huang
Master of Science in Electrical & Computer Engineering
University of California, Los Angeles, 2022
Professor Yahya Rahmat-Samii, Chair

The Global Positioning System (GPS) opened for general usage in the 1980s and has since led to the rapid growth of GPS receivers in both types and numbers. The proliferation prompted applications in various domains and on platforms with special requirements and constraints. Precise positioning is especially critical in drones and other autonomous platforms for tracking and fleet management. One primary source of error in the precision of GPS positioning is multipath interference; typically, antennas that resolve multipath issues involves large, expensive, and specially-designed ground planes, such as choke rings. Patch antennas are still among the most commonly used antenna designs for GPS receivers; their compact size, flexibility in design, low-cost, and low-complexity manufacturing give them unique advantages that are hard to rival by other antenna types. This thesis investigates the potential of a low-cost, shorted-annular-ring (SAR) reduced-surface-wave (RSW) patch antenna design, with multipath mitigation characteristics, for high-precision positioning of freight-carrying autonomous rail transport vehicles. General considerations for GPS antenna designs and SAR antenna design principles are presented. Several simulation studies are conducted for single- and dual-band (L1 & L2) antenna designs. A L1-band SAR-RSW antenna is prototyped and measured, and compared to simulation. The radiation characteristics of the proposed design is also assessed through simulations near a side plate that models the freight container.

The thesis of Tianjian Huang is approved.

Aydin Babakhani

Yuanxun Ethan Wang

Yahya Rahmat-Samii, Committee Chair

University of California, Los Angeles

2022

To my parents, who always offer the best they have throughout my journey

And to the LORD God, for in Him I do not labor in vain,

and with Him all things are possible

TABLE OF CONTENTS

1 Introduction **1**

1.1 Background 1

1.2 Overview of Research 2

1.2.1 Design Goals and Requirements 2

1.2.2 Organization of Work 4

2 Overview of GPS System **7**

2.1 Introduction to GPS and GNSS 7

2.2 GPS System Architecture 8

2.2.1 Space Segment 8

2.2.2 Control Segments 9

2.2.3 User Segments 10

2.3 Ranging Using GPS Signals 10

2.3.1 Pseudorange Measurement 11

2.3.2 Carrier-Phase Measurement 12

2.3.3 GPS Positioning Modes 13

2.4 GPS Signals 16

2.4.1 GPS Signal Characteristics 16

2.4.2 Effects on Signal Propagation 19

2.4.3 Signal Power 20

2.4.4 GPS Modernization 20

3 GPS Antennas: General Considerations **24**

3.1 Polarization 24

3.2 Axial Ratio . . . 25
3.3 Radiation Pattern . . . 26
3.3.1 Dilution of Precision . . . 30
3.4 Bandwidth . . . 32
3.5 Phase Center Stability . . . 32
3.6 Multipath Rejection . . . 34
3.7 GPS Antenna Designs . . . 36
3.7.1 FRPA: Patch Antenna . . . 36
3.7.2 Other FRPA Designs . . . 39
3.7.3 Mitigation of Multipath and Platform Effects . . . 40

4 Shorted-Annular-Ring Patch Antenna . . . 43
4.1 Surface Waves in Patch Antennas . . . 43
4.2 Circular Patch Antennas . . . 46
4.2.1 Cavity Model . . . 46
4.2.2 Resonant Frequencies . . . 49
4.3 Shorted-Annular-Ring Patch . . . 50
4.3.1 Conditions for Surface Wave Suppression . . . 50
4.3.2 Single-Band SAR Patch Design . . . 53
4.4 Dual-Band SAR Antennas . . . 55

5 Simulation and Performance Verification . . . 58
5.1 Single L1-Band SAR Patch: Design Investigation and Performance Verification 58
5.1.1 Substrate Considerations . . . 58
5.1.2 Simulated Performance . . . 61
5.1.3 Fabrication and Measurements: Feed Networks . . . 66

5.1.4 Fabrication and Measurements: Antenna S_{11} 69
5.1.5 Sensitivity Analysis and Tolerance Studies 71
5.1.6 Fabrication and Measurements: Antenna Radiation Characteristics . 76
5.1.7 Four-Feed L1-Band SAR Patch Antenna 78
5.2 Dual-Band SAR Antenna Characterizations 80
5.2.1 Stacked SAR Patch 81
5.2.2 Side-by-Side SAR Patch 85
5.3 Radiation Performance With Side Plate 89
5.3.1 Radiation Performance With Antenna Tilting 92

6 Conclusions and Future Works 96

A Derivation of the Fundamental Design Equation for a SAR Patch With No Surface Wave Excitation 98

B Platform Effects of SAR Patch Antenna: A Comparative Study 101

C Bandwidth Enhancement Using Air Gap 105

D Evaluation of UCLA Spherical Near-Field Antenna Measurement Range at L-Band GPS Frequencies 112

References 116

LIST OF FIGURES

1.1 (a) A simplified depiction of the autonomous railway vehicle system for freight transportation [5]. (b) Approximated antenna mount positions on an autonomous vehicle, showing the antenna size constraints. 3

1.2 A flow chart depicting the organization of the thesis. 6

2.1 The GPS constellation, showing the six, equally-spaced orbital planes around the equator. Note that the orbital radius is not-to-scale. The figure shows 27 GPS satellites; this is known as the expandable 24-slot constellation, which contains three extra satellites in addition to the baseline 24-slot constellation [10]. 9

2.2 A simplified depiction of pseudorange measurement [4]. The top waveform represents the navigation signal broadcast by the satellite and received by the receiver, and the bottom waveform represents the identical code generated by the receiver. Assuming perfect clock synchronization, the transit time of the broadcast signal is Δt, which can be found by performing autocorrelation of the broadcast signal and the receiver-generated signal; the transit time can then be used to determine the position of the receiver. 12

2.3 A simple, one-dimensional example of relative positioning utilizing carrier phase measurements [6]. (a) The initial measurement geometry. The goal is to determine the distance d between the two receivers A and B. There are two unknowns in this setup, which are the distance d and integer ambiguity N. An equation relating d and N is shown in the top left. (b) Ambiguity resolution is enabled after a change in the satellite position. In this case, we can obtain another equation shown in the top left of figure (b). Together with the equation in figure (a), we obtained two equations with two unknowns, thus enabling the solving for d and N. 15

2.4 Block diagram on the GPS L1 & L2 legacy signal structures and generations [2]. It is observed that the GPS signal components (e.g. the carrier, ranging code, and navigation data) are generated from a common reference frequency source at $f_0 \approx 10.22999999543$ MHz, shown on the bottom left, and then are combined and modulated to produce the broadcast signal. 18

2.5 Illustration of the frequency allocations for several GNSS, regional, and augmentation systems [16]. 23

3.1 (a) Idealistic GPS antenna radiation patterns with a relatively uniform RHCP directivity or gain value at the upper hemisphere and a sharp cut off at the elevation mask angle [1]. The elevation mask angle θ_m is equivalent to $90° - \theta_c$, where θ_c is the polar angle measured from the zenith shown in the figure. The figure on the right shows a higher elevation mask angle convention after 2010. (b) Example of a more realistic, though still some-what ideal, GPS antenna radiation pattern with a roughly 5° elevation mask [7]. 28

3.2 (a) Positioning accuracy as relates to the relative geometry of the user and signal emitters [2]. DOP is a common consideration used in time-of-arrival ranging, which calculates the transit time of signals from position reference sources such as foghorns, GPS satellites, or other signal emitters. As the figure on the right suggests, signal sources that are clustered together will result in higher uncertainties and errors in range measurements, and hence, poor DOP values. (b) Figures illustrating GPS satellite geometries that would result in good or poor DOP values. 31

3.3 (a) Ideal versus realistic equiphase front radiated by an antenna [19]. The actual phase front will not be perfectly spherical and is depended on the angular positions, which will introduce different phase offset errors for the signals received by the antenna. (b) Another illustration of ideal versus actual equiphase surface [7]. The apparent phase center is different at each angular observation positions; the average phase center is the mean location of the apparent phase centers for a defined angular range. (c) The antenna reference point (ARP) is the geometric center of the antenna. The difference between the ARP and the average phase center is the phase center offset (PCO). The difference between the phase center at a particular angular position and the average phase center is the phase center variation (PCV) at that position [20]. 33

3.4 Direct and multipath signals received by a GPS antenna. 35

3.5 (a) GPDF254.A, a L1-band dual-feed GNSS patch antenna by Taoglas [23]. (b) HX-CS7017A, a L1/L2 dual-band, four-feed, stacked patch antenna by Harxon [24]. (c) HX-CSX078A, another four-feed, stacked patch GNSS antenna design by Harxon, featuring circular patches [25]. 37

3.6 (a) HC771E, a L1-band quadrifilar helix antenna by Tallysman [26]. (b) A wide-band conical spiral antenna designed for GPS and Galileo reception [27]. (c) A GPS Block IIR-M satellite, showing its antenna farm [28]. Quadrifilar helix and spiral antenna elements are also commonly used in GNSS satellites for UHF communications and transmission of L-band navigation signals [3]. 39

3.7 (a) HX-CGX606A, a multi-band GNSS antenna by Harxon, featuring a 2D choke ring around the antenna element at the center [29]. (b) A typical EBG structure [30]. (c) A EBG structure installed around a GPS antenna [31]. 41

4.1 Surface waves propagate through the substrate, causing deleterious effects such as radiations from edge diffraction and increased mutual coupling between antenna elements [7]. 44

4.2 Basic geometry of a circular patch. 46

4.3 Top and side view of a representative shorted-annular-ring patch. Note the concentric ring boundary that shorts the top patch to the ground plane. 51

4.4 Magnetic ring current sheet $\mathbf{M}_s$ at the perimeter ($\rho = a$) of the patch. 52

4.5 The inverted shorted-annular-ring (ISAR) patch antenna configuration. Note that the shorting boundary can be constructed with a ring of vias. The two feed probes are excited with 90° relative phase to generate circular polarization. . . . 56

4.6 A dual-band configuration (top view) can be constructed with a SAR patch and an ISAR patch. 57

5.1 The top view (a) and bottom view (b) of the simulated L1-band SAR patch antenna model in HFSS. (c) Top and side views of the SAR patch with labeled design parameters, whose values are shown in Table 5.3. 61

5.2 Simulated S_{11} of the L1 SAR Patch Antenna. The results for both ports are near-identical. 63

5.3 (a) $\phi = 0°$-cut gain pattern in dB. (b) $\phi = 90°$-cut gain pattern in dB. 63

5.4 (a) Axial ratio in dB versus polar angle θ for ϕ-plane cuts. (b) Axial ratio in dB in the zenith direction ($\theta = 0°$) versus frequency. The AR = 3 dB level is shown as black dotted line. 64

5.5 (a) Determining the phase center of the SAR patch antenna. The global coordinate system is shown in black, and the varying local coordinate system is shown in red, offset by Δz from the global coordinate origin. (b) Simulated RHCP radiated far-field phase versus the polar angle θ for the L1-band SAR patch when the origin of the coordinate system is at the phase center of the antenna. The black dotted lines denote $\theta = 50°$. 65

5.6 (a) Test board and QCN-19+ on-chip quadrature hybrid. (b) Assembled test board with soldered QCN-19+ in the center and SMA connectors 66

5.7 VNA measurements for (a) S_{11}, (b) phase unbalance, and (c) isolation of the QCN-19+ on the test board. 68

5.8 (a) Top view and (b) bottom view of a prototyped L1-band SAR patch antenna manufacturerd by PCBWay. 69

5.9 Comparison between measured and simulated S_{11} of the antenna feed ports. The three measured results are shown to assess the manufacturing consistency of the prototypes. 70

5.10 (a) Tolerance studies with varying substrate permittivity ϵ_r. Original design is based on RO4003C substrate with $\epsilon_r = 3.55$. The tolerance value is based on the process tolerance of ± 0.05 provided by Rogers. (b) Tolerance studies with varying substrate thickness h. We choose the tolerance value of $\pm 10\%$ based on the vendor PCBWay. (c) Tolerance studies with varying patch radius a. The value is chosen to be ± 0.1 mm based on the trace tolerance of PCBWay. (d) Tolerance studies with varying shorting boundary radius b, which is the radius of the ring of vias. The value is also based on the trace tolerance of PCBWay. (e) Simulated S_{11} with the nominal "worst-case" tolerance of the studied parameters. For example, "worst-case" lower shows the nominal worst-case result using tolerances of all parameters that shift the resonant frequency to a lower value. 74

5.11 Comparison between measured and simulated S_{11} of the antenna feed ports. . . 75

5.12 Measurement setup of the L1 SAR patch antenna in the spherical near-field antenna measurement chamber at the Center for High Frequency Electronics (CHFE) of UCLA. 76

5.13 Comparison between measured and simulated normalized gain pattern for the SAR patch. 77

5.14 Comparison between measured and simulated AR for the SAR patch. The AR = 3-dB level is plotted in black dotted line. 78

5.15 A four-feed L1-band SAR antenna design. Note that the four feeding probes are spatially separated by 90° concentrically, and their excitations are also phase shifted by 90° successively to achieve CP. 79

5.16 (a) Simulated S_{11}; all four ports yield identical results. (b) Simulated gain pattern at $\phi = 0°$ and 90°. (c) Simulated AR versus polar angle θ at $\phi = 0°$ and 90°. . . 80

5.17 (a) 3D view of the stacked SAR patch antenna design model. (b) Top and side views of the stacked SAR patch with labeled design parameters, whose values are shown in Table 5.6. The red objects attached at the bottom are coaxial feed probes. 82

5.18 (a) Simulated S_{11} for the stacked SAR patch antenna; for each patch, the S_{11} results of the four feed ports are identical. (b) Simulated gain pattern at $\phi = 0°$ and 90° for the L1 patch. (c) Simulated AR versus polar angle θ at $\phi = 0°$ and 90° for the L1 patch. (b) Simulated gain pattern at $\phi = 0°$ and 90° for the L2 patch. (c) Simulated AR versus polar angle θ at $\phi = 0°$ and 90° for the L2 patch. 84

5.19 The dual-band side-by-side configuration, with the smaller L1-band SAR patch on the left, and the larger L2-band SAR patch on the right. 85

5.20 Side-by-side configuration simulation results at L1. Simulated gain pattern for (a) $\phi = 0°$ and (b) $\phi = 90°$. (c) simulated AR versus polar angle θ. (d) Simulated RHCP radiated far-field phase versus the polar angle θ when the origin of the coordinate system is at the phase center of the antenna. The phase center is found at around $z = 15$ mm, where $z = 0$ is defined at the ground plane layer, and the z-axis is through the L1 patch geometric center. The black dotted lines denote $\theta = 50°$. 86

5.21 Side-by-side configuration simulation results at L2. Simulated gain pattern for (a) $\phi = 0°$ and (b) $\phi = 90°$. (c) simulated AR versus polar angle θ. (d) Simulated RHCP radiated far-field phase versus the polar angle θ when the origin of the coordinate system is at the phase center of the antenna. The phase center is found at around $z = 15$ mm, where $z = 0$ is defined at the ground plane layer, and the z-axis is through the L2 patch geometric center. The black dotted lines denote $\theta = 50°$. 87

5.22 The freight container is modeled as a large PEC side plate. 89

5.23 Simulated (a) $\phi = 0°$-plane and (b) $\phi = 90°$-plane gain patterns of the setup in Fig. 5.22. 90

5.24 The ideal radiation patterns of the GPS antenna on the autonomous vehicle (a) with and (b) without a freight container installed 91

5.25 The SAR patch antenna is tilted by 45° away from the side plate. This is done by rotating the antenna by 45° around the y-axis based on the geometry of Fig. 5.22. 92

5.26 Simulated results for the setup in Fig. 5.25. Simulated gain pattern for (a) $\phi = 0°$ and (b) $\phi = 90°$. (c) Simulated AR versus the polar angle θ. (d) Simulated RHCP far-field phase versus the polar angle θ when origin of the coordinate syst em is placed near the phase center of the antenna. 94

A.1 Geometry used in the far-field surface wave phase approximation. 99

B.1 (a) A L2-band SAR patch antenna. The dimensions of the antenna is tuned so that it resonates at L2. Other characteristics, such as the substrate used, are the same as the L1-band SAR patch discussed in Section 5.1.2. (b) A conventional L2-band, dual-feed, circularly-polarized square patch antenna used for comparative purposes. The substrate material is F4B with the same dimensions of 150 mm × 150 mm × 1.524 mm as the substrate for the L2 SAR patch. 101

B.2 The platform modeled as a metallic circular plate with a diameter of 50 cm and is positioned 6 mm below the antenna. 102

B.3 Simulated (a) $\phi = 0°$ gain pattern and (b) AR of the L2-band SAR patch antenna shown in Fig. B.1a. Simulated (c) $\phi = 0°$ gain pattern and (d) AR of the L2-band square patch antenna shown in Fig. B.1b. 104

C.1 (a) A 3D view of the SAR patch antenna design with an air gap layer. The substrates maintain the dimension of 150 mm × 150 mm as the designs shown in Chapter 5. (b) The front view of the design, showing a 5-mm-thick air gap layer between the patch and ground plane substrates. 106

C.2 Simulated (a) S_{11}, (b) AR versus polar angle θ, (c) gain pattern at $\phi = 0°$, and (d) gain pattern at $\phi = 90°$ of the air gap design shown in Fig. C.1. 107

C.3 (a) A 3D view of the SAR patch antenna "all-metal" design with an air gap layer. (b) The top and front views of the design. Note the thickness of the metal is chosen to be 1 mm. The air gap is maintained at a 5-mm thickness. 109

C.4 Simulated (a) S_{11}, (b) AR versus polar angle θ, (c) gain pattern at $\phi = 0°$, and (d) gain pattern at $\phi = 90°$ of the "all-metal" design shown in Fig. C.3. 110

D.1 The WR770 OEWG probe used in the L1- & L2-band measurements at UCLA's CHFE. 112

D.2 The (a) top view, (b) bottom view, and (c) front view of the truncated-corner patch antenna for probe measurement assessment. 113

D.3 (a) Fabricated test antenna. (b) Fabricated antenna mounted in the measurement chamber. (c) The simulation model includes the mount plate that introduces platform effects and affects the measured results. 114

D.4 Measured and simulated radiation patterns at (a) L2 and (b) L1. 115

LIST OF TABLES

2.1 Minimum user-received GPS signal power specifications for Block IIR-M/II-F GPS satellites [6, 12]. 20

2.2 Frequency characteristics of GPS and Galileo constellations [2, 12, 14–16]. 22

3.1 Minimum gain requirement versus elevation mask angle for the L1–E5 bands [7]. 30

5.1 Characteristics of the substrates used in simulation and antenna performance metrics achieved using the substrates. All performance results are simulated at the L1 frequency (1.575 GHz). 59

5.2 Inquired prices, minus the shipping cost, for one PCB unit of size 150 mm × 150 mm from two different PCB vendors. *Layer refers to the number of conductive (e.g. copper) layers, not substrate layers, of one PCB unit. N/A indicates that the substrate is either unavailable for the vendor at the time of inquiry, or that the number of conductive layers is not supported for that particular substrate by the vendor. 59

5.3 Nominal design values used in the HFSS simulation for the antenna dimensions shown in Fig. 5.1c. 62

5.4 Some relevant parameters of the QCN-19+ quadrature-phase hybrid coupler. . . 67

5.5 Relevant information regarding the QCN-19+ test board. 67

5.6 Nominal design values used in the HFSS simulation for the antenna dimensions shown in Fig. 5.17. 84

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere gratitude to my advisor Professor Yahya Rahmat-Samii. His invaluable instructions, immense knowledge and insights, and passion for the field have guided me through out my Master's study. It is an honor to be one of his students.

The research presented in this thesis is motivated by the idea of GPS positioning for autonomous freight transportation, based on the concept developed by Parallel Systems and interactions with its experts, particularly Mr. John Howard.

In addition to my advisor, I want to thank the rest of my committee members, Professor Aydin Babakhani and Professor Yuanxun Ethan Wang, for their time and effort in reviewing this thesis.

Last but not least, this work would not have been possible without my colleagues at the UCLA Antenna Research, Analysis, and Measurement Laboratory, and in particular, Anastasios Papathanasopoulos, who has been a tremendous mentor. His unconditional supports, fruitful discussions, and patients with me were the foundations that brought this work into fruition, to say the least. It is a great pleasure to know and work with him.

CHAPTER 1

Introduction

1.1 Background

The Global Positioning System (GPS) was entered into service in the 1970s and since the 1980s, GPS applications in the civilian domains became more accessible and widespread, surpassing its originally intended purpose limited to the military arena. As the service extended globally, the performance and cost issues for GPS antennas have fundamentally changed. Today, the price of a GPS antenna can vary greatly, ranging from $4 to $8000 for distinct platforms and applications, with some even exceeding the price of a receiver [1].

Applications utilizing GPS are also being developed at an astonishingly fast rate, and their scale have increased dramatically over the past few decades. Currently, more than 53% of the satellite navigation market is dedicated to location-based services (LBS); these include embedded GPS receivers in cell phones, tablets, and other wearable devices that are now considered essential in daily lives. GPS usage in land transportation, such as vehicle navigation and tracking, has also become extremely proliferated, growing to a size perhaps second only to cell phones [2,3]. GPS receivers have become so ubiquitous that the positions of moving parties can be easily obtained. Besides the commonly-perceived navigation and routing applications, automobiles now incorporate standard GPS-enabled features such as roadside assistance, automatic crash notification, and automatics vehicle location. There is now growing GPS demands for systems such as commercial real-time fleet-monitoring; the data GPS provides are highly valuable for delivery, emergency response vehicles, and scheduled service fleet dispatch and control. Accurate GPS locations will also be essential for drones and other autonomous platforms for accurate positioning and fleet managements [4].

In other words, GPS has become an indispensable component in telematics technologies, which offers information-on-the-move.

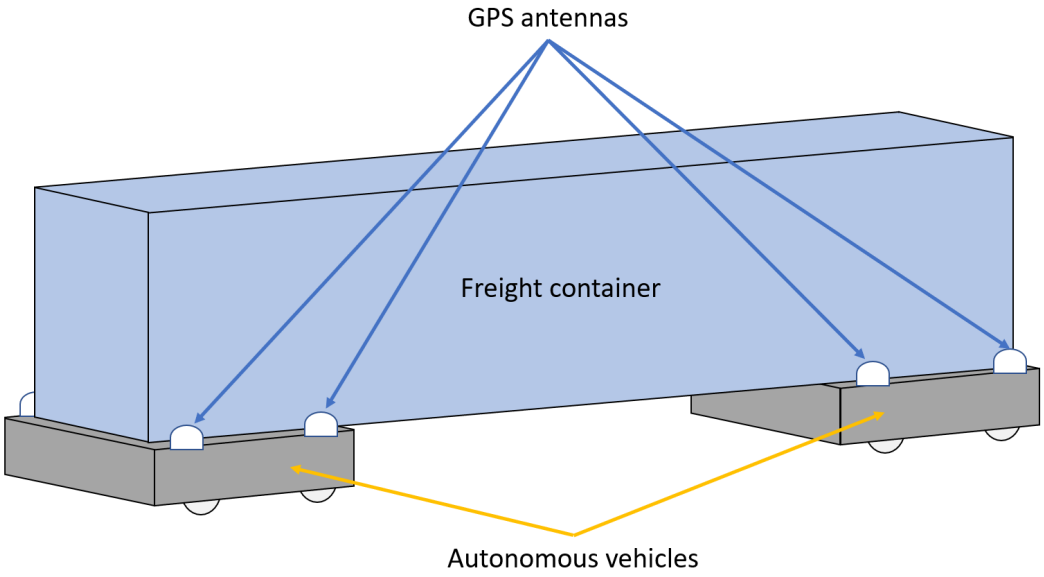
1.2 Overview of Research

1.2.1 Design Goals and Requirements

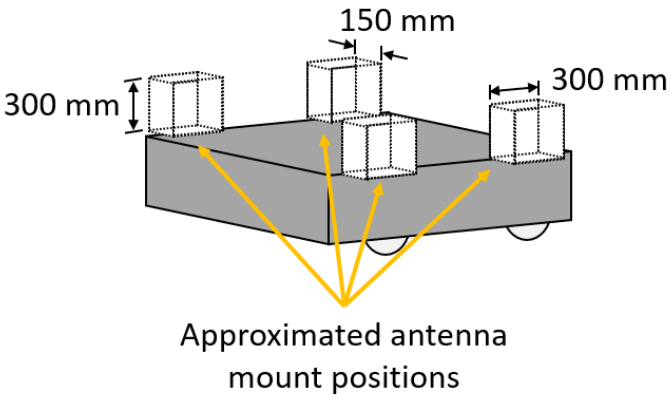
The objective of the thesis is to design, simulate, and investigate the performance of a potential GPS antenna design for autonomous rail vehicles (drones) for freight transportation, envisioned by Parallel Systems [5]. A generic depiction of such system is shown in Fig. 1.1. The freight container is to be carried by two identical autonomous vehicles; each vehicle is mounted with four GPS antennas. The representative design requirements and constraints of the GPS antenna are as follows:

- Low cost: $< \$25$/unit
- Polarization: right-hand circular polarization (RHCP)
- Cross-polarization discrimination (XPD): > 10 dB
- Front-to back ratio (D/U ratio): > 5 dB
- Operating band: L1 (1575.42 MHz) and L2 (1227.6 MHz).
- Antenna size constraint: 150 mm $\times$ 300 mm $\times$ 300 mm

These requirements are expected to be additions to the typical requirements for an antenna, such as good impedance match ($S_{11} < -10$ dB) over the operating bandwidth, and to the common requirement of a GPS antenna, such as multipath rejection and phase center stability, which will be discussed in details in later chapters. Many antenna design challenges are rooted in the platform constraints. For GPS antennas, the platform along with the platform's immediate environment dictates how well an antenna is capable of alleviating issues such as multipath and other interference from natural or man-made sources [1]. Each



(a)



(b)

Figure 1.1: (a) A simplified depiction of the autonomous railway vehicle system for freight transportation [5]. (b) Approximated antenna mount positions on an autonomous vehicle, showing the antenna size constraints.

autonomous vehicle can operate in three modes, namely freely (as an individual vehicle), in a transport unit (two vehicles carrying one freight container), or in a platoon (many transport units connected together). The GPS antenna design is expected to function in all three modes. In this case, the immediate radiation environment of the antenna is the autonomous vehicle with or without the freight container. For a changing radiation environment such as this, one may imagine that an adaptive antenna approach may be beneficial, such as a beamforming array. The idea is that the antenna can exhibit one radiation pattern when there is no freight crate and another when there is. However, due to the constraint on the required antenna size, the number of array elements that can fit within is questionable, especially when the required size is comparable to the wavelength of operation. The small number of array element then brings into question the beamforming capability. The array feed and beamforming networks will further increase the costs. High-permittivity substrate may allow the miniaturization of patch antenna elements, but it is at the expense of other factors such as decreased bandwidth and gain.

The proposed design choice of this thesis is the reduced-surface-wave (RSW) shorted-annular-ring (SAR) patch antenna. The low-cost requirement is most easily satisfied using patch antenna designs on cheap PCB substrates. Patch antennas are also exhibit low profiles, and hence the height constraint of 300 mm will not be an issue. As will be shown in later chapters, the SAR patch antenna possesses multiple desirable characteristics, such as reduce surface wave radiations and the resulting multipath rejection capabilities, that enables it to become a promising candidate for high precision GPS applications.

1.2.2 Organization of Work

In order to understand design requirement and choices for a GPS antenna, the operation principles of the GPS should be first investigated. Chapter 2 provides a foundational overview on the functioning of the GPS system, such as the methods of ranging calculations and GPS signal characteristics, including the operating frequency bands and bandwidth. Chapter 3 then provides the insights on the general design considerations for GPS antennas. This

chapter conceptually introduces desirable features expected from a typical GPS antenna with the reasoning of such. Several figures of merits, such as cross-polarization discrimination (XPD) and multipath rejection ratio, are defined to aid the assessments of the GPS antenna performance. A brief survey of several common and proven GPS antenna designs are then presented and their advantages and disadvantages are discussed. Chapter 4 then provides the reasoning of the selection and theoretical treatment of the shorted-annular-ring (SAR) patch antenna, which the design proposed for the autonomous vehicles. In this chapter, we will derive the mathematical model and fundamental design equation needed to construct such antenna that will exhibit a suppressed surface wave radiations. Based on the design given in Chapter 4, Chapter 5 first provides the results from the simulation studies done on a L1-band SAR patch antenna. The performance of a manufactured prototype is then measured and assessed comparing to the simulation. An error/tolerance studies is conducted to find the possible causes of errors between the measurements and simulations. Based on the L1-band SAR antenna designs, more simulation studies are presented including the four-feed designs to increase the impedance bandwidth, dual-band configurations, and lastly, radiation performance assessment when the antenna is placed near the freight container model. A flow chart representing the progression of the studies in this thesis is shown in Fig. 1.2.

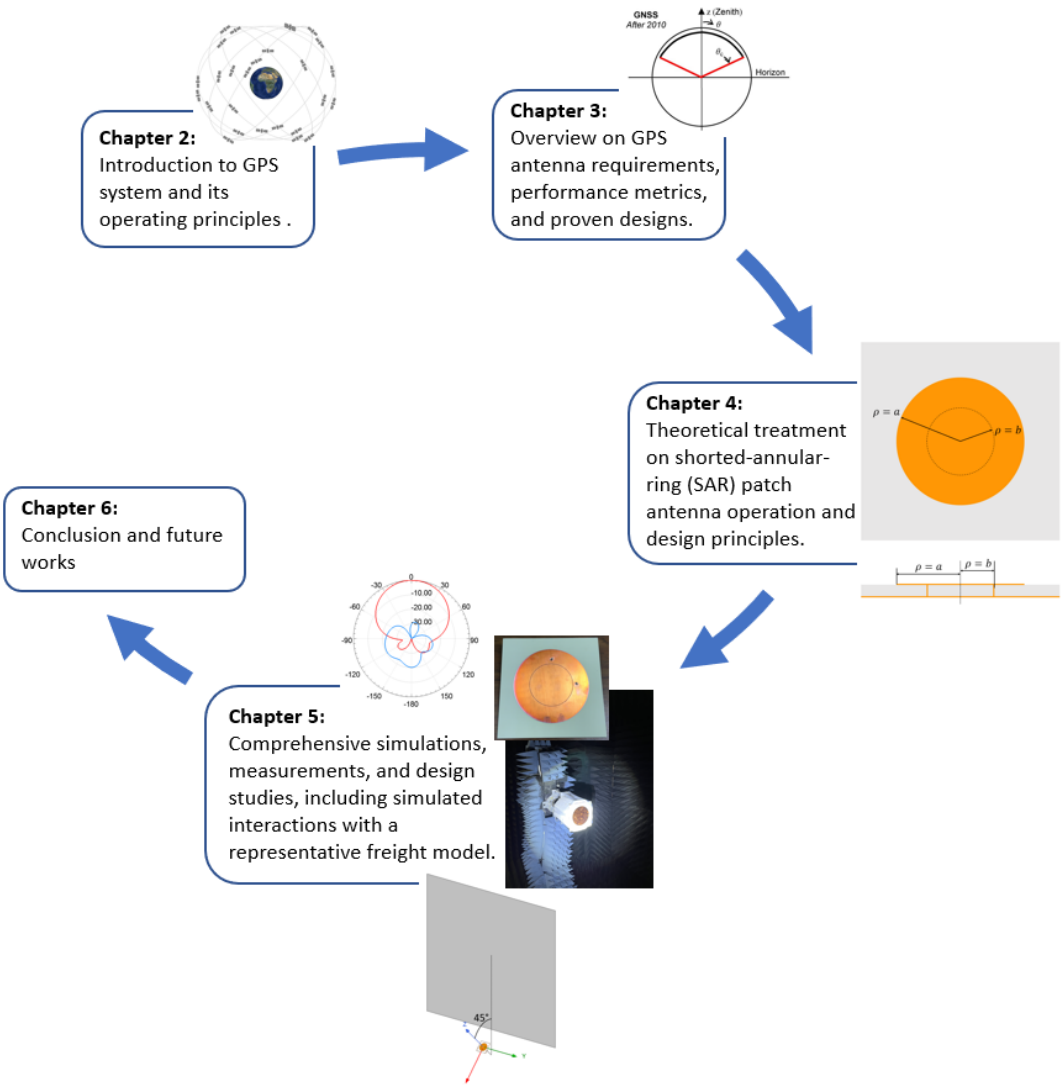


Figure 1.2: A flow chart depicting the organization of the thesis.

CHAPTER 2

Overview of GPS System

2.1 Introduction to GPS and GNSS

The Global Positioning System (GPS) is a satellite-based, passive radio navigation system operate by the U.S. Department of Defense (DoD). Originally developed for the U.S. military in the 1970s and was known as the Navigation Satellite Timing and Ranging Global Positioning System (NAVSTAR GPS), the service was declared fully operational by 1995 and has gradually become accessible to general use around the world since the 1980s [6]. Particularly, component miniaturization and large-scale manufacturing enabled GPS receivers to become more proliferated and inexpensive. In 2004, there was an estimate of eight million GPS receivers just in automobiles alone in the U.S [2]. The GPS system is continuously improving in terms of accuracy, availability, and integrity as new blocks of satellites and capabilities are added to the system.

Any navigation satellite constellation with global coverage can be called a global navigation satellite system (GNSS), though the term may also refer to the total collection of the all satellite navigation systems and their augmentation systems. As of January 2022, GPS along with Russia's Global Navigation Satellite System (GLONASS), China's BeiDou Navigation Satellite System (BDS), and European Union's Galileo are operational constellations that can be classified as GNSS [2, 7]. A GNSS provides accurate, continuous, and three-dimensional (3D) position, velocity, and time (PVT) information to users with appropriate receiving equipment [2]. Today, GNSS such as the GPS have become an indispensable, though less appreciated, part of daily lives. Other than providing high-precision positioning and navigational needs in land, maritime, airborne, and space domains, GPS is extensively

used in land surveying, mapping, and geographic information systems (GIS) [4]. Importantly, GPS is also used in commercial and industrial infrastructure as a source of precise time. GPS-derived time are widely used as a reference for scientific purposes and correlating events recorded at different geographic locations, with applications including synchronizing components of telecommunication networks, setting clocks at major internet nodes, and marking bank transactions, to name a few [6,8]. Potential applications of GNSS are continuing developing at a rapid rate. It is projected that there will be over 9 billion GNSS-employed devices by 2023, which is more than one unit for every person, with application in diverse areas in industrial, commercial, scientific, and personal domains [2].

2.2 GPS System Architecture

The GPS architecture consists of three main segments: the Space Segment, the Control Segment, and the User Segment. The Space Segment consists of the constellation of GPS satellites, and the Control Segment tracks the satellites and oversees satellite operations; both segments are managed by the DoD. The User Segment involves the development of GPS receivers and encompasses GPS users in both military and civilian sectors [6].

2.2.1 Space Segment

The baseline GPS constellation is composed of 24 satellites in near-circular medium Earth orbit (MEO) of approximately 20,200 km in altitude and with an orbital period of roughly twelve hours. This is known as the baseline 24-slot constellation [2, 6]. After GPS achieved full operational capability (FOC) in 1995, the constellation has always maintained more than 24 operational satellites for redundancy and improved coverage; as of Janurary 2022, the GPS constellation consists of 29 operational satellites [6, 9].

The baseline 24 satellites are divided into six orbital planes, inclined at 55° relative to the equatorial plane, with four main satellite slots distributed unevenly in each plane. There are spare satellite slots in each plane to support more satellites. The six orbital planes are equally spaced around the equatorial axis at 60° separation, shown in Fig. 2.1. The baseline

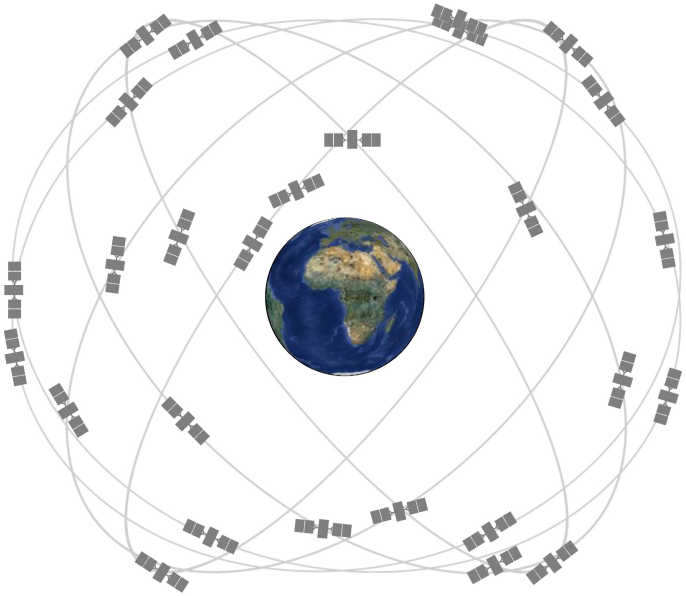


Figure 2.1: The GPS constellation, showing the six, equally-spaced orbital planes around the equator. Note that the orbital radius is not-to-scale. The figure shows 27 GPS satellites; this is known as the expandable 24-slot constellation, which contains three extra satellites in addition to the baseline 24-slot constellation [10].

constellation is designed to ensure a minimum of four satellites in view for users with a clear view of the sky at any point on Earth [6, 7, 9]. Realistically, a user can expect to see six to eight satellites most of the time [6]. The GPS constellations are continually improving system performance in terms of accuracy and jamming resistance as new blocks of GPS satellites are launched and older ones are phased out.

2.2.2 Control Segments

The Control Segment's primary responsibility is to receive satellite status and transmit commands to the satellites to ensure their proper functions. The segment comprises a global network of dedicated L-band monitor stations and S-band ground antennas for tracking and communicating with the satellites. At the core of the Control Segment is the Master Control Station (MCS), located at Colorado Springs in the U.S. The monitor stations are spread across the globe in longitude; they are unmanned and operated remotely from the MCS.

Some of the more detailed functions of the Control Segment include [2, 6, 11]:

- Tracking and predicting precise satellite locations
- Monitoring and maintaining satellite orbits, also called "stationkeeping," including relocation of satellites if necessary
- Monitoring satellite health and signal integrity to ensure functionality and accuracy, and performs anomaly resolution when necessary
- Maintaining GPS time
- Generating/updating navigation messages and uploading to the satellites

2.2.3 User Segments

The fast proliferation of GPS usage in the civilian portion of the User Segment is mainly owing to the development of integrated circuits, which enables light, compact, and inexpensive GPS receivers. It was estimated that more than one million GPS receivers have been produced yearly since 1997 [6]. GPS receivers are now embedded almost ubiquitously in cellphones, automobiles, robotic platforms, etc. The receiver designs are numerous, usually driven by specific applications and requirements, and may differ greatly. To meet the anticipated market demands, receivers that cover two or more GNSS systems have also received considerable attentions and developed at a relatively low costs [1]. GPS antenna remains to be an important component in any GPS receiver unit and will be discussed in details in Chapter 3.

2.3 Ranging Using GPS Signals

A GPS receiver receives the L-band navigation signals from at least four GPS satellites to calculate the user's PVT information using trilateration. The measurement of ranges using trilateration requires three position references, which are objects at known positions. For GPS receivers, these references are three GPS satellites. The satellites are moving with a

speed of roughly 4 km/s, however, their position can be estimated with an error no larger than a few meters based on prediction made a day or two earlier [6]. The distance between the user and a satellite can be determined using two main methods: pseudorange measurement and carrier-phase measurement.

2.3.1 Pseudorange Measurement

In Pseudorange measurement, the distance between the user and the satellites are determined from the transit time of the signal from the satellite to the receiver; this is an example of time-of-arrival (TOA) ranging. The term pseudorange will be explained below. To make use of the transit time for calculating the accurate distances from the satellites, the clocks in the transmitters and the receivers must be synchronized with each other [2, 4, 6]. For example, an error of 1 µs in synchronization will result in a deviation of 300 m in range measurement. The measurement accuracy is also affected by other sources such as measurement noise and propagation delay, but these are generally negligible when compared to the error produced by non-synchronized clocks [2]. The on-board atomic clocks on the GPS satellites are near-perfectly synchronized with each other, and they are constantly monitored by the Control Segment [6]. The main challenge lies in synchronizing the receiver clocks with those of the GPS satellites. Assuming this is the case, as the satellite transmits the navigation ranging code, the receiver generates an exact copy of this GPS code. The receiver will pick up the transmitted code after some time. By comparing the phases of the transmitted code and the generated copy at the time of reception, the receiver can calculate the signal transit time. Hence, the pseudorange measurement is also known as code-phase measurement. A simplified view of the pseudorange measurement is shown in Fig. 2.2.

Realistically, however, one should not expect GPS receivers, numbered by the billions, to be synchronized with the satellite. The requirements on the receiver clock should also be kept at a minimum to reduce cost for wide-spread general use. This seemingly extraordinary requirement can actually be avoided, allowing the use of inexpensive quartz oscillators in the receivers [6]. This is because any clock bias in the receiver clock affects the observed transit

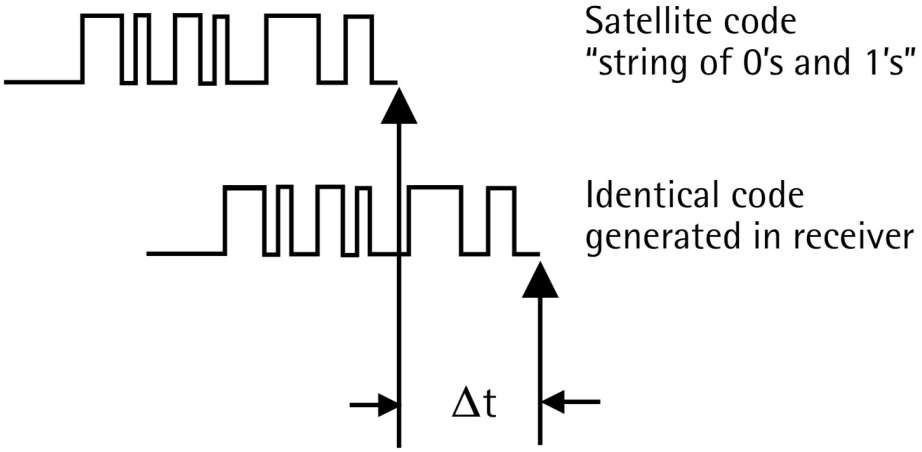


Figure 2.2: A simplified depiction of pseudorange measurement [4]. The top waveform represents the navigation signal broadcast by the satellite and received by the receiver, and the bottom waveform represents the identical code generated by the receiver. Assuming perfect clock synchronization, the transit time of the broadcast signal is Δt, which can be found by performing autocorrelation of the broadcast signal and the receiver-generated signal; the transit time can then be used to determine the position of the receiver.

time for all satellite signals equally. The resulting measured ranges, known as pseudoranges, are therefore all differ from the true ranges by a common amount [2, 4, 6]. The receiver clock bias then becomes the fourth unknown, in addition to the three position coordinates of the receiver. Thus, a fourth GPS satellite is needed to solve the four unknowns for the receiver: three spatial position coordinates and time [4, 6].

2.3.2 Carrier-Phase Measurement

Carrier-phase ranging is an alternative method to pseudorange measurement. It measures the difference between the phases of the receiver-generated carrier wave and the actual carrier wave received from the satellite at the measurement instance. In an ideal case where the clocks are perfectly synchronized and that there is no relative motion between the satellite and the receiver, the carrier phase measurement will be fixed at a fraction of a cycle. The-

oretically, the distance between the satellite and the receiver can be calculated as the sum of the total number of full carrier cycles plus the measured fractional cycle multiplies by the carrier wavelength [4, 6].

However, there is one major issue. The receiver will have no means to differentiate one cycle from another, and thus is unable to determine the total number of cycles between the satellite and the receiver. The unknown total number of whole cycles is known as the integer ambiguity or initial cycle ambiguity. The receiver, however, does have the ability to keep track of the phase changes, such as when the satellites move and the distances between them and the receiver change; this would ensure the initial ambiguity would remain the same as long as there is no signal loss/cut off, known as cycle slip. Since the carrier phase can be measured with great precision, the calculated range using the carrier-phase method is far more accurate than that using pseudorange method because the wavelength of the carrier is much smaller, and hence higher resolution, than that of the ranging codes. The error due to multipath and receiver noise in carrier-phase measurements is at millimeter- to centimeter-level, which is roughly one-hundredth of that for pseudorange method [2, 4, 6].

Nonetheless, one still needs to solve the integer ambiguity. There are several ways to solve this in order to utilize the great precision in measured carrier phase difference and determine the range. One method is to combine the pseudorange and carrier-phase measurements, in which the former is less precise but unambiguous, and the latter is precise but contains integer ambiguity. However, one typical method is to employ relative positioning technique, which will be discussed in the following section.

2.3.3 GPS Positioning Modes

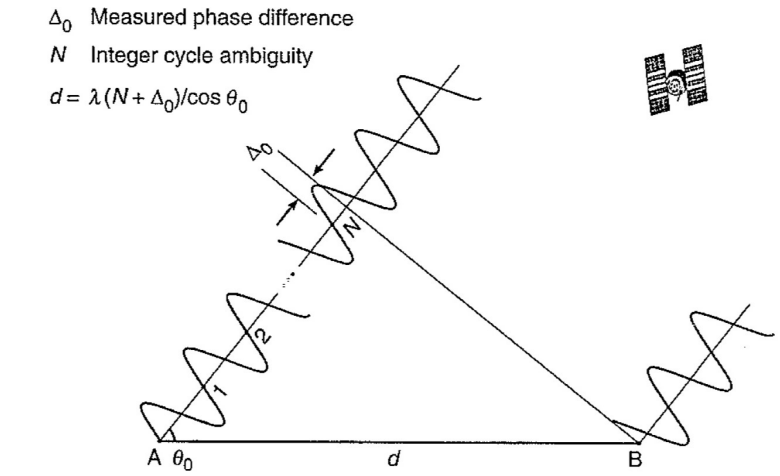
Positioning using GPS can be done in two main ways: point positioning and relative positioning. In point positioning, also called standalone or autonomous positioning, the user employs only one GPS receiver that uses pseudorange measurement to determine the user's instantaneous position [4]. Almost all commercially available GPS receivers are capable of point positioning. Point positioning is used primarily when a relatively low accuracy is

required [2, 4, 6].

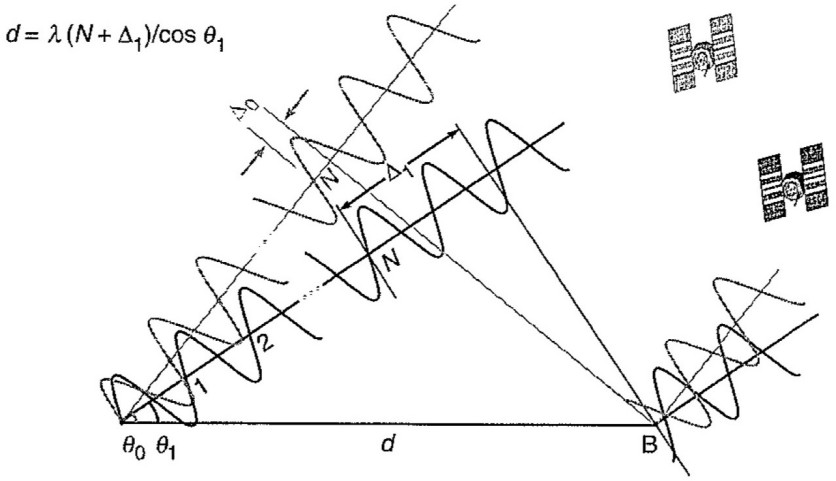
GPS relative positioning, also known as differential positioning, employs two (or more) GPS receivers to track the same satellites. A minimum of four common satellites in view for both receivers is required to perform relative positioning. One receiver is chosen to be a reference, or base, which is stationary with precisely known coordinates. The other receiver, called rover or remote receiver, has unknown coordinates and may or may not be stationary. Either pseudorange or carrier-phase measurement can be employed, and depending on which one is used, an accuracy range of meter- to subcentimeter-level can be achieved. Relative positioning provides a higher accuracy than point positioning does; this is because when two receivers are tracking the same satellites, the measurements contain roughly the same errors and biases, called common-mode errors; the closer the two receivers are with to each other, the more similar the common-mode errors will be. Thus, by taking a difference between two measurements, hence the name differential positioning, the common-mode errors can be removed [2, 4].

Traditionally, the integer ambiguity are usually solved with carrier-phase measurements carried out over a span of time, or at two different times, by the reference and rover receivers. By combining the recorded measurements over different times, one can construct an interferometry that produces a set of equations, which are then solved to determine the integer ambiguity value [4]. A one-dimensional example is illustrated in Fig. 2.3. An associated pseudorange measurement may aid the ambiguity resolution by limiting the size of the integer ambiguity to around $\pm 10\lambda$ [2]. A brute force method involves applying least squares solution to each time iteration and minimizing the residual. Brute force method is usually computationally intensive; when applying more sophisticated computational techniques, a positioning accuracy of around 20 cm in near real-time and millimeter-level accuracy post-processing are achievable [4].

Relative positioning using carrier-phase measurement is widely used in precise positioning applications. Depending on whether the rover receiver is stationary or mobile, one can conduct very precise static survey or kinematic survey [4]. In the 1980s, data collection for a GPS survey would require an hour or more from stationary antennas just for adequate integer



(a)



(b)

Figure 2.3: A simple, one-dimensional example of relative positioning utilizing carrier phase measurements [6]. (a) The initial measurement geometry. The goal is to determine the distance d between the two receivers A and B. There are two unknowns in this setup, which are the distance d and integer ambiguity N. An equation relating d and N is shown in the top left. (b) Ambiguity resolution is enabled after a change in the satellite position. In this case, we can obtain another equation shown in the top left of figure (b). Together with the equation in figure (a), we obtained two equations with two unknowns, thus enabling the solving for d and N.

ambiguity resolution [6]. To improve survey efficiency, the rover receiver needs to compute positions in real time while on the field in motion. This is known as real-time kinematic (RTK) mode, and a key requirement is the capability to estimate the integer ambiguity while the rover receiver is mobile, known as on-the-fly (OTF) ambiguity resolution [4]. Describing specific OTF techniques is beyond the scope of this section. If the ambiguity resolution can be done reliably and near-instantaneously, RTK can also be used as a basis for precise navigation.

2.4 GPS Signals

The GPS satellites broadcast one-way navigation signals, and GPS receivers operate in passive radio mode that only receive and not transmit signals. This arrangement enables GPS to service an indefinite number of users simultaneously [6].

2.4.1 GPS Signal Characteristics

Each GPS signal can be broken down into three components:

- Carrier
- Ranging code
- Navigation data

Currently, the GPS constellation broadcasts navigation signals using sinusoidal carrier waves at several L-band frequencies:

- L1: 1575.42 MHz
- L2: 1227.6 MHz
- L5: 1176.45 MHz

All current GPS satellites transmit two legacy signals, which are the carrier frequencies called Link 1 (L1), the primary frequency, and Link 2 (L2), the secondary frequency. L5 is added as part of the GPS modernization, which will be discussed in Section 2.4.4. Most GNSS constellations today use carrier frequencies in the L-band, with spectra for different constellations densely spread across 1146–1616 MHz, as shown in Fig. 2.5.

The GPS ranging codes are examples of pseudorandom noise (PRN) code, and is unique for each GPS satellite. The code modulation is different for each satellite, so that there are no signal interference even though all satellites transmit at the same L-band frequencies. In fact, GPS was the first broad-scale application of such signaling concept known as code division multiple access (CDMA). The ranging codes enable precise range measurements and alleviate the undesired effects from reflected or interfering signals received by a GPS receiver [6]. The GPS ranging codes are classified in to two types: coarse acquisition code (C/A-code) used in the standard positioning service (SPS) for general use, and precision code (P-code) used in the precise positioning service (PPS) for the U.S. military and DoD authorized users. As the names suggest, the C/A-code range measurement is less precise compared to that of the P-code, but it is less complex [4]. The P-code is encrypted and referred to as the Y-code when the satellites are operating in the antispoof (A-S) mode, with the main goal to protect users from spurious GPS signals transmitted by an adversary. It is common to refer to P-code in either mode of operation (A-S mode on or off) collectively as the P(Y)-code [2]. Originally, both C/A- and P(Y)-codes are transmitted on L1, whereas only P(Y)-code is transmitted on L2. For L1 and L2 signals, the ranging codes are then additionally modulated by a 50-bps navigation message data, which provides the necessary information to the receiver for computing the precise locations of the satellites and signal transit times [2, 4].

The three components of a GPS signal (carrier, ranging code, navigation data) are all generated coherently on board the satellites using a common frequency source that is driven by a rubidium or cesium atomic clock, known as the atomic frequency standard (AFS), with accuracy better than 1 ns [6]. The nominal reference frequency f_0 is 10.23 MHz for an observer on the ground. To compensate for the relativistic effects from the satellite orbiting

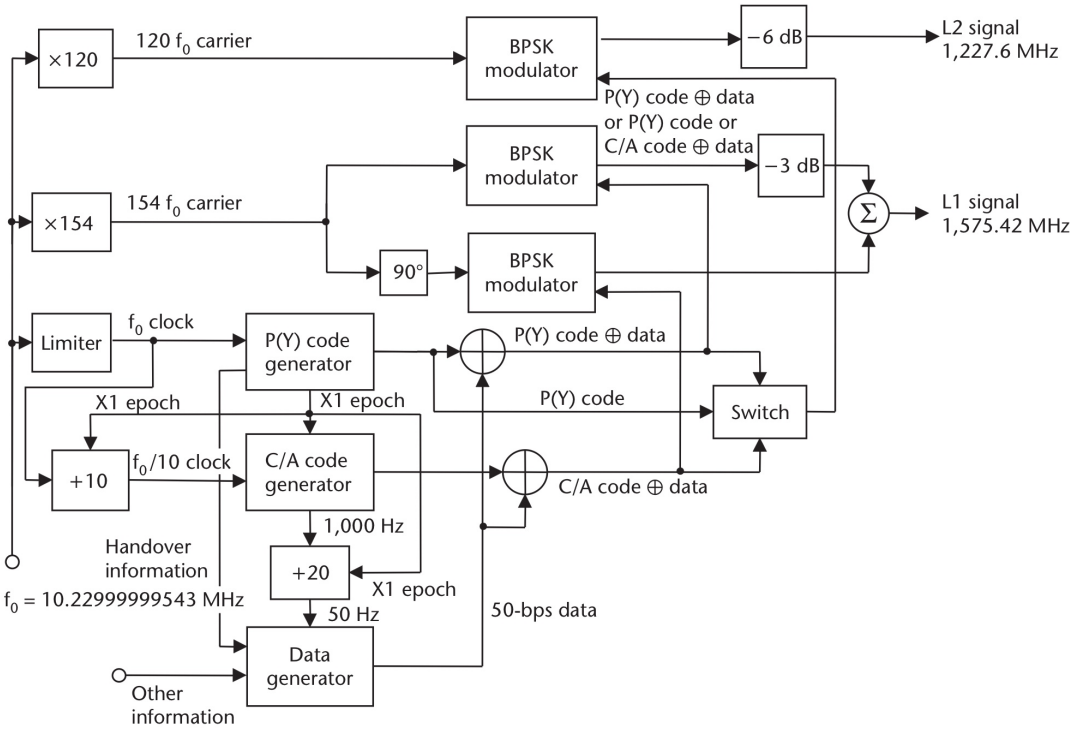


Figure 2.4: Block diagram on the GPS L1 & L2 legacy signal structures and generations [2]. It is observed that the GPS signal components (e.g. the carrier, ranging code, and navigation data) are generated from a common reference frequency source at $f_0 \approx 10.22999999543$ MHz, shown on the bottom left, and then are combined and modulated to produce the broadcast signal.

around the Earth, the actual output reference frequency from the satellite is set to a slightly lower frequency of 10.2299999954326 MHz [12]. The actual carrier frequencies are then generated at a set multiples of the reference frequency f_0, more specifically:

- $f_{\mathrm{L1}} = 154 f_0 = 1575.42$ MHz
- $f_{\mathrm{L2}} = 120 f_0 = 1227.6$ MHz

A block diagram on the generation of legacy L1- and L2-band GPS signals are presented in Fig. 2.4.

2.4.2 Effects on Signal Propagation

GPS signals transmitted from fast-moving satellites propagate through the Earth's atmosphere to reach the receivers. As it propagates, the signal will experience several impairments including but not limited to Doppler shift and atmospheric effects in the ionosphere and troposphere.

The GPS satellites in MEO orbit with a speed of around 4 km/s. This creates a relative motion between the satellite and the receiver, and as a result, a noticeable Doppler effect on the received GPS signal. It is estimated that the resulting maximum uncertainty in the frequency is roughly ± 10 kHz, and thus a receiver should search within this 20-kHz band for the transmitted GPS signal [3].

GPS satellites transmit signals in L-band frequencies. L-band was chosen to be the optimal spectrum range for GPS signals because it was relatively uncluttered in the early 1970s, but more importantly because it offered several advantage compared to other bands. At lower frequencies, the signals will experience longer delays and severe fading due to inhomogeneities in the atmosphere, and at higher frequencies, space loss and atmospheric attenuation will increase significantly [6]. One major effect the ionosphere has on the GPS signals is its dispersive characteristics, which causes frequency-dependent phase shift, or group delay, of the signals. The effect of the dispersion is resolved by using two widely spaced frequencies; in GPS, this is mostly done using the two main carrier frequencies, L1 and L2, which are separated by several hundred megahertz. This allows the receiver to estimate the delays experienced by the signals as they pass through the ionosphere and thus correct almost all ionospheric effects. The troposphere, though non-dispersive at GPS frequencies, can delay the signal waves by refraction due to different concentrations of water vapor. The resulting error from the troposphere effects is smaller than that from the ionosphere, but it cannot be resolved using calculations and can only be estimated using a general model [3].

Table 2.1: Minimum user-received GPS signal power specifications for Block IIR-M/II-F GPS satellites [6, 12].

Frequency band	C/A-code, L2C, or L5 (dBW)	P(Y)-code (dBW)
L1	−158.5	−161.5
L2	−160	−161.5
L5	−154.9	N/A

2.4.3 Signal Power

The RF power at the input port of a GPS satellite's antenna is around 50 W, roughly half of which is assigned to the C/A-code [6]. The satellite antenna is designed such that the RF power is spread roughly evenly over the Earth's surface. After the long distance travel from the more than 20,000-km medium-Earth orbit, the received GPS signals on the surface of Earth are extremely weak. The specifications on the minimum received GPS power level for users on Earth are given in Table 2.1. The table assumes Block IIR-M and II-F GPS satellites; the specifications will be slightly different for older and newer GPS satellite blocks. The specified power levels are based on signal received by an ideal right-hand circular polarized (RHCP) antenna with unity gain (0 dBic). Clearly, the GPS signals are well below the background RF noise level. A GPS receiver is able to extract such weak signal from the background noise owing to the knowledge of the PRN ranging codes; such signal boost is known as processing gain. Nonetheless, the low signal power remains a main concern for the resilience of GPS against RF interference and jamming, especially when more and more public infrastructure depends on GPS [6].

2.4.4 GPS Modernization

GPS modernization was launched in the late 1990s soon after the system was operational and the full potentials for military and civil applications became clear [6]. The modernization program was aimed to provide signal redundancy, improved positioning accuracy, signal

availability, and system integrity [13]. Dual-frequency measurements are crucial for precise positioning, which allowed users to correct for the effects of the ionosphere. As mentioned in sections above, originally, only the P(Y)-code is transmitted on both L1 and L2 for U.S. military use. In 1998, two new civilian signals were announced: L2C and L5. First, the L2C signal was the C/A-code added to L2 to improve civilian GPS accuracy. The L2C addition was found to be inadequate in the civil aviation safety-of-life applications, mainly due to potential interference from ground radars operating near the GPS L2 band. As a result, a third civilian signal L5, was introduced. As mentioned prior, since a main concern for GPS is the low power level, L5 is designed to be transmitted at a higher power. L5 signals will exhibit a wide broadcast bandwidth of 20 MHz minimum and will be more robust under noisy and multipath conditions. L5 signals will also implement new ranging and navigation code structures that are more efficient. A new military signal, known as M-code, was also introduced along with L2C and was designed to be more robust than the P(Y)-code [4]. M-code is transmitted on both L1 and L2 frequencies coexisting with the previous L1 and L2 codes, and is intend to eventually replace the P(Y)-code [2]. A fourth civilian signal, L1C, was also introduced; it is transmitted on the L1 frequency and enables the interoperability between GPS and international satellite navigation systems [13].

The frequency characterstics of the broadcast signal for GPS and Galileo are listed in Table 2.2. Since the spectrum shape of a particular signal depends on the modulation used, the null-to-null bandwidth shown in Table 2.2 can be defined as either the width of the main spectral lobe, or the width between the outer nulls of the split main lobes (such as in binary offset carrier (BOC) modulation used for M-code), of the signal power spectrum. The bandwidth of a particular GNSS signal (hence the required bandwidth of a receiver for that signal) is then typically referring to the null-to-null bandwidth. Figure 2.5 also illustrates the frequency allocation plans for the current GNSS and other regional and augmentation systems. It is worth noting that many frequency bands used by different constellations overlap with, or are situated relatively close to each other. This facilitates the development and proliferation of multi-band GNSS receivers and receivers that covers two or more GNSS systems.

Table 2.2: Frequency characteristics of GPS and Galileo constellations [2, 12, 14–16].

GNSS Constellation	Center Frequency (MHz)		Null-to-Null Bandwidth (MHz)
GPS	L1 C/A	1575.42	2.046
	L1 P(Y)	1575.42	20.46
	L2 P(Y)	1227.6	20.46
	L2C	1227.6	2.046
	L5	1176.45	20.46
	L1 M	1575.42	30.69
	L2 M	1227.6	30.69
	L1C	1575.42	4.092
Galileo	E1	1575.42	24.552
	E6	1278.75	40.92
	E5b	1207.14	20.46
	E5	1191.795	51.15
	E5a	1176.45	20.46

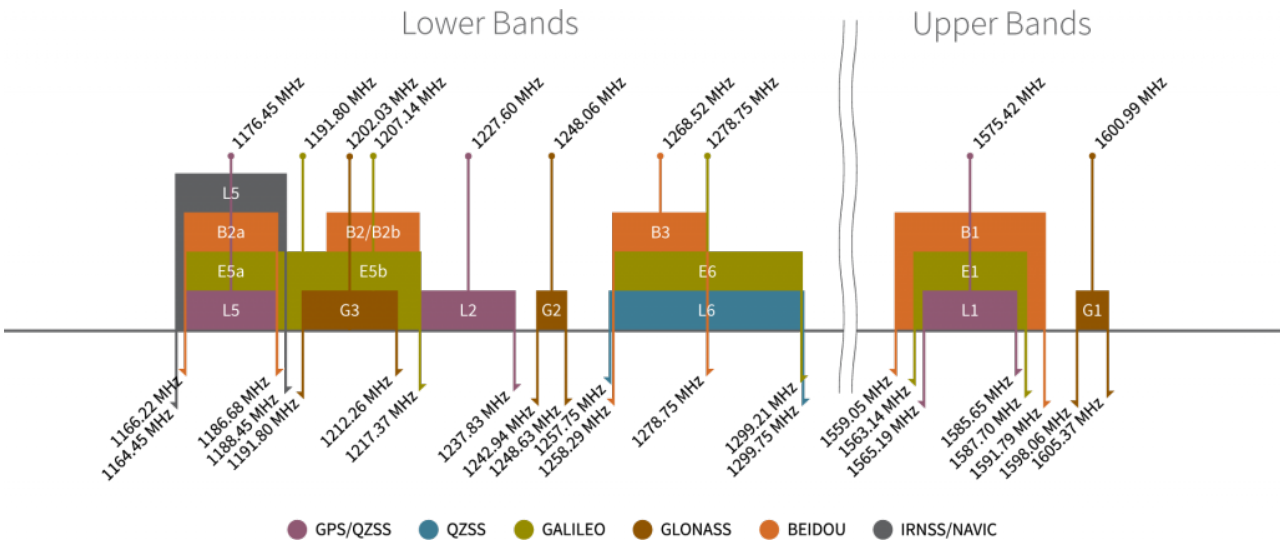


Figure 2.5: Illustration of the frequency allocations for several GNSS, regional, and augmentation systems [16].

CHAPTER 3

GPS Antennas: General Considerations

GPS ranging is accomplished using the concept of trilateration, as discussed in Chapter 2. Although the trilateration is conceptually straightforward, accurate range measurement is subjected to factors such as atmospheric effects in ionosphere and troposphere, receiver circuitry, multipath and other RF interference, and eventually, receiver antenna designs. Antenna is the first element in any GNSS receiver to process a user's positioning information. The antenna acts as both a frequency and spatial filter, therefore it will have a major impact on the overall performance of the receiver system, notably regarding coverage, positioning accuracy, and interference mitigation.

3.1 Polarization

The GPS satellite and receiver antennas exhibit right-hand circular polarization (RHCP), which offers several significant advantages. Circularly polarized signals are desired because linearly polarized signals undergo changes in polarization direction as they propagate through the ionosphere due to Earth's magnetic field, which is a phenomenon known as Faraday rotation [7]. Additionally, a linearly or elliptically polarized antenna must be properly aligned relative to the polarization of the incident wave to receive the maximum possible power, otherwise a misalignment will reduce polarization efficiency. This is another complication in the receiver design that can be avoided using circularly-polarized signals. Another benefit of using circular polarization (CP) is that it allows the receiving antenna to distinguish a direct signal from a strong multipath signal. Typically, the strongest multipath signals arrive at the receiving antenna from directions close to the horizon or below it, usually after reflection

from the surface of a good conductor, such as sea water or a building [7]. The polarization of a circularly-polarized signal will switch handedness upon reflection; in the case of GPS, the signal is changed to left-hand circular polarization (LHCP) from the original RHCP. Since RHCP and LHCP components of an electric field are orthonormal, a perfect RHCP receiving antenna is able to completely discriminate/reject a multipath signal that is perfectly LHCP.

3.2 Axial Ratio

The axial ratio (AR) is a metric for circularly and elliptically polarized antennas that measures the ratio of the major axis to the minor axis of the polarization ellipse. It can be used to assess a CP antenna's polarization efficiency and the ability to reject LHCP signals. AR is calculated as [2, 3, 7, 17]:

$$\begin{aligned} \mathrm{AR} &= \frac{\text{major axis}}{\text{minor axis}} = \left|\frac{E_{\max}}{E_{\min}}\right| \\ \mathrm{AR(dB)} &= 20\log \mathrm{AR} \end{aligned} \tag{3.1}$$

$E_{\max}$ and $E_{\min}$ are the maximum and minimum electric field magnitudes, which defines the major and minor axis of the polarization ellipse. For CP antennas, an axial ratio of 0 dB (unity) implies perfect CP. In practice, an antenna has an acceptable CP when its AR is less than 3 dB [3].

Axial ratio is inherently tied to the cross-polarization (X-pol) level; for GPS antenna, the X-pol is LHCP and the co-polarization (co-pol) is RHCP. A metric that measures the X-pol level is the cross-polarization discrimination (XPD), and can be determined from the AR as [7]:

$$\begin{aligned} \mathrm{XPD} &= \frac{|E_R|}{|E_L|} = \frac{\mathrm{AR}+1}{\mathrm{AR}-1} \\ \mathrm{XPD(dB)} &= 20\log \mathrm{XPD} \end{aligned} \tag{3.2}$$

where E_R and E_L denote the RHCP and LHCP components of the electric fields respectively. Using (3.2), for an acceptable CP with an AR less than 3 dB, the corresponding XPD

should be more than around 15.34 dB. For a GPS antenna, this means that the LHCP level should be more than 15.34 dB below the RHCP level across the desired angular range and range of frequencies for acceptable CP performance. In general, AR and XPD are strong functions in terms of the elevation and azimuth angles, and the characteristics may vary noticeably depending on the specific antenna design geometries and feed network used. For GPS antennas, maintaining a good AR over a beamwidth that can be very wide, extending from the zenith to a low elevation of 10 degrees for reasons that will be discussed in the following section, is usually difficult to achieve [7].

3.3 Radiation Pattern

Typically, a GPS antenna is placed where there is a clear view of the sky with ideally as little obstruction as possible. One of the key requirements then for such antennas is to provide a minimum gain needed and a sufficiently broad beamwidth for at least four GPS satellites in view. These factors dictate the radiation pattern for most traditional GPS antennas, and since gain and beamwidth vary inversely with each other, this creates an unavoidable trade-off.

The optimal radiation pattern for a typical GPS antenna is also limited by the required minimum elevation mask angle θ_m in the upper hemisphere [7, 14]. For GPS receivers, the elevation mask angle is a preset angle above the horizon; any satellite signal below the mask angle will be ignored, or "masked out," by the receiver [18]. This is depicted in Fig. 3.1. An ideal GPS antenna should receive radiations only above the minimum mask angle and have no radiation below it. The sharp cutoff at the elevation mask is thus a major design difficulty, since most realistic GPS antennas have smooth, cardioid-like patterns [1]. One consideration for setting the mask angle is related to receiver noise. The incoming GPS signal is already very weak; elevation mask helps to reduce potential atmospheric noise in the incoming signal by masking out satellites closer to the horizon, whose signals will travel longer, and thus degrade more, in the atmosphere than those of satellites closer to the zenith. Another consideration for elevation mask has to do with radiations below the mask

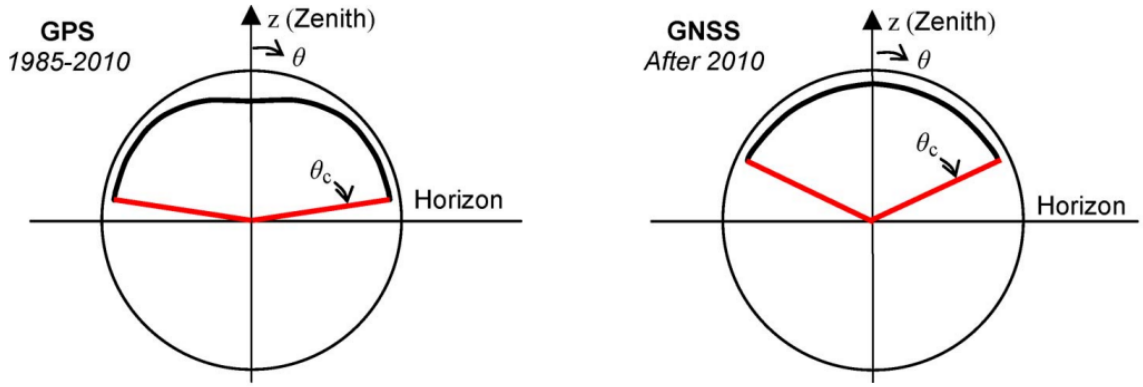
angle, which may subject the antenna and receiver to undesired effects such as multipath, interference, and jamming [1, 7].

Theoretically, for a 24-satellite constellation, an elevation mask of 5° will always ensure at least five satellites in view and at least seven satellites in view more than 80% of the time for a user at all altitudes [7]. Before 2010, the general convention is that for terrestrial applications, the elevation mask angle usually ranges from 10° to 15 or 20°, and −10° to 10° for airborne applications; a negative mask angle means it is below the horizon. After 2010, with the advent of multiple GNSS constellations in operation with more or less overlapping operating bands, in addition to receivers being developed capable of receiving from multiple constellations, there is a trend of moving toward a higher elevation mask angle. Traditionally, the concern was that a mask angle that is too high may exclude satellites needed for a minimum number of four. With a well-developed constellations in orbit, the likelihood of such scenario became increasingly low. Thus after 2010, for terrestrial applications, the elevation mask angle can be from 20° to even as high as 40°, and 10° to 30° for airborne applications. Other than the availability of more satellites near the zenith, factors contributing to the trend of higher elevation angles include the increase in interference signals coming from the horizon from other applications operating near the GPS L-bands, and the superior performance using near-zenith coverage enabled by GPS augmentation constellations such as Japan's Quasi-Zenith Satellite System (QZSS) [1].

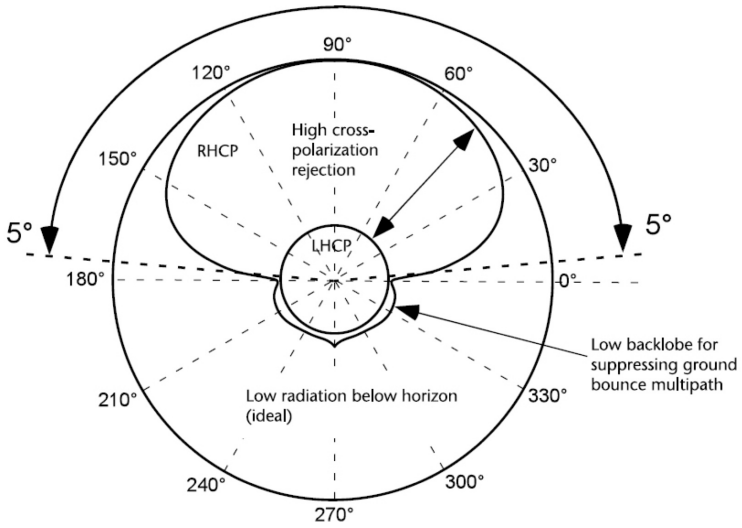
Ideally, a GPS antenna's RHCP far-field radiation gain pattern should exhibit a greater-than-unity (>0 dBic), constant (flat) gain value over the elevation angles above the elevation mask, and a sharp gain roll-off below the mask [7,14]. This is clearly depicted by the idealized radiation pattern shown in Fig. 3.1a. For an ideal, lossless, purely RHCP polarized antenna with such pattern, one can determine the directivity (or gain) as:

$$D = \frac{4\pi}{\Omega_A} = \frac{2}{1 - \sin\theta_m} \tag{3.3}$$

Suppose the elevation mask angle θ_m is set to 10°, the gain of this "ideal" radiation pattern is calculated to be around 3.8 dBic using (3.3); a gain below 3.8 dBic indicates there exists



(a)



(b)

Figure 3.1: (a) Idealistic GPS antenna radiation patterns with a relatively uniform RHCP directivity or gain value at the upper hemisphere and a sharp cut off at the elevation mask angle [1]. The elevation mask angle θ_m is equivalent to $90° - \theta_c$, where θ_c is the polar angle measured from the zenith shown in the figure. The figure on the right shows a higher elevation mask angle convention after 2010. (b) Example of a more realistic, though still some-what ideal, GPS antenna radiation pattern with a roughly 5° elevation mask [7].

radiations below the elevation mask, and a gain higher means a narrower antenna beamwidth and a more directive pattern. Obviously, realistic GPS antennas will not possess such ideal patterns or perfect RHCP polarization across the beamwidth, such as non-uniform gain above the elevation mask, non-zero backlobes and radiations below elevation mask, and degraded RHCP polarization at lower elevation angles that introduces polarization losses [7].

The pattern shown in Fig. 3.1 is for a generic baseline GPS antenna placed in idealized environment with open sky and no obstruction, and no further design constraints. Ultimately, the desired radiation pattern is still platform-specific, which may impose more, or less, specific operating requirements on the GPS antennas [1]. For example, though vast majority of GPS antennas would consider a sharp cutoff at low elevation mask angle desirable for reasons explained above, it may not be the case for antenna installed on platforms such as ships and aircraft. Since ships can experience considerable pitch and roll motions caused by heavy sea conditions, and similarly for aircraft during pitch, yaw, or roll maneuvers with significant orientation changes, a good gain at low elevation angles may actually be desired. This is in agreement with the mask angle convention discussed above, in which the mask angles are usually lower for airborne platforms, even extending to 10° below horizon before 2010, compared to terrestrial platforms [1, 7].

It is also noteworthy here to address the notion of beamwidth for GPS antennas. Typically, when mentioning the beamwidth of an antenna, it is referring to the 3-dB beamwidth, or half-power beamwidth (HPBW), defined as the angular range whose boundary is at a level 3 dB below the main lobe peak. However, this conventional definition, though important for point-to-point communications links, is regarded to be of little significance when evaluating the performance of a GPS antenna. As mentioned before, typical GPS antennas are required to have very broad beam, extending from the zenith to the elevation mask, to acquire as many satellites as possible in the upper hemisphere. Therefore the gain in this large angular range has to satisfy a minimum requirements that is considered crucial for GPS signal acquisition. The minimum gain beamwidth extends from the zenith down to the elevation angle as long as the gain of the GPS antenna satisfy the minimum gain requirement within this angular range [7]. The minimum gain requirements at different elevation angles for the

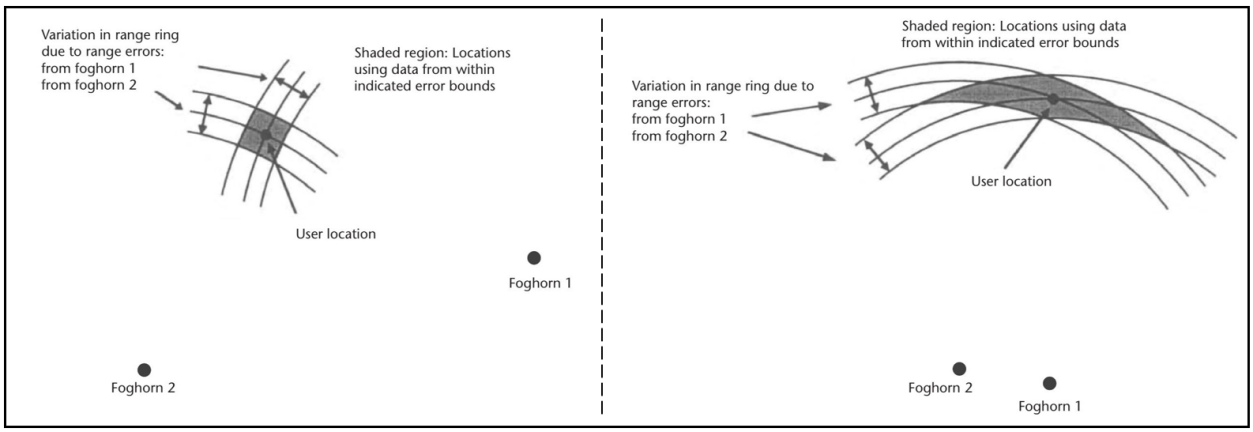
L1–E5 bands are listed in Table 3.1.

Table 3.1: Minimum gain requirement versus elevation mask angle for the L1–E5 bands [7].

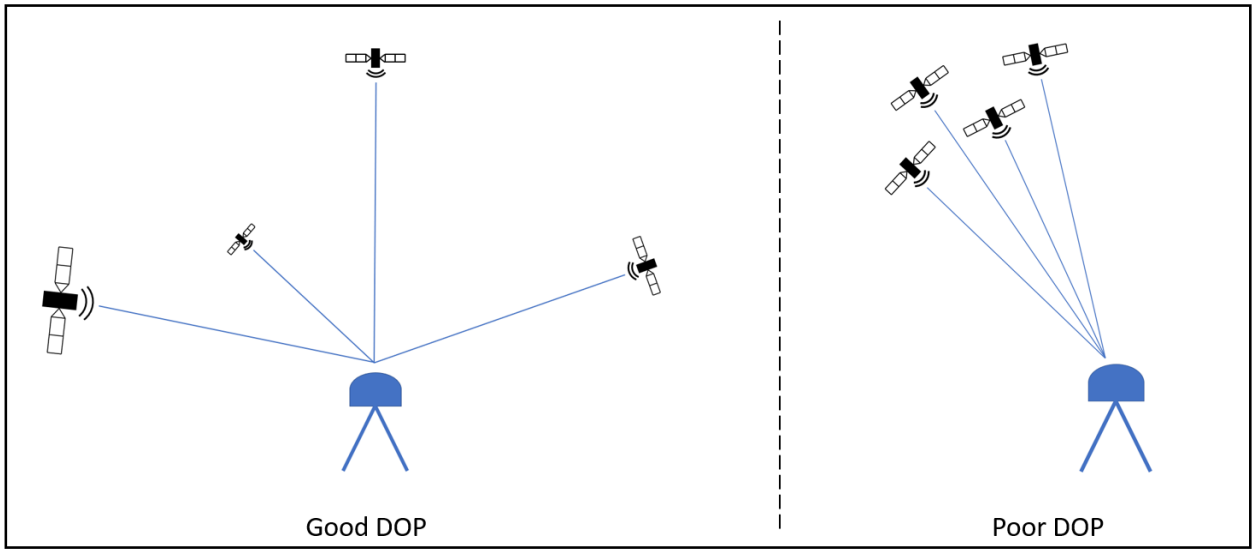
Elevation mask angle (°)	Minimum gain (dBic)
> 15	-2
10	-3
5	-4.5
0	-7.5

3.3.1 Dilution of Precision

It is worth mentioning here the notion of dilution of precision (DOP), which is a measure of the geometry of the visible satellites [18]. More specifically, DOP is a numerical value that indicates the confidence of the position measurement obtained by the receiver from the geometry of the visible satellites [7]. Ideally, one would want the four or more GPS satellites to be equally spaced around the receiver. However, this is not possible, since the receiver will only see the upper hemisphere, as the Earth itself is blocking the bottom hemisphere. Therefore, the second best orientation would be to have a satellite at the zenith directly above the receiver, and the other three equally spaced around the receiver, elevated to about 25 to 30 degrees above the horizon [18]. If the satellites are clustered together, it will result in poor DOP values; this can be illustrated in Fig. 3.2. This is another reason why an elevation mask that is too high may be undesirable, since it may exclude satellites at lower elevations for a good DOP. This is a particular concern for GPS precision in mountainous terrain and in large cities with densely packed high-rises, not to mention the fact that high-rise buildings also serve as near-perfect reflectors that produce multipath signals. For these cases, one would imagine the only visible satellites are the ones almost directly above. This is some what alleviated by the deployment of GPS augmentation constellations such as QZSS, which operates satellites in geosynchronous orbit (GSO); GSO satellites are situated near-zenith, which can directly aid in signal reception in urban environment [1, 18].



(a)



(b)

Figure 3.2: (a) Positioning accuracy as relates to the relative geometry of the user and signal emitters [2]. DOP is a common consideration used in time-of-arrival ranging, which calculates the transit time of signals from position reference sources such as foghorns, GPS satellites, or other signal emitters. As the figure on the right suggests, signal sources that are clustered together will result in higher uncertainties and errors in range measurements, and hence, poor DOP values. (b) Figures illustrating GPS satellite geometries that would result in good or poor DOP values.

3.4 Bandwidth

There are three main definitions used for antenna bandwidth. The first one, the impedance bandwidth, is by far the most common, which specifies the range of frequencies where adequate impedance matching is achieved. More specifically, it is defined as the frequency range within which over 90% of the incident power is delivered to the antenna, or equivalently, the S_{11} (the reflection coefficient) is below -10 dB; this is the typical impedance matching condition. This bandwidth can also be expressed as a percentage of the center frequency f_0, known as the fractional bandwidth %BW:

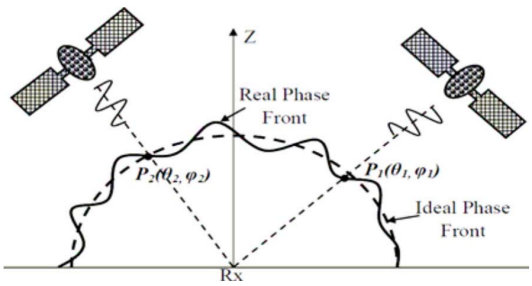
$$\%\text{BW} = \frac{\text{BW}}{f_0} \times 100\% \tag{3.4}$$

where BW on the right-hand side is the impedance bandwidth.

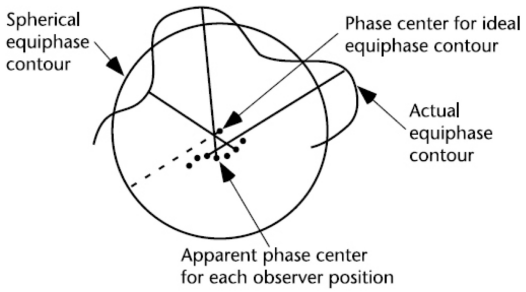
Two other less used, though important, definitions of bandwidth is gain bandwidth and axial ratio bandwidth. The gain bandwidth is defined as the frequency range over which the gain is above certain minimum threshold values. For GPS antenna specifically, the minimum gain values are required for the antenna to acquire the GPS satellites from the zenith to the elevation mask; they are specified in Section 3.3 above. The AR bandwidth is defined as the range of frequency whithin which the antenna is able to meet the maximum X-pol level requirement (AR <3 dB) for a desired angular range. Most typically, the impedance bandwidth is the limiting factor, which often can be almost an order of magnitude smaller than the gain bandwidth [7].

3.5 Phase Center Stability

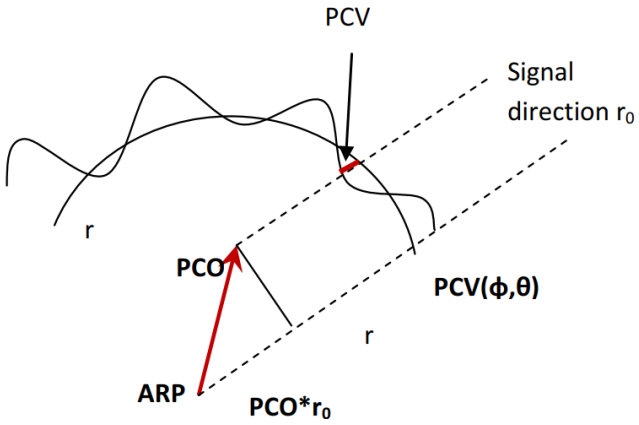
The concept of phase center is related to equiphase surface; for the far-field radiation pattern, the equiphase surface can be approximated as a spherical surface, and the center of the sphere is equivalent to the phase center of the antenna [3, 21]. More formally, the phase center of an antenna can be defined as the point at which the spherical phase fronts of the far-field radiation pattern seem to emanate from when the antenna is in transmit mode. This means



(a)



(b)



(c)

Figure 3.3: (a) Ideal versus realistic equiphase front radiated by an antenna [19]. The actual phase front will not be perfectly spherical and is depended on the angular positions, which will introduce different phase offset errors for the signals received by the antenna. (b) Another illustration of ideal versus actual equiphase surface [7]. The apparent phase center is different at each angular observation positions; the average phase center is the mean location of the apparent phase centers for a defined angular range. (c) The antenna reference point (ARP) is the geometric center of the antenna. The difference between the ARP and the average phase center is the phase center offset (PCO). The difference between the phase center at a particular angular position and the average phase center is the phase center variation (PCV) at that position [20].

that the phase center is also the point at which the electromagnetic energy seems to be received by the antenna, assuming reciprocity holds, which is true for most conventional antennas. Hence, the spatial location of the phase center of a receiving antenna is the actual 3D position measured by a GPS user/receiver. For an ideal antenna, the phase center would be at the center of the far-field spherical, equiphase surface surrounding the antenna; realistically, this is unattainable for any practical antenna designs simply because of the radiation mechanisms. The equiphase surface, and hence the phase center, of an antenna is then largely depended on the operating frequency and polarization, and has a range of valid solid angle around the antenna broadside for that phase center [3]. Fig. 3.3 illustrates the concept of antenna phase center. For a GPS receiver, particularly in high-precision applications, the stability and repeatability of the antenna phase center, when receiving satellite signals in different frequencies and elevation and azimuth angles, are therefore crucial features in order to achieve millimeter-level precision [7].

3.6 Multipath Rejection

Multipath remains a major source of error for both pseudorange and carrier-phase measurements. Multipath error in carrier-phase can reach a maximum value of a quarter of a cycle, which is about 4.8 cm for L1 carrier phase, and in C/A-code pseudorange measurements can be potentially up to several tens of meters in theory; in practice, this error can typically be reduced to several meters using mitigation techniques [4]. Methods of mitigating multipath effects for GPS receivers can be classified into two main approaches: antenna based or signal processing based. Most signal processing based approaches have some ways of comparing the direct and multipath signals. At a basic level, for example, one can recognize that the multipath signal travels a longer path, thus its signal will arrive slightly later with a slightly weaker power. Although signal processing methods have progressed substantially, multipath-induced errors still occur frequently, even with high-performance correlator receivers [3].

Antenna based approaches rely on reducing the reception of multipath signals through clever design choices. Choices such as the settings where the antenna is placed can be sig-

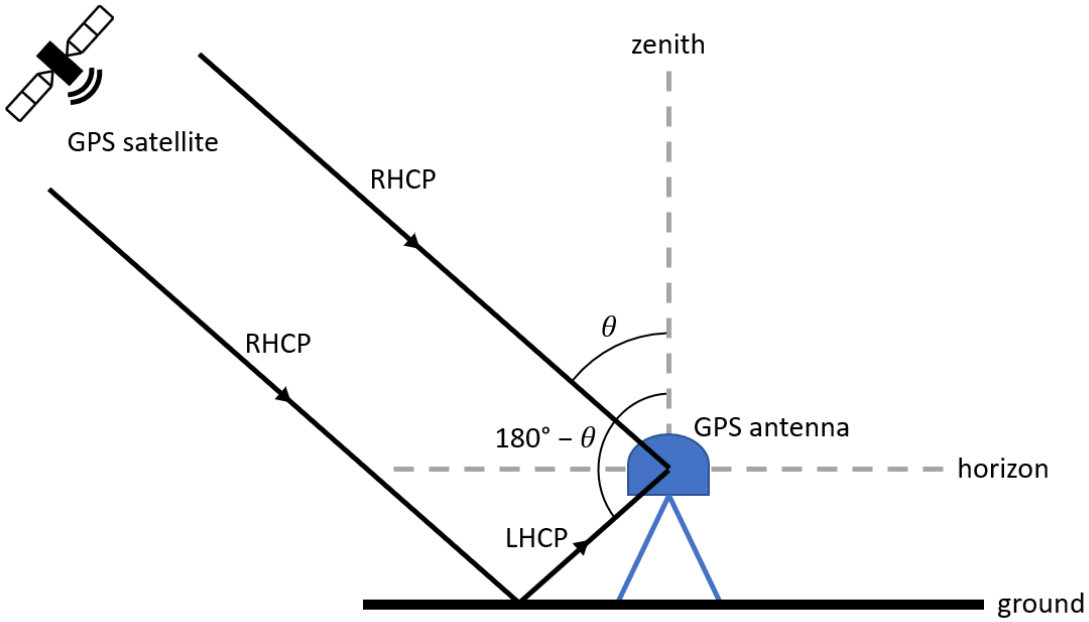


Figure 3.4: Direct and multipath signals received by a GPS antenna.

nificant; for example, applying coating with RF-absorptive materials of reflective structures close to the antenna, and placing the antenna closer to the ground that decreases the path lengths of multipath signals coming from below [2]. But more than often, antennas are designed inherently to reject or minimize multipath reception. Since the strongest multipath signals usually arrived from the bottom or low elevation angles, GPS Antennas should be designed with very small backlobes. Antenna designs that place choke ring, electromagnetic band gap (EBG) metamaterials, or resistively tapered ground planes around the rim of antennas are very successful in eliminating multipath signals arriving from the ground and low-elevation scatterers [7].

It is important then to discuss multipath rejection ratio (MPR) as a metric for an antenna's multipath rejection quality. The MPR is defined as [7, 22]:

$$\mathrm{MPR} = \frac{G_{\mathrm{RHCP}}(\theta)}{G_{\mathrm{RHCP}}(180^\circ - \theta) + G_{\mathrm{LHCP}}(180^\circ - \theta)} \tag{3.5}$$

where G_{RHCP} and G_{LHCP} are the RHCP and LHCP gains respectively, and θ is the polar angle measured from the zenith. These variables are illustrated in Fig. 3.4.

A related alternative metric for evaluating multipath rejection quality is the desired-to-undesired (D/U) ratio, also known as up/down ratio or front-to-back ratio, defined as [7]:

$$\text{D/U ratio} = \frac{G_{\text{RHCP}}(\theta)}{G_{\text{LHCP}}(180^\circ - \theta)} \tag{3.6}$$

For an ideal antenna possessing negligible gain in both right- and left-hand CPs in the backward direction (bottom hemisphere), the denominator in (3.5) or (3.6) will approach zero, and the MPR or D/U ratio of such antenna will approach infinity; in other words, this antenna will have a perfect multipath rejection ability. Though this is not realistically achievable, the gain below the elevation mask should be as low as possible to avoid multipath interference [7].

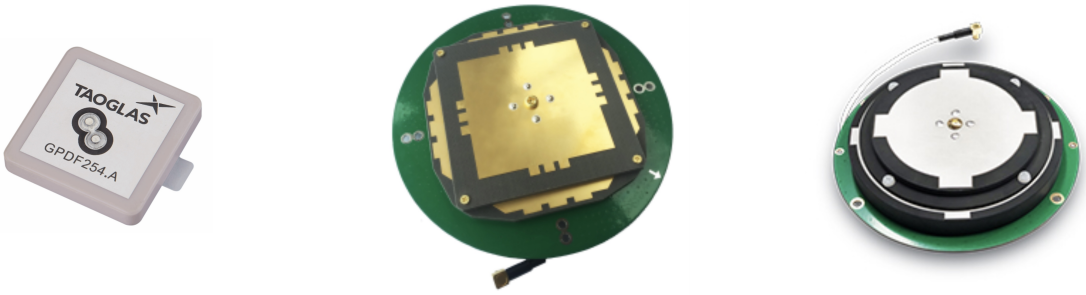
3.7 GPS Antenna Designs

Designs for GPS antenna vary greatly in accordance with specific requirements and imposed constraints. A detailed description on all of them will be well-beyond the scope of this section. Instead, a brief overview on some of the most widely used and proven antenna choices for GPS reception is presented.

3.7.1 FRPA: Patch Antenna

Fixed radiation pattern antennas, or FRPAs, have fixed radiation patterns as the name suggests, as opposed to controlled radiation pattern antennas (CRPAs), such as beamforming array antennas. FRPAs include some of the most popular and common GPS antenna choices. A FRPA possesses a more or less omnidirectional pattern that provides a better-than-minimum required gain in the upper hemisphere and is designed to acquire at least four GPS satellites [1, 7].

Microstrip patch antennas are by far the most universally employed FRPA antenna design used in both civil and military GPS applications. This is facilitated by their typical advantages such as low profile, medium-to-small size, possessing a structure that can conform



(a) (b) (c)

Figure 3.5: (a) GPDF254.A, a L1-band dual-feed GNSS patch antenna by Taoglas [23]. (b) HX-CS7017A, a L1/L2 dual-band, four-feed, stacked patch antenna by Harxon [24]. (c) HX-CSX078A, another four-feed, stacked patch GNSS antenna design by Harxon, featuring circular patches [25].

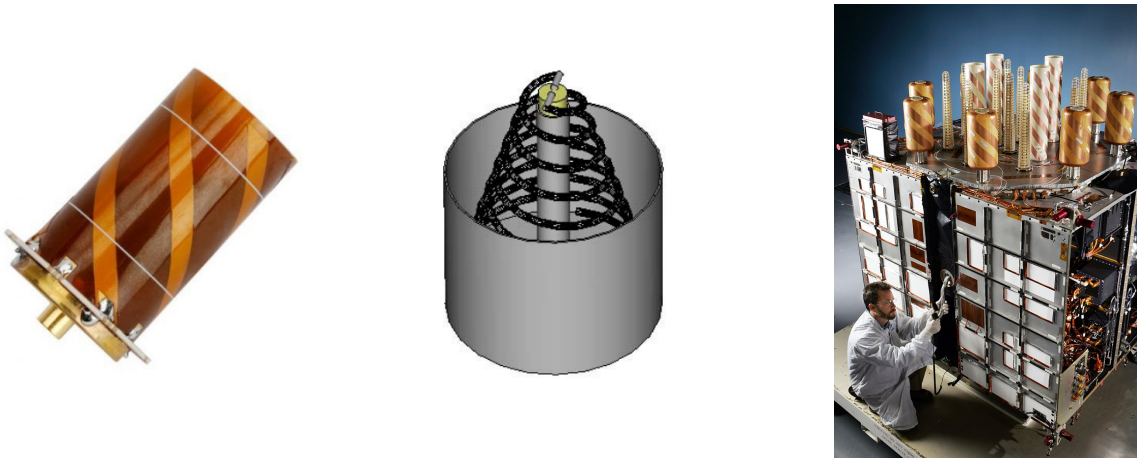
to the platform, relatively easy to obtain RHCP from simple feed networks, and importantly, low cost manufacturing [3,7]. CP can be obtained using specific patch structures, such as the truncated corner patch, or using two or four orthogonally placed feeds with 90° successive relative phase shifts between the feeds; some of the representative commercially-available patch antennas are shown in Fig. 3.5. The three most used patch shapes for achieving circular symmetry in azimuth are a square, a circle, or an annular ring; square and circular patch designs are shown in Fig. 3.5. The annular ring is the basis of the design for this thesis, and will be discussed in much more details in Chapter 4.

The dielectric substrate selection is usually one of the most crucial steps that affects characteristics such as patch size, height, bandwidth, and gain of the antenna. More than often, the substrate choice involves the trade-off between these characteristics. For example, a substrate with high permittivity may make the size of the patch more compact at the expense of reducing both gain and bandwidth and also increasing the radiation close to the horizon via surface waves [7]; a substrate with high dielectric loss tangent may be very low cost and enable a broader bandwidth at the expense of a decreased gain. Ultimately, the substrate choice would defer to the specific real-world, application-driven constraints.

Patch antennas are obviously not without their drawbacks. Being a type of resonant

antenna, a patch antenna is therefore inherently narrow-band. The fractional bandwidth for an ordinary patch antenna is only around a few percent. A majority of the GPS patch antennas used in low-cost, commercially available receivers operate only over a single frequency band, the L1 band, with a single-layer substrate; some have only an around 2 MHz bandwidth just enough to receive the L1 C/A-code. Many elaborate efforts have been made to increase the bandwidth or add more bands of operation. Additionally, the AR for patch antenna generally degrades significantly with decreasing elevation angles. This is because all microstrip patch antenna possesses a ground plane, on which the horizontal component of the electric field must become zero to satisfy the boundary condition. This makes the antenna linearly polarized, with polarization almost vertical near the ground plane. Hence, the AR can be observed to approach a very large value at smaller elevation angles. Due to this degradation of AR, the signal received from satellites at lower elevation angles can experience a significant polarization loss compared to those at higher elevation angles. The RHCP gain may decrease from the zenith to the horizon by as much as 10 dB [7]. There are ways to overcome this such as mounting a choke ring around the patch antenna, which will be discussed in the sections below.

As market demand grows, operating in dual L1 and L2 bands are increasingly desired for high measurement accuracy; the trend may be moving to even tripled band in the future as L5 matured and proliferated, or even broadband operation covering the entire GNSS spectrum of around 1146–1616 MHz [1]. Multiband or broadband patch antennas are usually much more difficult to achieve. Multiband patches usually combine two or more patches resonating at the desired bands and either stack them on top of or parasitically coupled to each other. Stacked multi-band GPS antennas are shown in Fig. 3.5b and 3.5c. Usually the smaller L1 patch is placed on top of the larger L2 patch. The larger bottom patch acts as the ground plane for the top patch; both patches have negligible reactance effect on the other patch when they are resonant on their own respective frequencies.



(a) (b) (c)

Figure 3.6: (a) HC771E, a L1-band quadrifilar helix antenna by Tallysman [26]. (b) A wide-band conical spiral antenna designed for GPS and Galileo reception [27]. (c) A GPS Block IIR-M satellite, showing its antenna farm [28]. Quadrifilar helix and spiral antenna elements are also commonly used in GNSS satellites for UHF communications and transmission of L-band navigation signals [3].

3.7.2 Other FRPA Designs

Other popular GPS FRPAs worth mentioning include but not limited to quadrifilar helix and spiral antennas, which are shown in Fig. 3.6.

The quadrifilar helix antenna (QHA) consists of four helical elements wrapped around a cylinder at the desired pitch angle; the CP is achieved by either adjusting the length of the four helical elements, known as self-phasing QHA, or fed the elements with successive 90° phase shifts, called externally-phased QHA,. The QHA possesses multiple advantages such as light weight, inexpensive to build, and have no need of a ground plane that enabled its compact size. But the main attractiveness is its ability to produce a broad RHCP beam, of up to 100° or even 180°, and is the only antenna design that can sustain a reasonable RHCP radiation at low elevation angles; it is typically used in handheld and mobile receivers [3, 7].

High precision geodetic-grade GPS antenna designs often uses some form of planar or conical spirals. Highly precise phase center accuracy and stability requirements are best

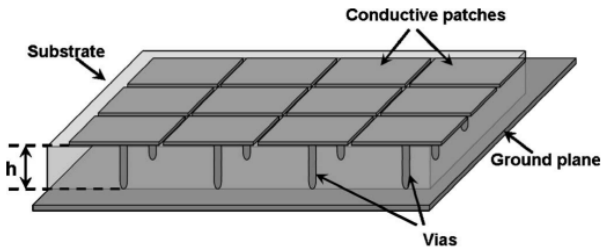
fulfilled with a well-designed spiral antenna, such as a conical spiral mounted on top of a ground plane. This is mainly because the RHCP characteristics is achieved intrinsically in the spiral structure itself, and the bandwidth is frequency-independent for such spirals, leading to a highly uniform gain pattern with sustained RHCP and rejection of LHCP at low elevation angles [7].

3.7.3 Mitigation of Multipath and Platform Effects

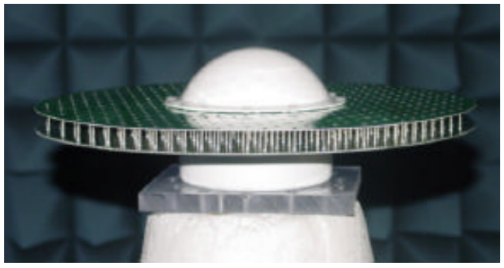
As described in previous sections, typical GPS antennas are designed with very wide beamwidth, which may cause significant illumination on the platform on which they are installed. The size, shape, and material of the platform can then heavily influence the resulting radiation patterns. Patch antenna's ground plane are able to suppress multipath signal reflecting from the ground. However, diffraction from ground plane edges may interfere with the primary radiated signal, causing ripples in the radiation patterns that affect the gain and phase center of the antenna as a function of observation angle, and also increase the antenna backlobes, which make the antenna more susceptible to multipath. A simple demonstration on the platform effects on the performance of two GPS antenna designs is shown in Appendix B. These platform effects (also called ground plane effects) are important in high-precision GPS applications when submeter to even subcentimeter accuracy is required [7]. Many types of modified ground planes are employed to mitigate edge diffraction from ground plane in additional to suppressing influence from multipath signals. These are choke ring, EBG ground plane, rolled edge ground plane, and resistively tapered ground plane. A brief survey of each will be given below.

Choke rings are widely used for shaping the radiation pattern to mitigate multipath effects and stabilize the phase center. Typically, they are constructed with three to five concentric metallic circular corrugations with the GPS antenna element placed at the center. A typical choke ring antenna is shown in Fig. 3.7a. The depth of the corrugations are roughly a quarter-wavelength at the desired operating bands, mostly L1 or L2. The corrugations are deigned to function as quarter-wavelength transformer, in which the short circuit at the

(a)



(b)



(c)

Figure 3.7: (a) HX-CGX606A, a multi-band GNSS antenna by Harxon, featuring a 2D choke ring around the antenna element at the center [29]. (b) A typical EBG structure [30]. (c) A EBG structure installed around a GPS antenna [31].

bottom/base of the corrugation is impedance transformed to an open circuit at the top, where the antenna element is placed. As such, they present a high impedance surface to the electric field polarized perpendicular to the corrugations. This way, they can effectively suppress the propagation of surface waves and prevent edge diffraction from ground planes. The resulting radiation pattern has a sharp cutoff near the horizon with low susceptibility to multipath, suppressed LHCP levels over the entire upper hemisphere, and a smooth main beam roll of down to the horizon with no ripples. Choke rings are heavily used in high-precision geodetic-grade surveys. Despite their performance, choke rings are generally bulky, heavy, and expensive. A typical choke ring can have a diameter of tens of centimeter and weights around 5 kg or more, which may limit their viability on many platforms [7].

Compared to choke rings, EBG ground planes are relatively compact and lightweight, able to function with thickness much less than a quarter-wavelength. A typical EBG ground plane surface is a frequency selective surface (FSS) comprises of a two-dimensional array

of subwavelength metal discs with pins, or subcells, connecting them to the ground plane beneath, as shown in Fig. 3.7b and 3.7c; these cells are aptly called "mushroom" structures [32]. Each subcell funcitons as an LC resonant circuit. The EBG surface can provide high impedance for both polarizations for all propagating directions [7].

Rolled edge ground plane prevents diffraction from sharp edge by simply rolling the edge under the ground plane. They are easier to manufacture compared to choke rings or EBG ground planes. However, they do not provide much backlobe suppression due to creeping wave diffraction. This can be mitigated by coating the rolled edge with radar absorbing materials (RAM), though it will increase the weight and cost [7].

Resistively tapered ground plane employs gradual increase of surface resistance from the center to the outer edge of the ground plane. The advantage of this design is its broad range of frequency use as opposed to limiting to certain frequency bands like the choke ring or EBG ground planes. This type of ground planes have demonstrate the ability to suppress the diffraction ripples and backlobe radiations [7].

CHAPTER 4

Shorted-Annular-Ring Patch Antenna

Multipath is one of the major factors that limits the accuracy of a high-precision GPS receiver, and it is one factor that can be more or less mitigated via antenna design, in addition to signal processing techniques and receiver designs [2, 6]. As discussed in Chapter 3, multipath-mitigating antennas typically involve specially-designed ground planes such as choke rings and EBG structures that also suppress deleterious ground plane effects. However, these ground planes are usually large and expensive, and are thus unable to satisfy our design requirements. Shorted-annular-ring (SAR) patch antenna is an attractive design with the unique characteristics of not exciting surface wave modes, which is favorable for multipath limitation. It also possess the advantages enjoyed by a typical patch antenna, which are compact size, relatively low complexity, and low cost manufacturing. These make it a promising design competitive with the much more expensive choke ring antennas in high-precision GPS applications.

4.1 Surface Waves in Patch Antennas

It is important to discuss the concept of surface waves and theirs effect on GPS antennas. All patch antennas possess a grounded dielectric substrate, and thus generate surface waves. Surface waves are EM modes that are "trapped" in the dielectric substrate layer and propagate via successive internal reflections between the dielectric-air boundary and the ground plane. The surface waves travels through the substrate and once they reach the edge of the substrate or the ground plane, they are reflected or diffracted into air, as shown in Fig. 4.1. The diffracted radiation may interfere with the main radiation from the antenna and cause

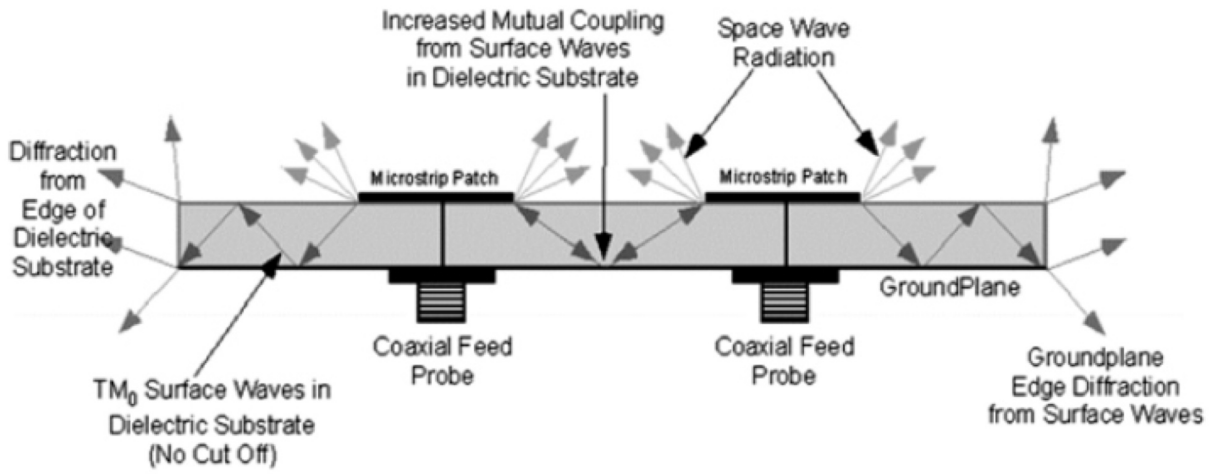


Figure 4.1: Surface waves propagate through the substrate, causing deleterious effects such as radiations from edge diffraction and increased mutual coupling between antenna elements [7].

several deleterious effects on the antenna performance; these include increased multipath from the backlobes, higher X-pol levels, and ripples in the antenna pattern. Additionally, the surface wave will be a source of increased mutual coupling and interactions between adjacent antenna elements or metallic objects positioned near the antenna, that will result in a change in the antenna pattern. This is a particular concern for antenna arrays with closely space antenna elements [7, 33].

Surface waves excited in a grounded dielectric substrate can be in either transverse magnetic (TM) or transverse electric (TE) modes. The cutoff frequencies of the surface wave modes are give by [33, 34]:

$$f_c = \frac{nc}{4h\sqrt{\epsilon_r - 1}}, \qquad n = \begin{cases} 1, 3, 5, ... & \text{for TE}_n \text{ mode} \\ 0, 2, 4, ... & \text{for TM}_n \text{ mode} \end{cases} \tag{4.1}$$

where c is the speed of light, ϵ_r is the relative permittivity of the substrate, and h is the substrate thickness. Equation (4.1) can be obtained by solving two simultaneous transcendental equations, which are derived from applying the wave equations on a grounded dielectric substrate structure, detailed in [34]. It should be noted that for any nonzero substrate thickness

with permittivity greater than 1, there is always a propagating mode, the TM_0 mode. This lowest order, dominant TM_0 mode has no cutoff frequency, as shown in (4.1), and can be excited at any frequency, and therefore has the greatest effect on GPS antennas [7, 34]. The propagation of the next lowest mode, TE_1 mode, can be mitigated by using a thin substrate, such that h satisfies the condition:

$$\sqrt{\epsilon_r - 1} k_m h < \frac{\pi}{2} \tag{4.2}$$

where k_m is the free-space wave number of the maximum frequency f_m in the intended operating frequency band; again, this condition can be derived from [34]. Equation (4.2) can also be reexpressed as:

$$h < \frac{c}{4 f_m \sqrt{\epsilon_r - 1}} \tag{4.3}$$

An alternative practical formula for the condition of substrate thickness h is given by [7, 35] as:

$$h < \frac{0.3c}{2\pi f_m \sqrt{\epsilon_r}} \tag{4.4}$$

In general, decreasing the substrate thickness and/or permittivity, less energy is coupled into the surface waves. However, these are done at the expense of a reduction in antenna bandwidth and radiation efficiency, or an increase in the patch size [33].

For ordinary GPS patch antennas that are placed ideally where there are no obstruction, the surface wave may be beneficial for patches that use high permittivity substrate. This is because high permittivity substrate enhances the surface wave radiation in direction close to the horizon, which can improve the directivity of the patch antenna at low elevation angles for acquiring GPS satellites [7]. Nonetheless, these theoretical advantages need to be evaluated with consideration of the negative impact resulting from using high permittivity substrate, such as increased backlobes and susceptibility to multipath, as discussed previously.

4.2 Circular Patch Antennas

In order to understand the operating principles of a shorted-annular ring patch antenna, which is the design studied in this thesis, an ordinary circular patch antenna should be studied first.

4.2.1 Cavity Model

A generic circular patch on a dielectric substrate is shown in Fig. 4.2. The modes supported by a conventional circular patch can be found by treating the patch, the ground plane, and the materials in between together as a circular cavity. The theoretical treatment in finding the modes of a circular patch can then be done using cavity model, which approximates the cavity as bounded by perfect electric conductors (PECs) on the top and bottom, and by perfect magnetic conductors (PMCs) along the perimeter of the patch ($\rho = a$). PMC can be used for modeling the "side walls" because of the assumption of negligible current flow around the edge of the patch from the bottom to the top side of the patch metal layer, and that most of the charges and current flows are concentrated at the bottom side of the patch metal. By this assumption, zero current flow to the top will create no tangential magnetic field components at the perimeter of the patch, and hence, one can model the side as a

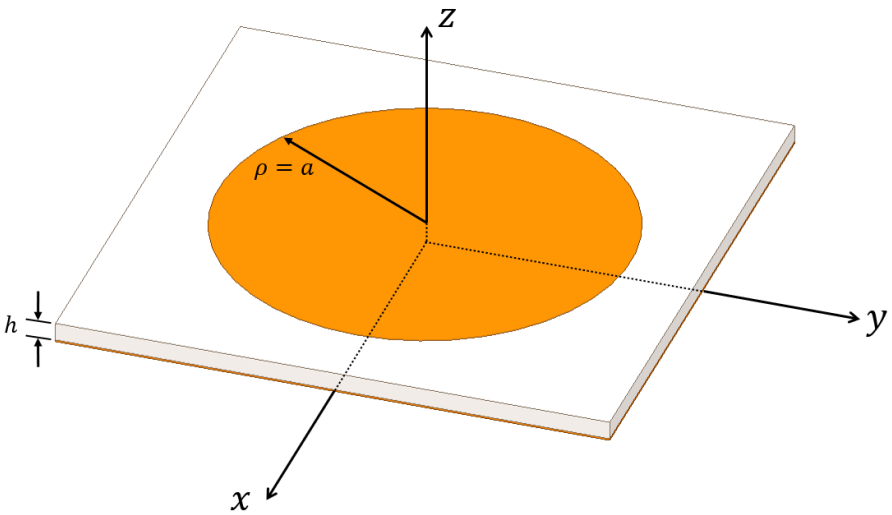


Figure 4.2: Basic geometry of a circular patch.

perfect magnetic wall. Though realistically there will not be exactly zero current flow to the top, this is a good approximation for thin substrates. There are other valid approximations that can be made for thin substrates ($h \ll \lambda$). First, the field variation along the height (z-direction) will be considered as constant. Additionally, because the substrate is thin, fringing fields will be very small, and therefore the electric fields will be nearly normal to the surface of the patch along the patch perimeter. Hence, the primary modes of interest are TM^z, where z is the direction perpendicular to the patch. In general, the cavity model becomes more accurate as the substrates becomes thinner [21].

The TM^z modal fields for the cavity model can be determined once the magnetic vector potential A_z is found. It is sufficient to work with only the z-component of the vector potential because we have established that the electric fields are in the z-direction perpendicular to the patch; the other components of the vector potential can be set to zero. With this simplification, the fields within the cavity, in cylindrical coordinates, can be found as [17,21]:

$$\begin{aligned} E_\rho &= -j\frac{1}{\omega\mu\epsilon}\frac{\partial^2 A_z}{\partial\rho\partial z} & H_\rho &= \frac{1}{\mu}\frac{1}{\rho}\frac{\partial A_z}{\partial\phi} \\ E_\phi &= -j\frac{1}{\omega\mu\epsilon}\frac{1}{\rho}\frac{\partial^2 A_z}{\partial\phi\partial z} & H_\phi &= -\frac{1}{\mu}\frac{\partial A_z}{\partial\rho} \\ E_z &= -j\frac{1}{\omega\mu\epsilon}\left(\frac{\partial^2}{\partial z^2}+k^2\right)A_z & H_z &= 0 \end{aligned} \tag{4.5}$$

These fields are subject to the following boundary condition:

$$\begin{cases} E_\rho\left(z=0,h\right)=0 \\ E_\phi\left(z=0,h\right)=0 \\ H_\phi\left(\rho=a\right)=0 \end{cases} \tag{4.6}$$

The boundary conditions in (4.6) is consistent with the cavity model, with PECs on the top ($z=h$) and bottom ($z=0$) and a PMC wall along the patch perimeter ($\rho=a$).

The vector potential A_z within the cavity, in cylindrical coordinates, must obey the scalar homogeneous wave equation:

$$\nabla^2 A_z(\rho, \phi, z) + k^2 A_z(\rho, \phi, z) = 0 \tag{4.7}$$

The solution form for A_z can be obtained using the method of separation-of-variables:

$$A_z(\rho, \phi, z) = f(\rho)\, g(\phi)\, h(z) \tag{4.8}$$

One can apply (4.8) into (4.7), then follow the well-established steps of solving the scalar homogeneous wave equation using the method of separation-of-variables in cylindrical coordinates to obtain a general solution form of A_z. The solution form of A_z satisfying (4.7), with physical considerations for the fields within the cavity, is expressed as [17, 21]:

$$\begin{aligned} A_z = &\left(A_1 J_m(k_\rho \rho) + B_1 Y_m(k_\rho \rho)\right) \cdot \left(A_2 \cos(m\phi) + B_2 \sin(m\phi)\right) \\ &\cdot \left(A_3 \cos(k_z z) + B_3 \sin(k_z z)\right) \end{aligned} \tag{4.9}$$

where J_m is the Bessel function of the first kind of order m, and Y_m is the Bessel function of the second kind of order m. Solution (4.9) is also in accordance with the constraint, or dispersion, equation:

$$k_\rho^2 + k_z^2 = k^2 = \omega^2 \mu \epsilon \tag{4.10}$$

The solution form of (4.9) can be further simplified if we consider one more physical characteristics of the field. The fields within the cavity must be finite, hence we can determine that B_1 must be zero, since Y_m will approach infinity at the center of the patch ($\rho = 0$). Applying (4.9) into (4.5) and enforcing the boundary conditions of (4.6), we can additionally determine $B_3 = 0$, and the following relations:

$$k_\rho = \frac{\chi'_{mn}}{a}, \qquad \begin{cases} m = 0, 1, 2, ... \\ n = 1, 2, 3, ... \end{cases} \tag{4.11}$$

$$k_z = \frac{p\pi}{h}, \qquad p = 0, 1, 2, ... \tag{4.12}$$

where χ'_{mn} is n^{th} zero of the derivative of the Bessel function $J_m(x)$ and determines the order of the resonant frequencies. The first four values of χ'_{mn}, in ascending order, are:

$$\begin{aligned} \chi'_{11} &= 1.8412 \\ \chi'_{21} &= 3.0542 \\ \chi'_{01} &= 3.8318 \\ \chi'_{31} &= 4.2012 \end{aligned} \tag{4.13}$$

We can then write our solution of A_z for a specific TM^z_{mnp} mode as:

$$A_z = B_{mnp} J_m\left(k_\rho \rho\right)\left(A_2 \cos\left(m\phi\right) + B_2 \sin\left(m\phi\right)\right)\cos\left(k_z z\right) \tag{4.14}$$

where B_{mnp} represents the coefficient after merging coefficients A_1 and A_3 in (4.9), and k_ρ and k_z obey (4.10)–(4.12).

Equation (4.10) and (4.12) also entail why electric field variation along the z-direction has to be constant for thin substrate (small h). If the electric field varies in the z-direction, then $p \neq 0$, which also means $k_z \neq 0$. In this case, small h may make k_z too large to satisfy the dispersion equation of (4.10).

4.2.2 Resonant Frequencies

As discussed previously, since typical patch antenna substrates are thin ($h < 0.05\lambda_0$), the fields are constant along the z-direction, which implies $p = 0$ and $k_z = 0$. Thus, the resonant frequencies of TM^z_{mn0} modes are determined to be:

$$f_{mn0} = \frac{k}{2\pi\sqrt{\mu\epsilon}} = \frac{1}{2\pi\sqrt{\mu\epsilon}}\sqrt{k_\rho^2 + k_z^2} = \frac{1}{2\pi\sqrt{\mu\epsilon}}\left(\frac{\chi'_{mn}}{a}\right) \tag{4.15}$$

As seen from (4.13), the dominant mode for a circular patch antenna is the TM^z_{110} mode; using (4.15), it's resonant frequency is then:

$$f_{110} = \frac{1.8412}{2\pi a\sqrt{\mu\epsilon}} = \frac{1.8412c}{2\pi a\sqrt{\epsilon_r}} \tag{4.16}$$

Note that (4.16) is based on the assumption of no fringing fields. Realistically, fringing fields will make the effective radius of the circular patch to be slightly larger than the nominal value of a. The effective radius a_e given as [7, 21]:

$$a_e = a\left(1 + \frac{2h}{\pi a\epsilon_r}\left(\ln\left(\frac{\pi a}{2h}\right) + 1.7726\right)\right)^{1/2} \tag{4.17}$$

The nominal values of a for calculating the resonant frequencies in (4.15) should then be replaced with the effective radius a_e. For example, the resonant frequency for the dominant TM^z_{110} mode becomes:

$$f_{110} = \frac{1.8412}{2\pi a_e\sqrt{\mu\epsilon}} = \frac{1.8412c}{2\pi a_e\sqrt{\epsilon_r}} \tag{4.18}$$

4.3 Shorted-Annular-Ring Patch

A shorted-annular-ring (SAR) patch antenna is essentially a circular patch with a concentric ring that shorts the top patch to the ground plane, shown in Fig. 4.3; the lowest order TM_0 surface wave mode can only be mitigated by such SAR patch antenna design [7]. Both the annular ring and patch radii, which are b and a respectively, of the antenna are adjusted to provide suppression of TM_0 surface wave and lateral wave propagation while operating at the desired frequency [36, 37].

4.3.1 Conditions for Surface Wave Suppression

From (4.5), we can determine the general form of the equivalent surface magnetic current density $\mathbf{M}_s$ at the patch perimeter ($\rho = a$) for TM^z_{110} mode:

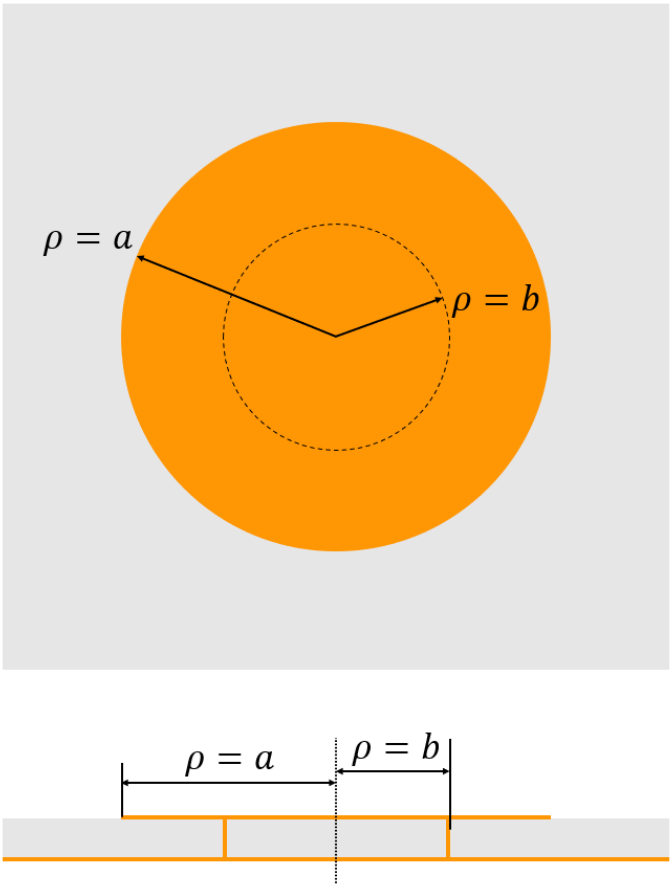


Figure 4.3: Top and side view of a representative shorted-annular-ring patch. Note the concentric ring boundary that shorts the top patch to the ground plane.

$$\begin{aligned}\mathbf{M}_s(\phi) &= -\hat{n} \times \vec{E}\Big|_{\rho=a} = -\hat{\rho} \times \hat{z} E_z(\rho = a, \phi) = \hat{\phi} E_z(\rho = a, \phi) \\ &\propto \hat{\phi} \cos\phi\end{aligned} \tag{4.19}$$

For thin substrates, the ρ and ϕ components of the electric field is zero for reasons explained in Section 4.2.1. Note that we choose the variation of $\mathbf{M}_s$ with respect to ϕ as $\cos\phi$ for simplicity; realistically, it will depend on the precise feed location with respect to the coordinate system and can be a combination of $\cos\phi$ and $\sin\phi$. This sheet of equivalent magnetic currents is depicted in Fig. 4.4.

Using superposition principle, the radiation of this sheet of magnetic current can be constructed from the radiations of single magnetic current loops $K(\phi)$ at various heights z_0

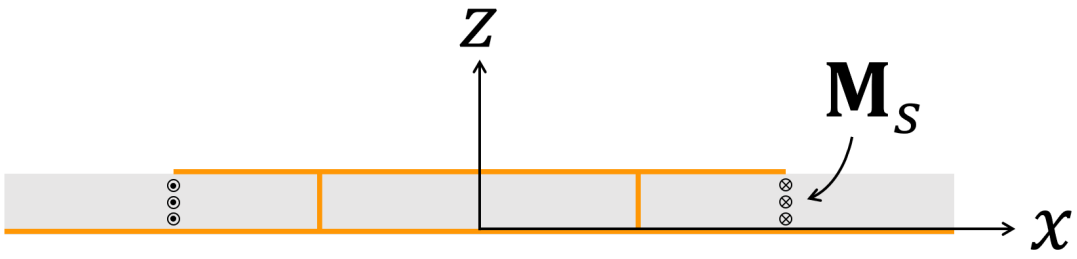


Figure 4.4: Magnetic ring current sheet $\mathbf{M}_s$ at the perimeter ($\rho = a$) of the patch.

above the ground plane, where the z-direction is shown in Fig. 4.4. Now consider a single Hertzian dipole of magnetic current, oriented in the x-direction, parallel to and at a height z_0 above the ground plane; the dipole will excite TM_0 mode surface waves [36]. Let ψ denotes any particular component of this surface-wave field, the general form of the field is given as [36]:

$$\psi = A\left(z, z_0\right) H_1^{(2)}\left(k_{\mathrm{TM}_0}\rho\right) \sin\phi \tag{4.20}$$

where $H_1^{(2)}$ is the Hankel function of the second kind of the first order representing outward traveling waves in cylindrical coordinates, k_{TM_0} is the propagation constant of the TM_0 surface wave in the substrate, and $A\left(z, z_0\right)$ is the amplitude factor as a function of the height of the source z_0 and the height of the observation point z. The $\sin\phi$ factor comes from the radiation pattern of a Hertzian dipole oriented in the x-direction; this factor will change according to the orientation of the dipole. For the single magnetic current loop $K\left(\phi\right)$, its surface-wave radiation can be constructed from that of the Hertzian dipole integrating over the ring current. The procedure will be detailed in Appendix A. The resulting surface wave field of the single magnetic current loop is found to be [36]:

$$\psi \propto H_1^{(2)}\left(k_{\mathrm{TM}_0}\rho\right) \cos\phi J_1'\left(k_{\mathrm{TM}_0}a\right) \tag{4.21}$$

where J_1' is the derivative of the Bessel function of the first kind of the first order.

Equation (4.21) is considered to be the fundamental design equation. Essentially, it

indicates that for the magnetic current loop to not excite TM_0 mode surface waves ($\psi = 0$), the radius a of the current loop, and thus of the patch, should be such that:

$$J_1'\left(k_{\mathrm{TM}_0} a\right) = 0 \tag{4.22}$$

which yields:

$$k_{\mathrm{TM}_0} a = \chi_{1n}', \qquad n = 1, 2, 3, ... \tag{4.23}$$

For the smallest possible patch size, we choose $n = 1$. For thin substrates ($h \ll \lambda_0$), another feature of a magnetic current loop satisfying (4.22) is that it will also exhibit a reduced space-wave radiation along the air-substrate interface. Space-waves along the interface are dominated by the radially-propagating lateral waves with free-space wave number k_0. For thin substrate, it can be shown that $k_0 \approx k_{\mathrm{TM}_0}$. Therefore, (4.22) is also a condition for suppressing lateral wave propagation [36].

Thus, the smallest patch radius a, such that the TM_0 surface wave mode and lateral wave propagation are suppressed, satisfies the following condition:

$$k_0 a \approx k_{\mathrm{TM}_0} a = \chi_{11}' = 1.8412 \tag{4.24}$$

or:

$$a \approx \frac{\chi_{11}'}{k_0} = \frac{1.8412}{2\pi}\lambda_0 \tag{4.25}$$

Note that (4.24) and (4.25) indicate that the patch radius a for maximum surface wave suppression is nearly independent of the substrate permittivity for thin substrates.

4.3.2 Single-Band SAR Patch Design

A circular patch with radius a chosen according to (4.25) will suppress surface wave and lateral wave propagation. However, the patch at this size is generally not resonant at the

desired operating frequency [36]. This can be shown by examining (4.15) from the cavity model; a circular patch resonating at TM^z_{110} mode will have a radius of:

$$a = \frac{\chi'_{11}}{k} \tag{4.26}$$

where k is the wave number in the substrate at the desired operating frequency. However, surface wave suppression requires that the radius of the patch should also satisfy (4.25), and therefore we have:

$$k_{TM_0} \approx k_0 < k \tag{4.27}$$

This implies that a conventional circular patch cannot simultaneously satisfy both (4.25) and (4.26), and therefore cannot be resonating at a desired frequency while exhibiting surface wave suppression. Equation (4.27), together with (4.25) and (4.26), indicates that a circular patch that suppresses the TM_0 mode surface waves will be larger than the desired resonant patch size, and resonating at a lower frequency as a result. In order for the patch to resonant at the desired frequency without changing the patch radius a, somehow the resonant frequency has to be raised.

A SAR patch antenna exhibits a concentric short-circuit boundary at radius b as shown in Fig. 4.3; such shorting ring can accomplish the task of decreasing the effective patch radius so that the antenna can operate at the desired frequency, while maintaining the actual patch size of radius a which provides suppression of TM_0 mode surface waves. The radius b of the shorting ring is adjusted to achieve the desired resonant frequency. Due to the short-circuit, the region of $\rho < b$ is not important, and thus the metal within this region can be removed, and hence the name "annular-ring" [36]. The shorting boundary can be realized by either an actual metal cylinder or a wall of vias, which makes the design practical to conventional PCB manufacturing [7, 36, 37].

To obtain the proper short-circuit boundary radius b for the resonant condition, we must derive the field expression for a SAR patch. The short-circuit annular ring established additional boundary conditions at $\rho = b$:

$$\begin{cases} E_z\left(\rho = b\right) = 0 \\ E_\phi\left(\rho = b\right) = 0 \end{cases} \tag{4.28}$$

Note that these additional boundary conditions are added to the existing boundary condition of (4.6). Again, for thin substrates, the E_ϕ component of the electric field is zero for reasons explained in Section 4.2.1. For the original vector potential expression in (4.9), we determined that B_1 has to be zero, since the field must be finite at $\rho = 0$ for an typical circular cavity. For annular ring patch, however, we are only interested in the region of $b < \rho < a$, and therefore, B_1 does not need to be zero. Using (4.5) and (4.9), the z-component of the electric field for the TM^z_{110} mode can then be expressed as:

$$E_z \propto \left(A_1 J_1\left(k\rho\right) + B_1 Y_1\left(k\rho\right)\right)\cos\phi \tag{4.29}$$

The variation with respect to ϕ is again chosen to be $\cos\phi$ for simplicity. Enforcing the boundary conditions of (4.6) and (4.28) at $\rho = a$ and $\rho = b$, we can obtain the following relationship:

$$\frac{J_1(kb)}{Y_1(kb)} = \frac{J_1'(ka)}{Y_1'(ka)} \tag{4.30}$$

The transcendental equation of (4.30) is the design equation for determining the radius b of the shorting ring at the desired resonant frequency.

4.4 Dual-Band SAR Antennas

Modern GPS receivers typically operate in both L1 and L2 bands. Typical dual-band patch antenna designs involve stacked or parasitically coupled patches. However, these designs may not be easily realizable due to the existence of the shorting boundaries. As the resonant frequency is very sensitive to the radius of the shorting boundary, the construction or manufacturing quality and precision of such boundary is especially crucial. Nonetheless, the shorting boundary can be realized using a ring of vias, this increases the feasibility of

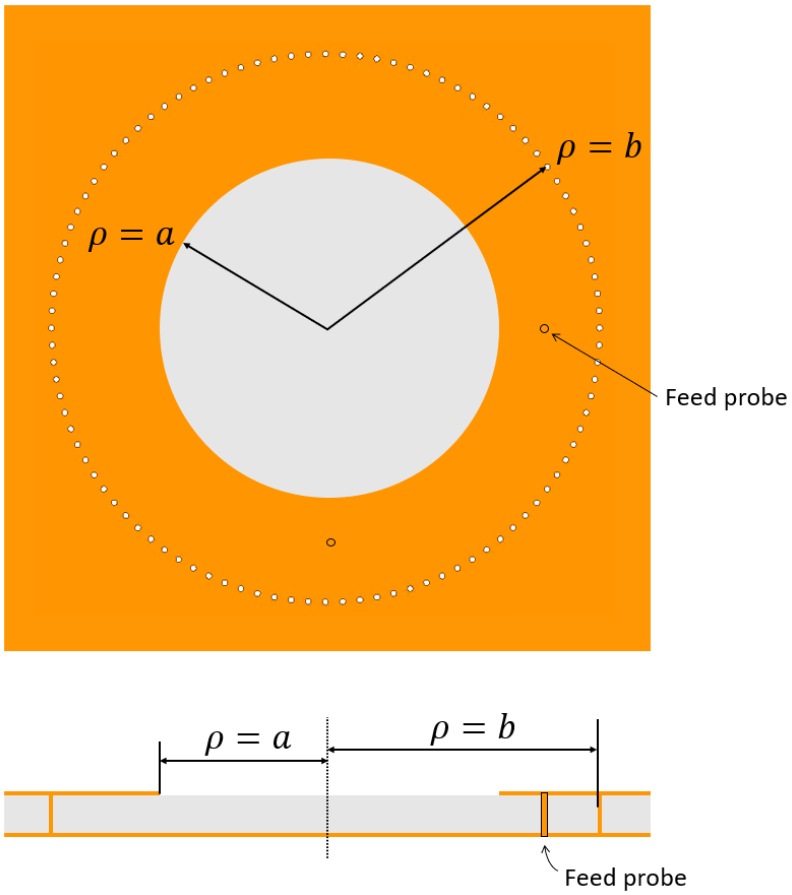


Figure 4.5: The inverted shorted-annular-ring (ISAR) patch antenna configuration. Note that the shorting boundary can be constructed with a ring of vias. The two feed probes are excited with 90° relative phase to generate circular polarization.

these designs using conventional, low-cost PCB manufacturing techniques for the SAR patch design.

One clever method of realizing a dual-band SAR design recognizes that a resonant circular cavity can also be formed by placing the shorting boundary outside the magnetic current loop [38]. Such configuration is shown in Fig. 4.5. This configuration is known as the inverted shorted-annular-ring (ISAR) patch antenna, and can be considered as a complementary structure to the SAR patch. The design also obeys the transcendental equation (4.30); the only difference is that the ISAR design has $b > a$ as opposed to $b < a$ for the SAR design. It is worth noting here that the ISAR design will intrinsically not excite surface waves due to

the existence of the top layer metal that extends to the edge of the substrate and the shorting boundary positioned outside the magnetic current loop [38]. Nonetheless, the radius a of the magnetic current loop is still chosen using (4.25) to reduce lateral wave propagation.

A Dual-band GPS antenna can then be constructed with a smaller L1-band SAR patch placed concentrically within a larger L2-band ISAR patch [38]; such antenna is shown in Fig. 4.6.

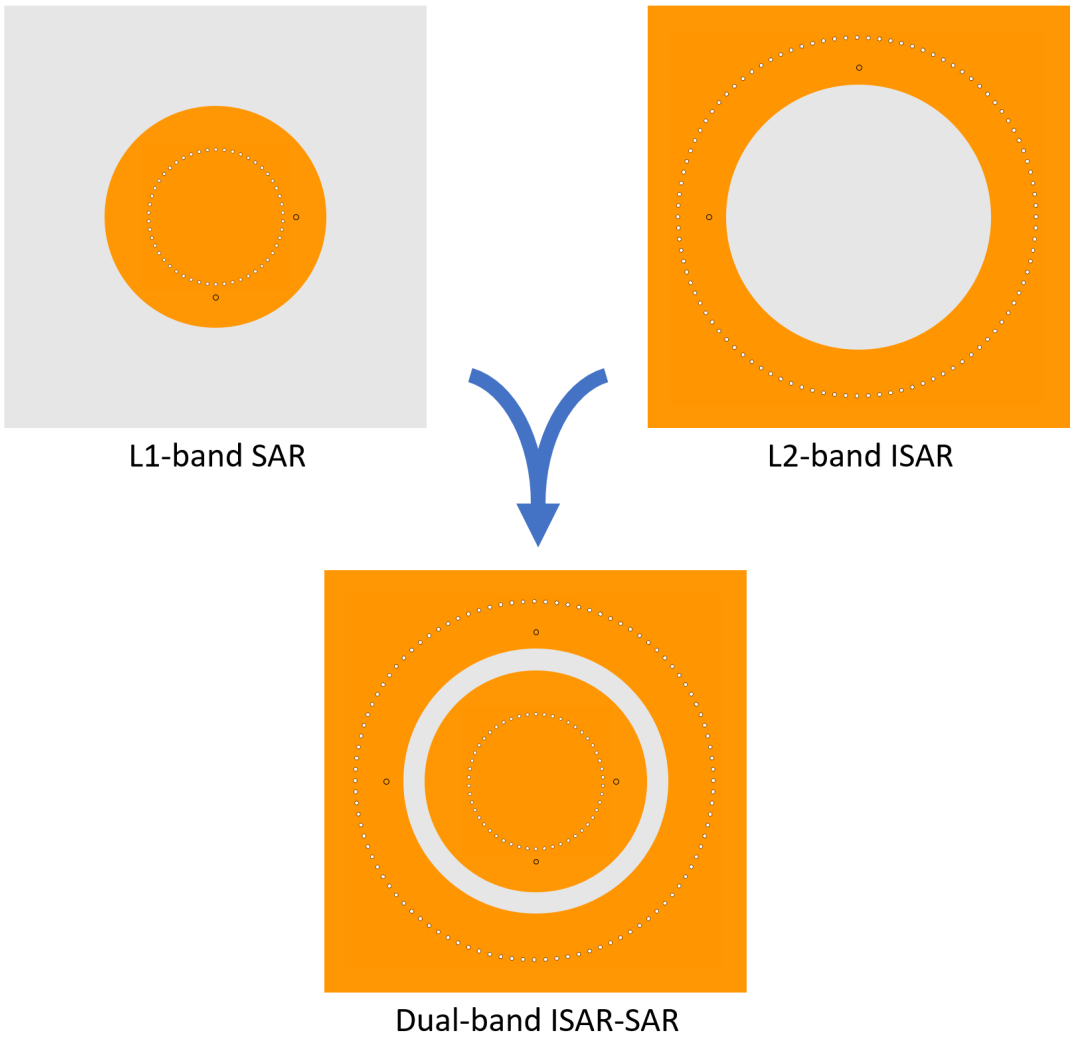


Figure 4.6: A dual-band configuration (top view) can be constructed with a SAR patch and an ISAR patch.

CHAPTER 5

Simulation and Performance Verification

5.1 Single L1-Band SAR Patch: Design Investigation and Performance Verification

The design of GPS SAR patch antenna starts with validating the performance of a single-band circularly-polarized SAR antenna. The measurement of a fabricated prototype was carried out in the spherical near-field antenna measurement chamber at the Center for High Frequency Electronics (CHFE) of UCLA; the measurement accuracy at L1- and L2-bands using the open-ended waveguide probe is discussed in Appendix D.

5.1.1 Substrate Considerations

Substrate selection is one of the first crucial steps in microstrip patch antenna design. As detailed in Chapter 3, substrate characteristics such as permittivity, thickness, loss tangent can profoundly affect antenna performance. One important note here is that typically in patch antenna design, high-permittivity substrates, such as ceramics substrates, allow the miniaturization of the patches; the size reduction can be desirable for some applications such as patch antenna arrays. However, as described in Chapter 4, the size of a SAR patch has a well-defined radius such that it will result in reduced surface waves, and this is almost independent of the substrate used. Therefore, high permittivity substrates will likely not offer much benefits for SAR patches. Several common substrates are simulated using HFSS during design investigation for the L1-band SAR patch. The SAR antenna performance of four of the substrates studied are shown in Table 5.1.

Table 5.1: Characteristics of the substrates used in simulation and antenna performance metrics achieved using the substrates. All performance results are simulated at the L1 frequency (1.575 GHz).

Substrate Material		FR4	RT/duroid 6002	RO4003C	Taconic RF-41
Material Parameters	ϵ_r	4.4	2.94	3.55	4.1
	$\tan\delta$	0.02	0.0012	0.0027	0.0038
	Thickness (mm)	1.524	1.524	1.524	3.175
Antenna Performance	Directivity (dBic)	9.1	9.4	9.4	9.3
	Gain (dBic)	3.5	8.8	8.2	8.5
	Bandwidth (MHz)	≈30	≈12	≈11	≈21

Table 5.2: Inquired prices, minus the shipping cost, for one PCB unit of size 150 mm × 150 mm from two different PCB vendors. *Layer refers to the number of conductive (e.g. copper) layers, not substrate layers, of one PCB unit. N/A indicates that the substrate is either unavailable for the vendor at the time of inquiry, or that the number of conductive layers is not supported for that particular substrate by the vendor.

Substrate Material		FR4	RT/duroid 6002	RO4003C	Taconic RF-41
Vendor 1: PCBWay	2 Layers*	≈$10	N/A	≈$100	N/A
	4 Layers	≈$20	N/A	N/A	N/A
Vendor 2: Fast Turn PCB	3 Layers	≈$200	≈$2000	N/A	N/A

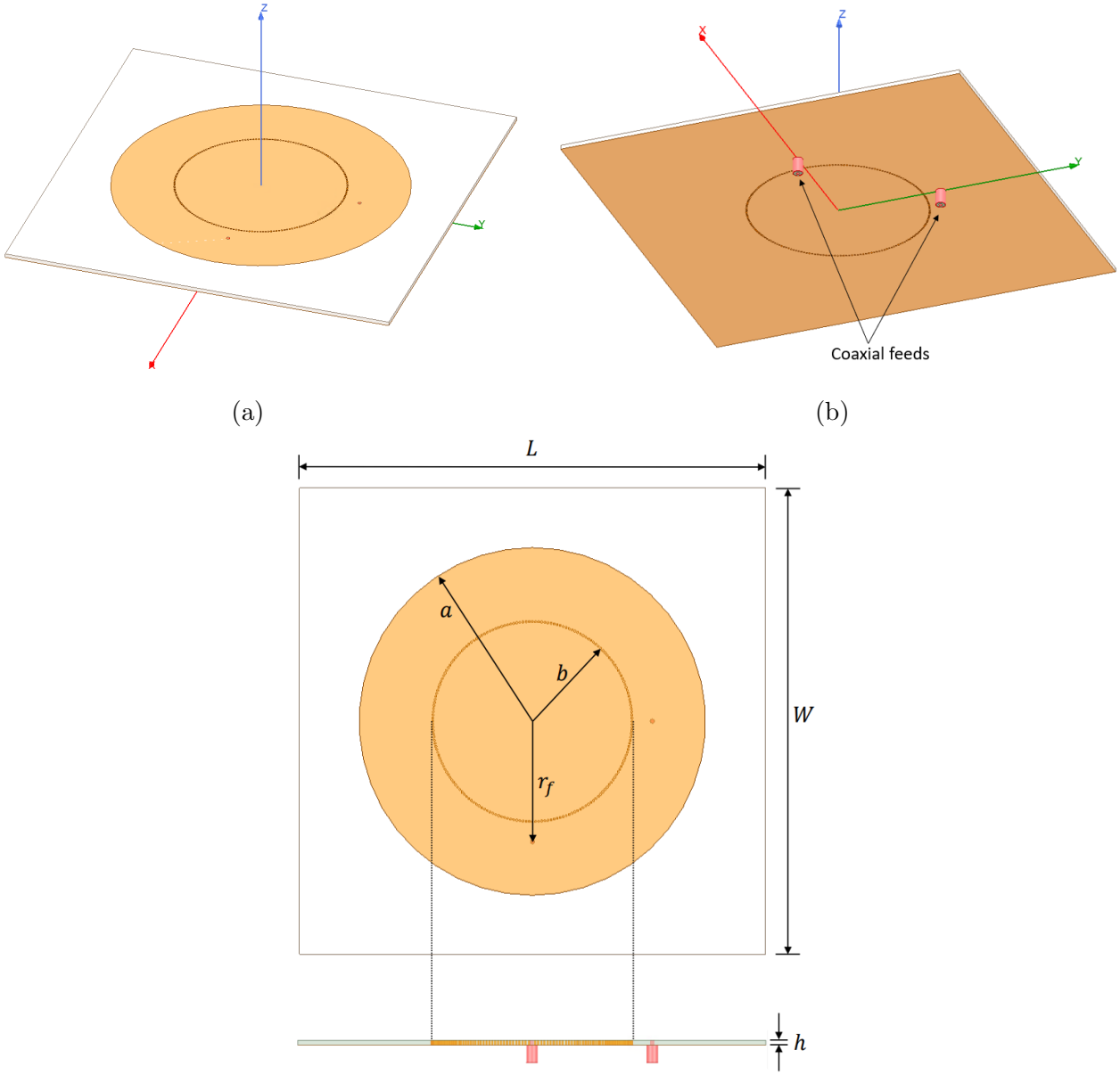
For each substrate that yields the results in Table 5.1, some of the antenna parameters are tuned such that the antenna is resonant at the L1 frequency for that substrate.

During the simulation and price inquiry from vendors, the size of the antenna substrate is 150 mm $\times$ 150 mm, which adheres to the design constraints. The inquired prices listed in Table 5.2 are for substrate with thickness of around 1.5 mm per layer of substrate, and thicker substrates are proportionally more expensive. We show approximated prices in Table 5.2 because different substrates and vendors have different offers of standard thickness. For example, Fast Turn PCB is able to work with a FR4 substrate with thickness of 1.524 mm, whereas PCBWay can only work with 1.6 mm, among other standard thickness it offers. As shown in Table 5.2, we also inquired the PCB prices for different number of metal layers. This is for price comparison and in consideration of the potentially multi-layered design accounting for the feeding networks. It is worth pointing out that two conductive layers correspond to one substrate layer, and three conductive layers correspond to two substrate layers, and so on; each substrate layer is around 1.5 mm as mentioned above.

Since one of our design constraint of for the GPS antenna is low cost, this will factor into substrate and vendor considerations. FR4 is a very common PCB substrate while being comparatively low cost, as shown in Table 5.2. However, by examining Table 5.1, FR4 can be eliminated as a candidate since its relatively high loss tangent results in a large loss of efficiency at L1, which can be seen in the reduction of 5.6 dB from the directivity to gain. The two substrates from Rogers Corporation, RT/duroid 6002 and RO4003C, results in similar antenna performance. Antenna that uses the Taconic RF-41 substrate also shows comparable directivity and gain; since the RF-41 substrate has an available thickness of almost twice of those of the two Rogers substrates, the bandwidth is almost doubled of the Rogers ones. Unfortunately, the RF-41 substrate is not available for the two vendors considered. For the remaining two Rogers substrates, RO4003C is offered at a much lower price at at one vendor, and hence will be selected for our manufactured prototype for performance assessment.

5.1.2 Simulated Performance

A preliminary L1-band SAR patch antenna design using the RO4003C substrate is simulated using HFSS, shown in Fig. 5.1.



(a) (b)

(c)

Figure 5.1: The top view (a) and bottom view (b) of the simulated L1-band SAR patch antenna model in HFSS. (c) Top and side views of the SAR patch with labeled design parameters, whose values are shown in Table 5.3.

Table 5.3: Nominal design values used in the HFSS simulation for the antenna dimensions shown in Fig. 5.1c.

Design parameters	Nominal values (mm)
Substrate length (L)	150
Substrate width (W)	150
Substrate thickness (h)	1.524
Patch radius (a)	55.8
Shorting boundary radius (b)	32.1025
Feed location (r_f)	38.6

The patch radius a is first calculated using (4.25) at L1 frequency of 1.575 GHz, which yields $a = 55.8$ mm. The shorting boundary radius b can then be determined by solving (4.30), which produces a preliminary value of around $b = 31.91$ mm for the given patch radius of 55.8 mm at L1. The value of b is then further tuned during simulation such that the resonant is almost exactly at the L1 frequency. The two coaxial feed points, with 90° relative phase to achieve circular polarization, are located at r_f, which can also be adjusted to obtain a good matching. The final nominal values for the design parameters are listed in Table 5.3. The shorting boundary is realized by a ring of 200 vias, each with a diameter of 25 mils, or 0.635 mm, and they are closely spaced with each other. The discrete spacing between the vias will produce a reactive surface impedance, and not an exact short-circuit, at the desired shorting boundary. According to [38], this effect can be assumed to be negligible for closely-spaced vias, though no definitive rule is addressed on how many vias is considered sufficient. Actual HFSS simulations using different number of vias, ranging from 20 to 200, produce no significant differences in antenna performance such as radiation patterns, return loss, and gain, etc.

The simulation results of the L1 SAR patch antenna is shown in Fig. 5.2–5.4. From Fig. 5.2, the bandwidth ($S_{11} < -10$ dB) is around 11.4 MHz. Though the SAR patch antenna is inherently very narrowband as with many conventional patch antennas, such bandwidth is more than sufficient to receive the civilian L1 C/A and L1C signals, as shown in Table 2.2.

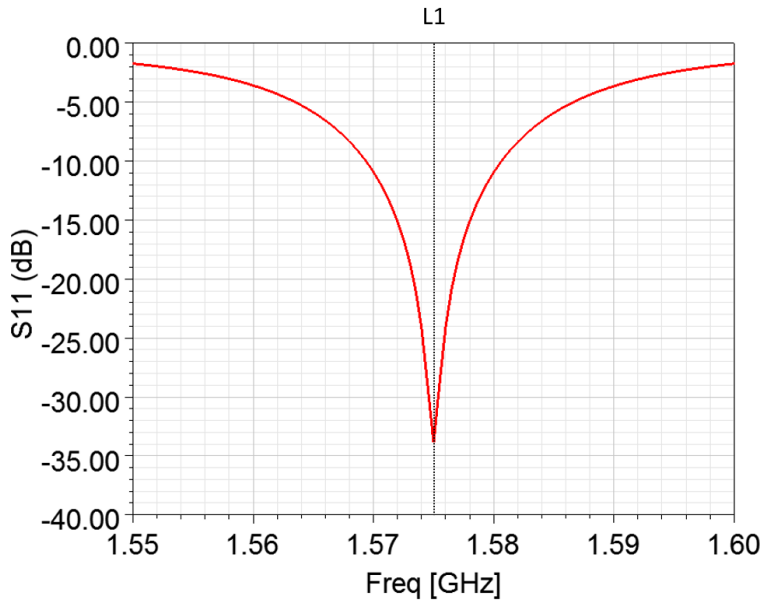
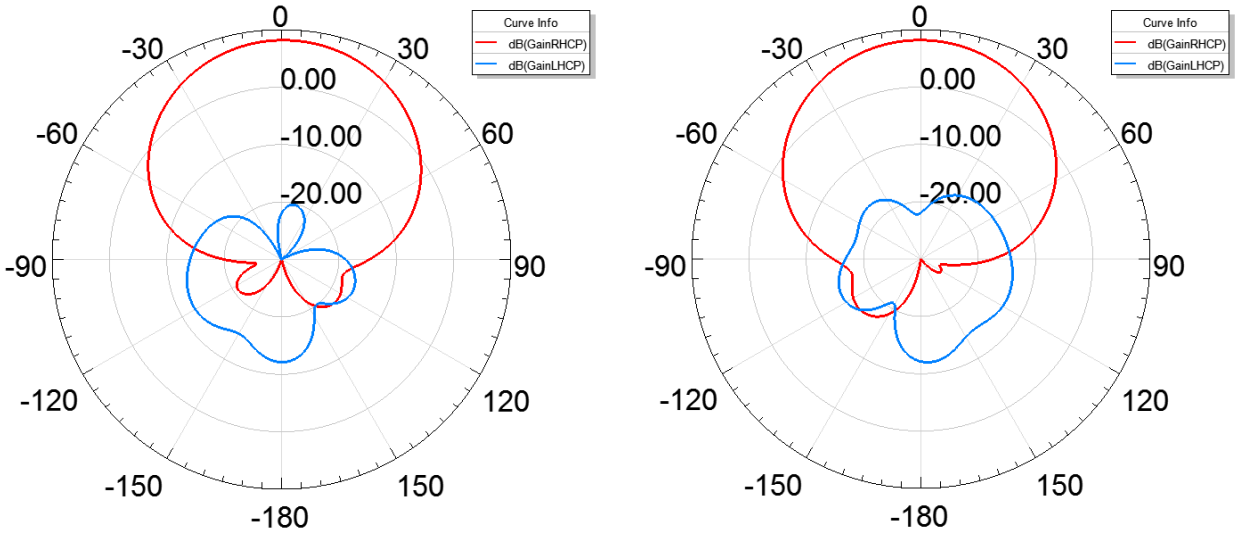


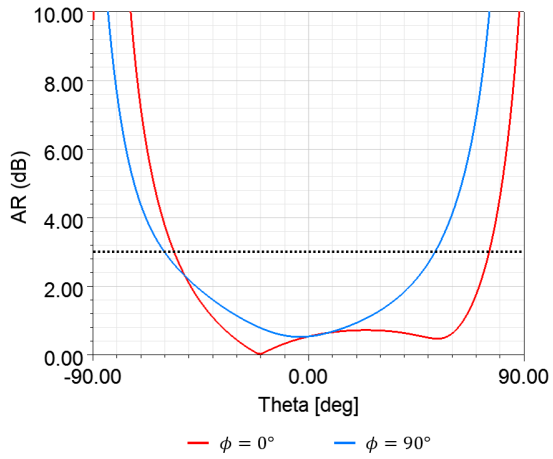
Figure 5.2: Simulated S_{11} of the L1 SAR Patch Antenna. The results for both ports are near-identical.



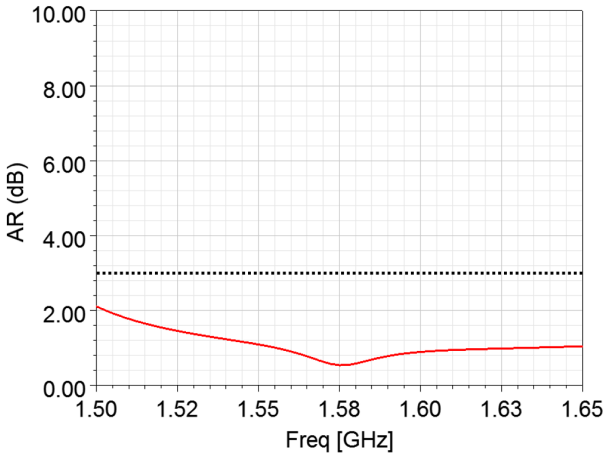
(a) (b)

Figure 5.3: (a) $\phi = 0°$-cut gain pattern in dB. (b) $\phi = 90°$-cut gain pattern in dB.

However, wider-bandwidth designs should also be considered if one also wants to receive the military L1 P(Y)- or M-codes that has ten times or more the bandwidth of the C/A-codes. Another practical reason for investigating higher-bandwidth designs or bandwidth enhance-



(a)



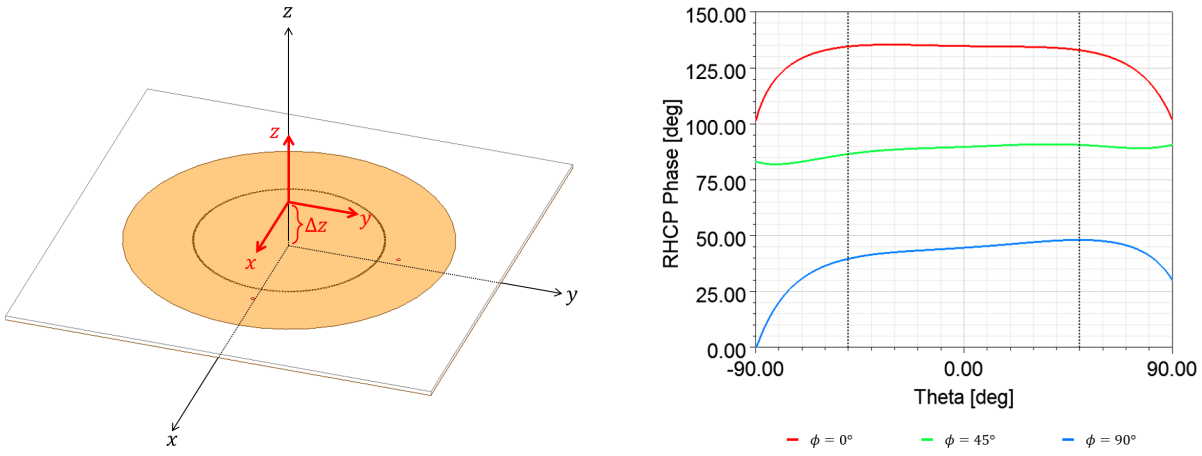
(b)

Figure 5.4: (a) Axial ratio in dB versus polar angle θ for ϕ-plane cuts. (b) Axial ratio in dB in the zenith direction ($\theta = 0°$) versus frequency. The AR = 3 dB level is shown as black dotted line.

ment techniques is to consider practical issues such as manufacturing/material tolerances, which can shift the resonant frequencies; these will be addressed later sections.

The gain patterns shown in Fig. 5.3 shows several desired features such as smooth broadside radiation pattern, low backlobes with a D/U ratio of more than 20 dB for angles close to the zenith, high XPD value of more than 20 dB to almost 30 dB near broadside, and low gain values (or large gain roll-off) near the horizon. The peak directivity and gain are 9.4 dBic and 8.2 dBic respectively, and a wide beamwidth that satisfies the GPS minimum-gain requirements; these are more than sufficient for typical GPS applications. The radiation pattern of a SAR patch antenna is usually more directive than a typical GPS antenna, but this should not be a concern as the recent trend for mask angle if to shift to higher elevations. The AR results shown in Fig. 5.4 shows an acceptable CP (AR $<$ 3 dB) for more than 100° in beamwidth, and good AR levels across frequencies.

As discussed in Section 3.5, antenna phase center stability is important in high-precision GPS applications. The phase center can be found by first varying the position of a local coordinate system, then identifying the point at which the radiated RHCP far-field, defined in the local coordinate, will exhibit the least amount of phase variation. Since the SAR



(a) (b)

Figure 5.5: (a) Determining the phase center of the SAR patch antenna. The global coordinate system is shown in black, and the varying local coordinate system is shown in red, offset by Δz from the global coordinate origin. (b) Simulated RHCP radiated far-field phase versus the polar angle θ for the L1-band SAR patch when the origin of the coordinate system is at the phase center of the antenna. The black dotted lines denote $\theta = 50°$.

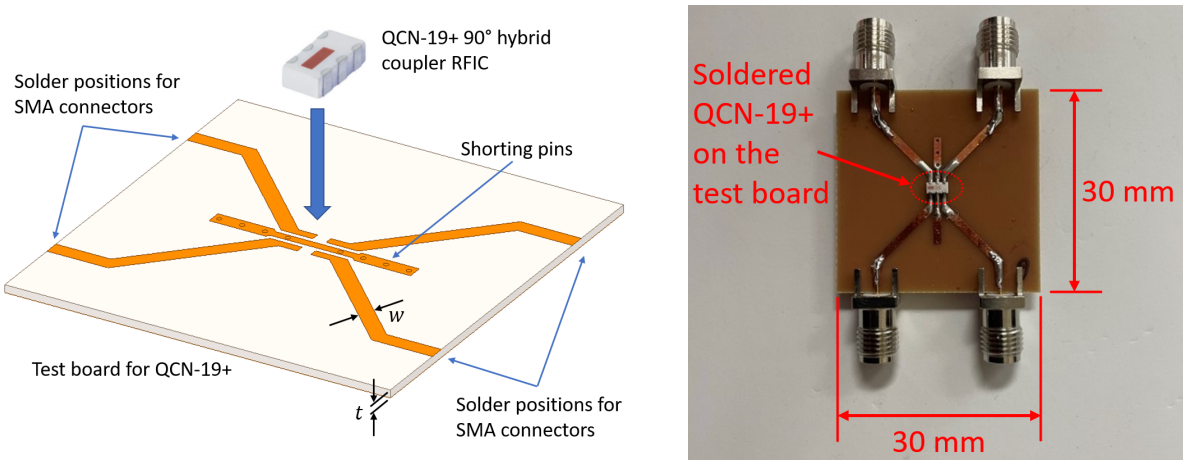
patch is circularly symmetric around the z-axis (though not perfectly symmetric due to the dual-feeds), it is reasonable to assume that the phase center will be on, if not close to, the z-axis.

We will then only vary the local coordinate along the z-axis by Δz to find the phase center, where Δz is the z-direction offset of the local coordinate system's origin from the origin of the global coordinate system as shown in Fig. 5.5a. The origin of the global coordinate is defined at the center of the ground plane of the SAR antenna, as shown in Fig. 5.1a and 5.1b. When $\Delta z = 0$, the local and global coordinate systems coincide and are identical. HFSS simulations found that the phase center is around $\Delta z = 15$–16 mm for $\phi = 0°$, $45°$, and $90°$ plane cuts. The co-pol (RHCP for our case) radiated far-field phase the three ϕ-plane cuts are plotted in Fig. 5.5b. The absolute values of the phase does not matter. It is clear that within the $100°$ beamwidth (the beamwidth at $\theta = 50°$), the RHCP far-field phases for the three plane cuts are spaced by roughly $45°$ apart as expected, and their variation within this beamwidth is consistently within $10°$. At L1, a $10°$ phase variation

translates to a variation of about 5 mm in length measurement; such precision is more than enough for our application.

5.1.3 Fabrication and Measurements: Feed Networks

A low-cost feed network for testing purposes was constructed for the L1 SAR patch antenna prototype measurements. The two feed points of the antenna have 90° relative phase to realize RHCP, which require a 90°-phase hybrid coupler. The coupler network can be designed using typical microstrip lines, however, an alternative is to use a low-cost commercial off-the-shelf component. One promising candidate is QCN-19+, an ultra-small RFIC manufactured by Mini-Circuits. QCN-19+ is a two-way quadrature-phase power combiner/splitter; some of the important parameters of QCN-19+ is shown in Table 5.4. QCN-19+ has an operating frequency range of 1.1–1.925 GHz, which covers both L1 and L2 bands. The RFIC has a maximum dimension is around 3.2 mm, which is small enough to be incorporated in a microstrip feeding network. QCN-19+ has a typical phase unbalance of around 2° across its operating frequency, and a maximum of 4°. The L1 SAR patch antenna is simulated with the two feeds having relative phase of 86° and 94° to account for the phase unbalance; the



(a) (b)

Figure 5.6: (a) Test board and QCN-19+ on-chip quadrature hybrid. (b) Assembled test board with soldered QCN-19+ in the center and SMA connectors

Table 5.4: Some relevant parameters of the QCN-19+ quadrature-phase hybrid coupler.

Price ($ per unit)	4.89
Dimensions (mm)	$3.2 \times 1.6 \times 0.89$
Frequency Range (GHz)	1.1–1.925
Input Power Rating (W)	15
Isolation (dB)	25
Phase Unbalance (°)	$< \pm 4$

Table 5.5: Relevant information regarding the QCN-19+ test board.

Manufacturer/vendor	PCBWay
Substrate	FR4
Price ($ per piece)	$\approx$1
Size (mm)	30×30
Thickness t (mm)	0.6
Transmission line width W (mm)	1.147

result shows that a 4° phase unbalance will not significantly impact the RHCP performance such as gain and AR.

For the potential finalized design in the future, the antenna will consists of multiple layers of substrate, one of which will be used for the feeding networks. However, for the initial testing purposes, a separate test board will be constructed for the 90° hybrid coupler chip, which allows its performance to be measured and assessed independently from the antenna.

As shown in Fig. 5.6a, the test board consists of four 50-Ω transmission lines that will be soldered to four SMA connectors, and the on-chip coupler QCN-19+ will be soldered at the center of the board. QCN-19+ is a 6-pin RFIC; the two center pins are ground pins, which will be connected to the ground plane via the shorting pins, shown in Fig. 5.6a. The other four pins of QCN-19+ are the coupler ports, which will be soldered to the transmission lines;

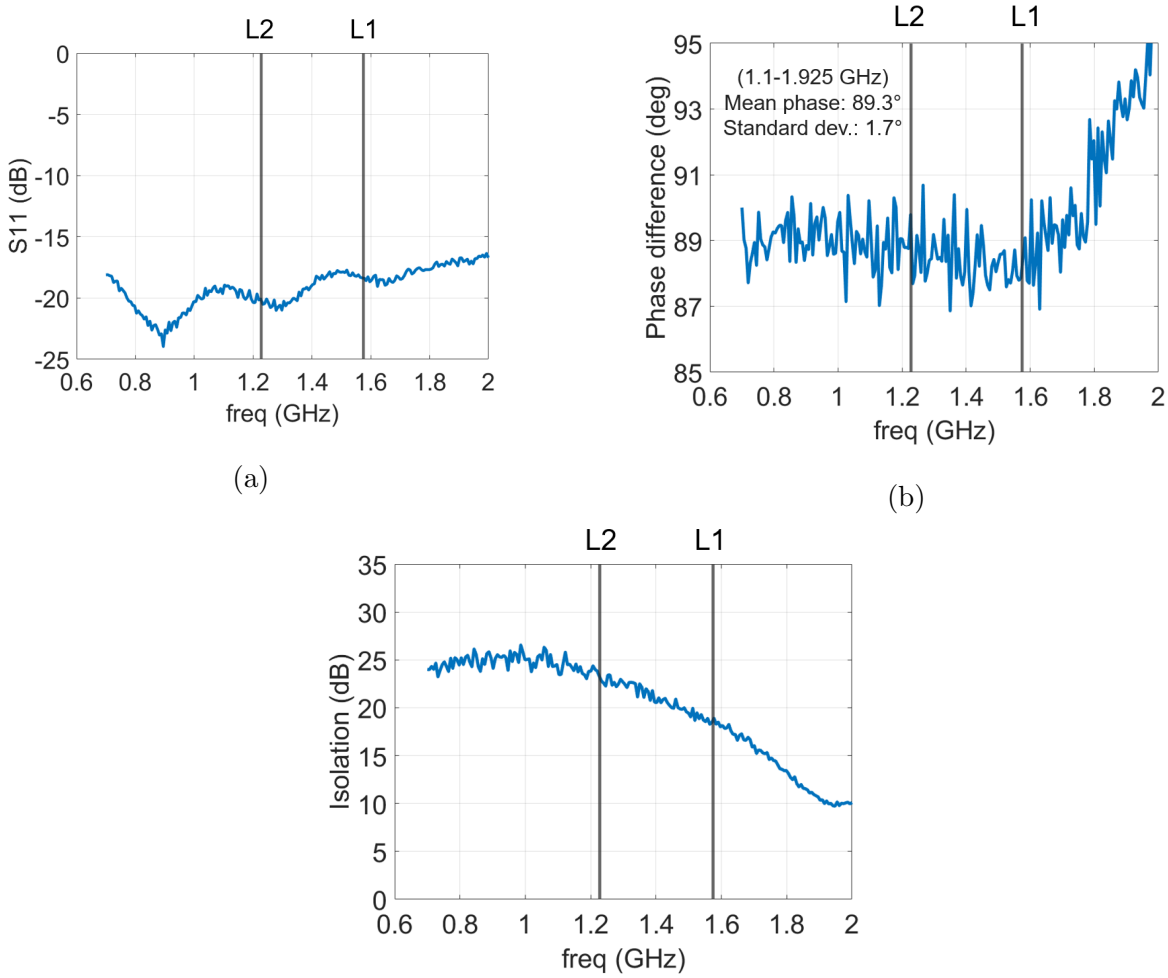


Figure 5.7: VNA measurements for (a) S_{11}, (b) phase unbalance, and (c) isolation of the QCN-19+ on the test board.

the transmission lines will then be soldered to SMA connectors at the edge of the board. The fully assembled test board with soldered SMA connectors and QCN-19+ is shown in Fig. 5.6b. The test board is manufactured on FR4 substrate through the very low-cost vendor PCBWay, which is also the SAR patch manufacturer; other relevant information of the test board are listed in Table 5.5.

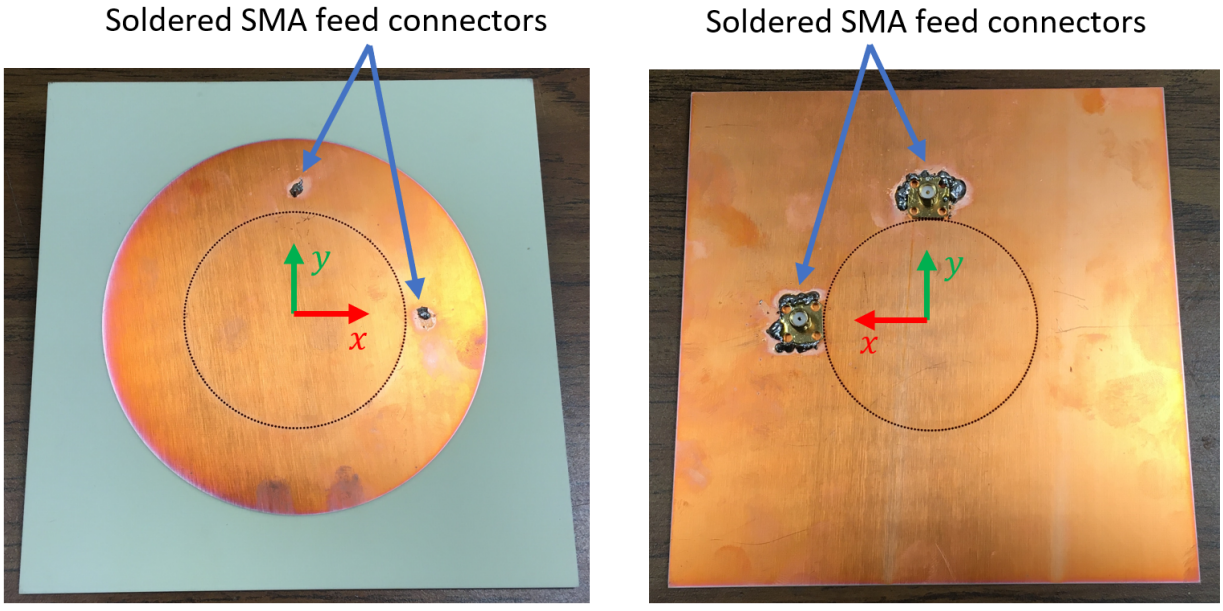
The performance of QCN-19+ is measured using vector network analyzer (VNA) connected to the assembled test board. The measured performance is shown in Fig. 5.7. From Fig. 5.7a, the S_{11} is consistently below -15 dB across the operating frequencies. Fig. 5.7b

shows a calculated mean value of 89.3° and a standard deviation of 1.7°, and the unbalance is almost consistently within ±2° across the operating frequencies. The isolation is measured between the two output (or input if functioned as a power combiner) ports, and it is around 20 dB at L1 and L2, which is acceptable for an extremely low-cost unit.

5.1.4 Fabrication and Measurements: Antenna S_{11}

The preliminary L1-band SAR patch antenna design shown in Section 5.1.2 is manufactured. Part of the purpose of the preliminary design is to assess the performance of the manufactured prototype through a low-cost vendor such as PCBWay.

The design uses a circular array of 200 vias as the required shorting boundary. As noted in Section above, although HFSS simulations for different number of vias produce no significant difference in the performance, the manufacturing cost for different number of vias is the same for the vendor PCBWay. Hence, the manufactured prototype will feature 200 vias. It is known that the resonant frequency of the SAR patch antenna is particularly



(a) (b)

Figure 5.8: (a) Top view and (b) bottom view of a prototyped L1-band SAR patch antenna manufacturerd by PCBWay.

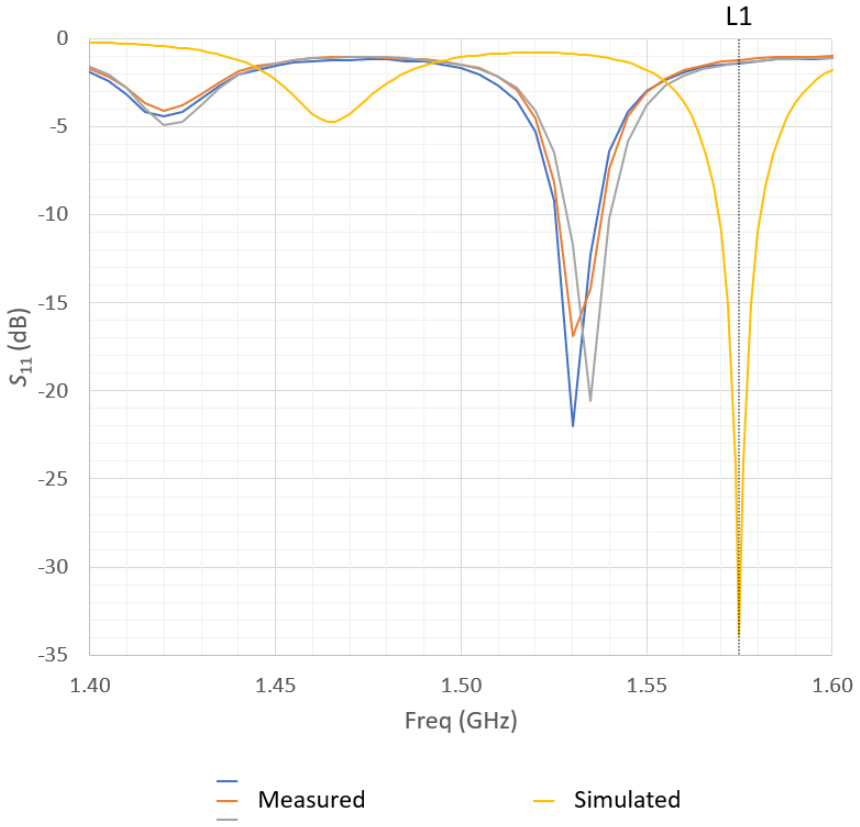


Figure 5.9: Comparison between measured and simulated S_{11} of the antenna feed ports. The three measured results are shown to assess the manufacturing consistency of the prototypes.

sensitive to the shorting boundary radius b, and since the design is also inherently narrow-band, a measured prototype will assess the viability of such narrow-band, sensitive design through a commercial low-cost vendor.

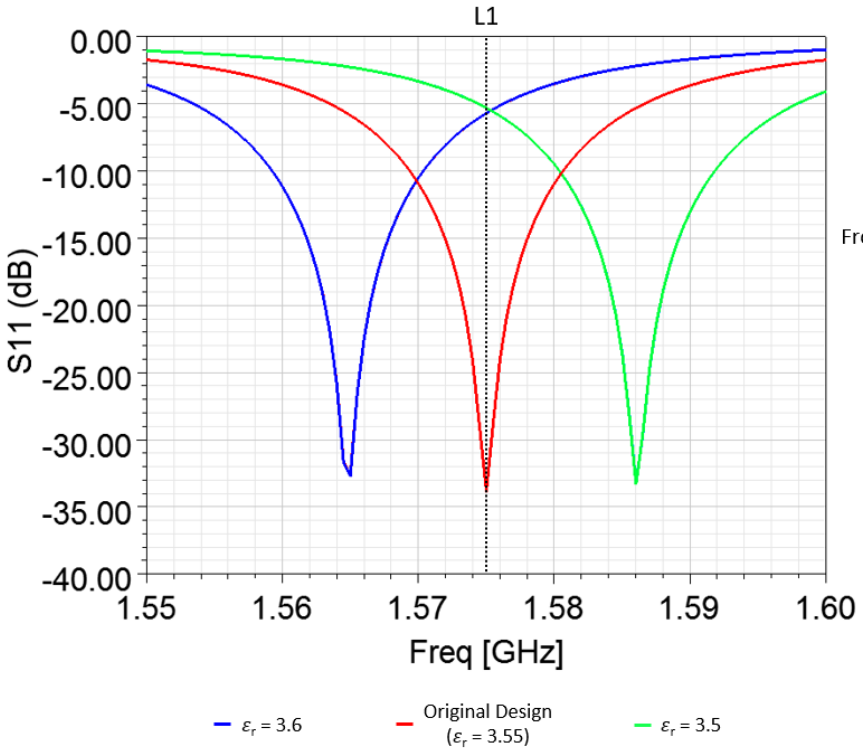
Five manufactured antenna units is ordered from PCBWay; one unit with soldered SMA feed connectors is shown in Fig. 5.8. The S_{11} of the antenna input ports are measured using a VNA, and the measured results are obtained at the input ports of several prototyped units. The simulated and three measured results are plotted in Fig. 5.9. Clearly, the measured S_{11} exhibits a resonant shift of around 40–45 MHz to a lower frequency of roughly 1.53 GHz compared to the simulated design, which is resonant at L1 (1.575 GHz). All measurements show a very similar frequency shift, which means the cause is likely to be systematic in nature. Both simulated and measured results show a bandwidth of around 11–12 MHz; though minor frequency shift between simulation and manufactured prototype is common in patch antenna

designs, the observed frequency shift is outside the bandwidth of the original design centered at L1. This further suggest that a wider-band design is desirable. A sensitivity and tolerance studies will be conducted in the following section.

5.1.5 Sensitivity Analysis and Tolerance Studies

There are several main parameters that can affect the resonant frequency of a SAR patch, which are: the permittivity of the substrate ϵ_r, substrate thickness h, patch radius a, and the shorting boundary position b, which is the radius of the ring of vias. There are certainly other potential factors, but these four have relatively more influence. The tolerance and sensitivity studies are done by varying each of the above parameters with the maximum nominal tolerances and observe the resulting resonant frequency shifts in the S_{11}, while other parameters remain the original nominal values in Table 5.3.

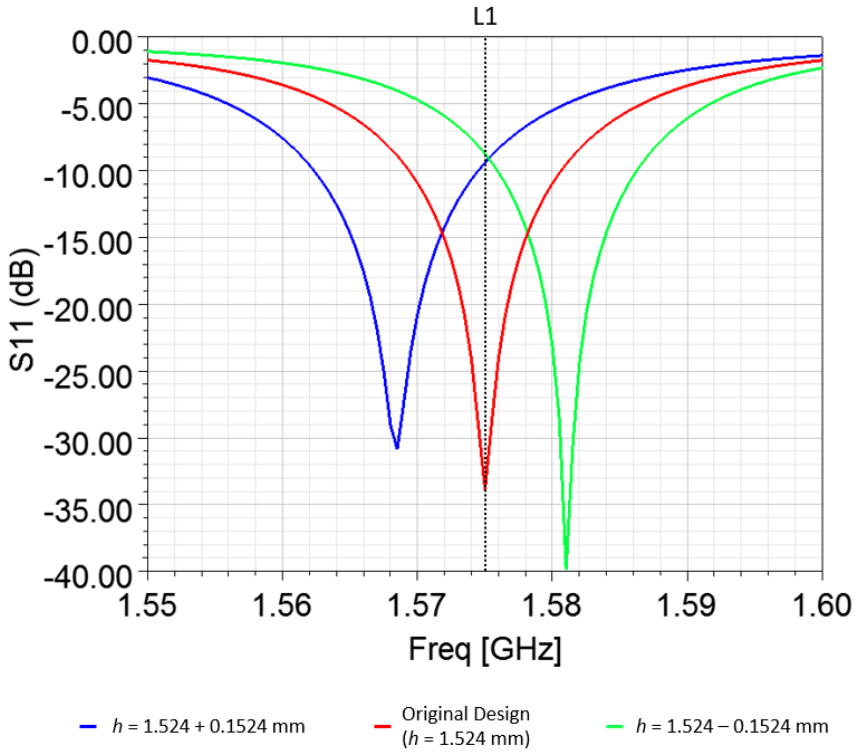
The tolerance and sensitivity study results are shown in Fig. 5.10. The variation in the parameters are based on the nominal tolerance values from the manufacturer or vendor of the material such as PCBWay and Rogers. Since the SAR patch is inherently narrowband with a bandwidth of around 11–12 MHz, the nominal tolerances can shift the resonant frequency outside the bandwidth centered at L1, as seen from the simulated results. A rough measurement from the manufactured patch shown in Fig. 5.8 suggests that the patch radius a and via ring radius b are manufactured fairly accurately. HFSS simulations also determined that for the resonant frequency to shift by around 40–45 MHz to the left of L1 as shown in Fig. 5.9, the via ring radius b will be roughly 0.7 mm smaller than the nominal value in Table 5.3. A 0.7 mm deviation seems unlikely given the rough measurements performed on the actual unit and the manufacturing tolerance specified by PCBWay. Similarly, the substrate height is also fairly accurate, and unlikely to contribute to the observed resonant shift. The "worst-case" scenario shown in Fig. 5.10e suggests that it is possible for the resonant frequency to shift by more than 30 MHz, based on the nominal tolerances given by the vendors. Realistically, however, the tolerances are unlikely to reach their maximum values, let alone all shifting the resonant frequency in one direction. Additionally, the actual



From Rogers datasheet:

Dielectric Constant, ε_r Process	3.38 ± 0.05
[2] Dielectric Constant, ε_r Design	3.55

(a)



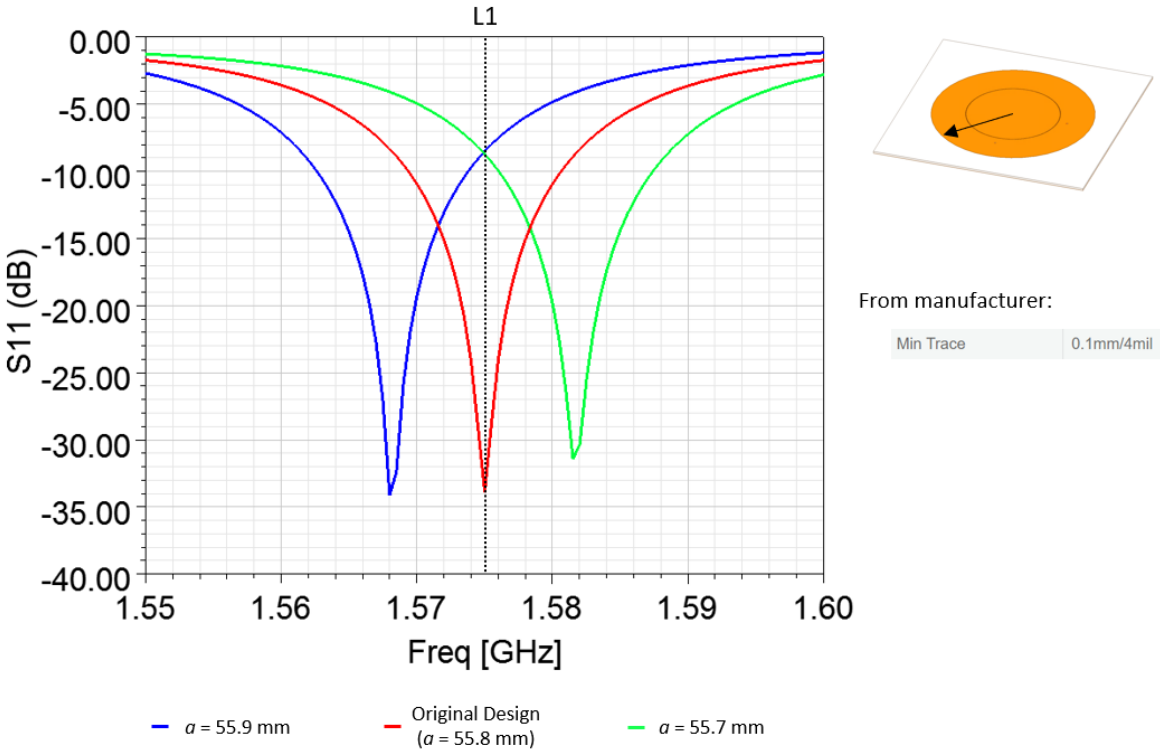
From Rogers:

*RO4003C™ RO4360™	0.008" (0.203mm) ± 0.001"
	0.012" (0.305mm) ± 0.001"
	0.016" (0.406mm) ± 0.0015"
	0.020" (0.508mm) ± 0.0015"
	0.032" (0.813mm) ± 0.002"
	0.060" (1.524mm) ± 0.004"

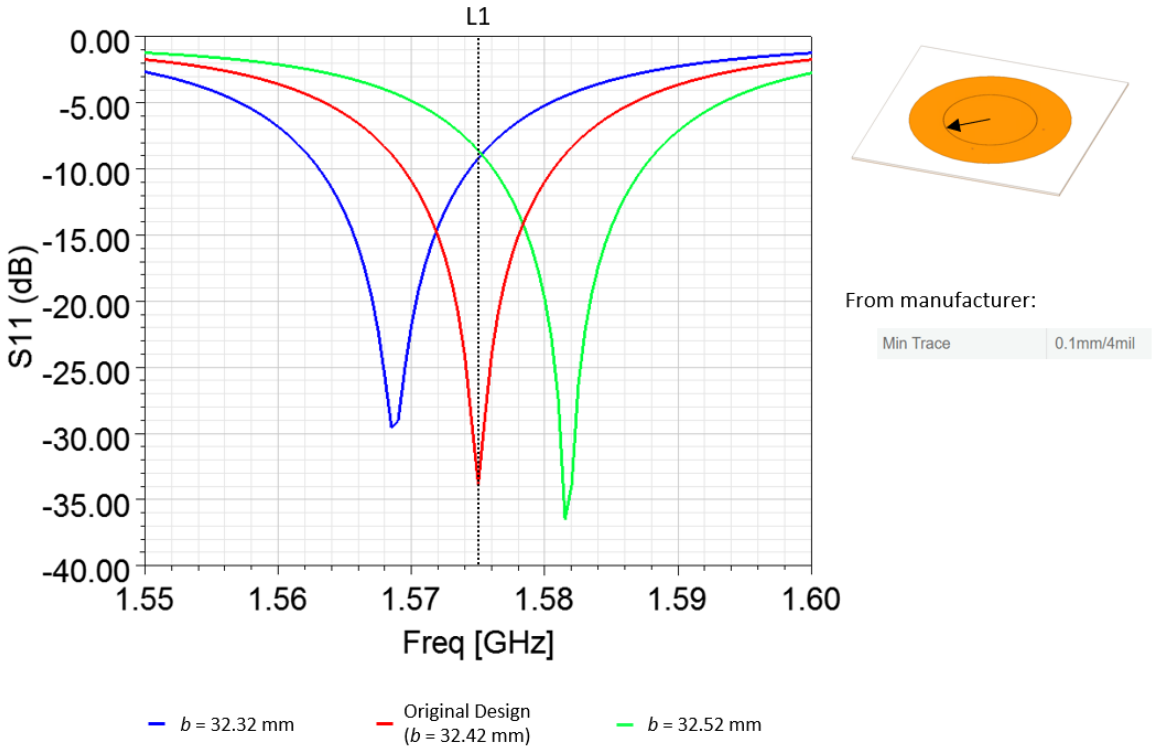
From PCBWay:

Board Thickness Tolerance(t≥1.0mm)	±10%

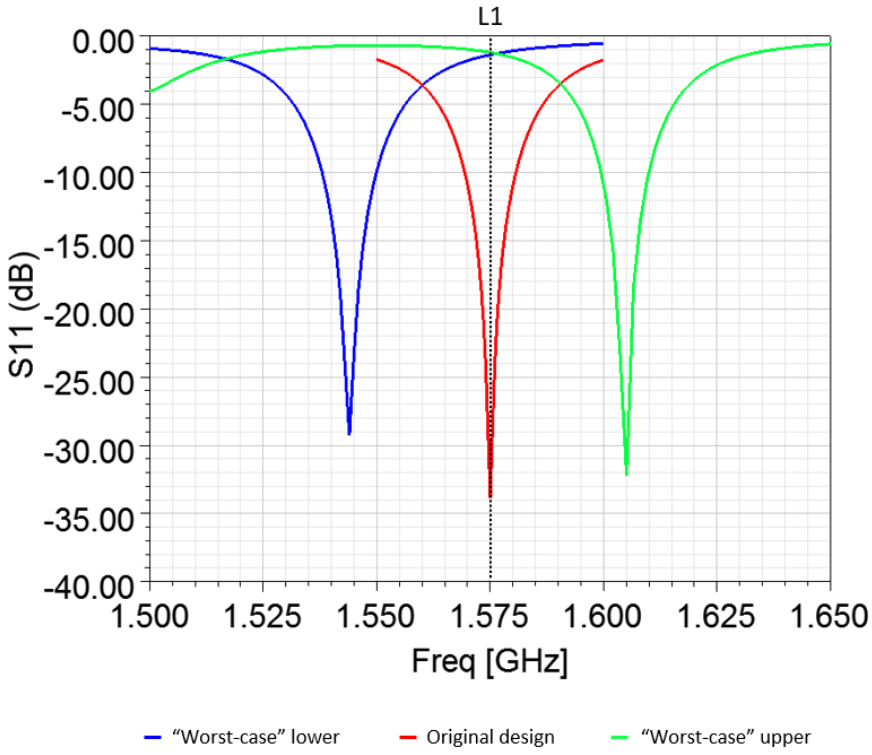
(b)


L1
0.00
-5.00
-10.00
-15.00
-20.00
-25.00
-30.00
-35.00
-40.00
S11 (dB)
1.55
1.56
1.57
1.58
1.59
1.60
Freq [GHz]
a = 55.9 mm
Original Design
(a = 55.8 mm)
a = 55.7 mm
From manufacturer:
Min Trace
0.1mm/4mil

(c)


L1
0.00
-5.00
-10.00
-15.00
-20.00
-25.00
-30.00
-35.00
-40.00
S11 (dB)
1.55
1.56
1.57
1.58
1.59
1.60
Freq [GHz]
b = 32.32 mm
Original Design
(b = 32.42 mm)
b = 32.52 mm
From manufacturer:
Min Trace
0.1mm/4mil

(d)



(e)

Figure 5.10: (a) Tolerance studies with varying substrate permittivity ϵ_r. Original design is based on RO4003C substrate with $\epsilon_r = 3.55$. The tolerance value is based on the process tolerance of ± 0.05 provided by Rogers. (b) Tolerance studies with varying substrate thickness h. We choose the tolerance value of $\pm 10\%$ based on the vendor PCBWay. (c) Tolerance studies with varying patch radius a. The value is chosen to be ± 0.1 mm based on the trace tolerance of PCBWay. (d) Tolerance studies with varying shorting boundary radius b, which is the radius of the ring of vias. The value is also based on the trace tolerance of PCBWay. (e) Simulated S_{11} with the nominal "worst-case" tolerance of the studied parameters. For example, "worst-case" lower shows the nominal worst-case result using tolerances of all parameters that shift the resonant frequency to a lower value.

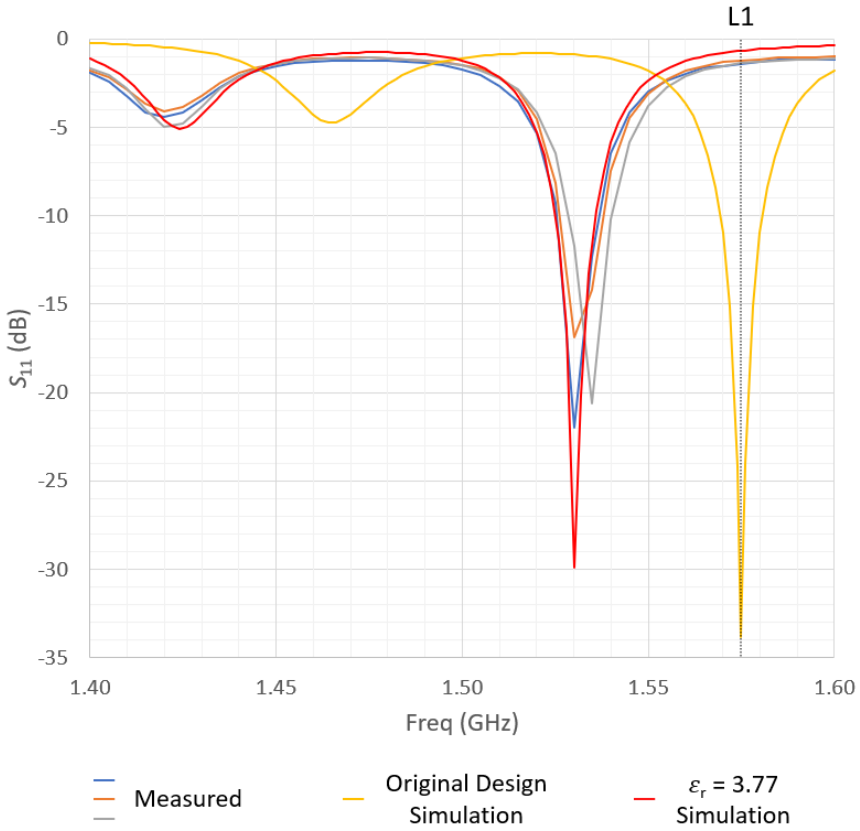


Figure 5.11: Comparison between measured and simulated S_{11} of the antenna feed ports.

frequency shift in the fabricated unit is more than the nominal "worst-case" scenario, and since the observed frequency shifts are all some-what consistent, as suggested by Fig. 5.9, the cause of which is more likely to be systematic rather than due to tolerances, which are likely to be random.

One important note is that the nominal permittivity of RO4003C substrate specified by Rogers is 3.55 from 8 to 40 GHz [39]; there is no simple way to verify that 3.55 is the actual permittivity value at L1 for the fabricated antenna substrates. Additionally, 3.55 is only the average value for a number of tested lots of materials by Rogers, and for a 20-mil (or 0.508 mm) thick RO4003C substrate, the permittivity is around 3.68 at L1, measured using differential phase length method by Rogers. HFSS simulations found that using a substrate with permittivity of 3.77 (other parameters stay the same as RO4003C) will shift the resonant frequency to around 1.53 GHz; this is plotted in red, shown in Fig. 5.11, alone with the previously simulated and measured results of Fig. 5.9. Therefore, future L1 design

using RO4003C substrate may require simulations with a substrate with permittivity of 3.77, with other parameters staying the same as RO4003C. Nonetheless, the tolerance study also shows the benefit of a wider-band design, which will potentially experience less impact from such resonant shifts.

5.1.6 Fabrication and Measurements: Antenna Radiation Characteristics

Since the fabricated antenna shows a frequency shift and resonates at around 1.53 GHz, the radiation measurements will be performed at 1.53 GHz instead of L1; the measured results will then be compared to the simulated results at L1. HFSS simulation verifies that the radiation characteristics for a design resonant at 1.53 GHz is very similar to those for the original design at L1; the SAR antenna design resonant at 1.53 GHz is the same one that

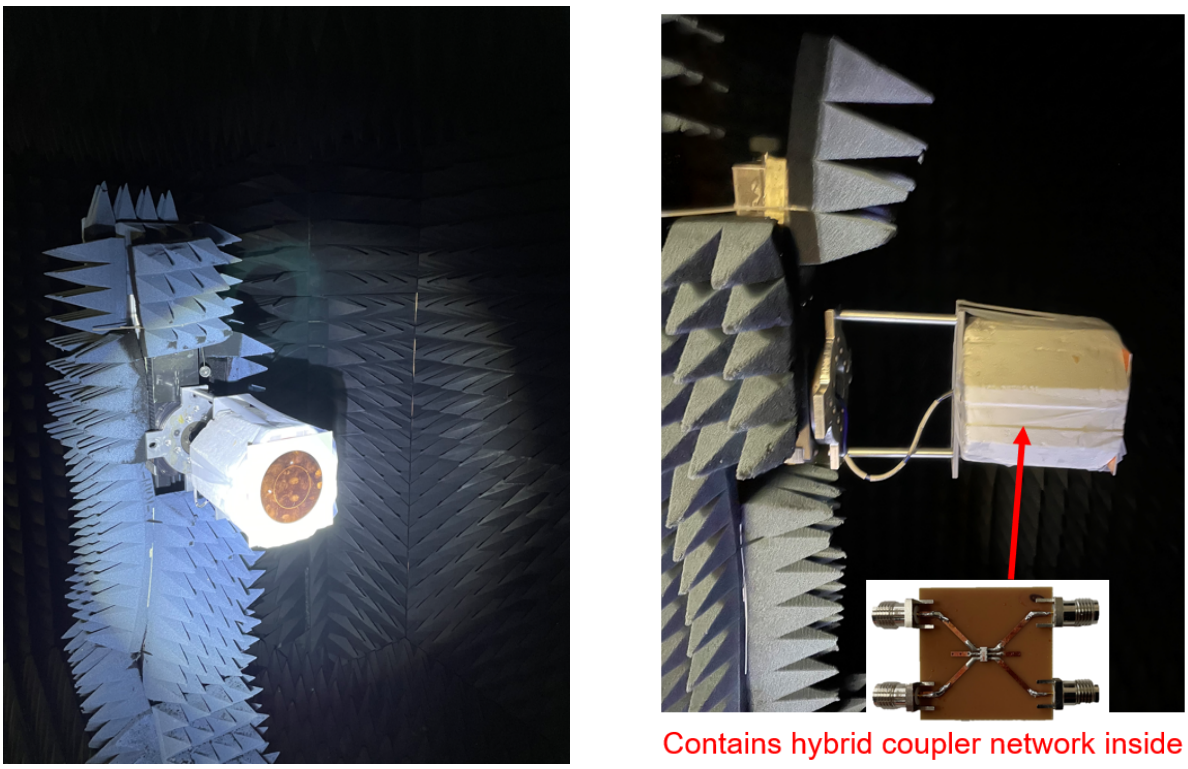


Figure 5.12: Measurement setup of the L1 SAR patch antenna in the spherical near-field antenna measurement chamber at the Center for High Frequency Electronics (CHFE) of UCLA.

produces the red curve shown in Fig. 5.11. Therefore, the comparison between the measured result and the simulated one at L1 is adequate.

The fabricated SAR patch is measured in the spherical near-field antenna measurement chanmber at CHFE of UCLA, shown in Fig. 5.12. The measurement is only done for the upper hemisphere, or $0° < \theta < 90°$, since the lower hemisphere is occupied by the antenna mount. The measured normalized gain pattern and AR at 1.53 GHz are plotted compared to the simulated ones of L1-band SAR patch, shown in Fig. 5.13 and 5.14. The measured RHCP radiation patterns follow rather closely with the simulated results. The measured directivity is around 9.003 dBic, compared to 9.4 dBic simulated. The X-pol levels, however, are observed to be significantly higher than the simulation, which means that the measured AR levels is also worse than the simulated one, as shown in Fig. 5.14. One important thing to consider is that the electrical connections between the test board, which contains the quadrature hybrid coupler, and the antenna can be relatively crude. The connections are

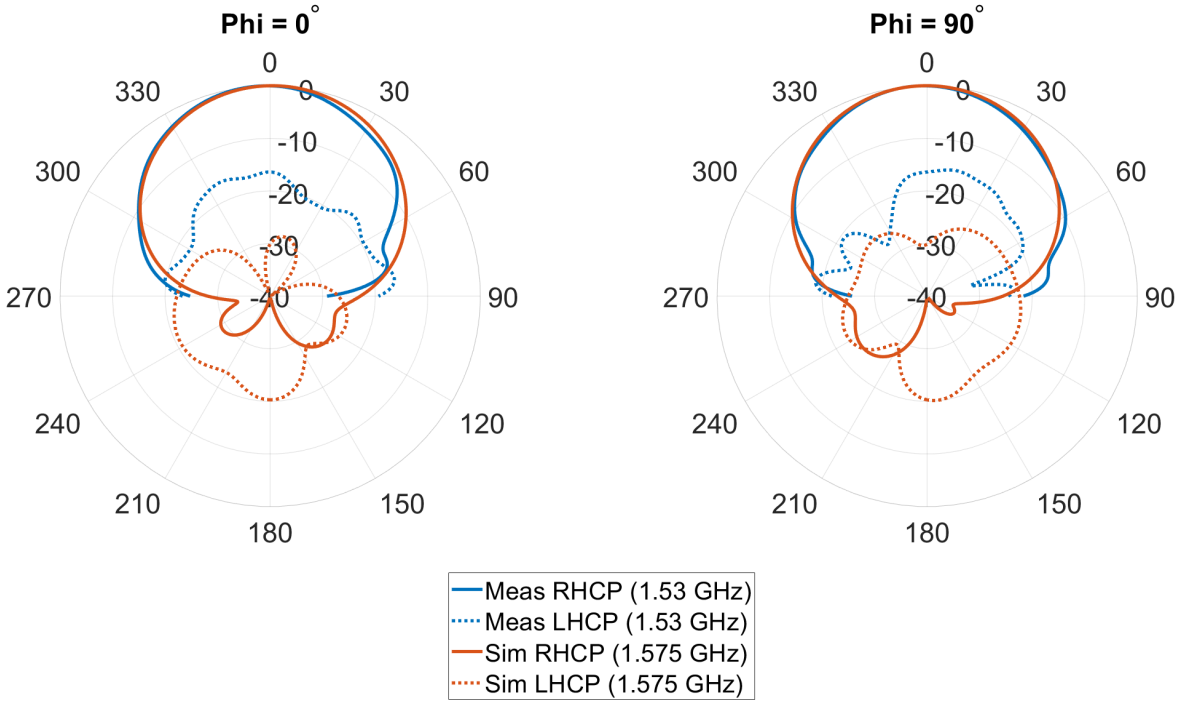


Figure 5.13: Comparison between measured and simulated normalized gain pattern for the SAR patch.

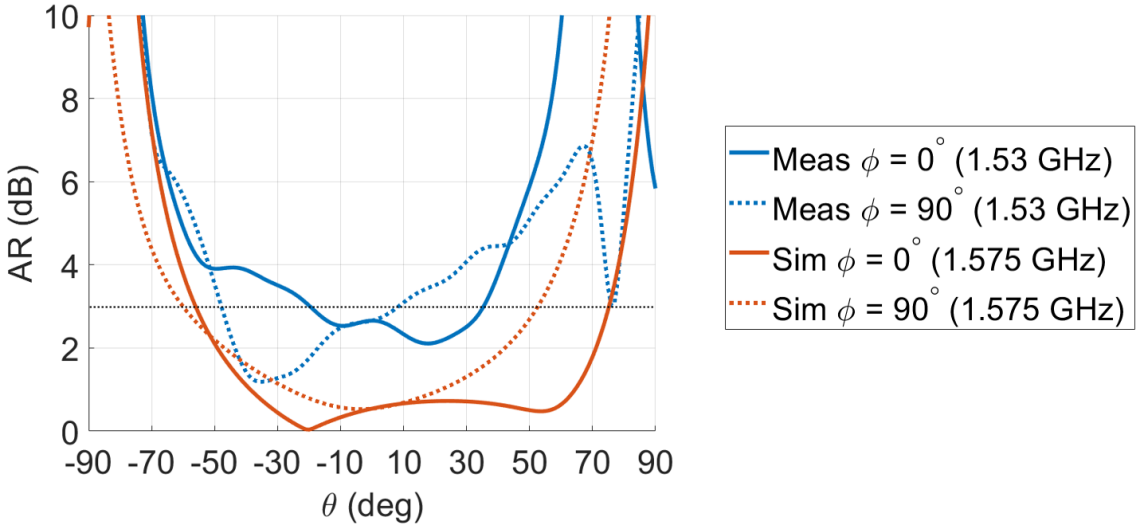


Figure 5.14: Comparison between measured and simulated AR for the SAR patch. The AR = 3-dB level is plotted in black dotted line.

made with hand-soldered SMA connectors and cables, which are more or less stretched using the assembly of the antenna and the test board. In the potential finalized design, the coupler will be soldered on a layer of substrate beneath the patch layer, and their connections will be realized with plated vias. These can be done with standard manufacturing procedures and therefore are not subject to nonidealities associated with hand-soldering and connections using SMA cables. Though the measured results are far from ideal, they nonetheless provided a rough assessment of the viability of the low-cost narrowband L1 patch design and the QCN-19+ hybrid coupler.

5.1.7 Four-Feed L1-Band SAR Patch Antenna

A potential four-feed L1-band SAR patch antennas is also simulated during design investigation. This configuration is large the same as the dual-feed design, except it uses four feeding probes with successive 90° phase shifts, shown in Fig. 5.15. The radius b of the via ring and the feed probe positions r_f are adjusted so that good resonance and matching at L1 are achieved.

The simulated S_{11}, gain pattern, and AR results are shown in Fig. 5.16. Clearly, the

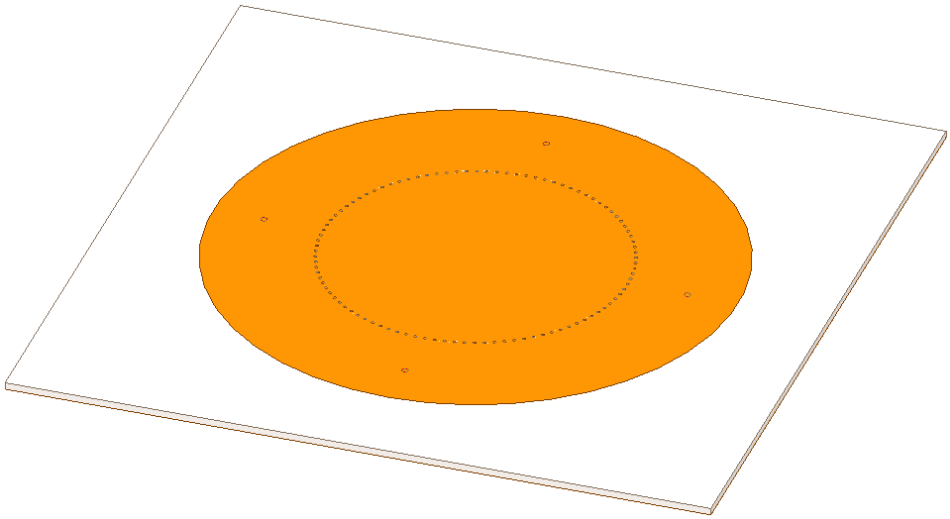
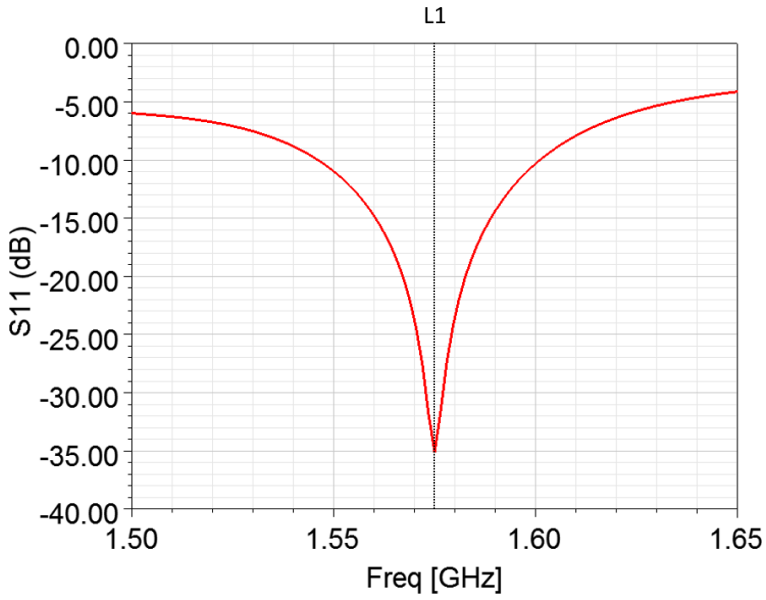


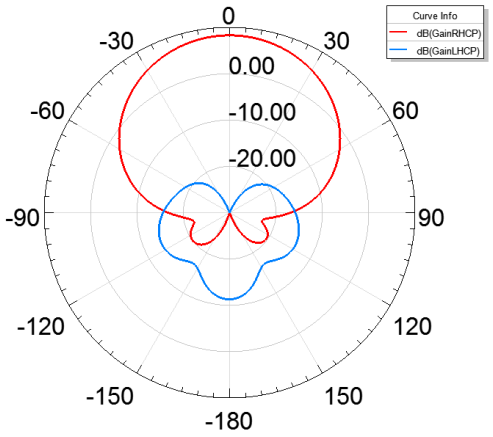
Figure 5.15: A four-feed L1-band SAR antenna design. Note that the four feeding probes are spatially separated by 90° concentrically, and their excitations are also phase shifted by 90° successively to achieve CP.

four-feed design inherits all the desirable performance features of the dual-feed design, such as low radiation levels near the horizon and very high XPD. Additionally, using four feeding probes saw a large improvement in the impedance bandwidth, with the additional benefits of a more symmetric radiation pattern. Unlike the dual-feed design shown in the previous sections, the four-feed configuration is circularly symmetric around the z-axis, hence the results at $\phi = 0°$ and 90° are expected to be identical.

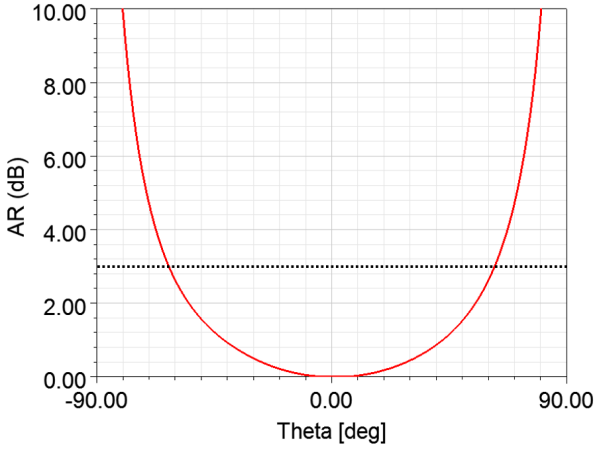
The notable feature of the four-feed design is its larger bandwidth of more than 50 MHz, which is roughly five times that of the dual-feed design. Although 50 MHz may still be regarded as narrow-band in the conventional sense, this bandwidth is more than enough for receiving both the civilian and military L1-band signals. The wider bandwidth will also enable the design to be more resilient against potential frequency shifts in the manufactured unit due to nonidealities. Nonetheless, one drawback of the four-feed design is the more complicated feeding (or receiving) networks, which can potentially increase the cost.



(a)



(b)



(c)

Figure 5.16: (a) Simulated S_{11}; all four ports yield identical results. (b) Simulated gain pattern at $\phi = 0°$ and 90°. (c) Simulated AR versus polar angle θ at $\phi = 0°$ and 90°.

5.2 Dual-Band SAR Antenna Characterizations

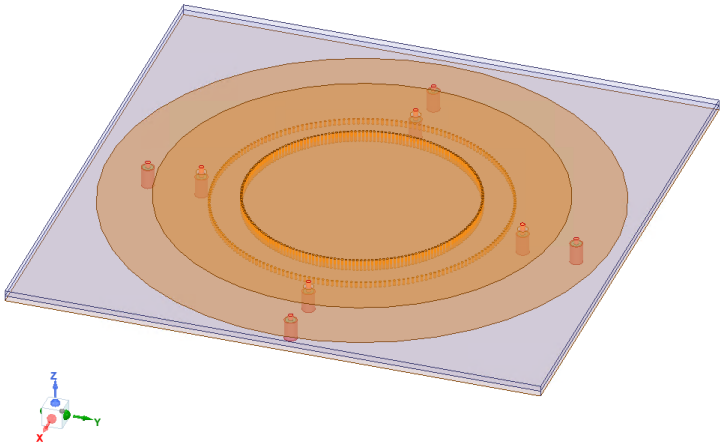
As discussed in Section 4.4, a popular configuration of dual-band SAR antenna uses a concentric pair of L2-band ISAR antenna and L1-band SAR antenna, as shown in Fig. 4.6. However, since the ISAR antenna also follows the same design equation as a SAR patch, we can calculate the radii of the magnetic current loop and the shorting boundary required

at L2-band using Eq. 4.25 and 4.30 respectively. Since the L2-band antenna is ISAR, the radius of the shorting boundary b is larger than the radius of the magnetic current loop a. The numerical value for the magnetic current loop radius a at L2 is determined to be 70.7 mm. This approaches close to the design constraint for this project, which limits one dimension of the antenna to be within 150 mm. A shorting boundary then will definitely exceed the constraint and therefore not realizable. Hence, an alternative dual-band configuration is needed.

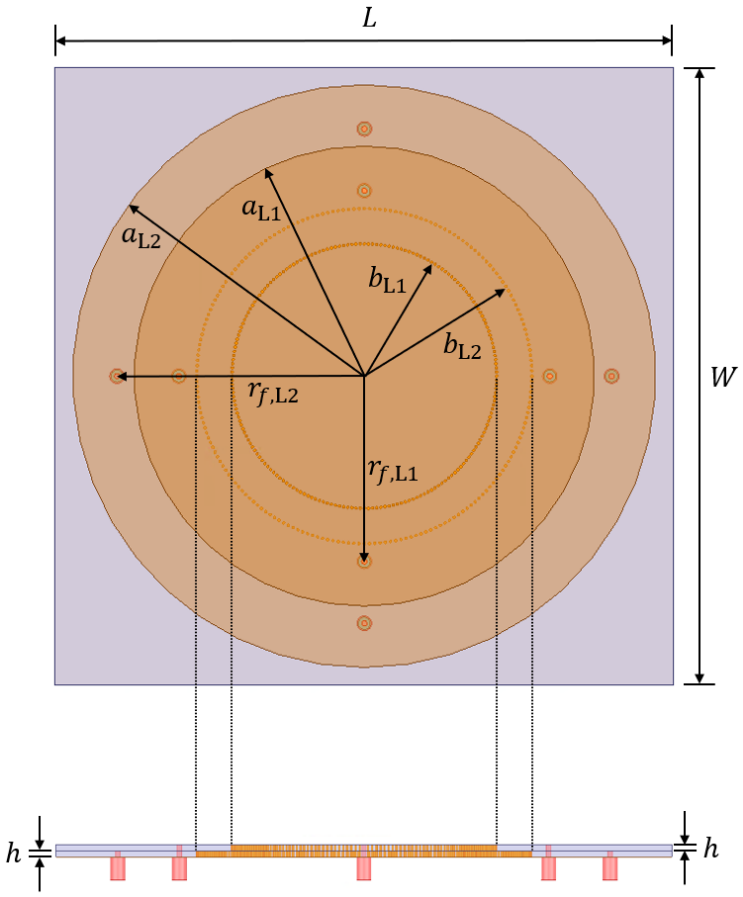
5.2.1 Stacked SAR Patch

Stacked patch antenna is a common configuration to realize dual-band operation. A dual-band stacked SAR antenna configuration operating at the GPS L1 and L2 bands is simulated; the geometry is shown in Fig. 5.17. This configuration consists of two layers of RO4003C substrate with three conductive layers. The smaller L1-band SAR patch is positioned at the top conductive layer and the larger L2-band SAR patch is positioned at the middle conductive layer, sandwiched between the two substrates. Each patch is fed by four coaxial probes, which increases the impedance bandwidth comparing to dual-feed designs. The dimension of the designs are shown in Table 5.6 for the parameters shown in Fig. 5.17b. One advantage of the stacked dual-band design is its compatibility with the conventional PCB manufacturing process. As shown in Fig. 5.17, the L1-band shorting boundary is formed using a ring of plated through-hole vias, similar to the single band design in Section 5.1.7; the L2-band shorting boundary consists of a ring of plated blind vias, connecting only the middle and bottom conductive layers.

The simulated results are shown in Fig. 5.18. The impedance bandwidths at L1 and L2 frequencies are around 40 MHz and 50 MHz respectively, as shown in Fig. 5.18a, which are more than enough to cover the entire L1- and L2-band signals. The radiation patterns and AR at both frequencies both suggests good performance characteristics, consistent with the single-band design discussed in Section 5.1.7.

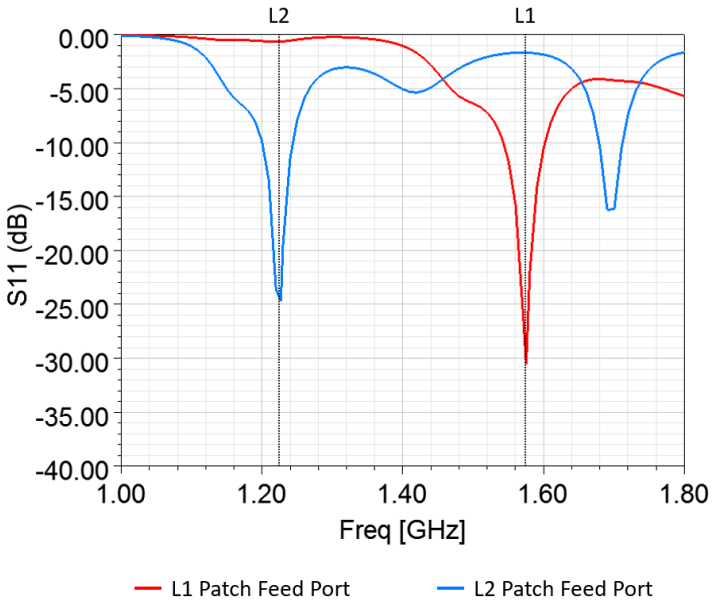


(a)

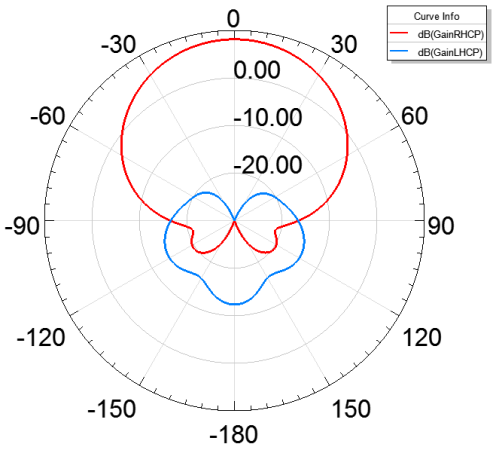


(b)

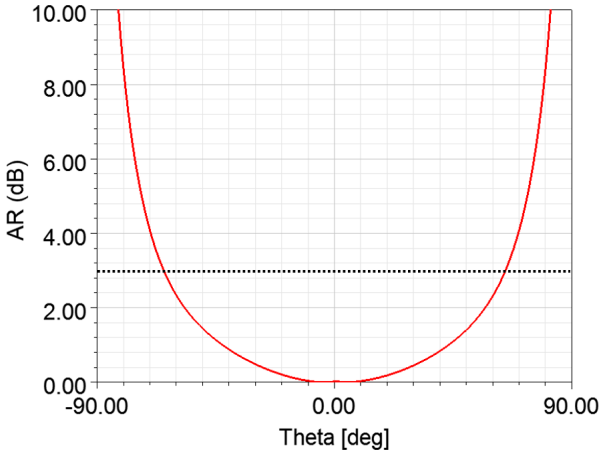
Figure 5.17: (a) 3D view of the stacked SAR patch antenna design model. (b) Top and side views of the stacked SAR patch with labeled design parameters, whose values are shown in Table 5.6. The red objects attached at the bottom are coaxial feed probes.


L2
L1
0.00
-5.00
-10.00
-15.00
-20.00
-25.00
-30.00
-35.00
-40.00
S11 (dB)
1.00
1.20
1.40
1.60
1.80
Freq [GHz]
L1 Patch Feed Port
L2 Patch Feed Port

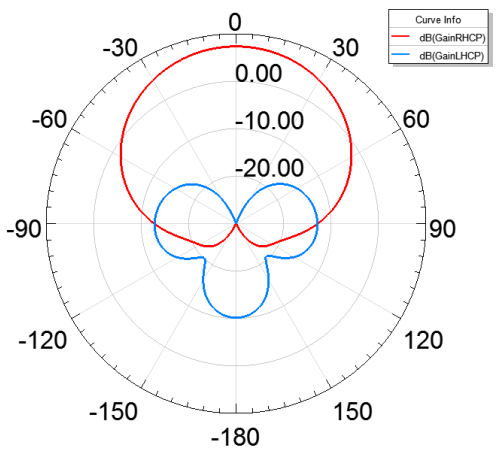
(a)


Curve Info
dB(GainRHCP)
dB(GainLHCP)
0
30
60
90
120
150
-180
-150
-120
-90
-60
-30
0.00
-10.00
-20.00

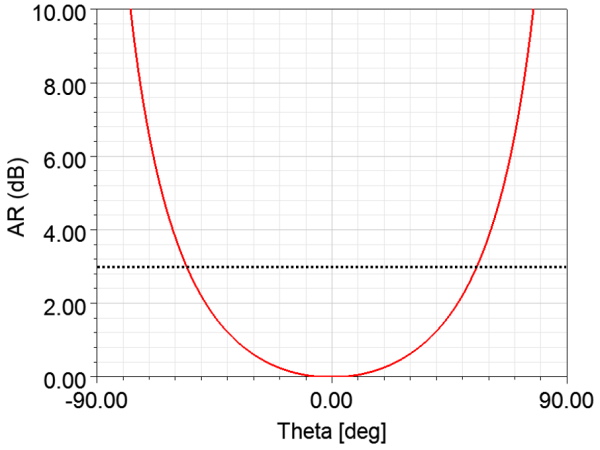
(b)


10.00
8.00
6.00
4.00
2.00
0.00
AR (dB)
-90.00
0.00
90.00
Theta [deg]

(c)



(d)



(e)

Figure 5.18: (a) Simulated S_{11} for the stacked SAR patch antenna; for each patch, the S_{11} results of the four feed ports are identical. (b) Simulated gain pattern at $\phi = 0°$ and $90°$ for the L1 patch. (c) Simulated AR versus polar angle θ at $\phi = 0°$ and $90°$ for the L1 patch. (b) Simulated gain pattern at $\phi = 0°$ and $90°$ for the L2 patch. (c) Simulated AR versus polar angle θ at $\phi = 0°$ and $90°$ for the L2 patch.

Table 5.6: Nominal design values used in the HFSS simulation for the antenna dimensions shown in Fig. 5.17.

Design parameters	Nominal values (mm)
Substrate length (L)	150
Substrate width (W)	150
Single-layer substrate thickness (h)	1.524
L1 patch radius (a_{L1})	55.8
L1 shorting boundary radius (b_{L1})	32.1025
L1 feed location ($r_{f,\mathrm{L1}}$)	45
L2 patch radius (a_{L2})	70.7
L2 shorting boundary radius (b_{L2})	40.6425
L2 feed location ($r_{f,\mathrm{L2}}$)	60

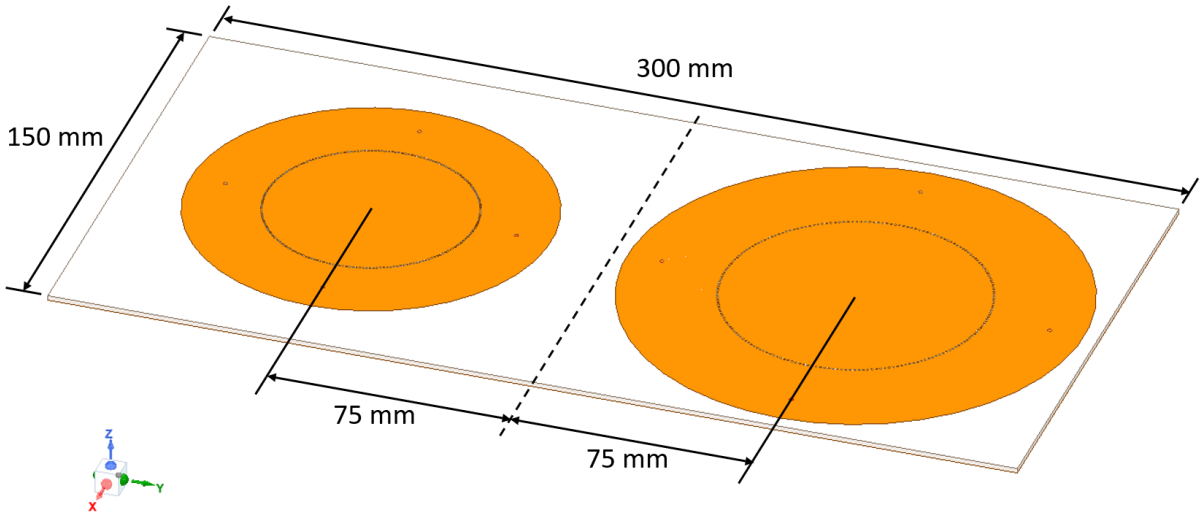


Figure 5.19: The dual-band side-by-side configuration, with the smaller L1-band SAR patch on the left, and the larger L2-band SAR patch on the right.

5.2.2 Side-by-Side SAR Patch

Since the design requirement has a maximum size constraint of 150 mm × 300 mm, the previously shown single-band and stacked dual-band design, having the same profile of 150 mm × 150 mm, only utilize half of the available space. The extra space enables us to alternatively place the L1- and L2-band patches in a side-by-side configuration, as shown in Fig. 5.19, instead of a stacked configuration.

Under typical GPS operating scenarios, this side-by-side placement may not be desired, since it will introduce a considerable offset between the phase centers of the two antennas. The realistic impact of this offset on the positioning operation by the GPS receiver will need to be determined experimentally, and if necessary, compensated through signal processing or software means. However, one radiation environment of concern for this project is when the antenna is placed close to a large metallic body, i.e. the freight container, which essentially introduces blockage of half of the upper hemisphere. In this case, at least two antennas are needed to place on each side of the freight-carrying vehicle in order to obtain acceptable GPS signal reception. This means that the receiver should readily account for the position measurements from two spatially-separated antennas. In this way, studying the performance of

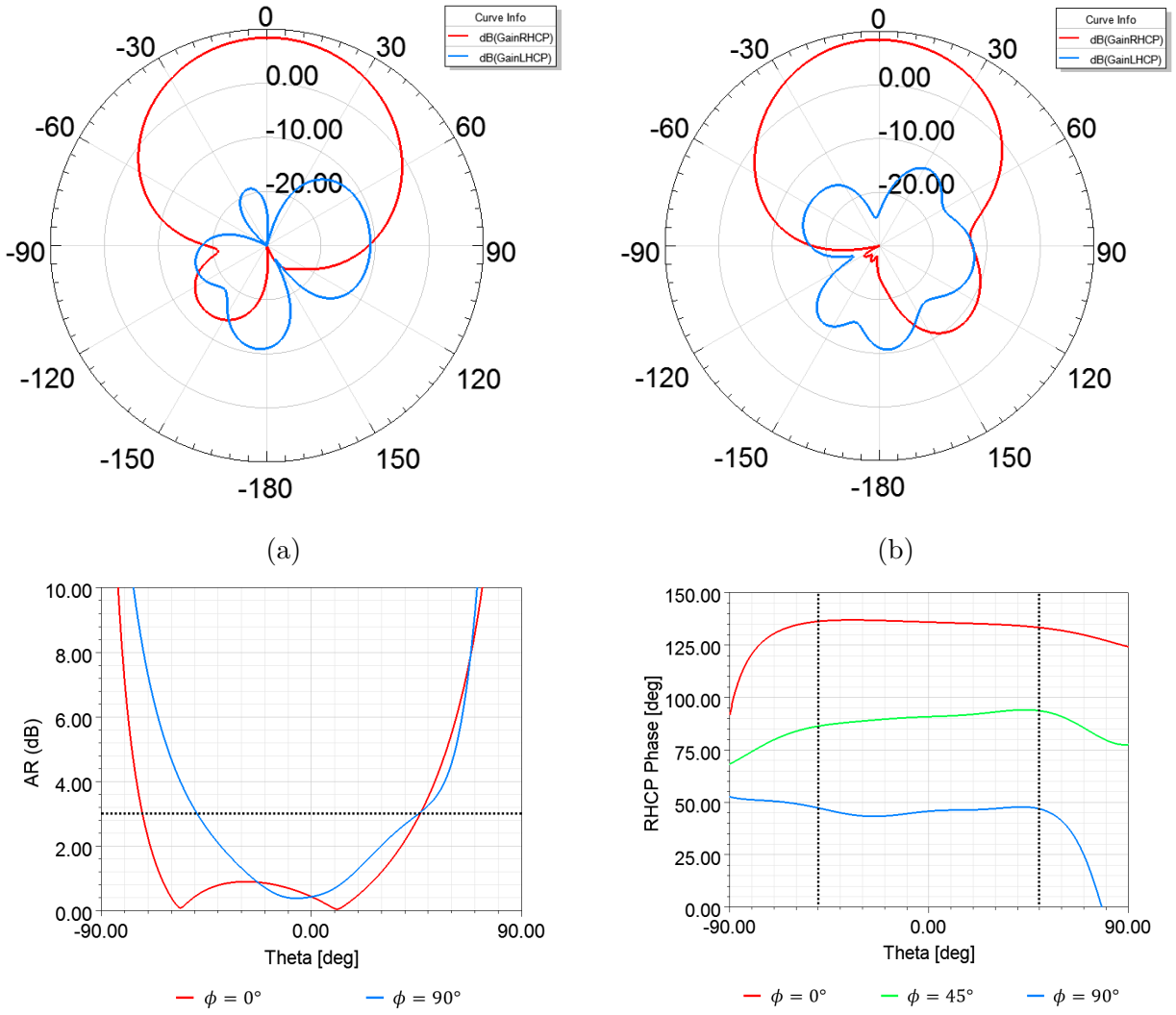


Figure 5.20: Side-by-side configuration simulation results at L1. Simulated gain pattern for (a) $\phi = 0°$ and (b) $\phi = 90°$. (c) simulated AR versus polar angle θ. (d) Simulated RHCP radiated far-field phase versus the polar angle θ when the origin of the coordinate system is at the phase center of the antenna. The phase center is found at around $z = 15$ mm, where $z = 0$ is defined at the ground plane layer, and the z-axis is through the L1 patch geometric center. The black dotted lines denote $\theta = 50°$.

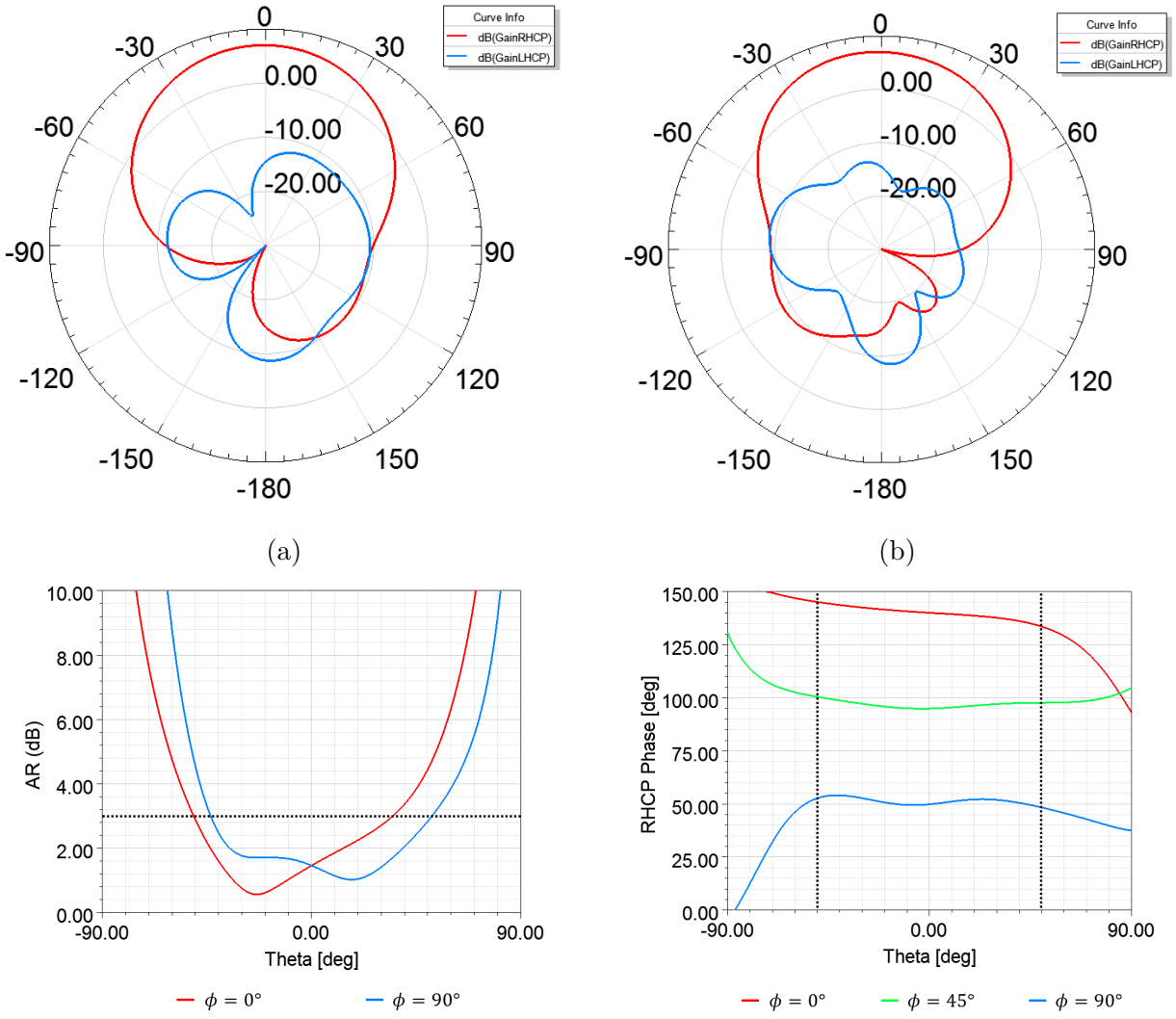


Figure 5.21: Side-by-side configuration simulation results at L2. Simulated gain pattern for (a) $\phi = 0°$ and (b) $\phi = 90°$. (c) simulated AR versus polar angle θ. (d) Simulated RHCP radiated far-field phase versus the polar angle θ when the origin of the coordinate system is at the phase center of the antenna. The phase center is found at around $z = 15$ mm, where $z = 0$ is defined at the ground plane layer, and the z-axis is through the L2 patch geometric center. The black dotted lines denote $\theta = 50°$.

a side-by-side configuration is not unrealistic. Nonetheless, the practical performance under said scenario should be verified via field tests. In this section, we assume that the receiver is able to perform positioning through multiple spatially-separated antennas. Though the stacked patch design has the advantage of having a smaller size, it is worth while to investigate the the side-by-side configuration, since if plausible, it has the potential advantage of being more flexible in design for each of the L1- and L2-band antennas. In this section, we focus on the SAR design for both bands.

A typical concern for placing microstrip patch antennas close together is mutual coupling, which is a function of edge separation between adjacent elements [7]. Mutual coupling can have potential adverse effects on important performance parameters for GNSS applications such as impedance mismatch and degradation in RHCP gain and AR. Mutual coupling is possible through both surface wave, particularly the TM_0 mode, and space wave radiations. For SAR patches, the surface wave and lateral waves can be significantly reduced, as explained in Chapter 4, and hence should alleviate the deleterious effect of mutual coupling in principle. The simulated radiation characteristics at L1- and L2-bands are shown in Fig. 5.20 and 5.21 respectively. The S_{11} results are virtually identical to Fig. 5.18a, since the patch designs are the same. Some design dimensions are slightly tuned so that the patches resonant at L1 and L2. With the side-by-side configuration, the radiation patterns do not exhibit the symmetry as those of the stacked design. As expected, it is caused by coupling effects from adjacent elements, though the effects are not too severe such that the radiation characteristics become undesirable. The highest LHCP backlobes gain are around the same level as the stacked design, with a D/U ratio of around 15 dB or more for angles close to the broadside. The radiations toward the horizon are 15 dB or more below the zenith. The AR results show a considerable beamwidth at both L1 and L2 for good CP performance (AR < 3 dB). The far-field phase simulations shown in Fig. 5.20d and 5.21d indicate that within the 100° beamwidth (the beamwidth at $\theta = 50°$), the RHCP far-field phase vary by around 10° at the most, which corresponds to sub-centimeter-level variation in length measurement; such precision is more than sufficient for our application.

5.3 Radiation Performance With Side Plate

Ultimately, the potential SAR antenna design needs to function on the autonomous vehicle when a shipping container is attached. This presents an atypical radiation environment, in which roughly half of the upper hemisphere will be blocked by a conductor positioned very close to the antenna. Such environment is expected to be very adverse to the radiation performance of any GPS antenna. A simulation using a full-sized container model will be very unrealistic for HFSS to run, since it will require extremely large mesh operations. Hence, we will approximate the freight container using a large side plate with a size of 2 m $\times$ 2 m, shown in 5.22.

The SAR antenna used for the simulation will be the L1-band dual-feed SAR patch discussed in Section 5.1.2; the simulated results shall provide insights to the potential radiation

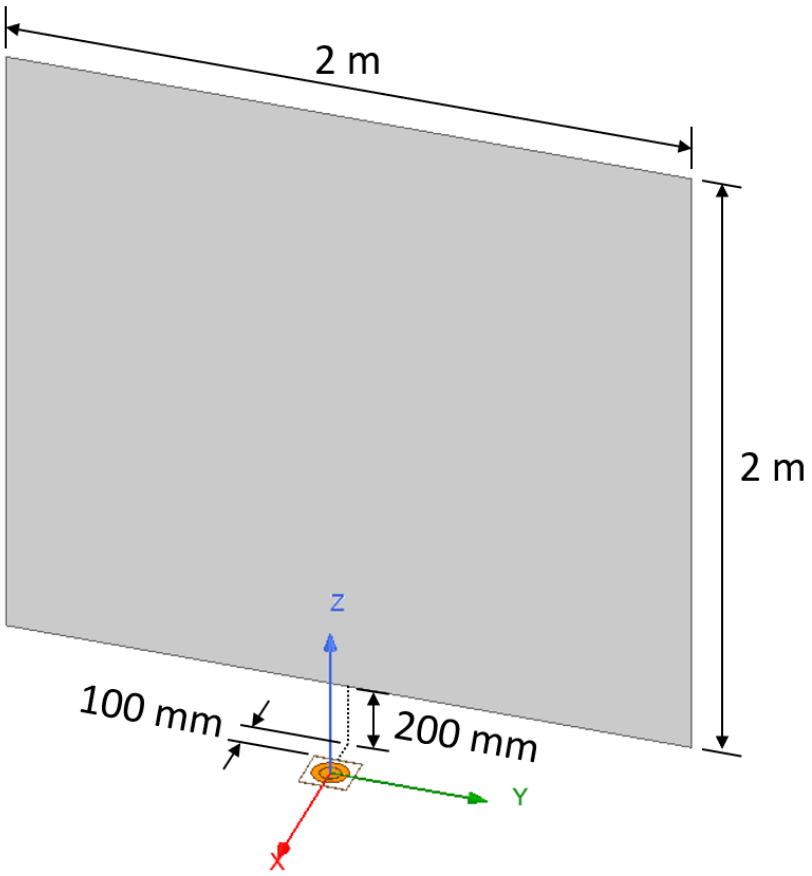
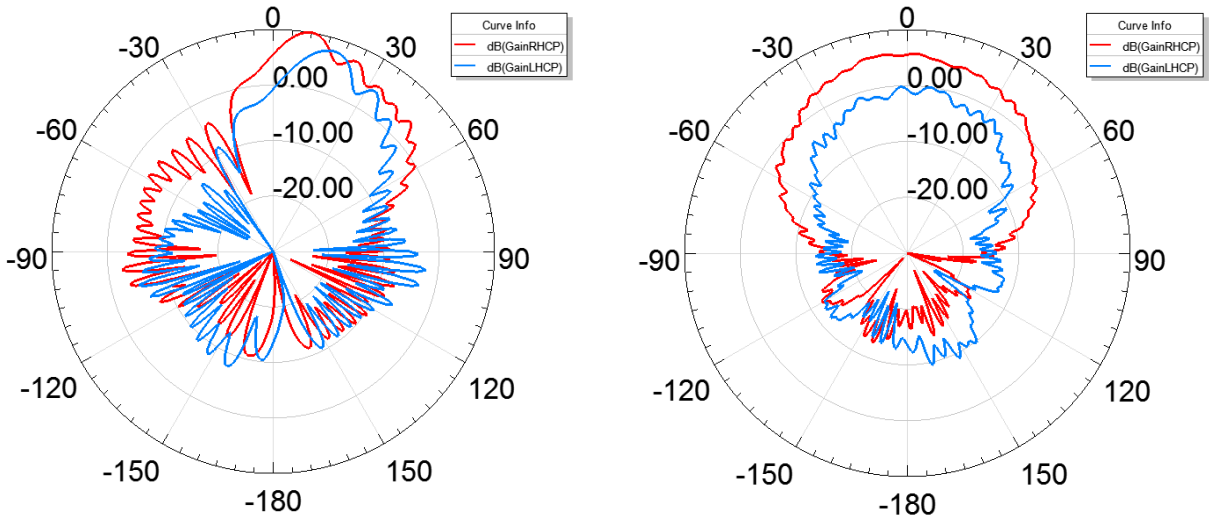


Figure 5.22: The freight container is modeled as a large PEC side plate.

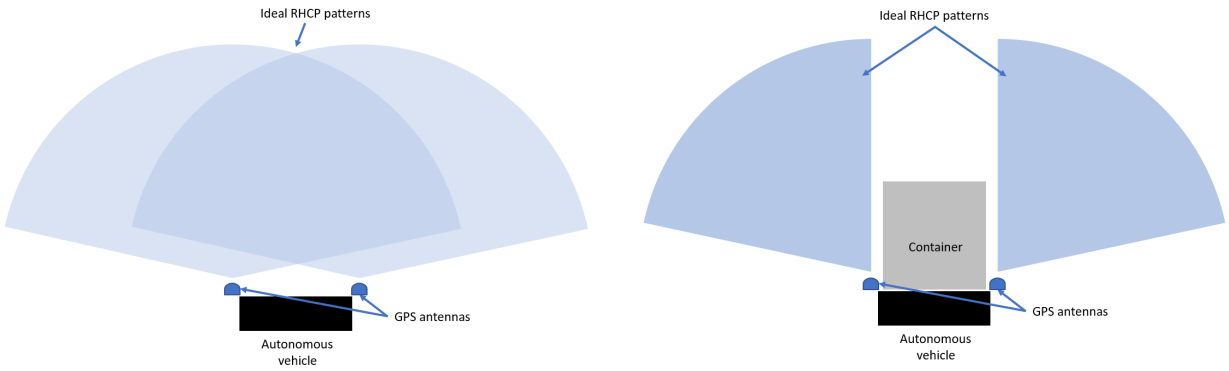


(a) (b)

Figure 5.23: Simulated (a) $\phi = 0°$-plane and (b) $\phi = 90°$-plane gain patterns of the setup in Fig. 5.22.

impact on other SAR antenna designs, such as the L2-band patch and the stacked or side-by-side dual-band patch, since their radiation patterns are similar. The side plate will be positioned 100 mm from the edge of the antenna substrate in the x-direction and offset by 200 mm in the z-direction; this is shown in Fig. 5.22. These are reasonable values chosen based on the expected installed location of the antenna relative to the freight container; the exact values are not a concern for this simulation study, since the primary objective is to gain insight on the impact on the radiation performance. The simulated radiation patterns of the setup of Fig. 5.22 is shown in Fig. 5.23. Clearly, the side plate results in significant platform effects on the radiation pattern. The radiation patterns in both plane-cuts exhibits significant ripples as expected, which are resulting from the diffraction from the side plate edges. More importantly, the RHCP radiation is completely destroyed in the $\phi = 0°$-plane, since the co-pol and X-pol are are similar levels, which will realistically result in polarization efficiency loss. The $\phi = 90°$-plane cut shows more symmetric pattern as expected from the geometry, but also shows poor XPD values of around 5 dB throughout the beamwidth.

The drastic increase in LHCP radiations as shown in the radiation patterns is expected



(a) (b)

Figure 5.24: The ideal radiation patterns of the GPS antenna on the autonomous vehicle (a) with and (b) without a freight container installed

from the reflections from the side plate. There is no easy way of avoiding radiation interaction between the antenna and the side plate, as the antenna is installed very close to such object. Since half of the upper hemisphere is blocked by the side plate, for ideally no radiation interactions, an antenna placed in the orientation shown in Fig. 5.22 must exhibit asymmetric radiation pattern or a radiation pattern that changes when the freight container is installed. The depictions of the ideal radiation patterns for the GPS antennas installed on the autonomous vehicle with and without a freight container is shown in Fig. 5.24. Needless to say, the radiation patterns shown in 5.24b is not easily achievable on conventional, low-cost antenna designs, since the ideal radiation pattern is not only relatively directive but also uniform in the directivity/gain value across a wide beamwidth. To achieve such specialized radiation patterns, excitation distribution design, array or adaptable antenna architecture may need to be considered. These solutions are expected to be costly, and additionally, their feasibility within our space constraint of 150 mm $\times$ 300 mm is questionable. Non antenna-oriented solutions such as covering the freight container with radiation-absorbent material may alleviate the undesirable effects, but they will not be our focus.

5.3.1 Radiation Performance With Antenna Tilting

To reduce radiation interaction between the antenna and the side plate, one potential simple solution is to tilt the antenna by 45° away from the container; this is depicted in Fig. 5.25. The tilting may be realized through low-cost actuators, which activates when the freight container is installed on the vechicle. For the tilted antenna, though, radiation interaction with the ground will now become an issue. The 45° tilt angle is chosen since it is the mid point between the zenith and the ground/horizon, and that the SAR patch radiation pattern is symmetric. However, the ground is not modeled in the simulation due to additional mesh size considerations on simulation time and size, and that a finite sized reflective surface will introduced additional edge diffraction that is not realistic for the ground. Nonetheless,

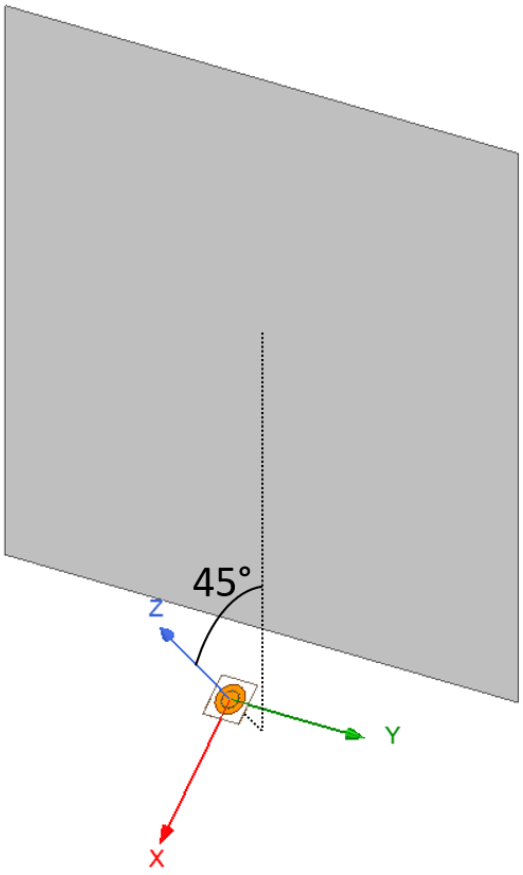


Figure 5.25: The SAR patch antenna is tilted by 45° away from the side plate. This is done by rotating the antenna by 45° around the y-axis based on the geometry of Fig. 5.22.

effects from the ground merit further investigation in the future.

The simulated antenna performance is shown in Fig. 5.26. Note the orientation of the coordinate system with respect to the side plate as shown in 5.25; we chose the z-axis to be pointing out of the antenna geometric center. In this case, $\phi = -45°$ is the zenith and $\phi = 45°$ is the horizon, as labeled in Fig. 5.26a. It is clear that tilting the antenna can potentially improve the radiation characteristics; namely, the patterns exhibit relatively less diffraction ripples and lower X-pol levels. For $\phi = 0°$, the AR is relatively high for the angles close to the zenith, which still suffers from considerable radiation interaction with the side plate, though much less so than without tilting. For $\phi = 90°$, the AR is less than 3 dB for a wider and more symmetric angular range around the antenna main beam, as expected from the geometry.

The radiation interaction with the side plate also affects the antenna far-field radiated phase characteristics. Since it will take a considerable amount of simulation time to find the antenna phase center with the side plate, we assumed it is not far from the phase center for the L1-band SAR patch of Section 5.1.2, which is at $\Delta z = 15$ mm along the z-axis pointing out of the antenna geometric center. We then vary the local coordinates by a small amount around the $\Delta z = 15$ mm point and determined that the phase variations around this point is relatively more uniform, and hence it is in the vicinity of the actual phase center. The phase characteristics shown in Fig. 5.26d all display uneven ripples resulting from radiation interactions with the side plate.

For the $\phi = 0°$-plane, we will mainly focus on the phase characteristics within the $-45° < \theta < 45°$ range, which is the angular space between the zenith and horizon; the phase will vary drastically outside this range as expected. The largest phase variation is close to $\theta = 40°$, which is near the side plate. Nonetheless, it is still within 25∘ from the phases at other angle θ within the range of interests; this corresponds to a maximum ranging error of around 1.3 cm, which is still relatively small. For the $\phi = 45°$-plane, the largest phase variation is close to $\theta = 50°$ with a maximum deviation of 70∘ from the phases at other angle θ; this will result in a maximum ranging error of around 3.7 cm. Lastly, for the $\phi = 90°$-plane, the phase variation is the most uniform within the angular range of $-40° < \theta < 40°$; this is

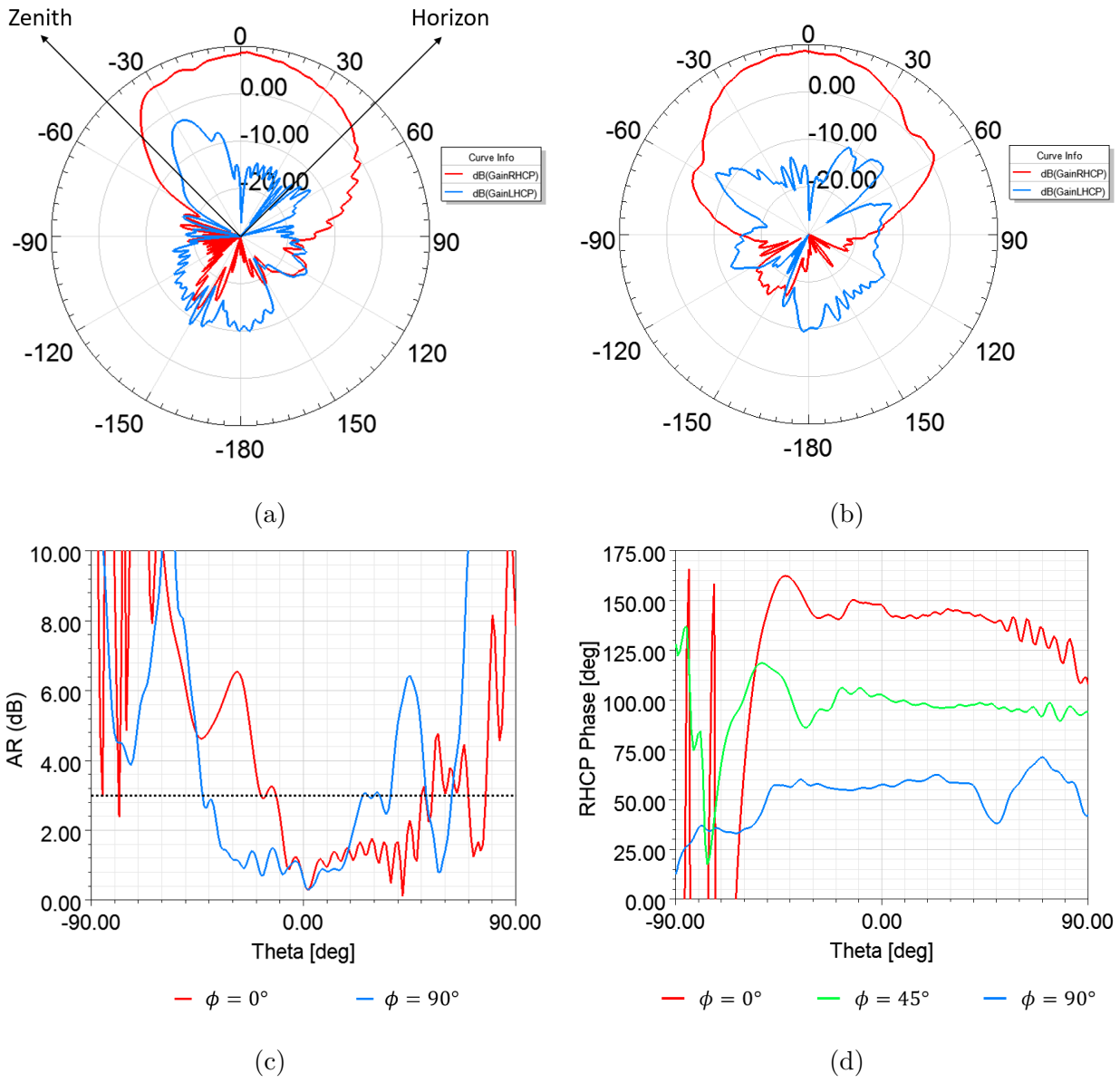


Figure 5.26: Simulated results for the setup in Fig. 5.25. Simulated gain pattern for (a) $\phi = 0°$ and (b) $\phi = 90°$. (c) Simulated AR versus the polar angle θ. (d) Simulated RHCP far-field phase versus the polar angle θ when origin of the coordinate syst em is placed near the phase center of the antenna.

expected since it is the plane that will experience the least amount of platform effects from the side plate.

These simulated results provide a rough expectation on the antenna radiation characteristics when it is installed near a large metallic object such as the freight container. However, it should be stressed that for the tilting setup, the analysis are done without incorporating the effects from the ground. Without the ground element, the radiation pattern near the horizon is not expected to be accurate as shown in Fig. 5.26a, and the actual radiation characteristics are expected to be worse than Fig. 5.25 shows. More importantly, the tilting will subject the antenna to more multipath signals; the effects of which should be duly evaluated in real-world performance of the antenna. Additionally, the setups shown in Fig. 5.22 and 5.25 are certainly simplistic models compared to the actual antenna mount platforms and freight containers. Ultimately, field tests of the antenna radiations with a side plate and exposed ground should be conducted to assess the actual antenna performance in real-world conditions. This merits further investigation in the future.

CHAPTER 6

Conclusions and Future Works

Starting in the 1980s, the Global Positioning System (GPS) became accessible for general use and since then, it has continued to be indispensable in many aspects of modern life. Multitude of GPS antenna and receiver designs are developed and installed on distinct platforms, enabling a wide range of applications. New applications with unique platforms and requirements are continued to be devised. This thesis aimed to investigate a potential GPS antenna design that would enable the precise navigational needs of a novel autonomous, freight-carrying, railway vehicle. The antenna should feature dual-band, and right-hand circular polarization (RHCP) with sufficient gain and beamwidth while remaining low-cost, limited in size, and possess multipath rejection capability. In the first part of this work, we presented the basic operation and navigation principle of the GPS system, detailed the typical desired features a GPS antenna possesses, and presented several proven GPS antenna designs. The autonomous vehicle also imposes atypical radiation environment when a freight container is installed. It is a challenge to simultaneously satisfy the constraints while providing optimal performance with and without a freight container installed on the autonomous vehicle.

Precise GPS receivers require robust multipath signal rejection capabilities, since multipath is one of the main issues affecting the ranging precision of a GPS receiver. Fortunately, multipath can be mitigated through antenna design. As the main focus of this work, a reduced-surface-wave (RSW) shorted-annular-ring (SAR) patch antenna is proposed and studied. The theoretical treatment using cavity model suggests that a circular patch with a radius at a particular value would result in suppressed surface wave radiations. This provides several advantages desirable for a GPS antenna, including low radiation toward horizon, low backlobes, reduced platform effects, and limited mutual coupling between adjacent SAR

antenna elements. These radiation characteristics enable the SAR antenna to satisfy multipath requirement without the need of bulky and expensive specialized ground planes such as a choke ring. A L1-band SAR antenna prototype was manufactured through a potential low-cost vendor and its performance was measured using facilities at UCLA's CHFE. A quadrature hybrid coupler feed network is also constructed to provide the required circular polarization. The low-cost prototypes exhibits adequate performance; deviations from the simulated results are likely due to manufacturing/assembly tolerances. More precisely manufactured and assembled units are desired in the future.

The minor frequency shifts in the S_{11} of the measured prototype may be alleviated using wider band designs, such as using four feed probes that increases the impedance bandwidth by almost five times that of the dual-feed design. Dual L1-L2 band SAR patch antennas are also simulated. The stacked configuration provides good radiation characteristics at both bands. The side-by-side approach utilizes all the available space in the constraint, and can provide more flexibility than the stacked configuration in the design of the individual antenna at each band if desired, though further investigations are needed on the positioning effectiveness using separate L1- and L2-band antennas. Dual-band antennas and feed network designs for the four-feed configuration should be prototyped in future works, after the performance assessments on the more precisely fabricated single-band antennas.

The L1-band SAR patch antenna is also simulated in the presence of a large side plate modeling the freight container. The side plate will severely affect the radiation characteristics of the SAR patch antenna placed perpendicular to it. The tilting concept may alleviate the radiation interactions with the side plate. However, the simulated container models are simplified and limited in size and scope. A more accurate characterizations of the antenna performance, with and without tilting, near the container should be investigated; the performance of a tilted antenna should be field-tested to assess the realistic impact from the side plate, the ground, and multipath signals in order to evaluate the feasibility of the concept in actual operations. The studies performed in this thesis provide the necessary theoretical foundation and design basis for developing a complete and robust GPS antenna for the autonomous vehicles in the future.

APPENDIX A

Derivation of the Fundamental Design Equation for a SAR Patch With No Surface Wave Excitation

The fundamental design equation, as expressed in (4.21) of Chapter 4, for a SAR patch is reproduced here:

$$\psi \propto H_1^{(2)}\left(k_{\mathrm{TM}_0}\rho\right)\cos\phi J_1'\left(k_{\mathrm{TM}_0}a\right) \tag{A.1}$$

where ψ represents any particular component of the surface-wave field produced by the magnetic current ring sheet, as shown in Fig. 4.4. This appendix will detail the derivation procedure from (4.20) to (4.21), described in Section 4.3.1. The derivation procedure is originally outlined in [36], and will be reproduced below.

As discussed in Section 4.3.1, the surface-wave radiation from a single magnetic current loop $K\left(\phi\right)$ can be obtained by integrating the fields of the magnetic Hertzian dipole over the ring current. The far-field approximation of the phase delay over the path length s of the surface wave radiation from each particular point on the ring yields the following factor:

$$e^{jk_{\mathrm{TM}_0}s} = e^{jk_{\mathrm{TM}_0}a\cos\left(\phi-\phi'\right)} \tag{A.2}$$

where ϕ is the observation angle and ϕ' is the angle of the particular source point on the current ring, shown in Fig. A.1. Note that the time convention $e^{j\omega t}$ is not explicitly shown, which is typical when dealing with time-harmonic fields. Unlike phase, we can assume the magnitude variation from points along the ring is negligible at far-field. Using (4.20) and (A.2), the surface wave fields of the magnetic current ring sheet can be obtained by integrating along the ring:

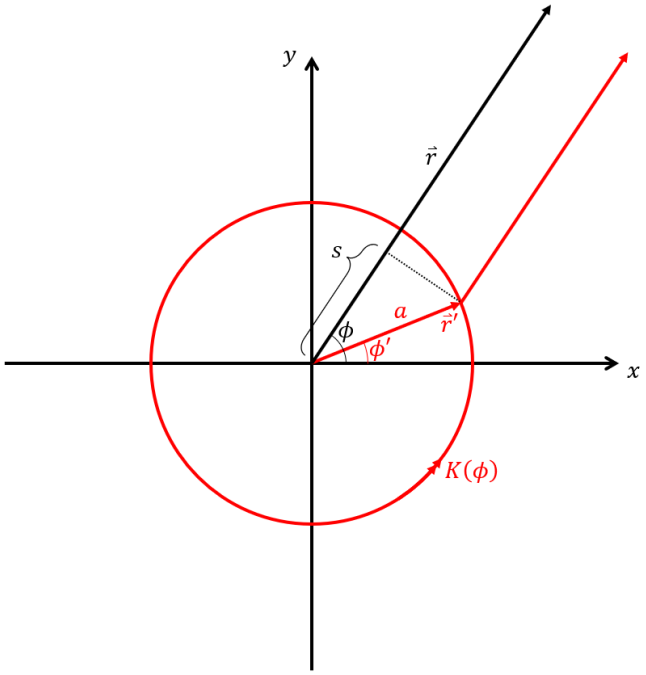


Figure A.1: Geometry used in the far-field surface wave phase approximation.

$$\psi \propto H_1^{(2)}\left(k_{\mathrm{TM}_0}\rho\right) B(z) \int_0^{2\pi} \cos(\phi') \cos(\phi - \phi') e^{jk_{\mathrm{TM}_0} a \cos(\phi - \phi')} a \, d\phi' \tag{A.3}$$

where $B(z)$ is a z-dependent amplitude factor defined as:

$$B(z) \equiv \int_0^h A(z, z_0)\, dz_0 \tag{A.4}$$

In (A.3), the $\cos(\phi')$ factor in the integrand comes from the magnetic current distribution $K(\phi)$ along the ring. The factor $\cos(\phi - \phi')$, which can be obtained from $\hat{r} \cdot \hat{r}'$ using the geometry shown in Fig. A.1, accounts for the orientation of the current at each point ϕ' along the ring with respect to the observation point at ϕ; this factor can be thought of as how the radiations from a Hertzian dipole at a particular ϕ' on the ring will appear to the observation point at ϕ.

To evaluate the integral of (A.3), apply the substitute variable $\alpha \equiv \phi' - \phi$ into the integral, we then have:

$$\psi \propto H_1^{(2)}\left(k_{\mathrm{TM}_0}\rho\right) B(z) \int_0^{2\pi} \left(\cos\phi\cos^2\alpha - \sin\phi\cos\alpha\sin\alpha\right) e^{jk_{\mathrm{TM}_0}a\cos\alpha} a\, d\alpha \tag{A.5}$$

We further apply the Jacobi-Anger expansion to the phase delay factor:

$$e^{jk_{\mathrm{TM}_0}a\cos\alpha} = J_0(k_{\mathrm{TM}_0}a) + 2\sum_{n=1}^{\infty} j^n J_n(k_{\mathrm{TM}_0}a)\cos(n\alpha) \tag{A.6}$$

Applying (A.6) into the integral and carrying out the integration in (A.5), it can be shown that only the $n = 2$ term in (A.6) will yield a non-zero result; we then obtain:

$$\psi \propto \pi a H_1^{(2)}\left(k_{\mathrm{TM}_0}\rho\right) B(z)\cos\phi\left(J_0(k_{\mathrm{TM}_0}a) - J_2(k_{\mathrm{TM}_0}a)\right) \tag{A.7}$$

Using the following property of the derivative of Bessel function:

$$J_1'(x) = \frac{1}{2}\left(J_0(x) - J_2(x)\right) \tag{A.8}$$

then applying it to (A.7), we have:

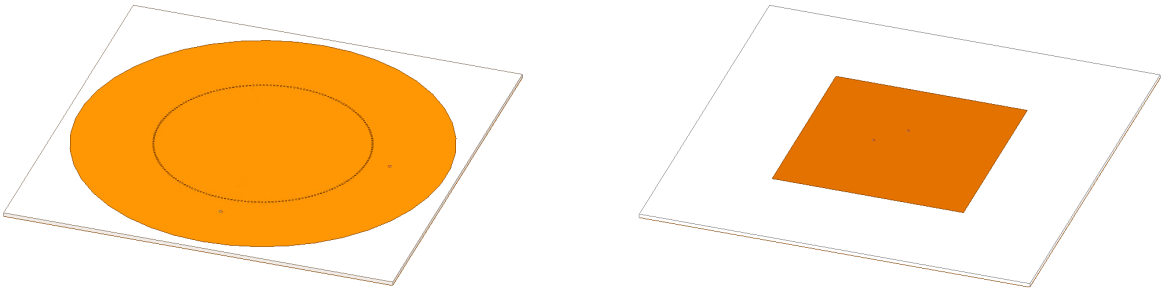
$$\psi \propto 2\pi a B(z) H_1^{(2)}\left(k_{\mathrm{TM}_0}\rho\right)\cos\phi J_1'(k_{\mathrm{TM}_0}a) \tag{A.9}$$

We thus complete the derivation of the surface wave field radiated by the equivalent magnetic ring current sheet, which is also the fundamental design equation of the SAR antenna as expressed in (4.21) of Chapter 4. Since only the $J_1'(k_{\mathrm{TM}_0}a)$ term contains a SAR antenna design parameter, the rest of the factors in (A.9) can be ignored. This terms suggests that for an ideal SAR antenna to radiate no surface wave, its patch radius a must be chosen such that the term $J_1'(k_{\mathrm{TM}_0}a)$ becomes zero; this is detailed in (4.22)–(4.25).

APPENDIX B

Platform Effects of SAR Patch Antenna: A Comparative Study

One of the main desirable features of a SAR patch antenna is its suppressed surface wave radiations. This allows less platform effects, discussed in Section 3.7.3, such as diffracted radiations from ground plane and substrate edges that may cause uneven radiation patterns (ripples) and higher backlobe and cross-polarization levels [7, 33]. We will compare the radiation characteristics of a L2-band SAR antenna with that of a conventional L2-band square patch antenna, shown in Fig. B.1, when they are installed on a larger metallic platform.



(a) (b)

Figure B.1: (a) A L2-band SAR patch antenna. The dimensions of the antenna is tuned so that it resonates at L2. Other characteristics, such as the substrate used, are the same as the L1-band SAR patch discussed in Section 5.1.2. (b) A conventional L2-band, dual-feed, circularly-polarized square patch antenna used for comparative purposes. The substrate material is F4B with the same dimensions of 150 mm $\times$ 150 mm $\times$ 1.524 mm as the substrate for the L2 SAR patch.

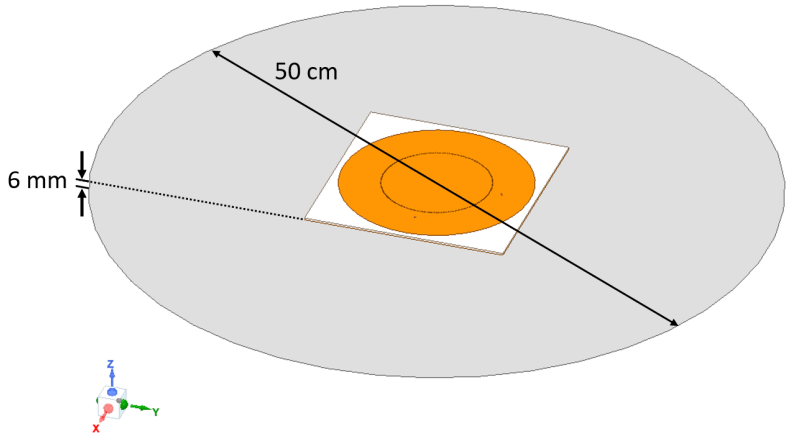
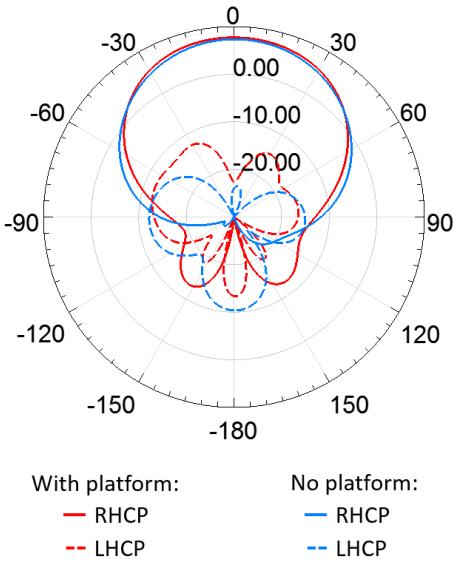


Figure B.2: The platform modeled as a metallic circular plate with a diameter of 50 cm and is positioned 6 mm below the antenna.

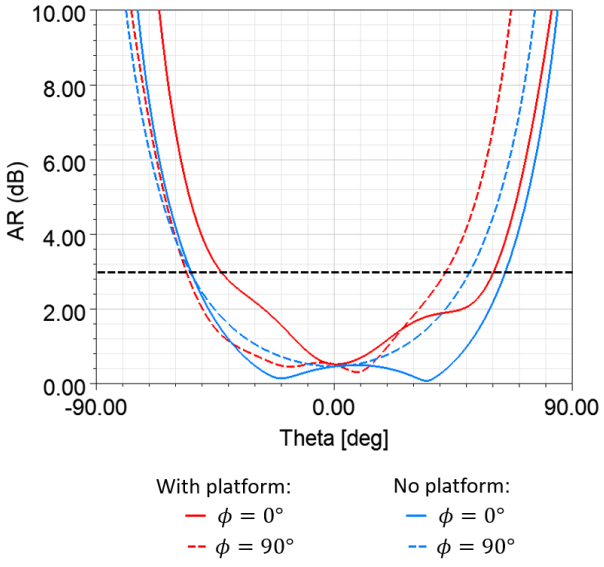
We will model the platform, on which the antenna is installed, as a metallic circular plate with a diameter of 50 cm and is positioned 6 mm below the antenna, shown in Fig. B.2. The simulated radiation patterns and axial ratios for the two antennas are shown in Fig. B.3; the plots compared the results with and without the platform effects. Compared to the conventional patch antenna, though not insignificant, the SAR patch results are much less affected by the platform effects in the presence of the large metallic ground plane platform. The theoretical treatment conducted in Chapter 4 gives a perfect TM_0 mode surface wave suppression, since we assumed a PMC boundary at the patch perimeter and therefore an ideal model of the equivalent magnetic ring current distributions. Ideally with no surface waves present, the expected results with and without the platform should differ very little. Realistically, however, this is not the case, since the patch perimeter is not a PMC and fringing fields exists, not to mention the existence of higher order surface wave modes depending on the substrate thickness. Thus, the formulated magnetic ring current model described in Chapter 4 is only an approximation. As a result, surface wave excitation will always exist.

Nonetheless, the simulated results shows that the conventional patch antenna is more severely affected by the platform effects, including a more reduced gain near the broadside, a much higher X-pol level, and a significantly more reduced AR beamwidth. Although this

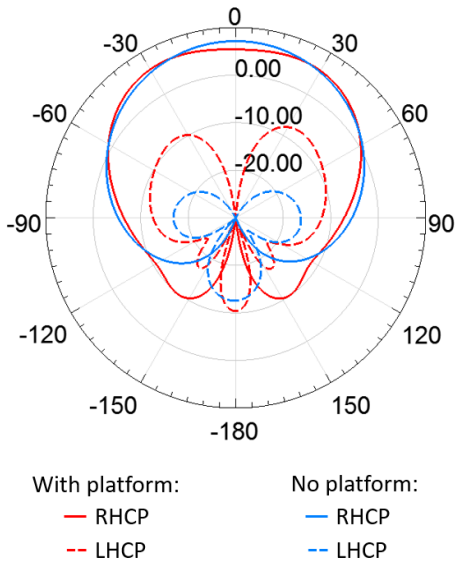
simulation is by no means comprehensive, since the antenna platform can be of a variety of sizes and shapes, it does demonstrate the advantage enjoyed by a SAR patch antenna when radiating in the presence of such particular metallic platform and likely similar ones.



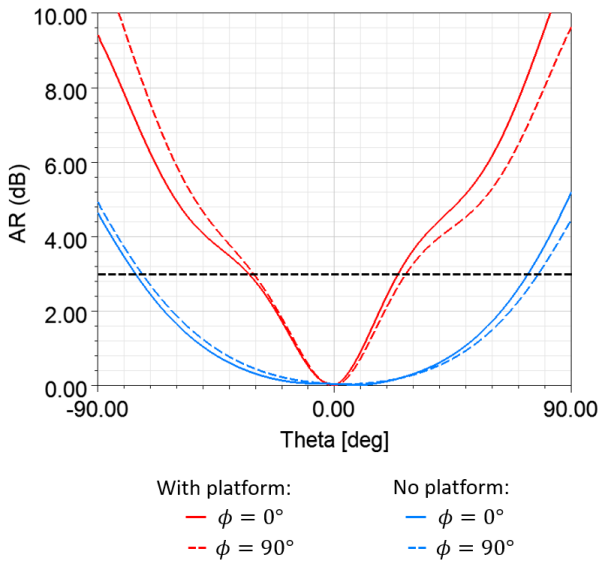
(a)



(b)



(c)



(d)

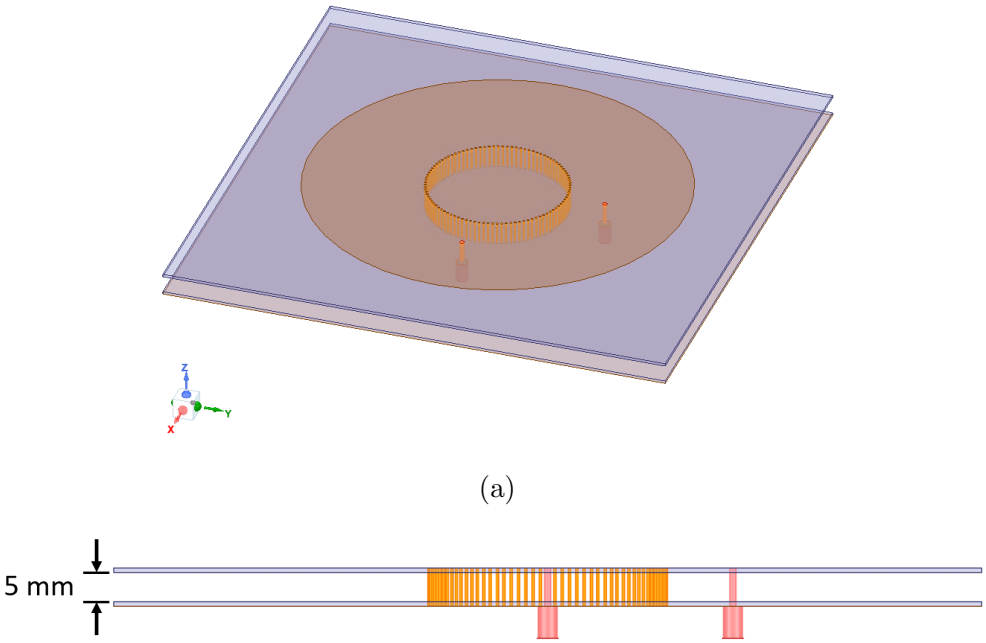
Figure B.3: Simulated (a) $\phi = 0°$ gain pattern and (b) AR of the L2-band SAR patch antenna shown in Fig. B.1a. Simulated (c) $\phi = 0°$ gain pattern and (d) AR of the L2-band square patch antenna shown in Fig. B.1b.

APPENDIX C

Bandwidth Enhancement Using Air Gap

Typically, increasing the thickness of the substrate is a reliable method of increasing the impedance bandwidth of a patch antenna. This, however, comes at a price of increasing the manufacturing cost. The cost of a patch antenna is expected to be proportional to the substrate size and thickness, which is confirmed by several vendors; a substrate with twice the thickness will then be roughly twice as expensive. Decreasing the substrate permittivity will also increase the impedance bandwidth. However, the current design uses the RO4003C substrate with a permittivity of 3.55, which is already not very high. Therefore, using a substrate material with a lower permittivity (for the same substrate thickness) will likely not result in a drastic improvement in the impedance bandwidth.

Another important note is that the SAR antenna is not a structurally flexible design, since it requires a fix patch size to satisfy the condition of (4.22) for surface wave suppression, and a shorting boundary for resonant frequency tuning. Hence, any potential design modifications on the patch for bandwidth improvement will be very limited without destroying the basic operation principles of a SAR patch. Nonetheless, incorporating an air gap layer is a known technique to improve the patch antenna impedance bandwidth [40]. Some SAR antenna designs studied in [41–43] have employed air gap layer for bandwidth enhancement purposes. This will provide an alternative method of bandwidth enhancement without resorting to the use of four feed probes, which is associated with a more complex feeding/receiving network. The air gap layer essentially acts as a low permittivity "substrate" and its thickness can be chosen arbitrarily. The circular patch can then be manufactured on a cheap substrate that is somehow suspended above the ground plane to realize the air gap layer. A representative structure is shown in Fig. C.1.



(a)

(b)

Figure C.1: (a) A 3D view of the SAR patch antenna design with an air gap layer. The substrates maintain the dimension of 150 mm × 150 mm as the designs shown in Chapter 5. (b) The front view of the design, showing a 5-mm-thick air gap layer between the patch and ground plane substrates.

The design shown in Fig. C.1 uses two FR4 substrate layers with thickness of 0.8 mm, one for the patch and one for the ground plane; they are separated by a 5-mm-thick air gap layer. The shorting boundary used in the simulation is still realized using a ring of vias; the vias connect the patch to the ground plane through both substrate and air gap layers as shown in the figure. In Table 5.1, we noted that FR4 substrate possesses a loss tangent that is too high for the L1-band single-layer design, such that the resulting gain is too low compared to other substrate options. However, as we will show later, FR4 substrate is viable for the air gap design, since the effective substrate loss is dropped considerably when factoring in the air gap layer. Additionally, the electric field strength between the patch and the ground plane is decreased as the thickness of the substrate (factoring in the air gap thickness) increases; this will also decrease the dielectric loss, which is proportional to the

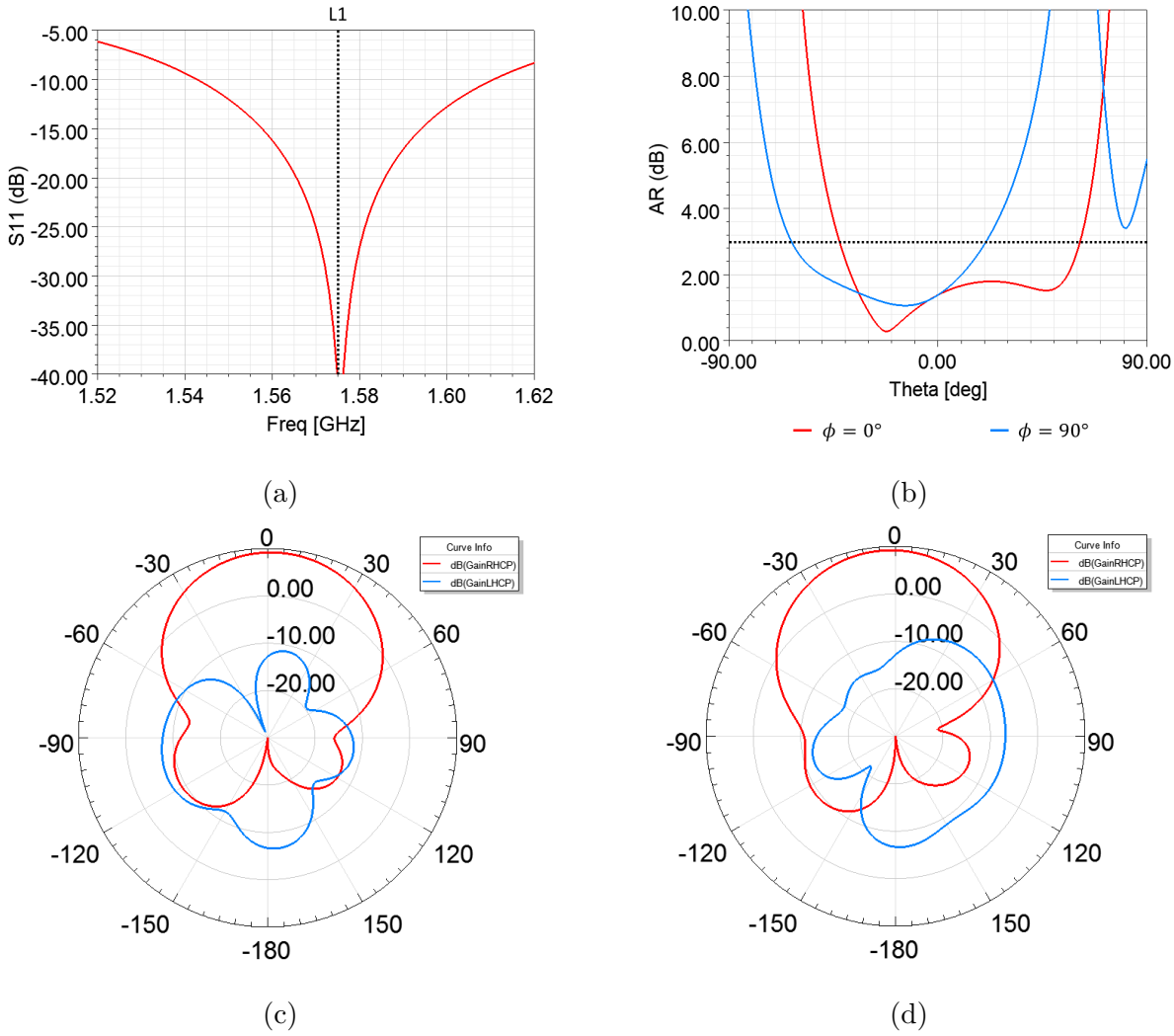


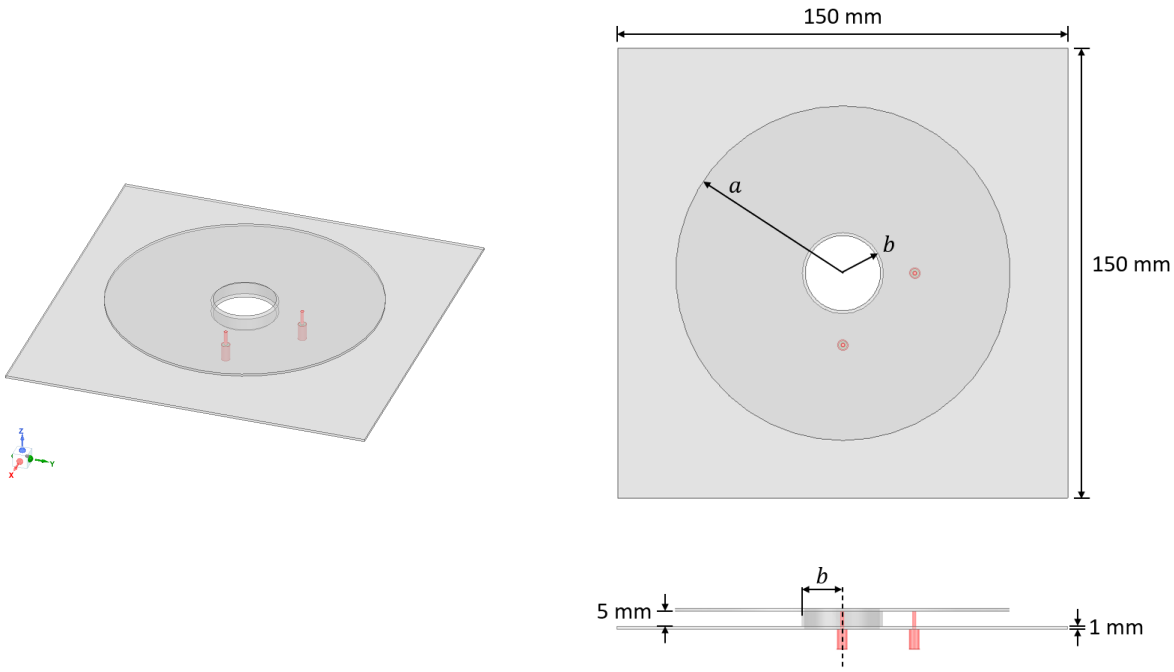
Figure C.2: Simulated (a) S_{11}, (b) AR versus polar angle θ, (c) gain pattern at $\phi = 0°$, and (d) gain pattern at $\phi = 90°$ of the air gap design shown in Fig. C.1.

field strength. Hence, the resulting gain will not exhibit a large drop from the directivity value as shown in Table 5.1.

The simulated results are shown in Fig. C.2. Clearly from Fig. C.2a, the air gap layer improved the impedance bandwidth to more than 70 MHz. However, the increased thickness affects the radiation characteristics in some undesirable way. For example, the antenna exhibits more backlobe radiation and radiations toward the horizon compared to the previous designs studied in Chapter 5. The cross-polarization levels are also significantly higher than the previous designs, though the level is below the minimum required gain for GPS application. The axial ratio still exhibits an appreciable beamwidth. However, it is much less symmetric, especially in the $\phi = 90^\circ$ plane caused by the rather asymmetrical X-pol level. The asymmetry in the radiation and AR results can be resolved by using the symmetric four-feed design. However, using the four-feed configuration can increase the impedance bandwidth as shown in Chapter 5, which defeats the purpose of the air gap layer.

There is a potential risk of increasing the substrate thickness to a point that it will introduce higher order surface wave modes, such as the TE_1 mode, as discussed in Chapter 4. For the air gap design, the equivalent substrate (including the air gap and the two FR4 substrate layers) has a total thickness of 6.6 mm. A simple estimation of the equivalent relative permittivity using series capacitance calculation yields a value of 1.2626; the value is expected to be closer to 1 since the air gap occupies a majority of the space. We then choose a maximum operating frequency of 1.61 GHz according to Fig. C.2a, then using (4.4), the maximum height for the equivalent substrate to not excite higher order modes is roughly 7.9 mm. The total thickness of 6.6 mm for the air gap and substrates is thus acceptable. Nonetheless, it should be noted that this calculation is based on the ideal theoretical treatment using the cavity model. Realistically, as explained in Appendix A, there will always be surface wave excitation since the patch perimeter is not a PMC and exhibits fringing fields. Thicker substrates will also result in more fringing, which implies larger deviations from the ideal model, and hence the analysis becomes less accurate.

The presence of an air gap layer also poses a significant problem to traditional PCB manufacturing procedures. Consider the structure shown in Fig. C.1b, for the most part,

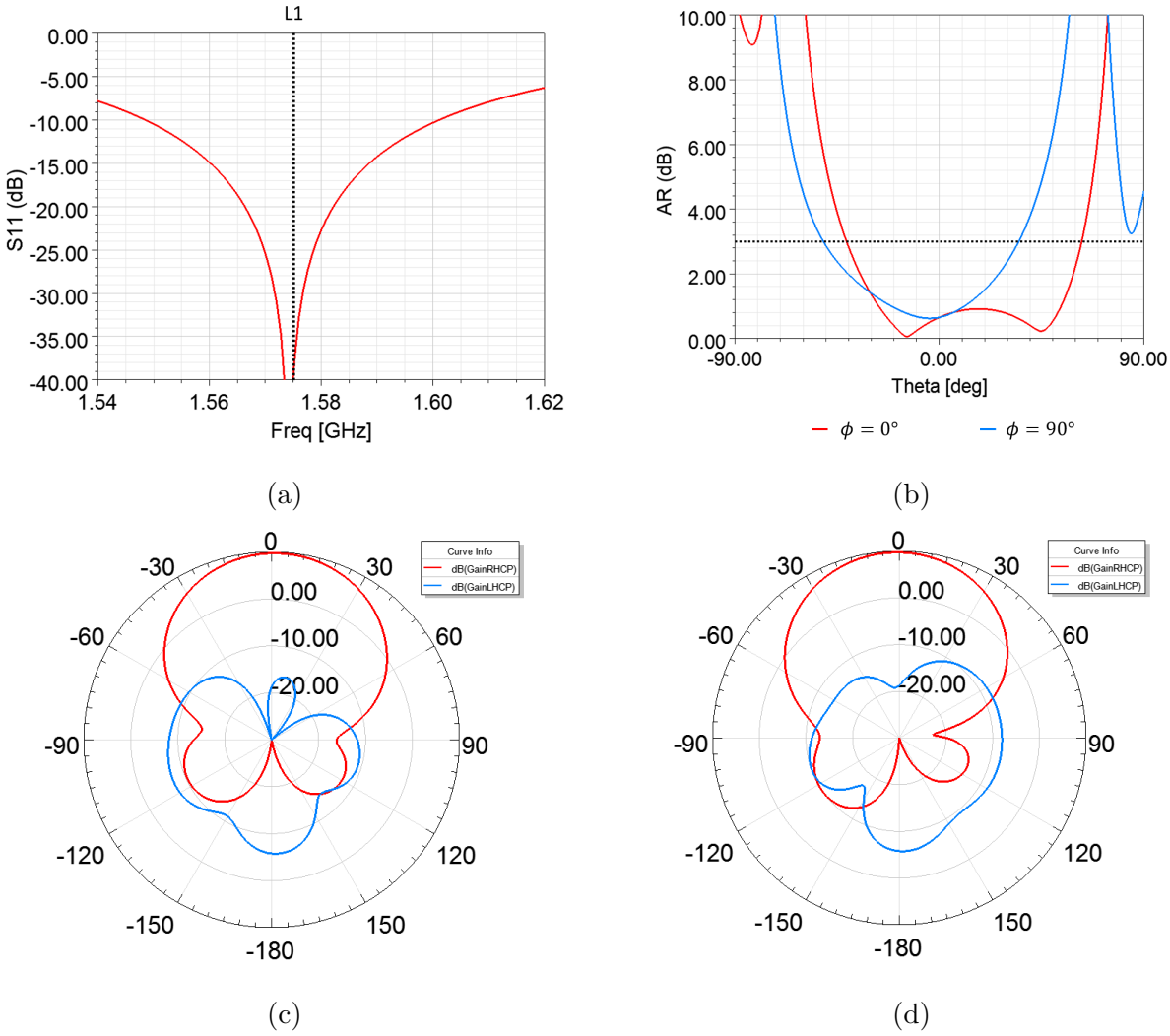


(a) (b)

Figure C.3: (a) A 3D view of the SAR patch antenna "all-metal" design with an air gap layer. (b) The top and front views of the design. Note the thickness of the metal is chosen to be 1 mm. The air gap is maintained at a 5-mm thickness.

it is highly unlikely that the through-air-gap vias are manufacturable through conventional PCB vendors, since traditional vias are not metal cylinders but are instead plated on the PCB through-holes. It is not explicitly explained in [42, 43] how the through-air-gap shorting vias are added, but they are very likely to be hand-placed and soldered cylindrical metal pins. Alternatively, there are SAR antenna designs that uses a cylindrical metal surface as the shorting boundary, such as in [41, 44, 45]. Again, such metal cylinder of the required radius and height is not realizable through conventional PCB vendors. In other words, realizing either the through-air-gap vias or the metal cylinder design will require either hand-made or alternative manufacturing processes. Since the PCB process is not viable, a SAR patch antenna design without using PCB substrates is simulated; such "all-metal design" is shown in Fig. C.3.

The thickness of the patch, ground plane, and shorting cylinder are chosen to be 1 mm,



(a) (b)

(c) (d)

Figure C.4: Simulated (a) S_{11}, (b) AR versus polar angle θ, (c) gain pattern at $\phi = 0°$, and (d) gain pattern at $\phi = 90°$ of the "all-metal" design shown in Fig. C.3.

and the air gap is maintained at 5-mm. The simulated results are shown in Fig. C.4. An impedance bandwidth of more than 50 MHz is achieved. The simulated radiation pattern and AR results are largely similar to those of the air gap design, but with relatively lower X-pol levels. The manufacturing options of such all-metal antenna structure should be further investigated. Regardless of the manufacturing process, one of the key requirement is high precision, since the resonant frequency is highly sensitive to the radius of the shorting boundary b. One potential option is through metal 3D printing, though the potential high cost may be a concern and thus vendor options should be carefully evaluated in the future.

APPENDIX D

Evaluation of UCLA Spherical Near-Field Antenna Measurement Range at L-Band GPS Frequencies

The L1-band SAR patch antenna measurements shown in Section 5.1.6 are conducted in the spherical near-field antenna measurement chamber at the Center for High Frequency Electronics (CHFE) of UCLA. The model of the open-ended waveguide (OEWG) probe used in the measurements is WR770 by NSI-MI, shown in Fig. D.1. The WR770 probe has a nominal frequency range of 0.96–1.45 GHz, and therefore, the L1 frequency of 1575.42 MHz is out of range. Sadly, the L1 frequency is also outside the operating range of the other available probes at the measurement chamber. The WR770 probe was then chosen for its frequency range being the closest to the L1 frequency. Operating beyond the specified upper frequency limit may have the adverse effect of exciting unwanted higher-order modes, and hence, producing inaccurate measurement results. Although L1 frequency is not drastically

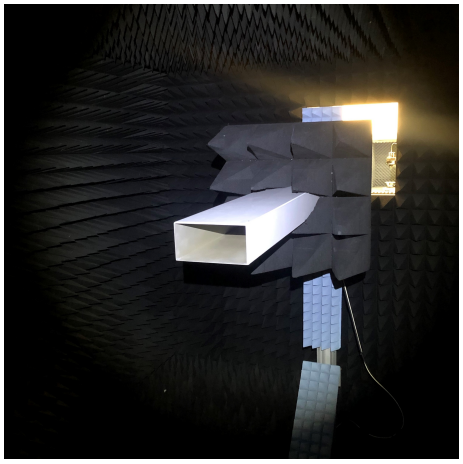


Figure D.1: The WR770 OEWG probe used in the L1- & L2-band measurements at UCLA's CHFE.

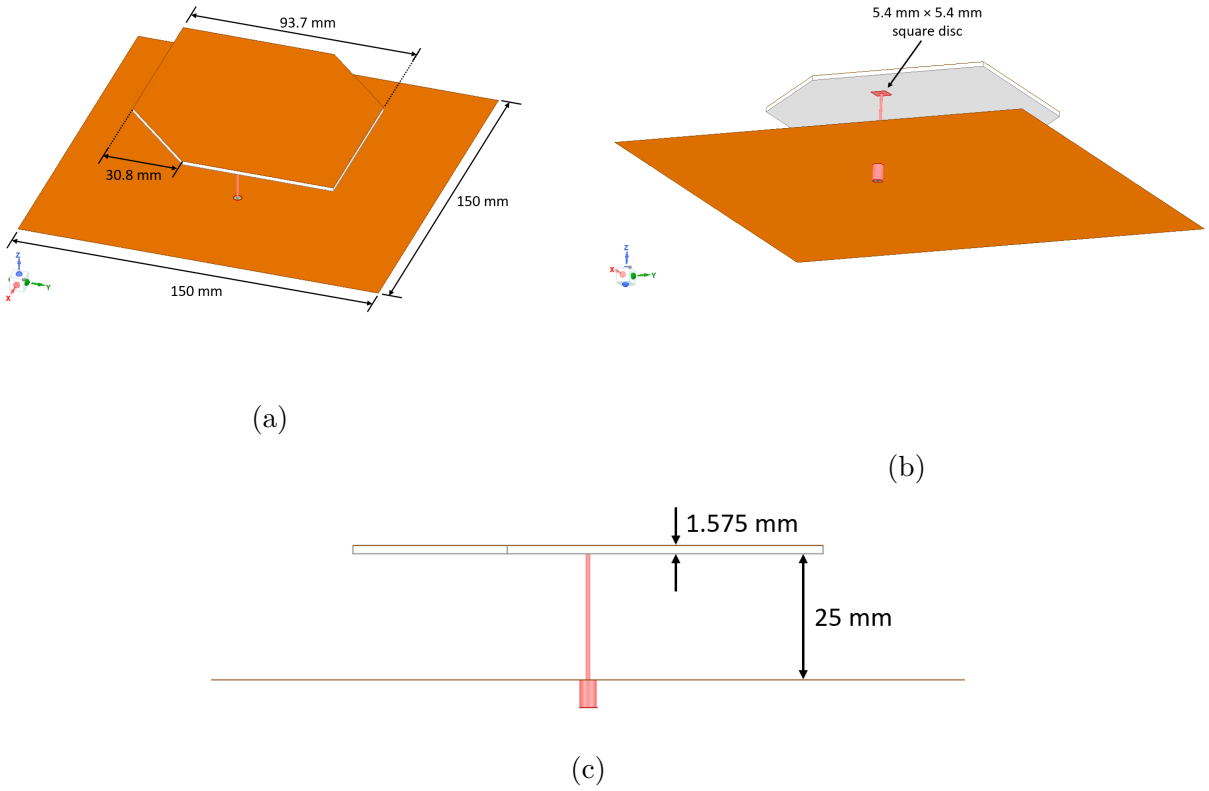
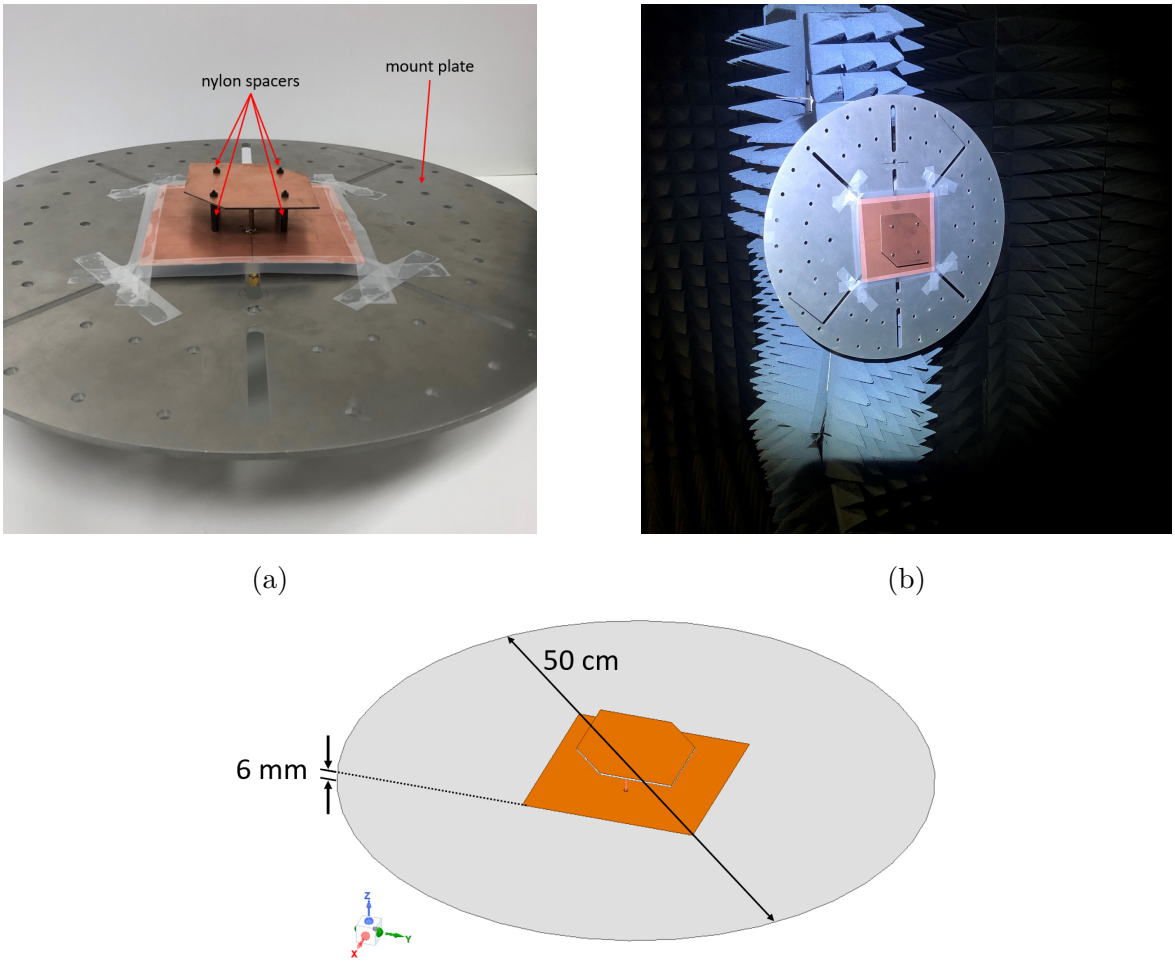


Figure D.2: The (a) top view, (b) bottom view, and (c) front view of the truncated-corner patch antenna for probe measurement assessment.

higher than the upper limit, it is still prudent to verify the measurement accuracy.

A simple patch antenna design is then constructed, as shown in Fig. D.2, to verify the WR770 probe measurement accuracy at L-band GPS frequencies, in particular L1. The antenna is a L2-band RHCP truncated-corner patch antenna. It features a 25-mm-thick air gap layer that enables a wide impedance bandwidth. The thick air gap layer necessitates the use of a long coaxial probe, which introduces significant inductance. The impedance matching can be achieved through a proximity-coupled feed using a square disc at the bottom of the antenna substrate, shown in Fig. D.2b; the disc introduces a series capacitance that compensates for the series inductance of the long coaxial probe [46, 47]. One primary reason for this design choice is that the antenna can be constructed entirely by hand, without using external vendors, and there are no precision etching or drilling involved. The wide impedance



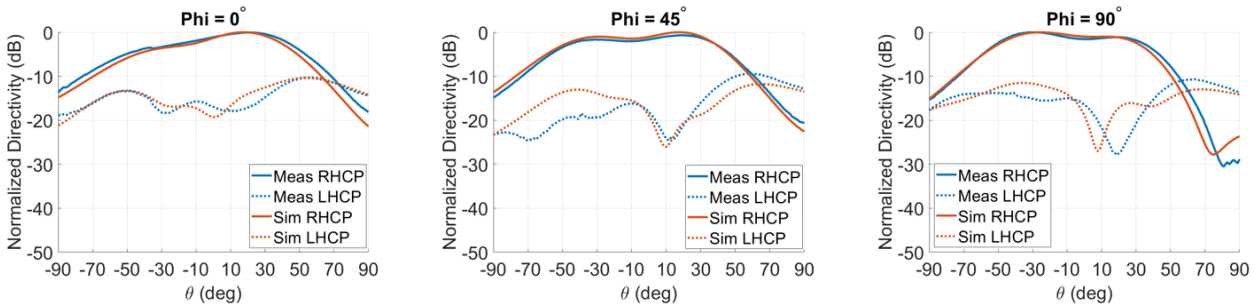
(a) (b)

(c)

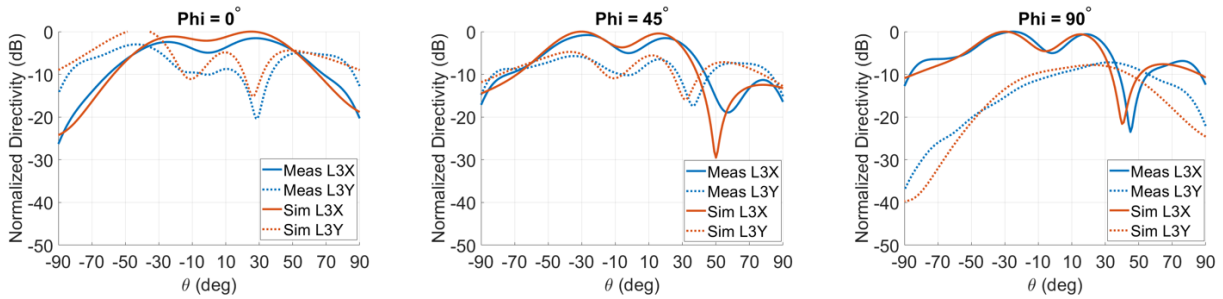
Figure D.3: (a) Fabricated test antenna. (b) Fabricated antenna mounted in the measurement chamber. (c) The simulation model includes the mount plate that introduces platform effects and affects the measured results.

bandwidth also ensures any potential errors during hand-fabrication will not drastically affect the antenna resonance. One important thing to note is that the performance of the antenna is not a concern, since it is used for assessment of the measurement accuracy. In other words, we are only concerned with the close agreement between simulated and measurement results, and not the performance itself.

The hand-constructed antenna is shown in Fig. D.3. The truncated-corner patch is directly cut from a piece of 1.575-mm-thick RT/duroid 5870 substrate, and the copper on



(a)



(b)

Figure D.4: Measured and simulated radiation patterns at (a) L2 and (b) L1.

the back side is removed entirely using chemical etching solutions. The square disc on the back side of the patch was realized using copper tape. The patch structure is supported above the ground plane using four 25-mm-long nylon spacers, which also realizes the air gap. The mounting platform available in the measurement chamber is a circular plate with a diameter of 50 cm, which is accounted for in the HFSS simulation model.

The measured and simulated radiation patterns at both L2- and L1-bands are shown in Fig. D.4. At L2, the measured and simulated results matches very well, which also suggests that any error introduces during hand-fabrication process is not significant. Note that the L1 measurements are shown in linear polarization in Ludwig-3 definition, since circular polarization is not expected to be supported at L1 for the L2-band patch. The measured and simulated results at L1 also closely agree with each other, even though L1 frequency is slightly beyond the nominal operating range of WR770 probe. The measurement results thus suggest that the use of WR770 probe for measurements at L1 frequency is acceptable.

REFERENCES

[1] J. J. H. Wang, "Antennas for global navigation satellite system (gnss)," *Proceedings of the IEEE*, vol. 100, no. 7, pp. 2349–2355, 2012.

[2] E. D. Kaplan and C. J. Hegarty, *Understanding GPS/GNSS: Principles and Applications*, 2nd ed. Artech House, 2006.

[3] X. Chen, C. G. Parini, B. Collins, Y. Yao, and M. U. Rehman, *Antennas for Global Navigation Satellite Systems*, 2nd ed. John Wiley & Sons, 2012.

[4] A. El-Rabbany, *Introduction to GPS: The Global Positioning System.* Artech House, 2002.

[5] "Parallel systems," https://moveparallel.com/, accessed: 01-30-2022.

[6] P. Misra and P. Enge, *Global Positioning System: Signals, Measurements, and Performance*, 2nd ed. Ganga-Jamuna Press, 2006.

[7] B. Rama Rao, W. Kunysz, R. Fante, and K. McDonald, *GPS/GNSS Antennas.* Artech House, 2013.

[8] "Gps applications," https://www.gps.gov/applications, accessed: 01-30-2022.

[9] "Space segment," https://www.gps.gov/systems/gps/space, accessed: 01-30-2022.

[10] "Expandable 24-slot satellite constellation, as defined in the sps performance standard." https://www.gps.gov/systems/gps/space/, accessed: 01-30-2022.

[11] "Control segment," https://www.gps.gov/systems/gps/control, accessed: 01-30-2022.

[12] "Is-gps-200m," https://www.gps.gov/technical/icwg/IS-GPS-200M.pdf, accessed: 01-30-2022.

[13] "New civil signals," https://www.gps.gov/systems/gps/modernization/civilsignals, accessed: 01-30-2022.

[14] W. A. Imbriale, S. S. Gao, and L. Boccia, *Space Antenna Handbook.* John Wiley & Sons, 2012.

[15] "Galileo frequency bands," https://galileognss.eu/galileo-frequency-bands, accessed: 01-30-2022.

[16] "Gnss constellations, radio frequencies and signals," https://www.tallysman.com/gnss-constellations-radio-frequencies-and-signals/, accessed: 01-30-2022.

[17] C. A. Balanis, *Advanced Engineering Electromagnetics.* John Wiley & Sons, 2012.

[18] "Global positioning system (gps)," https://www.got.nv.gov/home/showpublisheddocument/9212/636451509988270000, accessed: 01-30-2022.

[19] A. Kumar, A. D. Sarma, E. Ansari, and K. Yedukondalu, "Improved phase center estimation for gnss patch antenna," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 4, pp. 1909–1915, 2013.

[20] W. Kunysz, "Antenna phase center effects and measurements in gnss ranging applications," in *2010 14th International Symposium on Antenna Technology and Applied Electromagnetics the American Electromagnetics Conference*, 2010, pp. 1–4.

[21] C. A. Balanis, *Antenna Theory: Analysis and Design*. John Wiley & Sons, 2016.

[22] G. Moernaut and G. Vandenbosch, "The shorted annular patch as a multi-path suppressing antenna," in *The Second European Conference on Antennas and Propagation, EuCAP 2007*, 2007, pp. 1–6.

[23] "Gpdf254.a – passive dual pin gnss patch antenna 25*25*4 mm," https://www.taoglas.com/product/passive-dual-pin-gnss-patch-antenna/, accessed: 01-30-2022.

[24] "Oem gnss antenna hx-cs7017a," https://en.harxon.com/products-detail.php?Prold=129, accessed: 01-30-2022.

[25] "Oem gnss antenna hx-csx078a," https://en.harxon.com/products-detail.php?Prold=136, accessed: 01-30-2022.

[26] "Hc771e single-band helical antenna," https://www.tallysman.com/product/hc771e-embedded-single-band-helical-antenna/, accessed: 01-30-2022.

[27] S. Yun, D. Chang, and I. Yom, "Wideband receive antennas of sensor stations for gps/galileo satellite," in *2008 4th Advanced Satellite Mobile Systems*, 2008, pp. 210–212.

[28] D. G. DeGryse, "Gps modernization and the path forward: Bringing new capabilities to military and civil users worldwide," https://www.afspc.af.mil/Portals/3/documents/HF/AFD-080522-087.pdf, pp. 15–17, 2008, accessed: 01-30-2022.

[29] "2d choke ring antenna hx-cgx606a," https://en.harxon.com/products-detail.php?Prold=101, accessed: 01-30-2022.

[30] B. Mohajer-Iravani, S. Shahparnia, and O. Ramahi, "Coupling reduction in enclosures and cavities using electromagnetic band gap structures," *IEEE Transactions on Electromagnetic Compatibility*, vol. 48, no. 2, pp. 292–303, 2006.

[31] R. Hurtado, W. Klimczak *et al.*, "Artificial magnetic conductor technology reduces size and weight for precision gps antennas," in *Proceedings of the 2002 National Technical Meeting of The Institute of Navigation*, 2002, pp. 448–459.

[32] F. Yang and Y. Rahmat-Samii, *Electromagnetic Band Gap Structures in Antenna Engineering*. Cambridge University Press, 2008.

[33] R. B. Waterhouse, *Microstrip Patch Antennas: A Designer's Guide*. Springer Science+Business Media, 203.

[34] W. A. Imbriale, S. S. Gao, and L. Boccia, *Space Antenna Handbook*. John Wiley & Sons, 2012.

[35] R. A. Sainati, *CAD of Microstrip Antennas for Wireless Applications*. Artech House, 1996.

[36] D. Jackson, J. Williams, A. Bhattacharyya, R. Smith, S. Buchheit, and S. Long, "Microstrip patch designs that do not excite surface waves," *IEEE Transactions on Antennas and Propagation*, vol. 41, no. 8, pp. 1026–1037, 1993.

[37] L. I. Basilio, J. T. Williams, D. R. Jackson, and M. A. Khayat, "A comparative study of a new gps reduced-surface-wave antenna," *IEEE Antennas and wireless propagation letters*, vol. 4, pp. 233–236, 2005.

[38] L. I. Basilio, R. L. Chen, J. T. Williams, and D. R. Jackson, "A new planar dual-band gps antenna designed for reduced susceptibility to low-angle multipath," *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 8, pp. 2358–2366, 2007.

[39] "Ro4000® series high frequency circuit materials," https://rogerscorp.com/-/media/project/rogerscorp/documents/advanced-electronics-solutions/english/data-sheets/ro4000-laminates-ro4003c-and-ro4350b---data-sheet.pdf, accessed: 01-30-2022.

[40] V. Sharma, B. Sharma, V. Saxena, K. Sharma, and D. Bhatnagar, "Modified rectangular patch antenna with air-gap for improved bandwidth," in *2008 International Conference on Recent Advances in Microwave Theory and Applications*, 2008, pp. 227–229.

[41] Boccia, Amendola, and D. Massa, "A dual frequency microstrip patch antenna for high-precision gps applications," *IEEE Antennas and Wireless Propagation Letters*, vol. 3, pp. 157–160, 2004.

[42] Z. Wang, H. Liu, S.-J. Fang, and Y. Cao, "A low-cost dual-wideband active gnss antenna with low-angle multipath mitigation for vehicle applications," *Progress In Electromagnetics Research*, vol. 144, pp. 281–289, 2014.

[43] H. Liu, S. Fang, and Z. Wang, "Novel dual-band antenna for multi-mode gnss applications," *Journal of Systems Engineering and Electronics*, vol. 26, no. 1, pp. 19–25, 2015.

[44] L. Boccia, G. Amendola, G. Di Massa, and L. Giulicchi, "Shorted annular patch antennas for multipath rejection in gps-based attitude determination systems," *Microwave and Optical Technology Letters*, vol. 28, no. 1, pp. 47–51, 2001.

[45] L. Boccia, G. Amendola, and G. Di Massa, "Performance evaluation of shorted annular patch antennas for high-precision gps systems," *IET Microwaves, Antennas & Propagation*, vol. 1, no. 2, pp. 465–471, 2007.

[46] K. Fong, H. Pues, and M. Withers, "Wideband multilayer coaxial-fed microstrip antenna element," *Electronics letters*, vol. 21, no. 11, pp. 497–499, 1985.

[47] J. Ma, A. B. Kouki, and R. Landry Jr, "Wideband circularly polarized single probe-fed patch antenna," *Microwave and Optical Technology Letters*, vol. 54, no. 8, pp. 1803–1808, 2012.