2015-11-01

A Compact Phased Array Radar for UAS Sense and Avoid

Jonathan Cullinan Spencer
Brigham Young University - Provo

Follow this and additional works at: https://scholarsarchive.byu.edu/etd

Part of the Electrical and Computer Engineering Commons

BYU ScholarsArchive Citation
https://scholarsarchive.byu.edu/etd/5837

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in All Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.
ABSTRACT
A Compact Phased Array Radar for UAS Sense and Avoid

Jonathan Cullinan Spencer
Department of Electrical and Computer Engineering, BYU
Master of Science

As small unmanned aerial systems (UAS) are introduced into the national airspace, measures must be introduced to ensure that they do not interfere with manned aviation and other UAS. Radar provides an attractive solution because of its inherent range accuracy and because it works in diverse weather and lighting conditions. Traditional radar systems, however, are large and high power and do not meet the size, weight and power (SWaP) constraints imposed by UAS, and fully integrated automotive solution do not provide the necessary range. This thesis proposes a compact radar system that meets both the SWaP and range requirements for UAS and can act as a standalone sensor for a sense and avoid system (SAA).

The system meets the field of view requirements motivated by the UAS sensing problem ($120^\circ \times 30^\circ$) and tracks targets in range and azimuthal angle using a four element phased array receiver. The phased array receiver implements real time correlation and beamforming using a field programmable gate array (FPGA) and can track multiple targets simultaneously. Excluding antennas, the radar transceiver and signal processing platform weighs approximately 120g and is approximately the size of a whiteboard eraser ($2.25in \times 4in \times 1in$), which meets the payload requirements of many small (<25kg) UAS. To our knowledge, this is the first real time phased array radar that meets the sensing and SWaP requirements for small UAS.

Our testing was done with the radar system on the ground, aimed at airborne UAS targets. Using antennas with a gain of 12 dB, and 800 milliwatts of transmitted power, the system detects UAS targets with a radar cross section of less than 0.1 m$^2$ up to 150 meters away. The ground based system demonstrates radar detectability of extremely small UAS targets, and is scalable to further ranges by increasing antenna gain or adding additional elements. Based on our success in detecting airborne UAS, we conclude that radar remains a feasible option for a UAS collision avoidance sensor.

Keywords: FMCW, Radar, Phased Array, UAV, UAS, Sense and Avoid
ACKNOWLEDGMENTS

The research published in this thesis is the result of collaboration with several other students. I’m ever so grateful for their involvement and friendship. To name a few: Kaleo Roberts, Michael Boren, Laith Sahawneh, Jared Wikle, and Luke Newmeyer.

I’m also grateful for other current and past students who offered friendship and advice on both personal and technical topics. Notably: James Mackie, James Eck, Kyle Browning, Matt Morin, Ben Arnold, Richard Black, Zhenchao Yang, and Junming Diao.

I’m grateful for the BYU Electrical Engineering faculty who have instructed me, given me counsel and treated me like a peer. It’s been so fun developing friendships that extended outside of campus, especially cycling with my committee member and friend Randy Beard.

I’m particularly grateful for my advisor Karl Warnick. His leadership style helped me grow as a student and a person, and his light hearted demeanor made learning a joy.

My biggest cheerleaders have been and always will be my dear family. Bless them.
TABLE OF CONTENTS

LIST OF TABLES ............................................................... vi

LIST OF FIGURES .......................................................... vii

Chapter 1 Introduction ..................................................... 1
  1.1 Motivation ............................................................ 1
  1.2 Literature Review .................................................... 3
  1.3 Contributions ........................................................ 5
  1.4 Overview ............................................................. 5

Chapter 2 Background ...................................................... 7
  2.1 FMCW Radar .......................................................... 7
     2.1.1 Stationary Point Target ...................................... 11
     2.1.2 Moving Targets ................................................ 14
     2.1.3 Received Power ................................................ 18
     2.1.4 Noise ............................................................... 20
  2.2 Phased Array Antennas .............................................. 27

Chapter 3 Radar System Design .......................................... 33
  3.1 Parameter Selection ................................................ 33
     3.1.1 Carrier Frequency ............................................. 34
     3.1.2 Chirp Bandwidth ................................................. 37
     3.1.3 Intermediate Frequency Bandwidth ......................... 38
     3.1.4 Chirp Period ...................................................... 39
  3.2 Transmitter .......................................................... 40
     3.2.1 Direct Digital Synthesizer .................................... 40
     3.2.2 Current Architecture: Phase Locked Loop .................. 42
     3.2.3 Power Amplifier and Antennas .............................. 48
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3</td>
<td>Receiver</td>
<td>50</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Homodyne Architecture</td>
<td>51</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Radio Frequency Subsystem</td>
<td>51</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Intermediate Frequency Subsystem</td>
<td>54</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Link Budget</td>
<td>56</td>
</tr>
</tbody>
</table>

**Chapter 4  Digital Signal Processing**  
4.1 Data Acquisition, Windowing and Fourier Transformation  
4.2 Correlation and Integration  
4.3 Beamforming  
4.4 Thresholding and Clustering  
4.5 Target Tracking, Risk Assessment, and Collision Avoidance

**Chapter 5  Experimental Results**  
5.1 Experimental Setup  
5.2 Stationary Target - Single Channel  
5.3 Moving Target - Single Channel  
5.4 Moving Target - Phased Array

**Chapter 6  Conclusions**  
6.1 Future Work

REFERENCES

Appendix A  Noise Figure in Op Amps
LIST OF TABLES

3.1 Radar Transceiver Parameters ................................................. 33
3.2 Calculated link budget for a single receiver channel of the system. ........... 57
2.1 The triangle FMCW chirp (a) and the sawtooth FMCW chirp (b) are the most common waveforms used in FMCW. They consist of a sine wave whose frequency increases or decreases at a constant rate. The frequency chirp is seen in the spectrograms (a) and (b) as a straight line with a constant slope. The frequency chirp is shown in the time domain for the sawtooth chirp in Fig. (c). .............................................................. 8

2.2 A typical test setup for the radar system. The radar system is stationary, and is aimed either at the ground or the sky, depending on what is being detected. In this case, the targets are two simple UAVs hovering in a stationary position. The transmitted waves are shown in blue, and the reflected echoes off the targets are shown in red. This figure corresponds with the data shown in Fig. 2.4. 9

2.3 A simplified block diagram of an FMCW radar system. .......................... 9

2.4 The signal flow through an FMCW radar system for two point targets in the setup shown in Fig. 2.2. The plots are scaled to show the returns from a single transmitted chirp. .............................................................. 10

2.5 Doppler shifted FMCW waveforms. Figs. (a) and (b) show the RF spectra, with the transmitted waveform in blue and the received waveform in red. Both received waveforms experienced a positive Doppler shift \( f_d \), denoting a target that is approaching the radar. Figs. (c) and (d) show the homodyne IF output after the mixer. Fig. (c) shows the ability of triangle FMCW pulses to determine instantaneous range and velocity, while Fig. (d) shows that sawtooth FMCW pulses are left with an unresolvable range/velocity ambiguity. 15

2.6 Signal leakage paths in a homodyne radar system. (1) is the desired primary signal path that provides the local oscillator signal for downmixing. (2) shows bleedthrough from poor amplifier or antenna matching, and (3) shows bleedthrough between the transmit and receive antennas. In most CW systems, (3) is the primary leakage path and it is blocked by absorbers or reflectors. 26

2.7 Noise sources in an FMCW radar system. (1) is phase noise introduced by close-up signal leakage. (2) represents the signal waveforms from two targets in addition to the external noise which is present at the antenna outputs. (3) is the added thermal noise from the receiver electronics. .......................... 27

2.8 Standard architecture of a phased array receiver that implements analog beamforming. .............................................................. 28
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.9</td>
<td>Radiation pattern for a 10-element phased array. The antenna weights ( w_i ) have a linear phase increment that steers the pattern 40 degrees off boresight. The radiation pattern is the product of the array factor ( A(\theta, \phi) ) and the pattern of the identical individual elements, indicated in red.</td>
</tr>
<tr>
<td>2.10</td>
<td>Standard architecture of a phased array receiver that implements digital beamforming.</td>
</tr>
<tr>
<td>3.1</td>
<td>This graph shows the one way gaseous attenuation of electromagnetic waves in the atmosphere due to oxygen and hydrogen. Figure from [1].</td>
</tr>
<tr>
<td>3.2</td>
<td>The typical output frequency spectrum of a DDS. (1) indicates the fundamental, desired waveform, which is a frequency swept sine wave. (2) indicates all harmonics and sub harmonics, which are duplicates of the fundamental chirp at different frequencies, some sweeping in the opposite direction. (3) indicates spurious tone bleedthrough from the reference oscillator. (4) indicates the extremely tight filtering necessary to remove all spurs from a DDS waveform generator.</td>
</tr>
<tr>
<td>3.3</td>
<td>Spectrogram of the actual transmitted chirp generated by the PLL. Although the main portion of the chirp in (a) is extremely linear, the chirp does require approximately 80 ( \mu s ) of recovery time before the next chirp. This is shown in (b), which is a magnified version of the last portion of (a).</td>
</tr>
<tr>
<td>3.4</td>
<td>Final transmitter architecture based on a phase-locked loop waveform generator.</td>
</tr>
<tr>
<td>3.5</td>
<td>Wilkinson splitter design realized on a PCB using CPWG. This splitter performs within 0.5 dB of an ideal split (including connector losses) and provides 25dB of isolation at the frequency of interest.</td>
</tr>
<tr>
<td>3.6</td>
<td>HMC952 off-board power amplifier with heat sink and fan to help with thermal dissipation.</td>
</tr>
<tr>
<td>3.7</td>
<td>Planar Vivaldi antenna used as both the transmit antenna and the receiver antennas for the majority of testing.</td>
</tr>
<tr>
<td>3.8</td>
<td>Final homodyne receiver architecture.</td>
</tr>
<tr>
<td>3.9</td>
<td>Architecture and layout of intermediate frequency subsystem.</td>
</tr>
<tr>
<td>4.1</td>
<td>Block diagram of the radar digital signal processing. The BOARAC board acts as an interface between the radar transceiver and the DSP. The operations that involve a high data rate are performed on the MicroZed FPGA, and the operations that involve a relatively slow data rate are performed on the ARM processor.</td>
</tr>
</tbody>
</table>
4.2 The MicroZed board (left) is a commercially available off-the-shelf development board based on the Xilinx Zynq-7000 FPGA/Processor and performs all back-end processing for this system. The BOARAC board (right) is a custom interface board that includes the ADCs and SPI module routing to interface the radar transceiver with the MicroZed DSP board.

4.3 The signal flow of the correlation process after the $m^{th}$ chirp has been recorded. The signals from the $i^{th}$ FFT bins of each channel are gathered into the vector $x_i[m]$. The correlation matrix $R_i$ is formed using the outer product, then integrated over $N_{corr}$ time steps. The result is a collection of integrated correlation matrices, one for each FFT bin.

4.4 The simulated output of a four-channel phased array radar system with three targets.

4.5 The thresholding process takes the maximum value from each row of the beamformed image. These values are collected over multiple time steps to estimate the noise properties from which a CFAR threshold is calculated. The clusters which rise above the threshold are combined into a single target location estimate for each cluster.

5.1 The fully integrated radar system. The boards from top to bottom are: MicroZed processor board, BOARAC ADC and DAC board, IF filter and amplifier board, RF transceiver board.

5.2 Radar testing sites. The radar location is indicated with a white star and the approximate beam orientation is indicated with the orange gradient.

5.3 Radar setup configurations.

5.4 Different targets used in radar testing.

5.5 Comparison of radar return data when the antenna is angled at the sky vs. at a sphere reflector. Both plots are given for a single chirp and for $N=100$ chirps averaged.

5.6 Spectrogram of radar aimed at moving pedestrian target ($\sigma_{RCS} = 1m^2$) on the Clyde Building lawn. Faint targets can be seen between 100m and 120m, which are from a group of people playing catch. Strong targets can be seen consistently at 130m and 140m, which correspond to cars turning the street corner where the radar was angled. The bright flash corresponds to the strong return when the car is angled perpendicular to the radar beam.
5.7 Radar imagery of the flight path of an X8 UAV flying at Rock Canyon park. The radar was angled at 10° above the horizon and the UAV flew a slightly diagonal path in order to stay within the antenna beam. The pre-whitening filter used two frames of radar data in the Levinson algorithm to generate filter coefficients.

5.8 Received power surface map of the phased array data for a single time index.

5.9 Phased array radar response for an X8 UAV flying a GPS waypoint pattern of a straight line with a 10m × 10m square deviation at a velocity of 1m/s. The phased array images were thresholded and reduced to a single location estimate for every time step. The plots show the location estimates transformed from range and angle to cartesian coordinates where the x direction is cross range and the y direction is down range from the phased array boresight.

A.1 Ideal op-amp in inverting configuration with added noise elements.
CHAPTER 1. INTRODUCTION

1.1 Motivation

Radar has now existed in some form or another for a little over a century. At its core, radar is any system that reflects electromagnetic waves off objects in order to learn something about them. Radar systems can use echoes to determine the distance between the object and the observer, the velocity at which the object is traveling, or physical properties of the object like its shape or material makeup. The word radar initially derived from an acronym for RAdio Detection And Ranging, but the ubiquity of the acronym transformed radar into a word of its own.

Radar did not receive much attention until WWI and WWII, when it was discovered that radar could provide an early warning for intruding aircraft. During this period, it became the focus of major research by all the world powers, each hoping to gain a competitive advantage over the other. After the world wars, however, it stabilized and became the basis for modern air traffic control. Until recently, the majority of modern radar research was done by the military, developing very large phased array radar systems capable of detecting small, incoming missiles from hundreds of miles away.

Within the last decade, however, radar has experienced somewhat of a resurgence due to the advent of monolithic microwave integrated circuits or MMICs. Microwave circuits of the past were made up of bulky waveguides and heavy ferrite components, but now most microwave systems can be implemented on a single printed circuit board (PCB) with a few MMICs. The miniaturization of RF components and computing devices has now opened up potential applications for radar beyond military and commercial aeronautics. Radar is
now being used in short range applications, primarily in the automotive industry, but also in surveillance and motion detection.

The field of aeronautics has experienced a similar technological revival with the advent of unmanned aerial vehicles (UAVs). UAVs or drones have been used in military applications for the past decade, but the most recent applications of UAVs have been in the civilian sector. Amazon caught the world’s attention in December 2013 when it unveiled a plan to use drones to deliver packages to customers’ doorsteps. Since then, several companies including Google and Facebook have started projects to develop unmanned aerial systems (UAS) that integrate UAVs with guidance technology to perform specialty tasks. Low cost civilian UAS offer a viable solution to many problems of scale, including search and rescue, pipeline surveillance, agriculture, disaster site exploration, and more. However, in order for UAS to integrate smoothly into current manned air traffic, they must have a way of reliably avoiding collision with current air traffic and other UAS.

Aeronautical navigation falls within two classes: cooperative detection and non-cooperative detection. Cooperative detection systems such as the Traffic Collision Avoidance System (TCAS) and Automatic Dependent Surveillance-Broadcast (ADS-B) rely on communication between aircraft. Approaching aircraft use pre-defined communication schemes to transmit information such as their position, heading, and velocity. In TCAS, this information is given when interrogated, and in ADS-B, this information is openly broadcast on a regular interval. For the majority of manned air traffic there are regulations in place that require the implementation of a cooperative system, but most UAVs currently fall within a size and weight class that does not require them to implement these safety measures. This situation requires non-cooperative aircraft detection, which relies on sensors to provide information about intruding aircraft rather than communication from the aircraft themselves. Non-cooperative systems tend to offer less accuracy than cooperative systems, but they guarantee safe flight under more general circumstances [2].

Sensor requirements for non-cooperative navigation systems vary widely and are highly dependent on the application. For UAS sense and avoid (SAA) systems, the sensor should
be able to detect targets in diverse lighting conditions within a few kilometers. The need to operate the sensor in diverse weather and lighting conditions limits the potential solution to active sensors. An active sensor is any system that excites its surroundings with some sort of pulse, then uses the response to that excitation to determine properties of its surroundings. Ultrasound, LIDAR and radar are examples of active sensors, while optical or infrared cameras are considered passive sensors. Ultrasonic sensors are cost effective and can provide an efficient solution for indoor navigation, but their range is limited to 5-10 meters, making them unsuitable for outdoor UAS. LIDAR is also a promising solution that offers high range and angular accuracy, but systems designed for more than 100 meters do not meet the size, weight, and power (SWaP) requirements for small UAS. The advent of autonomous cars that use LIDAR as a primary sensor may drive research in this area, making LIDAR more feasible for aeronautical applications.

Radar has performed so well for aeronautics over the past fifty years because it is inherently good at giving precise range estimates in diverse lighting and weather conditions. However, the current aeronautical radar systems are much too large to be used in UAS and automotive radar systems do not provide sufficient range [3]. The goal of this research is to show that radar is a good option for UAS by creating a radar system that meets the sensing and payload requirements of small UAS.

1.2 Literature Review

Radar is just one of many viable sensor options for UAS. Contarino, et al. provide a nice summary of sensor possibilities for larger aircraft and compare the advantages and disadvantages associated with each modality [2]. The majority of work in UAV collision avoidance is done using vision based or electro-optical sensors because they meet the SWaP restrictions imposed by UAS. Computer vision techniques allow the vision data to be exploited both for flight stability [4] and collision avoidance [5]. However, radar is a very attractive alternative to vision based sensors because it provides long range sensing in diverse weather and lighting conditions.
Since it is more than 50 years old, radar is a very mature technology that is well represented in the literature [1, 6]. However, progress in microwave integrated circuits has introduced fresh work on the subject, producing a full frequency modulated continuous wave (FMCW) radar transceiver on a single integrated circuit [7]. In addition to miniaturization, the quality of FMCW radar systems has also improved because of advancements in data converter technology, which make use of direct digital synthesis and phase locked loops to produce linear chirps with low phase noise [8].

Because autonomous unmanned aerial systems have become a widely researched topic in just the last decade, the use of radar in this application is a relatively young field of research. Using radar simulations, Kwag, et al. determined that FMCW radar is a potentially good option for UAV collision avoidance [9]. The group led by Kemkemian at Thales Airborne Systems in France did an in-depth study of the different radar possibilities and concluded that an FMCW radar at X band provides the most ideal characteristics for UAV sense and avoid systems [10]. Kemkemian, et al. proposed a MIMO radar architecture in order to determine angle of arrival information in elevation and azimuth [11, 12].

Some of the first airborne test results for sense and avoid were shown by Itcia, et al., who implemented an SAA radar on a Cessna that was capable of detecting larger aircraft [13]. Their work further validated the use of X-band, and confirmed that flood light illumination of the transmitter with digital beamforming on the receiver is an effective way to determine angle of arrival in a fully static system. The first radar system mounted on a UAV in the small UAS weight class (<25 kg) was the system built by Moses, et al. Their system was a CW radar that identified aircraft using their Doppler signature, but provided no range data [14]. Shi, et al. built an FMCW radar for use in the same weight class using only off-the-shelf components to provide range detection for sense and avoid. Their system processed data using a 2D-FFT to provide instantaneous range and Doppler information [15] and demonstrated good performance in bench top testing scenarios [16].

The current work demonstrates both the feasibility and potential for a UAV collision avoidance radar sensor, however to this point there is little in the way of published data
that demonstrates detection of small UAS from a platform that meets the SWaP constraints of small UAS.

1.3 Contributions

This radar builds upon the previous work in the field and to our knowledge is among the smallest fully portable phased array radar systems with a full digital beamforming back end. It has a form factor that allows it to fit on very small UAS and can be run off of a battery. With it, we have produced some of the first flight results that provide range and angle estimations of small (<25kg) UAS using a radar.

Some unique elements of our design are:

- A phased array digital back end based on a real-time correlator and beamformer.
- A CPWG Wilkinson splitter that performs an efficient $0^\circ$ split using minimal board space.
- An IF filter design that helps to mitigate the phase noise introduced by coupling between the transmitter and receiver and reduces the dynamic range of the system.
- A clutter rejection filter that uses the Levinson algorithm to subtract background reflections.

Although many of the contributions from the radar perspective are incremental, this work also provides insight to the UAS detect and avoid community. Together with the BYU MAGICC lab we have produced several papers that use radar models in sense and avoid simulations to demonstrate the feasibility of radar as a standalone SAA sensor[17]. The flight results in this work validate those findings and encourage further work on the subject.

1.4 Overview

This paper is organized as follows. Chapter 2 gives an introduction to the fundamentals of radar, noise, and phased arrays. Chapter 3 explains the hardware aspects of the system,
both for the transmitter and the receiver. It explains many of the revisions of the design and lessons learned in arriving at the final design. Chapter 4 explains the software aspects of the system and walks the reader through each step of the digital signal processing. The experimental results and analysis are given in Chapter 5, and conclusions are drawn in Chapter 6.
CHAPTER 2. BACKGROUND

Radar systems can be broadly categorized as either pulsed radar systems or continuous wave (CW) radar systems. Pulsed radar systems are covered in great detail by many textbooks because they make up the majority of commercial and military radar systems. The air traffic control system that has existed for the last half a century is based entirely on pulsed radar systems which use a basic coded pulse waveform to request identification information from the aircraft. Pulsed radar systems also use other types of modulation, such as linear frequency modulation (LFM) and methods such as pulse compression to achieve a higher signal to noise ratio (SNR).

Although modulated pulsed radar shares some of the same mechanisms and propagation effects as CW radar, there are some key differences. Continuous wave radar is generally used in small, short range, low-cost applications. Perhaps the most well known continuous wave radar application is that of police radar, which transmits a constant tone and relies on the Doppler shift of the signal to determine its velocity. While CW radars are inherently powerful detectors of velocity, they are unable to detect range without some sort of modulation to the transmitted waveform. Like pulsed radar, CW radar is able to apply coded phase modulation or linear frequency modulation in order to extract this additional information.

2.1 FMCW Radar

Frequency Modulated Continuous Wave (FMCW) radar is the most common type of CW radar. Frequency modulation introduces time variation to the waveform, which allows the radar to track range as well as velocity. Because FMCW radars transmit and receive simultaneously, they are limited in the amount of maximum transmit power that is used in order to prevent bleedthrough. As a result, FMCW radars are generally used in small,
short range applications. Many auto manufacturers are beginning to integrate FMCW radar systems as part of adaptive cruise control systems, which could soon make FMCW radar the most common type of radar system in use. FMCW radar is ideal for use in UAV collision avoidance because it can provide the required detection radius with high range accuracy in a small, low-cost sensor.

Figure 2.1: The triangle FMCW chirp (a) and the sawtooth FMCW chirp (b) are the most common waveforms used in FMCW. They consist of a sine wave whose frequency increases or decreases at a constant rate. The frequency chirp is seen in the spectrograms (a) and (b) as a straight line with a constant slope. The frequency chirp is shown in the time domain for the sawtooth chirp in Fig. (c).

The most common type of FMCW waveform, and the one used in this system, is a linear frequency modulation waveform. This is generally achieved by either a sawtooth modulation or a triangular modulation, as is shown in Fig. 2.1.
Figure 2.2: A typical test setup for the radar system. The radar system is stationary, and is aimed either at the ground or the sky, depending on what is being detected. In this case, the targets are two simple UAVs hovering in a stationary position. The transmitted waves are shown in blue, and the reflected echoes off the targets are shown in red. This figure corresponds with the data shown in Fig. 2.4.

Figure 2.3: A simplified block diagram of an FMCW radar system.
(a) RF Spectrogram of the transmitted (blue) and received (red) chirp for two targets. The target’s distance determines the time delay $\tau$, so the echo time for the farther target (B) is longer than for the closer target (A). For a linear frequency chirp, the beat frequency $f_b$ is constant over the duration of the chirp.

(b) Spectrogram of the received signal after the mixer and low-pass filter. This plot shows the frequency difference between the transmitted and received waveforms in plot (a). In a spectrogram, a constant frequency sine wave appears as a horizontal line, so the beat frequencies for A and B are horizontal lines.

(c) Time domain plot of the received signal on the output of the mixer after it has passed through the low-pass filter. This is the same data from plot (b), but represented in the time domain. Targets A and B each produce sine waves of constant frequency, and their sum, shown in the bottom plot, is the time domain signal seen at the low-pass filter output.

(d) Frequency spectrum of low-pass filter output. This plot is the result of taking the Fourier transform of the “Sum” waveform in plot (c). The Fourier transform shows the frequency content of a signal, which now distinguishes targets A and B. Each frequency corresponds to a different range, so from this plot the targets’ ranges can be extracted.

Figure 2.4: The signal flow through an FMCW radar system for two point targets in the setup shown in Fig. 2.2. The plots are scaled to show the returns from a single transmitted chirp.
Radar determines distance using the speed of light time delay from when the waveform is transmitted to when it reflects off the target and returns. This echolocation mechanism is easily apparent in pulsed radar systems where the time delay corresponds directly with the output of the receiver antenna, but in FMCW radar the relationship is more subtle. In FMCW radar, the received waveform overlaps almost entirely with the transmitted waveform. While the time delay is not readily apparent as in pulsed radar, the time delay in FMCW radar leads to a shifted version of the transmitted waveform. If the slope of the frequency change is linear, then a constant time delay will result in a constant frequency difference between the transmitted and received waveforms for the duration of the chirp. This relationship is visible in Fig. 2.4a and 2.4b. The frequency difference is a linear function of the time delay and the range to the targets. Through the operations of mixing, low pass filtering, and taking the fast fourier transform, the output is readily visible. This simple block diagram is shown in Fig. 2.3 and a basic test configuration is shown in Fig. 2.2.

2.1.1 Stationary Point Target

The following is a mathematical derivation of key FMCW equations for a stationary point target separated from the radar by a distance $R$. Consider a frequency that is increasing linearly in time with a rate of increase $\alpha$,

$$f(t) = f_0 + \alpha t. \quad (2.1)$$

It can be seen from Fig. 2.4a that $\alpha$ is the ratio of the radio frequency (RF) modulation bandwidth $B_{RF}$ and the chirp period $T_c$, i.e. $\alpha = B_{RF}/T_c$. Phase is found by integrating frequency, so the changing phase of the chirp is given by

$$\phi(t) = 2\pi \int_0^t f(t)dt = 2\pi \left( f_0 t + \frac{\alpha t^2}{2} \right) + \phi_0, \quad (2.2)$$
where $\phi_0$ is the initial phase of the oscillator at the beginning of the chirp. A single chirp in the continuous train can be represented by this time gated cosine.

$$ s(t) = \sqrt{P_t} \cos(\phi(t)) = \sqrt{P_t} \cos(2\pi f_0 t + \pi \alpha t^2 + \phi_0), \quad \text{for} \quad 0 \leq t \leq T_c \quad (2.3) $$

The pulse travels the distance $R$ to the target and is reflected resulting in a two-way time delay

$$ \tau = \frac{2R}{c} \quad (2.4) $$

or if the pulse is traveling inside a dielectric with dielectric constant $\epsilon_r$, then the time delay becomes

$$ \tau = \frac{2R\sqrt{\epsilon_r}}{c} \quad (2.5) $$

The received waveform can be seen as an attenuated, time delayed version of the transmitted waveform as is shown in Fig. 2.4a with additive white Gaussian noise (AWGN) denoted by $n_1(t)$ and attenuation factor denoted as $\sqrt{P_r/P_t}$. Substituting in the time delay $\tau$, we get

$$ r(t) = \sqrt{\frac{P_r}{P_t}} s(t + \tau) + n(t) $$

$$ = \sqrt{P_r} \cos(2\pi f_0 (t + \tau) + \pi \alpha (t + \tau)^2 + \phi_0) + n_1(t), \quad 0 \leq t \leq T_c. \quad (2.6) $$

The mixing operation multiplies the two signals. Using the law of cosines, signal at the output of the mixer is

$$ m(t) = s(t)r(t) = \frac{\sqrt{P_rP_t}}{2} \cos(4\pi f_0 t + 2\pi f_0 \tau + 2\pi \alpha t^2 + 2\pi \alpha t \tau + \pi \alpha \tau^2 + 2\phi_0) $$

$$ + \frac{\sqrt{P_rP_t}}{2} \cos(2\pi f_0 \tau + 2\pi \alpha t \tau + \pi \alpha \tau^2) + n_2(t), \quad 0 \leq t \leq T_c. \quad (2.7) $$

The mixer output $m(t)$ has two signal terms, one corresponding to the sum of the frequencies and the other corresponding to the difference between the two frequencies, which, as is shown in Fig. 2.4a, contains the echo time delay information. The term $n_2(t)$ is the downmixed
version of the noise, which after filtering becomes \( n_3(t) \) and will be explored later in greater detail. The operation of low pass filtering extracts the wanted information and results in the final expression, which includes the total system power gain including mixing losses from the antenna output to the ADC input, \( G_s \). The final expression for the received signal is

\[
x(t) = \sqrt{G_s P_r} \cos\left(2\pi \frac{f_b}{c_0} t + 2\pi f_0 T + \pi \alpha T^2 \right) + n_3(t), \quad \text{for} \quad T \leq t \leq T_c.
\]  

(2.8)

Notice that \( x(t) \) is a time gated cosine with constant frequency \( f_b = \alpha T \) and constant phase \( \phi_b = 2\pi f_0 T + \pi \alpha T^2 \). The initial phase offset of the transmitted waveform is not present in \( x(t) \), and the phase term is only a function of the time delay and the system parameters. The term \( f_b \) is referred to as the FMCW beat frequency for a single target.

The distance to the target can be found by substituting Eq. (2.4) into the equation for \( f_b \),

\[
f_b = \alpha T = \frac{2RB_{RF}}{c_0 T_c}
\]  

(2.9)

and

\[
R = \frac{f_b c_0 T_c}{2B_{RF}}.
\]  

(2.10)

The signal can be analyzed using the Fourier transform to produce the following:

\[
X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt = \frac{\sqrt{G_s P_r}}{2} (\delta(f - f_b)e^{j\phi_b} + \delta(f + f_b)e^{-j\phi_b}).
\]  

(2.11)

The frequency domain representation of \( x(t) \) is a set of complex valued delta functions. Since the same information is contained in both positive and negative frequencies, we will examine only the positive frequency spectrum, as shown in Fig. 2.4d.

In early FMCW radar systems, after the mixing and filtering operations the analog beat cosine would pass through a set of narrowband filters called a filter bank, followed by a set of power detectors. The power in each narrow frequency block corresponds to a target at a range given by Eq. (2.10). Current systems work in a similar way, but the filter banks are
realized digitally via an analog to digital converter (ADC) and fast Fourier transform (FFT). By examining the FFT of the downconverted signal $x(t)$, the target locations are readily apparent.

The FFT operation in FMCW radar is often referred to as range compression or pulse compression. This term originates from pulsed radar systems that use frequency modulation, but provides insight into its use in FMCW radar. The Rayleigh resolution for an unmodulated pulse is the inverse of the pulse duration. For short pulses, this can result in very poor resolution. Pulsed radar systems address this issue by introducing frequency modulation into the pulse. By convolving the received signal with a matched filter, the time domain pulse is compressed, and has improved range resolution and SNR by a factor of $B_{RF}T_c$. Because FMCW radar transmits and receives simultaneously for the duration of the pulse, it has no way of resolving targets based on their envelope response. FMCW radar relies completely on matched filter pulse compression in order to resolve targets. The matched filter in FMCW radar is the mixer and FFT. Through the FFT operation, the overlapping beat cosines are compressed into single range bins. This results in a strong SNR improvement because the FFT signals are only compared against noise from a single frequency bin rather than the full IF bandwidth.

### 2.1.2 Moving Targets

While the above relationships hold for stationary targets, there are some interesting effects introduced when the target is in motion. As was mentioned previously, any waves experience a Doppler shift when they reflect off a moving target. The relationship between Doppler shift and relative velocity is given by

$$f_d = \frac{2v f_0}{c_0},$$  \hspace{1cm} (2.12)

where $v$ is the relative radial velocity between the radar and the target, $f_0$ is the radio frequency of operation, and $c_0$ is the speed of light. The factor of two accounts for the
two-way propagation involved in radar as opposed to the one-way propagation experienced by listening to a speeding ambulance.

Figure 2.5: Doppler shifted FMCW waveforms. Figs. (a) and (b) show the RF spectra, with the transmitted waveform in blue and the received waveform in red. Both received waveforms experienced a positive Doppler shift \( f_d \), denoting a target that is approaching the radar. Figs. (c) and (d) show the homodyne IF output after the mixer. Fig. (c) shows the ability of triangle FMCW pulses to determine instantaneous range and velocity, while Fig. (d) shows that sawtooth FMCW pulses are left with an unresolvable range/velocity ambiguity.

For FMCW radar, the Doppler effect is a frequency shift of the received waveform in addition to the time delay. This can be seen in Fig. 2.5b. This has the effect of a frequency shift in the received waveform which is apparent in Fig. 2.5d. With a sawtooth waveform, the frequency shift creates an unresolvable ambiguity, because the resulting lower beat frequency could either come from a closer target, or a fast moving target that is further in range. With a triangle waveform, the Doppler shift has an interesting effect that the up-chirp and down-chirp produce different beat frequencies. The resulting beat frequencies are shown in
Fig. 2.5c. The up-chirp has a beat frequency of

\[ f_{up} = f_b + f_d = \alpha \tau - \frac{2vf_0}{c}, \]  

(2.13)

and the down-chirp has a beat frequency of

\[ f_{up} = f_b + f_d = \alpha \tau + \frac{2vf_0}{c}. \]  

(2.14)

For a single target, this relationship can be exploited to extract both the range and velocity instantaneously. For multiple targets, especially those that are tightly spaced, the process is more complex and requires a more sophisticated waveform in order to provide enough degrees of freedom to resolve all possible ambiguities.

This application uses the sawtooth chirp pattern, despite the range/velocity ambiguity that the Doppler effect introduces. While the Doppler shift introduces a potential range error of a few meters for fast moving targets, the sawtooth chirp enables coherent processing, which greatly improves system performance. This is explained in Section 3.2.2 in greater detail.

A mathematical derivation helps to see what effect Doppler has on the received waveform. Consider the instantaneous expression for the received waveform

\[ r(t) = \sqrt{P_r} \cos(2\pi f_0(t + \tau) + \pi \alpha (t + \tau)^2 + \phi_0) + n_1(t), \quad 0 \leq t \leq T_c. \]  

(2.15)

It was assumed previously that the \( \tau \) term was a constant, but that assumption no longer holds for targets in motion. It is now assumed that \( \tau \) has a constant first derivative proportional to the target’s velocity,

\[ \frac{d\tau}{dt} = \frac{2dR}{c} = \frac{2v}{c}. \]  

(2.16)

The Doppler shift can be observed by taking the time derivative of the instantaneous phase of the received signal, Eq. (2.15). Rather than the \( \tau \) terms disappearing as constant phase terms
in the derivative, they reveal the Doppler shift within the instantaneous received frequency,

\[ f = \frac{1}{2\pi} \frac{d}{dt} \left[ 2\pi f_0(t + \tau) + \pi\alpha(t + \tau)^2 + \phi_0 \right] \]

\[ = f_0 + f_0 \frac{2v}{c} + \alpha(t + \tau)(1 + \frac{2v}{c}). \quad (2.17) \]

By noticing that the term \((1 + 2v/c)\) is a constant, it is lumped with \(\alpha\) to create a new slope constant \(\alpha' = \alpha(1 + 2v/c)\). Making this substitution and gathering terms, the expression for instantaneous frequency reaches a recognizable form,

\[ f(t) = \text{chirp} \left( f_0 + \alpha' t + \alpha' \tau + \frac{f_d}{c} \right). \quad (2.18) \]

The most noteworthy aspect of Eq. (2.18) is that the Doppler effect does not only have the effect of a frequency shift, as is seen in Fig. 2.5, but it also modifies the slope of the chirp as well. For most practical applications, this slope change is very small. For an airplane flying at 350 mph, the chirp slope only increases by 0.0001%.

A different chirp slope between the transmitted and received waveforms, however, may introduce unwanted phase offsets. To determine the net effect, a similar analysis is performed on the low-pass filtered signal described in Eq. (2.8). The phase of \(x(t)\) is

\[ \phi(t) = 2\pi\alpha\tau t + 2\pi f_0\tau + \pi\alpha\tau^2. \quad (2.19) \]

The frequency is found by taking the derivative and dividing by \(2\pi\),

\[ f(t) = \alpha \frac{d\tau}{dt} t + \alpha\tau + f_0 \frac{d\tau}{dt} + \alpha\tau \frac{d\tau}{dt}. \quad (2.20) \]

After making substitutions and rearranging terms, this becomes

\[ f(t) = \frac{f_0}{\alpha\tau} \text{chirp} + \frac{f_d}{c} + \frac{2(f_0 + \alpha t)v}{c} + B_{RF} \frac{T_c}{T_c} \frac{2v}{c}. \quad (2.21) \]
This form reveals some interesting effects of the Doppler shift. Note that the Doppler shift is a function of the instantaneous chirp frequency, and not simply the average. For ultra wide-band chirps, this may result in the target being spread over several frequency bins. There is also an additional Doppler term that is a function of the time delay $\tau$. This means that for linear FMCW waveforms the Doppler shift is a function of both velocity and distance. In essence, two targets of equal velocities at different ranges will introduce different Doppler shifts to the system.

For short range systems, many of these effects can safely be ignored. Consider a radar with the following parameters: $f_0 = 10$ GHz, $B_{RF} = 500$ MHz, $T_c = 2$ ms, $f_{samp} = 2$ Msamp/s, $N_{FFT} = 4096$. This results in an FFT bandwidth of 488 Hz. For a target at a distance of 500 m with a fast closing velocity of 150 m/s (335 mph), the beat frequency is 833 kHz. The Doppler term introduces a frequency shift that starts at 10 kHz and finishes at 10.5 kHz. The range dependent Doppler term introduces a shift of 0.833 Hz. Clearly the range dependent term can be completely ignored in the majority of applications. The velocity dependent frequency spreading of the principal Doppler term is 500 Hz, which is only slightly larger than the width of a single bin. In most applications, chirp non-linearity and multi-path will contribute more target spreading than the Doppler term, so it is safe to use the average carrier frequency rather than the instantaneous carrier frequency in many Doppler calculations. Finally, if the entire Doppler effect is ignored and all frequencies present in the baseband cosine are attributed to range, this introduces a range error of about 1%, or 5 m for this target. For many applications, even this is an acceptable amount of error.

### 2.1.3 Received Power

The power received by a radar system from a target is a well defined function of the system parameters and the distance to the target. Once the system is well calibrated, the system can predict properties of targets based on the amount of received power. If the radar system is operating in a pulsed mode, targets with moving parts, like a helicopter or a jet
turbine, will give off different Doppler signatures, allowing for even more aspects of the target to be identified.

Although there are many different forms of the radar equation, FMCW uses the continuous wave version that assumes a constant value of transmitted power. To derive it, consider an omnidirectional antenna that illuminates a patch of sky at a distance $R$. The power density at that patch decays in the predictable $R^2$ fashion for an expanding sphere and is

$$S_u = \frac{P_t}{4\pi R^2}. \quad (2.22)$$

If an antenna with a gain greater than unity is used, then the power density in the direction of peak gain is

$$S_g = S_u G_t = \frac{P_t G_t}{4\pi R^2}, \quad (2.23)$$

where $G_t$ is the peak gain of the transmitting antenna.

When the transmitted energy reaches a target, the amount of power reflected is determined by a combination of its geometrical and material properties of the object. This figure is called the radar cross section (RCS or $\sigma_{RCS}$) and has units of m$^2$. The RCS determines how much of the incident power density is re-radiated by the target. RCS varies by angle and frequency of operation. RCS can also be measured in bistatic mode, where the radar transmitter and receiver are in different locations. In this case, the radar is operating in monostatic mode, where the transmitter and receiver are at the same location, and the RCS is a measure of how much power is reflected back at the angle of incidence. The reflected power is found by multiplying the incident power density by the RCS.

The amount of echo power density that reaches the radar $S_e$ is found by multiplying the reflected power by the isotropic decay term in Eq. (2.22) to produce

$$S_e = S_g \sigma_{RCS} \frac{1}{4\pi R^2} = \frac{P_t G_t \sigma_{RCS}}{(4\pi)^2 R^4}. \quad (2.24)$$
The amount of power received by the radar system \( P_r \) is found by multiplying the echo power density by the effective aperture size \((m^2)\) of the receiver antenna, which gives

\[
P_r = A_r S_e = \frac{P_t G_t A_r \sigma_{RCS}}{(4\pi)^2 R^4}.
\]  

(2.25)

It is more useful to characterize the antenna in terms of its gain than in terms of its aperture size, so the final form of the radar equation is found by making the aperture gain substitution to produce

\[
P_r = \frac{P_t G_t G_r \lambda^2 \sigma_{RCS}}{(4\pi)^3 R^4 L_s},
\]  

(2.26)

where \( \lambda \) is the wavelength of operation in meters given by \( \lambda = \frac{c}{f} \) (speed of light / frequency of operation). This equation predicts the amount of power at the output of the receiver antenna. When the electronics are constructed and the amount of noise is fixed, the only way to receive more power from a target is by modifying one of these parameters. It includes a loss term \( L_s \), which includes signal losses such as atmospheric losses, antenna pattern attenuation for targets angled off the peak gain of the antenna, or signal processing losses where predicted and actual signal processing gains deviate.

2.1.4 Noise

In the absence of noise, a radar system could theoretically detect targets at an infinite distance as long as it had a clear line of sight. Unfortunately, noise is inevitable, and it is the primary limiting factor in the detection range of a radar system. Noise comes from external factors such as thermal radiation and clutter as well of internal thermal electronic noise.

External Noise

One of the primary categories of noise in an RF system is the noise external to the system that is present at the antenna, then later amplified. Anything that is not the signal
of interest falls into this category. A few common sources are thermal noise, interference, and clutter.

Thermal Noise is the dominant natural external noise source at radio frequencies. Objects whose temperature is greater than absolute zero experience random electron motion, which radiates wideband electromagnetic energy. This is referred to as blackbody or thermal radiation. Its contribution depends on the radiation efficiency of the antenna $\eta_{\text{rad}}$ and the physical temperature $T_a$ of whatever the antenna beam is angled at, which at X-band is approximately 280 Kelvin for anything on or below the horizon or approximately 10 Kelvin when angled directly at the sky.

Interference can be the result of non-malicious users of the same spectrum, or it can come from malicious users attempting to jam the signal. FMCW radar is similar to technologies such as Wi-Fi or Bluetooth that spread the signal power over a wide bandwidth, which makes them relatively immune to interference, but if the interferer is close enough and radiates enough power, it may compromise the linearity of the receiver front end. This effect explains why the operation of a high power microwave oven degrades Wi-Fi connection quality.

Clutter is a broad term that encompasses unwanted radar reflections off of ground, vegetation, and buildings. Sometimes clutter itself is the signal of interest. Generally, however, clutter occurs when the antennas are oriented such that either the main lobe or one of the side lobes is at least partially angled at the ground. Detecting a target in the presence of clutter is a bit like shining a flashlight into the bushes to look for something. Although the object may be illuminated, the illumination of the surrounding vegetation overwhelms the senses and makes it difficult to discriminate between the object and the clutter. This analogy offers some insight into effective ways to mitigate clutter. Because clutter is comprised of real but unwanted radar returns, adjusting the power of the transmitter is ineffective since it illuminates both the signal and the clutter by the same ratio. Where possible, clutter is avoided by reducing or eliminating the squint
illumination angle so as to minimize direct illumination, i.e. point the flashlight at the sky or just above the horizon. When clutter is unavoidable, signal processing techniques can be used to calibrate out background returns or to only detect targets in motion.

**Internal Thermal Noise**

Once the signal is captured at the antenna input, the amount of noise present can only increase in relation to the signal. Most receiver systems are designed so that the system adds about the same amount of internal electronics noise as the amount of external noise. In applications like radio astronomy, the external noise is very small ($T_a = 10 \text{ K}$), so the receiver systems are cryogenically cooled to reduce electronics noise. In radar, the external noise is much higher, but care must still be taken to ensure that internal noise does not overpower the signal.

A resistor disconnected from any voltage source will still generate a voltage across its terminals due to the random motion of warm electrons. The voltage is additive, white (uniform spectrum), Gaussian (AWG) noise with an RMS voltage variance equal to $\overline{v_n^2} = \sqrt{4k_BT_RB}$ where $k_B$ is Boltzmann’s constant and equals $1.380648813 \times 10^{-23} \text{ J/K}$, $T$ is the temperature of the resistor in Kelvin (generally $T_0 = 290\text{K}$ is used for room temperature), $R$ is the resistance value in ohms and $B$ is the observation bandwidth in hertz.

When a resistor is connected to another resistor of equal value, half of the noise voltage is delivered to the matched load. This produces

$$P_n = \frac{(\overline{v_n}/2)^2}{R} = k_BT_B \quad (W).$$  \hspace{1cm} (2.27)

This equation serves as a reference for both the noise power delivered by a resistor as well as any AWG noise source. Once the noise power is measured, an effective noise temperature is determined by solving Eq. (2.27) for $T$.

Although Eq. (2.27) was derived for a warm resistor, the noise mechanisms in an RF system are more complex than a single resistor. Much of the noise originates from amplifiers that produce both voltage and current noise that combine in complex ways. The amount
of noise added by an RF module can be represented simply using the “noise factor” \( F \) or the “noise figure” \( NF \), which share the relationship \( NF = 10 \log_{10}(F) \), though the term noise figure is often used to refer to both quantities.

Noise figure is the ratio of total noise power at the output to amplified input noise power when the device input noise temperature is equal to \( T_0 = 290K \). The output noise power is the sum of the input noise power amplified by the gain \( G \) and the output referred noise added by the amplifier \( N_a \). This results in the expression

\[
F = \frac{N_{out}}{N_{in}G} = \frac{N_a + k_B T_0 BG}{k_B T_0 BG}.
\]  

(2.28)

Once the noise figure has been measured, the output-referred added noise can be found by rearranging the previous equation and becomes

\[
N_a = (F - 1) k_B T_0 BG.
\]

(2.29)

When multiple amplifiers are combined, the output noise term includes the amplified output noise from the first stage added to the amplified output noise of the second stage. For three stages, the noise figure is

\[
F = \frac{N_{out}}{N_{in}G_{sys}} = \frac{N_{a2} + (N_{a2} + (N_{a1} + k_B T_0 BG_1)G_2)G_3}{k_B T_0 BG_1 G_2 G_3}.
\]

(2.30)

From Eq. (2.30) it is apparent that the first amplifier in the chain contributes the most noise because its noise is amplified by the gains of the other two stages. This leads to an important principle in receiver design that the first stage establishes the noise figure of the entire system. This becomes more visible in the expression for the overall system noise figure

\[
F_{sys} = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + ... + \frac{F_n - 1}{G_1 G_2 ... G_{n-1}},
\]

(2.31)

where \( G_i \) and \( F_i \) are the gain and noise figure of the i-th stage. If the gains of early stages are large, then noise contribution from later stages can be insignificant. In some cases, such as a
filter or a mixer, the stage incurs a signal loss rather than a signal gain. Lossy components are often characterized by their loss factor $L$, which is either referred to as insertion loss for a filter or conversion loss/gain for a mixer. The value for noise factor and gain of a lossy component used in Eq. (2.31) are the inverse of the loss factor, $F = G = 1/L$. In that situation, the noise can be minimized by either introducing higher gain before the lossy component, or minimizing the noise figure of the amplifier following the lossy component.

After the external noise has traveled through the receiver electronics, the total amount of noise power at the output is given by

$$P_n = k_B BT_0 F_{sys} G_{sys}. \quad (2.32)$$

where $G_{sys}$ is the total gain (including loss) of the system from input to output. After the gain is set to establish appropriate signal levels, SNR is maximized by optimizing the noise figure.

For pulsed radar systems, the receiver bandwidth is equal to the pulse bandwidth, so the noise bandwidth $B$ is the bandwidth of the transmitted pulse. For CW radar systems, the noise bandwidth is independent of the transmitted waveform and $B$ is the bandwidth of the receiver electronics at any point in the receiver chain. In the case of this system, the RF electronics are wideband, and there are no RF filters so in the RF stage $B$ has a value of several GHz. The IF stage has an anti-aliasing filter with a cutoff frequency of a few MHz. In the IF, the noise is still AWGN even though the LO signal used for downmixing was a frequency sweep, but its bandwidth is reduced significantly and is now equal to the IF bandwidth. After digitization and the FFT, the bandwidth is reduced even further. The FFT acts as a filter bank, and the noise power bandwidth $B$ in each frequency bin is the FFT bin width. Thus, the noise terms $n_1(t)$ and $n_2(t)$ from Eqs. (2.6) and (2.7) are identical and their powers are calculated from Eq. (2.32) using a noise bandwidth $B$ equal to the bandwidth of the RF electronics. The low-pass filtered analog noise term $n_3(t)$ from Eq. (2.8) has a noise bandwidth $B$ equal to the anti-aliasing filter bandwidth $B_{IF}$, and the noise power in each
FFT bin has a noise bandwidth $B = B_{IF}/N_{FFT}$, the IF noise bandwidth split equally across all FFT bins.

**Signal to Noise Ratio**

The key figure of merit in radar and communications systems is the signal to noise ratio (SNR). At the end of the signal chain, a target’s SNR is what determines whether or not it is detectable. Although SNR is a ratio, it is usually expressed in terms of decibels (dB). The dB relationship is $\text{SNR}_{dB} = 10 \log_{10}(\text{SNR})$. Generally for a target to be detectable, it needs to have about 10 dB of SNR, though systems are able to detect targets with lower SNR using sophisticated filtering mechanisms. The SNR at any point in the system is found by dividing the signal power by the noise power. The SNR at the the system output is found by multiplying the received signal power in Eq. (2.26) by the receiver gain $G_{sys}$ and dividing by the output noise power in Eq. (2.32). In so doing, the receiver gain cancels out and the expression becomes

$$\text{SNR} = \frac{P_t G_t G_r \lambda^2 \sigma_{RCS}}{(4\pi)^3 R^4 k_B B T_0 F_{sys} L_s}. \quad (2.33)$$

One of the primary means of increasing SNR in pulsed radar systems is by modulating the transmitted waveform, and implementing a matched filter receiver of that same modulation. This operation increases SNR by a factor of $\tau \beta$, where $\tau$ is the pulse width and $\beta$ is the modulation bandwidth. A similar effect is seen in FMCW in the reduction of the noise bandwidth $B$ through downmixing and taking the FFT. The receiver originally receives signal and noise power over the entire bandwidth $B_{RF}$ of the transmitted chirp. In the FFT output, however, the signal is compressed into a spike in a single frequency bin, and the noise power used in calculating the SNR is the FFT bin width.

**Phase Noise**

For most analyses it is assumed that the radar waveform generated by the voltage controlled oscillator (VCO) in the transmitter is a perfect tone that shifts in frequency. In
the frequency domain, this would appear as a perfect delta function that shifts along the frequency axis. In dealing with actual circuit components, however, the phase of the output experiences random fluctuations due to thermal effects in transistors. The random phase fluctuation is called phase noise, and has the effect of spreading signal power into adjacent frequencies in a tent-like spectrum.

Phase noise is manifest in the system through signal leakage between the transmitter and receiver. Poorly matched antennas, antenna coupling, and short range reflections from the ground or a radome all create paths where the signal could bleed through, as shown in Fig. 2.6. If the transmitted waveform had no phase noise, then the short range reflections would show up as narrow spikes. However, when the VCO has significant phase noise, the short reflections translate the entirety of the phase noise spectrum to the IF of the radar. This is shown in Fig. 2.7. When left unmitigated, the phase noise can exceed the levels of both the electronics noise and the external noise. Ideally, the phase noise would be reduced below the level of the electronics noise.

Phase noise is a function of the noise correlation, transmit power, and isolation. In a true homodyne system where the transmitter is used for downconversion, much of the phase
noise can be eliminated. This was seen in Eq. (2.8) where the phase term $\phi_0$ disappeared in the baseband signal. If the leakage path lengths are short and the system is a true homodyne system, then much of the phase noise will be coherently cancelled in the downmixing process. Some systems even implement a phase shifter in a feedback loop on the receiver in order to coherently cancel the phase noise power. This is discussed in detail in [6]. Because phase noise occurs from leakage, it can be reduced by either improving isolation or reducing transmit power. Isolation can only be improved to a certain extent, so once it has reached its limit, the only other option is to decrease transmit power. Pulsed radar systems are able to transmit hundreds of kilowatts of power by turning off receiver electronics during the transmitted pulse. CW radar systems, however, are generally limited to 1-2 watts of transmit power because they transmit and receive simultaneously.

### 2.2 Phased Array Antennas

A single antenna is much like a microphone. Although it can tell if a source of sound is near or far, it has no way of determining the location of the source. The angular discrimination
of an antenna is limited to detecting whether or not a target is within its beam, or antenna pattern. Some radar systems make use of this by designing large dish antennas that have extremely focused “pencil beam” antenna patterns. They obtain angular information by rotating the antenna on a mechanical gimbal. These systems have the advantage that the narrow antenna pattern helps them to inherently see much further than a wide antenna pattern would; however, they are limited by their large size and the speed of rotation of the gimbal. Gimbal rotation speeds are generally 1-2 second per revolution at a minimum, which can be inadequate in applications where a fast update rate is needed.

The alternative to a mechanically steered antenna is an electronically steered antenna array. Electronically steered arrays or phased array antennas are made up of a collection of small elements with wide antenna patterns. Although the elements of a phased array can be placed in any configuration, they are generally arranged in a uniform linear array (ULA) where the antennas are spaced evenly along a line, which is the case considered in this research. In a phased array, the beam is formed based on the time difference that results as a signal arrives at different elements. For a narrow bandwidth of operation, this time delay is manifest as a phase delay on the carrier frequency. The phased array forms an antenna pattern by reversing the phase shift or time delay for each antenna, and combining their outputs.
Because a signal arriving broadside to the array would arrive at all elements at the same time, in order to form a beam in that direction the antenna outputs are summed with no phase shift. A signal arriving from an angle would have a linear phase shift between each element. In order to form a beam at that angle, a linear phase shift is applied to each element and the output is summed as is shown in Fig. 2.8. The direction of the beam can be changed by simply changing the phase weighting at each element.

Many phased arrays apply amplitude weighting in addition to phase weighting. The most general form for the output of a phased array with N elements is

$$v_{out} = v_{oc} \sum_{n=1}^{N} w_n E_n(\vec{r}),$$

(2.34)

where $n$ indicates the index of the element, $E_n(\vec{r})$ is the value of the reflected electric field at each element, $\vec{r}$ is the location of the element relative to the array origin, $w_n$ is the complex weight for that element that includes both amplitude and phase weighting, and the open circuit voltage scale factor is

$$v_{oc} = \frac{4\pi jre^{jk\vec{r}}G_s}{\omega \mu},$$

(2.35)

where $G_s$ is the system gain of the receiver and $k = 2\pi/\lambda$ is the wavenumber of the carrier.

If all elements are identical and relatively closely spaced, the electric field at each element will only differ in phase, as a function of its location and the direction of arrival. As a result, Eq. (2.34) can be rewritten to extract the total electric field $\bar{E}_{el}$ for an element at the origin of the array and add the phase shift to each element as a function of the direction of propagation $\hat{r}$. This gives

$$v_{out} = v_{oc} \bar{E}_{el}(\vec{r}) \sum_{n=1}^{N} w_n e^{jK \cdot \vec{r}_n},$$

(2.36)

where $\bar{K} = k\hat{r}$. 

29
The combination of antenna locations \( \bar{r}_1, \bar{r}_2, \ldots, \bar{r}_N \) and corresponding weights \( w_1, w_2, \ldots, w_N \) leads to the array factor \( A(\theta, \phi) \), which indicates the pattern of the formed beam as a function of angle of arrival either in terms of \( \bar{k} \) or in terms of \( \theta \) and \( \phi \). The array factor is not the lone contributor to the beam that is formed. A phased array beam is the product of both the antenna pattern from a single element and the array factor, as is shown in Fig. 2.9. If the system has full control over the weight values \( w_n \), then it can form completely arbitrary beams, but the beams that it forms will always be attenuated to fall within the element pattern.

![Figure 2.9: Radiation pattern for a 10-element phased array. The antenna weights \( w_i \) have a linear phase increment that steers the pattern 40 degrees off boresight. The radiation pattern is the product of the array factor \( A(\theta, \phi) \) and the pattern of the identical individual elements, indicated in red.](image)

An intuitive explanation for the way an array forms beams can be drawn from the discrete time Fourier transform (DTFT). The antenna locations are analogous to discrete time samples. The antenna weights are related to amplitudes of the time samples, and the beam pattern is analogous to the DTFT of the time samples or weighted antenna locations.
A shift in the Fourier domain, in this case an angular shift of the beam, is the result of a linear phase increment on the elements, or a linear phase term in the weight values. Fourier domain characteristics such as sidelobe levels can be controlled by applying an amplitude weighting window to the element domain, and a finer beam in the Fourier domain is the result of more time samples, or more elements. Concepts such as undersampling and aliasing apply as well, and explain the appearance of spurious grating lobes, sidelobes with the same amplitude as the main beam that appear when the antennas are spaced too far apart.

A phased array that applies the complex antenna weights and performs the summing operation using analog hardware is only capable of steering the beam in a single direction at any given time. Because the steering is controlled electronically, an analog beamformer can scan the antenna beam much more quickly than a mechanical gimbal, but it is still limited to pointing at one patch of sky at any given moment. In order to form multiple beams and track multiple targets simultaneously, the system must implement digital beamforming. Shown in Fig. 2.10, digital beamforming creates a full receiver chain for each antenna, and performs the operations of phase shifting and summing in the digital domain. A digital beamformer can scan its beam in all directions simultaneously, but it does so at the cost of extra computation and RF hardware. As computation has become less expensive, digital beamforming has become more popular in phased array systems.
Figure 2.10: Standard architecture of a phased array receiver that implements digital beamforming.
CHAPTER 3. RADAR SYSTEM DESIGN

The majority of the work of this thesis was the development of a radar sensor to be used in UAV detection. The detection scenario presented by UAVs is the short-to-medium range (∼1km) detection of small airborne targets (\(\sigma_{RCS} < 1\) m) from either a ground-based or aerial platform. This section describes the analog and RF design of the radar transceiver and how it applies to the detection of small UAS. The discussion highlights the major lessons learned over the course of the several years spent in the design and debugging process.

3.1 Parameter Selection

Although the design of a radar system entails dozens of small decisions, there are a few key parameters that have a major effect on the way that the transmitter and receiver function. Some of these are flexible and can be changed with a few lines of code, while changing others would entail a complete re-design of the system. The major system parameters are listed below in Table 3.1. This section provides a brief discussion of some of the key parameters and their associated trade-offs.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>2.25in x 4in x 1in</td>
<td>Weight</td>
<td>120 g (0.26 lbs)</td>
</tr>
<tr>
<td>Consumed Power</td>
<td>8 W</td>
<td>Carrier Frequency</td>
<td>10.25 GHz</td>
</tr>
<tr>
<td>Transmitted Power ( (P_t) )</td>
<td>5 mW</td>
<td>Chirp Bandwidth ( (B_{RF}) )</td>
<td>500 MHz</td>
</tr>
<tr>
<td>Chirp Period ( (T_c) )</td>
<td>2.048 ms</td>
<td>IF Bandwidth ( (B_{IF}) )</td>
<td>1 MHz</td>
</tr>
<tr>
<td>System Noise Figure ( (F) )</td>
<td>6 dB</td>
<td>ADC sample rate</td>
<td>2 Msamp/s</td>
</tr>
<tr>
<td>Range Resolution</td>
<td>0.3 m</td>
<td>Maximum Range</td>
<td>614 m</td>
</tr>
</tbody>
</table>
### 3.1.1 Carrier Frequency

The choice of carrier frequency affects the physical size of the system and its antennas, the cost of manufacturing, and the propagation losses of the signal. The choice of carrier frequency may end up being the one over which the engineer has the least control. The electromagnetic spectrum is finite, and telecommunications companies spend billions of dollars for the right to use different parts of it. There are certain parts of the spectrum that are designated as “unlicensed” bands. These bands are free for use by industrial, scientific, and medical (ISM) industries, provided that users abide by certain limitations. Other bands, while not commercialized, are reserved for certain purposes, such as radiolocation or radionavigation. In deciding on the frequency of operation of a radar system, the engineer may analytically determine an approximate choice, but make the final decision based on the closest legally available band. In many cases, the navigation of the Federal Communications Commission (FCC) regulations may prove more difficult than the design of the system itself.

The relationship between the carrier frequency, propagation loss, and the antennas can be seen by analyzing the standard radar equation presented in Eq. (2.26). Another form of this equation is given in terms of the effective aperture area of the transmitting and receiving antennas \( A_t \) and \( A_r \) using the aperture gain substitution \( G = 4\pi A/\lambda^2 \) to give

\[
P_r = \frac{P_t A_t A_r \sigma_{RCS}}{(4\pi)^3 R^4 \lambda^2}.
\]  

(3.1)

In the gain form of the radar equation, Eq. (2.26) on page 20, \( \lambda \) is in the numerator, implying that an increase in frequency results in an decrease in received power, all else held constant. However, in the aperture form of the radar equation, Eq. (3.1), the opposite is implied. This discrepancy is understood by observing what is being held constant. When the antenna gain is held constant and the frequency is increased, the antenna can shrink in size. When the effective antenna aperture area is held constant and the frequency is increased, the antenna stays the same size, but the gain of the antenna and possibly its complexity increase.

Which viewpoint is more useful? It depends on the application. In the case of designing a sensor, it is most practical to assume that the gain is a fixed variable. This is because the
application often dictates a required field of view. Because antennas have an approximate inverse relationship between field of view and gain, by constraining the system to a fixed field of view the designer is essentially fixing the gain of the antennas regardless of the frequency of operation. There is another practical reason why it is more appropriate to maintain a fixed antenna gain. In the case of reflector and horn antennas almost any aperture size can be realized; however for small, low cost applications, antennas are generally planar systems integrated on a printed circuit board (PCB) or integrated circuit (IC). Most planar antenna designs consist of a collection of individual elements that are combined into a single input using a feed network. A planar antenna can be viewed as an aperture antenna with the same approximate physical area, but there is a physical limit on how large the planar antenna can be made. The larger the planar antenna, the more complex the feed network must be, which in some cases renders the design unrealizable. In addition, the size of the feed network is a function of the substrate thickness more than the frequency of operation, so as the frequency increases and the elements shrink the feed network remains the same size and adds additional loss.

Tied to the carrier frequency and antenna selection is the field of view. For manned aircraft the FAA mandates a “see and avoid” field of view of approximately 30 degrees by 120 degrees. This results in an approximate antenna gain of 10dB. The approximate planar antenna size needed to realize this field of view and gain can be determined by treating the planar antenna as a rectangular aperture. The relationship between gain and aperture area is given by

\[ \Delta \theta = \frac{50.76}{D} \lambda \]  \hspace{1cm} (3.2)

\[ G = \frac{32383 \text{deg}^2}{\Delta \theta_{el} \Delta \theta_{az}} \]  \hspace{1cm} (3.3)

where \( \Delta \theta \) is the half power beam width in degrees for a given dimension, \( D \) is the length in meters of the aperture for a given dimension, and \( G \) is the gain of the aperture [18].

At 1 GHz, the antenna would measure 12cm × 50cm. At 10 GHz, 1.2cm × 5cm, and at 100 GHz, approximately 0.12cm × 0.5cm. This analysis shows that, given the field of view
requirements imposed by the FAA, any system with an operating frequency much lower than 10 GHz would result in antennas that present a payload too large for most UAVs. Thus, for small UAVs, the required beam widths and available payload size imposes a soft lower limit on the frequency of operation at approximately 10 GHz.

In addition to the propagation losses incurred at higher frequencies as predicted by Eq. (2.26), there are also additional losses incurred by the atmosphere at higher frequencies. The total loss is a combination of gaseous attenuation from oxygen and water vapor, attenuation due to rain, and attenuation due to clouds. The gaseous attenuation shown in Fig. 3.1 demonstrates that lower frequencies are more favorable for avoiding atmospheric attenuation. Although radar can penetrate clouds and rain, there is also attenuation associated with both of those, which follows the same general trend of an upward increase with frequency. The attenuation is a function of distance traveled in the medium, usually dB/km.

Figure 3.1: This graph shows the one way gaseous attenuation of electromagnetic waves in the atmosphere due to oxygen and hydrogen. Figure from [1].

In short range applications where the radar maximum range is only a few meters, a high frequency is desirable because it allows for miniaturization and integration of the antennas. This application dictates a maximum range of 1-2 km, a distance where propagation effects and atmospheric absorption begin to degrade the signal. Target RCS is also a strong
function of frequency. As a general rule, radio waves only produce strong backscatter when the size of the target is equal to or larger than the wavelength of operation. High frequency systems can see small targets that low frequency systems may not. It should also be noted that the prices of system components and test equipment also follow a general upward trend with increasing frequency.

The examination of free space path loss, propagation effects, and manufacturing costs demonstrated that a radar system for UAV SAA should operate at the lowest possible frequency. The analysis of the antenna dimensions demonstrated that lowering the frequency increased the size of the antennas required to meet the field of view requirements. In addition, the FAA regulations limit the choice of frequency to a few discrete bands. The 24GHz band is a good option, and a compact 24GHz radar for UAS applications was developed in a collaborative work [19, 20]. This sensor met the SWaP requirements for UAS, however it did not have adequate range for detecting small UAS. Taking all these into consideration, this system was designed at 10GHz as a good trade-off between propagation effects, antenna size, cost of development and maximum desired range.

3.1.2 Chirp Bandwidth

The chirp bandwidth or RF bandwidth $B_{RF}$ refers to the actual range of frequencies transmitted. While the bandwidth is restricted by FCC regulations, the bandwidth also experiences limitations from a hardware standpoint. Chirp bandwidth, chirp period, and intermediate frequency (IF) bandwidth all combine to determine the visible range of the radar, but chirp bandwidth is the least flexible of the three and should be chosen first.

Reducing chirp bandwidth allows for narrower bandwidth RF components, which can improve the quality and cost of components available, achieving higher powers and lower noise. Many RF components are characterized by their quality factor or Q factor. There is an inverse relationship between the Q factor and the bandwidth over which the device operates. Antennas, amplifiers, mixers, oscillators, and splitters that work over a narrow bandwidth
can achieve higher performance in terms of efficiency and noise. If a system must work over a wider bandwidth, the performance metrics tend to decrease and the price tends to increase.

Increasing chirp bandwidth improves range resolution and lowers the radiated power spectral density, which mitigates interference. Range resolution is the separation distance required to resolve two different point scatterers. For pulsed radar systems, range resolution is determined by the time duration of the pulse. For CW radar systems, range resolution is determined by the receiver filter bank frequency separation, which in this case is the FFT bin width formed by the digital signal processing block. The bin width is the ADC sample rate divided by the size of the FFT,

\[ f_{\text{bin}} = \frac{f_{\text{adc}}}{N_{\text{FFT}}} = \frac{1}{T_c}. \]  
(3.4)

Substituting the bin width into the FMCW range equation Eq. (2.10) gives

\[ R_{\text{res}} = \frac{c_0}{2\alpha T_c} = \frac{c_0}{2B_{\text{RF}}}. \]  
(3.5)

Equation (3.5) demonstrates that as RF chirp bandwidth is increased, the range resolution becomes finer.

In high altitude applications, range resolution may not be important, but in situations where targets must be detected in the presence of clutter, range resolution becomes more important. Range resolution was a moderate priority for this sensor, so we elected to use all the available bandwidth at our chosen center frequency. The system has a bandwidth of 500 MHz that extends from 10 GHz to 10.5 GHz, occupying the full band allotted by the FCC for radiolocation, and has a resulting range resolution of approximately 0.3 meters.

3.1.3 Intermediate Frequency Bandwidth

Intermediate frequency (IF) bandwidth \( B_{\text{IF}} \) is established by the filters, amplifiers, and ADC after the mixer on the receiver. The IF bandwidth can be set independently of the RF bandwidth, although they both affect the maximum range of the radar through the
equation
\[ R_{\text{max}} = \frac{c B_{\text{IF}}}{2\alpha}. \] (3.6)

Once the slope \( \alpha \) of the chirp is set, the IF bandwidth determines the maximum range because any target with a longer time delay and beat frequency will be filtered out by the IF low pass filter.

If both high range resolution and long range are needed, then the IF bandwidth must be relatively large (10+ MHz). From an analog perspective, this is difficult because active filters require operational amplifiers with approximately 100x the filter bandwidth to function properly. If wideband RF amplifiers are used, they generally have a low frequency cutoff that make the radar blind to close targets. From a digital perspective, finding an ADC to operate at high sample rates is both difficult and costly. In addition, the digital back end must be capable of supporting the data rate established by the ADC.

In general, it is almost always more desirable to have a larger IF bandwidth because it enables both a fast update rate and fine range resolution. However, the analog and digital considerations have the greatest influence on this parameter, and the chirp period is chosen after the IF bandwidth in order to establish appropriate values of maximum range. This system has an ADC sample rate of 2 MHz. It utilizes a set of active analog filters that establish the usable IF Bandwidth at the Nyquist rate of 1 MHz.

3.1.4 Chirp Period

The chirp period \( T_c \) is the duration of frequency sweep and is often the last parameter chosen because it is the easiest to modify. The chirp period follows the relationship in Eq. (3.6), and once both \( B_{\text{RF}} \) and \( B_{\text{IF}} \) have been fixed, it shifts the maximum range. Some systems may switch between different values of \( T_c \) during normal operation in order to shift the radar field of view between shorter and longer fields with finer and coarser resolution.

The chirp period can also be thought of as the integration time of the radar. If \( T_c \) is decreased, then the radar has a faster update rate and is able to detect faster moving targets with less smearing. If \( T_c \) is increased, the radar has an increased coherent processing interval,
which improves range resolution and gives an SNR advantage for stationary and slow moving targets.

Most FMCW radars in the literature have a chirp period on the order of a few milliseconds. We chose $T_c$ after we had chosen all the other parameters to give us the desired detection range of 1-2 km. For most testing, we used a value of $T_c = 2.048$ ms, which establishes a maximum range of about 600m. When different range profiles are needed, we can switch between values of $T_c = 4.096$ ms and $T_c = 8.192$ ms to detect up to 2400m. In order to improve efficiency, $T_c$ is chosen so that the corresponding number of ADC samples in a frame is equal to a power of two.

### 3.2 Transmitter

The primary goal of the transmitter is to create a clean chirp with high linearity. While the theoretical transmitter architecture shown in Fig. 2.3 of the background section looks simple, realizing such a system has many practical challenges. Over the course of this research we implemented several different waveform generation architectures, discovering their advantages and disadvantages for the UAV sense and avoid application. The following section describes the pros and cons of each architecture we implemented.

#### 3.2.1 Direct Digital Synthesizer

A direct digital synthesizer (DDS) is a type of digital-to-analog converter (DAC) that specializes in creating sine waves. It is the digital counterpart to an analog circuit known as a phase locked loop (PLL). Whereas a standalone DAC can create arbitrary waveforms, a DDS combines a DAC with frequency controlled registers and numerically controlled oscillators to create a system that produces sine wave outputs with high precision.

Within the frequency band of interest, a DDS has outstanding performance. It has better frequency agility and less phase noise than a PLL, which enables it to produce FMCW waveforms with high linearity. Although a DDS can generate the desired waveform extremely well, it generates many additional spurious signals, which are referred to as spurs. A typical
Figure 3.2: The typical output frequency spectrum of a DDS. (1) indicates the fundamental, desired waveform, which is a frequency swept sine wave. (2) indicates all harmonics and sub harmonics, which are duplicates of the fundamental chirp at different frequencies, some sweeping in the opposite direction. (3) indicates spurious tone bleedthrough from the reference oscillator. (4) indicates the extremely tight filtering necessary to remove all spurs from a DDS waveform generator.

DDS output is shown in Fig. 3.2. Even with an analog reconstruction filter, there are still significant reference spurs and image spurs, harmonics and subharmonics. DDS systems work best when the chirp bandwidth is small, so as to provide sufficient room for filter rolloff. Even with a small bandwidth, the amount of filter stages needed is almost prohibitive, and some amount of spurs must simply be accepted. These issues are exacerbated when an already cluttered spectrum is mixed to higher frequencies. Inherent nonlinearities in mixers and amplifiers amplify the unwanted signals and add additional spurs that must be filtered using RF filters. Some techniques such as RF coupled line filters and single sideband mixers can aid in the pursuit of a clean waveform, but the amount of additional circuitry and complexity required is substantial. After designing a connectorized version of the DDS-based transmitter, it was determined that a clean version of the architecture could not be made small enough to meet SWaP requirements for UAS.

**Tuned Oscillator**

Since most of the issues surrounding the generation of the chirp using a DDS involve harmonics, mixers, amplifiers and filtering, a way around those issues is to generate the
waveform directly at the frequency of operation. For a system operating at 10GHz, harmonics and subharmonics at 5 GHz, 15 GHz and 20 GHz can be filtered much more easily than the tightly spaced spurs and image frequencies at 10.5 GHz and 9.5 GHz from the DDS.

One solution for this architecture is a single chip radar transceiver produced by ViaSat that operates a 24GHz. The ViaSat radar chip operates by locking an oscillator to a base frequency, then fine tuning the oscillator control voltage using a DAC. A chirp is produced by successively incrementing the DAC output code, which increases the oscillator tuning voltage resulting in a frequency chirp at the oscillator output. Because the oscillator operates in an open loop configuration with the DAC, the user can produce arbitrary modulation patterns with high agility. In an unlocked state, however, the system can experience significant drift due to temperature and a high degree of phase noise. Because this system relies on a DAC to tune the oscillator, it is also subject to the non-linear tuning curve of the oscillator. As was discussed previously, the predominant effect of nonlinearity is to effectively decrease the range resolution and range.

Although the tuned oscillator solution does not provide a range sufficient for UAS, it is a practical solution for short range applications. The decreased complexity and high frequency of operation allow for significant miniaturization. After designing a functional 24 GHz single chip radar system, we determined that the chirp non-linearities established range resolution and maximum range that were insufficient for UAV applications.

3.2.2 Current Architecture: Phase Locked Loop

Although both of the previous architectures included phase locked loops (PLLs) at some point in the design, the PLL is the core of this architecture. This architecture is one that has only become possible at high frequencies due to very recent developments. The essential component in this architecture, the ADF4159 Fractional-N Frequency Synthesizer by Analog Devices, was only released in the year 2012, just one year before this architecture was explored in this research.
This architecture has four main blocks which are connected to form a phase locked loop: the reference oscillator and phase comparator, the loop filter, the voltage controlled oscillator (VCO) and the frequency divider. Similar to the tuned oscillator architecture described previously, the PLL architecture generates the waveform directly at RF, but the PLL architecture remains in a closed-loop locked configuration during operation and is inherently more stable in regards to frequency drift and phase noise.

The operation of a PLL is similar to the way a car adjusts to the speed limit. Initially, the driver notices a difference between the posted speed limit and the reading on the car’s speedometer and pushes on the gas to eliminate the difference. As the car accelerates, the speedometer reading increases and the driver applies less gas. The car may exceed the marked speed, in which case the driver applies the brakes until the speed limit is achieved. The driver maintains a constant speed by constantly comparing the car’s speedometer to the posted speed limit and making fine-tuned adjustments using the gas and brake pedals.

The gas pedal in a PLL is the tuning voltage of the VCO. The loop compares the divided down output of the VCO to the reference oscillator. If the VCO output frequency is lower than the oscillator, the phase comparator applies a positive voltage to the VCO through the loop filter and the VCO adjusts to match the frequency of the reference crystal. The loop filter determines how quickly the VCO will lock to the reference and how stable it will be once it gets there. A PLL that can quickly hop from one frequency to another will jitter at every frequency it hops to. This jittering is referred to as phase noise. A PLL that slowly transitions, however, will be very stable once it reaches a given frequency. These dynamics are controlled by the bandwidth of the loop filter, and work was done to ensure the right balance between frequency agility and stability. The magic of this system is in the Fractional-N frequency divider. This section allows the user to compare a 10GHz oscillator output to a 100MHz reference crystal by dividing by 100. Because the phase comparator locks against the output of the frequency divider to force the equality

\[ f_{\text{ref}} = \frac{f_{\text{VCO}}}{N}, \]  

(3.7)
any desired VCO output frequency can be created by changing the division ratio. A division ratio $N = 100$ forces an output frequency of 10GHz. A division ratio $N = 99$ forces an output frequency of 9.9GHz. The system generates a frequency sweep by making small, fast adjustments to the division ratio.

Because the sweep is generated in the frequency domain as opposed to sweeping the tuning voltage, the PLL architecture automatically corrects for the non-linearities in the VCO transfer curve. Although the circuitry is slightly more complex than the tuned oscillator circuitry, it can generate a clean, linear waveform and still provide adequate SWaP specifications. Figure 3.3a shows a power spectrogram of the actual chirp being generated on a log color map. Some faint harmonics are visible due to the measurement setup for capturing this waveform, but they are approximately 40dB below the fundamental. The nonlinearity was found to be below the quantization noise floor introduced by the FFT, which is approximately 100kHz.

![Figure 3.3: Spectrogram of the actual transmitted chirp generated by the PLL. Although the main portion of the chirp in (a) is extremely linear, the chirp does require approximately 80 µs of recovery time before the next chirp. This is shown in (b), which is a magnified version of the last portion of (a).](image)

The chirp is triggered by the same circuitry that controls the analog-to-digital converter (ADC). Because the chirps are triggered synchronously, the received waveforms can be
integrated coherently to improve SNR. After the chirp is complete, the PLL drops to the starting frequency of 10 GHz to wait for the next triggering signal. The settling characteristics of the PLL are shown in greater detail in Figure 3.3b. It takes approximately $80\mu s$ for the PLL to settle and be ready for the next chirp. While this delay is small, it does represent lost target dwell time, which could be averaged to increase SNR. While a continuous triangular chirp waveform would solve this issue, it was determined experimentally that asynchronous sampling of a triangular waveform introduces a phase ambiguity of the IF output. For a single channel radar, the phase ambiguity creates incoherence between chirps, lowering the achievable SNR gain from averaging. For phased array radar, phase ambiguities are catastrophic and inhibit the system’s ability to extract angle of arrival. After the necessity of phase coherence was discovered, the architecture switched from a continuous triangle chirp to the triggered sawtooth shown, despite the small losses in dwell time.

The final transmitter architecture is shown in Figure 3.4. The system uses a Hittite HMC512LP5 VCO in conjunction with the Analog Devices ADF4159 Fractional-N Frequency Synthesizer chip which controls the frequency divider and the phase comparator. The loop filter was designed in-house and has a loop bandwidth of 125 kHz. The optimal loop bandwidth was determined in simulation using the ADIsimPLL tool available online. Although the rest of the RF subsystem has strong frequency dependencies, the frequency of operation of the transmitter is determined predominantly by the VCO. Because the input bandwidth of the frequency divider chip is from 0.5 GHz to 13 GHz, this transmitter could easily be switched to the 5.8 GHz ISM band or up to 50GHz using different VCOs with built in frequency dividers.

This design is realized using ground-backed co-planar waveguide (CPWG) on 32 mil (.813mm) Rogers 4003C dielectric. To reduce fabrication costs, the same substrate was used for both the RF transceiver and the antennas. While RF design is made easier using thin dielectric, antenna design is better accomplished using thick dielectric, so 32 mil dielectric was chosen to satisfy the needs of both. In RF design, the choice of dielectric affects the size of the traces since the width of a microstrip line scales linearly with substrate thickness and dielectric constant. For this substrate the width of a 50 Ω microstrip line is approximately 72
Figure 3.4: Final transmitter architecture based on a phase-locked loop waveform generator.

mil, which is wider than some of the ICs used in this system and a major drawback for its implementation. CPWG traces can be made very small depending on the size of the gap between the center conductor and the side ground planes. Since most monolithic microwave integrated circuits (MMICs) have RF input/output transitions that are inherently CPWG, CPWG is a good choice for most RF PCB designs. A drawback of using CPWG is that many passive RF filters are realized in microstrip, so in order to realize a microstrip filters, a high quality transition from CPWG to microstrip is needed.
One important part of the transmitter is the splitter used to couple some of the transmitted energy over to the receiver in the homodyne configuration to act as the local oscillator for the receiver. Some systems use directional couplers or different configurations involving extra circulators in order to minimize excess bleedthrough from the transmitter to the receiver. We implemented the splitter using a Wilkinson power divider with an equal power split. Since the LO needed to be split several times to accommodate each phased array receiver channel, the Wilkinson splitter was able to be reused in the design.

![Wilkinson Power Splitter](image)

**Figure 3.5:** Wilkinson splitter design realized on a PCB using CPWG. This splitter performs within 0.5 dB of an ideal split (including connector losses) and provides 25 dB of isolation at the frequency of interest.

This Wilkinson splitter uses the standard topology of matching $\lambda/2$ segments with characteristic impedance $Z_0\sqrt{2}$. Because the dielectric thickness produced microstrip lines whose width was large enough to introduce coupling, the splitter was realized in CPWG. The design was inspired by an integrated circuit design by [21] and was optimized using HFSS. The parameters of optimization were: the side length, which controls center frequency; the trace widths and gaps, which control impedances and matching; and the size of the gap where the 100Ω resistor sits, which contributes to isolation. It was found that using a smaller resistor and less solder improved overall performance and matching. The final design exhibits
a total loss of 3.9 dB including the connectors. A separate measurement of a through line using the same connectors exhibits a loss of 0.4 dB, meaning that the designed splitter only deviates from the ideal by 0.5 dB. The splitter has an isolation of 25 dB over the bandwidth of interest, which mitigates some of the effects of phase noise.

The prototype of the system that is shown does not utilize any RF filters on the transmitter. The only spurs generated by this transmitter architecture are at well defined harmonic and subharmonic values of 5GHz, 15GHz, and 20GHz. Because the power amplifier and antennas are only designed for the band of frequencies around 10GHz, complexity and physical area can be saved by eliminating RF filters from the transmitter with little degradation in performance. If this were a commercial system attempting to pass stringent FCC limitations on out-of-band emissions, additional filtering stages may be necessary to ensure the harmonics do not rise above strict legal limits.

### 3.2.3 Power Amplifier and Antennas

The output power of the VCO is approximately 5 dBm, which for many of our testing scenarios was sufficient. However, in some cases there was a need for additional output power in order to increase the SNR of small targets. The additional output power was provided by either an on-board power amplifier (PA) or a higher powered external power amplifier.

The on-board power amplifier can be seen in the left side of Fig. 3.4b. After a splitter, the signal is either sent directly to an output SMA connector, or is directed through the on-board PA to a separate output SMA connector. The on-board power amplifier is the HMC451, and can be powered off when not in use. It has a gain of approximately 20 dB, which after the splitter supplies a total output power of 22 dBm (159 mW). The off-board power amplifier shown in Fig. 3.6 is the HMC952 and has a gain of approximately 30 dB. It is used on the unamplified VCO output to produce a total output power of 29.3 dBm (850 mW). It was maintained off the primary radar board for reasons of thermal dissipation.

We used a set of 4 × 1 planar endfire Vivaldi antennas designed by BYU student James Eck. The antennas and their radiation pattern are shown in Fig. 3.7. The antennas have
Figure 3.6: HMC952 off-board power amplifier with heat sink and fan to help with thermal dissipation.

(a) Single Vivaldi antenna, a $4 \times 1$ endfire array.

(b) Vivaldi gain patterns at 10.25 GHz

Figure 3.7: Planar Vivaldi antenna used as both the transmit antenna and the receiver antennas for the majority of testing.

A peak gain of approximately 12 dB and a beam width of approximately $30^\circ \times 110^\circ$. The optimal spacing between the transmit and receive antennas is a minimum of 15 cm, which
results in an average of -50dB of coupling between the transmit and the receive antennas over the frequency band of 10 to 10.5 GHz.

Coupling turned out to be a significant issue in the system, but it only became apparent once high power amplifiers were used in an effort to increase range. One effect of coupling was the phase noise at lower frequencies seen in Fig. 2.7. Although this raised the noise floor at low frequencies, the power received by close targets was large enough that it did not affect detectability of targets. Another effect of coupling was clipping in the IF waveform. The short path through the cables and between the antennas was seen by the radar as a target with a very low beat frequency on the order of a few kilohertz. The addition of this low frequency beat cosine to the rest of the radar returns introduced periodic clipping in the IF output, which effectively modulated all the signals with a square wave. In the FFT output, the modulation created equally-spaced ghost targets surrounding each true target. The number of ghost targets varied depending on the severity of the clipping, or the amount or power coupled in. The spacing of the ghost targets varied depending on the length of the bleedthrough path, which was a function of the length of the cables and spacing of the antennas. We were able to eliminate the coupling by increasing the spacing of the transmit and receive antennas and by placing absorbing material in between them.

3.3 Receiver

The key considerations in receiver design are noise and gain. These figures drive decisions of which amplifiers to use; and other design considerations such as power consumption, size and cost are balanced against how they might contribute to the signal-to-noise ratio (SNR). As soon as the signal has been captured by the antenna the amount of noise in the signal can only be increased. In this case, optimizing for size, weight, power and cost, a reasonable amount of signal integrity is sacrificed in order to achieve a system that will fit on a micro-sized UAS. This section describes the receiver architecture in detail and explains the SNR tradeoffs involved in each decision.
### 3.3.1 Homodyne Architecture

There are three potential receiver architectures for a continuous wave (CW) radar system: Autodyne, Homodyne, and Heterodyne. Autodyne systems are those where the oscillator and the mixer are the same component and are generally seen in simple doppler radar systems like police radar guns. Homodyne systems are ones where the transmitted signal is used as the local oscillator (LO) signal for downmixing the received waveform. Heterodyne systems are ones that use separate oscillators for upconverting the transmitter and downconverting the receiver. The primary difference among these three architectures is in the amount of isolation they provide between the transmitter and the receiver. Isolation is key to ensuring that the phase noise from the transmitter does not contribute significantly more noise than the receiver noise or the sky noise.

This system implements a homodyne architecture with separate antennas for the transmitter and receiver. While the homodyne architecture is favorable in terms of making the system much more compact as well as lowering the requirements of both ADC and processing needs, it does introduce significant local oscillator (LO) bleed-through. Some systems with a single antenna use adaptive cancellation techniques that subtract bleedthrough using a feedback loop. This system uses separate transmit and receive antennas, good antenna matching, and good splitter isolation in order to achieve levels of isolation that do not saturate the RF amplifiers or mixers.

### 3.3.2 Radio Frequency Subsystem

The receiver is divided into two parts, the radio frequency (RF) subsystem and the intermediate frequency (IF) subsystem. The layout of the RF subsystem and the overall block diagram are shown in Fig. 3.8. The primary focus of the RF subsystem is to establish a reasonable noise figure, minimize coupling between receiver channels, and apply sufficient gain to make up for downconversion losses. An effort was made to make the RF system as small as possible. This serves two purposes. By using fewer components in the RF subsystem,
the cost of the system decreases. This also reduces the physical area of each receiver, which helps lower the coupling between receiver channels.

![Block diagram of a single receiver channel](image)

(a) Block diagram of a single receiver channel

![PCB Layout of Receiver RF Subsystem](image)

(b) PCB Layout of Receiver RF Subsystem

Figure 3.8: Final homodyne receiver architecture.

This system uses the Avago AMMP-6222 LNA as the front-end amplifier. Because the first amplifier in a receiver establishes the noise figure for the system, extra care was taken in the layout of the amplifier. Although the amplifier is matched internally to 50Ω, the geometry of the RF input pin only matches that impedance for one specific PCB substrate material.
and thickness. Three different configurations of various tapering mechanisms were tested to determine the optimal interface in the presence of geometric mismatch. We determined that the transition method that introduced the least mismatch and loss was an abrupt transition that maintained the 50Ω CPWG trace geometry all the way to the MMIC footprint, then abruptly changed to the footprint. The same process was used with the other RF amplifiers in order to determine the optimal connection method and the same result was verified. Abrupt transitions result in the lowest mismatch loss because they minimize the electrical length over which the MMIC and the transmission lines are mismatched.

Since this is a homodyne phased array system, that means that the transmitted signal must be split equally and sent to each of the receiver channels. This is known as the clock distribution network or LO distribution network. We implemented the LO distribution network by using a series of cascaded Wilkinson splitters. The result, shown in Fig. 3.8b, is only somewhat compact, but has the advantage of being totally lossless and has high isolation among all the channels. In order to have equal LO phase delays among all channels, this solution is ideally suited to situations where the number of receiver channels is a power of two. An alternative solution is the resistive power splitter. A resistive power splitter accomplishes LO distribution in a much more compact manner, operates on a very wide bandwidth, and can be used for an arbitrary amount of channels. However, it does so at the cost of power dissipation. A resistive splitter has double the dB loss of a lossless splitter because of the dissipation through the resistors. Since our application is low-power airborne operation, we use a lossless divider; but for a ground based system with looser power constraints, a resistive power divider could be a better option.

The mixers used are the Mini Circuits MAC-12GL+ double balanced mixers. They require 4 to 7 dBm of LO drive power, so our LO distribution network integrates a preliminary amplifier in order to compensate for splitting losses. The choice of a high quality mixer is important because of the homodyne architecture. If a mixer introduces harmonics, these will appear as ghost targets at two and three times the distance of actual targets. Even if the harmonics are attenuated by 20 or 30dB, the wide dynamic range of the receiver system may
detect the harmonics for close targets. Harmonic distortion is improved by ensuring that the IF output is well matched to 50Ω, but the best way to ensure good performance is to pay the price for a high quality mixer.

### 3.3.3 Intermediate Frequency Subsystem

At the output of the mixer begins the intermediate frequency (IF) portion of the system. The goal of the IF system is to provide anti-aliasing filtering for the analog-to-digital converters (ADCs) and to amplify the signal to the necessary levels for digitization. This system adds an additional filter to mitigate the effects of antenna coupling.

All the amplification and filtering in the IF is accomplished using low-cost operational amplifiers (op-amps). Op-amps are ideal for amplification because their gain can be easily set and modified using simple resistor ratios. Op-amps are ideal for low frequency filtering because they can eliminate inductors, and use much smaller capacitor values than equivalent passive filters. This allows for more precise filter tuning, and occupies less board space than passive filters, especially at low frequencies.

The IF has three stages: the high-pass coupling rejection filter, the amplification, and the low-pass anti-aliasing filter. As is the case with all the RF receiver chains, all the IF receiver chains are identical. The PCB layout for all four is shown in Fig. 3.9a, and the schematic of each individual chain is shown in Fig. 3.9b.

The first amplifier is a high-gain, low noise amplifier using the ADA4895 op-amp. Noise is still an important consideration at this point because the mixer that precedes this stage is a lossy component, as is the impedance matching resistor R7. The noise contribution of the op-amp is derived in Appendix A and is more complex than the noise figure calculation of an RF amplifier. The amplifier is meant for high-gain applications, and has a gain of approximately 2000 V/V or 66dB. Because it has such high gain, care must be taken in the layout. One key principle to improve stability of high gain op-amps is to minimize the parasitic capacitance on the inverting pin of the op-amp. We accomplished this by removing a patch from the ground plane underneath the amplifier, and minimizing the physical size...
of the inverting node. Using these techniques, our amplifiers were stable over the full IF bandwidth.

The coupling rejection filtering was accomplished by implementing a high-pass filter. The received power is inversely proportional to the distance to the fourth power. This equates to a 40 dB/decade increase in received power as a target approaches the radar. By implementing a two-pole high-pass filter with an equal and opposite slope, targets will appear to have the same amplitude as they approach the radar. Since the primary amplitude fluctuation of the receiver is due to changes in range, this high-pass filter eliminates that difference and reduces the necessary ADC dynamic range considerably. The primary issue with coupling and close-up clutter is that since their effective range is so small, they have a huge received power that swamps other targets. By implementing this high-pass filter, the effects of coupling and close clutter returns are mitigated so that they appear to have a similar amplitude as other target returns.

The ideal implementation of the clutter rejection filter would be a filter that rises at precisely 40 dB/decade until it reaches half the sampling frequency of the ADC, where it would encounter a perfect “brick wall” low-pass anti-aliasing filter. In practice, 40 dB/decade rise corresponds with a two-pole high-pass filter. In this system, those poles are controlled by the values of the DC-blocking capacitors C8 and C10 in Fig. 3.9b. Since these are both real valued poles, the corner of the filter is not as sharp as it could be; and it may take about a decade for the filter to transition into the 40 dB/decade slope. In addition, building a system with the ideal clutter rejection response may introduce stability issues because it would have peak gain at a single frequency. As a result, this system uses a filter corner of approximately 50 kHz, which is somewhat lower than the anti-aliasing filter corner and creates a flat response region of the filter.

The anti-aliasing filter is realized using the Sallen-Key filter topology and is a two-pole low-pass Bessel filter with linear phase response in order to improve the time domain response. Since the sample rate is 2 Msamp/sec, the filter cutoff frequency is 1 MHz. It is usually advisable to place the anti-aliasing filter cutoff slightly lower than half the sample frequency,
but because the natural falloff of the radar amplitude due to the $1/r^4$ relationship is 40 dB/decade, the filter has an effective falloff of 80 dB/decade and does a good job of reducing the effect of any aliasing.

The two primary sources of signal degradation in the IF were amplifier instability and switching regulator noise. The amplifier instability occurred most in multi-stage amplifier designs where feedback between stages introduced oscillations. Once we transitioned to a design that accomplished all the gain in a single stage, stability improved significantly and resulted in a white noise spectrum. The voltage regulator noise resulted from using on-board switching voltage regulators. These regulators introduced switching noise into both the IF spectrum and the VCO tuning line. Although the IF switching noise can be removed using a digital filter, spurious noise on the VCO tuning line creates additional chirps that create ghost targets. In order to mitigate the noise introduced by switching regulators, we moved them to a separate regulator board off the main board that plugged into the battery or primary DC supply. In ground-based situations where size is not a problem, voltage regulator issues can be avoided completely by using multiple DC power supplies or linear voltage regulators that burn extra power.

3.3.4 Link Budget

The overall receiver performance can be characterized using a link budget analysis. A link budget shows the gain and noise levels at every stage of the design, and can be a helpful way of identifying trouble spots in the design. Using a link budget analysis, the designer can make decisions concerning how much gain is necessary at each stage and thus save money on expensive RF amplification. Table 3.2 provides the link budget for the system. The system gain $G_{sys}$ is the product of all the previous gains, and the system noise figure $F_{sys}$ is the combination of the measured gains and noise figures from every previous stage using Eq. (2.31). These calculations are performed using natural numbers, but the numbers listed in Table 3.2 are given in terms of dB for convenience.
Table 3.2: Calculated link budget for a single receiver channel of the system.

<table>
<thead>
<tr>
<th>Component</th>
<th>$G$ (dB)</th>
<th>$F$ (dB)</th>
<th>$G_{sys}$ (dB)</th>
<th>$F_{sys}$ (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Noise Amplifier (AMMP-6222)</td>
<td>25.3</td>
<td>3.1</td>
<td>25.3</td>
<td>3.1</td>
</tr>
<tr>
<td>General Purpose Amplifier (VMMK-2503)</td>
<td>10.2</td>
<td>5.4</td>
<td>35.5</td>
<td>3.1</td>
</tr>
<tr>
<td>Double Balanced Mixer (MAC-12GL+)</td>
<td>-8</td>
<td>8</td>
<td>27.5</td>
<td>3.3</td>
</tr>
<tr>
<td>Operational Amplifier (ADA4895)</td>
<td>50</td>
<td>17</td>
<td>77.5</td>
<td>3.3</td>
</tr>
<tr>
<td>Operational Amplifier (LMP7718)</td>
<td>15</td>
<td>17</td>
<td>92.5</td>
<td>3.3</td>
</tr>
</tbody>
</table>

One tricky aspect in calculating the link budget is the transition between the RF and the IF because of different impedance requirements. The gain and noise figure for all the RF components, including the mixer, are defined using 50 Ω termination or matching at all device inputs. This is because maximum power transfer occurs when components are matched. Op-amps are designed for voltage gain and not power gain, so they have a large input impedance and a small output impedance. Their voltage gain is a well defined ratio of the source and feedback resistors. When calculating or measuring the gain of the op-amp in the system, it is necessary to take into account the location of the op-amp in the system. If the op-amp directly follows an RF component, the resistor used to match impedance must be included on the input of the op-amp when measuring its gain. This results in an apparent decrease in expected gain of $-6$ dB due to the additional voltage drop over the matching resistor. All op-amps that follow the initial impedance transition, however, will have a measured gain that matches the analytic one.

The link budget indicates that the portion of the design that introduces the most additional noise is the low noise amplifier at the beginning of the chain. The receiver performance can be improved by using lower bandwidth amplifiers that are matched specifically to this frequency rather than using packaged, wideband amplifiers. However, since the dominant noise source in the receiver is generally clutter, excess electronics noise is acceptable inasmuch as it does not exceed the levels of the clutter.
Figure 3.9: Architecture and layout of intermediate frequency subsystem
CHAPTER 4. DIGITAL SIGNAL PROCESSING

Digital phased array beamforming improves system performance significantly. By sampling all channels, the system can form beams in multiple directions simultaneously or create a two-dimensional image from a previously one-dimensional sensor. This performance, however, comes at a cost. The cost for digital beamforming is the computational cost of a back-end processing framework powerful enough to handle the massive amounts of data produced by a digital beamformer. The back end processing framework is shown in block diagram form in Fig. 4.1 along with brackets showing where the computation is performed.

Figure 4.1: Block diagram of the radar digital signal processing. The BOARAC board acts as an interface between the radar transceiver and the DSP. The operations that involve a high data rate are performed on the MicroZed FPGA, and the operations that involve a relatively slow data rate are performed on the ARM processor.
In the initial stages of the research, the raw data was recorded, then processed in post-processing using MATLAB and Python. Using MATLAB, many of these steps were accomplished using just a few lines of code. However, performing these operations in real time on a small platform requires the use of an FPGA or some other mechanism for real-time digital signal processing. We used the MicroZed, a small, low-cost off-the-shelf development board based on the Xilinx Zynq-7000 System-on-Chip that includes an FPGA and ARM processor. As is shown in Fig. 4.1, the FPGA performs the repetitive, parallelizable processing like the FFT and correlation while the processor handles the more variable tasks of target estimation, tracking, and avoidance planning. This was where the BYU CHREC group made significant contributions to the research. They took our MATLAB code and converted it into efficient FPGA firmware capable of handling the data in real time and implemented Linux on the ARM processor so our processor codes could be written in C. They also developed the BOARAC board, a PCB to interface between the MicroZed and the radar transceiver that has eight channels of ADCs, two digital-to-analog converters (DACs), and all the necessary connections for sending the digital control commands to the radar transmitter using the serial-to-parallel interface (SPI). Although the miniaturization of the radar transceiver is impressive, the BOARAC and MicroZed boards play an equally important role in accomplishing phased-array processing in real time.

4.1 Data Acquisition, Windowing and Fourier Transformation

Although this system is a continuous wave radar system, the data acquisition is performed in discrete blocks of time synchronized with each chirp. This is demonstrated in part in Fig. 3.3a, where the chirp is triggered with a slight delay between chirps to allow time for the PLL to settle. The triggering mechanism for the transmit chirp is synchronized so that the ADC begins collecting samples at the beginning of the chirp and stops collecting samples upon completion of the chirp. This method of data acquisition enforces a relationship

\[
 f_{\text{samp}} = \frac{N_{\text{FFT}}}{T_c}, \tag{4.1}
 \]

60
where $f_{samp}$ is the sample rate of the ADC in Hz, $T_c$ is the chirp period in seconds, and $N_{FFT}$ is the number of samples gathered in a single time block. This nomenclature implies that the FFT size is the same as the number of samples collected. If the number of samples collected in a single period is less than the desired FFT size, the FFT can be performed using samples from multiple consecutive chirps or use zero padding. It is convenient to choose a power of 4 for the value of $N_{FFT}$ because it improves computational efficiency. This system uses a chirp period $T_c$ that results in $N_{FFT} = 4096$. The ADC triggering process collects the raw samples into a first-in first-out (FIFO) buffer until all $N_{FFT}$ samples have been collected for each channel so they can be loaded into the FFT block at the same time.

Optionally, before the Fourier transform operation is performed, the time samples can be multiplied by amplitude weights known as a window function. Because the ADC data is a discrete time sequence of a cosine, its discrete Fourier transform (DFT) will be a sinc function. Window functions are useful because they can reduce the sidelobes of the sinc function and reduce scalloping loss. Scalloping loss is the reduction in signal power
when the cosine frequency lands between two DFT frequency bins and its power is split between them. A signal’s amplitude can vary up to 40% due to scalloping loss. A window function expands the main lobe of the sinc, which spreads the signal across multiple bins and reduces the potential variation, but also the total signal strength and SNR. When the system’s range resolution and FFT bin width are coarse, implementing a window function can be very helpful. This system has a fine range resolution, meaning that a single target will physically occupy multiple frequency bins, so scalloping does not occur. In this situation, windowing only results in a reduction of SNR and is not helpful, though in other situations it might improve reliability.

Another potential step before the FFT is a digital filter to reduce clutter. For ground-based radar, clutter is really just a bunch of stationary scatterers the system wants to ignore. Stationary clutter has a coherent structure that allows it to be removed by a pre-whitening filter. A pre-whitening filter is a way of subtracting out a constant background in a way that allows for slight amplitude and phase variations. The filter is implemented by recording some background data and generating an autocorrelation sequence from the time-domain data. Then, using a speech compression tool such as the Levinson algorithm, the autocorrelation sequence is converted to a set of digital filter coefficients that can be implemented to invert the clutter response and remove it from future data. For clutter rejection when the radar is in motion, pre-whitening is more involved, and requires updating the filter coefficients often.

The Fourier Transform used in this system is an $N_{FFT}$-point fast Fourier transform (FFT). After a full frame of samples is loaded into the FIFO buffer and optionally multiplied by a window function or filtered, the data is loaded into an FFT module. The ADCs are 12-bit, and operate at a frequency of $f_{samp} = 2.03$ Msamp/sec. This establishes an input data rate of 3 MBps for each channel, or 12 MBps for all four channels. The FFT is performed for each channel independently and performs the 4096-point FFT in 165 $\mu$s, leaving plenty of additional time before the next frame of data since the frame interval is 2 ms. Since the ADC data is real valued, the FFT spectrum is conjugate symmetric for positive and negative frequencies, and the negative frequencies can be discarded with no loss of information.
The FFT operation produces 24-bit integers, but because most targets are incredibly weak, only the 16 least significant bits are preserved. The output of the FFT block is a set of 2048 ($N_{\text{FFT}}/2$) complex valued 16-bit integers that are passed into another buffer to await correlation.

4.2 Correlation and Integration

Correlation is a statistical operation that compares the signals from different antennas and helps to separate the signal from the noise. Although phased array beamforming is possible without correlation, correlation is the most mathematically effective way to average out the noise. Correlation is performed for each frequency bin, and initially increases the data rate, multiplying it by a factor equal to the number of antennas. However, since it is coupled with the operation of integration, the data rate can then be decreased to a rate much lower than the original rate.

The correlation process for a single time step is shown in Fig. 4.3. We use $m$ to denote the time index from one chirp to the next and $i$ for the fft bin index at a given time step. In radar literature, $m$ is often referred to as the slow time index. Using this notation, the received signal from all receiver channels for a single time step and frequency bin is

$$x_i[m] = [A_i[m] \ B_i[m] \ C_i[m] \ D_i[m]]^T,$$  \hspace{1cm} (4.2)

where $A$, $B$, $C$ and $D$ denote the FFT vectors from each of the antenna channels. The vector $x_i[m]$ can be separated into

$$x_i[m] = s_i[m] + \eta_i[m],$$  \hspace{1cm} (4.3)

where $s_i[m]$ is the deterministic signal component and $\eta_i[m]$ is the stochastic noise component.
The signal flow of the correlation process after the $m$th chirp has been recorded. The signals from the $i$th FFT bins of each channel are gathered into the vector $x_i[m]$. The correlation matrix $R_i$ is formed using the outer product, then integrated over $N_{corr}$ time steps. The result is a collection of integrated correlation matrices, one for each FFT bin.

The signal vector $s_i[m]$ is complex valued and is given by

$$s_i[m] = \begin{bmatrix}
\sqrt{G_A P_i} e^{j\phi_b} e^{jK \cdot r_A} \\
\sqrt{G_B P_i} e^{j\phi_b} e^{jK \cdot r_B} \\
\sqrt{G_C P_i} e^{j\phi_b} e^{jK \cdot r_C} \\
\sqrt{G_D P_i} e^{j\phi_b} e^{jK \cdot r_D}
\end{bmatrix}, \tag{4.4}
$$

where $P_r$ is the received signal power at that range, $\phi_b$ is the bulk phase shift derived in Eq. (2.8), the term $e^{jK \cdot r_A}$ is the additional phase shift to each antenna, and $r_A$, $r_B$, $r_C$, $r_D$ are the locations of each of the antennas with respect to the array origin. Because correlation involves conjugate multiplication, the bulk phase shift is removed from the signal leaving only the phase variation due to angle. This means that even if the bulk phase shift changes from pulse to pulse, it will have no effect on the coherence of the integration.

The noise component $\eta_i[m]$ is random in both space and time. The noise is uncorrelated in time because the thermal input noise has a white power spectral density. The relationship
between the noise at different antenna locations in space is written as the covariance matrix

\[
R_{\eta,i} = \begin{bmatrix}
\sigma_A^2 & \sigma_{AB} & \sigma_{AC} & \sigma_{AD} \\
\sigma_{BA} & \sigma_B^2 & \sigma_{BC} & \sigma_{BD} \\
\sigma_{CA} & \sigma_{CB} & \sigma_C^2 & \sigma_{CD} \\
\sigma_{DA} & \sigma_{DB} & \sigma_{DC} & \sigma_D^2 
\end{bmatrix}
\]  

(4.5)

The diagonal variance terms for each channel are the noise powers previously defined as

\[
\sigma^2 = k_B T \frac{f_{\text{amp}}}{N_{\text{FFT}}} FG,
\]

where \( F \) and \( G \) are the noise figure and gain for that particular channel. The off-diagonal terms are the covariances between each pair of antennas. Spatial noise correlation at antenna inputs arises from mutual coupling, where the presence of one antenna affects the receiver characteristics of another or from noise sources coming from a particular direction. Often the assumption is made that the noise is spatially white, meaning that the noise comes from all directions and the off-diagonal terms are zero. When this is not the case, an estimate of the noise covariance matrix is measured when no signal is present and used to calibrate out antenna coupling from future measurements.

If we assume that the signal is zero mean, then the correlation matrix is the time expectation of the received signal vector

\[
R_i = E \left[ x_i x_i^H \right].
\]

(4.7)

The expectation is approximated by integrating the correlation matrix over several time steps to produce the maximum likelihood estimate

\[
R_i \approx \frac{1}{N_{\text{corr}}} \sum_{m=1}^{N_{\text{corr}}} x_i[m] x_i[m]^H,
\]

(4.8)
as is shown in Fig. 4.3. A good value for the number of chirps to correlate, $N_{\text{corr}}$, is determined by the amount of time a target is expected to stay within an FFT bin. This is calculated using range resolution $3.5$, the expected radial velocity, and the chirp period as

$$N_{\text{corr}} = \frac{R_{\text{res}}}{vT_c}. \quad (4.9)$$

Using a value for $v$ of $50\text{m/s}$ (33.5mph) and the radar parameters listed in Table 3.1, $N_{\text{corr}}$ is calculated to be approximately 3 pulses. If the range resolution is coarser and the pulses are shorter, $N_{\text{corr}}$ can be much larger.

Correlation on the FPGA is implemented in a slightly modified way to improve computational efficiency. The correlation matrix $R_i$ is Hermetian symmetric, meaning that the off-diagonal entries in the upper right half are equal to the complex conjugates of the off-diagonal entries in the lower right half. The system exploits this by only computing and storing the diagonal entries and the upper right half and stores them in a row vector rather than in matrix form. This reduces the number of entries from 16 to 10 for a four-channel system.

Integration is interleaved with correlation to reduce the total number of operations and memory required. Rather than storing the correlation data from every pulse and integrating it after all $N_{\text{corr}}$ pulses have been acquired, the integration is done using an accumulator. The accumulator maintains a rolling sum of the pulse correlations and keeps track of the sum index. Once the sum is complete, the accumulator divides the result by $N_{\text{corr}}$. In order to perform this process more efficiently, our system rounds $N_{\text{corr}}$ to the nearest power of two so the division step can be accomplished using a bitshift operation.

The operations of correlation and integration are done in parallel for every frequency bin simultaneously. The FPGA handles both of these operations very efficiently, accomplishing correlation in just 6 clock cycles for correlation and 3 clock cycles for integration. This introduces a total latency of just 45 ns for the operations of correlation and integration. Although correlation introduces a slight increase in data rate, the integration reduces it significantly. For a value of $N_{\text{corr}} = 16$, the data rate on the output of the correlator and
integrator block is reduced from 12 MBps to 1.9 MBps as the data output interval slows from 2 ms to 32 ms.

4.3 Beamforming

Beamforming extracts direction of arrival information from the correlation matrices. This system uses beamforming to form a range/angle image of the radar field of view and applies image processing techniques in the estimation of target locations. The image is comprised of a set of equally spaced beams in angle for every range bin. This method creates rich radar visualizations, and is helpful in debugging but is not the only method of phased array beamforming. This section details the method used in this system and mentions a few alternatives.

The formulation of the weight vector begins with the normalized signal steering vector

\[
d = \begin{bmatrix} e^{j \overline{k} \tau_A} & e^{j \overline{k} \tau_B} & e^{j \overline{k} \tau_C} & e^{j \overline{k} \tau_D} \end{bmatrix}^T.
\] (4.10)

The vector \(d\) is a normalized version of the signal vector \(s_i[m]\) in Eq. 4.4, but includes only the phase components that correspond to the phase delays between each antenna position. \(d\) is formed using the relative antenna locations \(\overline{r}_A, \overline{r}_B, \overline{r}_C, \overline{r}_D\), and the incident wave vector \(\overline{k}\). Although \(\overline{k}\) in \(s_i[m]\) is the actual incident wave vector, \(d\) can be formed for any desired incident angle. For a one-dimensional linear array with \(N\) elements spaced at distances of \(d\), the signal steering vector is written as

\[
d = e^{jkd \cos(\theta)[0,1,...,N-1]},
\] (4.11)

where \(\theta\) is the angle between the incident wave and the vector normal to the array direction.

The SNR at the output of the beamformer is written as

\[
\text{SNR} = \frac{w^H R_s w}{w^H R_n w},
\] (4.12)
where $R_s$ is the correlation matrix of the signal vector $s_i$ and $R_n$ is the correlation matrix of the noise vector $\eta_i$. If we want to maximize the SNR at the beamformer output, the optimal weight vectors are found by determining the solution to the generalized eigenvalue problem

$$R_s w = \lambda R_n w. \quad (4.13)$$

One potential option for a phased array system is to stop at this point and use it as the basis for the beamformer. Rather than form a range/angle image, the system could solve this eigenvalue problem by estimating $R_s$ and $R_n$ at every range. If the solution to Eq. (4.13) produces a single dominant eigenvalue, then the corresponding eigenvector is used as the weight vector to estimate the signal strength. The angle of arrival is estimated by finding the signal steering vector $d$ that most closely approximates the eigenvector. If there are multiple dominant eigenvalues corresponding to multiple targets at that range, then the same thing can be done for their corresponding eigenvectors.

The solution to Eq. (4.13) is simplified if we make the assumption that $R_s$ is a rank one matrix of the form $\sigma_s^2 dd^H$. This means we assume that there is only one target per range bin. If the range resolution is small and there is little clutter in the environment, this is a valid assumption, even in the presence of multiple intruders. Under this assumption, the weight vector solution simplifies to

$$w = R_n^{-1} d. \quad (4.14)$$

If the noise at the array outputs is thermally isotropic and there is no mutual coupling between elements, then $R_n$ is the identity matrix and $w = d$. If this is not the case, then $R_n$ can be estimated by pointing the radar system at the sky and recording the noise correlation matrices for every range bin.

In this system, the weight vectors were formed by determining a desired number of beams, $N_{\text{beams}}$, then creating a vector of incident angles linearly spaced between $-90^\circ$ and $90^\circ$. These angles are substituted into Eq. (4.11) and Eq. (4.14) to create a matrix $W$ of $N_{\text{beams}}$ different weight vectors. For a two-dimensional phased array, the weight vectors are
formed by choosing $k$ vectors equally spaced over a hemisphere, which are substituted into the more general Eq. (4.10). These weight vectors are determined once during calibration, then used for the duration of testing. If the system has strong interference from a known source that it wishes to remove via calibration rather than processing, it can use the method of subspace projection to modify the weight vectors. This is given by

$$w_{SP} = \left(I - \frac{1}{4}d_id_i^H\right)w,$$

(4.15)

where $d_i$ is the signal steering vector for the interferer. In radio astronomy, the interferer might be a satellite or cell phone tower. In radar, the interferer might be a source of clutter (building, vehicle) that is overwhelming the signal.

Figure 4.4: The simulated output of a four-channel phased array radar system with three targets.

A range/angle image is created by forming a set of beams for every range bin. We will call the image matrix $S(i,a)$ where $i$ is the range/frequency index that ranges over $i = 1, 2, \ldots, N_{FFT}/2$ and $a$ is the angle index that ranges over $a = 1, 2, \ldots, N_{beams}$. If we let $w_a$ refer to the weight vector $w$ at angle index $a$, then the pixel intensity (power) using the
The full image $S(i,a)$ is $N_{\text{FFT}}/2$ by $N_{\text{beams}}$ pixels. If the number of elements in the array is smaller than the number of beams formed there will be smearing between column elements in a given row because the beams will overlap. This smearing is shown in Fig. 4.4. If the image is transformed to XY coordinates, then the smearing occurs radially in a banana shape. The final image may be complex valued, but at this point, only the magnitude is important.

An ideal system would have identical paths from the antenna output to the ADC input for all receiver channels. In reality, there are gain and phase variations in addition to coupling between channels, which may affect phased array performance. These effects can be accounted for and calibrated out using the transfer matrix $Q$. When the system is stable, $Q$ can be measured once, then used for the life of the system. $Q$ is measured experimentally by placing a scatterer at a known distance with the array boresight angled directly at the scatterer. Each row in $Q$ is the signal vector $x_i$ when all antenna inputs are capped except for the antenna corresponding to that row. Once all the measurements have been obtained, we can normalize the matrix to the first element since we are only concerned about relative variations between antennas. Since gain and coupling may vary as a function of frequency, in order to fully calibrate the system a $Q$ matrix must be measured for every single frequency bin. In practice, this process is simplified by measuring $Q$ at a few distances, and then averaging them. $Q$ can be simplified further to account for just the path delays by removing all off-diagonal elements and normalizing the diagonal elements to unit magnitude. The calibration is applied by modifying the analytically computed beamformer weights $w$ to create the effective open circuit beamformer weights

$$w_{oc} = Q^H w.$$  \hfill (4.17)

Although beamforming is well suited to FPGA implementation, our system implemented it on the MicroZed processor due to lack of available reconfigurable hardware.
Upgraded versions of the MicroZed have more available FPGA space, which would allow this block to be shifted off the processor. In the debugging and development phases of the project, the entire range/angle image was formed at every time step in post processing to create a video effect. In the real time version of the system, beamforming is performed only on the FFT bins that exceed a preliminary threshold in order to save computing resources.

4.4 Thresholding and Clustering

Thresholding is the point where decisions are made on what is actually signal and what is actually noise. Because the principal noise variation occurs on the range data, thresholding is done on a single range-indexed vector that combines the inputs from all antennas which we will call $q$, which has a total of $N_{FFT}/2$. In order to extract the maximum possible SNR, thresholding should be done after beamforming. In this case, the values in $q$ are the maximum values in each row

$$q_i = \max(S(i,:)). \tag{4.18}$$

If thresholding is done before beamforming, then the system suffers an SNR penalty of approximately the square root of the number of elements in the array. In this method, possible values of $q_i$ are the trace of the correlation matrix $R_i$ or some matrix norm such as the Frobenius norm.

The threshold established is a constant false alarm rate (CFAR) threshold. The CFAR threshold works by estimating the probability density function (PDF) of the noise, and then setting a threshold $k$ that will yield the desired probability of false alarm ($P_{FA}$) according to the equation

$$P_{FA} = \int_k^\infty f_{\text{noise}}(z)dz, \tag{4.19}$$

where $f_{\text{noise}}(z)$ is the PDF of the noise. If the noise is assumed Gaussian, then the noise in the range/angle image will be Rayleigh (power) and the integral is easy to solve for the appropriate threshold given an estimated variance. This variance is estimated by storing $q$ for every time step. However, the noise sources may be more complex, and a more efficient
method is the cell averaging technique. Because the receiver includes many non-ideal filters, a different CFAR threshold is calculated for every range bin as is shown in Fig. 4.5. This way, the user sets a value for $P_{FA}$ and the system automatically determines thresholds for every range bin that account for gain and noise figure variations across frequency.

Figure 4.5: The thresholding process takes the maximum value from each row of the beamformed image. These values are collected over multiple time steps to estimate the noise properties from which a CFAR threshold is calculated. The clusters which rise above the threshold are combined into a single target location estimate for each cluster.

Because the actual targets are not point scatterers, and because their beat frequency will not always fall perfectly in an FFT bin, a single target will occupy multiple frequency bins. After thresholding, this will result in clumps of targets. The location estimate $\hat{L} = (\hat{r}, \hat{\theta})$ of the target is found by taking the weighted average of the range and angle clustering, which simplifies to

\[
\hat{r} = \frac{c_2}{N} \sum_{i} q_i \hat{i} \tag{4.20}
\]

\[
\hat{\theta} = \frac{c_3}{N} \sum_{i} q_i a_s(i, \text{max(\theta)}) \tag{4.21}
\]
where the summation over \( N \) denotes summing over all indices within a given cluster. The index \( a_{S(i, \text{max}(\theta))} \) corresponds to the angle index of the maximum value in the \( i^{th} \) row of the image \( S(i, \theta) \). Because the weighted averages yield an index estimate of the true target location (which may be fractional), the constants

\[
e_2 = \frac{f_{\text{samp}} C}{N_{\text{FFT}} 2\alpha} \tag{4.22}
\]

\[
e_3 = \frac{180^\circ}{N_{\text{beams}}} \tag{4.23}
\]

convert the index estimates to numerical estimates of range and angle respectively.

After the targets have been thresholded, clustered, and centroided, the data rate is reduced significantly. Rather than large matrices that occupy megabytes of space, the data consists of a collection of ordered pairs that apply very little computational stress. Although the current system implements beamforming, thresholding, and target estimation on the MicroZed processor, future iterations will shift this load to the FPGA so that there is no burden of raw radar data on the processor and it can be dedicated to the avoidance and navigation algorithms.

4.5 Target Tracking, Risk Assessment, and Collision Avoidance

The operations of target tracking, collision risk assessment and avoidance path planning fall within the domain of control theory. They comprise the Avoid part of the Sense and Avoid system. The BYU MAGICC lab provided invaluable expertise and developed these aspects of the system. This collaboration provided a new application for radar and an interesting solution to open questions in the field of autonomous navigation. This research also produced multiple journal and conference publications on the subject [17]. This section will provide a short explanation giving the overall function and context of these operations, which are explained in much greater detail in the referenced works and the theses and dissertations of the students who worked on them.
The target tracking in this system was accomplished using the Recursive Random Sample Consensus Algorithm (R-RANSAC) developed at BYU as a way to track multiple targets in the presence of clutter [22]. The R-RANSAC algorithm works by creating a number of models that fit new data with groupings from past data over a specified time window. At each time step, the models are passed through a Kallman filter to predict the future state. If the new data fits the prediction of any models, those models are updated to include the new data. If the new data does not fit any of the models, a new model is formed by randomly selecting several past points and performing a linear extrapolation between the new data point and each of the randomly selected points. The quality of each potential model is determined by the inlier ratio: the number of past points that fit the model within a certain threshold. The randomly generated model is added to the set of existing models. The set is ordered according to the inlier ratio of each model, and the model with the lowest inlier ratio is removed from the set. Models that have been alive past a certain threshold and have the required inlier ratio are passed to the output as valid target tracks. The R-RANSAC algorithm performs well in systems with a low probability of detection by maintaining a long active window so that many measurements can be missed without discarding a good model.

After radar return is deemed to be a valid target, the system determines if the intruder actually imposes a collision risk. This assessment is made using the planned trajectory of the primary aircraft and the estimated trajectory of the intruder. In one method, the intruder’s future location is estimated by using linear extrapolation and the closest point of approach is calculated. If the closest point of approach falls within a certain safety threshold, the scenario is flagged as a collision course, and avoidance maneuvering is initiated. Another method uses a probabilistic model of the intruder location, rather than a linear extrapolation. The probabilistic model creates a sort of spatial PDF of where the intruder might fly in the future based on the uncorrelated encounter model developed by MIT Lincoln Laboratory and the flight dynamics of the aircraft in question. An assessment of the probability of collision is determined by how much the primary aircraft overlaps the intruder PDF at the closest point of approach. The probability of collision is continuously evaluated, and if it exceeds
a threshold, then an avoidance maneuver is executed. Although the probabilistic method requires more computation than the linear extrapolation method, it offers an improvement in the overall probability of correctly detecting a collision [17].

The avoidance path planning is accomplished by using a graph search algorithm. The algorithm divides the space into a three-dimensional grid of locations where it can travel. At each time step, a cost function is assigned to each of the closest nodes. The cost function includes the cost of traveling the distance to that node, the cost of deviating from the original course, the cost of assuming a trajectory that will increase collision risk, and the cost of being close to the intruder. The system chooses the node that minimizes cost at every time step. If the node discretization is coarse, then the planned avoidance path may be somewhat clunky, and it can be passed through additional processing to create a smooth flight path [23].

The target tracking, risk assessment, and path planning algorithms were all developed in the MATLAB scripting language. At the time of this publication, they were successfully integrated with the radar sensor using laptop computers. Work in the coming months will focus on the transitioning of these algorithms into C, C++, and FPGA firmware that can run on the MicroZed processor.
CHAPTER 5. EXPERIMENTAL RESULTS

The final configuration of the radar system is shown in Fig. 5.1. Although this radar was part of a larger collaborative research project, the results provided in this chapter focus only on the radar’s standalone performance.

5.1 Experimental Setup

All of the data presented in this section was collected in a ground based configuration. Since clutter ended up being the dominant source of noise of the ground based radar, testing was done in wide open fields. The stationary tests and tests involving moving pedestrians were done on the grassy hill south of the Clyde Engineering Building on the BYU campus, shown in Fig. 5.2a. The lawn was adequate for short-range tests; however the nearby hills and cars introduced clutter that overwhelmed small targets. Tests with a remote controlled UAV were done in Rock Canyon Park, near the BYU campus in Provo, Utah, shown in Fig. 5.2b. Rock Canyon Park is a bowl-shaped park devoid of trees and shrubs. It provided adequate range for initial flight testing (\(\approx 200\) m); however, the radar still received clutter returns from the grass and hills on either side and from drainage equipment. The ultimate clutterless test environment for future testing would be to use the radar on an aerial platform. Testing in both locations was generally performed early in the morning so as to minimize pedestrian traffic; however at times pedestrian traffic was a useful way to validate radar performance. As a safety precaution, none of the flight tests performed at Rock Canyon Park used fully autonomous path planning in case something went wrong with pedestrians present. Fully autonomous tests are performed in a controlled, remote location in Elberta, UT in order to minimize risk in case of system failure.
The radar test setups are shown in Fig. 5.3. For ground-based testing, the radar was always connected with a laptop computer that initiates the transmitter, logs the data, and performs target tracking, collision avoidance, and avoidance path planning. The power for the radar systems was supplied by DC laptop power supplies that were regulated to the appropriate voltages by the external regulator. In some tests, the external power amplifier for the transmitter was used, as can be seen in Fig. 5.3b. This was especially useful for the UAV flight tests since the octorotor aircraft we were using had a very small RCS ($\approx 0.05m^2$).

The majority of the time, the antennas used were the endfire Vivaldi antennas; however for some single channel tests we used horn antennas as can be seen in Fig. 5.3a. The horn antennas were useful for system calibration since their patterns were more directive and better able to avoid clutter returns. For phased array testing, the Vivaldi antennas were placed in the 3D printed cage seen in Fig. 5.3b in order maintain proper spacing and orientation between the receiver antennas. In some tests, the transmit antenna was removed from the
Figure 5.2: Radar testing sites. The radar location is indicated with a white star and the approximate beam orientation is indicated with the orange gradient.

cage and separated at an even further distance ($\approx 30\text{cm}$) in order to reduce coupling from high transmit power.

In testing, the targets used were corner reflectors, spherical reflectors, pedestrians and UAVs. Corner reflectors were used primarily in the initial stages of development since they had a very high RCS that could be detected even in the presence of clutter and high
system noise. Since initially some aircraft were below the detection threshold of the radar, corner reflector cubes, shown in Fig. 5.4a, were attached to the aircraft to enhance its RCS. The spherical reflector shown in Fig. 5.4b was used in calibration once the system reached a stable state because it has a more precise RCS that does not vary with angle. When the circumference of a sphere is greater than ten times the wavelength of operation, its RCS is independent of wavelength and is well approximated by its cross-sectional area $\pi r^2$, which was $0.073 m^2$ for the sphere used in testing. In order to reduce the effect of ground clutter, the sphere reflector was raised onto a cardboard post that provided support without contributing to the RCS. The primary aircraft we used was the 3DR X8 radio-controlled octocopter. The X8 is outfitted with a PixHawk autopilot which allows it to fly pre-programmed paths, and can be manually overridden by the observing pilot. Because the X8 is a complex structure, computing its RCS analytically is cumbersome. Its RCS was estimated by comparing it to known targets and it was found to be approximately $0.05 m^2$. The most readily available targets for testing were humans. They were helpful in measuring the progress of the system because they have a relatively low RCS ($1 m^2$ at X-band [24]) so as the system improved, human targets became more and more visible.
5.2 Stationary Target - Single Channel

Since the radar system initially began as a single-channel system, the very first tests were single channel. Additionally, the lack of data-streaming capabilities of the system made it impossible to record data for any tests with a moving target. Stationary targets were helpful as a validation tool in gauging the system’s performance, and determining areas that needed work.

The plot in Fig. 5.5 compares the radar response due to a spherical reflector on the ground with the response when the antennas are angled at the sky. Angling the antennas at the sky is a way of empirically determining the system’s internal noise contribution and validating the predicted SNR. Included in the plot are also the responses when the output is averaged over N chirps, which should lower the power of Gaussian white noise by a factor of N. The averaging is useful in determining the source of the noise.

The internal noise contribution measured by the sky pointings is a very good demonstration of the predicted internal noise sources in Fig. ??.
Figure 5.5: Comparison of radar return data when the antenna is angled at the sky vs. at a sphere reflector. Both plots are given for a single chirp and for N=100 chirps averaged.

dominant at low frequencies, and it drops beneath the thermal noise floor at high frequencies. When the thermal noise floor is lowered by averaging, the phase noise dominance extends to higher frequencies since it is coherent and unaffected by averaging. The sky pointings also demonstrate that the low frequency noise is VCO phase noise and not nearby ground clutter. This insight is helpful since phase noise is more avoidable than clutter. More careful oscillator design or a closed loop phase correction scheme as proposed by (ref/richards) would reduce the overall contribution of phase noise. If the primary bleedthrough is reflected through the LO path rather than antenna coupling as seen in Fig. ??, shifting some of the gain from IF to RF would also reduce the amount of phase noise. Since clutter is range dependent and phase noise is frequency dependent, phase noise could also be avoided by adjusting the chirp properties of the radar so that the beat frequency for a desired distance does not fall within the phase noise region. Despite its strong presence, however, phase noise was not an issue in most of our testing since the natural $1/R^4$ term provided enough gain for targets to be easily detected.
The averaging was also telling in examining external noise sources. While the averaging lowered the thermal noise in the sky pointing by an expected 15-20dB, the response of the sphere reflector was nearly unaffected. In the initial revisions of the system, we incorrectly attributed the noise seen in in the reflector pointing to poor receiver design and extremely noisy amplifiers. The fact that the noise does not decrease with averaging, and the fact that the sky pointing is devoid of this noise provides good evidence that the noise in the reflector pointing originates from clutter.

The test in Fig. 5.5 also serves as an indirect noise figure measurement and a validation of the SNR equation derived in Eq. (2.33). Noise figure can be written as

$$ F = \frac{\text{SNR}_{\text{out}}}{\text{SNR}_{\text{in}}} $$

(5.1)

when the input is at 290 Kelvin. Fig. 5.5 gives the SNR out, and the SNR in is the analytical expression Eq. (2.33) excluding F. The calculation of SNR for the sphere reflector is written out as

$$ \text{SNR} = \frac{P_t G_t G_r \lambda^2 \sigma_{\text{RCS}}}{(4\pi)^3 R^4 k_B B T_0 F_{\text{sys}} L_s} $$

(5.2)

$$ = \frac{.001 \cdot 43.2 \cdot 112.2 \cdot .03^2 \cdot .073}{(4\pi)^3 \cdot 19.8^4 \cdot 1.38 \times 10^{-23} \cdot 1.981 \cdot 290 \cdot 6 \cdot 1} = 20,852(43.2dB). $$

The SNR from Fig. 5.5 is estimated by comparing the signal strength of the sphere to the sky noise strength in that frequency bin. However, since the sky noise in that bin is dominated by phase noise, we can estimate the noise level by looking at the flat region of the noise, and taking a visual average of the level in that region. Using that technique, the plot verifies that the SNR for the sphere reflector is approximately 43 dB, which means that the system is performing as expected.
5.3 Moving Target - Single Channel

The plots for stationary targets show that the system performance is degraded severely by clutter. Often, the only method we had for distinguishing between signal and clutter was by bringing a large reflector near the target of interest in order to mark it for post processing. This is because of the fundamental issue that clutter consists of real, but unwanted radar returns. One way to distinguish the signal from the clutter is by exploiting the motion of the targets.

Figure 5.6: Spectrogram of radar aimed at moving pedestrian target ($\sigma_{RCS} = 1m^2$) on the Clyde Building lawn. Faint targets can be seen between 100m and 120m, which are from a group of people playing catch. Strong targets can be seen consistently at 130m and 140m, which correspond to cars turning the street corner where the radar was angled. The bright flash corresponds to the strong return when the car is angled perpendicular to the radar beam.
Fig. 5.6 is the radar response of a pedestrian target jogging away from and towards the radar on the Clyde Building lawn. Each column of pixels in the image is the unprocessed FFT of a single chirp, which combined form a spectrogram image of the data. Although the pedestrian target has a much larger RCS than the sphere reflector, clutter is still a visibly dominant source. However, the motion of the target makes it more easily distinguishable from the clutter.

Since the clutter in Fig. 5.6 is relatively constant from chirp to chirp, the human eye recognizes it as background and ignores it. The background subtraction performed by the human eye can be performed in software using the pre-whitening filter mentioned in Section 4.1. Fig. 5.7 shows a small airborne target before and after pre-whitening. Where previously the far section of the flight pattern is lost in the clutter, the filter makes it possible to track the target for the duration of the flight.

Figure 5.7: Radar imagery of the flight path of an X8 UAV flying at Rock Canyon park. The radar was angled at 10° above the horizon and the UAV flew a slightly diagonal path in order to stay within the antenna beam. The pre-whitening filter used two frames of radar data in the Levinson algorithm to generate filter coefficients.
5.4 Moving Target - Phased Array

The final goal of this research was to be able to track the location of a moving airborne target in both range and angle. The phased array data was recorded in Rock Canyon Park using the X8 octotor as a target. The data was recorded, and phased array data was generated in post processing using the digital signal processing scheme described in Chapter 4, excluding the avoidance navigation portion. The flight path was set by GPS waypoints and consisted of a straight line with a 10m \times 10m square deviation at a velocity of 1m/s.

Fig. 5.8 shows the phased array data from a single time index. The data was plotted as a surface rather than a 2D image to highlight the amplitude variation. The variation along the range axis is very coarse because it corresponds to the white noise variation between frequency bins. The variation along the angle axis is very smooth since we formed 100 equally spaced beams for just a four-element array. That oversampling results in very smooth angular data, even when there is just noise in a given bin. The difference in power between the signal and the noise is approximately 30-40dB, whereas the difference in power between the signal and the first sidelobe is down by approximately 8dB. This validates the approach of thresholding first in range, then angle, because the strongest noise term would correspond to the sidelobe if thresholding were performed on the whole image at once.

This data was also pre-whitened using the technique described in the previous section. The filter was an 8192 order filter that used two frames of data from every channel. After the filter coefficients were generated for every channel using the Levinson algorithm, the filter coefficients were averaged. The data from every channel was passed through the average pre-whitening filter rather than the individual filters to ensure that the pre-whitening did not affect the relative phases.

Phased array data is not visualized easily because it has three dimensions (range, angle, power) for every time step. If power is represented as a colormap, then the most representative visualization is a video with a single 2D radar image for every time frame. In printed form, however, the data must be reduced significantly in order to adequately represent it. For the following figures, the data for every time step consisted of a 2D radar
image like the one in Fig. 5.8 that is compressed to a single point of aircraft location using the clustering algorithm described in Chapter 4. Doing this reduces the dimensionality of the data and allows it to be represented in printed form. The information that is lost is the complete visualization of the noise seen in Fig. 5.8.

The radar data follows well the waypoint path described. There is some distortion of the square, which could be the result of process noise (wind), measurement noise from the GPS unit guiding the aircraft or measurement noise from the radar. It is likely that all three contributed. Calibration testing could be done by placing a target on a metal track controlled by a motor in order to ensure that there is no undesired movement of the target. In spite of the noise, the data in Fig 5.9 demonstrates successful phased array detection of a small UAV in the presence of clutter.
Figure 5.9: Phased array radar response for an X8 UAV flying a GPS waypoint pattern of a straight line with a 10m × 10m square deviation at a velocity of 1m/s. The phased array images were thresholded and reduced to a single location estimate for every time step. The plots show the location estimates transformed from range and angle to cartesian coordinates where the x direction is cross range and the y direction is down range from the phased array boresight.
CHAPTER 6. CONCLUSIONS

The goal of this research was to determine if radar is feasible as a standalone collision avoidance sensor for small UAS. Part of answering the question posed by this research involved predicting the maximum visible distance of the radar. The SNR plot in Fig. 5.5 demonstrates that the SNR of a target, and hence its maximum range is predictable in the presence of electronics and thermal noise, but not in the presence of clutter. Clutter was one of the primary obstacles in this research.

An airborne radar system doesn’t have the same clutter issues as a ground based radar system. As a result, we can conclude that this radar would work well as a mobile airborne sensor since its output SNR is predictable. However, predicting system functionality for a ground-based system is much more difficult because the clutter in ground based scenarios is unique to every environment. We demonstrated that ground clutter could be reduced by high order filtering. The test data showed that clutter limits the visible range of ground-based radar systems, which could rule it out as a detection method of small UAS. However, since the constraints of size weight and power are much more relaxed for ground based systems, it is still possible that given different system parameters it could work.

Although the UAV application presented many unique research questions, the design also represents major innovation. This system is one of the first full phased-array radar systems of its size that implements real-time correlation and digital beamforming. This success was the result of collaboration combining talent from many different fields.

6.1 Future Work

This research showed a radar successfully tracking an airborne target, so the next step in that direction would be to complete the integration of the radar and the avoidance
algorithms, and close the loop for full autonomy. The next step after that would be to demonstrate system performance when the radar is mounted on an airborne platform.

Another direction of future work would be on further miniaturization of the radar hardware. The system started as a collection of connectorized parts that we integrated onto a PCB. The next level of miniaturization would be to integrate the entire system onto a single chip. This would require significant innovation in order to maintain levels of coupling that would result in a working system. Single chip phased-array radar would make radar a viable option for any size of UAV.

More work could also be done in the phased array. This system implemented the simplest form of a phased array, the uniform linear array. This work could be extended by increasing the complexity of the array and forming a two-dimensional phased array capable of tracking targets in elevation, azimuth, and range. This may also require innovation in the antennas used, perhaps making antenna systems that are more conformal to the UAV shape. As the array size increases, it may make more sense to shift to the next ISM band at 24 GHz in order to maintain the system size within the payload restrictions of a UAV.

There is also much more work that could be done in the processing of radar data. Clutter rejection was shown briefly with a pre-whitening filter, but this could be explored in more depth to better characterize the limits of clutter rejection. More sophisticated processing could also be done by combining data from multiple sensors. Multi-sensor radar data fusion could be done by combining radar data from multiple ground sensors, multiple airborne sensors, or a mix of both.

A functioning phased array radar system was the finishing point of the research in this thesis, but my hope is that it can be the starting point of much more research.
REFERENCES


APPENDIX A. NOISE FIGURE IN OP AMPS

As was described in the introduction, when noise power is considered in a matched system, the noise power that a noisy resistor transfers to its load is given by $P_n = k_B T B$. This works well for the majority of RF components as the systems are generally matched to 50Ω. After a signal is mixed down to the intermediate frequency, however, these stages can be low enough frequency where inexpensive op-amps can be used instead of expensive amplifier ICs. Op-amps are a great option where they can be used because they allow for the design of powerful active filters and flexible gains that can be tuned to the exact desired amount. However, one drawback of using op-amps is that the analysis of the noise they add into the system is not so straightforward as with matched components. Op-amps can sometimes directly follow the mixer (a lossy component), so care must be taken to minimize the noise figure, which requires first an understanding of the causes of noise in an op-amp.

Op-amp noise figure is defined the same way as microwave components. It is the ratio between SNR at the input and SNR at the output when the input noise temperature is 290 K. Simplified, it is the ratio of total noise (input and added) to input noise as is shown below.

$$F = \frac{S_i/N_i}{S_o/N_o} \quad (A.1)$$

$$= \frac{S_i/N_i}{GS_i/ (GN_i + GN_{ai})} \quad (A.2)$$

$$= \frac{N_I + N_A}{N_I} \quad (A.3)$$

$$= 1 + \frac{N_A}{N_I} \quad (A.4)$$
There is an important distinction between this definition and the previous definition for the noise figure. This analysis uses slightly different noise terms than in the matched impedance case. Because the op-amp is a voltage driven device, \( N_A \) and \( N_I \) are put in terms of voltage noise density (rms, squared) rather than power, and have units of \( \frac{V^2}{Hz} \). This makes the analysis more straightforward but yields an identical result since \( F \) is unitless. Also, because op-amp gain is variable, the added noise is referenced to the input rather than being amplified and referenced to the output, which is what makes the gain term cancel out in equation A.3.

Op-amps have two important figures of merit that are listed on every data sheet: input voltage noise and input current noise. As is shown in the figure below, when defining these specifications, op-amp manufacturers lump all noise voltage generated by the device to the non-inverting terminal. Current noise is sometimes defined for each terminal individually, but generally one figure is given which represents a current noise source at each terminal inside the op-amp that excites a noise voltage in the equivalent resistance seen at each terminal. Additionally, each surrounding resistor contributes voltage noise to the system. The amount that each contributes is determined by how much of that noise voltage reaches the input terminals after it has been divided by the other resistors. Thus, for an op-amp there are three sources of noise: resistor voltage noise, op-amp voltage noise, and voltage noise that the op-amp current noise excites across the resistors. At high resistance values, the voltage noise from the resistors themselves in addition to the voltage noise created by the op-amp noise currents become the dominant sources of noise. At low resistance values, the voltage noise from the op-amp becomes the dominant noise source. The optimum approach in design is to analyze the contributing noise sources and reduce the dominant source to the level of the others.

We will provide the equations for noise in an inverting amplifier, much of which is drawn from[25].. The inverting and differential amplifiers entail a much more involved analysis because noise gain and signal gain are different and must be referred to the point of signal input to the amplifier. For a more detailed explanation of this subject, refer to that
document. Although the entire equation may seem rather daunting, each of the noise terms is rather easy to come by and can be easily re-derived.

- The op-amp noise voltage density is immediately present at the input terminal so it is added without scaling to the total added noise term $N_{aii}$ in the non-inverting case, but becomes a bit trickier for the inverting case. Generally, this number is given in units of $\frac{nV}{\sqrt{Hz}}$, so this term is squared.

- The op-amp noise current density is modeled as a current source at each terminal. Hence, op-amp current noise contributes two terms to the total added noise. These terms are equal to the noise current density squared times the equivalent resistance seen at each terminal. The equivalent resistance is the parallel combination of all resistors that enter that node. The Miller Effect come into play because this is a voltage-sensing current-feedback amplifier, so the current is not being amplified.

- There will be as many resistor noise terms as there are resistors; however, the actual resistor noise voltage that arrives at the input terminals of the op-amp is only a fraction of the noise of each of the resistors, as some of that noise voltage will be dissipated in other resistors before it reaches the op-amp input terminals. To determine the amount that reaches the input terminals, treat the noisy resistor as a voltage noise source in
series with noiseless resistor. Then, perform the voltage division between the noiseless resistor and the rest of the resistors to see how much of the voltage noise source arrives at the terminal.

\[ N_I = 4k_B T R_S \left( \frac{R_T}{R_S + R_T} \right)^2 \] Source thermal noise present at non-inverting terminal

(A.5)

\[
N_A = \begin{cases} 
  v_n^2 & \text{Op-amp noise voltage} \\
  + i_n^2 \left( \frac{R_S R_T}{R_S + R_T} \right)^2 & \text{Op-amp noise current at non-inverting terminal} \\
  + i_n^2 \left( \frac{R_F R_G}{R_F + R_G} \right)^2 & \text{Op-amp noise current at inverting terminal} \\
  + 4k_B T R_T \left( \frac{R_S}{R_S + R_T} \right)^2 & \text{Noise from } R_T \text{ seen at non-inverting terminal} \\
  + 4k_B T R_G \left( \frac{R_F}{R_F + R_G} \right)^2 & \text{Noise from } R_G \text{ seen at inverting terminal} \\
  + 4k_B T R_F \left( \frac{R_G}{R_F + R_G} \right)^2 & \text{Noise from } R_F \text{ seen at inverting terminal} 
\end{cases}
\]

(A.6)

In our system, the op-amp that directly follows the mixer in the IF uses the ADA4841 chip and the following values:

- \( R_S = 50\Omega \) - The IF output of the mixer has an impedance of 50\( \Omega \).
- \( R_T = 50\Omega \) - Since we are using a double-balanced passive mixer, IF impedance matching is imperative in order to ensure optimal performance.
- \( R_G = 470\Omega \) and \( R_F = 4.7k\Omega \) - Noise performance also improves as these values are lowered, but since we were using a single-supply configuration and needed a DC decoupling capacitor connected to \( R_G \), we had to strike a balance so that the capacitance wasn’t too large without having too high of a corner frequency.
- \( v_n = 2.1 \frac{nV}{\sqrt{Hz}} \) and \( i_n = 1.4 \frac{nA}{\sqrt{Hz}} \) - Per spec of ADA4841 datasheet.
These values gave a simulated noise figure of $17.8\,dB$. The dominant noise terms were the op-amp noise voltage term and the thermal noise from $R_G$. An important nuance to note is that by increasing the value of $R_T$ closer to the value of the input impedance of the op-amp (or removing it altogether), the noise figure drops by $6\,dB$. This is because the op-amp itself is a high impedance load. By adding a matching resistor, the voltage of both the signal and the noise drop to $1/2$ of the value that was output from the previous stage, and hence the power drops to $1/4$. If the input impedance of the op-amp were $50\,\Omega$, then having a matched configuration would provide optimal power transfer. However, since the op-amp is a high impedance load, the addition of a $50\,\Omega$ matching resistor acts as a -$6\,dB$ attenuator between the two stages and most of the power is burnt in the resistor. In our application, the amplifier directly follows a double-balanced passive diode mixer. Because of the nature of that type of mixer, it is imperative to match the IF output to $50\,\Omega$. Because of this, we are obligated to use a $50\,\Omega$ matching resistor and accept the $6\,dB$ of degradation in noise figure. Because of the design of our system, however, an IF noise figure of $20\,dB$ is acceptable, although anything much higher than that would begin to degrade performance.