SIGNAL PROCESSING METHODS FOR THE DETECTION & LOCALIZATION OF ACOUSTIC SOURCES

VIA

UNMANNED AERIAL VEHICLES

by

© Brendan Harvey

A dissertation submitted to the School of Graduate Studies

in partial fulfillment of the requirements for the degree

of

Doctor of Philosophy (Ph.D.)

Faculty of Engineering and Applied Science

Memorial University of Newfoundland

March 2018

ABSTRACT

This thesis provides an in-depth examination of utilizing acoustic sensing to form the basis of a non-cooperative collision avoidance system for Unmanned Aerial Vehicles (UAVs). Technical challenges associated with the development of such a system in the areas of acoustics, kinematics, statistics, and digital signal processing are clearly identified, along with the requirements of such a system to be commercially viable. Theoretical developments in the areas of adaptive filtering, signal enhancement, signal detection, and source localization are proposed to overcome current limitations of the technology and ultimately establish a practical and viable sensing system. Each of the proposed methods were also evaluated using both computer generated and experimental data.

A number of techniques to adaptively filter harmonic narrowband noise without using any reference signal or producing any phase distortions are proposed. These included: 1) A distortionless FIR notch filtering method via the use of a second-order IIR notch filter prototype, 2) A distortionless notch filtering method via the use of FIR Comb filters, and 3) Multichannel adaptive filtering methods for systems containing multiple harmonic noise sources.

Several signal processing techniques to enhance the detection of continuous harmonic narrowband signals are proposed. These methods included: 1) A generalized spectral transform to exploit the periodic peak nature of harmonic signals in the frequency domain, 2) A series of processors which exploit the phase acceleration properties of continuous periodic signals, and 3) Modifications to the generalized coherence function for multichannel systems to include phase acceleration information. In addition to the proposed signal enhancement processors, Constant False Alarm Rate (CFAR) detection relationships for unknown signals residing in noise with fixed bandwidth regions and unknown properties are also provided. These include: 1) The establishment of distribution-free CFAR relationships for non-independent testing scenarios, 2) Development of a distribution-free CFAR detector through frequency tracking of consecutive windowed spectra, 3) Development of a Robust Binary Integration scheme to better facilitate the detection of non-stationary signals, and 4) A CFAR-Enhanced Spectral Whitening technique to facilitate the accurate use of distribution-free CFAR detectors with non-identically distributed noise. An examination of the statistical and kinematic requirements to establish a reliable UAV collision avoidance system is also provided. This include a brief analysis to determine minimum required detection probability rates, and the development of an analytical model to approximate minimum required detection distances.

A beamforming method is proposed to enhance the localization accuracy of harmonic continuous source signals via the Steered Response Power (SRP) method. In addition, algorithms are developed to reduce computational loads associated with the SRP localization technique. These include a Crisscross Regional Contraction method, and an adaptive approach which utilizes the steepest ascent gradient search. In addition, it was also shown that

by performing signal detection prior to beamforming, greatly reduced computational loads and increased localization accuracy can be obtained.

Finally, a number of experiments were conducted to establish the overall viability of an acoustic-based collision avoidance system and verify the performance of the proposed signal processing techniques. These included: 1) The detection of a continuous ground-based stationary source from a moving fixed-wing UAV, 2) The detection and localization of a moving unmanned aircraft from a moving fixed-wing UAV, and 3) The detection and localization of a moving manned aircraft via a moving multi-rotor UAV. Based on the results obtained, it was found that both manned and unmanned aircraft were detected and localized with sufficient range and accuracy to avoid a collision. Thus, it was finally concluded that acoustic sensing does in fact appear to be a viable technology to establish a non-cooperative collision avoidance system for UAVs.

ACKNOWLEDGEMENTS

I would like to express my most sincere gratitude to my supervisor Dr. Siu O'Young for his guidance, support, and providing me with the opportunity to pursue this research endeavour; without him none of this would have been possible.

To my supervisory committee: Dr. Michael Hinchey and Dr. Reza Shahidi, I thank you for taking the time to review this admittedly long thesis and provide me with the necessary insight and guidance to achieve this academic goal.

To my fellow Ph.D. research students of the RAVEN project: Jordan Peckham, Kevin Murrant, Yake Li, and Iryna Borshchova, and the RAVEN staff: Scott Fang, Dilhan Balage, Stephen Crewe and Noel Purcell, I am thankful for all the support you have provided and facilitating the conducting of various field experiments.

I would also like to thank all my family and friends for their love, support, and encouragement throughout these past few years; it has been a very long and hard road but as my mother always used to say, "if the goal isn't hard to achieve then its probably not worth your time to attain it".

Finally, I would like to thank my father Larry Harvey for instilling me with the value and importance of knowledge and education at very young age; if were not for you I would have never been able to attain this level of academic achievement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

A-AVC	Adjusted Acceleration Vector Coherence
A-SAC	Adjusted System Acceleration Coherence
AVC	Acceleration Vector Coherence
BI	Binary Integration
CA-CFAR	Cell Average CFAR
CDF	Cumulative Distribution Function
CDF-CFAR	Constrained Distribution Free CFAR
CFAR	Constant False Alarm Rate
CHST	Complex Harmonic Spectral Transform
CPA	Closest Point of Approach
CPLE	Coherent Phase Line Enhancer
DF-CFAR	Distribution-Free CFAR
DFT	Discrete Fourier Transform
DOA	Direction of Arrival
FBP	Fluctuation Based Processing
FD	Frequency Domain
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FT	Fourier Transform
GASC	Generalized Acceleration Squared Coherence
GCC	Generalized Cross-Correlation
GHST	Generalized Harmonic Spectral Transform
GLRT	Generalized Likelihood Ratio Test
GMSC	Generalized Magnitude Squared Coherence
HPBW	Half Power Beam Width
HPS	Harmonic Product Spectrum
HSB	Harmonic Spectral Beamformer
HST	Harmonic Spectral Transform
IDFT	Inverse Discrete Fourier Transform
IFT	Inverse Fourier Transform
IIR	Infinite Impulse Response
IRT	Impulse Response Truncation
MAP	Maximum a Posteriori
MCP	Modified Coherence Processor
MFD	Mean Frequency Deviation
MHST	Multichannel Harmonic Spectrum Transform
MIMO	Multiple Input / Multiple Output
ML	Maximum Likelihood
MLE	Maximum Likelihood Estimates
MMSE	Minimum Mean Square Error
MSC	Magnitude Squared Coherence
MSE	Mean Squared Error
MSNR	Maximum Signal-to-Noise Ratio
MVDR	Minimum Variance Distortionless Response
NP	Neyman-Pearson
OS-CFAR	Order Statistic DFAR
PAC	Phase-Aligned Coherent Processor
PAP	Phase Acceleration Processor
PAV	Phase-Aligned Vector Processor
PDF	Probability Density Function
PSD	Power Spectral Density
PVC	Phase Vector Coherence
RBI	Robust Binary Integration
RMS	Root Mean Square
ROC	Receiver Operating Characteristic
SAC	System Acceleration Coherence
SAA	Sense-and-Avoid
SC	Single Cell
SCDF-CFAR	Selective-cell Constrained Distribution Free CFAR
SIMO	Single Input / Multiple Output
SNR	Signal-to-Noise Ratio
SPL	Sound Pressure Level
SRP	Steered Response Power
ST	Single Trial
TDOA	Time Difference of Arrival
TPT	Total Processing Time
UAV	Unmanned Aerial Vehicle
ULA	Uniform Linear Array
UMPT	Uniformly Most Powerful Test

NOMENCLATURE

	Chapter 3		
θ	Normalized frequency	G	Shifting gain
N	Number of filter coefficients	f_o	Fundamental frequency
f_s	Sampling frequency	M	Number of harmonics
ω	Angular frequency	S	Number of sources
K	Number of signals	μ	Adaptive step size
r	IIR notch radius	β	LMS Gradient
BW	Notch bandwidth	σ	Variance
	Chapter 4		
R	Number of signal harmonics	ϕ	Phase acceleration
W	Number of FFT windows	β	Modulo 2π scaling factor
S	Number of signals	Ψ	Modulo 2π exponential factor
$\overline{\mathrm{H}}_a[\,]$	HST	$\vec{\Theta}(f)$	PVC
$\overline{\mathrm{H}}_{<a,b,c>}^{<R,S,W>}[\,]$	GHST	$\vec{\Phi}(f)$	AVC
$\Gamma(f)$	MSC	$\vec{\Phi}^{\Psi}(f)$	A-AVC
$\tilde{\Gamma}(f)$	GMSC	Φ_λ	SAC
$\tilde{\Gamma}^{\Psi}(f)$	GASC	Φ_λ^{Ψ}	A-SAC
H_o	Null Hypothesis	H_1	Alternative Hypothesis
P_{fa}	False alarm probability	P_{FA}	Cumulative false alarm probability
P_d	Detection probability	P_D	Cumulative detection probability
P_{FA}^{SC}	Single cell false alarm	P_{FA}^{ST}	Single trial false alarm
P_{FA}^{BI}	Binary integration false alarm	P_{FA}^{RBI}	Robust binary integration false alarm
η	CFAR threshold	α	CFAR scaling factor
α_{ca}	CA-CFAR scaling factor	α_{os}	OS-CFAR scaling factor
k	Order statistic	$\bar{k}$	Reversed order statistic
X_c	Test cell value	X_k	Order statistic cell value
$\vec{N}$	CFAR noise sample set	N	Number of CFAR noise samples
$\vec{B}$	CFAR test sample set	B	Number of CFAR test samples
$\vec{G}$	CFAR guard cell set	G	Number of CFAR guard cells
T	Number of trials	D	Number of detections
I	Number of interfering targets	F	Number of fractional harmonic peaks
G_F	Fractional peak guard cells	M	Number of tracked maxima
Δ	RBI cell deviation	R	Number of signal harmonics
ξ	Recursive mean forgetting factor	δ	Spectral whitening flooring factor
f_{res}	Frequency spectra resolution	f_o	Source fundamental frequency
	Chapter 5		
S	Number of signals	R	Number of Harmonics
$\hat{k}$	Steering vector	τ	Time delay
ϑ	Azimuth	φ	Elevation
f_s	Sampling frequency	J	Cost function
μ	Adaptive step size	ε	Error function
j-$\overline{\mathrm{H}}_a[\,]$	CHST	d	Array element spacing

- 1 - Introduction

1.1 - Problem Statement & Motivation

Unmanned Aerial Vehicles (UAVs) are a rapidly advancing technology with many applications in the private, commercial, and government sectors. Currently, no safeguards exist to facilitate the safe operation of these devices in populated uncontrolled airspace without posing potential hazards to other manned or unmanned aircraft. Thus, the successful integration of these devices within the constructs of a commonly shared aviation system will ultimately require a level of safety equivalent to that of manned aircraft [1]. Conventional anti-collision systems require the successful communication between neighboring aircraft (cooperative system) with the pilot acting as the last line of defense in the event of a system failure. However, current regulations only require passenger aircraft greater than 5,700 kg to be equipped with such avoidance systems [2]. Since autonomous UAVs do not meet these requirements and do not have the benefit of an onboard pilot, non-cooperative systems must be established to facilitate the detection and subsequent avoidance of other approaching aircraft.

A number of technologies are currently being investigated to develop a UAV based Sense-and-Avoid (SAA) system. The most popular include electro-optic, infra-red, and radar. However, each of these technologies currently has major drawbacks which has limited successful development thus far [3]. For example, electro-optic and infra-red both suffer from a narrow field-of-view and their performance is greatly reduced in situations where fog or cloud cover may be present. This can be a serious problem since most mid-air collisions do not occur head-on but rather from behind, the side, above, or below; often in unfavourable weather conditions [4]. Radar does not suffer from the drawbacks of optical methods. However, in order to achieve the required detection distances, a great deal of power is required making the device and supporting equipment too large and heavy for most UAVs.

It is believed that acoustic sensing can facilitate a non-cooperative SAA system without being subject to the drawbacks associated with current conventional technologies. In theory acoustic sensing is capable of omni-directional detection and localization in all weather conditions. It is also a passive technology, with low power, size, and weight requirements. Thus, the overall objective of this research project is to establish the viability of utilizing this technology to form the basis of a UAV aircraft anti-collision system. In order to achieve this goal, two major criteria must be met: 1) The intruding aircraft must be detected at a distance adequate to facilitate an avoidance maneuver, and 2) The sensing aircraft must be able to establish a basic spatial position (azimuth & elevation) and trajectory of the intruder once detection has been achieved. The remainder of this dissertation will therefore examine the kinematic, acoustic, and signal processing requirements to establish such a system, which will also be verified through physical experimentation.

1.2 - General System Description

Consider the general case pertaining to a potential mid-air collision between two aircraft as depicted below in Figure 1-1. The system consists of an intruding aircraft which emits some unknown acoustic signal, and a detecting aircraft fitted with a microphone array. Both aircraft are assumed to be in continuous motion with constant headings and velocities, and are separated by distances large enough such that incident waves arriving at the sensors may be treated as planar with linear fronts. Translating coordinate frames are fixed to both aircraft and give locations with respect to the GPS coordinate system. Each translating frame also contains a rotated coordinate reference which provides kinematic information relative to the respective aircraft orientation (yaw, pitch, roll). Typically, only information projected along a direct line-of-sight vector connecting the two aircraft is of concern since this component ultimately governs characteristics of the received source signal. Relative separation distances will largely dictate source attenuation levels, with atmospheric effects also providing some typically unknown contribution. Relative motion between the aircraft will produce variations in the observed source frequency due to Doppler effects as will be later described. Self-noise generated by the sensing aircraft will effectively corrupt acquired signals and thus influence the ability to perform detection and localization operations. The ultimate goal of the system is to utilize the known kinematic information of the sensing aircraft in conjunction with measured data such as acquired signal amplitude, spatial phase characteristics, and perceived source frequency, to determine information about the unknown acoustic source such as location and velocity. With respect to the above considerations, the overall system may therefore be described in terms of its acoustic properties, kinematic relationships, and signal acquisition/processing requirements. Each of these areas is further discussed below to provide the reader with the basic background knowledge required to effectively interpret the methods and solutions proposed throughout the remainder of this dissertation.

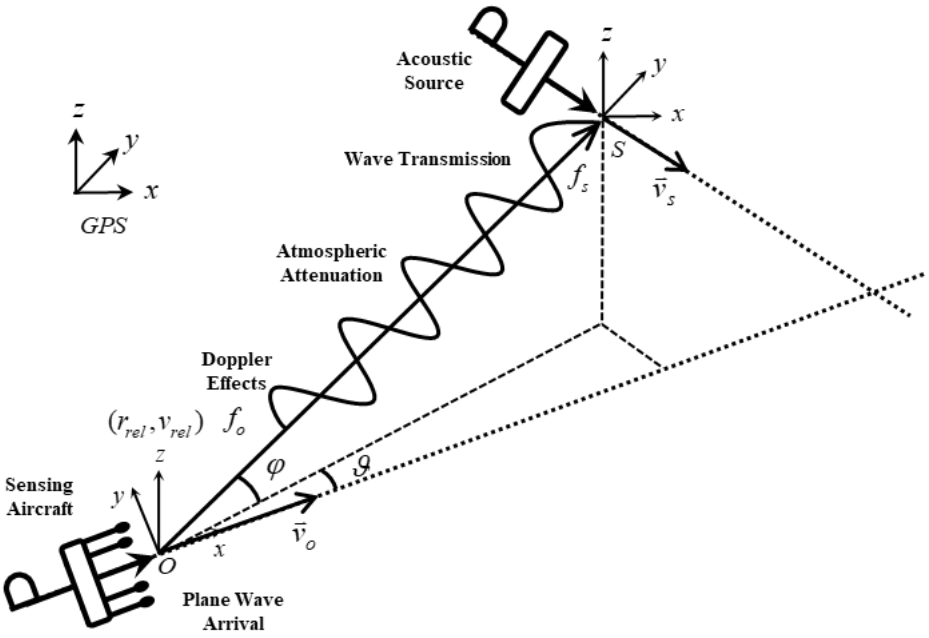

Figure 1-1: Depiction of the kinematic and acoustic properties of a mid-air encounter.

1.3 - Technical Challenges

Much like other potential SAA technologies, acoustic sensing does have its drawbacks and associated challenges. Perhaps the greatest problem is the inherent self-noise generated by the sensing system as a whole. Strongly correlated narrowband and randomly distributed broadband components often produce very low signal-to-noise ratios (SNRs). This greatly limits the ability to achieve adequate detection distances and perform more advanced operations such as spatial localization. Moreover, self-generated noise signals tend to be highly non-stationary, greatly increasing the difficultly in removing them. Thus, utilizing UAVs as an acoustic sensing platform is generally considered a very difficult task. The main reasons why this is the case can be summarized as follows:

1) The acquired signal power of the self-generated noise components is much larger than that of the source to be detected. Thus, unfiltered signals will have very large negative SNR values making the problem of unknown source detection extremely difficult.
2) The sensing system has a very dynamic in nature since engine RPM and airspeeds often vary considerably over time. These changes produce highly non-stationary self-noise components which must be removed via some active filtering approach. However, the properties of the system prevent the establishment of a noise-only reference sensor to facilitate standard active filtering methods.
3) All signal processing operations performed must be done so without producing any phase distortions between acquired signals since this information is required for source localization operations.
4) The acoustic source signal to be detected will often have similar component frequencies and will also be non-stationary if engine speeds vary. If source component frequencies equal that of the self-generated noise, they will not be detectable.
5) Relative velocities between the detecting and intruding aircraft will effectively cause frequency shifts producing a perceived non-stationary signal regardless of the true level of stationarity at the emitting source.
6) Atmospheric conditions such as humidity, wind, and temperature differentials cause unpredictable acoustic attenuation and directivity properties which will ultimately affect detection distances to some degree.

Despite these technical challenges, it will be shown that a combination of physical noise reduction steps in conjunction with advanced digital signal processing techniques can be implemented to achieve appreciable detection and localization capabilities. Each of the identified technical challenges and proposed solutions will be further discussed in future sections.

1.4 - Research Overview & Contributions

The presented research consists of theoretical developments in the area of digital signal processing which are evaluated using simulated data and verified through physical experimentation. Developments are made to address the issues outlined in the previous section and maximize the performance of the proposed technology for the specific application at hand. Table 1-1 displayed below provides an overview of the theoretical contributions made in the areas of filtering, enhancement, detection, and localization.

Table 1-1: Research overview and list of contributions.

Adaptive Filtering	
Referenceless adaptive notch filtering methods for multichannel systems.	1) Multichannel IIR notch filtering methods for SIMO and MIMO systems. 2) A distortionless zero-phase FIR notch filtering method via the use of a second-order IIR notch filter prototype. 3) A distortionless zero-phase notch filtering method via the use of FIR Comb filters.
Signal Enhancement & Source Detection	
Frequency domain spectral enhancement processors to increase source detectability.	1) A generalized spectral transform to exploit the periodic peak nature of harmonic signals. 2) A series of processors which exploit the phase acceleration properties of continuous periodic signals. 3) Modifications to the generalized coherence function for multichannel systems to include phase acceleration information.
Statistical and kinematic requirements of a non-cooperative collision avoidance system.	1) A statistical evaluation of the collision avoidance requirements for UAVs. 2) A kinematic model to approximate the minimum required detection distances to avoid a collision.
CFAR detection relationships for unknown signals residing in fixed bandwidth regions of unknown noise properties.	1) Establishment of distribution-free CFAR detector relationships for non-independent testing scenarios. 2) Establishment of a distribution-free CFAR detector through frequency tracking of consecutive windowed spectra. 3) Development of a Robust Binary Integration scheme to better facilitate the detection of non-stationary signals. 4) Modifications for CFAR detectors to facilitate accurate usage with harmonically transformed spectra. 5) A CFAR-Enhanced Spectral Whitening technique to facilitate the accurate use of distribution-free CFAR detectors with non-identically distributed noise.
Source Localization	
Efficient frequency domain source localization and tracking methods.	1) Development of a criss-cross regional contraction method for efficient 3D source localization via the Steered Response Power beamformer. 2) Establishment of an adaptive localization and tracking method via the steepest ascent approach. 3) Development of a frequency domain beamforming technique which exploits properties of harmonic source signals to enhance localization capabilities.

1.5 - Dissertation Outline

The remainder of this dissertation is organized as follows: Chapter 2 provides the necessary background information required to gain an understanding of the demands, technical challenges, and analytical techniques required to effectively develop and test an acoustic-based collision avoidance system. In addition, a review of the current literature in the area of non-cooperative collision avoidance technologies is also provided. Chapter 3 presents several techniques to adaptively filter harmonic narrowband signals without the use of any reference signal and without producing any phase distortions. The proposed methods are evaluated and validated using simulated and experimental data. Chapter 4 provides developments made in the areas of signal enhancement and source detection. The chapter is partitioned into two main sections: The first presents several spectral enhancement processors which utilize phase acceleration and coherence while exploiting the properties of harmonic signals to increase source detectability. The second provides an analysis and development of CFAR detection schemes for unknown signals in noise of unknown properties. Simulation studies are also provided to evaluate the proposed enhancement processors and detection schemes. Chapter 5 provides a discussion of common source localization methods and outlines the approach best suited for the application at hand. Computationally-efficient 3D localization methods are proposed along with a frequency domain beamforming method, which exploits properties of harmonic signals to enhance localization capabilities. Chapter 6 provides details of the experimental studies conducted to evaluate the overall suitability of acoustic sensing to form the basis of a non-cooperative UAV collision avoidance system. Results are also provided to validate the proposed signal processing developments presented throughout the dissertation. Finally, Chapter 7 provides the overall conclusions based on the experiments conducted and provides recommendations regarding future work in this area.

- 2 - Background Information

The following chapter provides a review of current literature available on the topic of non-cooperative UAV collision avoidance systems. A brief review of conventional technologies is provided along with an in-depth examination of related work conducted in acoustic sensing and localization. A clear technical description of the investigated system is presented in context of acoustics, kinematics, signal acquisition, and digital signal processing. The technical challenges associated with developing an acoustic-based aircraft anti-collision system are presented along with a clear outline of the proposed solutions.

2.1 - Unmanned Aerial Vehicles

The term Unmanned Aerial Vehicle (commonly termed drones), simply refers to any aircraft which may be piloted remotely or fly autonomously and does not carry any human operator onboard. They range from simple electric powered hand operated short-range systems, to turbofan powered long endurance high altitude systems that require a traditional airstrip for operation. In addition, UAVs may consist of both fixed-wing and rotary-wing design configurations.

Contrary to popular belief, UAVs are not a new technology with historical accounts of operational systems dating back to the early 1900's. First developments in the field date back to WWI when the U.S. developed a pilot-less aircraft known as the "Kettering Bug", which essentially acted as a timed flying bomb that would release its wings and fall to earth after some pre-programmed period of time. During the 1930's, the British developed and produced more than 400 unmanned vehicles for target practice purposes. These vehicles where known as "Queen Bees" and would later coin the popular UAV term "drone". However, it wasn't until the 1990's that UAVs became familiar to the general public as they gained acceptance as a useful military tool. The conflicts in the first Iraqi war and later in the Balkans ushered in a new era for UAVs giving them mass media exposure; this exposure further increased during the most recent conflicts in Afghanistan and Iraq [1].

Today, UAVs have reached unprecedented levels of growth as interest continues to expand worldwide. Recent advances in computer technology, software development, lightweight material manufacturing, advanced data links, and sensing technologies are strengthening capabilities and further fueling demand through increased application potential. Many countries across the globe are now developing UAVs for military, civil, and commercial uses with hundreds of diverse models now having been produced. Civil government function will probably compose the majority of future UAV usage. These applications would address many of the functions currently provided by manned aircraft but offer greater endurance and lower operating costs. Typical applications may include: emergency response, law enforcement surveillance, search and rescue, forest fire monitoring, illegal hunting, communications relay, flood mapping, high altitude imaging, nuclear, biological, chemical (NBC) sensing/tracking, traffic monitoring, humanitarian aid, land use mapping, chemical/petroleum spill monitoring, border patrol, monitoring of sensitive sites, drug trafficking surveillance and prevention, domestic traffic surveillance, and coastal port security. The commercial industry will also see an increased

number of potential applications for UAVs once better regulatory infrastructure and more affordable systems are established. Potential commercial uses for both large and small UAVs may include: crop monitoring, utility inspection, news and media support, aerial advertising, urban cargo delivery, surveying and mapping, commercial imaging, and business security to name a few [5].

UAVs can generally be categorized as being either fixed-wing or rotary-wing (multi-rotor), with conventional or pusher style propulsion system configurations. Figure 2-1 provides a depiction of each aircraft type including a typical array configuration. Pictures of the actual aircraft used for experimental studies are later presented in Chapter 6. Studies were conducted using both forms of aircraft since each type provides different associated benefits and technical challenges.

Fixed-wing UAVs generally contain a single propeller-based propulsion system which may be located at either the front (conventional) or rear (pusher) of the aircraft. A continuous forward motion is required to achieve flight which in turn produces the need for some form of airstrip to facilitate landing and take-off operations. Since these aircraft are typically much larger than the multi-rotor type, the establishment of acoustic arrays are generally much less constrained in terms of possible geometric configurations; a property which ultimately governs array performance for a given signal frequency and fixed sensor quantity. In addition, larger spatial availability permits the placement of microphones further from the propulsion system which is the major contributing noise source. In such respects, the pusher configuration would thus generally be preferred over the conventional style. Although fixed-wing aircraft allows greater variability in array configuration, the continuous forward motion required to maintain flight also generates high velocity airflow past the microphone sensors. This in turn may generate considerable amounts of noise in the acquired acoustic signals.

In contrast, multi-rotor UAVs contain multiple vertically oriented and equally opposed lifting fans to remain airborne. They have the benefit of not requiring any directional velocity to achieve flight and do not require an airstrip to facilitate takeoff and landing operations. Moreover, they may traverse and/or rotate in essentially any desired direction creating a much higher degree of maneuverability. It is because of these reasons that multi-rotor UAVs are becoming much more popular and widely utilized than the fixed-wing variety. As with fixed-wing aircraft, multi-rotors may be configured in the conventional lifting or alternative pusher style configurations. Since these aircraft do not require continuous motion to produce flight and velocities present during typical operations are relatively low, flow-generated noise is typically of much less concern. However, size constraints and the presence of multiple lifting fans mean microphones will inherently be located relatively close to multiple high-level noise sources which may also be operating at different frequencies. As with fixed-wing aircraft, multirotors consisting of the pusher configuration will generally be preferred since there is greater freedom in microphone placement, and sensors can be located further away from the high-speed downward airflow generated by the lifting propellers.

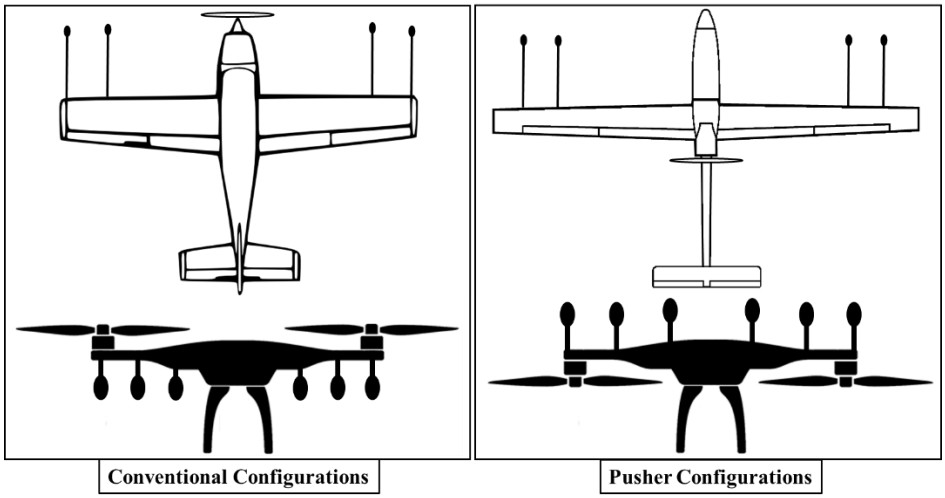

Figure 2-1: Depiction of various UAV configurations.

2.2 - Collision Avoidance Systems

2.2.1 - Cooperative Systems

Cooperative SAA technologies are those that require the successful transmission of positional information between aircraft and/or ground-based air traffic control systems to avoid midair collisions. The most popular cooperative detection systems include the Traffic Collision Avoidance System (TCAS) and Automatic Dependent Surveillance-Broadcast (ADS-B) systems. TCAS is the primary collision avoidance system currently utilized by industry and has been progressively implemented since the mid 1950s [6]. It actively interrogates open airspace for the presence of other aircraft on a 1030 MHz channel via a transponder. The presence of another aircraft within the transmission range will trigger a response from the TCAS system of the aircraft subject to the interrogation. The pilot will then be notified of the aircraft's presence via a 1090 MHz radio frequency. However, aircraft that are not equipped with a TCAS transponder will not recognize another aircraft in its vicinity regardless of whether or not that particular aircraft has TCAS; both aircraft will be effectively blind to one another.

ADS-B is a relatively new technology that allows both pilots and ground stations to detect other equipped aircraft in the surrounding airspace with much more precision than has previously been possible with older systems such as TCAS. Making use of GPS, it determines the aircraft position along with other information such as altitude, speed, heading, flight number, etc. This information is digitized and broadcast several times a second via a discrete frequency data link through a universal access transceiver which allows communication between aircraft within a 240 km radius [7]. Using this information, the pilot is then able to easily make decisions on how best to avoid an approaching aircraft well in advance of it ever becoming a threat. As with TCAS, an aircraft that is not equipped with a traditional ASD-B device will not recognize another aircraft in its vicinity regardless of whether or not that particular aircraft is ASD-B equipped. However, new system developments which involve augmentation of the communication signal with random biphase modulation

allows operation in a manner similar to that of a radar, and may be capable of detecting non-cooperative targets [6].

2.2.2 - Non-cooperative Technologies

As previously mentioned in Chapter 1, UAVs do not meet the regulatory requirements to carry cooperative collision avoidance systems and do not have the benefit of an onboard pilot to act as the last line of defence. These aircraft must therefore utilize some form of non-cooperative system to facilitate the detection and subsequent avoidance of other approaching aircraft. In contrast to cooperative systems, non-cooperative avoidance does not require communication between approaching aircraft. Each aircraft would instead utilize some form of independent sensor system to detect airborne threats and perform an avoidance maneuver if required. Sensor systems currently being investigated for this purpose include electro-optical (EO), infrared (IR), radar, laser, sonar, and acoustic; each of which are discussed in further detail below.

2.2.2.1 - *Electro-Optical*

In the context of collision avoidance systems, EO sensing refers to the use of Charged Coupled Device (CCD) or Complementary Metal Oxide Semiconductor (CMOS) cameras to detect and localize nearby aircraft. These devices operate by converting light intensity or a change in light intensity into an electronic signal. Although very similar to a passive IR sensor, they cannot detect target intensity (energy emitted) [8]. Various studies have been conducted using a number of sensing methods to evaluate the technology's potential for collision avoidance purposes. The underlying principle is to use multiple cameras placed at different locations to create multiple view angles which may effectively determine target vectors through image/pixel differentiation [9]. The most common processing techniques include stereo vision and optical flow methods [10-14]. Stereo based systems are relatively simplistic and computationally efficient. However, the effective detection range is ultimately governed by image resolution and camera spacing which is very limited onboard most all UAVs. Optical flow methods often provide increased detection distances compared to the stereo based approach, but the technique requires target motion across the reference image frame. Thus, a stationary target or aircraft on a head-on collision course would not be detected since it would not appear to be moving [15]. Other technical challenges inherent with all EO methods include high computational requirements for real-time operation, and the need to estimate and compensate for any motion of the sensing aircraft in order to achieve accurate results [8, 9, 16-19].

In addition to technical challenges, there are a number of severe drawbacks inherent with EO-based sensing which greatly limits its potential use. The main being the requirement of good atmospheric visibility during operation. This effectively prevents usage during nighttime or reduced light conditions and limits capabilities when fog or cloud cover is present. EO systems also suffer from a relatively narrow field of view with spherical detection coverage being almost impossible to achieve. Thus, the exclusive use of this technology would not constitute an effective anti-collision system.

2.2.2.2 - *Infrared*

IR imaging is a passive sensing technology that makes use of thermographic cameras. Often called IR cameras, these detect radiation in the infrared region of the electromagnetic spectrum which has a wavelength of anywhere between 0.78–12 µm. IR imaging is generally separated into four categories based on the detected wavelength. Near IR (NIR) detects wavelengths between 0.78–1 µm, Short Wave IR (SWIR) between 1–3 µm, Mid-Wave IR (MWIR) between 3–7 µm, and Long Wave IR (LWIR) between 7–12 µm. Near IR (NIR) is a red wavelength that is just beyond human eye sensitivity. NIR and SWIR behave similarly to visible wavelengths and can be treated the same as CCD technology. MWIR and LWIR detect primarily the thermal emission spectra of an object rather than the reflected emission, and can be used in either day or night conditions [3].

IR images of detected radiation are called thermographs and closely resemble that taken by a standard optical camera. Since infrared radiation is emitted and reflected by all objects above absolute zero according to the black-body radiation law, thermography makes it possible to see ones environment with or without visible illumination. The amount of radiation emitted by an object increases with temperature; therefore, thermography allows one to see variations in the temperature of a body. When viewed through a thermal imaging camera, warm objects such as the engine of a UAV stand out well against cooler backgrounds such as the sky. Thus, IR cameras are most effective during night time operation which is in contrast to that of EO sensing. Depending on the type of IR camera used (SWIR for example), radiation reflected from the UAV fuselage and wings may also be measured to further increase detectability. Currently, little research has been conducted on the use of IR technologies for SAA operations with the exception of one study which utilizes a hybrid EO and IR system [20].

As with EO sensing, IR technologies also suffer from similar drawbacks limiting potential usage. IR cameras have a narrow field of view, are susceptible to visibility conditions such as fog or cloud cover and require high computing capabilities to enable real-time operation. In addition, most commercially available systems offer poor resolution, low frame rate, and a narrow detectable wave band. Cameras best suited for SAA operations are controlled by the International Traffic in Arms Regulations (ITAR) because of their potential use in military applications. Thus, the ability to acquire such devices for research, development, and commercial distribution is greatly hindered.

2.2.2.3 - *Radar*

Radar is an active detection system that uses radio waves to determine the range, altitude, and/or velocity of objects. It works on the principle of transmitting electromagnetic waves that reflect off essentially any object in the transmission path. The reflected waves are collected via a dish or antenna and processed to construct a picture-like representation of the signal-impeding object. Radar technologies such as Continuous-Wave and

Pulse Doppler Radar are currently used extensively in target detection and collision avoidance for ground and sea-based applications.

Unlike EO and IR technologies, radar is not significantly affected by lighting or atmospheric conditions. However, the device requires a great deal of power to achieve reasonable detection distances, and large antennas are required for beam localization; both of which essentially exclude their use on-board most all UAVs. Amphitech has developed a compact radar-based collision avoidance system with a detection range up to 5 nautical miles [21]. However, the device weighs approximately 55 lbs which excludes it from most all non-military UAVs. Synthetic Aperture Radar (SAR) offers a potential solution to this problem by using the sensing aircraft's motion to create a synthetic aperture or window through which electromagnetic waves may be sent and collected. The result is a drastically reduced antenna size and the ability to achieve higher resolutions compared to standard radar forms. Thus, SAR systems may be fitted to small UAVs for sensing purposes such as target detection and low altitude elevation tracking [22]. Research in other SAR applications such as 3D imaging and motion detection is currently being conducted and may eventually allow SAR to be used as a SAA technology [23].

2.2.2.4 - *Laser*

Light amplification by stimulated emission of radiation (LASER) sensing is an active technology that is receiving a great deal of attention. The recent decision by the Danish Air Force to equip its Elicottero Helicopter Industries-01 search-and-rescue helicopters with the SELEX Communications Laser Obstacle Avoidance and Monitoring (LOAM®) system has pushed laser technology to the forefront of SAA solutions [24]. It was also announced that Lockheed-Martin will be working in collaboration with SELEX on a new SAA system for civilian and military applications, including the U.S. Army's Black Hawk Utility Helicopter-60 (UH-60) [23].

Laser systems such as LOAM scan the immediate airspace at regular intervals while processing the data through echo-analysis software. Obstacles in the flight path of the aircraft will be detected if illuminated by the laser source [24]. Promising research has been conducted using high-resolution laser scanners on larger UAVs in cluttered environments [25, 26]. However, these aircraft weighed more than 75 kg and had to use most of their payload capability to lift the laser scanner. Systems have been miniaturized for use on small UAVs by sacrificing both resolution and sensing directions. Compact and lightweight laser scanners that measure target distances via a 2D plane have been successfully utilized by small UAVs for indoor operations [27, 28]. However, the range of these scanners is less than 30 m which would not be sufficient for aircraft anti-collision purposes.

2.2.2.5 - *Sonar*

Sonar is typically an active sensing technology that works on similar principals as Radar, with the major difference being that Sonar utilizes acoustic waves rather than electromagnetic waves. The use of sonar technology is generally ill-suited for UAV SAA purposes due to inherent range limitations and environmental

susceptibility when being used in air [16, 29]. Also, detection times would not be adequate in most cases if aircraft are traveling near sonic levels. Although generally deemed inappropriate for aircraft anti-collision applications, sonar has been successfully utilized as a high-accuracy near-range UAV altimeter for the purpose of autonomous landing operations [30].

2.2.2.6 - *Acoustic*

Acoustic sensing is a passive technology that involves the detection of acoustic wave energy produced by some oscillating body. The most common form involves using sensors known as microphones which detect pressure fluctuations produced during wave transmission. Acoustic sensing has many potential benefits over more traditional non-cooperative technologies such as EO, IR, and Radar. Since sensors are typically omni-directional, complete spherical sensing coverage can be achieved. This is a very important feature as the bulk of midair collision takes place either from behind, the side, above, or below; locations which would typically fall outside the field of view for most other sensing technologies [4]. Sensing systems are typically very small and lightweight since they consist of only a few microphones and a data recording/processing unit. Data acquisition and processing requirements are also much less than that of EO or IR due to decreased sensor data rates. By simultaneously using a number of spatially separated microphones in an array configuration, the detection, localization, and tracking of an acoustic source such as an aircraft can be achieved [31-36]. In some instances, analyzing the Doppler-induced frequency shift of the source signal over a period of time may also allow one to determine the velocity and heading of the sound source [36-40]. Recently, alternative sensing technologies have been developed that measure particle velocities in the transmission medium instead of pressure fluctuations [41]. These are known as Acoustic Vector Sensors (AVS) and have also been successively employed to detect and localize aircraft [42-44]. Although shown to be effective, these sensors are very expensive, fragile, and not well suited for high airflow applications.

The general use of acoustic sensing to detect, localize, and track moving targets has been well studied and documented in the literature [31-36]. However, there have been very few accounts of using this technology for UAV SAA purposes. There have been no reports of air-to-air detection and localization of another aircraft from either a fixed-wing or rotary-wing UAV. There have also been no previous accounts of air-to-ground detection of continuous acoustic sources from fixed-wing UAVs, air-to-ground detection of impulse-based sources from rotary-wing UAVs, or air-to-ground tracking from a moving rotary-wing UAV. Work in this area has either involved the detection of airborne UAVs from stationary platforms such as ground-based microphone arrays or tethered aerial balloons [4, 38, 45-48], simulating a UAV based detection system [4, 49-51], detecting ground-based impulse sources from fixed-wing UAVs, [52, 53], and localizing continuous ground-based sources from stationary low-altitude rotary-wing UAVs [54, 55].

Ferguson [52] utilized a small UAV (Aerosonde) fitted with two microphones to detect and localize acoustic impulses from a propane cannon located on the ground. Detection distances of up to 300 m were said to have been achieved with a localization bearing angle error of 3 degrees; although evidence for these claims was not

clearly presented. Robertson [53] also conducted experiments where a ground-based propane cannon was detected and localized from a small UAV fitted with four microphones. Detection distances of up to 180 m were said to be achieved with an average localization error of 8 degrees. However, evidence for these claims was again not clearly presented. Ohata used a multirotor fitted with a large number of MEMS microphones to detect and localize an acoustic source located on the ground while hovering at altitudes less than 5 m [54, 55]. Although high localization accuracy was achieved, detection distances were far too low to constitute any form of practical anti-collision system. Finally, Harvey showed that a ground-based loudspeaker emitting a 119 dBC audio recording of a small gasoline powdered UAV should be detectable at distances up to 1 km by a fixed-wing UAV fitted with four microphones (original work pertaining to this thesis) [56].

Scientific Applications and Research Associates Inc. (SARA) has developed a compact acoustic sensor system for use on small UAVs known as the Passive Acoustic Non-Cooperative Collision Alert System (PANCAS) [57]. This system provides a means of detecting aircraft on a collision course by observing and tracking the sound of their engines, propellers, and/or rotors. The PANCAS sensor array consists of four microphones mounted in a configuration that allows the bearing and elevation of an acoustic source to be localized. The acoustic probes employ proprietary windscreen and mounting technology to reduce the effects of wind noise and platform vibration. The complete system weighs only 250g and consumes about 7 watts of 6-volt DC power due to its custom dedicated signal processing board and specially designed probes. It has been integrated on a number of small gasoline powered UAVs and supposedly obtained detection ranges of up to 2 km [58]. However, there are currently no accounts published in the scientific literature regarding this technology to support this particular claim. One publication did however briefly indicate that the system was capable of detecting a shockwave emitted from a ground based propane cannon from a distance of approximately 180 m via a fixed-wing UAV [53].

Although acoustic sensing is a relatively new idea in the context of collision avoidance technologies, it is evident that results obtained from studies conducted thus far appear promising. However, significant further work is required to fully evaluate capabilities for various scenarios and system configurations. Work presented in this dissertation will attempt to fill this void by providing a complete overview of the technology and its expected performance characteristics through a combination of theoretical considerations and experimental studies.

2.3 - Acoustic Considerations

2.3.1 - Defining and Quantifying Sound

Acoustic energy or “sound” is defined as a combination of pressure, particle velocity, and particle displacement oscillations in an elastic medium such as air or water. It is typically produced through an oscillating body, or the production of unsteady fluid flow in the medium. The local medium in which the sound waves propagate through is known as the sound field. Sound fields can generally be classified as being either near field, far field,

free field, and reverberant field. The near field is the region close to a source where the sound pressure and acoustic particle velocity are not in phase. This region is limited to a distance from the source equal to about one wavelength or three times the largest dimension of the sound source (whichever is the larger) [59]. The far field is divided into the free field and reverberant field. In the free field the sound behaves as if in open air without reflecting surfaces to interfere with its propagation. The reverberant field is defined as a region which experiences at least one reflection from a boundary surface. In this region, reflected sound waves interfere with one another in both destructive and constructive ways (depending on their phase). For the detection scenario pertaining to this thesis, it is assumed that the acoustic source is located in the far free field such that wave fronts arriving at the observer may be approximated as planar, while multipath effects occurring from surface reflections may typically be ignored.

Acoustic fields are often described by Sound Pressure Level (SPL), Sound Power Level (PWL), and Sound Intensity Level (SIL). Of these, SPL is the most commonly used quantity. It is given by the ratio of Root Mean Square (RMS) wave pressure to some reference value given in decibels:

$$SPL = 10\log_{10}\left(\frac{p_{rms}^2}{p_{ref}^2}\right) = 20\log_{10}\left(\frac{p_{rms}}{p_{ref}}\right) \tag{2.1}$$

where $p_{ref} = 20$ μPa is the standard reference pressure which represents the lowest level detectable by the human ear. The decibel conversion is used since the human ear does not perceive sound level changes in a linear manner but rather in a more logarithmic one. For spherical waves, pressure is inversely proportional to propagation distance from the source [59]. Thus, two arbitrary points in the radial direction can be related via the SPL according to:

$$SPL_2 = SPL_1 - 20\log_{10}\left(\frac{r_2}{r_1}\right) \tag{2.2}$$

where SPL_1 and SPL_2 are the sound pressure levels at some reference distances r_1 and r_2 away from the source. From the above equation it can be shown that in the free field, sound pressure level decreases by 6 dB for each doubling of the distance away from the source.

If the source is directional, an additional term can be used to account for the uneven distribution of the sound level as a function of direction. The Directivity Index (DI) is defined the difference between the actual sound pressure, and the sound pressure from a omni-directional point source with the same total acoustic power [60]. It is thus given by:

$$\mathrm{DI} = 10\log_{10}\left(\frac{P_{mes}^2}{P_{ref}^2}\right) \tag{2.3}$$

For an omni-directional source radiating into free 3D space DI = 0 dB. If the source is placed on a perfectly reflecting surface, hemispherical radiation will occur effectively doubling the field energy density giving DI = 3 dB.

Sound propagates in the form of longitudinal compression waves which may be described using the generalized wave equation:

$$\nabla^2 p(\vec{r},t) - \frac{1}{c^2}\frac{\partial^2 p(\vec{r},t)}{\partial t^2} = 0 \tag{2.4}$$

where c is the speed of sound in the medium, and $p(\vec{r},t)$ is a function representing the acoustic pressure at some point in time t and space $\vec{r}$ where $\vec{r} = [x,y,z]^T$. Using a separation of variables approach, the solution for a plane wave is given by [61]:

$$p(\vec{r},t) = Ae^{j(\omega t - \mathbf{k}\cdot\vec{r})} \tag{2.5}$$

where A is the wave amplitude, $\omega = 2\pi f$ is the frequency in radians per second, and $\mathbf{k}$ is the wavenumber vector which indicates the speed and direction of wave propagation:

$$\mathbf{k} = \frac{2\pi}{\lambda}[\sin\vartheta\cos\varphi, \sin\vartheta\sin\varphi, \cos\vartheta] \tag{2.6}$$

where $\lambda = c/f$ is the wavelength, ϑ and φ are the 3D azimuth and elevation angles respectively, and f is the frequency in Hz.

For the case of a point source emitting in 3D space, acoustic waves may be described via the spherical wave equation according to:

$$\frac{\partial^2 (r\cdot p)}{\partial r^2} - \frac{1}{c^2}\frac{\partial^2 (r\cdot p)}{\partial t^2} = 0 \tag{2.7}$$

which has a solution of the form:

$$p(r,t) = \frac{A}{4\pi r}e^{j(\omega t - kr)} \tag{2.8}$$

where $r = |\vec{r}|$ and $k = |\mathbf{k}|$. From the above result it is evident that wave pressure is now a function of distance away from the emitting source. Although sound waves are typically spherical in nature, they can be approximated as plane waves for large distances away from the source. This approximation is often used to simplify mathematical analysis for operations such as acoustic beamforming.

For an ideal fixed-point source, the resulting acoustic field may be described by the inhomogeneous wave equation:

$$\nabla^2 p(\vec{r},t) - \frac{1}{c^2}\frac{\partial^2 p(\vec{r},t)}{\partial t^2} = q(\vec{r},t) \tag{2.9}$$

where $q(\vec{r},t)$ is the source function. Through application of Greens theorem, it can be shown that the solution to the above form is given by [62]:

$$p(\vec{r},t)=\frac{q(t-\left|\vec{r}-\vec{r}_s\right|/c)}{4\pi\left|\vec{r}-\vec{r}_s\right|} \tag{2.10}$$

where $\vec{r}_s$ is location of the acoustic source.

Now consider the case for a moving monopole source. The sound emitted by the source at some time τ will arrive at the observer at some later time given by:

$$t=t_p+\tau \tag{2.11}$$

where t_p is the propagation delay time. The delay time will be a function of the distance between the source and observer at the time of emission:

$$t_p=\frac{\left|\vec{r}-\vec{r}_s(\tau)\right|}{c} \tag{2.12}$$

Substitution of (2.11) and (2.12) into (2.9) and applying Green's theorem again leads to the following solution [62]:

$$p(\vec{r},t)=\sum_i\frac{q(\tau_i)}{4\pi\left|\vec{r}-\vec{r}_s(\tau_i)\right|\left|1-M_{so}(\tau_i)/c\right|} \tag{2.13}$$

where M_{so} is the component of the source velocity (in Mach) along the direct transmission path from the source to the observer:

$$M_{so}(\tau_i)=\frac{1}{c}\frac{\vec{r}-\vec{r}_s(\tau_i)}{\left|\vec{r}-\vec{r}_s(\tau_i)\right|}\bullet\vec{v}_s \tag{2.14}$$

From the above analysis, one very important feature should be noted with regards to the localization of a moving acoustic source. Since a time delay exists between the transmission and receipt of acoustic information, the position of the source with respect to the observer at some time t, will actually be that at some point τ back in time. That is, the current position estimate will be delayed by t_p seconds.

For the case of a moving source and moving observer, the propagation delay will become:

$$t_p=\frac{\left|\vec{r}_o(t)-\vec{r}_s(\tau)\right|}{c} \tag{2.15}$$

while M_{so} is now obtained via the relative velocity between the source and observer along the direct transmission path:

$$M_{so}(\tau_i)=\frac{1}{c}\frac{\vec{r}_s(\tau_i)-\vec{r}_o(t)}{\left|\vec{r}_s(\tau_i)-\vec{r}_o(t)\right|}\bullet\vec{v}_s^{\,o} \tag{2.16}$$

where $\vec{v}_s^{\,o}$ is the velocity of the source relative to the observer and τ_i is obtained via the roots of:

$$t = \frac{\left|\vec{r}_o(t) - \vec{r}_s(\tau_i)\right|}{c} + \tau_i \tag{2.17}$$

It should be noted that the above solutions are only valid for the case of a source and/or observer moving with a constant velocity.

2.3.2 - Atmospheric Propagation

In most instances sound propagation rarely adheres to perfect free-field conditions as various environmental factors are often at play. For the case of atmospheric transmission, sound propagation may be affected by wind, temperature gradients, density variations, humidity, and the presence of any absorbing or reflecting surfaces [63]. We may define a complete atmospheric attenuation factor by summing the effects of each individual factor as follows:

$$AA = A_{abs} + A_{wind} + A_{temp} + A_{surf} \tag{2.18}$$

where A_{abs}, A_{wind}, A_{temp}, and A_{surf} are the attenuation levels due to atmospheric absorption, wind effects, temperature effects, and surface interaction effects.

Atmospheric absorption is caused by viscous frictional losses and relaxational effects associated with wave induced particle motion. Empirical models have been developed to predict attenuation levels based on source frequency and thermodynamic properties of the medium [64]. For a given propagation distance r, the total attenuation caused by atmospheric absorption is given by:

$$A_{abs} = \alpha r \tag{2.19}$$

where α is the absorption coefficient with units of dB/100 m obtained via empirical equations or data plots such as that displayed below in Figure 2-2. Equation (2.2) displayed above for sound propagation can now be written in the following form to account for atmospheric losses and source directionality:

$$SPL_2 = SPL_1 - 20\log_{10}\left(\frac{r_2}{r_1}\right) - A_{abs} + DI \tag{2.20}$$

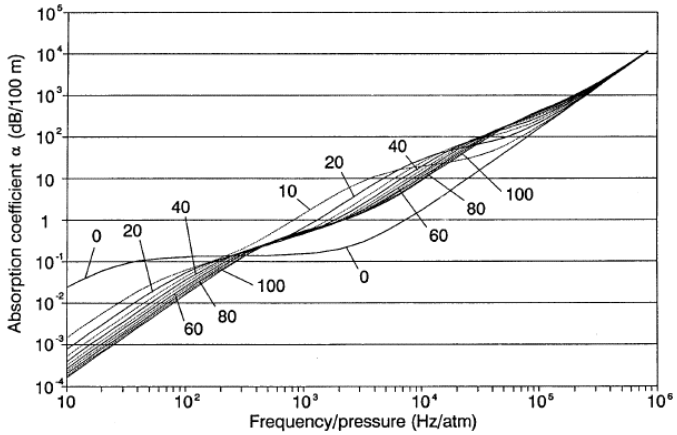

Figure 2-2: Absorption coefficient as a function of frequency and humidity at 20 °C [65].

The presence of reflecting bodies or surfaces may attenuate or amplify the acoustic energy through a combination of surface/body absorption and reflection/multipath effects. Consider a source and separated receiver both located at some distance above a perfectly reflecting surface as depicted in Figure 2-3. Sound will arrive at the receiver via the direct path, and some reflected path with angle of incidence given by ϕ. Sound arriving at the receiver will constructively or destructively interfere according to the phase of each path signal. Ultimately this will be dictated by wave frequency and path length differences. The sound pressure at the receiver due to the direct path p_d and reflected path p_r wave interaction is given by the following equation [60]:

$$p_{tot}^2 = p_d^2 + p_r^2 + 2\,p_d\,p_r\cos[\Delta\varphi] \tag{2.21}$$

where $\Delta\varphi$ is the phase difference between the two waves at the receiver. For the case of a non-ideal reflecting surface such as the earth, the reflected wave will be further attenuated through surface absorption effects. Empirical equations have been developed to predict attenuation levels based on surface, wave, and geometric properties [65].

In the context of acoustic sensing, properties of multipath reflection may be exploited to determine information about the source location [66, 67]. With respect to the receiver, sound will appear to be arriving from both the true source location and the ground incidence location. If beamforming techniques can be applied to determine the elevation angles to these points relative to the receiver, the source altitude and direct path distance can be determined. Consider the geometric depiction of the propagation scenario displayed in Figure 2-3, where ϕ is the angle of incidence and θ is the angular source location with respect to horizontal plane. It is assumed that ϕ and θ are obtained via some localization operation while h_r is also known. The direct path length L_D can be obtained via the law of sines according to:

$$L_D = L_{r_2}\frac{\sin(\pi - 2\phi)}{\sin(\phi - \theta)} \tag{2.22}$$

where L_{r_2} is given by:

$$L_{r_2} = \frac{h_r}{\sin(\phi)} \tag{2.23}$$

The altitude of the source can then be found according to:

$$h_s = h_r + L_D\sin(\theta + \phi) \tag{2.24}$$

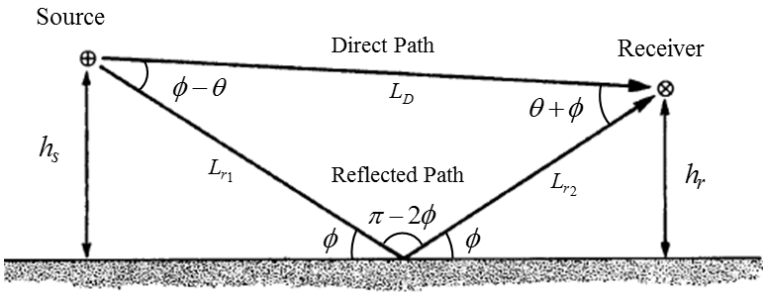

Figure 2-3: Multipath sound propagation.

Wind and temperature gradients may also affect sound transmission to a significant extent, typically through the production of refraction like effects. Over open ground, substantial velocity gradients may exist due to friction between the moving air and ground. In the absence of turbulence, boundary layer effects typically cause air speeds to vary logarithmically up to a height of approximately 100 meters [65]. As a consequence of this, sound traveling against the wind direction will be refracted or bent upwards, while sound moving away will be bent downwards. This effect is depicted below in Figure 2-4. Similar to wind, temperature will also have a refractive effect on sound propagation. In the presence of a temperature gradient (typically in the vertical direction), sound waves are refracted to direction of lower sound velocity (lower temperature region).

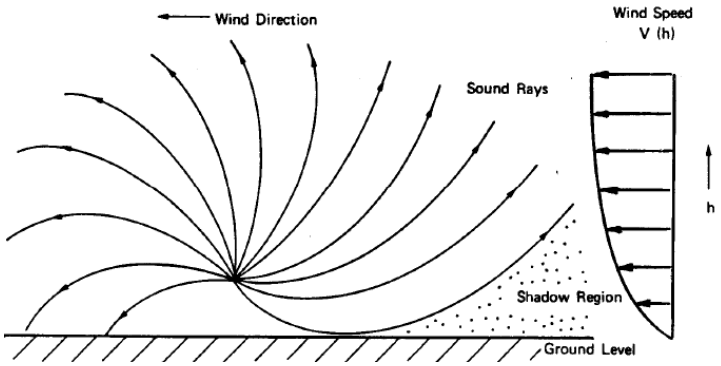

Figure 2-4: Refraction effects due to wind [68].

Determining attenuation levels due to refraction effects is extremely difficult and often impossible for many weather conditions [69]. Methods such as ray tracing may be utilized to form approximations, but accurate results require knowledge of the medium thermodynamic properties as a function of altitude. Thus, no attempts will be made to incorporate these effects into any aspects of the proposed acoustic detection system presented in this thesis.

2.3.3 - Doppler Effects

The acoustic Doppler effect is the apparent change in observed frequency of an acoustic source due to a relative motion between the source and observer. For advancing bodies (decreasing distances), perceived frequencies increase, while retreating bodies (increasing distances) produce a decreased frequency perception. The general equation relating the observed Doppler shifted frequency f_o and the true source frequency f_s is given by the following [70]:

$$f_o = f_s \frac{(1+\hat{r}_s^o \bullet \vec{v}_o / c)}{(1+\hat{r}_s^o \bullet \vec{v}_s / c)} \tag{2.25}$$

where c is the speed of sound in air, $\vec{v}_o$ is the velocity of the observer, $\vec{v}_s$ is the velocity of the source and $\hat{r}_s^o$ is the unit vector in the direction of the source with respect to the observer. If velocities are much less than the speed of sound, the above form can be approximated as [71]:

$$f_o = (1-\hat{r}_s^o \bullet \vec{v}_s^o / c)\, f_s \tag{2.26}$$

where $\vec{v}_s^o$ is the velocity of the source relative to the observer. If velocities are greater than or equal to the speed of sound, the above Doppler model is no longer valid, since the emitted wavelength will approach zero giving rise to a shockwave. However, very few aircraft operate at these velocities and thus the issue is of little concern here. For the case of a medium with uniform flow velocity $\vec{v}_m$, the observed frequency is given by [70]:

$$f_0 = f_s \frac{(1 + \hat{r}_s^o \bullet [\vec{v}_o - \vec{v}_m]/c)}{(1 + \hat{r}_s^o \bullet [\vec{v}_s - \vec{v}_m]/c)} \tag{2.27}$$

Since both aircraft are in constant motion, some degree of doppler shifting will exist at almost all times. From the above equation however, it is evident that the observed frequency will equal that of the emitted signal when the relative velocity is zero. At this point, the distance between the source and observer will be at its minimum and is termed the Closest Point of Approach (CPA). It has been found that the Doppler effect may be exploited to predict kinematic source parameters such as velocity, heading, and altitude for an aircraft flying past a ground-based array [36-40]. Using the CPA as a boundary condition, a closed form solution to the system kinematic equations can be achieved. Parameter values may then be estimated through data fitting of observed frequency values via methods such the least squares approach. Unfortunately however, these methods can only be implemented in a post processing form since the full data record of the observed flight is required.

2.3.4 - Aircraft Noise Generation

Acoustic generated noise is the greatest limiting factor in achieving detection distances required to establish a viable SAA system. Due to the physical nature of the system, large amounts of self-noise is generated by the sensing aircraft. This in turn is captured by the sensing system effectively corrupting acquired acoustic signals. The aircraft propulsion system constitutes the major noise component with unsteady airflow past the sensing elements and airframe contributing to a lesser extent. Figure 2-5 provides a depiction of these noise sources for a fixed-wing pusher style aircraft fitted with four microphone sensors. However, depending on the aircraft configuration (fixed-wing or rotary-wing), these components may vary largely in terms of their properties and contribution levels. Details regarding the exact mechanisms by which acoustic energy is produced along with its specific properties is complex and constitutes its own field of study. Thus, topics in this area will only be discussed to a degree which maintains a practical significance to work presented in this thesis.

During flight operations, large pressure fluctuations and unsteady fluid flows are created by the propeller motion. Although most of the turbulent airflow is swept downstream, pressure waves radiate outwards in all directions and would be sensed by any microphone located in the vicinity. Propeller noise can generally be decomposed into rotational and vortex-based components [72]. Vortex based noise is broadband in nature and is produced by the unsteady fluid interaction with the propeller blade surfaces and trailing edges. This noise is directional along the propeller axis and typically contributes to a much lesser extent [73]. The rotational noise component consists of discrete harmonic frequencies which are a function of the blade rotation rate. It is primarily produced through the pressure distributions required to generate lift, periodic air volume

displacement, and a combination of impulse based effects known as blade slap [72]. The fundamental component provides the largest power contribution with successive harmonic frequencies decreasing almost linearly with increasing order. The fundamental frequency value is determined by the product of the propeller rotation rate and number of blades present.

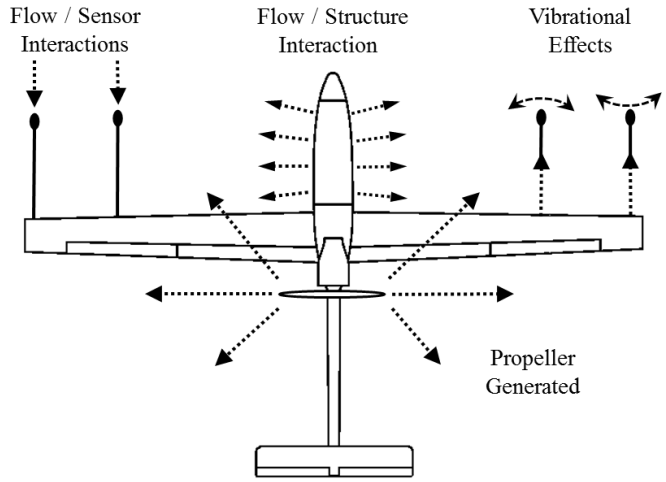

Figure 2-5: Depiction of basic fixed-wing aircraft noise sources.

For most fixed-wing aircraft, the above description provides an accurate account of noise generation mechanisms. For the case of rotary-wing aircraft however, multiple rotational sources are present operating at different frequencies thus giving rise to a more complex acoustic signature. For the minimum configuration (helicopter), two propeller-based sources are present which operate at different frequencies. Most helicopters have a main rotor frequency between 5 and 10 Hz, and a tail rotor frequency between 15 and 50 Hz with anywhere from 2 to 4 blades present on each rotor [74]. This contrasts with most fixed-wing aircraft which typically operate in the 50 to 100 Hz range. Figure 2-6 provides the power spectra generated from experimental data for a fixed-wing Cessna 185 airplane, Bell 206 helicopter, and a Sikorsky S-92 helicopter during a fly-by. From the plots it is evident that strong narrowband components are present combined with a frequency dependent broadband component. It is also apparent that the number of harmonic components is much greater for the two helicopter spectra and extend to a much lower frequency as expected. In addition, most all the narrowband components are attenuated to the broadband noise floor at frequencies above 1000 Hz for all of the aircraft. It will later be shown in Chapter 4 that the presence of theses harmonic components may be exploited to enhance aircraft detection capabilities.

For the case of UAVs, more advanced rotary-wing aircraft are often used which employ multiple equal lifting propellers without the use of any stabilizing rotor. Typically referred to as multi-rotor aircraft, these devices usually contain anywhere from 4 to 8 separate lifting fans with two propeller blades on each rotor. Figure 2-7 provides the power spectra of self-generated noise obtained from experimental acoustic data for three UAVs. The aircraft consisted of a fixed-wing pusher (Delta X-8), fixed-wing conventional (Giant Big Stik), and an eight engine multi-rotor (Kraken). Further details regarding these aircraft is later provided in Chapter 6. From

the plots it is evident that all aircraft exhibit strong narrowband components with a frequency dependent broadband component as expected. However, the conventional style and multi-rotor aircraft both exhibit stronger broadband components than the fixed-wing pusher configuration. This is expected since the sensing microphones were located further from the propeller axis for this particular aircraft.

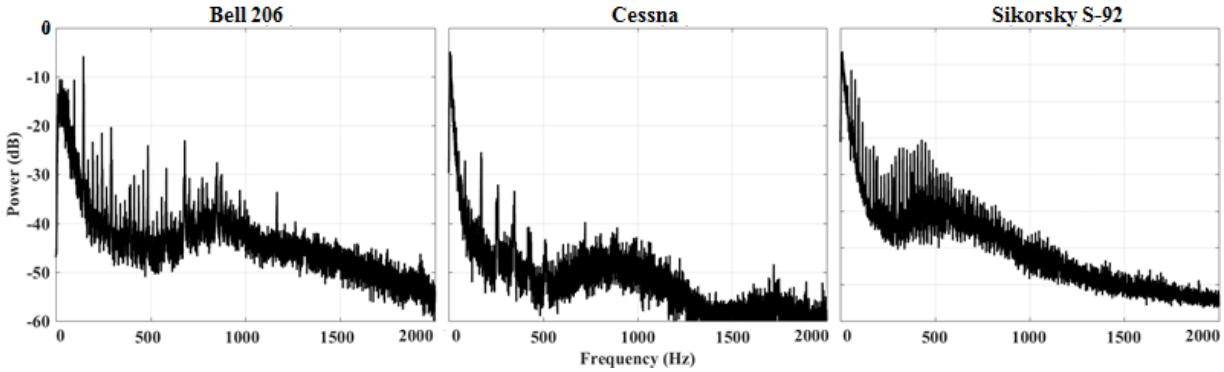

Figure 2-6: Power spectra of various aircraft during fly-by.

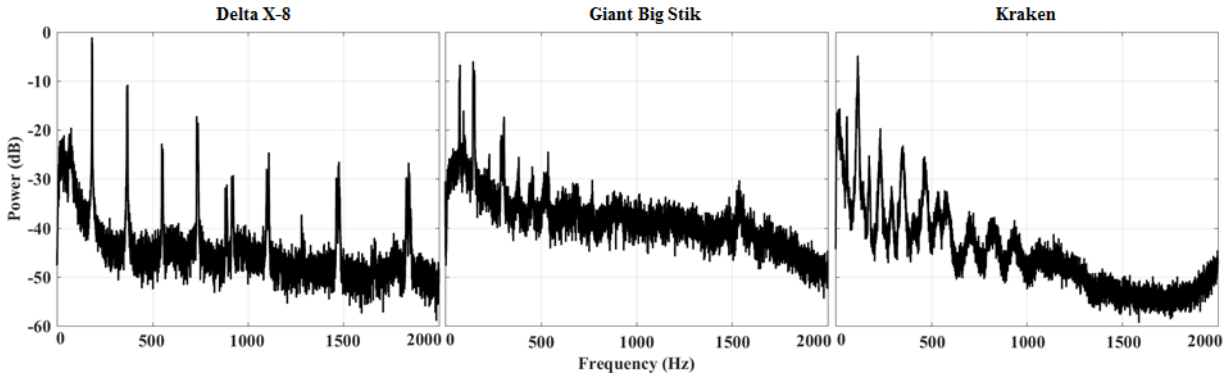

Figure 2-7: Self-noise from various UAV aircraft.

In addition to propeller generated noise, vibrations produced by the aircraft engine may also produce narrowband noise components. Similar to the propeller-based effects, these components also occur at discrete harmonic frequencies with the fundamental frequency being equal to the engine rotational rate. However, unlike propeller generated noise, vibrational effects can be effectively mitigated through incorporating vibration dampening materials in the sensor mounting system. Such a system is later described in the experimental details section (Chapter 6).

Components of the aircraft structure may also contribute to acoustic noise levels through the creation of unsteady flow phenomena such as vortex shedding and wake formation [73]. These effects are typically generated along surface edges such as wing tips, or by non-aerodynamic components such as the landing gear. Structural based noise is typically broadband in nature and extremely hard to predict since it is completely dependent on the aircraft structural configuration and component geometry.

Another obvious and often significant noise source is that generated by the microphone sensors. Placing microphones on a fixed-wing aircraft would typically mean they are operating in a high velocity fluid-flow field. The inherent shape of a microphone (cylindrical with a flat membrane top) creates unsteady flow

conditions across the sensing diaphragm effectively generating pressure fluctuations, which in turn produces noise on the recorded sensor signal(s). Depending on the physical characteristics of the microphone (dimensions and shape) and the fluid flow velocities involved, the noise generated may be quite severe. It will later be shown that utilizing specialized microphones fitted with noise cones to promote the production of laminar fluid flow across the sensing diaphragm can reduce this noise considerably. For the case of multi-rotor aircraft, flow generated noise is often of less concern since transit velocities are typically much lower. Because the aircraft is not constrained to any particular motion direction (with respect to its orientation), the laminar flow style microphones used for fixed-wing aircraft would provide little benefit. However, since fluid velocities are relatively low, noise levels can be reduced considerably by the use of standard foam windscreen covers.

2.4 - Kinematic Assessment

As previously demonstrated, the acoustic field at the observing aircraft will be a function of the relative distance and velocity along the direct transmission path (line of sight) between the two aircraft. Thus, we wish to obtain a kinematic model which describes the relative motion based on these parameters. Consider the vector representation for the potential collision system as depicted in Figure 2-8. The position of each aircraft is given in a universal Cartesian coordinate frame such as the GPS system. Translated coordinate systems are attached to the source and observing aircraft indicated by **S** and **O** respectively. The speed and heading for the two aircraft are defined in these frames via the spherical coordinate values (speed, azimuth, elevation) v_s, θ_s, ϕ_s and v_o, θ_o, ϕ_o respectively. In addition, a rotated coordinate system is attached to the observing aircraft given by **A** which defines its orientation via the roll, pitch, yaw angles χ, ψ, γ. The angular position of the source aircraft is given in this reference frame by the azimuth and elevation angles ϑ, φ. Note that for fixed-wing aircraft, the direction of motion is dictated by the aircraft orientation. That is, the aircraft always moves in the direction it is pointed such that $\theta = \gamma$ and $\phi = \psi$. However, for multirotor aircraft this is not the case. Its motion is completely independent of yaw and only partially dependent on the roll and pitch angles. Motion in the vertical direction is completely independent of orientation while motion in the horizontal plane is dependent on pitch and roll angles. The coordinate systems for the aircraft orientation frame and spherical heading values are defined in Figure 2-9. Note that the defined orientation angle system is different from the traditional convention. This was done to maintain uniformity between the standard universal, array directivity, and aircraft heading systems.

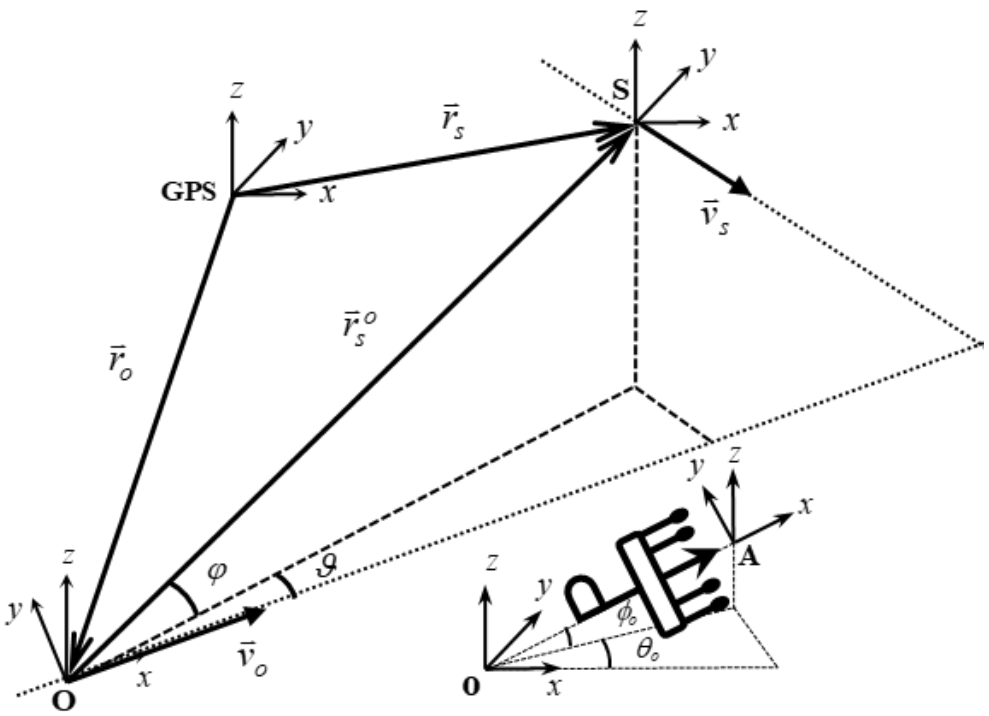

Figure 2-8: Vector representation of detection scenario kinematics

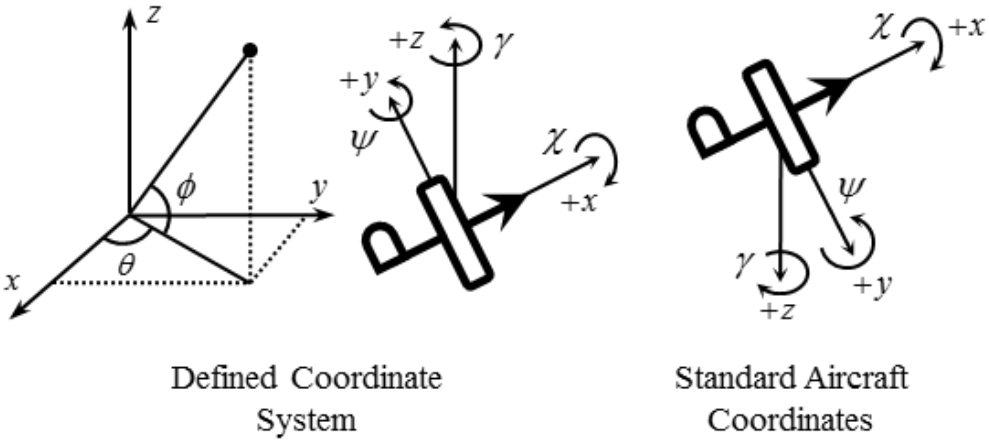

Figure 2-9: Defined coordinate systems

From the kinematic vector representation given above, it is evident that the relative distance and velocity along the direction transmission path are given by:

$$r_{rel} = \left|\vec{r}_s^{\,o}\right| = \vec{r}_s^{\,o} \bullet \hat{r}_s^{\,o} \tag{2.28}$$

$$v_{rel} = \vec{v}_s^{\,o} \bullet \hat{r}_s^{\,o} \tag{2.29}$$

where $\vec{r}_s^{\,o}$ and $\vec{v}_s^{\,o}$ are the displacement and relative velocity vectors respectively for the source position relative to the observer in the translated coordinate system **O** given by:

$$\vec{r}_s^{\,o} = \vec{r}_s - \vec{r}_o \tag{2.30}$$

$$\vec{v}_s^{\,o} = \frac{d}{dt}\vec{r}_s^{\,o} \tag{2.31}$$

The unit displacement or direction vector $\hat{r}_s^o$ may be approximated via the angular position ϑ, φ as obtained through beamforming methods. Conversion of theses angles from the sensing aircraft orientation frame to the universal coordinate system may be achieved according to:

$$\hat{r}_s^o = {}_{o'}^{o}R\,\hat{r}_s^{o'} = {}_{o'}^{o}RM^{\alpha,\beta} \tag{2.32}$$

where ${}_{o'}^{o}R$ is the rotation transform matrix given by:

$${}_{o'}^{o}R = \begin{bmatrix} \cos\chi\cos\psi & -\cos\gamma\sin\chi - \cos\chi\sin\gamma\sin\psi & \sin\gamma\sin\chi - \cos\gamma\cos\chi\sin\psi \\ \cos\psi\sin\chi & \cos\gamma\cos\chi - \sin\gamma\sin\chi\sin\psi & -\cos\chi\sin\gamma - \cos\gamma\sin\psi\sin\psi \\ \sin\psi & \cos\psi\sin\gamma & \cos\gamma\cos\psi \end{bmatrix} \tag{2.33}$$

and $M^{\vartheta,\varphi}$ angular source position vector given by

$$M^{\vartheta,\varphi} = \left[\cos\vartheta\cos\varphi, \sin\vartheta\cos\varphi, \sin\varphi\right]^T \tag{2.34}$$

Thus, the position vector $\vec{r}_s^o$ may also be written as:

$$\begin{aligned} \vec{r}_s^o &= \left|\vec{r}_s^o\right| \cdot \hat{r}_s^o = r_{rel} \cdot \hat{r}_s^o \\ &= r_{rel}\,{}_{o'}^{o}RM^{\vartheta,\varphi} \end{aligned} \tag{2.35}$$

For the case of constant velocities, the position vector for each aircraft will be given by:

$$\vec{r}_s = \vec{r}_{s,0} + \vec{v}_s t \tag{2.36}$$

$$\vec{r}_o = \vec{r}_{o,0} + \vec{v}_o t \tag{2.37}$$

where $\vec{r}_{s,0}$ and $\vec{r}_{o,0}$ are the initial positions at the first point of detection. The source and observer velocity vectors are defined by the speed and heading values according to:

$$\vec{v}_o = v_o M^{\theta_o,\phi_o} \tag{2.38}$$

$$\vec{v}_s = v_s M^{\theta_s,\phi_s} \tag{2.39}$$

where $M^{\theta,\phi}$ is the heading vector for the respective aircraft given by:

$$M^{\theta,\phi} = \left[\cos\theta\cos\phi, \sin\theta\cos\phi, \sin\phi\right]^T \tag{2.40}$$

Substitution of equations (2.36) and (2.37) into (2.30) gives:

$$\begin{aligned} \vec{r}_s^o &= \vec{r}_{s,0} - \vec{r}_{o,0} + t(\vec{v}_s - \vec{v}_o) \\ &= \vec{r}_{s,0}^o + \vec{v}_s^o t \end{aligned} \tag{2.41}$$

Rearranging and combining equations (2.41), (2.35), and (2.29) finally gives:

$$\begin{aligned} v_{rel} &= \frac{1}{t}\left(\vec{r}_s^o - \vec{r}_{s,0}^o\right)\bullet\hat{r}_s^o = \frac{1}{t}\left(\left|\vec{r}_s^o\right| \cdot \hat{r}_s^o - \left|\vec{r}_{s,0}^o\right| \cdot \hat{r}_{s,0}^o\right)\bullet\hat{r}_s^o \\ &= \frac{{}_{o'}^{o}R}{t}\left[{}_{o'}^{o}RM^{\vartheta,\varphi}\right]\bullet\left[M^{\vartheta,\varphi}r_{rel} - M^{\vartheta_0,\varphi_0}r_{rel,0}\right] \end{aligned} \tag{2.42}$$

For the case of non-constant aircraft velocities, the above equation does not hold. Instead the position vectors will be given by:

$$\vec{r}_s = \vec{r}_{s,0} + \int_0^t \vec{v}_s d\tau \tag{2.43}$$

$$\vec{r}_o = \vec{r}_{o,0} + \int_0^t \vec{v}_o d\tau \tag{2.44}$$

for which the relative velocity will be:

$$\begin{aligned} v_{rel} &= \frac{d}{dt} r_{rel} = \frac{d}{dt} \left| \vec{r}_s^{\,o} \right| = \frac{d}{dt} \vec{r}_s^{\,o} \bullet \hat{r}_s^{o} \\ &= \frac{d}{dt} \left[\int_0^t \vec{v}_s d\tau - \int_0^t \vec{v}_o d\tau \right] \bullet {}_o^o R M^{\alpha,\beta} \end{aligned} \tag{2.45}$$

2.5 - Signal Processing Requirements

Acoustic signals are acquired via multiple microphones located at various positions on the sensing aircraft to facilitate array processing operations. Each sensor contains a pre-amplification circuit and is supplied with a polarization voltage to enable operation. Analog signals are passed through an Analog-to-Digital Converter (ADC) operating at some fixed sampling rate before finally entering the digital processing unit as depicted below in Figure 2-10.

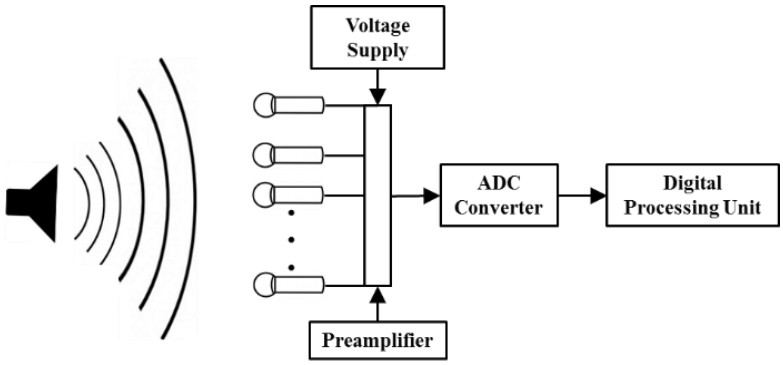

Figure 2-10: Acoustic data acquisition system.

Once signals have been digitally acquired, various processing steps are then implemented to enable source detection and spatial localization. The required processing steps may be sectioned into three main stages with each stage containing multiple sub operations as depicted in Figure 2-11. The major processing operations consist of: 1) Conditioning & Filtering, 2) Enhancement & Detection, and 3) Localization & Tracking. Each of these areas is the subject of its own chapter and is thus discussed later in much greater technical detail. The description given below simply provides a brief overview of the general signal processing operations required to establish an acoustic based collision avoidance system.

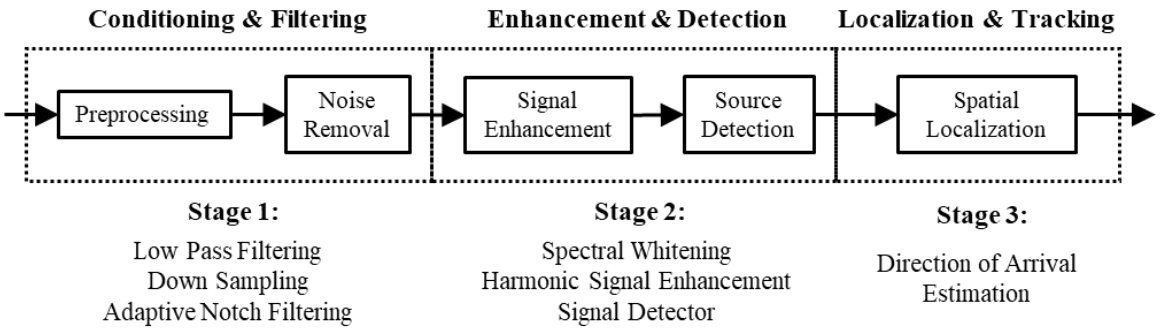

Figure 2-11: Overview of required signal processing steps.

The first processing stage conditions signals such that the performance of detection and localization algorithms are maximized. It consists of two main components: 1) Preprocessing and 2) Narrowband Self-Noise Removal. Preprocessing consists of basic operations to remove any signal components outside the bandwidth of interest, down-sample signals if possible to minimize computational requirements, and adjust gain values to maintain uniformity across signals. Filtering operations are performed on the conditioned signals to remove harmonic narrowband noise generated by the aircraft propulsion system. Since engine speeds often vary considerably during flight operations, adaptive filtering techniques must be employed. All processing operations performed in this stage are completed in the time domain.

Upon conditioning and removing all major noise components, signals are then subject to enhancement processors to further discriminate between random noise and potential source components. Since no information regarding the source properties (frequency, phase, amplitude, etc.) are known, signals must first be transformed to the frequency domain to optimize detectability. By exploiting the harmonic narrowband structure associated with aircraft acoustic emissions, operations may be performed to enhance the presence of potential narrowband components relative to the surrounding broadband noise. Threshold based detection algorithms are then implemented to verify the presence of any source signals with some predetermined probability. To maintain a constant false alarm rate, essentially all detection schemes require independent and identically distributed (IID) noise across the bandwidth of interest. Thus, spectral whitening techniques must be implemented prior to any detection operations to ensure probability requirements are accurately satisfied.

If a target signal is detected, array processing techniques may then be employed to determine the angular location of the source relative to the detecting aircraft orientation. Using this information, the detecting aircraft can then change course accordingly to avoid a collision. Although a successful avoidance maneuver can be performed using only angular positions, the resulting flight path will be unnecessarily long and thus require larger minimum detection distances. In order to determine the optimal course change for a given kinematic configuration, information regarding the target trajectory must also be known. This may be achieved by inputting the observed source frequency, amplitude, and spatial location values into a model which incorporates the kinematic, acoustic, and signal properties of the collision system as previously described. Details regarding the optimal path change for a given target trajectory is outside the scope of this thesis.

2.6 - Signal Processing Fundamentals

The following section provides a brief description of various fundamental signal processing concepts and operations utilized throughout the remainder of this thesis.

2.6.1 - Signal Statistics

Signals may be broadly classified as being either deterministic or random. Deterministic signals are fixed and can be completely described by analytical expressions. All past, present, and future values of these signals can be determined or predicted with certainty. For example, a signal containing only pure sinusoidal wave(s) would be deterministic since it can be described entirely through some combination of sinusoidal functions. Random or stochastic signals however cannot be characterized by a simple well-defined mathematical function, and their future values cannot be predicted with absolute certainty. For these signals we must utilize probabilistic techniques to describe their behavior. Most real-world signals are random in nature.

Random signals may be further classified as being either stationary or non-stationary. A stationary random process is one whose ensemble statistics do not depend on time; its probability distribution is the same at all times. In contrast, a non-stationary system is one whose characterizing statistics such as mean, variance, etc. changes over time. An ergodic process is one in which the time average and ensemble statistics are equal. Often, stationary processes are ergodic in nature.

Random stationary signals are often classified in terms of their expected value, variance, standard deviation, and Probability Density Function (PDF). The PDF of a stationary process is generally defined as:

$$f_X(x) = \frac{dF(x)}{dx} \tag{2.46}$$

where $F(x)$ is the Cumulative Distribution Function (CDF) of a process involving some random variable X. The probability that the random variable X will have a value between a and b is represented as $P[a \leq X \leq b]$ and is given by:

$$P[a \leq X \leq b] = F(b) - F(a) = \int_a^b f_X(x)\,dx \tag{2.47}$$

For any continuous single-valued function $g(x)$ with corresponding PDF $f_X(x)$, the expected value of $g(x)$ is defined by:

$$E[g(x)] = \int_{-\infty}^{\infty} g(x)\, f_X(x)\, dx \tag{2.48}$$

Commonly used statistical measures such as the mean, variance, standard deviation, etc. are defined by taking various moments of the expected value according to:

$$\rho_k = E[x^k] = \int_{-\infty}^{\infty} x^k f_X(x)\, dx \tag{2.49}$$

where the mean of a signal μ is defined by the fist moment:

$$\rho_1 \rightarrow E[x^1] = \mu = \int_{-\infty}^{\infty} x^1 f_X(x)\, dx \tag{2.50}$$

The mean square value ψ^2 which is analgous to signal power is given by the second moment:

$$\rho_2 \rightarrow E[x^2] = \psi^2 = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx \tag{2.51}$$

where the positive square root of ψ^2 is known as the Root Mean Square (RMS) value ψ.

The variance of a process σ^2 is defined as the second moment taken about the mean according to:

$$\rho_2^{\mu} \rightarrow E[(x-\mu)^2] = \sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)\, dx \tag{2.52}$$

where the positive square root of the variance is termed the standard deviation σ. It can be shown that the above quantities also satisfy the following condition [75]:

$$\psi^2 = \sigma^2 + \mu^2 \tag{2.53}$$

For stationary (ergodic) discrete time signals, the above statistical quantities can be obtained according to:

$$\mu_x = \lim_{N\rightarrow\infty} \frac{1}{N} \sum_{n=0}^{N} [x(n)] \tag{2.54}$$

$$\psi_x^2 = \lim_{N\rightarrow\infty} \frac{1}{N} \sum_{n=0}^{N} [x(n)]^2 \tag{2.55}$$

$$\sigma_x^2 = \lim_{N\rightarrow\infty} \frac{1}{N} \sum_{n=0}^{N} [x(n) - \mu_x]^2 \tag{2.56}$$

In reality, only a finite number of samples are available to determine the above statistical quantities. If the sample space is small (typically < 100 samples) values obtained are often said to be biased with respect to the true population statistics. That is, the presence of an outlier for example, would highly influence the calculated parameter from the true value for the population. For the case of variance, applying Bessel's Correction leads to more accurate result [76]:

$$\sigma_x^2 = \frac{1}{N-1} \sum_{n=0}^{N} [x(n) - \mu_x]^2 \quad for \;\; N < 100 \tag{2.57}$$

Often the PDF of a random process cannot be directly measured or obtained. For these cases, the process is assumed to have a PDF that can be approximated by a known standard distribution. Such common distributions include the Uniform, Gaussian, and Exponential to name a few. The Uniform distribution is the most simplistic

distribution form encountered. It simply states that a random variable X has an equal probability of representing values in the interval (a,b). It is thus given by:

$$f_X(x) = \frac{1}{b-a} \quad for \;\; a \le x \le b \tag{2.58}$$

Often the Uniform distribution is used to represent the phase of a deterministic periodic signal such as a sine wave. For such cases, the PDF may be written as:

$$f_\Theta(\theta) = \frac{1}{2\pi} \quad for \;\; 0 \le \theta \le 2\pi \tag{2.59}$$

A Gaussian random variable is defined as one having the following probability density function (PDF):

$$f_X(x) = \frac{1}{\sqrt{2\pi\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2.60}$$

Many random physical phenomena tend to produce density functions characteristic of the Gaussian PDF. Thus, this PDF is typically used to model random signal noise for a wide range of applications.

The most commonly found deterministic data are periodic in nature and can be decomposed into a collection of harmonically related sine waves. Consider a single sine wave with random amplitude X and uniformly distributed phase the given by:

$$x(t) = X \sin(2\pi ft + \theta) \tag{2.61}$$

The PDF is given by:

$$f_X(x) = \frac{1}{\pi\sqrt{2\sigma^2 - x^2}} \quad for \;\; |x| < X \tag{2.62}$$

which has zero mean and variance of $\sigma^2 = x^2/2$.

Consider a complex Gaussian noise signal of the form:

$$z(t) = r(t)\, e^{j\theta(t)} = x(t) + jy(t) \tag{2.63}$$

where $r(t)$ is the magnitude or envelope and $\theta(t)$ is the phase. If X and Y are statistically independent Gaussian random variables with equal variance and zero mean, the PDF of the signal magnitude defined by $Z = \sqrt{X^2 + Y^2}$ produces the Rayleigh distribution [77]:

$$f_Z(z) = \frac{ze^{-\frac{z}{2\sigma^2}}}{\sigma^2} \tag{2.64}$$

Since the Fourier transform operation does not modify the random variable distribution type, the above PDF is representative of the magnitude spectra for a real time valued Gaussian noise signal [78]. To obtain an

expression for the power spectra PDF of Gaussian noise, we simply take the square of a Rayleigh distributed random variable according to $Y = X^2$. This produces the Exponential distribution given by [79] :

$$f_X(x) = \lambda e^{-\lambda x} \tag{2.65}$$

where the scale parameter λ is given by $\lambda = 1/2\sigma^2$ producing the following form:

$$f_X(x) = \frac{e^{-\frac{x}{2\sigma^2}}}{2\sigma^2} \tag{2.66}$$

The Exponential distribution can also be completely described using the more generalized Gamma distribution which is given by:

$$f_X(x) = \frac{\beta^{-\alpha} x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)} \tag{2.67}$$

where α and β are scale parameters, and $\Gamma(\,)$ is the Gamma function. Using $\alpha = 1$ and $\beta = 1/\lambda$ gives the from previously given by equation (2.65). Note that the sum of N Exponential distributions can also be described by the Gamma distribution using $\alpha = N$ and $\beta = 1/N\lambda$. Expressing this distribution with respect to the original time domain Gaussian parameters gives:

$$f_X(x) = \frac{2^{-N} e^{-\frac{Nx}{2\sigma^2}} x^{-1+N}}{\Gamma(N)} \left(\frac{\sigma^2}{N}\right)^{-N} \tag{2.68}$$

Consider the case of a complex signal containing a sine wave of amplitude A in Gaussian noise. It can be shown that the resulting distribution for the signal magnitude is given by the Rice distribution [78]:

$$f_X(x) = \frac{x}{\sigma^2} e^{\frac{-(x^2+A^2)}{2\sigma^2}} I_0\left(\frac{Ax}{\sigma^2}\right) \tag{2.69}$$

where I_0 is the modified Bessel function of the 0$^{\text{th}}$ order. The above PDF is representative of the Fourier transform magnitude spectra for a real-valued sinusoidal signal in Gaussian noise [78]. Taking the square of a Rician random variable produces an expression for the PDF associated with the Fourier transform power spectra which can be modeled by the Nakagami distribution [80]:

$$f_X(x) = \frac{1}{2\sigma^2} e^{\frac{-(x+A^2)}{2\sigma^2}} I_0\left(\frac{A\sqrt{x}}{\sigma^2}\right) \tag{2.70}$$

Often, one may require the PDF resulting from some mathematical operation regarding one or more independent random variables. For two independent random variables X and Y, the resulting distribution of the sum according to $Z = X + Y$ is given by the following convolution integral:

$$f_Z(z) = \int_{-\infty}^{\infty} f_Y(z-x) f_X(x) dx = \int_{-\infty}^{\infty} f_Y(y) f_X(z-y) dy \tag{2.71}$$

Similarly, the PDF for the difference between two random variables $Z = X - Y$ is:

$$f_Z(z) = \int_{-\infty}^{\infty} f_Y(z+x) f_X(x) dx = \int_{-\infty}^{\infty} f_Y(y) f_X(z+y) dy \tag{2.72}$$

The PDF for the product of two independent random variables X and Y according to $Z = X \cdot Y$ is given by:

$$f_Z(z) = \int_{-\infty}^{\infty} f_Y(z/x) f_X(x) \frac{1}{|x|} dx = \int_{-\infty}^{\infty} f_X(z/y) f_Y(y) \frac{1}{|y|} dy \tag{2.73}$$

The division of two independent random variables X and Y according to $Z = X/Y$ is given by:

$$f_Z(z) = \int_{-\infty}^{\infty} y\, f_X(y \cdot z) f_Y(y) dy \tag{2.74}$$

2.6.2 - Spectral Estimation

The conversation of a stationary random signal to the frequency domain may be achieved via the Discrete Fourier Transform (DFT) [76]:

$$X(m) = \sum_{n=0}^{L-1} x(n) e^{\frac{-j2\pi mn}{L}} \tag{2.75}$$

where $\{m = 0,1,2, \dots, L-1\}$ is the DFT output bin and L is the input signal length. Alternately, transformation from the frequency domain to the time domain can be achieved via the Inverse Discrete Fourier Transform (IDFT) according to:

$$x(n) = \sum_{m=0}^{L-1} X(m) e^{\frac{j2\pi mn}{L}} \tag{2.76}$$

Often scaling factors are used to maintain amplitude or variance uniformity when transforming to and from the frequency domain. Table 2-1 displayed below provides common scaling factors and their effect on signal magnitude and variance for real-valued time domain signals.

Table 2-1: Effect of FFT scaling on signal amplitude and variance.

Scale Factor (*SF*)	**Magnitude (*A*)**	**Variance (σ^2)**
Nothing	$LA/2$	$L\sigma^2$
$1/L$	$A/2$	σ^2/L
$2/L$	A	$4\sigma^2/L$
$1/\sqrt{L}$	$\sqrt{L}A/2$	σ^2

In addition to scaling, the DFT operation may incorporate some form of windowing function to minimize spectral leakage caused by the abrupt termination of values at the signal endpoints. Thus, the general form for the scaled and windowed DFT is given by:

$$X(m) = SF \cdot \sum_{n=0}^{L-1} w(n)x(n)e^{\frac{-j2\pi mn}{L}} \tag{2.77}$$

where SF is the scale factor and w is the window function. Common window functions include the Hann, Hamming, Blackman-Harris, and Kaiser windows to name a few [76]. Typically, the Fast Fourier Transform (FFT) is utilized for frequency transformations rather than the DFT since it greatly reduces computational load by eliminating redundancies found in the DFT calculation process. It can be shown that a reduction of approximately $L/2 \cdot \log_2(L)$ can be achieved [76]. Often it is more intuitive to the express the DFT or FFT output in terms of frequency values rather than bin number. This may be accomplished via the following transformation:

$$f = \frac{mf_s}{L} \tag{2.78}$$

where f_s is the digital sampling frequency, and the resolution or spacing between each frequency bin is given by f_s/L. Thus, the DFT output given above may be rewritten as:

$$X(f) = \sum_{n=0}^{L-1} x(n)e^{\frac{-j2\pi fn}{f_s}} \tag{2.79}$$

From this point forth, signals in the frequency domain will be expressed in terms of their frequency bin value rather than discrete bin number.

The output of the FFT is a complex signal which is often written in the following form to facilitate algebraic manipulation:

$$X(f) = |X(f)|e^{j\theta(f)} \tag{2.80}$$

where $\theta(f)$ represents the signal phase as a function of frequency. The magnitude and phase are thus given by the following:

$$|X(f)| = \sqrt{X_{rel}(f)^2 + X_{img}(f)^2} \tag{2.81}$$

$$\theta(f) = \text{Arg}[X(f)] = \tan^{-1}\left(\frac{X_{img}(f)}{X_{rel}(f)}\right) \tag{2.82}$$

The scaled one-sided power distribution also known as the Power Spectral Density (PSD) can be approximated by the following equation which is commonly referred to as the Periodogram [75]:

$$S_{xx}(f) = \frac{2}{L}|X(f)|^2 \tag{2.83}$$

One issue encountered with performing FFT operations on finite random signals is increased noise variance in the final spectral estimation. Typically, a longer signal segment is desired when computing the FFT since it will maximize component detection through a combination of side lobe reduction and increased bin resolution. However, in doing so the variance of random noise component(s) will also increase [78]. To reduce these levels, multiple spectra may be averaged together in either a coherent or incoherent manner. This forms the basis of the well-known Welch and Bartlett Periodograms which were developed specifically to address the problem of Fourier-induced variance [81].

2.6.3 - Similarity Measures

The cross-correlation is a temporal processing procedure that can be utilized to determine the similarity between signals for the purpose of extracting features, recognizing patterns, localizing sound sources, and minimizing signal noise [82-89]. It is a measure of the similarity between signals as a function of the time lag between them. For two discrete signals $x(n)$ and $y(n)$ it is defined as:

$$R_{xy}(\tau) = E[x(n)y(n+\tau)] = \frac{1}{L}\sum_{n=0}^{L-\tau-1} x(n)y(n+\tau) \tag{2.84}$$

where τ is the time lag index, and L is the signal segment length. The above operation may also be performed on a single signal to establish the autocorrelation which is given by:

$$R_{xx}(\tau) = \frac{1}{L}\sum_{n=0}^{L-\tau-1} x(n)x(n+\tau) \tag{2.85}$$

In the context of acoustics, the cross-correlation may be utilized to obtain time difference of arrival (TDOA) between two or more sensing elements subject to either a continuous or impulsive acoustic source. The autocorrelation may be used for finding repeating patterns within a signal, such as the presence and frequency of a periodic component obscured by noise.

In addition to the definition presented above, the cross-correlation function can also be approximated by taking the IDFT of the two-sided cross power spectrum:

$$R_{xy}(\tau) = \sum_{f=0}^{L-1} S_{xy}(f)\, e^{\frac{j2\pi f \tau}{f_S}} \tag{2.86}$$

where the cross-power spectrum S_{xy} is defined as

$$S_{xy}(f) = S_x(f) \cdot S_y^*(f) \tag{2.87}$$

where S_x and S_y are the two-sided complex frequency spectra and * denotes complex conjugation. The cross-power spectrum is a complex valued function where the magnitude and argument provides the power shared by and phase difference between the two signals as a function of frequency.

2.6.4 - Digital Filtering

The filtering of digital signals is perhaps the most basic yet most important operation commonly performed. It is typically implemented in the time domain in order to remove undesired components of some specified frequency value. Digital filters may be generally classified according to their impulse response function (FIR or IIR) and their implementation form (fixed or adaptive). Finite Impulse Response Filters (FIR) filters constitute the most basic filtering operation. They are given by the convolution of some finite coefficient vector with the time domain signal of concern:

$$y(n) = \sum_{k=0}^{N-1} h(k)\, x(n-k) \tag{2.88}$$

where $h(k)$ is the impulse response (also known as the coefficient vector) and N is the vector size or number of filter taps. FIR filters are inherently stable and often provide a linear phase response. The frequency response is defined by:

$$H(\omega) = \sum_{k=-\infty}^{\infty} h(k) e^{-j\omega k} \tag{2.89}$$

which may be represented in the z-domain as:

$$H(z) = \sum_{k=-\infty}^{\infty} h(k) z^{-k} \tag{2.90}$$

The Z-transform converts a discrete-time signal, which is a sequence of real or complex numbers, into a complex frequency domain representation. It can be considered as a discrete-time equivalent of the Laplace transform, and provides a means of examining properties of systems such as stability and convergence which may otherwise be impossible using standard time or frequency domain methods. This aspect will be illustrated in Section 3.4 where an FIR filter is constructed from a IIR prototype using Z-transforms.

Given the frequency response $H(\omega)$, the filter impulse response may also be determined via the inverse Fourier transform according to:

$$h(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(\omega) \cdot e^{j\omega k} d\omega \tag{2.91}$$

where equations (2.89) and (2.91) form a Fourier transform pair.

The second class of filters are referred to as Infinite Impulse Response Filters (IIR). Unlike the FIR form, these filters are recursive in nature. As a result, they often produce a non-linear phase response and may become unstable depending on the choice of coefficient values. The IIR filter may be implemented according to the following convolution operation:

$$y(n) = \sum_{k=0}^{N_a-1} a(k)\, x(n-k) + \sum_{i=1}^{N_b-1} b(i)\, y(n-i) \tag{2.92}$$

where a and b is the feedforward and feedback filter coefficients respectively, with N_a and N_b giving the length of each coefficient vector. The transfer function and frequency response are given by the following equations respectively:

$$H(z) = \frac{\sum_{k=0}^{N_a-1} a(k)\, z^{-k}}{1 - \sum_{i=1}^{N_b-1} b(i)\, z^{-i}} \tag{2.93}$$

$$H(\omega) = \frac{\sum_{k=0}^{N_a-1} a(k)\, e^{-j\omega k}}{1 + \sum_{i=1}^{N_b-1} b(i)\, e^{-j\omega i}} \tag{2.94}$$

2.6.5 - Signal Quantification

The signal-to-noise ratio (SNR) is perhaps the most commonly used parameter to quantify random and deterministic signals. It is defined as the ratio of the noise-free signal power to the total noise power, and is typically expressed in decibels (dB) according to:

$$SNR = 10\log_{10}\left[\frac{Signal\ Power}{Noise\ Power}\right] = 10\log_{10}\left[\frac{\psi_s^2}{\psi_n^2}\right] = 10\log_{10}\left[\frac{\sigma_s^2 + \mu_s^2}{\sigma_n^2 + \mu_n^2}\right] \tag{2.95}$$

For the case of a sinusoidal signal in zero mean Gaussian noise, the above form produces the following familiar expression:

$$SNR = 10\log_{10}\left[\frac{A^2}{2\sigma_n^2}\right] \tag{2.96}$$

where A is the amplitude of the sinusoidal function. For a signal composed of multiple sinusoids or harmonic components, the SNR is given by:

$$SNR = 10\log_{10}\left[\frac{\sum_{r=1}^{R} \frac{A_r^2}{2}}{\sigma_n^2}\right] \tag{2.97}$$

where A_r is the amplitude of the r^{th} harmonic component. For discrete time signals having unknown signal and noise distribution properties, the SNR may be calculated directly from the sampled data according to:

$$SNR = 10\log_{10}\left[\frac{\sum_{n=0}^{N} [x_s(n)]^2}{\sum_{n=0}^{N} [x_n(n)]^2}\right] \tag{2.98}$$

In addition to the time domain realization, we may also calculate SNR values in the frequency domain according to:

$$SNR = 10\log_{10}\left[\frac{\sum_{f=0}^{f_s/2}|X_s(f)|^2}{\sum_{f=0}^{f_s/2}|X_n(f)|^2}\right] \tag{2.99}$$

where f_s is the sampling frequency. In many instances, statistical information regarding the noise and/or signal is unknown and the two components cannot be separated into two separate data streams. In such cases, the SNR of the mixed signal is calculated by approximating the contributions of each component to the total signal power. This is most easily accomplished by simply summing across specified bandwidth regions of the signal power spectrum:

$$SNR = 10\log_{10}\left[\frac{\sum_{f\in\hat{S}}|X_s(f)|^2}{\sum_{f\in\hat{N}}|X_n(f)|^2}\right] \tag{2.100}$$

where the signal and noise containing bandwidth regions are specified by the discrete index sets $\hat{S}$ and $\hat{N}$ respectively. For narrowband signals we may simply take the spectral peak value (and a small number of surrounding points to account for spectral leakage) and compare it to the remainder of the spectrum. Here, spectral leakage refers to the "smearing" of signal energy across frequency multiple bins. It occurs when the true frequency of the detected signal does not coincide with the center of one the discrete FFT bins. In many instances however, we may obtain a more practical or meaningful value by only considering noise in a localized region surrounding the signal peak. In doing so, any spurious noise peaks in the region which highly influence signal detectability will largely dictate the minimum noise level. Below Figure 2-11 illustrates the SNR and Effective SNR for a narrowband signal in broadband noise with spurious noise peaks. As will later be shown in Chapter 4, the Effective SNR typically provides a more meaningful measure when performing signal detection operations since it provides the dynamic range for which a signal may be detected.

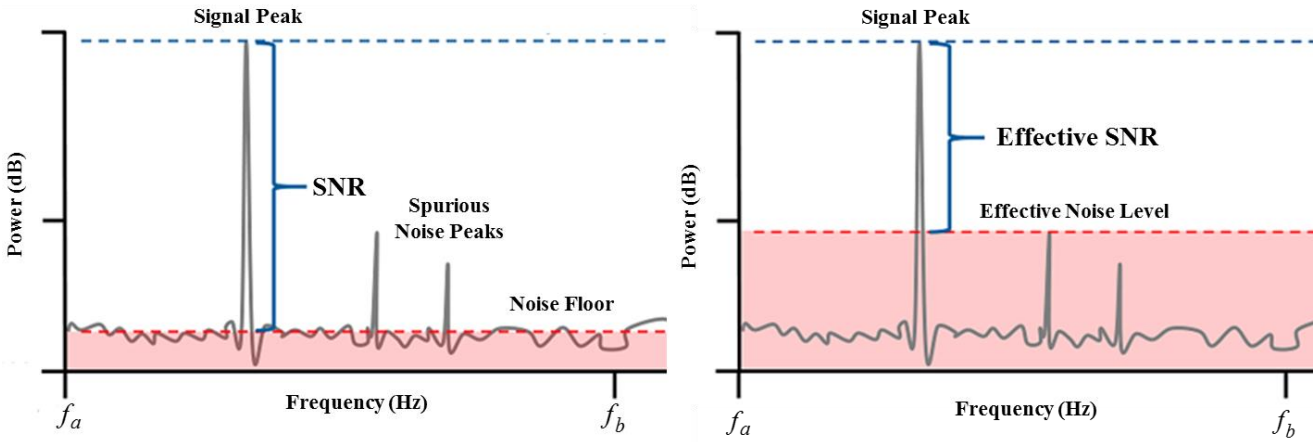

Figure 2-12: Illustration of standard SNR and effective SNR values.

- 3 - Narrowband Noise Removal

The following chapter presents numerous techniques to adaptively filter harmonic narrowband noise without using any reference signal or producing any phase distortions. Extensions to the IIR notch filtering method proposed by Tan [90-93] are presented to include systems containing multiple signals corrupted with multiple harmonic noise sources. An FIR approximation of this filter using the impulse response truncation method is presented which is capable of producing similar performance without the non-linear phase response inherent with IIR filters. In addition, a simplistic method of transforming any arbitrary FIR filter to achieve a true zero-phase response is provided. Using the proposed method, it is shown that an FIR Comb filter can be zero-phase transformed and implemented in a referenceless adaptive form to notch filter harmonic narrowband noise components. The performance of the presented filtering methods is evaluated using both computer-generated data and that obtained from experimental studies previously described.

3.1 - Introduction

As previously discussed, the major noise component present in the acquired acoustic signals is the harmonic narrowband noise generated from the aircraft propulsion system. Often during flight operations, engine speeds may vary significantly with time while the aircraft preforms maneuvers as required to meet mission objectives. This results in highly non-stationary noise components which must therefore be removed via some adaptive filtering method. However, a number of problems are encountered when attempting to use standard adaptive methods for this particular application.

The first problem arises from the requirement of a noise only reference signal to ensure the filter converges to an optimal or near optimal configuration [94]. For the application at hand, this is impossible using a microphone-based reference since any microphone placed on the aircraft will also be exposed to the acoustic source to be detected. An alternative approach would be to use a non-acoustic sensor such as a tachometer to obtain the noise approximation via frequency tracking of the aircraft motor. However, the use of such devices often requires more advanced onboard data acquisition systems and supporting hardware since these sensors will have different electrical requirements than the corresponding microphones. In addition, the acquired tachometer signal will most often need some degree of conditioning and preprocessing before it can be utilized directly with the acoustic signals. Thus, for instances in which size, weight, and power consumption of the sensing system is very limited this approach may not be a viable option.

The second issue arises from the unconstrained manner in which conventional adaptive filters operate, often producing non-linear phase distortions. Since the acquired signals will ultimately be utilized in some combined form for detection and/or localization purposes, it is imperative that phase distortions are not produced in the spatial or temporal domain. Phase information in the spatial domain is utilized for source localization operations by examining the phase shift present across an array of sensors. In theory, phase variations in the temporal domain will not affect most localization operations provided the variations are produced equally across all

signals in time. In some instances, array processing methods can be utilized to from an adaptive linear spatial filter to remove narrowband noise [95]. However, these methods cannot effectively be utilized for the particular application at hand since the noise source is essentially located within the array spatial domain.

Temporal phase variations are often a concern when performing signal detection operations. Many signal enhancement methods such as the phase acceleration and coherence-based processors to be later discussed rely on a constant phase progression of source components to discriminate unwanted noise components. Thus, in order to maximize detection and localization capabilities, all filtering should be performed without producing any phase distortions either in the temporal or spatial domain. The processes of filtering without producing any phase distortions, either linear or non-linear, will be referred to as zero-phase filtering from this point forth.

3.2 - Background Information

3.2.1 - Adaptive Filtering

3.2.1.1 - *Standard Form*

Adaptive filtering is an effective method to actively remove unwanted noise from non-stationary signals and was largely pioneered by Widrow in the mid 1970's [94]. Adaptive filters are used when the signal of concern is non-stationary and/or prior information regarding the unwanted noise component(s) is unknown. The principle behind the approach is to continuously adjust filter coefficients according to some cost function which establishes how well the undesired component is being removed. A very common method, known as the Least Mean Squares (LMS) approach utilizes the gradient of the instantaneous squared error between the filtered signal and a reference signal to modify filter coefficients. In addition to the LMS method, many other adaptive algorithms have been developed and presented in the literature such as the Normalized LMS, Sign LMS, Normalized Sign LMS, Leaky LMS, Recursive LMS and filtered-x LMS [81, 96-98]. Figure 3-1 displayed below provides a block diagram of the adaptive filtering process.

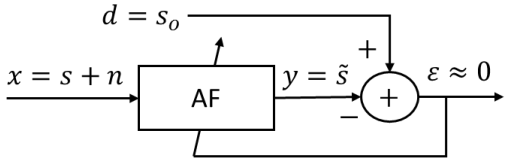

Figure 3-1: Block diagram of standard adaptive filtering model.

For a causal FIR filter, the filtered output is given by the convolution of the input signal and the filter weights:

$$y(n) = \sum_{k=0}^{N-1} a_k(n)\, x(n-k) = \mathbf{w^T X}(n) \tag{3.1}$$

$$\mathbf{w}^T = \left[a_0(n), a_1(n), \ldots, a_{N-1}(n)\right]$$

$$\mathbf{X}(n) = \left[x(n), x(n-1), \ldots, x(n-N+1)\right]^T$$

The signal error ε is given by:

$$\varepsilon(n) = d(n) - y(n) \tag{3.2}$$

For the LMS algorithm, the filter coefficients are updated according to:

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu \nabla J(n) \tag{3.3}$$

where μ is the step size and J is the cost function which is given by the mean square error. Typically, this is approximated by the instantaneous error squared:

$$J(n) = E[\varepsilon^2(n)] \approx \varepsilon^2(n) \tag{3.4}$$

for which the gradient is:

$$\nabla J(n) = -2\varepsilon(n)\mathbf{X}(n) \tag{3.5}$$

Thus, the final form is given by:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + 2\mu\ \varepsilon(n)\mathbf{X}(n) \tag{3.6}$$

In many instances, a reference approximation of the clean signal cannot be obtained but an approximation of the undesired noise component is available. For this case, the filter input is now the noise approximation and the reference becomes the noise-corrupted signal. The error signal will then converge to the clean desired signal rather than zero.

Similar to the FIR filter, recursive filters may be implemented via an adaptive filtering approach. The general form for a recursive IIR filter is given by the following:

$$y(n) = \sum_{j=0}^{N-1} a_j(n)\ x(n-j) + \sum_{k=1}^{M-1} b_k(n)\ y(n-k) \tag{3.7}$$

which may also be represented in vector form via the following equations:

$$y(n) = \mathbf{a}^{\mathbf{T}}\mathbf{x}_n + \mathbf{b}^{\mathbf{T}}\mathbf{y}_{n-1} = \mathbf{w}^{\mathbf{T}}\mathbf{u}_n \tag{3.8}$$

$$\mathbf{w}^T = [\mathbf{a}_n^T, \mathbf{b}_n^T] = [a_0(n)\ a_1(n) \ldots a_{N-1}(n), b_1(n)\ b_2(n) \ldots b_{M-1}(n)]$$
$$\mathbf{u}_n^T = [\mathbf{x}_n^T, \mathbf{y}_{n-1}^T] = \left[x(n)\ x(n-1) \ldots x(n-N+1), y(n-1)\ y(n-2) \ldots y(n-M+1)\right]$$

The coefficient update equation is then given by:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + 2\mu\ \varepsilon(n)\frac{\partial y(n)}{\partial \mathbf{w}(n)} \tag{3.9}$$

and the gradient can be obtained via the following:

$$\begin{aligned}\frac{\partial y(n)}{\partial \mathbf{w}(n)} &= \left[\frac{\partial y(n)}{\partial a_0(n)}, \frac{\partial y(n)}{\partial a_1(n)}, \ldots, \frac{\partial y(n)}{\partial a_{N-1}(n)}, \frac{\partial y(n)}{\partial b_1(n)}, \frac{\partial y(n)}{\partial b_2(n)}, \ldots, \frac{\partial y(n)}{\partial b_{M-1}(n)}\right]^T \\ &= \left[\alpha_0(n), \alpha_1(n), \ldots, \alpha_{N-1}(n), \beta_1(n), \beta_2(n), \ldots, \beta_{M-1}(n)\right]\end{aligned} \tag{3.10}$$

$$\alpha_j(n) = \frac{\partial y(n)}{\partial a_j(n)} \approx x(n-j) + \sum_{k=1}^{N-1} a_k(n)\ \alpha_j(n-j) \tag{3.11}$$

$$\beta_j(n) = \frac{\partial y(n)}{\partial b_k(n)} \approx y(n-k) + \sum_{k=1}^{M-1} b_k(n)\ \beta_j(n-k) \tag{3.12}$$

It is obvious from the above equations that adaptive implementation of the general IIR filter form is much more complex than that of FIR filters. It will be later shown however, that for instances in which no reference signal is utilized the cost function and subsequent gradient becomes greatly simplified.

3.2.1.2 - *Referenceless Form*

There are two basic approaches that may be utilized to construct a referenceless adaptive filter for periodic noise removal. The first and perhaps most popular involves constructing a reference signal that is correlated to the undesired noise component (or alternately desired signal component), and then apply standard adaptive filtering methods. Adaptive Line Enhancement (ALE) is an example of such an approach. The basic concept behind ALE is to use a delayed version of the noisy signal to serve as the reference input. By doing so, random broadband noise components will decorrelate, while the correlation between periodic components will remain. FIR based ALE methods have been successfully utilized for many applications and are capable of maintaining linear-phase characteristics [99]. However, the technique performs poorly when both the desired and undesired components are of a narrowband periodic form [100]. In addition, no constraints are present to ensure a zero-phase or even linear-phase response. Methods have been proposed to deal with this scenario by actively notch filtering any periodic components prior to ALE [101-104]. However, these methods attempt to directly measure the instantaneous frequency of noise components using either a zero-crossing average or the method proposed in [105]; both of which tend to perform poorly for non-stationary signals with harmonic components. Other methods of tracking periodic components exist such as the Hilbert transform, Polynomial Phase Modeling, and Adaptive Phase Locked Loops. However, these methods are only effective for single component signals or determining the instantaneous frequency of modulated signals [106-110].

The second approach involves using a filter that has a highly constrained response and may be defined by a single parameter variable such as that found with IIR notch and FIR Comb filters. Consider some arbitrary notch filter whose notch location is defined by the normalized frequency value θ. For adaptive implementation, coefficient values are updated such that the mean square of the output error is minimized. However, since no reference signal is being utilized the filtered output will now serve as the error signal $\varepsilon^2(n) = y^2(n)$. Thus, the filtering objective now becomes minimization of the filtered signal power which will be achieved when all narrowband noise components are removed. Application of the LMS algorithm previously given by equation (3.3) now produces the following form:

$$\theta(n+1) = \theta(n) - 2\mu y(n)\beta(n) \tag{3.13}$$

where μ is the step size and $\beta(n)$ is the gradient function which can be determined directly from the filter output equation according to:

$$\beta(n) = \frac{\partial y(n)}{\partial \theta(n)} \tag{3.14}$$

This technique has been employed with great success using second order IIR notch filters [90-93]. It will later be shown that this approach may also be utilized to construct referenceless adaptive Comb and FIR notch filters. A depiction of this basic filtering setup is displayed below in Figure 3-2.

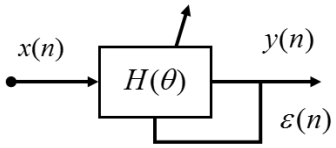

Figure 3-2: Constrained referenceless adaptive filtering.

3.2.2 - Zero-Phase Phase Filtering

3.2.2.1 - *Distortionless IIR Filtering*

Zero-phase filtering is not a new area of study with accounts of distortionless recursive filtering dating back to the early 1970's [111-116]. The field has generally been focused around the use of IIR filters since they often achieve a more desirable magnitude response with much less computation but will inherently produce a non-linear phase response. Traditionally, distortionless IIR filtering has been reserved for offline processing applications since the phase correction processes requires a time reversal of the input sequence [111]. However, more recent methods have been proposed to facilitate real-time implementation by various overlapped [113, 117, 118] and non-overlapped block-based processing methods [112, 116]. The general approach for two-pass IIR filtering is displayed below in Figure 3-3.

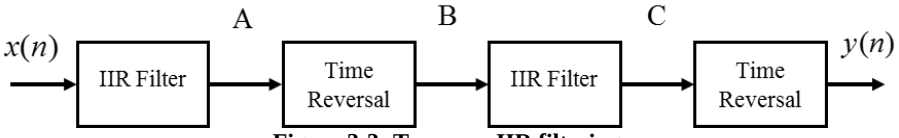

Figure 3-3: Two-pass IIR filtering.

Consider the case where some spectral component of $x(n)$ has an arbitrary initial phase response given by $\alpha(\omega)$ in degrees or radians. If the IIR filter has a phase response given by $\beta(\omega)$, then the phase of the filtered signal at stage A will now be given by $\alpha(\omega) + \beta(\omega)$. The first time reversal step will now conjugate this phase response and introduce some constant phase shift θ such that the final signal phase is now $-\alpha(\omega) - \beta(\omega) + \theta$. Applying the IIR filter again will once more apply a phase shift of $\beta(\omega)$ giving the following phase at point C: $-\alpha(\omega) - \beta(\omega) + \theta + \beta(\omega) = -\alpha(\omega) + \theta$. Application of one final time reversal will again conjugate the signal phase apply another constant phase shift of θ giving: $(-\alpha(\omega) + \theta) \rightarrow -(-\alpha(\omega) + \theta) + \theta = \alpha(\omega)$, which is the initial phase response of the input signal prior to any filtering. For the case of adaptive IIR filtering,

the time-reversed signal is simply filtered again using the time-reversed sequence of filter weights established in the forward direction (initial stage) filter.

Due to the inherent nature of recursive filtering, there will be transient effects at the beginning and end of the filtered signal segment(s). To avoid these effects, filtered segments are typically truncated at each end before reconstructing the final signal. As depicted below in Figure 3-4, the discrete time signal is initially broken into overlapping blocks of length L_i. Upon filtering in the forward and reverse directions, each block is then truncated to a final length of L_f effectively removing any transient effects. Typically, the truncation length will be at least 4 to 5 times the filter order [76]. It should also be noted that since the filtering operation is being performed twice, the magnitude response will become squared meaning a doubling (in dB) of passband ripple and stopband attenuation values.

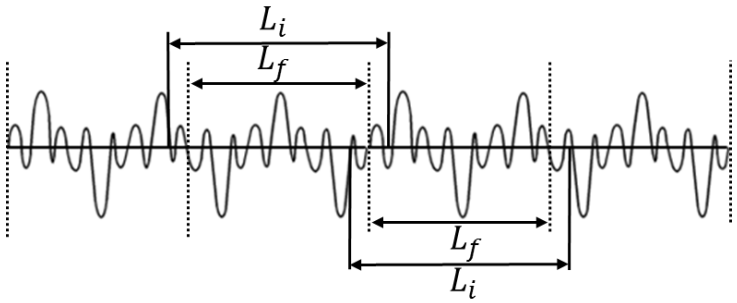

Figure 3-4: Overlapped block-based filtering approach.

3.2.2.2 - *Distortionless FIR Filtering*

Compared to IIR filters, zero-phase FIR filtering has received much less attention in the scientific community. Most relevant work has involved the development of adaptive linear-phase FIR filters, since these forms are adequate for most applications [119-126]. A linear phase response will be achieved when the filter impulse response satisfies either the symmetric or anti-symmetric requirements given by:

$$h_{sym}(n) = h(N-1-n) \tag{3.15}$$

$$h_{asym}(n) = -h(N-1-n) \tag{3.16}$$

where N is the number of filter coefficients. FIR filters that meet the above specifications can also be modified to achieve a zero-phase response by converting the filter to a non-causal form. The general procedure for doing so is to simply take the filter output with respect to the center tap position. Consider for example a causal FIR filter, which has the following arbitrary frequency response:

$$H\left(e^{j\omega}\right) = \sum_{k=0}^{N-1} b_k e^{-jk\omega} \tag{3.17}$$

For a symmetric filter with an odd number of coefficients, we may shift the filtering process about a zero-point time reference and perform the filtering as follows:

$$H\left(e^{j\omega}\right)=\sum_{k=-(N-1)/2}^{(N-1)/2} b_{k+(N-1)/2}e^{-jk\omega} \tag{3.18}$$

This produces a non-causal form that is symmetric about $k=0$. Evaluating the above summation produces the following frequency response, which has only real components:

$$H\left(e^{j\omega}\right)=b_0+2\sum_{k=1}^{(N-1)/2} b_k\cos(k\omega) \tag{3.19}$$

Although the filter is considered to have zero-phase, in actual fact its phase response $\phi_H(\omega)$ switches between 0 and π according to:

$$\begin{cases}\phi_H(\omega)=0\ ,\ \mathrm{Re}\left[H\left(e^{j\omega}\right)\right]\geq 0\\ \phi_H(\omega)=\pi\ ,\ \mathrm{Re}\left[H\left(e^{j\omega}\right)\right]<0\end{cases} \tag{3.20}$$

Thus, symmetric FIR filters do not exhibit a true zero-phase response but rather a π-phase response. Because of this, they are sometimes termed π-phase filters instead [127]. For odd impulse response functions where $h(n)=-h(-n)$, the phase switches between values of $\pm\pi/2$ rather than 0 and π.

Typically, fixed FIR filters are designed such that they adhere to the linear phase symmetry requirements previously given. However, unstrained adaptive implementation of such filters will most often result in a non-linear form since coefficient values are established without regard to symmetry (or anti-symmetry). The general solution to this problem is to place constraints on the operation such that symmetry is instead maintained. This general procedure has been successfully used to produce linear-phase adaptive filters using a variety of implementation algorithms [119-126]. For this scenario, the general filtering process is given by:

$$y(n)=\sum_{k=0}^{2N-2} b_{|N-k-1|}x(n-k)=\mathbf{B}^{\mathbf{T}}\,\mathbf{X}(n) \tag{3.21}$$

$$\mathbf{B}=\left[b_{N-1}(n)\ldots b_0(n)\ldots b_{N-1}(n)\right]^T$$
$$\mathbf{X}(n)=\left[x(n)\ x(n-1)\ldots x(n-2N+2)\right]^T$$

If an odd number of coefficients are used, the filter can be time-shifted to a non-causal form producing a π-phase filter as previously described. Implementation of the time-shifted form would thus be given by:

$$y(n)=b_0x(n)+\sum_{k=1}^{N-1}b_kx(n+k)+\sum_{k=1}^{N-1}b_kx(n-k)=\mathbf{B}^{\mathbf{T}}\,\mathbf{X}(n) \tag{3.22}$$

$$\mathbf{B}=\left[b_{N-1}(n)\ldots b_0(n)\ldots b_{N-1}(n)\right]^T$$
$$\mathbf{X}(n)=\left[x(n+N-1)\ldots x(n)\ldots x(n-N+1)\right]^T$$

Similar to the IIR case, non-linear FIR filters may also produce phase distortionless outputs via two-stage filtering operations. Since non-causal FIR filters are inherently stable however (unlike IIR filters), the operation

can be performed without time reversal or using block-based techniques. A linear phase output can be obtained by simply cascading the filter with its reflected self. By doing so, any phase distortions produced by the initial filter are linearized by the second stage. This process is depicted below by Figure 3-5.

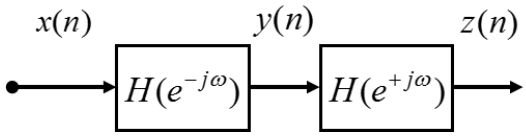

Figure 3-5: Zero-Phase FIR filtering.

The filtered output at the first and second stages will be given by the following equations respectively:

$$y(n) = \sum_{k=0}^{N-1} b_k \, x(n-k) \tag{3.23}$$

$$z(n) = \sum_{k=0}^{N-1} \bar{b}_k \, y(n-k) \tag{3.24}$$

where $b_k = [b_0 \; b_1 \ldots b_{N-1}]$ is the initial filter coefficients and $\bar{b}_k = [b_{N-1} \ldots \; b_1 \; b_0]$ are the reflected coefficients. The two filter stages may also be combined to give the following overall expression:

$$\begin{aligned} z(n) &= \sum_{k=0}^{N-1} \bar{b}_k \sum_{k=0}^{N-1} b_k \, x(n-k) \\ &= \sum_{k=0}^{N-1} b_{n-k-1} \sum_{k=0}^{N-1} b_k \, x(n-k) \end{aligned} \tag{3.25}$$

To achieve a zero-phase rather than linear-phase response we may simply convolve the initial causal filter with a reflected anti-causal twin to obtain a complete non-causal form. By doing so, any phase distortions that would be produced at the causal stage are cancelled by the second anti-causal portion. Consider the general frequency response function for the causal and anti-causal forms respectively:

$$H_c\left(e^{j\omega}\right) = \sum_{k=0}^{N-1} b_k e^{-jk\omega} \tag{3.26}$$

$$H_a\left(e^{j\omega}\right) = \sum_{k=0}^{N-1} b_k e^{+jk\omega} \tag{3.27}$$

where b is the filter coefficient vector. We may construct a new non-causal zero-phase filter by simply taking the product of the two forms:

$$\begin{aligned} H_0\left(e^{j\omega}\right) &= \sum_{k=0}^{N-1} b_k e^{-jk\omega} \sum_{k=0}^{N-1} b_k e^{jk\omega} \\ &= H_c\left(e^{j\omega}\right) H_c^*\left(e^{j\omega}\right) \\ &= \left|H_c\left(e^{j\omega}\right)\right|^2 \end{aligned} \tag{3.28}$$

It is apparent that the frequency response given above contains only real values and is always greater than zero. Thus, the phase response will always maintain a value of zero over all frequencies producing a true zero-phase

response. This is in contrast to the center tapped symmetric case previously given by equation (3.19) which may attain negative values producing a π-phase response.

To implement the filter, we simply convolve the initial coefficient vector with a flipped version of itself and apply the standard non-causal center tap approach previously described:

$$\begin{aligned} y(n) &= \sum_{k=0}^{N-1} b_k x(n-k) \circledast \sum_{k=0}^{N-1} b_k x(n+k) \\ &= \sum_{i=0}^{N-1} b_i \left(\sum_{k=0}^{N-1} b_k x(n-k+i) \right) \end{aligned} \tag{3.29}$$

Making use of symmetry, we may write the filtering operation in the following reduced computational form:

$$y(n) = x(n) \sum_{k=0}^{N-1} {b_k}^2 + \sum_{i=1}^{N-1} \left(\left[x(n-i) + x(n+i) \right] \sum_{k=0}^{N-i-1} b_k b_{k+i} \right) \tag{3.30}$$

Alternately, we may express the filter in terms of a single coefficient vector. Since convolving a vector with a flipped version of itself is equivalent to performing an autocorrelation, the combined coefficient vector is given by:

$$a_k = \sum_{n=0}^{N-k-1} b_n b_{n+k} \tag{3.31}$$

where $b_k = [b_0\ b_1 \ldots b_{N-1}]$ are the causal filter coefficients and $k = 1,2,\ldots,2N-1$. The filter can be thus implemented according to:

$$\begin{aligned} y(n) &= a_0 x(n) + \sum_{k=1}^{N-1} a_k x(n+k) + \sum_{k=1}^{N-1} a_k x(n-k) \\ &= \mathbf{A}^T\ \mathbf{X}(n) \end{aligned} \tag{3.32}$$

$$\mathbf{A} = \left[a_{N-1}(n) \ldots a_0(n) \ldots a_{N-1}(n) \right]^T$$

$$\mathbf{X}(n) = \left[x(n+N-1) \ldots x(n) \ldots x(n-N+1) \right]^T$$

By realizing that $\mathbf{A}$ will always be symmetric and contain an odd number of values, we may further reduce the number of computations required by only performing half of the convolutions such that $k = 1,2,\ldots,N-1$. The filtering operation can then be placed in the following reduced form:

$$y(n) = a_0 x(n) + \mathbf{A}^T \left[\mathbf{X_c}(n) + \mathbf{X_a}(n) \right] \tag{3.33}$$

$$\mathbf{A} = \left[a_1(n) a_2(n) \ldots a_{N-1}(n) \right]^T$$

$$\mathbf{X_c}(n) = \left[x(n-1)\ x(n-2) \ldots x(n-N+1) \right]^T$$

$$\mathbf{X_a}(n) = \left[x(n+1)\ x(n+2) \ldots x(n+N-1) \right]^T$$

Adaptive implementation via the LMS algorithm is thus given by:

$$a_o(n+1) = a_o(n) + 2\mu\ \varepsilon(n)x(n) \tag{3.34}$$

$$\mathbf{A}(n+1) = \mathbf{A}(n) + 2\mu\ \varepsilon(n)\left[\mathbf{X_c}(n)+\mathbf{X_a}(n)\right] \tag{3.35}$$

It should be noted that the convolved zero-phase filtering method can be implemented in a causal form if desired. Doing so will simply produce a linear-phase response rather than a zero-phase response.

3.2.3 - IIR Notch Filtering

The IIR notch filter is perhaps the most effective and efficient means of removing narrowband signal components. Compared to FIR filters, IIR filters are capable of achieving much tighter stopband regions while also reducing passband ripple and coefficient quantities [97]. The use of such filters has been previously studied by Tan for various applications such as filtering and tracking of harmonic signals with very good results [90-93]. Consider a second order IIR notch filter, which has the following transfer function:

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1 - 2\cos[\theta]z^{-1} + z^{-2}}{1 - 2r\ \cos[\theta]z^{-1} + r^2\ z^{-2}} \tag{3.36}$$

where θ is the normalized notch location parameter given in radians/sample, and r is the pole radius which governs the notch bandwidth. Zeros are located on the z-plane unit circle giving the notch infinite depth. In order to ensure the stability of the filter, the pole radius is constrained such that $r < 1$. The 3dB notch bandwidth may be approximated according to [128]:

$$BW = \frac{f_s}{\pi}(1-r) \tag{3.37}$$

where f_s is the signal sampling rate and $0.9 < r < 1$. The direct realization form is thus given by:

$$y(n) = x(n) - 2\cos[\theta(n)]x(n-1) + x(n-2) + 2r\cos[\theta(n)]y(n-1) - r^2 y(n-2) \tag{3.38}$$

Adaptive implementation via the LMS algorithm may be achieved by equation (3.13) previously presented, where the gradient function $\beta(n)$ can be determined directly from the filter output equation according to:

$$\beta(n) = \frac{\partial y(n)}{\partial \theta(n)} = 2\sin[\theta(n)]x(n-1) - 2r\sin[\theta(n)]y(n-1) + 2r\cos[\theta(n)]\beta(n-1) - r^2\beta(n-2) \tag{3.39}$$

If the noise is harmonic in nature, additional notches can be produced by simply cascading the filter with frequency-shifted versions of itself according to [93]:

$$H(z) = H_1(z)H_2(z)...H_M(z) = \prod_{m=1}^{M} H_m(z) \tag{3.40}$$

where m is the harmonic number, $H_m(z)$ denotes the m^{th} second-order IIR sub-filter whose transfer function is defined as:

$$H(z) = \prod_{m=1}^{M} \frac{1 - 2\cos[m\theta]z^{-1} + z^{-2}}{1 - 2r\ \cos[m\theta]z^{-1} + r^2\ z^{-2}} \tag{3.41}$$

Rather than expand out the above form and transform to the time domain to obtain a direct algebraic expression, we may instead implement in an iterative manner. This provides the benefit of maintaining one constant algebraic expression regardless of the number of stages included. The filter output at the m^{th} harmonic is thus given by the following iterative form:

$$y_m(n) = y_{m-1}(n) - 2\cos\left[m\theta(n)\right]y_{m-1}(n-1) + y_{m-1}(n-2) + 2r\cos\left[m\theta(n)\right]y_m(n-1) - r^2 y_m(n-2) \quad (3.42)$$

where $y_0(n) = x(n)$ is the initial input to the first IIR sub-filter. The gradient function can again be determined directly from the filter output equation in a recursive form according to:

$$\begin{aligned}\beta_m(n) &= \frac{\partial y_m(n)}{\partial\theta(n)} \\ &= \beta_{m-1}(n) - 2\cos\left[m\theta(n)\right]\beta_{m-1}(n-1) + 2m\sin\left[m\theta(n)\right]y_{m-1}(n-1)\ldots + \\ &\quad \beta_{m-1}(n-2) + 2r\cos\left[m\theta(n)\right]\beta_m(n-1) - r^2\beta_m(n-2) - 2mr\sin\left[m\theta(n)\right]y_m(n-1)\end{aligned} \quad (3.43)$$

where $\beta_0 = \frac{\partial y_0}{\partial\theta} = \frac{\partial x}{\partial\theta} = 0$

The notch location θ is then updated according to:

$$\theta(n+1) = \theta(n) - 2\mu y_M(n)\beta_M(n) \quad (3.44)$$

where $y_M(n)$ and $\beta_M(n)$ are the filtered output and gradient at the last harmonic stage as depicted below in Figure 3-6.

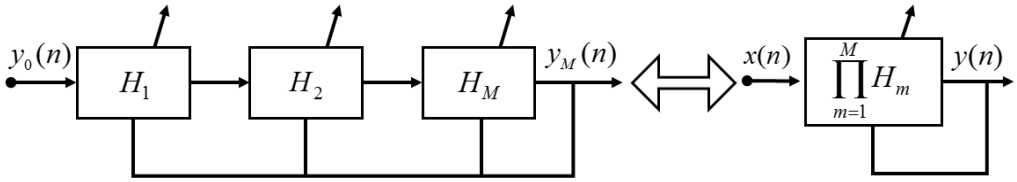

Figure 3-6: Single Input / Single Output filter system.

To increase stability and convergence performance, the current notch location can be updated based on a moving average of the past N samples according to:

$$\theta(n+1) = \frac{1}{N}\sum_{i=1}^{N}\theta(n-i) \quad (3.45)$$

where the initial starting values for θ may be obtained by simply performing an autocorrelation on the input signal and finding the frequency corresponding to maximum peak value.

It should be noted that as the pole radius r and/or number of harmonic components M increases, the transient effects associated with the filter also increase. If a considerable number of harmonics are included with radius values approaching one, roundoff errors produced during implementation may cause the filter to become marginally unstable. Stability prediction from these values however is a complex operation and is outside the scope of this thesis. Further information regarding performance of the filter can be found in [90-93].

3.3 - Adaptive Multichannel IIR Notch Filter

The previous section introduced the second order IIR notch filter and how it may be implemented in a referenceless adaptive form for systems consisting of a single signal and noise source. It will now be shown that the concepts developed by Tan [90-93] can be extended to facilitate the filtering of multiple parallel signals subject to multiple harmonic noise sources.

3.3.1 - Single Input / Multiple Output

Consider the case of a single engine fixed-wing aircraft equipped with an array consisting of K microphones. Such a scenario constitutes a Single Input / Multiple Output (SIMO) system as depicted in Figure 3-7. Using the indexed iterative form previously presented by equation (3.42), the filtered output of the m^{th} harmonic for the k^{th} sensor signal will be given by the following equation:

$$\begin{aligned} y_{m,k}(n) = {} & y_{m-1,k}(n) - 2\cos\left[m\theta_k(n)\right] y_{m-1,k}(n-1) + y_{m-1,k}(n-2) \ldots + \\ & 2r\cos\left[m\theta_k(n)\right] y_{m,k}(n-1) - r^2 y_{m,k}(n-2) \end{aligned} \tag{3.46}$$

where $\{m = 1,2,\ldots,M\}$ and $\{k = 1,2,\ldots,K\}$.

Since all signals are exposed to the same noise source, only one signal is required to obtain the noise frequency parameters. Typically, the sensor closest to the source would be chosen since the noise source power will be greatest and thus facilitating better tracking. For the closest microphone defined by the position $k = c$, the filtered output, gradient, and LMS notch update will now be given by the following equations respectively:

$$\begin{aligned} y_{m,k}(n) = {} & y_{m-1,k}(n) - 2\cos\left[m\theta_c(n)\right] y_{m-1,k}(n-1) + y_{m-1,k}(n-2) \ldots + \\ & 2r\cos\left[m\theta_c(n)\right] y_{m,k}(n-1) - r^2 y_{m,k}(n-2) \end{aligned} \tag{3.47}$$

$$\begin{aligned} \beta_m(n) = {} & \beta_{m-1}(n) - 2\cos\left[m\theta_c(n)\right]\beta_{m-1}(n-1) + 2m\sin\left[m\theta_c(n)\right] y_{m-1,k=c}(n-1) + \beta_{m-1}(n-2) \\ & \ldots + 2r\cos\left[m\theta_c(n)\right]\beta_m(n-1) - r^2\beta_m(n-2) - 2mr\sin\left[m\theta_c(n)\right] y_{m,k=c}(n-1) \end{aligned} \tag{3.48}$$

$$\theta_c(n+1) = \theta_c(n) - 2\mu y_{M,k=c}(n)\beta_M(n) \tag{3.49}$$

Figure 3-7 displayed below depicts the filtering scenario for the case of an K element microphone array with the second sensor acting as the closest reference (c=2).

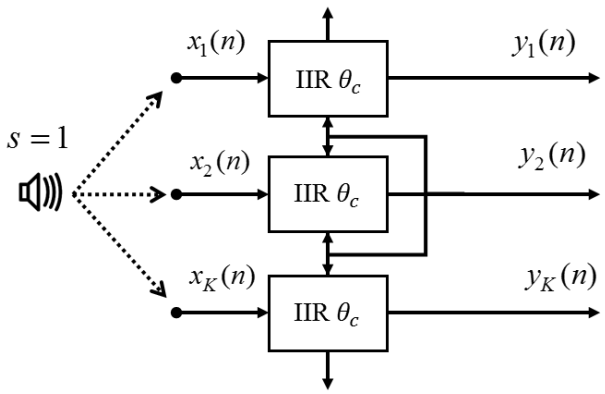

Figure 3-7: Single Input / Multiple Output filter system.

Filtering in the proposed manner offers decreased computational loads since only one adaptive operation is being performed instead of K operations which would normally be required. In addition, since all signals are processed in exactly the same manner, no phase distortions in the spatial domain (between signals) will be present. As previously discussed, this is essential to facilitate later operations such as beamforming.

3.3.2 - Multiple Input / Single Output

Now consider the case where a single sensor is subject to multiple harmonic sources as depicted below in Figure 3-8. Such as scenario constitutes a Multiple Input / Single Output (MISO) system. If S noise sources are present each having M harmonic components, all noise components may be removed by cascading S harmonic notch filters:

$$H(z)=\left[H_{1,1}(z)H_{1,2}(z)H_{2,1}(z)H_{2,2}(z)...H_{M,S}(z)\right]=\prod_{s=1}^{S}\prod_{m=1}^{M}H_{m,s}(z) \tag{3.50}$$

The cascaded transfer function will therefore be given by:

$$H(z)=\prod_{s=1}^{S}\left[\prod_{m=1}^{M}\frac{1-2\cos[m\theta_s]z^{-1}+z^{-2}}{1-2r\,\cos\left[m\theta_s\right]z^{-1}+r^2\,z^{-2}}\right] \tag{3.51}$$

The filtered output at the m^{th} harmonic for the s^{th} source will be:

$$\begin{aligned} y_{m,s}(n)=\;& y_{m-1,s}(n)-2\cos\left[m\theta_s(n)\right]y_{m-1,s}(n-1)+y_{m-1,s}(n-2)\\ &...+2r\cos\left[m\theta_s(n)\right]y_{m,s}(n-1)-r^2y_{m,s}(n-2)\end{aligned} \tag{3.52}$$

and the final harmonic output ($m=M$) of the current stage will become the initial input of the following stage according to $y_{0,s+1}(n)=y_{M,s}(n)$ for $\{s=1,2,..,S\}$ and $\{m=1,2,...M\}$. The gradient for m^{th} harmonic and s^{th} source stage is now given by:

$$\begin{aligned}\beta_{m,s}(n)&=\frac{\partial y_m(n)}{\partial\theta_s(n)}\\ &=\beta_{m-1,s}(n)-2\cos\left[m\theta_s(n)\right]\beta_{m-1,s}(n-1)+2m\sin\left[m\theta_s(n)\right]y_{m-1,s}(n-1)...+\\ &\quad\beta_{m-1,s}(n-2)+2r\cos\left[m\theta_s(n)\right]\beta_{m,s}(n-1)-r^2\beta_{m,s}(n-2)-2mr\sin\left[m\theta_s(n)\right]y_{m,s}(n-1)\end{aligned} \tag{3.53}$$

Note that $y_{1,0}(n)=x(n)$ is the initial input to the first IIR sub-filter as depicted in Figure 3-8. Notch placement locations are tracked at the output of each source stage according to:

$$\theta_s(n+1)=\theta_s(n)-2\mu_s y_{M,s}(n)\beta_{M,s}(n) \tag{3.54}$$

where μ_s is the LMS step size for the specified source filter.

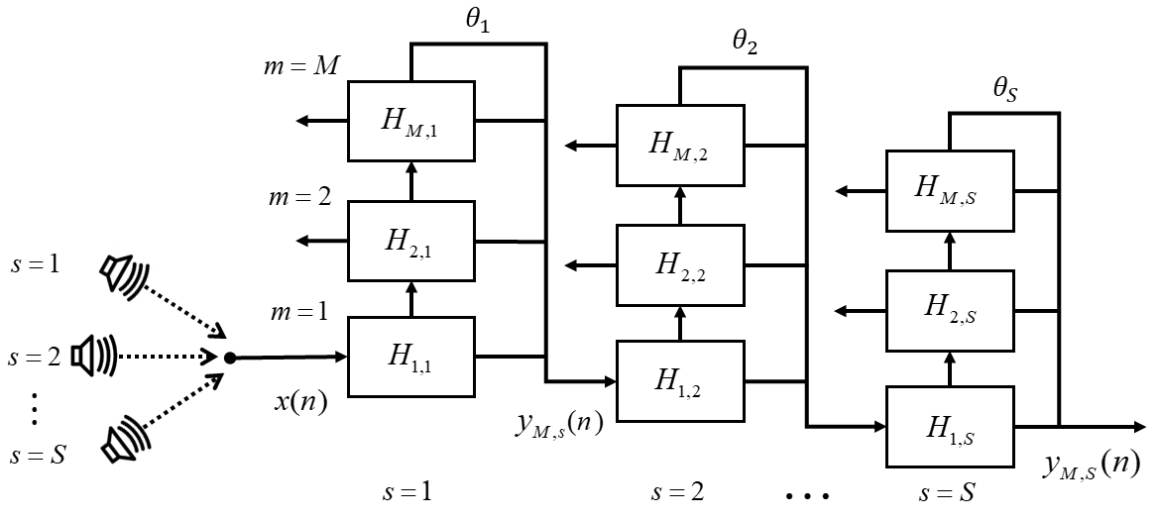

Figure 3-8: Multiple Input / Single Output filter system.

3.3.3 - Multiple Input / Multiple Output

Finally, consider the case where multiple sensors are all exposed to multiple harmonic narrowband noise sources. Such a scenario constitutes a Multiple Input / Multiple Output (MIMO) system and describes the case of using a multirotor or multi-engine fixed-wing aircraft. For multirotor aircraft, engines often run at differing frequencies as required to maneuver and maintain stability. Each noise source will therefore require a separate adaptive filter and each signal will require all filters in order to remove all noise components. Referring to the multirotor setup utilized for experiments presented in this thesis as displayed by Figure 6-8, it is apparent that each microphone will have its own primary noise source (engine that is closest). Thus, it is reasonable to assume that the most accurate frequency approximation for the source under consideration should be achieved by using the microphone closest to it. This scenario is depicted below in Figure 3-9 for the case of S sources (and corresponding stages) with M harmonic components recorded by K sensors where $m = 1,2,\dots,M$, $k = 1,2,\dots,K$, and $s = 1,2,\dots S$.

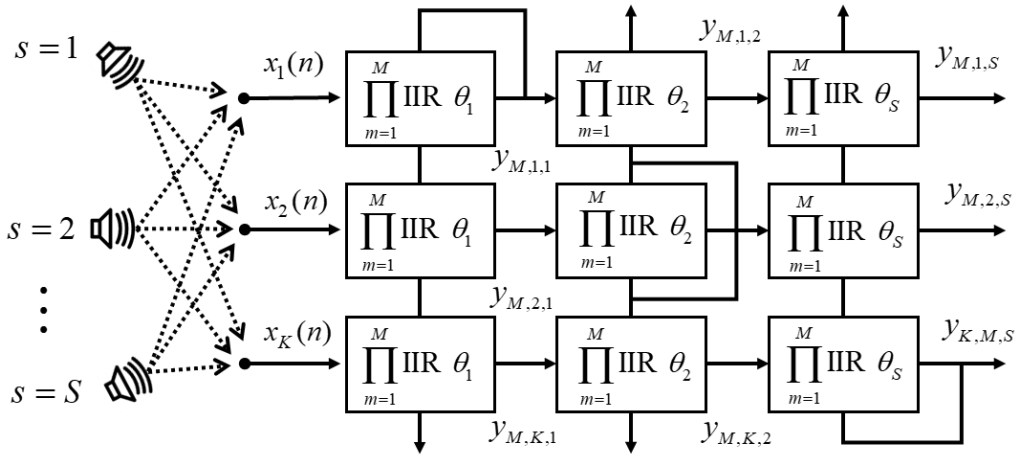

Figure 3-9: Multiple Input / Multiple Output filter system.

The transfer function for the k^{th} signal will thus be given by:

$$H_k(z) = \prod_{s=1}^{S}\left[\prod_{m=1}^{M}\frac{1-2\cos[m\theta_s]z^{-1}+z^{-2}}{1-2r\,\cos[m\theta_s]z^{-1}+r^2\,z^{-2}}\right] \tag{3.55}$$

The filtered output for the m^{th} harmonic of the k^{th} signal at the s^{th} source stage will be:

$$\begin{aligned} y_{m,k,s}(n) = {} & y_{m-1,k,s}(n) - 2\cos[m\theta_s(n)]\,y_{m-1,k,s}(n-1) + y_{m-1,k,s}(n-2) \\ & \ldots + 2r\cos[m\theta_s(n)]\,y_{m,k,s}(n-1) - r^2 y_{m,k,s}(n-2) \end{aligned} \tag{3.56}$$

Since each sensor will track the source nearest to it, only notch locations and corresponding gradients given by the sub-filter $k = s$ requires evaluation. We may facilitate a more general case by defining a source/signal designation vector $c[k]$ which identifies what source will attain the highest SNR value for each signal. For the case of the Kraken multirotor used for the presented experiments (described in Chapter 6), we would expect that each microphone will record the propeller located directly below it to a higher degree. Thus, with reference to Figure 6-9, the designation vector will be given by:

$$c[k] = [1,2,3,4,5,6] \tag{3.57}$$

However, the multirotor has eight noise-generating engines; the remaining two are located closest to microphones 6 and 1. To include these sources in the filtering process the following designation vector will be required:

$$c[k] = [1,2,3,4,5,6,6,1] \tag{3.58}$$

where $s = 1,2,\ldots,8$ and $k = 1,2,\ldots,6$. The gradient function and LMS notch locations are now given by the following recursive equations respectively:

$$\begin{aligned} \beta_{m,s}(n) = {} & \beta_{m-1,s}(n) - 2\cos[m\theta_s(n)]\,\beta_{m-1,s}(n-1)\ldots \\ & +2m\sin[m\theta_s(n)]\,y_{m-1,c[k],s}(n-1) + \beta_{m-1,s}(n-2)\ldots \\ & +2r\cos[m\theta_s(n)]\,\beta_{m,s}(n-1) - r^2\beta_{m,s}(n-2)\ldots \\ & -2mr\sin[m\theta_s(n)]\,y_{m,c[k],s}(n-1) \end{aligned} \tag{3.59}$$

$$\theta_s(n+1) = \theta_s(n) - 2\mu y_{M,c[k],s}(n)\beta_{M,s}(n) \tag{3.60}$$

As previously discussed in Section 2.3, the noise generated by an aircraft propeller may be modeled by a rotating dipole source with a fundamental frequency given by the product of the rotation rate and number of blades. However, in some circumstances physical features of the system may essentially cause the source to act as a dipole with an unequal power distribution. An example of such a case would be if a propeller blade becomes deformed, chipped, modified, etc. such that its aerodynamic properties no longer equal that of the opposing blade. For such instances, sub-harmonic and fractional harmonic components may appear. The sub-harmonic frequency will be given by the propeller rotation rate while the fractional harmonic frequency will be given by some multiple of this value. To facilitate the removal of such components, we may thus define a harmonic number vector H[ ] which is indexed according to the harmonic number m where:

$$\mathrm{H}[m] = [1\ ,\ 2\ ,...,\ M] \tag{3.61}$$

For the case of one sub-harmonic and one fractional harmonic, the harmonic number vector will become:

$$\mathrm{H}[m] = [0.5\ ,\ 1\ ,\ 1.5\ ,\ 2\ ,...,\ M] \tag{3.62}$$

where $m = 1,2,\ldots,M+2$.

Thus, the final filtering form for the most generalized case of K sensors exposed to S sources with indexed locations and harmonic numbers defined by $c[k]$ and $\mathrm{H}[m]$ respectively is given by the following equations:

$$\begin{aligned} y_{m,k,s}(n) = {} & y_{m-1,k,s}(n) - 2\cos\left[\mathrm{H}[m]\theta_s(n)\right] y_{m-1,k,s}(n-1) + y_{m-1,k,s}(n-2)... \\ & + 2r\cos\left[\mathrm{H}[m]\theta_s(n)\right] y_{m,k,s}(n-1) - r^2 y_{m,k,s}(n-2) \end{aligned} \tag{3.63}$$

$$\begin{aligned} \beta_{m,s}(n) = {} & \beta_{m-1,s}(n) - 2\cos\left[\mathrm{H}[m]\theta_s(n)\right]\beta_{m-1,s}(n-1)... \\ & +2\mathrm{H}[m]\sin\left[\mathrm{H}[m]\theta_s(n)\right] y_{m-1,c[k],s}(n-1) + \beta_{m-1,s}(n-2)... \\ & +2r\cos\left[\mathrm{H}[m]\theta_s(n)\right]\beta_{m,s}(n-1) - r^2\beta_{m,s}(n-2)... \\ & -2\mathrm{H}[m]r\sin\left[\mathrm{H}[m]\theta_s(n)\right] y_{m,c[k],s}(n-1) \end{aligned} \tag{3.64}$$

where the LMS update equation is given by (3.60) and $y_{M,k,S}$ is the final filtered output for the k^{th} sensor channel. Initial starting values for θ_s may be obtained by performing an autocorrelation on the input signals and finding the frequencies corresponding to maximum peak values.

Note that the filtering methods described for the above scenarios require all harmonics of each source component to be removed before proceeding on to the next source stage. Since the output power of the noisy signal serves as the adaptive control, failure to remove all noise elements originating from all noise sources will result in a steady-state error, which in turn will reduce the ability to achieve convergence in the notch placement location. Hence, all noise sources are simultaneously filtered by interconnecting the adaptive filters between parallel signals. To achieve a zero-phase response, the above form can be implemented in a block based manner with time reversal such as that described in [117] and previously depicted by Figure 3-3. The performance of this filtering method will be later evaluated using simulated and experimental data.

3.4 - Adaptive FIR Notch Filter

A method is now presented to construct an FIR notch filter using a second order IIR notch filter prototype. It will be shown that the initial non-linear phase response of the filter may be corrected using the zero-phase transforms previously presented. In addition, the filter may also be implemented in a referenceless adaptive form to remove harmonic narrowband noise components. Unlike other similar methods presented in the literature, Fourier approximations to the ideal impulse response can be avoided through the use of the inverse z-transform. The resulting form is therefore much simpler than that currently reported.

3.4.1 - General Description

Although FIR filters typically offer decreased performance compared to IIR filters in most all areas, they are often preferred for many practical applications due to an inherent stability and the phase response characteristics. There are essentially three well-known methods to design linear phase FIR filters: Frequency Sampling, Windowed, and Optimal designs. Frequency sampling is not typically used for notch filter designs since the desired frequency response changes drastically across the notch point leading to large interpolation errors. The Window design method offers a solution to this problem through direct evaluation of the prototype impulse response rather than sampling at discrete frequency points. However, in many instances an analytical solution for the required coefficients cannot be achieved since no direct Inverse Fourier Transform (IFT) exists. For cases in which a closed-form analytical solution is obtainable, the impulse response function may simply be truncated at some discrete value to obtain the FIR coefficients. This general procedure is known as the Impulse Response Truncation (IRT) method. Optimal designs such as the Parks-McClellan method typically offers better performance since the algorithm seeks to minimize pass and stop band errors through Chebyshev approximations. However, the design algorithm involves a numerical iteration procedure which cannot be fully expressed in analytical or adaptive form. A review of the above methods in the context of FIR notch filter design can be found in [129, 130].

Consider the generic problem of designing an FIR filter based on some desired frequency response criteria $H_d(\omega)$ known as the design prototype. For the ideal case, the desired impulse response can be obtained by taking the IFT of the prototype frequency response:

$$h(\nu) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_d(\omega) \cdot \ e^{j\omega\nu} d\omega \tag{3.65}$$

The resulting vector will be of infinite length and symmetrical about $\nu = 0$. To obtain the corresponding FIR coefficients we may simply truncate the impulse response at some desired length. However, doing so will produce ripples in the magnitude response from the well-known Gibb's effect. The common solution to this problem, known as the Window design method, acts to reduce this effect by incorporating the use of window functions. For such cases, the actual filter impulse is now given by:

$$h_w(\nu) = h(\nu) \cdot w(\nu) \tag{3.66}$$

where $w(\ )$ is the particular window function utilized (Hann, Hamming, etc.).

In many situations, a closed-form analytical solution cannot be achieved for the IFT. For such cases, coefficient values may be obtained via a numerical approximation of the ideal frequency response according to:

$$h(\nu) = \frac{1}{N} \sum_{m=-(N/2)+1}^{N/2} H_d(m)\, e^{j2\pi m\nu/N} \tag{3.67}$$

where $H_d(m)$ now consists of discrete points taken from the design prototype response.

If a large number of frequency points are chosen, the above expression may approximate the true desired impulse response to a high degree. However, the approach requires resampling and calculation of the IFT each time the desired frequency response is modified. For adaptive filtering in which the response will vary to some degree for essentially every time step, this would result in very large computational loads.

Another approach much less utilized involves direct evaluation of the filter transfer function rather than frequency response to obtain the direct implementation form. A number of methods have been presented to construct FIR notch filters using this method [130-133]. The general procedure for each involves using a second order IIR notch filter as the design prototype for which the transfer function is approximated through a power series expansion of the separated pole-zero components. Truncation of the series in the z-domain leads to algebraic expressions to calculate FIR coefficient values. However, the computation requires an iterative approach involving multiple equations, which is somewhat cumbersome and difficult to implement in digital form. Thus, it will now be shown however that a simplistic closed-form analytical solution to the problem may be achieved through the use of z-transform integrals rather than algebraic manipulations.

3.4.2 - Standard Form

Consider the frequency response function for a second order IIR notch filter given by:

$$H_d(\omega) = \frac{1 - 2\cos[\theta]e^{-j\omega} + e^{-2j\omega}}{1 - 2r\ \cos[\theta]e^{-j\omega} + r^2 e^{-2j\omega}} \tag{3.68}$$

where θ is the notch location in radians, and r defines the notch bandwidth where $0 < r < 1$. Using the standard Window design method, the ideal impulse response is thus given by:

$$h(\nu) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1 - 2\cos[\theta]e^{-j\omega} + e^{-2j\omega}}{1 - 2r\ \cos[\theta]e^{-j\omega} + r^2 e^{-2j\omega}}\ e^{j\omega\nu} d\omega \tag{3.69}$$

As in most instances, no closed form analytical solution can be achieved for the above integral. Now consider the z-domain transfer function for the desired response:

$$H_d(z) = \frac{1 - 2\cos[\theta]z^{-1} + z^{-2}}{1 - 2r\ \cos[\theta]z^{-1} + r^2 z^{-2}} \tag{3.70}$$

Similar to the IFT approach the impulse response may also be obtained through the inverse z-transform according to:

$$h(\nu) = \frac{1}{2\pi j}\oint H_d(z) z^{\nu-1} dz = \frac{1}{2\pi j}\oint \frac{1 - 2\cos[\theta]z^{-1} + z^{-2}}{1 - 2r\ \cos[\theta]z^{-1} + r^2 z^{-2}} z^{\nu-1} dz \tag{3.71}$$

Evaluating the above integral with the aid of the Mathematica software suite and performing extensive algebraic manipulation gives the following piecewise coefficient equation:

$$h(\nu)=\begin{cases}1 & \nu=0\\ (r-1)r^{\nu-2}\csc[\theta]\left(\sin\left[\theta(1-\nu)\right]+r\sin\left[\theta(\nu+1)\right]\right) & \nu\geq 1\end{cases} \tag{3.72}$$

Since the above expression is the exact solution for the IIR impulse response, the resulting FIR filter will equal the IIR form as $\nu \to \infty$. In reality however, only a finite number of coefficients values can be used to produce an approximation to the desired response. Thus, for a coefficient vector of length N the filtered output will be given by:

$$y(n)=\sum_{\nu=0}^{N-1}\begin{pmatrix}1 & \nu=0\\ (r-1)r^{\nu-2}\csc[\theta]\left(\sin\left[\theta(1-\nu)\right]+r\sin\left[\theta(\nu+1)\right]\right) & \nu\geq 1\end{pmatrix}x(n-\nu) \tag{3.73}$$

Adaptive implementation of the filter may be achieved via the LMS algorithm in a similar manner to that of the IIR form. The adaptive notch position will be given by equation (3.13) while the gradient function can be obtained through differentiation of the filter output equation $y(n)$ with respect to the notch location θ:

$$\beta(n)=\sum_{\nu=0}^{N-1}\begin{pmatrix}0 & \nu=0\\ (r-1)r^{\nu-2}\left((1+\nu)r\cos\left[\theta(n)(1+\nu)\right]+(1-\nu)\cos\left[\theta(n)(1-\nu)\right]\right)\csc\left[\theta(n)\right]... & \\ -(r-1)r^{\nu-2}\cot\left[\theta(n)\right]\csc\left[\theta(n)\right]\left(r\sin\left[\theta(n)(1+\nu)\right]+\sin\left[\theta(n)(1-\nu)\right]\right) & \nu\geq 1\end{pmatrix}x(n-\nu) \tag{3.74}$$

If the noise is harmonic in nature, additional notches can be produced by simply cascading the filter with frequency shifted versions of itself. As with the IIR filter, this may be accomplished using an iterative approach. The coefficient vector for the m^{th} harmonic is thus given by:

$$h_m(\nu)=\begin{cases}1 & \nu=0\\ (r-1)r^{\nu-2}\csc[m\theta(n)]\left(\sin\left[m\theta(n)(1-\nu)\right]+r\sin\left[m\theta(n)(\nu+1)\right]\right) & \nu\geq 1\end{cases} \tag{3.75}$$

while the filter output at the m^{th} harmonic stage is given by:

$$y_m(n)=\sum_{\nu=0}^{N-1}h_m(\nu)\,y_{m-1}(n-\nu) \tag{3.76}$$

where $y_0(n)=x(n)$ is the initial input to the first sub-filter. Taking the derivative of the above equation gives the gradient function at each stage:

$$\beta_m(n)=\frac{\partial y_m(n)}{\partial\theta(n)}=\sum_{\nu=0}^{N-1}\mathbf{A}_m\cdot y_{m-1}(n-\nu)+\sum_{\nu=0}^{N-1}\mathbf{B}_m\cdot\beta_{m-1}(n-\nu) \tag{3.77}$$

$$\mathbf{A}_m=\begin{pmatrix}0 & \nu=0\\ (r-1)r^{\nu-2}\left((1+\nu)mr\cos\left[m\theta(n)(1+\nu)\right]+(1-\nu)m\cos\left[m\theta(n)(1-\nu)\right]\right)\csc\left[m\theta(n)\right]... & \\ -m(r-1)r^{\nu-2}\cot\left[m\theta(n)\right]\csc\left[m\theta(n)\right]\left(r\sin\left[m\theta(n)(1+\nu)\right]+\sin\left[m\theta(n)(1-\nu)\right]\right) & \nu\geq 1\end{pmatrix} \tag{3.78}$$

$$\mathbf{B}_m=\begin{pmatrix}1 & \nu=0\\ (r-1)\csc\left[m\theta(n)\right]r^{\nu-2}\left(\sin\left[m\theta(n)(1-\nu)\right]+r\sin\left[m\theta(n)(\nu+1)\right]\right) & \nu\geq 1\end{pmatrix} \tag{3.79}$$

where $\beta_0 = \frac{\partial y_0}{\partial \theta} = \frac{\partial x}{\partial \theta} = 0$. The updated notch location is then given by equation (3.44) previously presented.

3.4.3 - Distortionless Form

Since the constructed filter will approximate the IIR prototype in both magnitude and phase, distortionless implementation methods are required to achieve a linear or zero-phase response. This may be accomplished via the direct convolution or the two-stage methods previously presented. Since the impulse response given by equation (3.72) is an exact solution as $\nu \to \infty$, the ideal response for the convolved zero-phase form is given by:

$$h_{zp}(\nu) = \int_{-\infty}^{\infty} h(\tau) h(\nu + \tau) \mathrm{d}\tau \tag{3.80}$$

However, a closed form expression for the above integral cannot be obtained since $h(\,)$ is a piecewise function that is discretely defined up to $\nu = 1$ and continuously defined thereafter. Using a numerical approach instead we obtain:

$$h_{zp}(\nu) = \sum_{m=0}^{N-\nu-1} h(m) h(m + \nu) \tag{3.81}$$

The distortionless filter process is then given by:

$$y(n) = h_{zp}(0) x(n) + \mathbf{h}_{zp}^T \left[\mathbf{X_c}(n) + \mathbf{X_a}(n) \right] \tag{3.82}$$

$$\mathbf{h}_{zp} = \left[h_{zp}(1)\ h_{zp}(2) \ldots h_{zp}(N-1) \right]^T$$

$$\mathbf{X_c}(n) = \left[x(n-1)\ x(n-2) \ldots x(n-N+1) \right]^T$$

$$\mathbf{X_a}(n) = \left[x(n+1)\ x(n+2) \ldots x(n+N-1) \right]^T$$

Unfortunately, the filter cannot be implemented in a referenceless adaptive form since an analytical expression for the zero-phase output with respect to θ is required to calculate the LMS gradient function. We may however use the causal two-stage approach previously presented. Application of the method yields the following final output form:

$$z(n) = \sum_{\nu=0}^{N-1} h(N-\nu-1) \sum_{\nu=0}^{N-1} h(\nu)\, x(n-\nu) \tag{3.83}$$

where the coefficient vector $h(\,)$ is defined by equation (3.72). To simplify the gradient function required for LMS implementation, the non-linear output at the first filter stage may be utilized instead of differentiating equation (3.83) above. This is illustrated in Figure 3-10 where $\bar{H}$ indicates the reflected coefficient filter. Adaptive implementation may then be achieved via the use of equations (3.13), (3.73), and (3.74), with the final linear-phase output given by:

$$z(n)=\sum_{\nu=0}^{N-1}h(N-\nu-1)y(n) \tag{3.84}$$

For the case of harmonic noise requiring multiple notches, adaptive implementation may be achieved via equations (3.44), (3.75), (3.76), and (3.77), with the final linear-phase output given by:

$$z_m(n)=\sum_{\nu=0}^{N-1}h_m(N-\nu-1)\,z_{m-1}(n-\nu) \tag{3.85}$$

where $z_{m-1}=y_M$ as depicted in Figure 3-10.

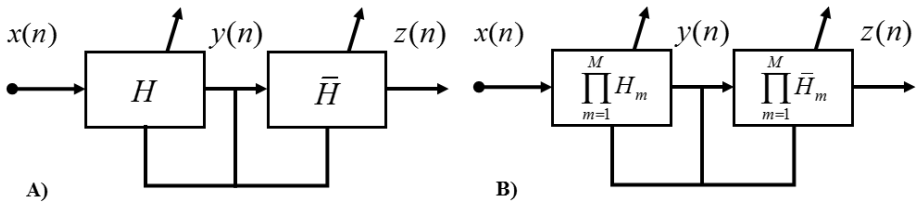

Figure 3-10: A) Adaptive two-stage FIR notch filter. B) Adaptive two-stage harmonic FIR notch filter.

3.4.4 - Multichannel Filtering

Using the procedure presented in the previous section for multichannel IIR notch filtering we may also establish a similar form for the multichannel FIR case. For the case pertaining to fixed-wing aircraft, we are presented with a SIMO system with one harmonic source and K sensor signals as previously described and depicted by Figure 3-7. Again, since all signals are exposed to the same noise source, only one signal is required to track the fundamental noise frequency which is defined by the position $k=c$. Thus, the coefficient equation, filtered output, gradient, and LMS update for the m^{th} harmonic of the k^{th} signal can be expressed by the following equations respectively:

$$h_m(\nu)=\begin{cases}1 & \nu=0\\ (r-1)r^{\nu-2}\csc\left[m\theta(n)\right]\left(\sin\left[m\,\theta(n)(1-\nu)\right]+r\sin\left[m\theta(n)(\nu+1)\right]\right) & \nu\geq 1\end{cases} \tag{3.86}$$

$$y_{m,k}(n)=\sum_{\nu=0}^{N-1}h_m(\nu)\,y_{m-1,k}(n-\nu) \tag{3.87}$$

$$\beta_m(n)=\frac{\partial y_{m,k=c}(n)}{\partial\theta(n)}=\sum_{\nu=0}^{N-1}\mathbf{A}_m\cdot y_{m-1,k=c}(n-\nu)+\sum_{\nu=0}^{N-1}\mathbf{B}_m\cdot\beta_{m-1}(n-\nu) \tag{3.88}$$

$$\mathbf{A}_m=\begin{pmatrix}0 & \nu=0\\ (r-1)r^{\nu-2}\left((1+\nu)mr\cos\left[m\theta(n)(1+\nu)\right]+(1-\nu)m\cos\left[m\,\theta(n)(1-\nu)\right]\right)\csc\left[m\,\theta(n)\right]... & \\ -m(r-1)r^{\nu-2}\cot\left[m\theta(n)\right]\csc\left[m\theta(n)\right]\left(r\,\sin\left[m\theta(n)(1+\nu)\right]+\sin\left[m\theta(n)(1-\nu)\right]\right) & \nu\geq 1\end{pmatrix} \tag{3.89}$$

$$\mathbf{B}_m=\begin{pmatrix}1 & \nu=0\\ (r-1)\csc\left[m\theta(n)\right]r^{\nu-2}\left(\sin\left[m\theta(n)(1-\nu)\right]+r\sin\left[m\theta(n)(\nu+1)\right]\right) & \nu\geq 1\end{pmatrix} \tag{3.90}$$

where $m = 1,2,\ldots,M$, $k = 1,2,\ldots,K$, and the LMS update equation is given by (3.49). The final linear-phase output will then be given by:

$$z_{m,k}(n) = \sum_{\nu=0}^{N-1} h_m(N-\nu-1)\, z_{m-1,k}(n-\nu) \tag{3.91}$$

where $z_{m-1,k} = y_{M,k}$

For the case pertaining to multirotor aircraft, we are presented with a MIMO system as previously described and depicted by Figure 3-9. As before, we may simply extend the concepts presented for IIR filters to obtain an expression for the FIR case. Thus, for the most generalized case consisting of *K* sensors exposed to *S* sources with indexed locations and harmonic numbers defined by $c[k]$ and $\mathrm{H}[m]$ respectively, the coefficient equation, filtered output, and gradient function will be given by the following equations:

$$h_{m,s}(\nu) = \begin{cases} 1 & \nu = 0 \\ (r-1)\, r^{\nu-2} \csc\left[\mathrm{H}[m]\theta_s(n)\right]\left(\sin\left[\mathrm{H}[m]\theta_s(n)(1-\nu)\right] + r\sin\left[\mathrm{H}[m]\theta_s(n)(\nu+1)\right]\right) & \nu \geq 1 \end{cases} \tag{3.92}$$

$$y_{m,k,s}(n) = \sum_{\nu=0}^{N-1} h_{m,s}(\nu)\, y_{m-1,k,s}(n-\nu) \tag{3.93}$$

$$\beta_{m,s}(n) = \sum_{\nu=0}^{N-1} \mathbf{A}_{m,s} \cdot y_{m-1,c[k],s}(n-\nu) + \sum_{\nu=0}^{N-1} \mathbf{B}_{m,s} \cdot \beta_{m-1,s}(n-\nu) \tag{3.94}$$

$$\mathbf{A}_{m,s} = \begin{pmatrix} 0 & \nu = 0 \\ (r-1)\, r^{\nu-2}\left((1+\nu)\mathrm{H}[m] r\cos\left[\mathrm{H}[m]\theta_s(n)(1+\nu)\right] + (1-\nu)\,\mathrm{H}[m]\cos\left[\mathrm{H}[m]\,\theta_s(n)(1-\nu)\right]\right)\csc\left[\mathrm{H}[m]\theta_s(n)\right]\ldots \\ -\mathrm{H}[m](r-1)\, r^{\nu-2}\cot\left[\mathrm{H}[m]\theta_s(n)\right]\csc\left[\mathrm{H}[m]\theta_s(n)\right]\left(r\,\sin\left[\mathrm{H}[m]\theta_s(n)(1+\nu)\right] + \sin\left[\mathrm{H}[m]\theta_s(n)(1-\nu)\right]\right) & \nu \geq 1 \end{pmatrix} \tag{3.95}$$

$$\mathbf{B}_{m,s} = \begin{pmatrix} 1 & \nu = 0 \\ (r-1)\csc\left[\mathrm{H}[m]\,\theta_s(n)\right] r^{\nu-2}\left(\sin\left[\mathrm{H}[m]\,\theta_s(n)(1-\nu)\right] + r\sin\left[\mathrm{H}[m]\,\theta_s(n)(\nu+1)\right]\right) & \nu \geq 1 \end{pmatrix} \tag{3.96}$$

where $c[k]$ and $H[m]$ were previously defined by equations (3.58) and (3.62), and the LMS update equation by (3.60). The final linear-phase output will then be given by:

$$z_{m,k,s}(n) = \sum_{\nu=0}^{N-1} h_{m,s}(N-\nu-1)\, z_{m-1,k,s}(n-\nu) \tag{3.97}$$

where $z_{m-1,k,s} = y_{M,k,s}$ and $z_{M,k,S}$ is the final filtered output for the k^{th} sensor channel.

3.4.5 - Performance Considerations

Figures 3-11 through 3-13 displayed below provides plots comparing the IIR notch filter response to that of the nonlinear and linear-phase FIR approximations. In addition, the effect of varying the pole radius, filter length, and use of window functions is also illustrated. From the plots, it is evident that the truncated FIR filter provides

a very good response approximation to the IIR filter, with no visual difference between the two. It is also apparent that the linearized FIR filter does in fact produce a linear-phase response. However, as predicted by equation (3.28), it is produced at the expense of squaring the magnitude response of the original non-linear form; a property that generates both favourable and unfavourable features. Undesired aspects involve increased notch bandwidth and passband ripple, while positive features involve increased notch depth. For the non-linear FIR form, bandwidth values will be given by equation (3.37) since the filter will directly approximate the IIR response. For the linear-phase form, notch bandwidth values may be approximated according to:

$$BW = \frac{\sqrt{3} f_s}{\pi}(1-r) \tag{3.98}$$

which is valid for $0.9 < r < 1.0$. The above expression was obtained following the general procedure presented [128] which was also utilized to produce the form previously given by equation (3.37).

It is also evident from the plots presented in Figures 3-12 and 3-13 that decreasing the filter length and/or increasing the notch pole radius both result in decreased approximation accuracy. The general effect will be an increase in passband ripple and a decrease in notch depth. Typically for FIR designs, window functions are applied to the impulse response to decrease passband ripples. For this particular instance however, doing so results in greatly decreased response characteristics since the notch becomes very shallow and wide. This effect can be observed in Figure 3-14 when using the Hann window.

To establish a quantitative performance comparison between the IIR and FIR forms, we may define a response approximation error according to the expression given below:

$$\varepsilon(\omega) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left| H_d(\omega) - H(\omega) \right|^2 d\omega \tag{3.99}$$

Using this definition, a plot was generated using a range of values for the filter length and pole radius, which is displayed in Figure 3-14. From the plot, it is again evident that decreasing the filter length and/or increasing the notch pole radius results in decreased approximation accuracy.

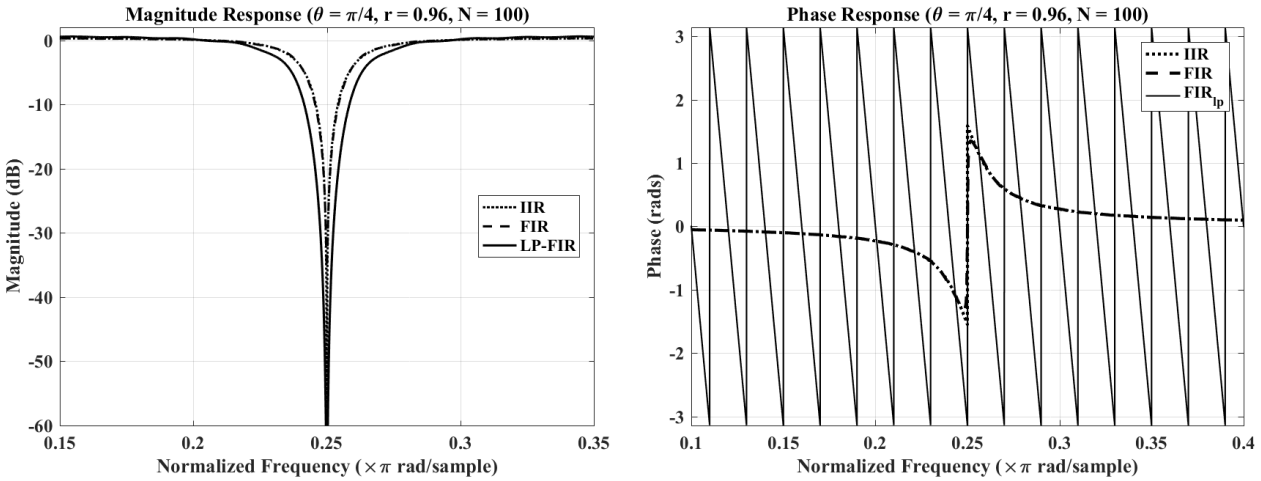

Figure 3-11: Magnitude and phase response of IIR and FIR notch filters.

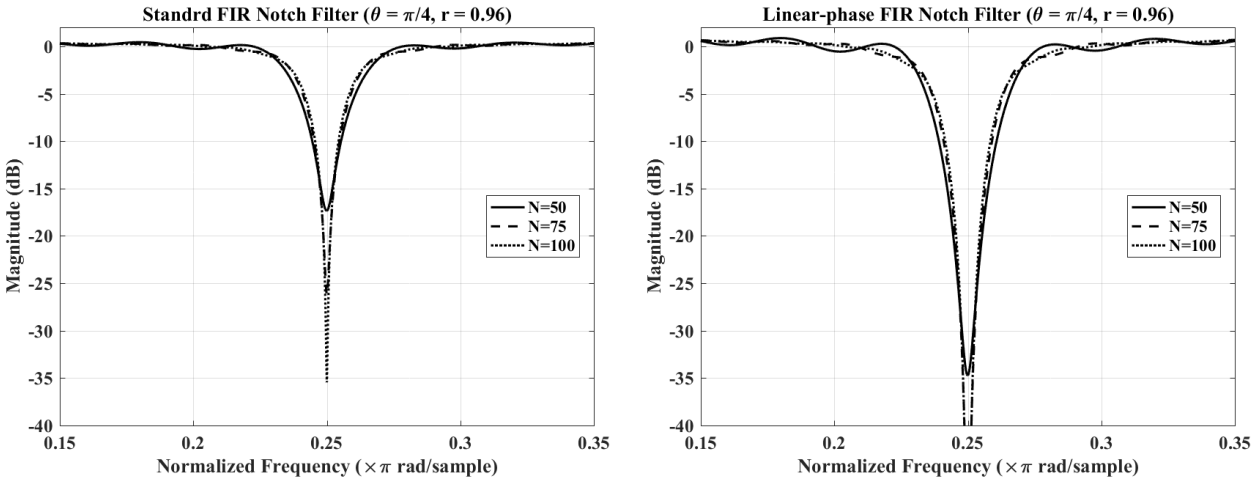

Figure 3-12: Magnitude response of FIR notch filters for varying filter length.

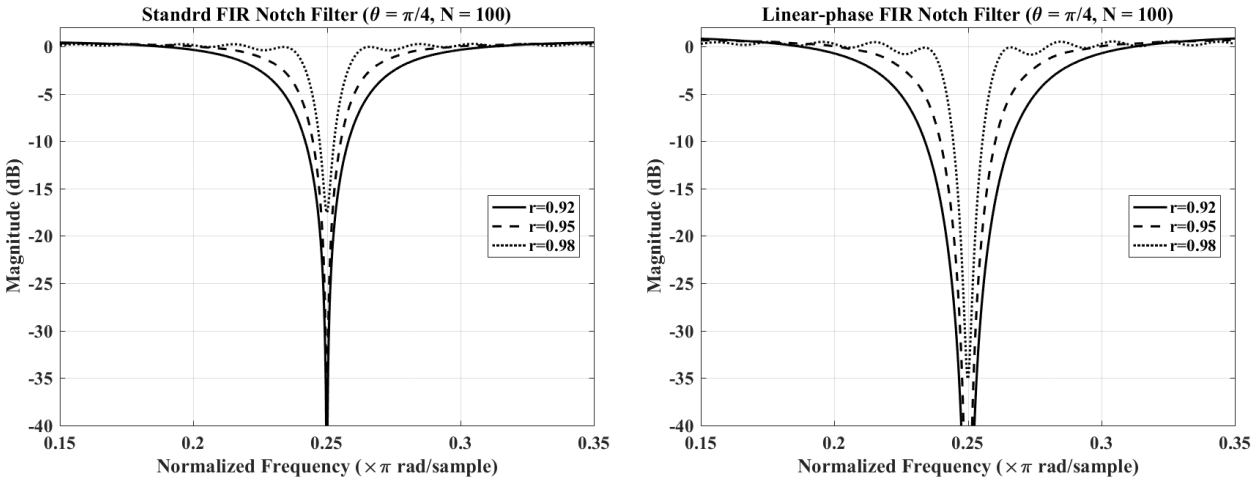

Figure 3-13: Magnitude response of FIR notch filters for varying pole radius.

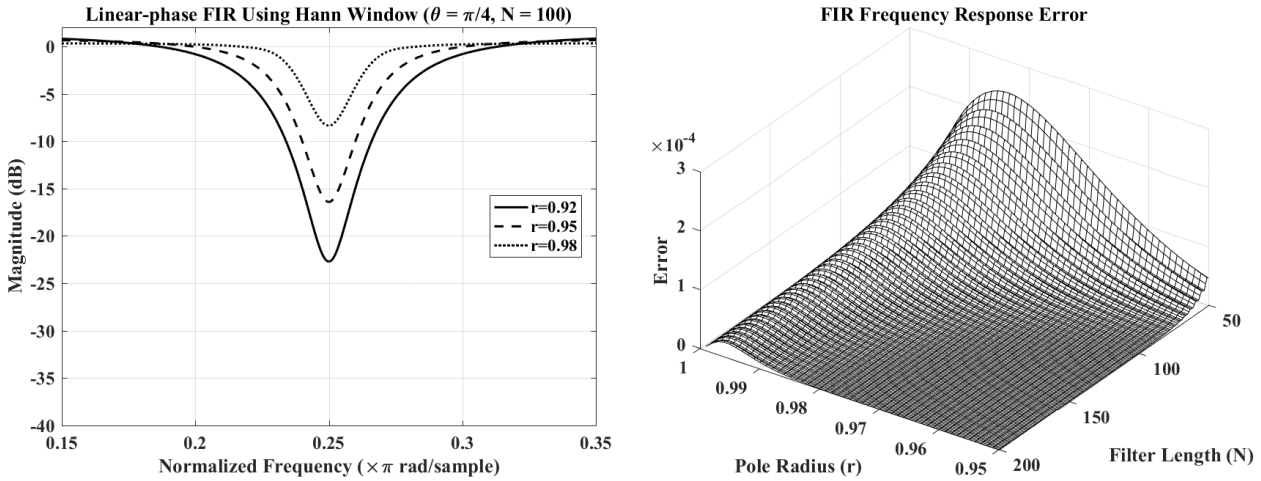

Figure 3-14: Left) Effect of applying window function to FIR impulse response. Right) Error between IIR and FIR frequency response.

It should be noted that although the filter is fully capable of facilitating MIMO systems such as that pertaining to multirotor aircraft, in some instances it might not be practical due to computational requirements. In general, the multi-stage iterative method previously presented may be replaced by a single stage filter, which is the result of convolving all stages together. Thus, the length of the final equivalent (linearized) filter required for a signal containing S noise sources each with M harmonic components is given by:

$$\bar{N} = 2SMN + 1 \tag{3.100}$$

where N is the fundamental filter length used at each stage. From the above expression, it is apparent that increasing the number of harmonics and noise sources produces an exponential increase in the final equivalent filter length. Consider for example the multirotor aircraft utilized for experiments presented in this thesis, which consists of $K = 6$ signals exposed to $S = 8$ sources. Typically, each aircraft engine operates at approximately 125 Hz. For a 1 kHz frequency band of interest, a total of $M = 8$ harmonics would therefore require removal. This would be equivalent to a single stage filter with $\bar{N} = 8 \cdot 1600 + 1 = 12{,}801$ coefficients which is extremely large. In addition, each of the $K = 6$ signals would require parallel processing using this filter. Thus, even when considering the processing power of modern day computers, it is highly unlikely that such a filter could be practically implemented in real-time on a dedicated system small enough to be located on-board most UAVs. However, for fixed-wing operations in which only one source is present, real-time implementation may easily be achieved.

3.5 - Adaptive FIR Comb Filter

A method is now presented to construct and implement an FIR Comb filter in a referenceless adaptive form. In addition, it will be shown that the filter may also be modified to operate in a phase distortionless form using the convolved zero-phase transform previously presented.

3.5.1 - General Description

Consider a standard *N*-delay Comb filter with shifting gain G as displayed in Figure 3-15, where the transfer function, magnitude response, and direct implementation form are given by the following equations respectively:

$$H(z) = G\left(1 - z^{-N}\right) \tag{3.101}$$

$$H(\omega) = G\left(1 - e^{-j\omega}\right) = G\left(1 - \cos(\omega) + j\sin(\omega)\right) \tag{3.102}$$

$$y(n) = G\left(x(n) - x(n-N)\right) \tag{3.103}$$

where N is the number of delay elements required to remove a narrowband signal of fundamental frequency f_o according to:

$$N = \frac{f_s}{f_o} \tag{3.104}$$

Note that the total filter length will now be $N+1$ coefficents.

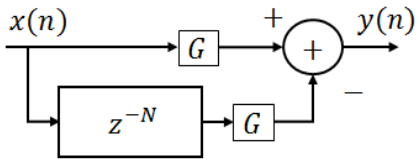

Figure 3-15: N-delay FIR Comb filter structure with shifting gain G.

The magnitude response for the filter may be obtained by taking the norm (magnitude) of the above frequency response equation:

$$\left|H(f)\right| = G\sqrt{2-2\cos\left[\frac{2\pi Nf}{f_s}\right]} \tag{3.105}$$

where the angular frequency ω in radians has been replaced by the more applicable form f given in Hertz.

The filter provides the benefit of being very simplistic and requires little computational load since it consists of only two multiplications and one addition. It will also produce a linear phase response since it satisfies the FIR anti-symmetric impulse response condition previously discussed. In addition, no cascaded sub-filters are required to remove harmonic-based noise since the filter will have periodic notches spanning the entire signal bandwidth located at multiples of f_s / N Hz. This feature can be observed from the magnitude response plot displayed in Figure 3-16.

3.5.2 - Distortionless Form

Unlike typical FIR filters which establish stop band regions through modifying the values of a constant length coefficient vector, the Comb filter is unique in that coefficient values remain constant while the vector length changes according to the desired stopband location. Thus, adaptive implementation will effectively produce a non-linear phase response since the linear-phase slope defined by the current delay number N will continuously vary with time. However, we may solve this problem by simply modifying the filter using the convolved zero-phase modification technique previously presented. Application of the transform gives the following equations for the transfer function, frequency response, magnitude response, and direct implementation form respectively:

$$\begin{aligned} H_o(z) &= G^2\left(1-z^{-N}\right)\left(1-z^{+N}\right) \\ &= G^2\left(2-z^{-N}-z^{+N}\right) \end{aligned} \tag{3.106}$$

$$\begin{aligned} H_o(\omega) &= G^2\left(2-e^{-j\omega}+e^{+j\omega}\right) \\ &= -2G^2\left(\cos[\omega]-1\right) \end{aligned} \tag{3.107}$$

$$\left|H_o(f)\right| = 2G^2\left(1-\cos\left[\frac{2\pi Nf}{f_s}\right]\right) \tag{3.108}$$

$$y(n) = G^2\left[2x(n)-x(n-N)-x(n+N)\right] \tag{3.109}$$

where the total filter length is now $2N+1$ since two filters of length $N+1$ are essentially convolved. Thus, the number of delay elements required to remove a signal of fundamental frequency f_o will now be given by:

$$N = \frac{2f_s}{f_o} \tag{3.110}$$

It is apparent from the frequency response given above that only real components are present, and values are always positive. Thus, the filter will always maintain a true zero-phase response as per the condition given by equation (3.20). This is in contrast to the standard form previously given by equation (3.102) which always has a real component less than zero except at the notch locations. Therefore, even though this filter is symmetric shifting it to a non-causal form would actually produce a π-phase filter where values alternate between ± π/2 at each notch location. Note that for a shifting gain value of $G = 1$ the standard and zero-phase filters attain a maximum magnitude of 6 and 12 dB respectively, effectively boosting any signals residing between the filter nulls. This gain can be further increased or removed by simply adjusting the gain factor. For instance, shifting the maximum magnitude down by -6 dB to 0 dB would require $G = 10^{-6/20} = 0.5012$. Figure 3-16 displayed below provides magnitude and phase response plots for the standard, zero-phase and π-phase filters.

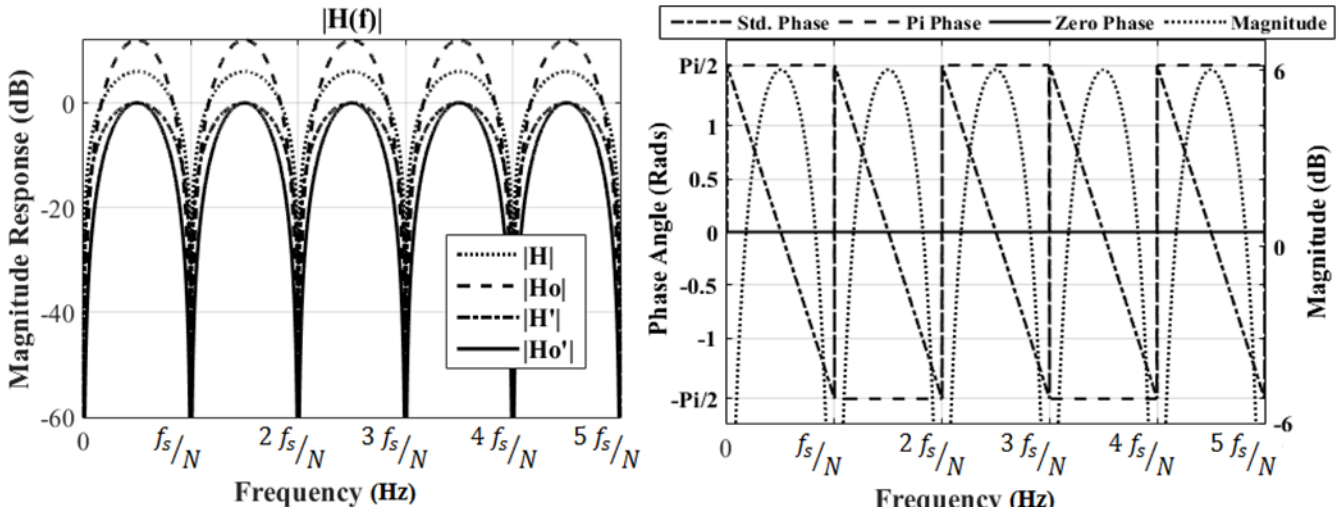

Figure 3-16: Left) Magnitude response for N = 10, where $|H|$ and $|H_o|$ are the standard and zero-phase filters with unity gain (G=1); $|H'|$ and $|H_o'|$ are the 0 dB shifted versions of these filters. Right) Phase and Magnitude response for a standard, π-phase transformed, and zero-phase transformed FIR Comb filter.

3.5.3 - Adaptive Implementation

Referenceless adaptive implementation of the Comb filter may be achieved using the approach previously described in Section 3.2.1.2 via the LMS update equation given by (3.13). Note that the delay number N may be expressed in terms of the normalized frequency θ according to:

$$N = \frac{2}{\theta} \tag{3.111}$$

To obtain the gradient function $\beta(n)=\frac{\partial y(n)}{\partial \theta(n)}$ we must first express the direct output form in terms of θ. Substitution of (3.111) into (3.109) and expressing in terms of the discrete time index n gives:

$$y(n)=G^2\left[2x(n)-x\left(n-\frac{2}{\theta(n)}\right)-x\left(n+\frac{2}{\theta(n)}\right)\right] \tag{3.112}$$

Differentiating the above equation with respect to the normalized notch location gives:

$$\beta(n)=\frac{2G^2}{\theta(n)^2}\left[x'\left(n+\frac{2}{\theta(n)}\right)-x'\left(n-\frac{2}{\theta(n)}\right)\right] \tag{3.113}$$

Since no analytical expression for the signal x exists, we must approximate the derivative $x'(\,)$ via numerical methods. This may be accomplished by taking the backwards finite difference, which produces the following approximation:

$$\beta(n)\approx\frac{2G^2}{\theta(n)^2}\left[x\left(n+\frac{2}{\theta(n)}\right)-x\left(n+\frac{2}{\theta(n-1)}-1\right)+x\left(n-\frac{2}{\theta(n)}\right)-x\left(n-\frac{2}{\theta(n-1)}-1\right)\right] \tag{3.114}$$

Thus adaptive implementation may now be achieved via the use of equations (3.13), (3.112), and (3.113). It should be noted that since N is a positive integer, implementation via standard methods would produce notch location errors due to rounding. This offset is also directly compounded for each harmonic component and increases as the signal sampling rate decreases. Figure 3-17 displayed below depicts the notch placement error for a harmonic signal with a fundamental frequency of 150 Hz. It is clear from the plot that low sampling rates produce substantial errors that is compounded for each harmonic component.

For situations in which higher sampling rates cannot be achieved to meet specified notch position error criteria, interpolation may be used to obtain fractional delay values. For the simplest case which utilizes linear interpolation, a fractional delay of $x(n\pm\theta)$ can be obtained for non-integer values of θ according to:

$$x(n\pm\theta)=x(n\pm\lfloor\theta\rfloor)+\left[x(n\pm\lceil\theta\rceil)-x(n\pm\lfloor\theta\rfloor)\right]\frac{(\theta-\lfloor\theta\rfloor)}{(\lceil\theta\rceil-\lfloor\theta\rfloor)} \tag{3.115}$$

where $\lfloor\ \rfloor$ and $\lceil\ \rceil$ indicates the floor and ceiling functions respectively.

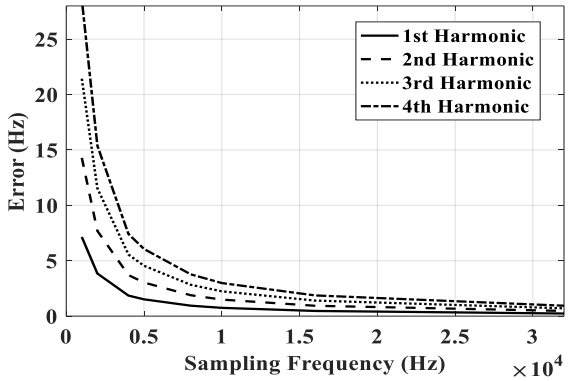

Figure 3-17: Notch location error due to rounding.

3.5.4 - Multichannel Filtering

For the case pertaining to fixed-wing aircraft in which K sensor signals are exposed to one harmonic source ($S = 1$), only one signal is required to track the fundamental noise frequency as previously discussed. Thus, defining the sensor closest to the noise source by $k = c$, the SIMO filtering case can be described according to the following equations:

$$y_k(n) = G^2\left[2x_k(n) - x_k\left(n - \frac{2}{\theta(n)}\right) - x_k\left(n + \frac{2}{\theta(n)}\right)\right] \tag{3.116}$$

$$\beta(n) \approx \frac{2G^2}{\theta(n)^2}\left[x_{k=c}\left(n + \frac{2}{\theta(n)}\right) - x_{k=c}\left(n + \frac{2}{\theta(n-1)} - 1\right) + x_{k=c}\left(n - \frac{2}{\theta(n)}\right) - x_{k=c}\left(n - \frac{2}{\theta(n-1)} - 1\right)\right] \tag{3.117}$$

where the LMS update equation is again given by (3.13).

For the MIMO case pertaining to multirotor aircraft in which K sensor signals are exposed to S sources, the filter may not be as appropriate. Similar to the FIR case previously discussed, multiple Comb filters may be placed in cascade form to remove multiple narrowband sources. However, since the filter requires strict non-causal operation to maintain a distortionless phase response for changing delay values, it cannot be implemented in an iterative form such as that used for the IIR and FIR notch filters. All stages must instead be convolved to obtain a final output expression. Although this may be achieved quite easily due to the simplicity of the governing filter equations, large data buffer requirements associated with the final output form may produce unacceptably large processing lags.

In order to process the filter in real-time, the signal is passed through a storage buffer of size $2N + 1$ such that the center of buffer serves as the zero-time reference point, with signal values being chosen in the forward and reverse directions at a distance of N samples. The real-time operation of the filter would then be subject to a minimum processing delay or latency of N/f_s seconds as required to introduce the anti-causal portion of the filter. Since the adaptive process requires changing the number of delay elements to track frequency components, the buffer must be chosen big enough such that it will be capable of handling the largest expected delay value. For example, consider the case of a signal subject to two narrowband noise sources. The transfer function for the zero-phase cascaded filter output will be given by:

$$\begin{aligned} H_o(z) &= G^2\left(2 - z^{-N_1} - z^{+N_1}\right) \cdot G^2\left(2 - z^{-N_2} - z^{+N_2}\right) \\ &= G^4\left(4 - 2z^{-N_1} - 2z^{N_1} + z^{-N_1-N_2} + z^{N_1-N_2} - 2z^{-N_2} - 2z^{N_2} + z^{-N_1+N_2} + z^{N_1+N_2}\right) \end{aligned} \tag{3.118}$$

where N_1 and N_2 are the number of delay elements required to track the first and second source respectively. From the above equation, it is evident that the data buffer must now be of size $2(N_1 + N_2) + 1$. In order to enable the anti-casual processing, a delay of $(N_1 + N_2)/f_s$ must therefore be present. This effect is compounded for each additional cascaded form such that the total delay required for K number of sources will be given by:

$$t_{delay} = (N_1 + N_2 + ... + N_K) / f_s \tag{3.119}$$

Thus, for scenarios in which a large number of sources are present such that found with the Kraken multirotor described in this thesis (8 sources), adaptive Comb filtering is not considered appropriate when considering detection distance requirements (discussed in the next chapter). However, for the case pertaining to fixed-wing aircraft which involve only one noise source, the filter performs reasonably well as will later be shown.

3.5.5 - Performance Considerations

Although the Comb filter is very simplistic and requires few computations, it does have a number of drawbacks related to notch bandwidth, notch location accuracy, and implementation latency as previously discussed. Perhaps the greatest of these is the notch bandwidth, which cannot be varied for a given notch frequency location. Analysis of the magnitude response previously given leads to the following -3 dB bandwidth approximation for the standard and zero-phase filters respectively:

$$BW = \frac{f_o}{\pi} \cos^{-1}\left(1 - \frac{0.251}{G^2}\right) \tag{3.120}$$

$$BW_o = \frac{f_o}{\pi} \cos^{-1}\left(1 - \frac{0.354}{G^2}\right) \tag{3.121}$$

From the above equations, it is clear that the bandwidth for each filter is directly proportional the fundamental frequency and completely independent of the sampling frequency. Thus, larger notch bandwidths will be obtained for higher fundamental frequency signals, which will ultimately deteriorate filtering performance. If frequency values are large enough, this aspect may actually render the filter unusable since large portions of the passband will become greatly attenuated. Figure 3-18 displayed below provides a comparison of the notch bandwidth for various fundamental frequencies. From the plot, it is evident that the filter is indeed ill suited for high frequency applications.

In addition to the operating latency previously discussed, another possible issue that may occur when using the filter in an adaptive form is the false convergence to a local minima rather than the global minimum. Figure 3-19 shows the Mean Squared Error (MSE) as a function of frequency and number of harmonics present in a sinusoidal signal with a fundamental frequency of 150 Hz. It is evident that as the number of harmonics increases, the number of local minima also increases, and the width of the frequency capture region decreases. This is because the Comb filter contains equally spaced notches that span the entire signal (up to the Nyquist frequency). A misplacement of the fundamental frequency notch may still produce a local minimum if one of the subsequent notches is located at an harmonic frequency. For example, a ten-harmonic signal (NH=10) with a fundamental frequency of 150 Hz has a local minimum at 133.3 Hz since the 9th Comb notch will be located at 1200 Hz, which also corresponds to the 8th signal harmonic. Thus, for signals containing large numbers of harmonics, it is essential that an accurate frequency starting point be chosen for the adaptive process. This may

be done through performing an autocorrelation or analyzing an initial signal segment in the frequency domain as previously discussed.

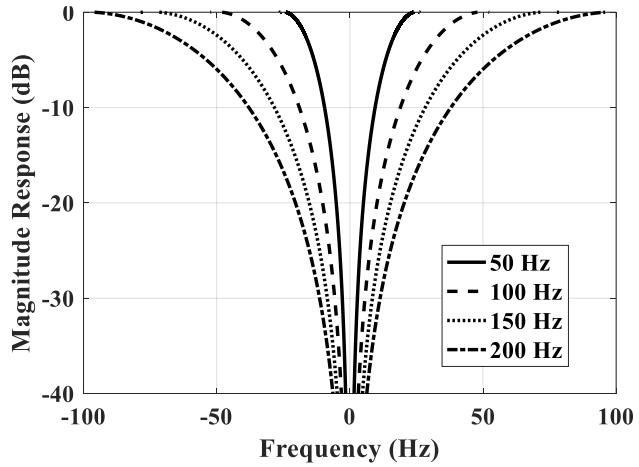

Figure 3-18: Notch bandwidth as a function of fundamental frequency.

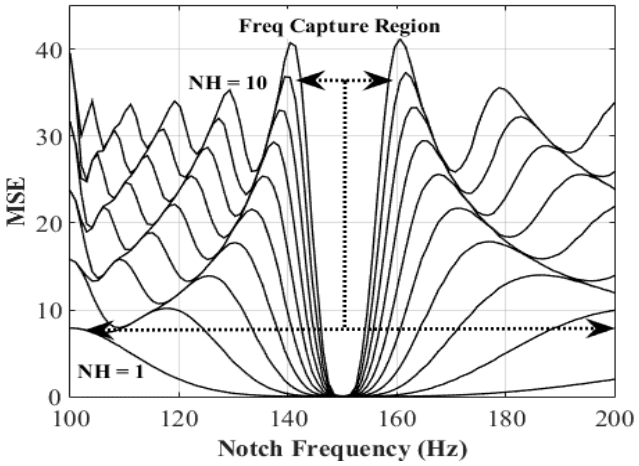

Figure 3-19: Frequency capture region as a function of the number of harmonic signal components (NH) present.

3.6 - Simulated Studies

To illustrate the performance of the presented methods, the filters are first implemented under ideal conditions using computer-generated signals. Two scenarios are investigated as it pertains to the fixed-wing and multi-rotor operations. A description of the simulation setup along with the results obtained for each filtering method is provided below.

3.6.1 - Simulation Description

To simulate the propeller generated noise produced by each aircraft, multiple non-stationary sinusoidal functions were combined with random Gaussian noise and attenuated according to a predefined source/sensor configuration geometry. The signal produced by the s^{th} source is defined as:

$$s_s(n) = \sum_{m=1}^{M} A_m \cos\left[m\phi_s(n)\right] \tag{3.122}$$

for $\{s = 1,2, \dots, S\}$ where A_m is the amplitude of the m^{th} harmonic component given by:

$$A_m = \frac{A_o}{\sqrt{m}} \tag{3.123}$$

with $A_o = 1$, and $\phi_s(n)$ is the phase function given by the cumulative sum of all past frequency values according to:

$$\phi_s(n) = \frac{2\pi}{f_s}\sum_{i=0}^{n} F_s(i) = \phi_s(n-1) + \frac{2\pi}{f_s}F_s(n) \tag{3.124}$$

where f_s is the sampling rate, and F_s is the frequency of the s^{th} source signal. The above amplitude attenuation function was utilized since it provides a good approximation to the harmonic attenuation properties of fixed-wing propeller driven aircraft [72].

To produce non-stationary signals, a modulating function was used to vary the frequency of each source with time. Thus, the time variant fundamental frequency is given by:

$$F_s(n) = B_s \cos[2\pi n F_{m,s} / f_s] + F_{o,s} \tag{3.125}$$

where B_s is the modulation amplitude, $F_{m,s}$ is the modulation frequency, and $F_{o,s}$ is the base or center fundamental frequency for each source signal. Finally, the source signal acquired by the k^{th} sensor is given by:

$$x_k(n) = \sum_{s=1}^{S} \alpha_{k,s}\, s_s(n) + w_k(n) \tag{3.126}$$

for $\{k = 1,2,\dots,K\}$ where $w(\,)$ is the random Gaussian noise component, and $\alpha_{k,s}$ is an attenuation factor to account for source/sensor spacing effects. Using the distance/SPL law for acoustic transmission previously given by equation (2.2) (converted to magnitude form), the amplitude attenuation factor for transmission between the s^{th} source and k^{th} sensor is given by:

$$\alpha_{k,s} = \frac{1}{|\vec{r}_s - \vec{r}_k|} \tag{3.127}$$

Filter performance is evaluated with regard to frequency tracking accuracy and the output MSE. It should be noted however, that since the error function is defined as the filtered output, MSE values for filters having different notch bandwidths cannot be directly compared. That is, a filter having a larger notch bandwidth will always attenuate more of the total signal bandwidth and thus produce lower MSE values. The frequency tracking performance of each filter is evaluated in terms of the Mean Frequency Deviation (MFD) which is defined as:

$$MFD = \frac{1}{L} \sum_{n=1}^{L} \left| \hat{F}_s(n) - F_s(n) \right| \tag{3.128}$$

where $\hat{F}_s$ is the fundamental frequency track for the filter of concern.

To simulate the SIMO system associated with a fixed-wing configuration, a total of two sensor signals $(K = 2)$ were constructed from a single source component $(S = 1)$. The sensors were located at equal distances from the source such as that $\alpha_{1,1} = \alpha_{2,1}$ depicted below in Figure 3-20. The source consisted of an $F_{0,1} = 150$ Hz signal with $M = 6$ harmonic components, modulated using amplitude and frequency values $B_1 = 5$ and $F_{m,1} = 0.2$ Hz respectively. The signals were constructed using a sampling rate of $f_s = 4$ kHz and combined with Gaussian noise of unity variance $(\sigma = 1)$ for a total duration of 10 seconds. Note that modulation values were chosen only to produce a smoothly transitioning nonstationary signal and do not reflect true modulation values associated with an aircraft engine. Figure 3-21 displayed below provides a spectrogram of the generated signal and a plot of the time variant fundamental frequency.

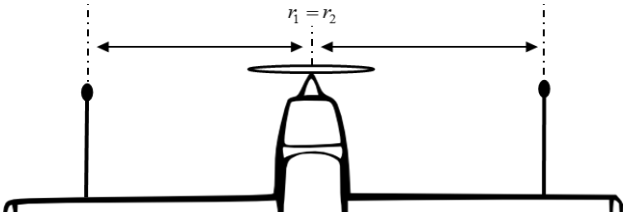

Figure 3-20: Source/sensor configuration for SIMO fixed-wing simulation.

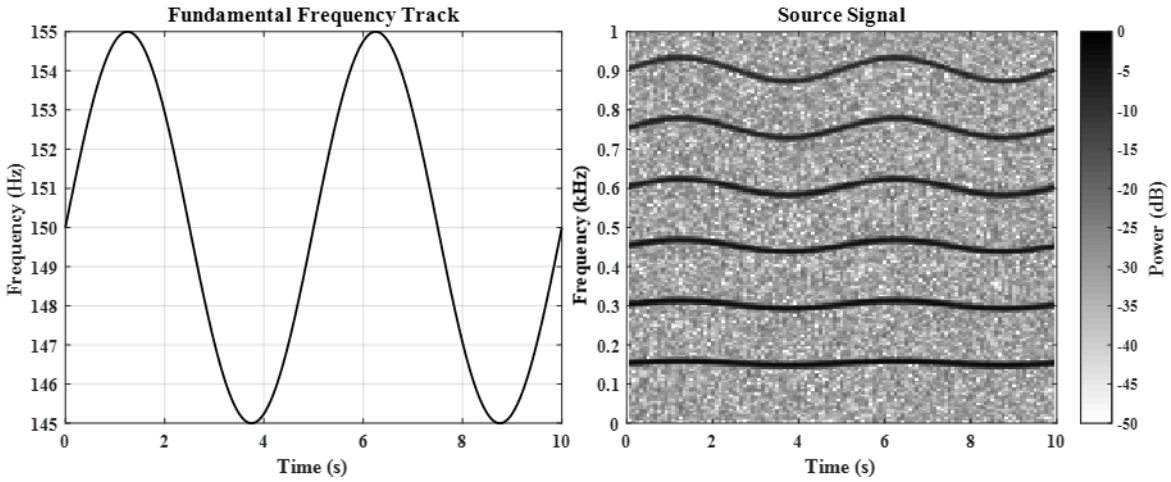

Figure 3-21: Source signal for fixed-wing simulation.

To simulate the MIMO system associated with the multi-rotor configuration, a total of three sensor signals ($K = 3$) was constructed from three generating source components ($S = 3$). Each sensor was equally spaced from its closest respective source as depicted below in Figure 3-22 such that $\alpha_{1,1} = \alpha_{2,2} = \alpha_{3,3}$. Each signal consisted of $M = 6$ harmonic components with fundamental frequencies of $F_{0,1} = 150$, $F_{0,2} = 151$ Hz and $F_{0,3} = 152$ Hz embedded in Gaussian noise of unity variance ($\sigma = 1$). All signals were constructed using a sampling rate of $f_s = 4$ kHz for a total duration of 10 seconds and modulated using amplitude and frequency values of $B_1 = 2$, $B_2 = 4$, $B_3 = 5$ and $F_{m,s} = 0.2$ Hz respectively. Figure 3-23 provides a spectrogram of the generated signals for each sensor, while Figure 3-24 provides a plot of the time variant fundamental frequency.

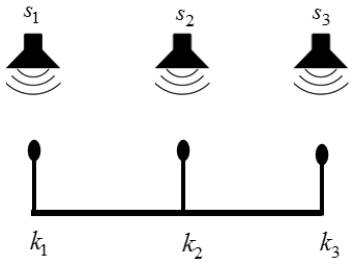

Figure 3-22: Source/sensor configuration for MIMO multirotor simulation.

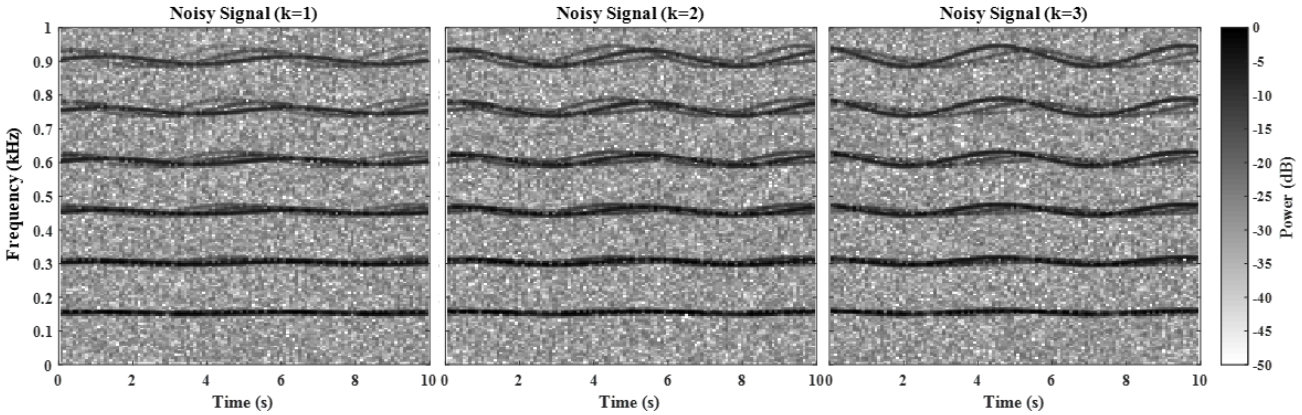

Figure 3-23: Source signal for multirotor simulation.

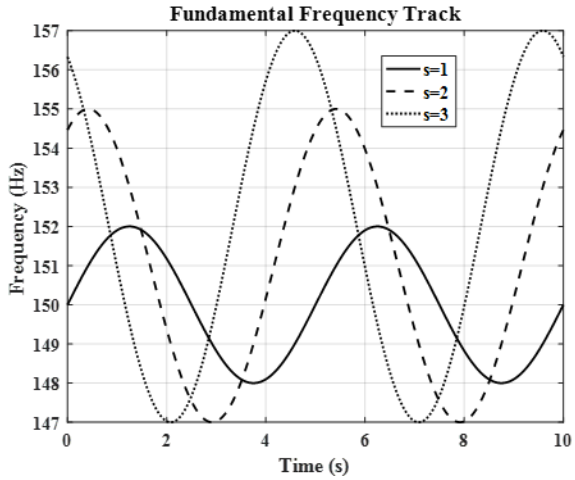

Figure 3-24: Source fundamental frequency track for multirotor simulation.

3.6.2 - Simulated Fixed-Wing Results

Since the filtered output for the two sensor channels are nearly identical, the results for only one channel are provided. The filtering methods evaluated include the IIR notch, FIR notch, and Comb filters. Zero-phase implementation of the IIR filter was achieved via the forward and reverse filter method previously described in Section 3.2.2, while the FIR and Comb filter were implemented in their presented linear and zero-phase forms respectively.

Table 3-1 provides the parameter values used for each filter type, while Figure 3-25 provides the magnitude response for each filter using these values. From the plot, it is evident that the IIR filter offers the most desirable response since it has a very narrow notch bandwidth and essentially no passband ripple. As expected, the FIR conversion of this filter provides a reasonable approximation, although ripples in the passband are present due to the Gibb's phenomena. This may be decreased by applying a windowing function such as the Hann or Hamming to the filter impulse response (coefficients). However, this will also result in decreased notch depth and increased notch width as previously discussed. The Comb filter provides the least desirable response since it has a relatively wide notch bandwidth and produces a non-linear passband gain.

Table 3-1: Filter parameters for fixed-wing simulation.

IIR	FIR	Comb
$r = 0.98$	$r = 0.98$	$G = 0.631$
$\mu = 2 \times 10^{-4}$	$\mu = 2 \times 10^{-7}$	$\mu = 0.5$
-	$N = 150$	-

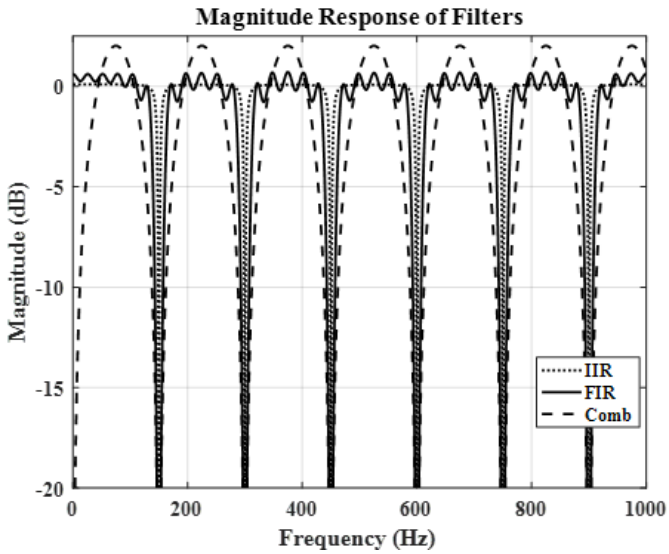

Figure 3-25: Magnitude response of evaluated filters.

The results obtained via each filtering method are displayed below in Table 3-2, while Figures 3-26 through 3-28 provide spectrograms and frequency tracking plots. From the results displayed, it is evident that each filter was successful in tracking and removing the non-stationary harmonic narrowband components. The IIR notch filter attained the highest performance with MFD values very close to zero and MSE values approximating the pure Gaussian noise power ($\sigma = 1$). This was closely followed by the FIR notch filter, which produced very similar results in terms of MSE values. MFD values were significantly higher, although a visual examination of the filtered spectrogram indicates all noise components were still adequately removed.

Table 3-2: Filter results for fixed-wing simulation.

	IIR	FIR	Comb
MSE	1.06	1.09	0.83
MFD	0.14	0.45	0.80

From examination of the FIR frequency tracking plot displayed in Figure 3-27, it is apparent that a lag exists between the actual and tracked noise frequency. This effect is caused by the cascaded iterative form in which the filter is implemented, since a change at the initial stage must propagate through all stages before it becomes realized at the final output. This propagation delay is essentially equal to the group delay that would be present if all nonlinear stages were combined into one complete filter. Since the group delay is equal to half the number of coefficients for a linear-phase filter, the adaptation delay may be approximated by:

$$t_{apt} = \frac{MN}{2 f_s} \tag{3.129}$$

For the current case in which $M = 6$, $N = 150$, and $f_s = 4000$ Hz, a delay value of approximately $t_{apt} = 0.11$ seconds is present. It should be noted that the above form is only an approximation to the true delay since the tracking filters are not linear and therefore do not produce a constant group delay. Since the actual delay value will be a function of frequency which also influences the LMS update algorithm, accurately determining these values and their effect on tracking performance would be very complex and is outside the scope of this thesis.

In addition to the adaptation delay, there is also a general implementation delay which is inherent in all FIR filters. This can be observed by the whitespace at the beginning of the spectrogram displayed in Figure 3-27. It is given by the group delay that would be present if all filter stages (initial and linear-phase correction) for all sources were combined:

$$t_{grp} = \frac{MNS}{f_s} \tag{3.130}$$

For the current scenario, a delay value of $t_{grp} = 0.22$ s is therefore present. It should be noted that the group delay of the filter does not affect its operation in any way. It is not a processing lag between the incoming real-time signal and filtered output, but rather a transient response only present at the beginning of the filtered output.

The Comb filter performed the worst of the three filters, which is apparent from examination of the spectrogram displayed in Figure 3-28. This was expected since the magnitude response offered the least desirable features. Frequency tracking accuracy was also much less with an average deviation of 0.83 Hz. However, this was also expected since the LMS gradient function required for tracking was simply approximated via a backwards finite difference approach. The filter did effectively remove all harmonic components, which is evident from the spectrograms displayed below, although a relatively large area surrounding the noise locations was also removed. This is generally undesirable since any target signal located relatively close to one of these components would be greatly attenuated.

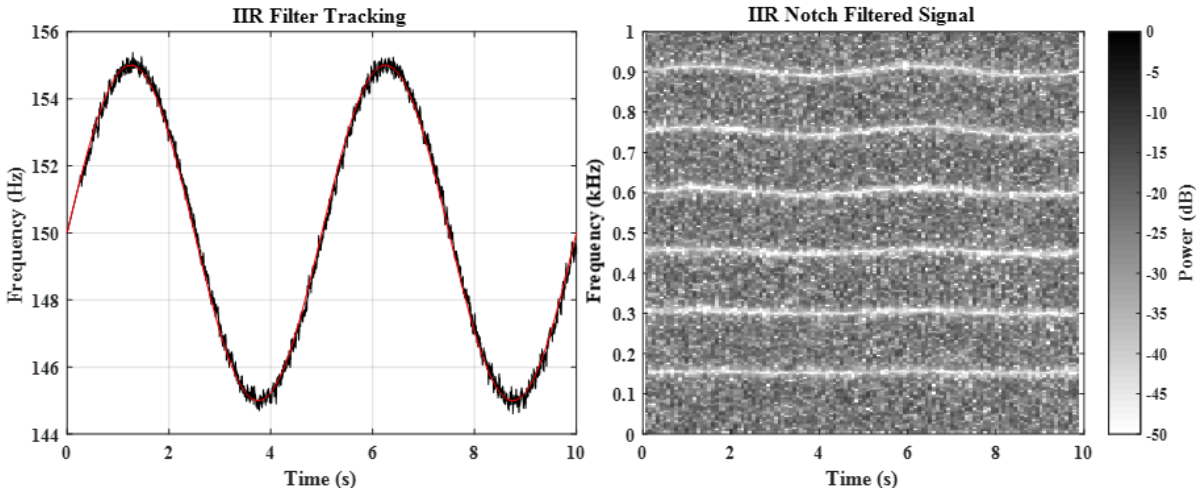

Figure 3-26: IIR notch filter results for fixed-wing simulation.

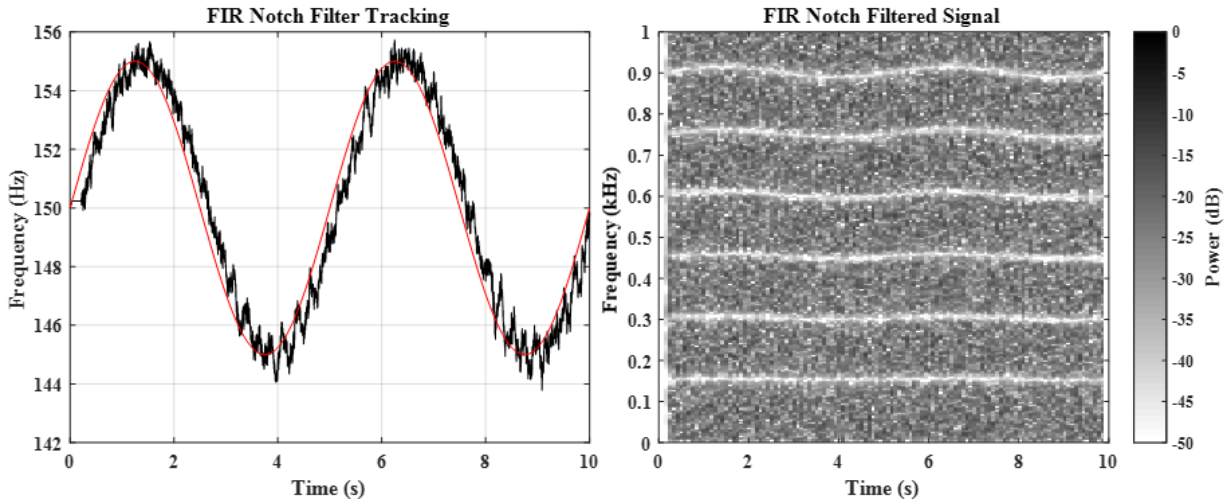

Figure 3-27: FIR notch filter results for fixed-wing simulation.

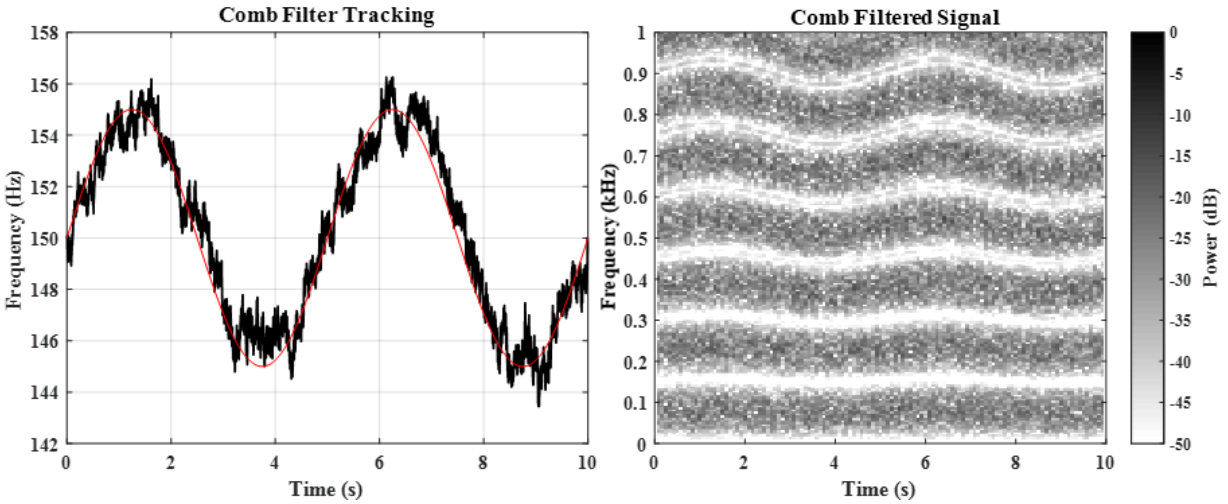

Figure 3-28: Comb filter results for fixed-wing simulation.

3.6.3 - Simulated Multirotor Results

The following section provides the results obtained from utilizing the IIR and FIR notch filters on signals constructed to simulate that obtained from a multirotor aircraft. The Comb filter was not evaluated for this case since its relatively poor frequency response in combination with required processing delays (as previously discussed) would render it impractical for experiments presented in this thesis. The FIR notch filter is also impractical for this application as previously discussed, due to the computational requirements associated with removing a high number of source signals with harmonic components. However, the filter is still evaluated to demonstrate that it is capable and practical for systems with similar dynamics but with fewer noise components.

Table 3-3 displayed below provides the filter parameter values used at each source removal stage. The results obtained are displayed in Table 3-4, while Figures 3-29 through 3-31 provide spectrograms and frequency tracking plots. Based on the results obtained, it is evident that both filters are effective in removing all harmonic components for each of the simulated signals. As before, the IIR filter attained the highest performance with MFD values very close to zero. The FIR form produced MFD values approximately double this, which is

apparent from examination of the frequency tracking plots presented in Figure 3-31. However, a comparison of the spectrograms obtained from each method produces little to no discernible difference. Higher tracking errors were obtained for the second and third source stages for both filters. This was expected since these stages had a larger degree of non-stationarity.

From the FIR frequency tracking plot, it is again evident that an adaptation lag exists between the actual and tracked noise frequency. This value is equal to that present for the fixed-wing simulation since the number of harmonics and filter length did not change. It is apparent from the spectrograms displayed below in Figure 3-30 however that the general processing delay has greatly increased. This increase is due to the presence of multiple noise sources which effectively increases the overall (combined stage) filter length. For the current case in which $M = 6, N = 150,\ S = 3$, and $f_s = 4000$ Hz, a delay value of $t_{grp} = 0.68$ seconds is present.

Table 3-3: Filter parameters for multirotor simulation.

IIR	FIR
$r = 0.99$	r = 0.98
$\mu_1 = 1 \times 10^{-4}$	N = 150
$\mu_2 = 2 \times 10^{-4}$	$\mu_1 = 8 \times 10^{-8}$
$\mu_3 = 3 \times 10^{-4}$	$\mu_2 = 8 \times 10^{-8}$
-	$\mu_3 = 16 \times 10^{-8}$

Table 3-4: Filter results for multirotor simulation.

	IIR			FIR		
	$S = 1$	$S = 2$	$S = 3$	$S = 1$	$S = 2$	$S = 3$
MSE	1.71	1.72	1.71	1.85	1.95	1.96
MFD	0.09	.017	0.21	0.32	0.45	0.36

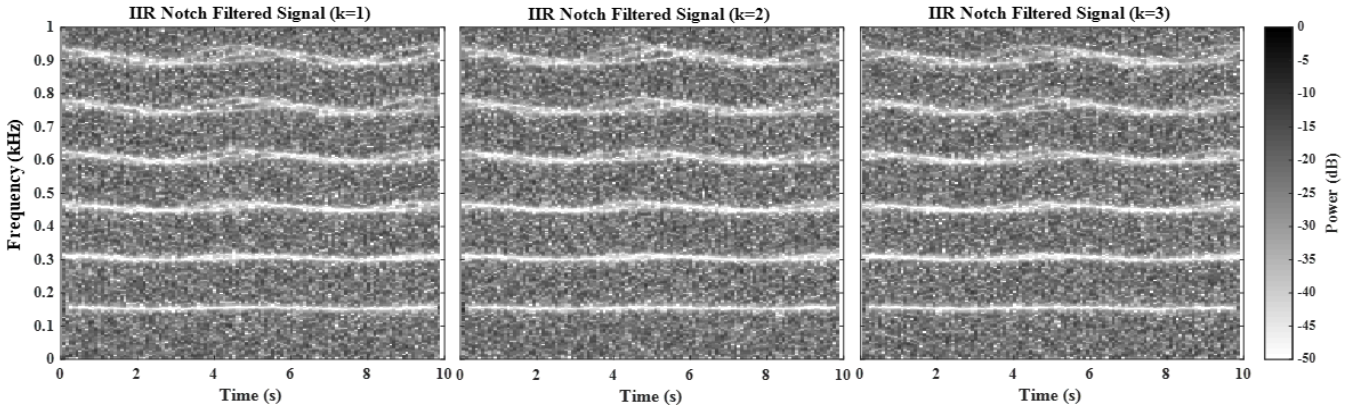

Figure 3-29: IIR notch filter results for multirotor simulation.

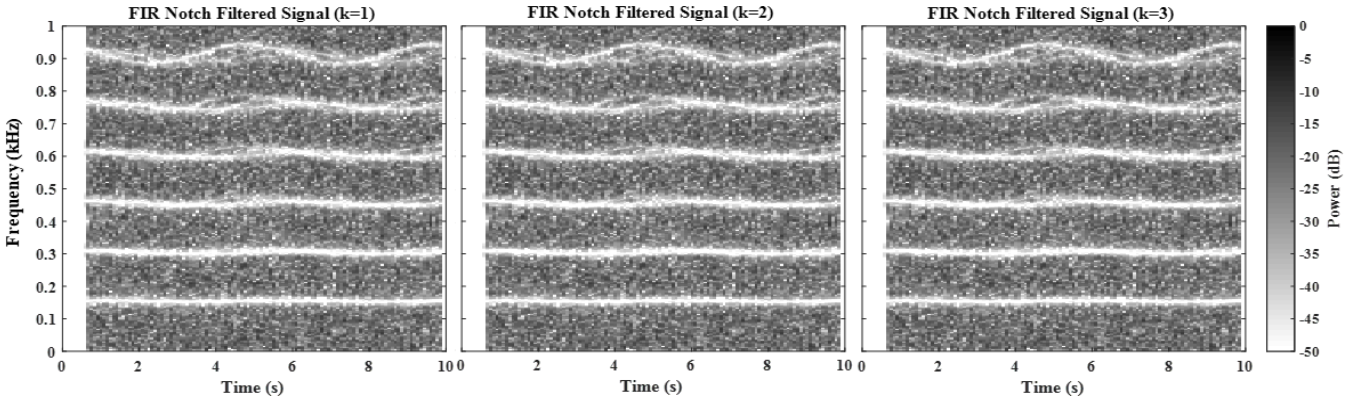

Figure 3-30: FIR notch filter results for multirotor simulation.

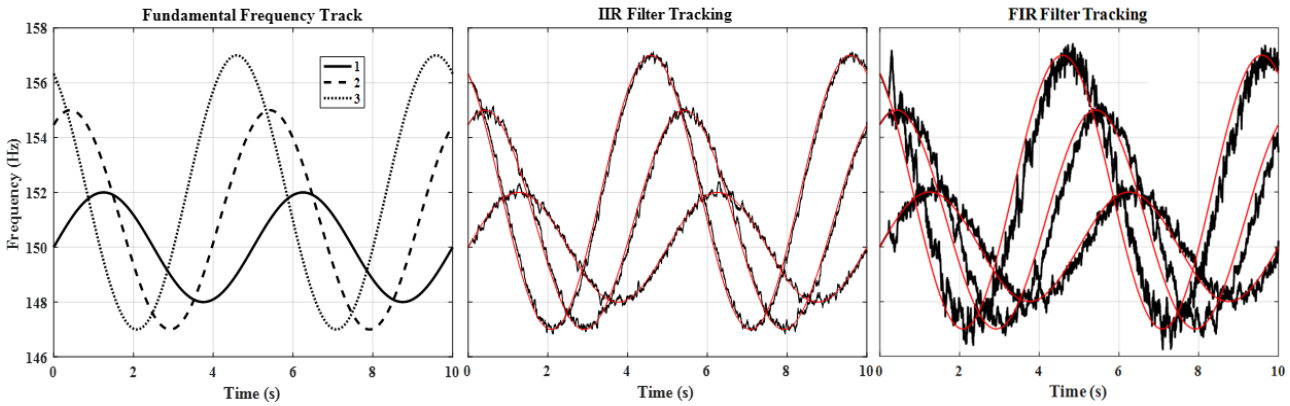

Figure 3-31: Frequency tracking results for multirotor simulation.

3.7 - Experimental Studies

The performance of the presented filters is now evaluated using data obtained from experimental studies involving both fixed-wing and multirotor aircraft.

3.7.1 - Fixed-wing Experiments

To validate the performance of the presented filters for an SIMO system, the methods were applied to data obtained from acoustic detection experiments conducted using a fixed-wing UAV. The recorded data was obtained from a fly-by of a Delta X-8 UAV at approximately 100 m above a ground-based loudspeaker emitting a 500 Hz tone. Data pertaining to this experiment is presented in further detail in Chapter 6 (TS#1). Acoustic signals were originally sampled at a rate of 48 kHz but were decimated to 4 kHz prior to filtering to reduce computational requirements. Figure 3-32 provides a spectrogram plot for a 30 s segment of the unfiltered noise corrupted signal. A plot of the fundamental frequency track for the engine noise is also displayed to aid in analyzing filter performance. The track was obtained by performing autocorrelations on consecutive 0.2 s signal segments. Thus, frequency values displayed by the plot are only an approximation as the true values at any given time are actually unknown.

Since the filtered outputs for the recorded signals are essentially identical, the results for only one of the four recorded channels is presented. The filtering methods evaluated include the IIR notch, FIR notch, and Comb filters. Zero-phase implementation of the IIR filter was achieved via the forward and reverse filter method previously described in Section 3.2.2, while the FIR and Comb filter were implemented in their presented linear and zero-phase forms respectively. The various parameters used for each filter are given in Table 3-5, while the MSE and MFD values obtained are displayed in Table 3-6. Figure 3-33 provides spectrogram plots for each of the filter output signals, while Figure 3-34 provides the corresponding frequency tracks.

From the results displayed, it is evident that each filter was successful in tracking and removing the non-stationary harmonic narrowband components. From a visual inspection of the spectrogram plots, it is apparent that the IIR notch filter again offered the best performance, closely followed by the approximated FIR form. The increased performance attained by the IIR filter can be explained by the fact that a larger pole radius was

utilized which effectively produced a narrower notch bandwidth. Although the same value could have been utilized for the FIR form, doing so would require an increase in the filter length and thus reduce its implementation viability from a computational standpoint. The Comb filter again performed the worst of the three filters, which is apparent from examination of the filtered signal spectrograms. Notch values were very wide compared to that of the IIR and FIR filters, while some harmonic components were barely attenuated to average floor levels.

It should be noted that the MSE and MFD values obtained cannot be directly utilized to assess and compare the performance of each filter, since the actual values for the ideal noise-free signal are unknown. This is evident when comparing the results obtained via the IIR and Comb filters for example. From the spectrogram plots, it is clear the IIR filter preforms better, however the values given in Table 3-6 suggest that the Comb filter is superior. The lower MSE value can be attributed to the fact that the Comb filter has notches spanning the whole signal bandwidth while the IIR filter only removed the first 7 harmonics as required for further processing operations (detection, localization, etc.). Differences in MFD values can also be attributed to the fact that the actual fundamental frequency values are unknown and only approximated through performing autocorrelations as previously mentioned.

Table 3-5: Filter parameters for fixed-wing experiment.

IIR	FIR	Comb
$r = 0.99$	$r = 0.98$	$G = 0.708$
$\mu = 1.8 \times 10^{-4}$	$N = 150$	$\mu = 5$
-	$\mu = 1.2 \times 10^{-7}$	-
$M = 7$	$M = 7$	-

Table 3-6: Filter results for fixed-wing experiment.

	IIR	FIR	Comb
MSE	0.43	0.59	0.21
MFD	0.51	0.67	0.31

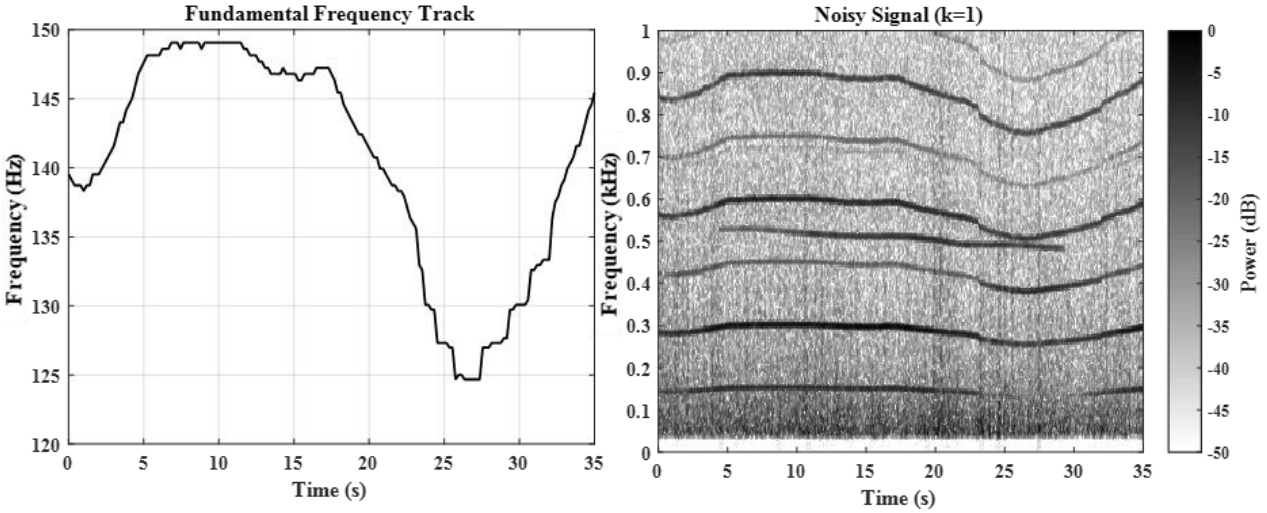

Figure 3-32: Spectrogram and fundamental frequency track plots for fixed-wing noisy signal.

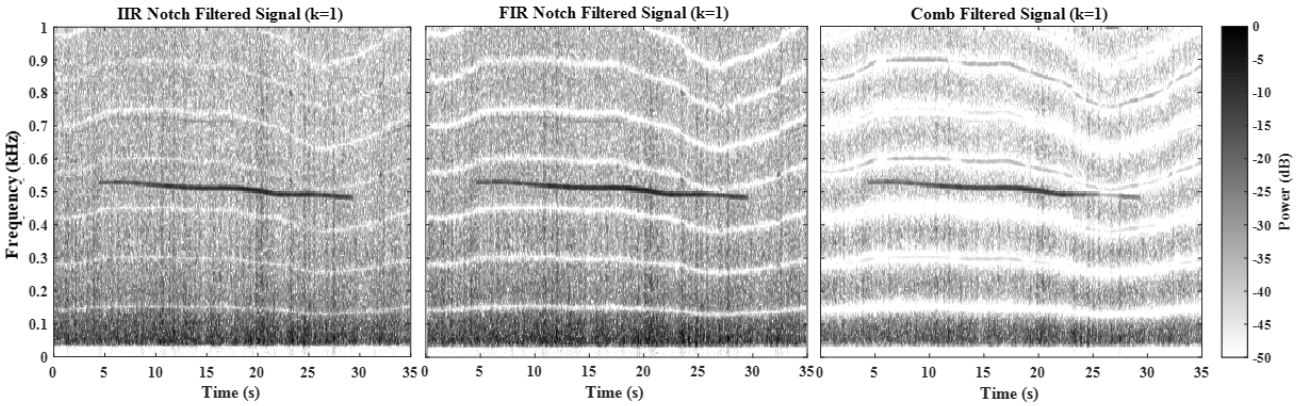

Figure 3-33: Spectrograms of filtered signals for fixed-wing experiment.

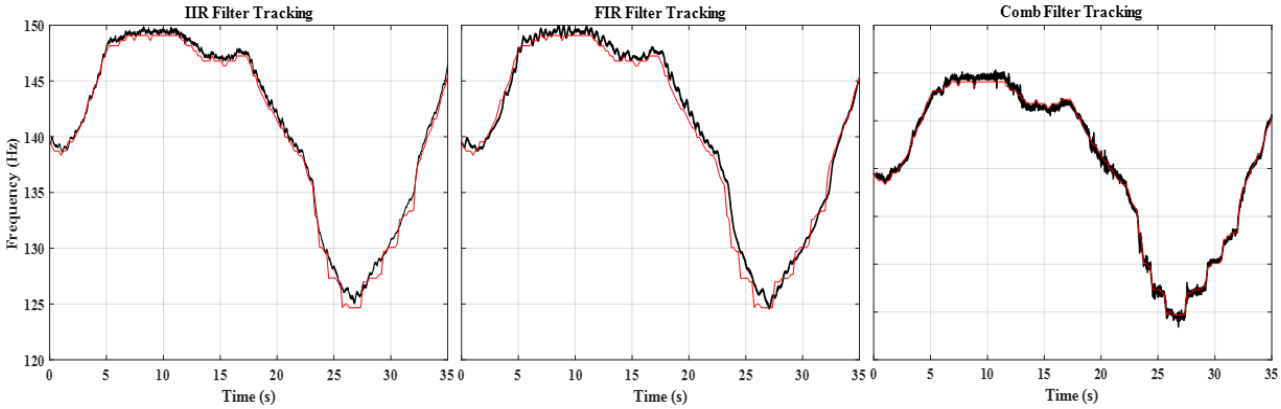

Figure 3-34: Filter frequency track plots for fixed-wing experiment.

3.7.2 - Multirotor Experiments

To validate the performance of the presented filters for a MIMO system, the methods were applied to data obtained from an acoustic experiment conducted using a multirotor UAV. The data was recorded during a flyby of the Kraken multirotor (described in Chapter 6) past a ground-based loudspeaker emitting an 83 Hz base tone with 6 harmonic components. Acoustic signals were originally sampled at a rate of 96 kHz but were decimated to 4 kHz prior to filtering to reduce computational requirements. Since the FIR and Comb filters were not considered appropriate due to computational requirements and processing delays as previously discussed, only the IIR notch filter is evaluated.

Table 3-7 provides the filter parameters used, while Table 3-8 provides the MSE and MDF results. Spectrogram plots for the noisy and filtered signals are displayed in Figures 3-35 to 3-37 for three of the six recorded channels, while Figure 3-38 displays frequency tracking plots. From the spectrograms, it is evident that the noise components are similar in frequency and highly non-stationary. In addition, the presence of sub and partial harmonics are clearly visible for the first and second channels. This is believed to be caused by a nonuniformity in the propeller located directly below the first microphone which was the result of minor surface damages incurred during a previous flight. It was unknown at the time of the experiments that the propeller would produce the observed acoustic effect since the damages were repaired and appeared negligible during visual inspection. Thus, the inclusion of the ability to remove such sub and partial harmonic components as previously

presented in Section 3.3.3 was vital to effectively filter the noise corrupted signal. From a visual inspection of the presented spectrograms, it can be concluded that the multichannel IIR notch filter was effective at removing all nonstationary noise components without attenuating the target source signal to any apparent degree.

Table 3-7: IIR filter parameters for multirotor experiment.

$r = 0.99$	$\mu_3 = 6 \times 10^{-4}$
$M = 10$	$\mu_4 = 8 \times 10^{-4}$
$\mu_1 = 2 \times 10^{-4}$	$\mu_5 = 10 \times 10^{-4}$
$\mu_2 = 4 \times 10^{-4}$	$\mu_6 = 12 \times 10^{-4}$

Table 3-8: IIR Filter results for multirotor experiment.

	$S = 1$	$S = 2$	$S = 3$	$S = 4$	$S = 5$	$S = 6$
MSE	1.18	1.19	1.19	1.17	1.20	1.18
MFD	0.16	0.24	0.21	0.19	0.22	0.18

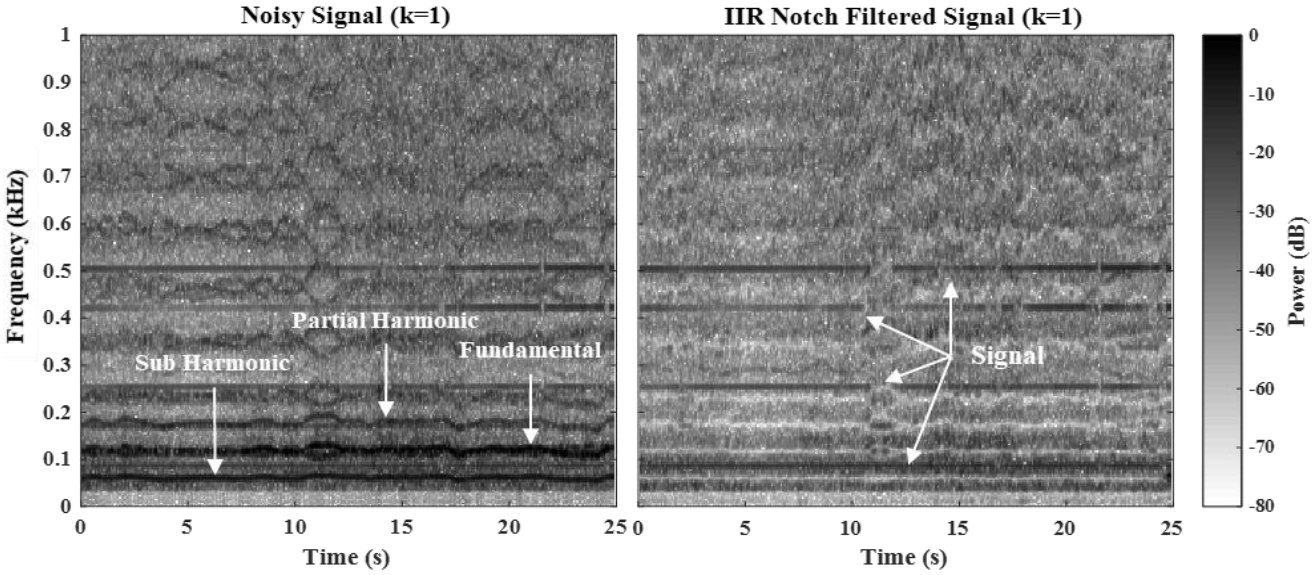

Figure 3-35: IIR notch filter results for multirotor experiment (k=1).

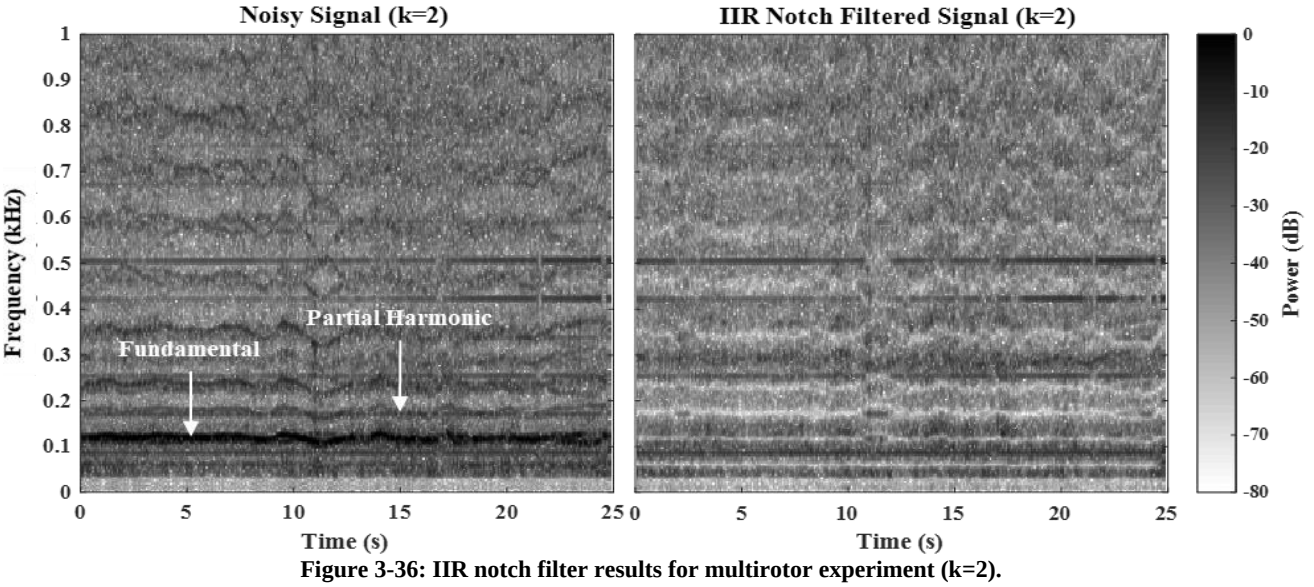

Figure 3-36: IIR notch filter results for multirotor experiment (k=2).

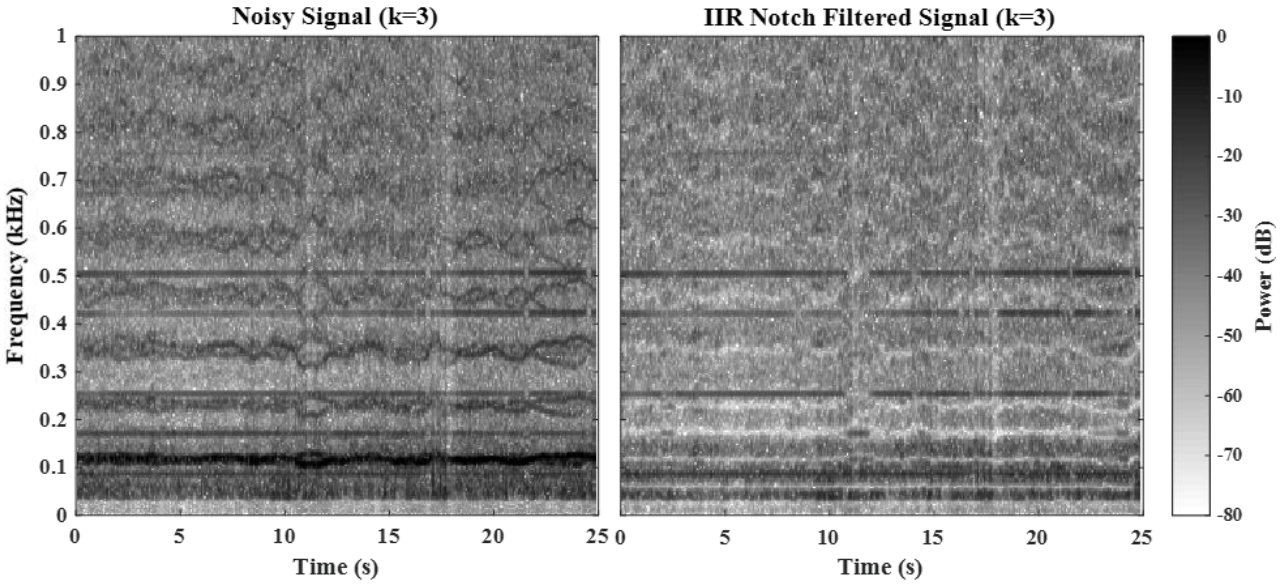

Figure 3-37: IIR notch filter results for multirotor experiment (k=3).

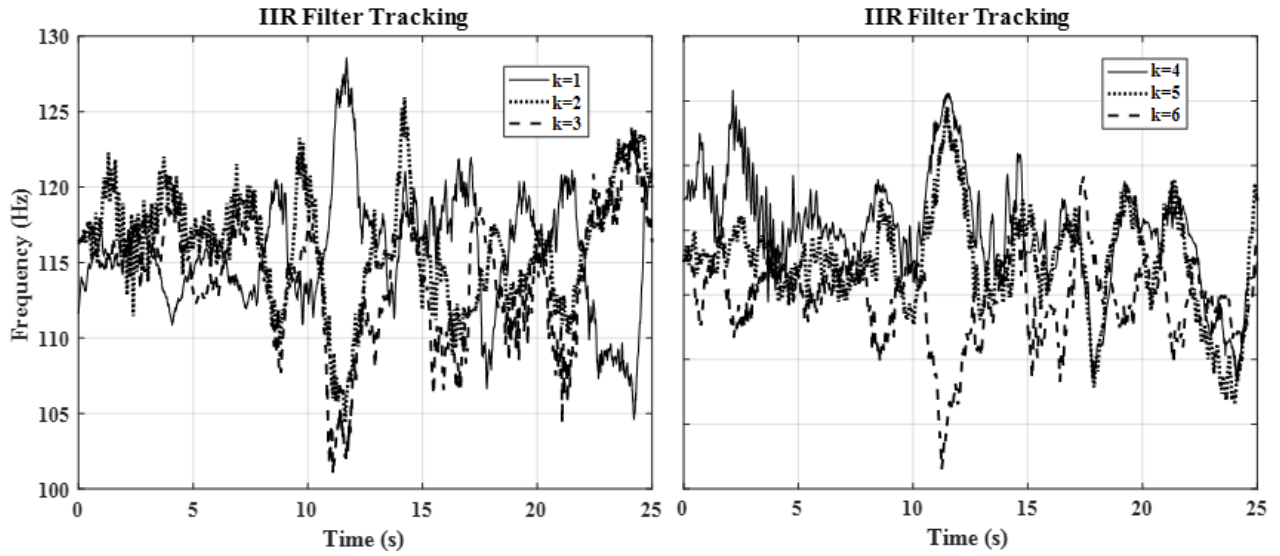

Figure 3-38: IIR frequency track for multirotor data.

It is apparent from the values displayed in the above filter parameters table that increasing LMS step sizes is required for each successive filtering stage. This is due to the fact that each of the propellers operate at similar frequencies and often coexist at the same value. For such instances, removal of a source component at the primary stage and associated signal (ex: $s = 1, k = 1$) may also remove parts of the subsequent source to be removed at the next stage by its primary signal (ex: $s = 2, k = 2$). Thus, for each subsequent stage there is less noise to be removed which requires a larger step size for adequate tracking. To minimize this effect, the notch bandwidth should be kept as small as possible. However, this effect can be completely avoided if required/desired by modifying the parallel filter configuration. Figure 3-39 displayed below provides a depiction of the configuration used to obtain the above results, and an alternate form which is unaffected by the conflicting location effects previously described. For such a configuration, the primary source of concern is removed from its maximally acquired signal prior to removing any other components. However, the main

drawback with this approach is that there is no guarantee that all source components will be removed. For example, if source s_1 has greater power than s_2 in both of the acquired signals, both notch filters will track this source and leave the other untouched. The method was evaluated using the above experimental data and generally found to perform less well especially during aircraft maneuvering operations where acoustic source levels are not of equal value.

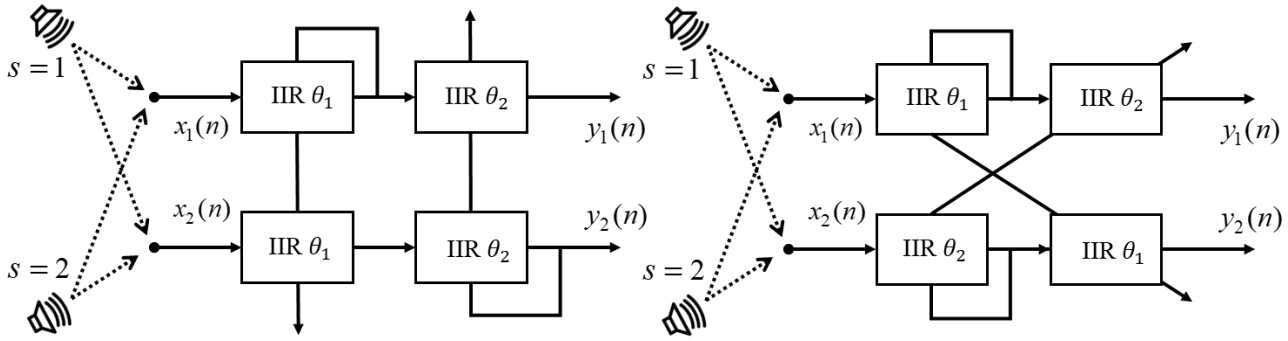

Figure 3-39: Left) Standard parallel configuration. Right) Alternate parallel configuration.

3.7.3 - Conclusions

Based on the results obtained from simulated and experimental data, it can be concluded that all of the proposed filtering methods provide an effective means to remove non-stationary harmonic noise without the use of any reference signal and without producing any phase distortions. The IIR notch filter offered the best performance in all filtering scenarios examined, with frequency tracking capabilities exceeding that of the proposed FIR notch and Comb filters. Modifications made to the filter to facilitate multichannel systems with partial harmonic components proved to be essential in order to effectively filter signals obtained via multirotor experiments.

The proposed FIR notch also provided similar results to that obtained via the IIR form, proving the filter is an effective alternative for situations in which a linear-phase inherently stable filter is required. It was shown that the filter may also be effectively used for multichannel systems, although computational requirements may limit the viability for applications in which a large number of noise sources are present.

The Comb filter generally performed the least well of the proposed methods. However, the filter was still effective in removing noise components for all scenarios examined. As with the FIR form, filtering large numbers of source components may not be viable for some applications due to the processing delays required to generate a distortionless output. However, the method does provide the advantage of requiring very few computational resources compared to the other proposed forms. Thus, for applications which demand low computational loads, a zero-phase output, and inherently stable operation, this filter may offer the best solution.

- 4 - Signal Enhancement & Detection

The following chapter presents several signal processing techniques which may be used to effectively enhance the detection of harmonic narrowband signals. Enhancement processors are proposed which exploit the presence of multiple channels in conjunction with the harmonic structure and phase acceleration properties of acquired signals to increase target detectability. In addition, a discussion of signal detection theory as it pertains to acoustic sensing is provided. A description of the Constant False Alarm Rate (CFAR) detection methodology is presented along with a distribution-free technique which may be applied to systems where the underlying noise properties are either unknown or changing with time. A theoretical analysis of the approach for the case of non-independent tests where noise approximations are constrained by bandwidth limitations is conducted. In addition, extensions to the method are proposed which minimize computational requirements while still maintaining operational performance levels. Finally, the proposed enhancement processors and detectors are evaluated using computer generated data.

4.1 - Introduction

The problem of detecting narrowband sinusoidal signals in noisy data is a very prominent one occurring in many fields such as sonar, radar, acoustics, and communications. In the context of developing an acoustic-based collision-avoidance system, the subject of signal detection can be partitioned into two main areas: The first is the development of signal processing techniques (processors) to condition signals for enhanced component detection. For example, a basic processor may consist of averaging multiple signal spectra in an attempt to increase SNR values through constructively combining coherent periodic components while destructively combining incoherent random noise. The second area corresponds to the specific detection statistic or algorithm used (detectors) to determine the presence of a target signal (or lack thereof). For example, it may be decided that a signal is present if a spectral peak value is simply greater than some threshold value.

In many instances, information regarding the target signal and/or noise is known in advance to aid in the detection problem. Such information may involve the expected frequency, phase, or amplitude of the target signal, and/or the underlying statistical distribution of the corrupting noise. For such cases, parametric hypothesis testing may employed to determine the test statistic and detection probability for a given system setup. However, in some situations variations in target properties and environmental conditions do not permit the use of prior assumptions. In such instances, parametric methods cannot be reliably employed and thus alternative non-parametric techniques must be used instead. Such methods are often termed distribution-free since they may be used without any prior knowledge of the underlying noise statistics.

Consider again the signal detection scenario pertaining to this thesis. We wish to determine the presence of some narrowband periodic component of unknown frequency, amplitude, and phase, within a signal containing random noise of unknown properties that may change from one instant to the next depending on environmental and operational conditions. For such scenarios, detection operations must typically be performed in the

frequency domain in order to achieve some constant pre-determined false alarm rate. This constraint however will not reduce the ability to detect the unknown target signal, since frequency domain methods generally offer better performance than the time domain alternatives. This is in part due to the fact that the Fourier based approach is considered an optimal receiver for narrowband signal detection [134]. Although signal parameters typically used to aid in detection are unknown, information pertaining to the physical acoustic source characteristics may be exploited to enhance detection capabilities. Such information would include: 1) the source signal is narrowband and periodic, 2) it typically contains a harmonic structure, and 3) it is emitted continuously with time producing a consistent phase progression. Using this information, processing algorithms may then be used to help dissociate between deterministic signals and random noise components. This will ultimately provide an additional level of detection sensitivity since traditional frequency domain methods rely solely on magnitude-based comparisons.

4.2 - Part I – Signal Enhancement

The following section presents a number signal enhancement processors which may be utilized to increase the detection of harmonic narrowband signals. Three general classes of processors are presented. These include: 1) Harmonic Spectral Transforms (HSTs), 2) Phase Acceleration Processors (PAPs), and 3) Modified Coherence Processors (MCPs). The performance of the processors is also evaluated in Section 4.4 using computer generated data, and the results are compared to those found in the literature using other similar techniques.

4.2.1 - Background Information

4.2.1.1 - *Harmonic Amplitude Processors*

The presence of harmonic components is an inherent property of many acoustic signals such as those generated by voiced speech, musical instruments, and aircraft propulsion systems [72, 135-137]. It arises from the presence of a physical boundary which establishes the condition for standing wave generation [138]. The process of examining harmonic signals to determine the fundamental frequency is known as pitch detection. It has been studied extensively, with the majority of developments produced in the context of voiced speech detection and classification. Pitch detection differs from the general detection of harmonic signals in that the former assumes the presence of the harmonic signal to an appreciable degree and attempts to estimate its component structure; while the latter simply attempts to determine the presence of such a signal without much regard to the accuracy of its structure. Thus, the performance of pitch detection algorithms is often evaluated using metrics such as the Gross Pitch Error (GPE) and Fine Pitch Error (FPE) rather than detection probability [139]. Although the problem of pitch detection is not directly relevant to the scope of this thesis, algorithm developments in this area may prove beneficial in the context of signal detection. That is, processing conducted to enhance pitch detection accuracy may also prove beneficial when applied for the purpose of signal detection. Thus, the concept will be further explored through investigation of potentially relevant pitch detection algorithms.

Pitch detection methods can generally be classified as being either parametric or non-parametric. Parametric algorithms define a stochastic model for the noisy signal then employ Maximum Likelihood (ML) or Maximum a Posteriori (MAP) techniques to estimate the model parameters [140]. Non-parametric algorithms avoid using explicit signal models and identify the pitch of a signal either from its harmonic structure in the frequency domain, or its periodicity in the time domain. The major benefit of utilizing a non-parametric approach is that no priori information or assumptions regarding the signal and/or noise is required; a property that typically degrades the performance of parametric models if deviations between these assumptions and actual signal conditions occur [139, 141-144]. Thus, methods proposed in this thesis will focus solely on non-parametric methods.

Pitch detection can be performed in either the time or frequency domain. Some popular time domain methods reported in the literature include: the modified autocorrelation method using clipping (AUTOC) [145], simplified inverse filtering technique (SIFT) [146], data reduction method (DARD) [147], parallel processing method (PPROC) [148], PRAAT [149], and the average magnitude difference function (AMDF) [150]. Popular frequency domain methods include: the harmonic product spectrum (HPS) [151], frequency histogram [151], harmogram [152, 153], cepstrum (CEP) [154, 155], PEFAC [141], subharmonic to harmonic ratio algorithm (SHRP) [156], BaNa [139], and various combinations of these above forms [157-160].

In general, frequency domain (FD) methods tend to perform better than the time domain (TD) alternatives. Since harmonic signals are expressed as a series of narrow equispaced peaks in the Fourier domain, the ability to manipulate and recognize patterns is greatly increased. In addition, FD methods are much less susceptible to unpredictable phase effects which in contrast will often cause the failure of many TD algorithms [144]. For example, both amplitude and power spectra do not utilize or contain phase information. Thus, algorithms such as the HPS, CEP, BaNa, and PRAAT, which use these spectral forms as part of their underlying foundation eliminate any of these phase-related problems [161]. In some circumstances however, this may also be considered a negative aspect since phase information may be used to assist in pitch detection operations. The phase fluctuation based processors developed by Wagstaff for example, utilize phase information contained in complex signal spectra to discriminate between random noise and periodic components [162-167]. Coherence based approaches also rely on the magnitude and phase similarity between signals to detect the presence of correlated components [168-172].

Although FD algorithms generally offer increased detection capabilities, they are not without their drawbacks. For example, the frequency histogram presented by Schroeder [151] has been found to be highly susceptible to octave errors [173], while the Cepstrum technique has been found to perform poorly under low SNR conditions [139]. Although very popular and well performing, the HPS method has been found to fail if any harmonic components are missing from the acquired signal [174]. This is a legitimate concern if notch filtering is used to remove narrowband self-noise components which may coexist with source harmonics. However, the non-parametric iPEEH method has been proposed as a potential solution to this problem. By preforming a self-

circular convolution on the FFT spectrum, the iPEEH method acts to enhance a degraded harmonic structure by "filling in" any missing components [175]. The approach has been found to increase the performance of the HPS, PEFAC, SHRP, and BaNa algorithms when applied to speech and music signals with poor harmonic structure [175].

The Harmonic Product Spectrum (HPS) was first developed by Schroeder as a means of pitch detection by exploiting the equispaced peak pattern found in the magnitude spectra of harmonic signals. It is found by taking the product of harmonically spaced components and is given by the following expression [151]:

$$HPS(f) = \prod_{r=1}^{R} |X(f \cdot r)| \tag{4.1}$$

where $X(f)$ is the magnitude spectrum of a discrete time signal obtained via the FFT and R is the number of harmonics being considered. The fundamental pitch frequency f_o is then indicated by the location of the maximum spectral value:

$$f_o = \text{ArgMax}\left[HPS(f) \right] \tag{4.2}$$

Typically, the HPS is calculated across a band of interest where the fundamental frequency is believed to reside. The maximum possible range is given by $[0\,, f_s/2R]$, where f_s is the sampling frequency. Similar to the HPS proposed by Schroeder, Hinch proposed a harmonic based Periodogram which he termed the Harmogram [152]. It is found by summing equispaced components of the power spectrum across a frequency band of interest according to:

$$HAR(f) = \sum_{r=1}^{R} |X(f \cdot r)|^2 \tag{4.3}$$

As previously stated, the HPS given by equation (4.1) has been found to perform poorly if harmonic components are missing or severely degraded [174]. The iPEEH method proposed by Wu offers a solution to this problem by enhancing the harmonic structure prior to processing through performing a self-circular convolution on the signal spectra [175]. The circular convolution for two discrete time sequences is given by [176]:

$$y = x * h = F^{-1}\{F\{x\} \cdot F\{h\}\} \tag{4.4}$$

where $F\{\}$ and $F^{-1}\{\}$ indicates the forward and inverse FFT respectively. For the case of self-circular convolution involving frequency spectra we have:

$$\widehat{Y} = X * X = F^{-1}\{F\{X\} \cdot F\{X\}\} = F^{-1}\{F\{X\}^2\} \tag{4.5}$$

where X represents the complex frequency spectra of the discrete time signal x. Since the values of X are complex numbers, the resulting signal will not attain true maximum values if subsequent harmonics are out of phase with one another. Thus, as shown by Wu, better performance can be obtained by using the magnitude spectra instead:

$$\hat{Y} = |X| * |X| = F^{-1}\left\{F\left\{|X|\right\}^2\right\} \tag{4.6}$$

The iPEEH enhancement procedure is performed as follows: First the signal is converted to the frequency domain using the FFT. The resulting spectrum is then circular convolved with itself to form the enhanced harmonic structure model. Prior to taking the IFFT of the self-convolution process, the signal baseband amplitude is removed by high pass filtering. This is accomplished by the setting the low frequency (ex: < 10Hz) bins of the $F\{|X|\}$ result to zero. The final step is to superimpose the self-convolved signal with the original by adding the two. In order to maintain appropriate scaling, both signals are normalized before the operation. Mathematically, this process may be represented according to the following:

$$\hat{Z} = \| |X| \| + \| \hat{Y} \| \tag{4.7}$$

where $\| \ \|$ indicates normalization ($\|x\| \equiv x/\max[x]$). Since the two forms are added, any harmonic components missing in the signal will be approximated using the convolution output. The downside to this approach is that noise contained in the convolution output will also be added to the original signal(s).

4.2.1.2 - *Fluctuation Based Processors*

Fluctuation Based Processing (FBP) is a signal processing technique that has currently received relatively little attention in the scientific community. In brief, the principle involves utilizing fluctuations in signal amplitude and/or phase to discriminate between periodic signals and random noise. The underlying concept is that random noise will produce large fluctuations between successive frequency domain realizations (windows), while continuous periodic signals will remain relatively stationary. Thus, processors may be developed to discriminate between the two by forming some basis of quantifying the relative degree of fluctuation present.

FBP has been primarily developed for underwater acoustic applications since associated signals are typically subject to fluctuations produced by a range of environmental factors [167]. However, many of the principles developed for this purpose may also be applied to aerial acoustics, since there are many similarities between the two. For aerial acoustics, signal fluctuations may occur for a number of reasons including:

- Temperature, density, and flow variations along the sound propagation path
- Turbulent flow fields
- Changing source-receiver range separation
- Interference from multipath arrivals

Fluctuation based processors may be broadly categorized as being amplitude-based, phase-based, or a combination of both. Amplitude-based processors typically track variations in magnitude or power spectra taken successively in time. Frequency components exhibiting larger variations are deemed undesired noise and thus attenuated, while components depicting less variation are amplified. Phase-based methods utilize the concept of signal phase acceleration for noise discrimination. Random signals will exhibit random phase fluctuations or acceleration with time, while continuous periodic components will produce a constant phase progression. The

term phase acceleration refers to the phase deviation of signal components from expected values (based on the component frequency) across some time interval.

Consider a single component acoustic wave which may be described by the following equation:

$$A(t) = p(t)e^{j\theta} \tag{4.8}$$

where $p(t)$ is the sound pressure amplitude, and θ is the phase function which is given by:

$$\theta = \left(2\pi ft + \kappa x + \theta_0 + \chi\right) \tag{4.9}$$

where κ is the wave number given by $\kappa = 2\pi/\lambda$, λ is the wavelength, x is the distance between the source and receiver which may also be a function of time, θ_0 is the initial phase of the signal at time $t = 0$, and χ is the phase shift caused by overlapping consecutive FFTs. For FFT windows of constant length and overlap spacing, the phase shift between adjacent windows will also be constant and is given by $\chi = 2\pi fwt_w$. Here w is the current FFT window number with respect to the beginning of the signal, and t_w is the time spacing between adjacent windows. For stationary systems of constant frequency, the phase θ will be a linear function with respect to time. Thus, to obtain maximum temporal coherence for a given signal, phase differences between adjacent times can simply be removed by linear phase shifting. However, applications such as the acoustic detection of aircraft do not satisfy these conditions. Relative motion between the sensing and intruding aircraft, multipath reflection and interference, and variations in the transmission medium would all lead to random fluctuations perceived phase values.

The proposed solution to address this problem is known as phase fluctuation or phase acceleration processing. Consider a series of phase measurements taken from a single frequency bin at equal time intervals. If no external phase influences are present, the angular rotation rate for the signal will remain constant with time:

$$\omega = \frac{\Delta\theta}{t} = constant \tag{4.10}$$

where ω is the angular velocity. For the discrete time case, the above form can be approximated as:

$$\omega_w = \frac{\theta_w - \theta_{w-1}}{w-(w-1)} = \theta_w - \theta_{w-1} \tag{4.11}$$

where w is the discrete FFT time index number which corresponds to w^{th} windowed segment. If the signal and its phase components remain stationary, then $\omega_w = \omega_{w-1}$. In terms of the phase angle this is given by:

$$\theta_w - \theta_{w-1} = \theta_{w-1} - \theta_{w-2} \tag{4.12}$$

For the case where components are not stationary with time, equation (4.12) will not hold. To account for the inconsistency, we may introduce a phase fluctuation factor ϕ such that the above equation now becomes:

$$\left[\left(\theta_w - \phi_w\right) - \theta_{w-1}\right] = \left[\theta_{w-1} - \theta_{w-2}\right] \tag{4.13}$$

Solving the phase fluctuation for the w^{th} windowed segment gives:

$$\phi_w = \left[\theta_w - \theta_{w-1}\right] - \left[\theta_{w-1} - \theta_{w-2}\right] \tag{4.14}$$

Notice that the above definition is actually the phase acceleration with respect to the discrete time index. Hence, the interchangeable use of the words phase fluctuation and phase acceleration.

The area of FBP has been largely pioneered by Wagstaff with numerous amplitude and phase-based processors having been presented [162-167, 177, 178]. Common amplitude-based processors include the Wagstaff Integration Silencing Processor (WISPR), Advanced WISPR Summation (AWSUM$_k$) filters, and the WISPR II processor [167, 177, 178]. The WISPR and AWSUM$_k$ filters are a class of processors which can be expressed via the following generalized functional form:

$$\bar{M}_k = \bar{M}_k(a) = \left[\frac{1}{W}\sum_{w=1}^{W} a_w^k\right]^{1/k} \tag{4.15}$$

where a is the data stream sequence:

$$a = \left\{a_1, a_2, a_3, ..., a_W\right\},\ a_w \geq 0\ \forall\ w \tag{4.16}$$

Using equation (4.15) with various integer values of k, a number of common statistical mean quantities can be defined which are displayed below in Table 4-1.

Table 4-1: Statistical mean definitions.

Harmonic Mean	Geometric Mean	Standard Mean	Root Mean Square
$k=-1$	$k=0$	$k=1$	$k=2$
$\bar{M}_{-1} = \left[\frac{1}{W}\sum_{w=1}^{W} a_w^{-1}\right]^{-1}$	$\bar{M}_0 = \lim_{k\to 0}\left[\frac{1}{W}\sum_{w=1}^{W} a_w^k\right]^{1/k} = \left[\prod_{w=1}^{W} a_w\right]^{1/W}$	$\bar{M}_1 = \frac{1}{W}\sum_{w=1}^{W} a_w$	$\bar{M}_2 = \left[\frac{1}{W}\sum_{w=1}^{W} a_w^2\right]^{1/2}$

If the sequence a represents the power value for a given FFT frequency bin across w consecutive windows, the AWSUM$_k$ class of filters may be thus defined as [167, 178]:

$$\text{AWSUM}_k = \bar{M}_{-k} = \left[\frac{1}{W}\sum_{w=1}^{W} X_w^{-k}\right]^{-1/k} \tag{4.17}$$

where X_w is the power spectra of the w^{th} time domain windowed segment, and W is the total number of segments. Using the above definition, the WISPR processor is given by [177, 178]:

$$\text{WISPR} = \text{AWSUM}_1 = \bar{M}_{-1} = \left[\frac{1}{W}\sum_{w=1}^{W} X_w^{-1}\right]^{-1} \tag{4.18}$$

which is also the harmonic mean of consecutive power values.

Gains on the order of 10 dB have been achieved using the WISPR processor compared to the average incoherent power [177]. Wagstaff showed that the processor performs somewhat independently of the FFT resolution but is highly dependent on the number of averaged windows and overlap percentage. Higher order AWSUM processors were shown to give much larger SNR gains than the WISPR processor, but at the expense of longer

required record lengths. The AWSUM_4 processor for example was found to achieve SNR enhancements in excess of 20 dB [167]. However, in order to achieve such gains an upwards of 700 windowed FFT segments were required using a 50% overlap, and 1000 with a 99% overlap. The WISPR processor also required a upwards of 50 segments with a 75% overlap to achieve gains on the order of 10 dB. In general, higher order AWSUM filters were found to produce larger SNR gains but required substantially higher record lengths to do so; lower order forms tend to perform better for short record lengths [167].

In addition to amplitude-based methods, a number of phase-based processors have also been proposed by Wagstaff [162-164, 166, 179]. Although originally developed for underwater acoustics, these processors have been successfully utilized for numerous relevant applications such as acoustic detection from an aerial balloon, and aircraft detection from ground-based arrays [163, 165]. Some common phase-based processors include: the phase-aligned vector average processor (PAV) [163], the scalar phase-aligned temporally coherent average processor (PAC) [163], and the AWSUM Environmentally Sensitive Phase processor (AWSUM ESP) [164]. Although found to give good results, the AWSUM ESP processor utilizes three empirical terms which are application specific and greatly affect overall performance. In addition, no details are given in how to choose appropriate values based on expected signal conditions [164]. Wagstaff states that the PAV and PAC processors can theoretically achieve enhancement gains on the order of $10G \cdot \log[W]$ and $15G \cdot \log[W]$ respectively; compared to the average incoherent power which would produce gains on the order of $5G \cdot \log[W]$, were W is the total number of windowed FFT segments and G is a scaling constant dependent on experimental conditions (typically $G < 1$) [163].

The PAV and PAC processors are given by the following equations respectively:

$$PAV(X,\phi) = \left[\frac{1}{W}\sum_{w=1}^{W}|X_w(f)|\cos[\phi_w(f)]\right]^2 + \left[\frac{1}{W}\sum_{w=1}^{W}|X_w(f)|\sin[\phi_w(f)]\right]^2 \tag{4.19}$$

$$PAC(X,\phi) = \left[\frac{1}{W}\sum_{w=1}^{W}|X_w(f)|\cos[\phi_w(f)]\right]^2 \tag{4.20}$$

In general, the fluctuation-based processing concept has been found to be very successful in producing significant SNR gains. However, current methods require a relatively stationary signal throughout the total time record which is often of considerable length. For applications such as those pertaining to this thesis where system kinematics produce dynamic signals, such processing methods may not be appropriate. However, it will be shown that concepts developed in this area may be utilized to construct phase acceleration-based processors which rely on the presence of multiple signals rather than multiple time realizations.

4.2.1.3 - *Coherence Processors*

Coherence-based processing is essentially a means of analyzing signals based on their phase and amplitude similarity. The standard definition for the complex coherence of a signal pair is given by [75]:

$$\gamma(f) = \frac{S_{xy}(f)}{\sqrt{S_{xx}(f)S_{yy}(f)}} \tag{4.21}$$

where S_{xy}, S_{xx}, and S_{yy} are the cross and auto spectral densities for the Fourier transformed signals given by:

$$S_{xy}(f) = X(f)Y^*(f) \tag{4.22}$$

$$S_{xx}(f) = X(f)X^*(f) \tag{4.23}$$

$$S_{yy}(f) = Y(f)Y^*(f) \tag{4.24}$$

where * indicates complex conjugation. Typically, the magnitude squared coherence (MSC) is utilized instead of the complex definition since it provides a real value ranging from 0 to 1 to indicate the degree of coherence [180]:

$$\Gamma(f) = |\gamma(f)|^2 = \frac{|S_{xy}(f)|^2}{S_{xx}(f)S_{yy}(f)} \tag{4.25}$$

The MSC is essentially a frequency dependent correlation coefficient which establishes the degree of linearity between two similar signals or the input/output of a system. For signals in which magnitude and phase differences remain constant with time, values of $\Gamma \approx 1$ will be obtained indicating a highly linear relationship. However, the presence of random noise will produce magnitude and phase fluctuations with time giving values of $\Gamma < 1$. If there is no linear relationship between the two signals, a value of $\Gamma \approx 0$ will be obtained.

From the above definition we can see that for one observation, the MSC will always maintain $\Gamma = 1$ across all frequencies. Thus, in order to obtain an accurate estimate, we must average a number of segments of the cross and auto spectral densities. This approach is often referred as the Welch method of coherence and is given by [181]:

$$\gamma(f) = \frac{\frac{1}{W}\sum_{w=1}^{W} S_{xy}(f,w)}{\sqrt{\frac{1}{W}\sum_{w=1}^{W} S_{xx}(f,w)\frac{1}{W}\sum_{w=1}^{W} S_{yy}(f,w)}} = \frac{\sum_{w=1}^{W} S_{xy}(f,w)}{\sqrt{\sum_{w=1}^{W} S_{xx}(f,w)\sum_{w=1}^{W} S_{yy}(f,w)}} \tag{4.26}$$

$$\Gamma(f) = \frac{\left|\sum_{w=1}^{W} S_{xy}(f,w)\right|^2}{\sum_{w=1}^{W} S_{xx}(f,w)\sum_{w=1}^{W} S_{yy}(f,w)} \tag{4.27}$$

where w is the segment number and W is the total number of segments averaged.

The use of coherence as a signal processing tool has been widely reported in the literature, with many applications in areas of filtering [182-185], signal detection [168-172], and spatial localization [186-191]. However, the standard definition as presented above is currently limited to the case of two signal systems. For

situations in which more than two signals are available, the standard definition does not facilitate a measure of the overall system coherence. For such cases, coherent combination pairs have been traditionally used to achieve processing gains [189]; although such an approach is not optimal in any sense. Other coherence based methods such as the Coherent Phase Line Enhancer (CPLE) developed by Jong is fully capable of simultaneously utilizing multiple channels [192-194]. However, as the author points out, this technique requires stationary signals across relatively large time segments. Some instances of non-stationary signals may be analyzed provided information regarding frequency and phase changes are known. Such information can then be used to actively modify sampling rates to effectively produce stationary signals. This general procedure is known as Order Tracking Analysis (OTA) and is typically applied to rotating machinery for the purpose of fault investigation in low SNR environments [195, 196]. The obvious downside to this approach is that even if the required information can be obtained, actively modifying sampling rates to examine one signal component will produce distortions in all other signal components unless they share the same dynamic features.

Recently, developments have been made to the address the multi-channel limitations of the standard coherence form. The Generalized Magnitude Squared Coherence (GMSC) proposed by Ramirez provides a means to establish an overall coherence value for a system of similar signals [197, 198]. In brief, the GMSC is given by the maximum eigenvalues of the complex coherence matrix.

Defining $\gamma_{i,j}(f)$ as the complex coherence spectrum between the i^{th} and j^{th} signals, the complex coherence matrix is thus given by:

$$\mathbf{C}_{\gamma}(f)=\begin{bmatrix} \gamma_{1,1}(f) & \gamma_{1,2}(f) & \cdots & \gamma_{1,S}(f) \\ \gamma_{2,1}(f) & \gamma_{2,2}(f) & \cdots & \gamma_{2,S}(f) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{S,1}(f) & \gamma_{S,2}(f) & \cdots & \gamma_{S,S}(f) \end{bmatrix} \tag{4.28}$$

where $\{i,j=1,2,\ldots,S\}$, $\gamma_{i,j}(f)=1$ for $i=j$, and S is the total number of signals. The GMSC is then given by the following:

$$\Gamma(f)=\frac{1}{S-1}\left(\lambda_{max}\left[\mathbf{C}_{\gamma}(f)\right]-1\right)^2 \tag{4.29}$$

where λ_{max} represents the maximum Eigenvalue of $\mathbf{C}_{\gamma}$ at each frequency. As with the MSC, the GMSC is also bounded between 0 and 1. A value of $\tilde{\Gamma}=1$ is obtained if all signals are perfectly correlated, and $\tilde{\Gamma}=0$ if no signals are correlated. For the case of two signals ($S=2$), it can be shown that the above definition reduces to the standard MSC form.

4.2.2 - Spectral Enhancement Processors

Several signal enhancement processors are now presented to aid in the detection of harmonic narrowband signals such as those produced by propeller driven aircraft. Extensions to previous developments made by Wagstaff [163-167, 177-179], Ramirez [197, 198], Schroeder [151] and Hinch [152], are proposed to exploit

the presence of multiple channels containing harmonic narrowband components with continuous phase functions. Using the proposed methods, it will be shown that increased signal detectability can be greatly achieved over that obtained via standard incoherent averaging methods and other relevant forms such as the PAV and PAC processors developed by Wagstaff [163]. Due to the large number of processors presented in this section, a list is provided below to give a summary of the various forms and proposed contributions.

Summary of developed processors and proposed contributions:

1) A generalized harmonic transform approach is proposed as an extension to that previously presented by Schroeder [151] and Hinch [152] to facilitate arbitrary spectral forms and include multi-channel systems.
2) The concept of phase vector coherence for multichannel systems is defined and utilized to establish a series of phase acceleration-based processors. In addition, a modulo 2π phase modification is proposed to eliminate the phase wrapping issues associated with the PAV and PAC processors developed by Wagstaff [163] and ultimately increase detection performance.
3) Modifications to the PAV and PAC processors are presented to exploit properties of harmonic signals and enable use with multichannel systems.
4) The use of self-circular convolution as proposed by Wu [175] is evaluated to determine if increased source detection can be achieved for harmonic signals with missing components.
5) A generalized multi-channel coherence function similar to that developed by Ramirez [197, 198] is proposed which utilizes phase acceleration in addition to standard phase and amplitude values.

4.2.2.1 - *Harmonic Spectral Transforms*

The following section presents a simplistic but effective means to enhance signal detectability in the frequency domain by exploiting the spectral peak periodicity associated with harmonic signals. Termed Harmonic Spectral Transforms (HSTs), the method is a generalization of the Harmonic Product Spectrum and Harmogram concepts developed by Hinch and Schroeder to include the processing of multi-channel systems with multiple realizations.

Consider some general harmonic signal transformed to the frequency domain via the FFT. As previously discussed, the frequency spectra of such a signal will exhibit peaks located periodically across the relevant bandwidth, such as that displayed below in Figure 4-1. Similar to the methods proposed by Hinch and Schroeder, we may exploit this pattern to enhance the fundamental frequency component for operations such as pitch detection and tracking. Referring back to equations (4.1) and (4.3), it is apparent that both the HPS and Harmogram are defined by taking either the sum or product of R harmonically spaced components across the relevant frequency spectrum (magnitude for the HPS and power for the Harmogram). In a similar manner, we may also define a more generalized spectral transform using the statistical means previously presented in Table 4-1. Given some arbitrary frequency spectra $X(f)$, the Harmonic Spectral Transform (HST) is defined as:

$$\bar{\mathrm{H}}_a[X(f)] = \left[\frac{1}{R}\sum_{r=1}^{R}\left|X(f\cdot r)\right|^a\right]^{1/a} \tag{4.30}$$

where a is an integer value which specifies the particular mean form as indicated below in Table 4-2. For pitch and/or signal detection, the HST would typically be applied across some frequency band of interest where the fundamental component is believed to reside. The maximum possible range is given by $[0\,, f_s/2R]$, where f_s is the sampling frequency. The fundamental frequency f_o is then indicated by the location of the maximum spectral value according to:

$$f_o = \mathrm{ArgMax}\left[\bar{\mathrm{H}}_a[X(f)]\right] \tag{4.31}$$

It is evident that the geometric mean is essentially equivalent to the HPS while the standard mean is equivalent to the Harmogram. Although the Harmogram was defined through the use of power spectra given by $|X(f)|^2$, it will later be shown that signal detectability is not affected by the choice of spectral units (magnitude or power). Figure 4-1 displayed below provides plots for the magnitude spectrum and standard mean HST for a harmonic signal. From this point forward, the magnitude spectrum will be considered the default unless indicated otherwise since this form will have a much lower dynamic range when performing signal processing operations.

Table 4-2: Harmonic frequency transforms.

Harmonic	Geometric	Standard	RMS
$a=-1$	$a=0$	$a=1$	$a=2$
$\left[\frac{1}{R}\sum_{r=1}^{R}\left\|X(f\cdot r)\right\|^{-1}\right]^{-1}$	$\left[\prod_{r=1}^{R}\left\|X(f\cdot r)\right\|\right]^{1/R}$	$\left[\frac{1}{R}\sum_{r=1}^{R}\left\|X(f\cdot r)\right\|\right]$	$\left[\frac{1}{R}\sum_{r=1}^{R}\left\|X(f\cdot r)\right\|^2\right]^{1/2}$

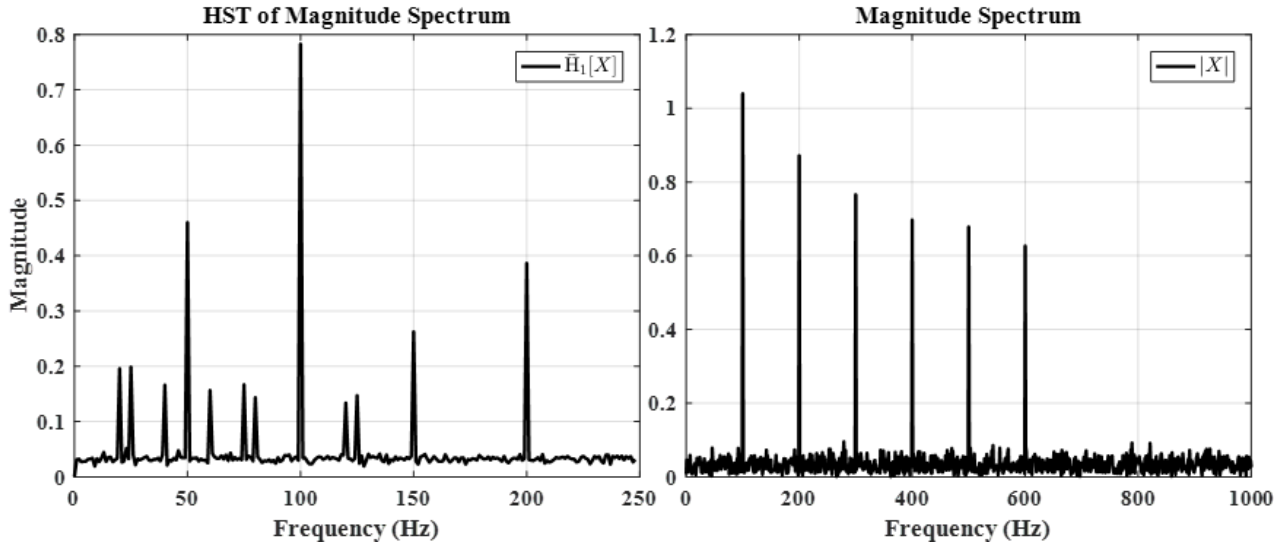

Figure 4-1: Magnitude spectrum and HST of magnitude spectrum.

If multiple channels are available for processing, the compounding operation may also be applied across channels to achieve further enhancement. For a system consisting of S signals, the Multichannel Harmonic Spectral Transform (MHST) is therefore defined according to:

$$\bar{\mathrm{H}}_{a,b}[X_s(f)] = \bar{\mathrm{H}}_b\left[\bar{\mathrm{H}}_a[X_s(f)]\right] \tag{4.32}$$

where $X_s(f)$ is the frequency spectrum for the s^{th} signal. Expanding out the above form gives:

$$\bar{\mathrm{H}}_{a,b}[X_s(f)] = \left[\frac{1}{S}\sum_{s=1}^{S}\left[\left[\frac{1}{R}\sum_{r=1}^{R}\left|X_s(f\cdot r)\right|^{a}\right]^{1/a}\right]^{b}\right]^{1/b} \tag{4.33}$$

It is evident from the above equation that a large number of processing combinations are possible. If each of the mean specification variables (a, b) are independent of one another and consist of integer values ranging between -1 and 2, a total of process 16 combinations is possible. If multiple time realizations are available, further enhancement may be achieved by applying the operation across all channels and realizations. Thus, for a system consisting of S signals with W windowed time realizations, the Generalized Harmonic Spectral Transform (GHST) may be defined as:

$$\bar{\mathrm{H}}_{a,b,c}[X_{s,w}(f)] = \bar{\mathrm{H}}_c\left[\bar{\mathrm{H}}_b\left[\bar{\mathrm{H}}_a[X_{s,w}(f)]\right]\right] \tag{4.34}$$

Expanding out the above form thus gives:

$$\bar{\mathrm{H}}_{a,b,c}[X_{s,w}(f)] = \left[\frac{1}{W}\sum_{w=1}^{W}\left[\left[\frac{1}{S}\sum_{s=1}^{S}\left[\left[\frac{1}{R}\sum_{r=1}^{R}\left|X_{s,w}(f\cdot r)\right|^{a}\right]^{1/a}\right]^{b}\right]^{1/b}\right]^{c}\right]^{1/c} \tag{4.35}$$

To obtain a more informative symbolic expression, the GHST defined above may be expressed using the following notational form instead, which now indicates the total number of harmonics R, signals S, and windowed segments W used:

$$\bar{\mathrm{H}}_{a,b,c}[\ ] \rightarrow \bar{\mathrm{H}}_{<a,b,c>}^{<R,S,W>}[\ ] \tag{4.36}$$

It is apparent from the above equations that the proposed transforms closely resemble that of the AWSUM_k processors proposed by Wagstaff [167, 178]. However, the transforms are in fact different since they are defined primarily for operations involving harmonic signals in which the peak periodicity of spectra is exploited. In contrast, the AWSUM_k processors do not exploit this spectral feature and are instead used primarily for single channel systems with large numbers of time realizations.

It should also be noted that the above transforms are not limited to standard spectra such as those obtained via the FFT operation. It can be applied to any frequency-based spectrum which exhibits peak periodicity across

the bandwidth of interest. It will later be shown for example that the approach may be applied to coherence spectra and those obtained from phase acceleration processors to enhance signal detection capabilities.

4.2.2.2 - *Phase Acceleration Processors*

The following section presents a number of processors which exploit the phase acceleration properties of periodic signals to enhance detectability. Modifications to the PAV and PAC processors proposed by Wagstaff are first presented which utilize the presence of multiple processing channels and exploit the spectral properties of harmonic signals. The concept of phase vector coherence for multichannel systems is then defined and used to establish a series of phase acceleration-based processors. A modulo 2π phase modification is also proposed to increase detection performance by eliminating phase wrapping issues associated with current processing techniques. Finally, the HSTs previously presented are applied to achieve further enhancement through exploiting the peak periodicity of the processed acceleration spectra.

4.2.2.2.1 - Modified Wagstaff Processors

Consider some general harmonic signal transformed to the frequency domain via the FFT, producing a complex valued spectrum represented by $X(f)$ which contains both magnitude and phase information. In some instances, it may be desired to express $X(f)$ in terms of phase acceleration rather than actual phase values. This may be accomplished according to:

$$X_{\phi}(f) = |X(f)|e^{j\phi(f)} \tag{4.37}$$

where ϕ is the phase acceleration as previously defined by equation (4.14). Using the above form, coherent operations may now be performed on multiple signals (or segments) without fear of destructively combing out-of-phase components since phase acceleration is a measure of progression rather than instantaneous values. That is, two signals which are completely out of phase will still constructively combine provided their phase progression with time are approximately equal for a given frequency. In this sense, phase acceleration operations are a form of phase coherence since the definition of coherence is a measure of similarity between signals with time (amplitude and phase). This concept was utilized by Wagstaff for the development of the PAV and PAC processors previously given by equations (4.19) and (4.20) respectively. For a single realization, the magnitude squared value of $X_{\phi}(f)$ is equivalent to the PAV processor, while the real part squared is equivalent to the PAC processor. For multiple realizations, $X_{\phi}(f)$ is simply coherently summed across all windowed segments.

Although the PAV and PAC processors are defined for the case of one signal with multiple realizations, the processors can also be applied to multichannel systems for a single realization. This may be achieved by simply summing across signals rather than windowed segments. Thus, for S signals the PAV and PAC processors will now be given by:

$$PAV(X,\phi)=\left[\frac{1}{S}\sum_{s=1}^{S}\left|X_s(f)\right|\cos\left[\phi_s(f)\right]\right]^2+\left[\frac{1}{S}\sum_{s=1}^{S}\left|X_s(f)\right|\sin\left[\phi_s(f)\right]\right]^2 \tag{4.38}$$

$$PAC(X,\phi)=\left[\frac{1}{S}\sum_{s=1}^{S}\left|X_s(f)\right|\cos\left[\phi_s(f)\right]\right]^2 \tag{4.39}$$

Since the processors utilize both amplitude and phase, a phase-only form may also be established to enable the use of a dual detection scheme. This would involve processing and evaluating amplitude and phase information separately to effectively produce two independent data streams. Removing the amplitude component and summing across signals gives the following phase-only forms:

$$PAV(\phi)=\left[\frac{1}{S}\sum_{s=1}^{S}\cos\left[\phi_s(f)\right]\right]^2+\left[\frac{1}{S}\sum_{s=1}^{S}\sin\left[\phi_s(f)\right]\right]^2 \tag{4.40}$$

$$PAC(\phi)=\left[\frac{1}{S}\sum_{s=1}^{S}\cos\left[\phi_s(f)\right]\right]^2 \tag{4.41}$$

For the case of harmonic signals, processed spectra will also exhibit peaks at each harmonic frequency. Thus, the HST previously presented may be employed to further increase signal detectability according to the following:

$$\bar{\mathrm{H}}_a\left[PAV(X,\phi)\right]=\left[\frac{1}{R}\sum_{r=1}^{R}\left|PAV\left[X(f\cdot r),\phi(f\cdot r)\right]\right|^a\right]^{1/a} \tag{4.42}$$

$$\bar{\mathrm{H}}_a\left[PAV(\phi)\right]=\left[\frac{1}{R}\sum_{r=1}^{R}\left|PAV\left[\phi(f\cdot r)\right]\right|^a\right]^{1/a} \tag{4.43}$$

$$\bar{\mathrm{H}}_a\left[PAC(X,\phi)\right]=\left[\frac{1}{R}\sum_{r=1}^{R}\left|PAC\left[X(f\cdot r),\phi(f\cdot r)\right]\right|^a\right]^{1/a} \tag{4.44}$$

$$\bar{\mathrm{H}}_a\left[PAC(\phi)\right]=\left[\frac{1}{R}\sum_{r=1}^{R}\left|PAC\left[\phi(f\cdot r)\right]\right|^a\right]^{1/a} \tag{4.45}$$

4.2.2.2.2 - Phase Vector Coherence

One of the downsides to the PAV and PAC processors is that both the amplitude and phase are used in a joint coherent form. Thus, two signals (or segments) which are completely incoherent and therefore reflective of random noise, may still combine to produce values greater than zero depending on their amplitudes. This would be undesirable since it ultimately reduces the level of discrimination between random noise and periodic components as defined through the use of phase acceleration. Thus, the concept of the Phase Vector Coherence (PVC) is first presented as a means of comparing signals based solely on phase information. The PVC is a measure of the phase similarity between a group of signals as a function of frequency. It is defined by taking the mean square of the vector sum of phase angles for a group of signals according to the following:

$$\vec{\Theta}(f) = \left| S^{-1} \sum_{s=1}^{S} e^{j\theta_s(f)} \right|^2 \tag{4.46}$$

where S is the total number of signals and $\theta_s(f)$ is the phase of the s^{th} signal. Expanding out the above form gives:

$$\vec{\Theta}(f) = S^{-2}\left(S + \sum_{n=2}^{S}\sum_{s=n}^{S} 2\mathrm{Cos}\left[\theta_{n-1}(f) - \theta_s(f)\right]\right) \tag{4.47}$$

It is evident from the above equation that the PVC provides a coherence measure by examining the phase difference between all possible signal combination pairs and assigns a value between 0 and 1. A value of one is achieved if all signal components are exactly in phase, and a value of zero is achieved if all components are exactly $2\pi/S$ out of phase.

4.2.2.2.3 - Acceleration Vector Coherence

Although possible, the PVC is not well suited for filtering and detection applications since it relies on the actual phase values at a given point in time rather than the progressive similarity in phase content with respect to time. To solve this problem, phase acceleration may be used instead. Substituting the phase acceleration ϕ in place of the phase value in equation (4.46) now defines the Acceleration Vector Coherence (AVC):

$$\vec{\Phi}(f) = \left| S^{-1} \sum_{s=1}^{S} e^{j\phi_s(f)} \right|^2 \tag{4.48}$$

Similar to the PVC, the AVC also provides an overall measure of the similarity in phase acceleration for a system as a function of frequency. For random noise components we would expect random phase acceleration values across signals, which would tend to produce AVC values close to zero. In contrast, for periodic signal components we would expect phase acceleration values of approximately zero, which will tend to produce AVC values of approximately one.

One issue regarding Wagstaff's PAC and PAV processors is the false detection of noise components caused by the modulo 2π nature of the Sin and Cos functions. Consider again the discrete time index definition of phase acceleration given by equation (4.14). If only white Gaussian noise is present, each phase angle will have a uniform probability distribution ranging from $-\pi$ and $+\pi$, while the phase acceleration will thus range between -4π and $+4\pi$. For coherent signals, phase acceleration values are expected to obtain values close to zero. Thus, values greater than $\pm\pi$ should obviously be considered purely noise. However, according to the definitions for the PAC, PAV, and AVC processors, acceleration values that are multiples of 2π will also attain equally high coherence values and thus indicate the false presence of a periodic signal.

To remedy this problem, a modulo 2π phase modification is proposed to attenuate processor outputs for acceleration values greater than $\pm\pi$. Consider again the AVC processor previously given by equation (4.48). Since the output is bounded between 0 and 1, we may utilize a product and power scaling factor of the form

aX^b without loss of functionality in the region of $X \approx 1$. Thus, the Adjusted Acceleration Vector Coherence (A-AVC) processor is defined by the following:

$$\vec{\Phi}^{\Psi}(f) = \left[\beta \vec{\Phi}(f)\right]^{\Psi(f)} \tag{4.49}$$

where β is the product scaling factor which may vary between 0 and 1, and Ψ is the exponential adjustment factor given by:

$$\Psi(f) = \frac{1}{S}\sum_{s=1}^{S}\left|\phi_s(f)\right| \tag{4.50}$$

In order to illustrate the effect of utilizing the proposed adjustment factors, the simple case of two signals is briefly analyzed. Using $S = 2$, the AVC, adjustment factor Ψ, and A-AVC now become:

$$\vec{\Phi}(f) = \left|2^{-1}\sum_{s=1}^{2} e^{j\phi_i(f)}\right|^2 = \cos\left[\frac{\phi_2(f) - \phi_1(f)}{2}\right]^2 = \cos\left[\frac{\Delta\phi(f)}{2}\right]^2 \tag{4.51}$$

$$\Psi(f) = \frac{1}{2}\sum_{s=1}^{2}\left|\phi_s(f)\right| = \left|\bar{\phi}(f)\right| \tag{4.52}$$

$$\vec{\Phi}^{\Psi}(f) = \left[\beta \vec{\Phi}(f)\right]^{\Psi(f)} = \left(\beta \cos\left[\frac{\Delta\phi(f)}{2}\right]^2\right)^{\left|\bar{\phi}(f)\right|} \tag{4.53}$$

Figure 4-2 displayed below provides a plot of the standard and adjusted AVC processor for various values of β. From the plot, it is evident that the standard AVC form produces coherence values of 1 at multiples of $\pm 2\pi$. For the modified form however, false coherence levels decrease exponentially with decreasing values of β while still maintaining unity for acceleration values around zero. Thus, the approach should be effective in reducing the likelihood of false detection while still maintaining the ability to correctly detect the presence of a continuous periodic signal. This will be later confirmed through the use of numerical simulation studies.

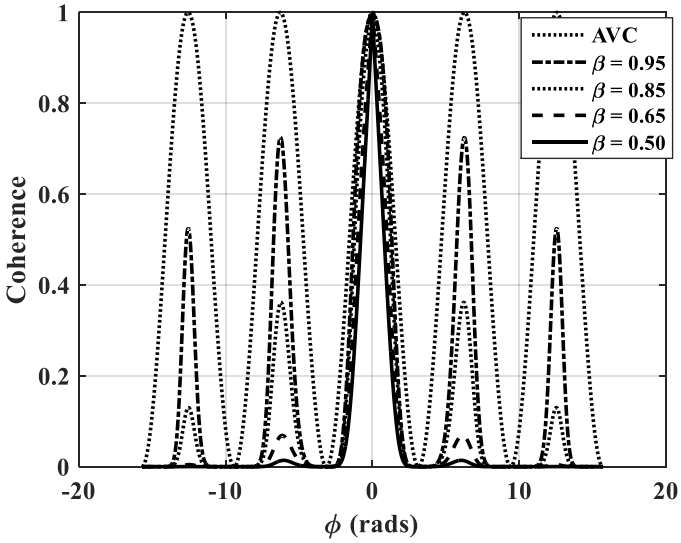

Figure 4-2: Comparison of the standard and adjusted AVC processors.

4.2.2.2.4 - System Acceleration Coherence

Another approach that may used to determine the overall phase acceleration coherence of a system is through the use of an eigenvalue decomposition. A similar approach was previously presented in Section 4.2.1.3 to extend the standard definition of coherence to multi-channel systems. By applying the AVC to all signal combination pairs, we may construct a phase acceleration coherence matrix according to:

$$\mathbf{C}_{\Phi}(f)=\begin{bmatrix}\vec{\Phi}_{i,j}(f) & \cdots & \vec{\Phi}_{i,S}(f)\\ \vdots & \ddots & \vdots\\ \vec{\Phi}_{S,j}(f) & \cdots & \vec{\Phi}_{S,S}(f)\end{bmatrix} \tag{4.54}$$

where

$$\vec{\Phi}_{i,j}(f)=\cos\left[\frac{\phi_i(f)-\phi_j(f)}{2}\right]^2 \tag{4.55}$$

and $\{i,j=1,2,\ldots,S\}$.

The overall acceleration coherence is now given by the maximum eigenvalue λ of $\mathbf{C}_{\Phi}$ which can be obtained by solving the following equation at each frequency bin:

$$\left|\mathbf{C}_{\Phi}-\lambda\mathbf{I}\right|=0 \tag{4.56}$$

where $\mathbf{I}$ is the $S\times S$ identity matrix and $||$ indicates the matrix determinant. Thus, the System Phase Acceleration Coherence (SAC) is defined by:

$$\Phi_{\lambda}(f)=\frac{1}{S-1}\left(\lambda_{max}\left[\mathbf{C}_{\Phi}(f)\right]-1\right) \tag{4.57}$$

By realizing that $\vec{\Phi}_{i,j}=1$ when $i=j$, $\lambda_{max}\left[\mathbf{C}_{\Phi}(f)\right]$ will attain a value of S if $\vec{\Phi}_{i,j}=1\ \forall\ i\neq j$ (perfect correlation), and $\lambda_{max}\left[\mathbf{C}_{\Phi}(f)\right]=1$ if $\vec{\Phi}_{i,j}=0\ \forall\ i\neq j$ (no correlation). Thus, the final value of Φ_{λ} will be bounded between 0 and 1. It should be noted that for the case of two signals ($S=2$), the above definition reduces to the standard AVC form.

As with the AVC processor, we may also apply an adjustment factor to the SAC processor to remedy phase wrapping issues. For this case the adjustment factor is defined by:

$$\Psi_{i,j}(f)=\frac{1}{2}\left[\left|\phi_i(f)\right|+\left|\phi_j(f)\right|\right] \tag{4.58}$$

and the adjusted-phase coherence matrix will be given by:

$$\mathbf{C}_{\Phi}^{\Psi}(f)=\begin{bmatrix}\vec{\Phi}_{i,j}^{\Psi_{i,j}}(f) & \cdots & \vec{\Phi}_{i,S}^{\Psi_{i,S}}(f)\\ \vdots & \ddots & \vdots\\ \vec{\Phi}_{S,j}^{\Psi_{S,j}}(f) & \cdots & \vec{\Phi}_{S,S}^{\Psi_{S,S}}(f)\end{bmatrix} \tag{4.59}$$

where

$$\vec{\Phi}_{i,j}^{\Psi}(f)=\left[\beta\vec{\Phi}_{i,j}(f)\right]^{\Psi_{i,j}(f)}=\left(\beta\cos\left[\frac{\phi_i(f)-\phi_j(f)}{2}\right]^2\right)^{\frac{|\phi_i(f)|+|\phi_j(f)|}{2}} \tag{4.60}$$

Thus, the modulo 2π phase Adjusted System Acceleration Coherence (A-SAC) is given by:

$$\Phi_{\lambda}^{\Psi}(f)=\frac{1}{S-1}\left(\lambda_{max}\left[\mathbf{C}_{\Phi}^{\Psi}(f)\right]-1\right) \tag{4.61}$$

Again, Φ_{λ}^{Ψ} will attain a value of 1 if all signals are perfectly coherent and 0 if completely incoherent.

It should be noted that unlike the AVC and SAC processors, the PAV and PAC processors are not bounded between 0 and 1. Therefore, the phase adjustment factors cannot be utilized without the possibility of decreasing performance since processor values larger than one having a phase acceleration approximating zero will be reduced to a value of approximately one. In addition, for larger acceleration values typical of noise, the occurrence of a processor value greater than one will actually become further amplified rather than attenuated. Ultimately, this will increase the likelihood of false signal detection.

4.2.2.3 - *Modified Coherence Processor*

The following section presents a coherence-based processor which may be used to enhance detection of periodic signals. The Generalized Magnitude Squared Coherence (GMSC) concept proposed by Ramirez [197] is used in conjunction with aspects of phase acceleration processing to produce an enhanced coherence form.

Consider the GMSC previously defined by equation (4.29). Since the coherence function is defined as a measure of similarity between signals as they progress in time, substituting phase values for phase acceleration will have little to no effect. However, because the GMSC is bounded between 0 and 1, an exponential phase acceleration factor may be utilized to achieve further enhancement. Such an approach was previously demonstrated with the AVC and SAC processors. If the coherence estimate is made from W overlapping windows, the exponential factor for the i^{th} and j^{th} signal pair is defined as:

$$\Psi_{i,j}(f)=\frac{1}{W}\sum_{w=1}^{W}\frac{|\phi_{i,w}(f)|+|\phi_{j,w}(f)|}{2} \tag{4.62}$$

Thus, the Generalized Acceleration Squared Coherence (GASC) is defined according to:

$$\Gamma^{\Psi}(f)=\frac{1}{S-1}\left(\lambda_{max}\left[\mathbf{C}_{\gamma}^{\Psi}(f)\right]-1\right)^2 \tag{4.63}$$

where,

$$\mathbf{C}_{\gamma}^{\Psi}(f)=\begin{bmatrix}\gamma_{i,j}^{\Psi_{i,j}}(f) & \cdots & \gamma_{i,S}^{\Psi_{i,S}}(f)\\ \vdots & \ddots & \vdots \\ \gamma_{S,j}^{\Psi_{S,j}}(f) & \cdots & \gamma_{S,S}^{\Psi_{S,S}}(f)\end{bmatrix} \tag{4.64}$$

and

$$\gamma_{i,j}^{\Psi}(f)=\left[\frac{\sum_{w=1}^{W}S_{i,j}(f,w)}{\sqrt{\sum_{w=1}^{W}S_{i,i}(f,w)\sum_{w=1}^{W}S_{j,j}(f,w)}}\right]^{\Psi_{i,j}(f)} \tag{4.65}$$

where $i,j\in\{1,2,\ldots,S\}$ and $\gamma_{i,j}^{\Psi_{i,j}}(f)=1$ for $i=j$.

By utilizing phase acceleration in addition to the magnitude and actual phase values, one would expect better discrimination between random noise and periodic components. Consider for example a periodic signal which attains a low GMSC value due to fluctuations in amplitude and/or signal phase. If the signal is indeed periodic, phase acceleration and consequently acceleration adjustment values (Ψ) should also give low values (approaching zero). Thus, the GASC will attain a higher coherence value (approaching unity) since the initial GMSC value is raised to a power approaching zero. Alternately, consider the case where a small number of windowed segments are used producing a poor approximation and giving falsely high GMSC values for noise components. If the components are in fact random noise, we would also expect high values for the adjustment factor Ψ. Thus, values for the GASC noise should now be less than the GMSC since the component is raised to a higher power. It will be later shown that the GASC provides enhanced detection capabilities compared to the GMSC through utilizing phase acceleration.

4.2.2.4 - *Combined Processor Forms*

As previously discussed, the HST is a generalized approached which may be applied to essentially any frequency spectrum that exhibits peak periodicity. Such is the case for the phase acceleration and coherence processors previously presented. Thus, HST forms for the AVC, A-AVC, A-SAC, SAC, and GASC processors are given by the following equations respectively:

$$\bar{\mathrm{H}}_{a}[\vec{\Phi}(f)]=\left[\frac{1}{R}\sum_{r=1}^{R}\left|\vec{\Phi}(f\cdot r)\right|^{a}\right]^{1/a}=\left[\frac{1}{R}\sum_{r=1}^{R}\left|S^{-1}\sum_{s=1}^{S}e^{j\phi_{s}(f\cdot r)}\right|^{a}\right]^{1/a} \tag{4.66}$$

$$\bar{\mathrm{H}}_{a}[\vec{\Phi}^{\Psi}(f)]=\left[\frac{1}{R}\sum_{r=1}^{R}\left|\left[\beta\vec{\Phi}(f\cdot r)\right]^{\Psi(f\cdot r)}\right|^{a}\right]^{1/a}=\left[\frac{1}{R}\sum_{r=1}^{R}\left[\beta\left|S^{-1}\sum_{s=1}^{S}e^{j\phi_{s}(f\cdot r)}\right|^{2}\right]^{a\Psi(f\cdot r)}\right]^{1/a} \tag{4.67}$$

$$\Psi(f\cdot r)=\frac{1}{S}\sum_{s=1}^{S}\left|\phi_s(f\cdot r)\right| \tag{4.68}$$

$$\bar{\mathrm{H}}_a[\Phi_\lambda(f)]=\left[\frac{1}{R}\sum_{r=1}^{R}\left|\Phi_\lambda(f\cdot r)\right|^a\right]^{1/a}=\left[\frac{1}{R}\sum_{r=1}^{R}\left|\frac{1}{S-1}\left(\lambda_{max}\left[\mathbf{C}_\Phi(f\cdot r)\right]-1\right)\right|^a\right]^{1/a} \tag{4.69}$$

$$\bar{\mathrm{H}}_a[\Phi_\lambda^\Psi(f)]=\left[\frac{1}{R}\sum_{r=1}^{R}\left|\Phi_\lambda^\Psi(f\cdot r)\right|^a\right]^{1/a}=\left[\frac{1}{R}\sum_{r=1}^{R}\left|\frac{1}{S-1}\left(\lambda_{max}\left[\mathbf{C}_\Phi^\Psi(f\cdot r)\right]-1\right)\right|^a\right]^{1/a} \tag{4.70}$$

$$\Psi_{i,j}(f\cdot r)=\frac{1}{2}\left[\left|\phi_i(f\cdot r)\right|+\left|\phi_j(f\cdot r)\right|\right] \tag{4.71}$$

$$\bar{\mathrm{H}}_a[\Gamma^\Psi(f)]=\left[\frac{1}{R}\sum_{r=1}^{R}\left|\Gamma^\Psi(f\cdot r)\right|^a\right]^{1/a}=\left[\frac{1}{R}\sum_{r=1}^{R}\left|\frac{1}{S-1}\left(\lambda_{max}\left[\mathbf{C}_\gamma^\Psi(f\cdot r)\right]-1\right)\right|^a\right]^{1/a} \tag{4.72}$$

$$\Psi_{i,j}(f\cdot r)=\frac{1}{W}\sum_{w=1}^{W}\frac{1}{2}\left[\left|\phi_{i,w}(f\cdot r)\right|+\left|\phi_{j,w}(f\cdot r)\right|\right] \tag{4.73}$$

If multiple time realizations are available, the HST may also be applied to the phase acceleration processors according to the GHST form previously given by equation (4.35). Note that because the functions already make use of multiple signals, the GHST form will be given by $\bar{\mathrm{H}}_{<a,1,c>}^{<R,1,W>}[\ ]$.

Since the proposed PAPs (standard and adjusted forms) only contain phase information, further enhancement may be possible by including magnitude information as well. This can be achieved through combining the HST magnitude and phase acceleration processors according to the following:

$$\prod[X,\Phi]=\bar{\mathrm{H}}_a[X(f)]\cdot\bar{\mathrm{H}}_a[\Phi(f)]=\bar{\mathrm{H}}_a[X(f)\cdot\Phi(f)] \tag{4.74}$$

where $\Phi(f)$ represents any of the proposed PAPs listed above.

Although this procedure effectively produces a processor which includes magnitude and phase information such as that of the PAV and PAC processors proposed by Wagstaff, the output form is considerably different since the two components are processed independently of each other. It will later be shown that a higher signal detectability can in fact be achieved by combining the two forms incoherently.

4.3 - Part II –Signal Detection

The following section provides a discussion on the detection of signals with unknown frequency, amplitude, and phase, in noise of unknown statistical properties. The basic concepts of threshold detection are first presented in addition to limitations for many real-world applications. A description of the Constant False Alarm Rate (CFAR) detection methodology is then provided along with a distribution-free technique which may be applied to systems in which noise properties are either unknown or changing with time. An analysis of the approach for the case of non-independent tests where noise approximations are constrained by bandwidth limitations is also conducted. Finally, extensions to the proposed distribution-free technique is provided to

minimize computational requirements while maintaining false alarm rates and detection sensitivity. The performance of the techniques is also evaluated in Section 4.4 using computer generated signals.

4.3.1 - Background Information

4.3.1.1 - *Threshold Detection*

Consider again the signal detection scenario pertaining to the application at hand. We wish to determine the presence of a periodic signal of unknown frequency, amplitude, and phase, within random noise of unknown properties that may change over time due to environmental and operational factors. At this point we assume that all narrowband self-noise has been removed leaving only random broadband components, and the signal(s) have also been transformed to the frequency domain via the FFT operation. Our task now becomes to examine the information contained in each frequency bin of the signal spectra and determine whether or not it belongs to noise or some unknown acoustic source. To facilitate this decision, we must devise some logical method or algorithm to determine which case is more likely based on the information at hand. In a statistical sense, this concept is known as hypothesis testing. The null hypothesis H_0 states that only noise is present, while the alternative H_1 states that some combination of signal and noise is present.

In order to evaluate the performance of our chosen detection algorithm, we must analyze the probability of correctly or incorrectly choosing each hypothesis. There are generally three probability areas we are concerned with: the probability of detection P_d, missed detection P_m, and false alarm P_{fa}. The basis of each are summarized in Table 4-3 below where D_1 and D_0 indicates whether the detection decision constraint has been satisfied (or not) respectively.

Table 4-3: Summary of possible detection decisions.

Accept	**Choose H_0**	**Choose H_1**
H_0 is true	Correct	Type I Error (False Alarm) $P_{fa} = P\left[D_1 \middle| H_0\right]$
H_1 is true	Type II Error (Miss) $P_m = P\left[D_0 \middle| H_1\right]$	Correct (Detection) $P_d = P\left[D_1 \middle| H_1\right]$

Depending on the particular application, we may place more emphasis on designing a detector that favours minimization of one error over another. For example, a detector for a missile defence system would seek to have an extremely low false alarm rate since falsely responding in retaliation to a perceived threat may have great consequences. For the present case of a collision avoidance system however, the consequences of not detecting another aircraft on a potential collision course would be much greater than falsely detecting one and initiating an avoidance maneuver. Unfortunately, P_{fa} and P_d have a positively correlated relationship for a given detection system; meaning one cannot be increased without consequently increasing the other. This relationship is known as the Receiver Operating Characteristic (ROC) and it is often used to analyse and

compare the performance of detection systems. The choice of acceptable values for the P_{fa} and P_d with respect to the presented detection problem will be discussed in detail in later sections.

Like the pitch detection problem previously discussed, signal detection may be performed in either the time or frequency domain. However, for situations in which information regarding the source signal to be detected is limited, frequency domain methods generally offer superior performance [134]. Perhaps the most common and basic form of frequency domain detection is a threshold based approach, which has been widely reported in the literature using various hypothesis testing methods [134, 168, 199-209]. Here, the detection decision D_1 is based on whether the magnitude or power value for a given frequency bin is greater than some predetermined threshold level. Test criterion have been established to optimize the threshold value based on the signal statistics for the null and alternative cases. The most popular of these include the Bayesian approach and the Neyman-Person (NP) criterion. The Bayesian approach seeks to minimize the total error while assigning costs to each possible event. However, it relies on priori probabilities regarding each hypothesis under investigation, which is usually not available for many real-world systems. Often situations may arise in which multiple signal parameters associated with an hypothesis such as amplitude and phase are unknown. Such cases constitute a composite hypothesis testing problem, which is typical for applications such as sonar and radar, where the signal and noise parameters may vary based on target type and environmental factors. For these scenarios, the NP criterion is typically employed instead. The optimal test is then performed by constructing the likelihood ratio and subjecting the result to a threshold established by the maximum acceptable false alarm probability. For the case of non-composite (simple) hypothesis tests it is given by the following:

$$\lambda(x) = \frac{p_1(x;H_1)}{p_0(x;H_0)} > \eta \tag{4.75}$$

where $p_0(x;H_0)$ and $p_1(x;H_1)$ are the PDFs under the null (H_0) and alternative (H_1) hypotheses respectively, and η is the threshold value. For the case of composite hypotheses, the likelihood ratio becomes:

$$\lambda(x) = \frac{p_1(x;H_1)}{p_0(x;H_0)} = \frac{\int p_1(x|\xi_1;H_1)\,p(\xi_1)d\xi_1}{\int p_0(x|\xi_0;H_0)\,p(\xi_0)d\xi_0} > \eta \tag{4.76}$$

where ξ_0 and ξ_1 are unknown vector quantities. If the above test maximizes the probability of detection for all alternatives, then it is considered a Uniformly Most Powerful Test (UMPT). However, to implement such a test, the statistic and its distribution under the null hypothesis must not depend on any unknown parameters [204]. Thus, for the detection problem pertaining to this thesis in which a signal of unknown amplitude, phase, and frequency must be detected in noise of unknown power or variance, a UMPT does not exist. For this case, the unknown quantities must be approximated by their Maximum Likelihood Estimates (MLE). The resulting likelihood ratio test is now given by:

$$\lambda(x) = \frac{p_1\left(x; \hat{\xi}_1, H_1\right)}{p_0\left(x; \hat{\xi}_0, H_0\right)} > \eta \tag{4.77}$$

where $\hat{\xi}_i$ is the MLE of ξ_i (the value that maximizes $p_1(x; \xi_1, H_1)$). The above form is known as the Generalized Likelihood Ratio Test (GLRT). For a threshold detector, the probability of detection and false alarm is thus given by:

$$P_d = \int_{\eta}^{\infty} p_1\left(x; \hat{\xi}_1, H_1\right) dx \tag{4.78}$$

$$P_{fa} = \int_{\eta}^{\infty} p_0\left(x; \hat{\xi}_0, H_0\right) dx \tag{4.79}$$

Consider the case of a single sine wave in Gaussian noise that has been transformed to the frequency domain and scaled by $1/L$, where L is the FFT length. The PDFs of the magnitude spectra for the noise only f_{X_0} and signal plus noise f_{X_1} cases are given by the Rayleigh and Rice distributions respectively [206]:

$$f_{X_0}(x) = \frac{Lx}{\sigma^2} e^{\frac{-Lx^2}{2\sigma^2}} \tag{4.80}$$

$$f_{X_1}(x) = \frac{Lx}{\sigma^2} e^{\frac{-\left(x^2+A^2\right)}{2\sigma^2}} I_0\left(\frac{xAL}{\sigma^2}\right) \tag{4.81}$$

where A is the time domain signal amplitude. The probability of false alarm is now given by:

$$P_{fa} = \int_{\eta}^{\infty} \frac{Lx}{\sigma^2} e^{\frac{-Lx^2}{2\sigma^2}} dx = e^{-\frac{L\eta^2}{2\sigma^2}} \tag{4.82}$$

which produces the following threshold:

$$\eta = \sqrt{-\frac{2\sigma^2 Log[P_{fa}]}{L}} \tag{4.83}$$

The probability of detection is thus given by:

$$P_d = \int_{\eta}^{\infty} \frac{Ly}{\sigma^2} e^{\frac{-\left(x^2+A^2\right)}{2\sigma^2}} I_0\left(\frac{xAL}{\sigma^2}\right) dx = Q\left[1, \frac{A\sqrt{L}}{\sigma}, \sqrt{-2Log\left(P_{fa}\right)}\right] \tag{4.84}$$

where Q is Marcum's Q-Function, and I_0 is the modified Bessel function of the first kind. A depiction of threshold-based detection is displayed in Figure 4-3.

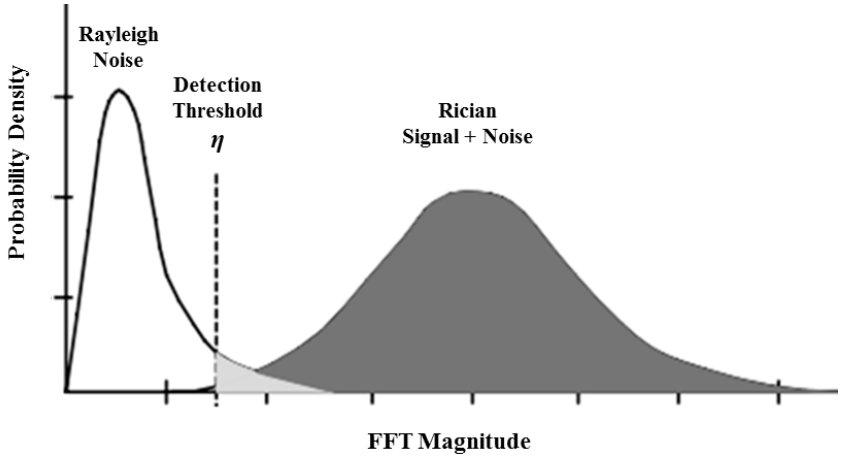

Figure 4-3: Depiction of threshold-based detection.

4.3.1.2 - *CFAR Detection*

One of the issues with the threshold approach presented above is that results are based on theoretical probabilities and are typically limited to white Gaussian noise with known variance (power). For real-world applications however, the noise is often colored, its power is unknown, and its level may change with time. For such cases, using the above theoretical approach with fixed thresholding does not provide good results. A solution to this problem is to use an adaptive approach in which the threshold value is obtained directly from the signal spectra under examination. This method is known as Constant False Alarm Rate (CFAR) detection and has been reported extensively in the literature [199, 202, 210-233]. Utilization of the CFAR approach in the frequency domain is very efficient because the ML estimates for the unknown signal parameters may be obtained directly from the FFT spectrum. It should be noted that no information about the target is used in deciding the threshold which means this detector will not have the same detection performance for different target distributions.

To determine whether a signal is present in a given frequency bin, the test cell is isolated, and the noise power is estimated from neighboring bins. Typically, bins immediately next to the test cell are not utilized to prevent spectral leaking from influencing the noise estimate; they are known as guard cells. The detection threshold for a CFAR detector is given by [79]:

$$\eta = \alpha \, \mathrm{P}_N \tag{4.85}$$

where P_N is the noise power estimate, α is a scaling factor also called the threshold factor, and η is the threshold value. The detection decision now becomes:

$$\begin{cases} Choose \ \mathrm{H}_1, \ if \ X_c \geq \eta \\ Choose \ \mathrm{H}_0, \ if \ X_c < \eta \end{cases} \tag{4.86}$$

where X_c indicates the spectral value of the cell under test.

A variety of CFAR detectors have been proposed, each with a slightly different approach to approximating the unknown parameters for the GLRT. Some common forms include the Cell Averaging CFAR (CA-CFAR) [214],

Greatest of Cell Average (GOCA-CFAR) [213], Smallest of Cell Average (SOCA-CFAR) [213], Ordered Statistic (OS-CFAR) [215], Censored Mean Level (CML-CFAR) [217], and the Trimmed Mean (TM-CFAR) [212]. Each or these detectors operate using the same principles, with differences only existing in the method in which the reference noise level is determined.

4.3.1.2.1 - Cell Average CFAR

The CA-CFAR is perhaps the most commonly utilized form due to its simplicity and the fact that it is considered an optimum detector for cases of homogeneous background noise with many reference samples [233, 234]. For this case, the noise estimate P_N is obtained from averaging the reference cells according to:

$$P_N = \frac{1}{N}\sum_{k=1}^{N} X_k \tag{4.87}$$

where N is the total number of frequency bins used for the noise estimate as defined by the set vector $\vec{N}$, and k represents the bin number. Note that $N = \left|\vec{N}\right|$, where $|\ |$ represents the cardinality (set length) rather than Euclidean norm (vector magnitude) of the set $\vec{N}$. For the case of Gaussian noise that has been transformed to the frequency domain via the FFT, the PDF for the noise power can be modeled by the exponential distribution as previously given by equation (2.66). Utilizing (2.66) in conjunction with the sum of random variables transform given by (2.71) and the false alarm threshold defined by (4.79), the following relationship can be obtained [79]:

$$\alpha_{ca} = N\left(P_{fa}^{-1/N} - 1\right) \tag{4.88}$$

where α_{ca} is the scaling factor used to calculate the threshold value according to equation (4.85). Figure 4-4 displayed below provides a pictorial illustration of the CA-CFAR detection scheme.

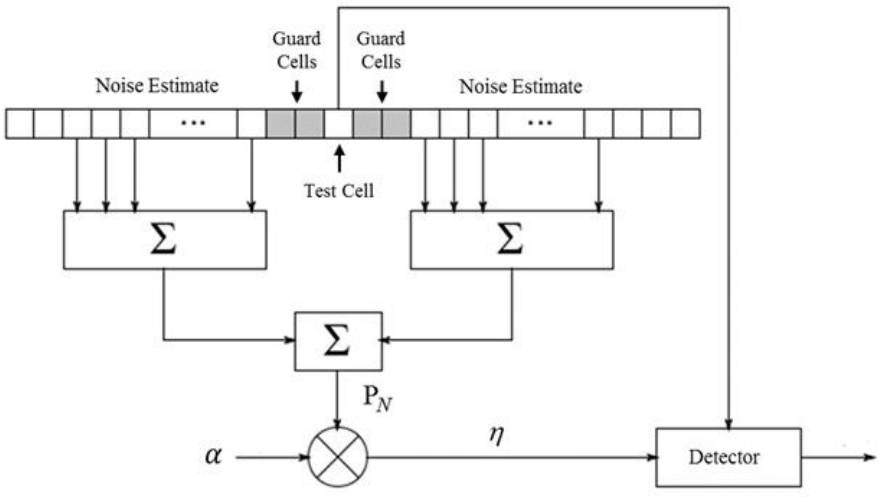

Figure 4-4: General CA-CFAR detector setup

Although the CA-CFAR detector is shown to be optimal for the case of homogeneous background noise, it performs very poorly if the assumption of identical statistics of the reference cells is not valid [212]. For the

case of acoustic detection, this may occur if multiple sources are present with closely spaced fundamental frequencies (target masking), edge clutter or spikes are produced from doppler shifted reflections, there is an outlier due to some impulsive interference or malfunctioning system component, or the noise simply has a non-even (colored) power distribution.

4.3.1.2.2 - Order Statistic CFAR

A commonly employed alternative to the CA-CFAR detector is Order Statistic or Rank Based CFAR (OS-CFAR). Proposed primarily for combating signal masking degradations [79], the OS-CFAR detector offers increased performance over the CA-CFAR for cases of high edge clutter and multiple target environments [213]. Unlike the CA-CFAR, the OS-CFAR does not utilize the noise sample average but rank orders the set defined by $\vec{N}$ in ascending order. The k^{th} element of the ordered list is termed the k^{th} order statistic. The detection threshold is now given by:

$$\eta = \alpha_{os} X_k \tag{4.89}$$

where X_k is the k^{th} order statistic and α_{os} is the scaling factor. For the case of exponentially distributed noise, the average probability of false alarm is now given by [79]:

$$P_{fa} = \frac{N!}{(N-k)!} \frac{\Gamma(\alpha_{os} + N - k + 1)}{\Gamma(\alpha_{os} + N + 1)} \tag{4.90}$$

where $\Gamma(\,)$ represents the Gamma function, N is the total number of noise estimate points, and k is the statistic number. For integer arguments the above from reduces to:

$$P_{fa} = \frac{N(\alpha_{os} + N - k)!}{(N-k)!(\alpha_{os} + N)!} \tag{4.91}$$

Nathanson showed that $k = 0.75N$ provides the best detector performance for most conditions [235]. Unfortunately, the above equation cannot be rearranged for the scaling factor α_{os} and thus numerical approximations must be utilized instead to solve for the desired P_{fa}. For the case of homogeneous noise without edge clutter or interfering targets, the OS-CFAR suffers from detection losses of about 0.3 to 0.5 dB since the threshold value will inherently be larger than that of the CA-CFAR detector [236]. However, for non-homogeneous noise, the OS-CFAR detector will generally perform much better. For the case of multiple interfering targets, the OS-CFAR is almost completely insensitive to masking provided the number of cells contaminated by interfering targets does not exceed $N - k$ [79]. Thus, for I number of interfering targets, the minimum order statistic will be given by:

$$k_{min} = N - I \tag{4.92}$$

4.3.1.2.3 - Distribution-Free CFAR

As previously discussed, the CFAR approach is an effective means of signal detection when little information in known regarding the contaminating noise. However, essentially all of the proposed techniques rely on the assumption that the noise component(s) follow some known probability distribution. In many instances, changing environmental factors modify noise conditions such that the initially assumed distribution is no longer valid, and the current distribution may not be obtainable. For such cases, standard CFAR methods are no longer capable of maintaining a constant false alarm rate and will often degrade performance to an unacceptable level [79, 212]. Moreover, even if the noise distribution is fully known, simple signal enhancement procedures applied prior to detection may modify these distributions to forms which cannot be established without resorting to numerical techniques. Consider for example the CA-CFAR detector previously presented, whose scaling factor was given by equation (4.88). This form is often referred to as a square-law detector since the noise component is squared to obtain the power spectra prior to applying the threshold analysis. For the case of a linear law (magnitude spectra) detector, the noise will exhibit a Rayleigh distribution as previously given by equation (4.80). For this case however, there is no closed-form mathematical expression for α as a function of P_{fa} [218]. Indeed, this is the situation for many types of noise distributions and one of the downfalls of using a parametric or semi-parametric approach [79]. Even for the case of power spectra with exponentially distributed noise, closed-form solutions typically cannot be achieved once enhancement processors such as those presented in the previous section are employed. For these scenarios, only numerical techniques or distribution-free detection methods can effectively be used.

Distribution-free detectors refers to a general class of detection algorithms that do not require prior knowledge or assumptions regarding the signal or noise statistics. This implies that the performance of these detectors is independent of the underlying distributions for type I and type II errors. They are designed to extract information present in observations to account for the gap of missing priori knowledge of the distribution for the null hypothesis. Although distribution-free methods provide obvious benefits over standard detection methods, work in this area has been relatively limited. Recently, Sarma and Tufts developed a DF-CFAR detector based on rank order statistics [237]. In short, the method is essentially that of the OS-CFAR but without the use of a scaling factor. The false alarm probability for this detector is given by the following:

$$P_{fa} = \frac{N+1-k}{N+1} \tag{4.93}$$

where N is the total number of noise estimate points, and k is the statistic number. From the above equation it is evident that the minimum P_{fa} will be achieved when $k = N$. For this case, the detector simply converges to a maximum value comparator, where the cell under test must be greater than all other cells across some bandwidth of interest. For a desired false alarm probability $\hat{P}_{fa}$, the required order statistic is thus given by:

$$k = \left\lceil (N+1)(1-\hat{P}_{fa}) \right\rceil \tag{4.94}$$

where ⌈ ⌉ indicates rounding up to the nearest whole number. The detection decision is now made according to:

$$\begin{cases} Choose\ \mathrm{H}_1,\ if\ \ X_c \geq X_k \\ Choose\ \mathrm{H}_0,\ if\ \ X_c < X_k \end{cases}$$

where X_c indicates the spectral value of test cell, and k is typically chosen to satisfy $N/2 < k < N$. One of the obvious downsides to this detector is that false alarm probabilities depend on the noise reference sample size, which may prevent the use of this approach in some systems. However, for applications which can facilitate the use of this detector, the performance is generally very good with detection losses on the order of 0.4 dB compared to an optimal detector in an homogeneous environment [237]. In addition, the P_{fa} can be further reduced if desired through Binary Integration as will be discussed.

4.3.1.3 - *Binary Integration*

To improve the overall reliability of detection, it may be required that a target be detected D times out of T trials before it is finally accepted as being valid. This process is called Binary Integration and it may be utilized to effectively increase the probability of detection P_d while simultaneously reducing the probability of false alarm P_{fa} [79]. If we assume P_d remains constant for each of the T threshold tests, then the probability of not detecting the target (probability of a miss P_m) on one trial is $1 - P_d$. If there are T independent trials, the probability of missing the target on all T trials is $(1 - P_d)^T$. Thus, the probability of detecting the target on at least one of the T trials, denoted as the cumulative probability of detection P_D is:

$$P_D = 1 - \left(1 - P_d\right)^T \tag{4.95}$$

where P_d is the detection probability for a single trial. Similarly, the cumulative probability of false alarm P_{FA} also follows this relationship:

$$P_{FA} = 1 - \left(1 - P_{fa}\right)^T \tag{4.96}$$

For the case of D detections out of T trials, the cumulative probability of detection or false alarm is given by the following [79]:

$$P = \sum_{r=D}^{T} \frac{\left(1-p\right)^{T-r}\, p^r T!}{\left(T-r\right)! r!} \tag{4.97}$$

where p is the probability of detection P_d or false alarm P_{fa} for a single independent test. Consider for example an application in which we require $P_D \geq 0.99$ and $P_{FA} \leq 1 \times 10^{-4}$. However, our detector will only achieve $P_d = 0.96$ and $P_{fa} = 0.01$ for a single test. If time constraints can facilitate multiple tests before making a decision, we may satisfy our detection requirements by requiring D detections out of T trials. For example, using three detections out of four tests $(D = 3, T = 4)$ now gives $P_D = 0.991$ and $P_{FA} = 4 \times 10^{-6}$.

4.3.2 - Detection Requirements

In order to develop an effective collision avoidance system, metrics must first be obtained regarding the statistical and kinematic properties associated with mid-air encounters. Thus, the following section provides an overview of aircraft close encounter statistics in order to determine appropriate detection and false alarm requirements. In addition, kinematic properties, which are also required to determine minimum detection distances required to facilitate an avoidance maneuver, are also analyzed.

4.3.2.1 - *Statistical Considerations*

Compared to manned aircraft operations, UAVs are considered to constitute a much higher risk for a mid-air collision. This is due to a number of factors:

1) They do not have a pilot with visual capabilities to act as a last line of defence.
2) They are not equipped with any form of communication equipment to broadcast location information to other neighbouring aircraft or air traffic control.
3) Their small size would prevent the pilot of a manned aircraft from obtaining visual sight at distances adequate to make an avoidance maneuver.

Due to UAVs being a relatively new technology in the field of aviation, little prior analysis has been conducted to determine the detection requirements of a collision avoidance system. Furthermore, relatively little statistics exist regarding current midair collision risks in Canadian airspace for either manned or unmanned aircraft to aid in the establishment of these requirements. Recently, Stevenson provided an analysis of current operational risks and mitigation strategies for UAVs operating in Canadian airspace [238]. He proposed that the probability of a mid-air collision between a manned aircraft and UAV can be approximated by:

$$P_C = P_E \cdot P_S \cdot (1 - P_{UAV}) \cdot (1 - P_{MAN}) \tag{4.98}$$

where P_E is the probability of a mid-air encounter that may lead to a collision, P_S is the separation loss probability due to a failure or absence of industry safeguards (failed ATC advisories, failed radio procedures, etc.), P_{UAV} is the probability that the UAV will detect the manned aircraft using some form of non-cooperative collision avoidance system, and P_{MAN} is the probability that the pilot of the manned aircraft will visually detect the UAV.

Rearranging the above equation, we obtain the following expression for the detection requirements for a non-cooperative UAV collision avoidance system:

$$P_{UAV} = \frac{P_C - P_E P_S + P_E P_S P_{MAN}}{P_E P_S P_{MAN} - P_E P_S} \tag{4.99}$$

where P_c probability for a collision between two manned aircraft.

If we assume the worst-case scenario in which the pilot will not see the UAV, and no safeguards are in place to provide notification of shared airspace, we obtain $P_S = 1$ and $P_{MAN} = 0$. Thus, the above equation reduces to:

$$P_{UAV} = \frac{P_E - P_C}{P_E} \tag{4.100}$$

In order to maintain an effective system, detection probabilities should be high enough to achieve a collision risk equivalent to that for manned aircraft. Currently, the generally accepted risk probability for a collision between two manned aircraft is $P_C = 1 \times 10^{-7}$, while the maximum expected probability that a mid-air encounter could occur given current air traffic levels is $P_E = 2.36 \times 10^{-5}$ [238]. Utilizing these values in conjunction with equation (4.100), we obtain a detection requirement of $P_{UAV} = 0.995$. Thus, in order to develop a successful UAV based SAA system which maintains current aviation safety standards, the system must be capable of correctly detecting the presence of another aircraft on a potential collision course 99.5% of the time. Although this probability is certainly achievable, it must also be realizable within the constraints of acceptable false alarm levels since the two are positively correlated.

Unlike detection probabilities, no clear requirement exists for acceptable false alarm rates. This is because requirements are generally subjective and depend on factors such as:

1) The outcome once a false detection and subsequent avoidance maneuver is performed (can the mission continue, does operator require absolute verification of incident, etc.).
2) Expected mission duration.
3) Rate at which detection decisions are made in conjunction with expected flight duration.

For sake of simplicity, we will ignore the first factor and instead assume that the operator simply specifies the average operational time between false alarms. Thus, with respect to the above design requirements, we obtain the following expression for the desired system false alarm rate:

$$P_{FA_{req}} = \frac{t_{scan}}{t_{fa}} \tag{4.101}$$

where t_{scan} is the scan time required before a detection decision is made, and t_{fa} is the amount of time desired between false alarms. For example, consider the hypothetical case of a UAV performing some surveying operation. A typical mission lasts approximately 30 minutes and the detection system completes a threat analysis each second during flight operations. If it is acceptable that on average one false alarm will occur every two missions, the required false alarm rate will be: $P_{FA_{req}} = \frac{1}{2 \cdot 60 \cdot 30} = 2.8 \times 10^{-4}$. Thus, parameters of the detector such as noise sample size N and order statistic number k should be chosen based on this value.

4.3.2.2 - *Kinematic Considerations*

In order to construct a viable acoustic-based SAA system, detection distances must be large enough to facilitate an avoidance maneuver. Currently no official requirements exist but it is generally accepted that 500 ft (152.4 m) is the absolute minimum that should be maintained between two aircraft at all times [2]. Direct analysis of the governing kinematic equations for two aircraft on a potential collision course leads to a series of equations

which cannot be solved directly using analytical methods. Geyer produced a solution for the direct head-on case by approximating the avoidance maneuver via a Taylor series expansion producing the following equation [2]:

$$d_o \approx 5.6 v_{rel} \sqrt{\frac{\pi}{2} - \phi} \tag{4.102}$$

where d_o is the minimum distance to initiate the avoidance maneuver to avoid a collision, v_{rel} is the relative approaching speed of the two aircraft, and ϕ is the bank (roll) angle initiated by the sensing aircraft upon detection. However, the approximation model is not appropriate for the small distances and velocities associated with UAV operations. For example, consider two aircraft with a closing velocity of $v_{rel} = 50$ knots, where a bank angle of $\phi = \pi/4$ rads is initiated by the sensing aircraft. Equation (4.102) gives a minimum detection distance of 127.6 m, which is less than the minimum accepted threshold of 152.4 m. Since the model proposed by Geyer is not appropriate for work pertaining to this thesis, a simplistic analytical two-dimensional model is presented that is more appropriate for UAV operations.

If we assume the detecting aircraft is small and travels at a relatively low speed (such as that with most UAVs), we may construct an accurate and analytically solvable model which accounts for scenarios not applicable by Geyer's model. Consider the case for a detecting and intruding aircraft both having constant speeds and headings of $|\vec{v}_\alpha|, \alpha$ and $|\vec{v}_\beta|, \beta$ respectively. Given the initial angular position γ of the intruding aircraft with respect to the detecting aircraft, we wish to find the minimum distance $|\vec{d}_o|$ which can facilitate an avoidance maneuver such that the separation range never falls below the collision boundary radius r_b. We assume that once the intruding aircraft is discovered, the detecting aircraft immediately initiates an instantaneous bank (roll) angle ϕ, and holds that angle until it reaches a new heading that is perpendicular to and away from the intruding aircraft flight path. Figure 4-5 displayed below provides a visual depiction of the system kinematic configuration. For sake of simplicity, the system is only defined in terms of two dimensions. Although the actual system is three dimensional, the 2-D model can provide a good approximation if the 3-D speed and heading values are projected on a 2-D plane connecting the two aircraft. In this regard, the 2-D model will produce a more conservative estimate, since one spatial degree of freedom is removed which may be used to perform the avoidance maneuver.

For the collision avoidance maneuver, there are three possible paths which can be taken by the detecting aircraft:

1) Head perpendicularly away from the intruder path by making the minimum bearing change.
2) Head perpendicularly away from the intruder path by making the maximum bearing change.
3) Adjust course to perpendicularly cross the intruder path.

Each of the cases are depicted below in Figure 4-6. The choice between Cases 1 and 2 may seem trivial at first but depending on the aircraft headings, choosing the minimum bearing change may steer the detecting aircraft towards the intruder rather than away. When considering the actual kinematic and geometric configuration however, this may still be the best course of action. The choice of which avoidance course to take for any given

scenario requires a more complex analysis that is outside the scope of this thesis. To simplify the analysis, we thus assume the aircraft adheres to avoidance Case 1.

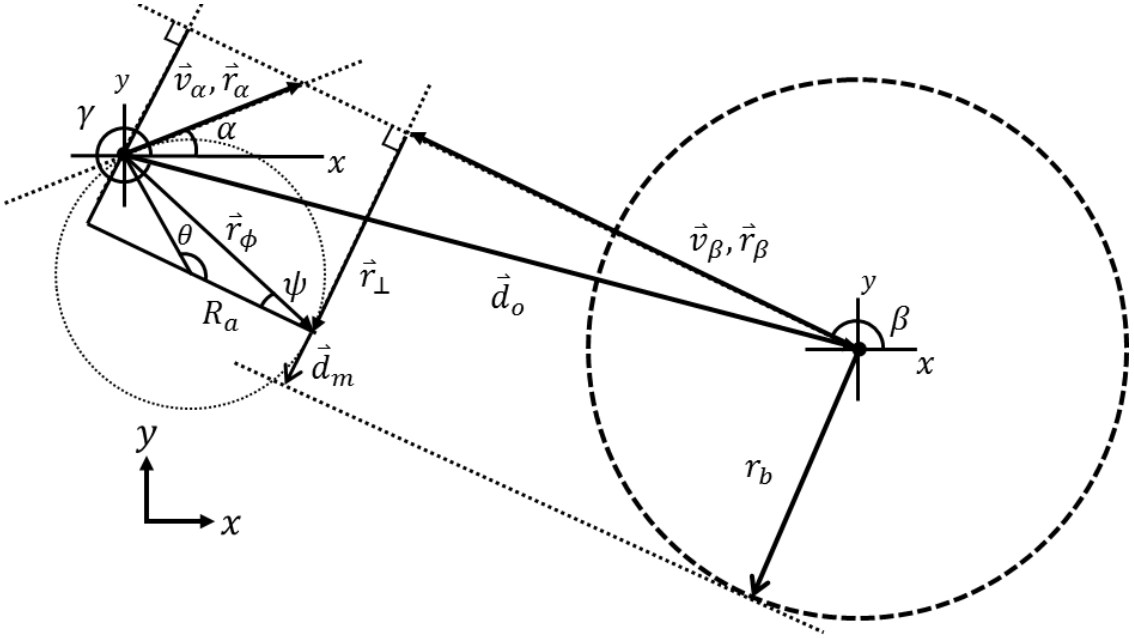

Figure 4-5: Kinematic illustration of two aircraft on a collision course.

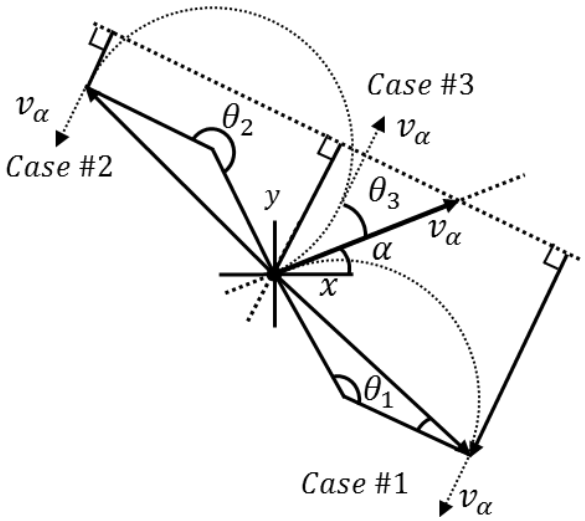

Figure 4-6: Depiction of various avoidance maneuvers.

As previously stated, upon detection, the sensing aircraft initiates an instantaneous bank angle ϕ which produces a turning radius R_a (assuming constant altitude) given by:

$$R_a = \frac{v_\alpha^2}{\tan(\phi) g} \tag{4.103}$$

Once the heading change is made, the sensing aircraft must then travel an additional distance $|\vec{d}_m|$ to avoid a collision:

$$d_m = r_b - |\vec{r}_\phi|\sin(\psi) - |\vec{r}_\alpha||\hat{r}_\alpha \bullet \hat{r}_\perp| \tag{4.104}$$

The following relationships can be obtained via geometry as illustrated in Figure 4-5.

$$\theta = \cos^{-1}\left(\hat{r}_\alpha \bullet \hat{r}_\perp\right) \tag{4.105}$$

$$\psi = \left(\frac{\pi - \theta}{2}\right) \tag{4.106}$$

$$|\vec{r}_\phi| = \sqrt{2R_a^2\left(1 - \cos[\theta]\right)} \tag{4.107}$$

$$|\vec{r}_\alpha| = -|\vec{d}_o|\csc(\alpha - \beta)\sin(\beta - \gamma) \tag{4.108}$$

where $\hat{r}_\perp$ is the direction perpendicular to $\hat{r}_\beta$ that minimizes θ as given by the following expression:

$$\hat{r}_\perp = \left\langle -\hat{r}_{\beta,y}\ , \hat{r}_{\beta,x} \right\rangle \vee \left\langle \hat{r}_{\beta,y}\ , -\hat{r}_{\beta,x} \right\rangle \tag{4.109}$$

The total time required to execute the heading change and travel the remaining distance $|\vec{d}_m|$ to reach a final perpendicular distance r_b away from the intruder path is given by:

$$t_{avoid} = \frac{R\theta + |\vec{d}_m|}{|\vec{v}_\alpha|} \tag{4.110}$$

During this time, the intruder will advance a distance $|\vec{r}_\beta|$ given by:

$$|\vec{r}_\beta| = |\vec{v}_\beta| t_{avoid} \tag{4.111}$$

Thus, the initial radial separation distance required is given by:

$$|\vec{d}_o| = \frac{|\vec{r}_\beta| - \hat{r}_\alpha \bullet \hat{r}_\beta |\vec{r}_\phi|\cos(\psi)}{|\hat{d}_o \bullet \hat{r}_\beta|} \tag{4.112}$$

Combining equations (4.104), (4.106), (4.108), (4.110) to (4.112), and solving for $|\vec{d}_o|$gives the following final form:

$$|\vec{d}_o| = \frac{|\vec{v}_\beta|\left(r_b + R_a\theta - |\vec{r}_\phi|\cos[\theta/2]\right) + |\vec{r}_\phi||\vec{v}_\alpha|\left[\hat{r}_\alpha \bullet \hat{r}_\beta\right]\sin[\theta/2]}{|\vec{v}_\alpha|\left|\hat{d}_o \bullet \hat{r}_\beta\right| - |\vec{v}_\beta||\hat{r}_\alpha \bullet \hat{r}_\perp|\csc[\alpha - \beta]\sin[\beta - \gamma]} \tag{4.113}$$

For the case of a head on collision course, the above equation reduces to:

$$|\vec{d}_o| = \frac{r_b|\vec{v}_\beta|}{|\vec{v}_\alpha|} + \frac{|\vec{v}_\alpha|\left(2|\vec{v}_\alpha| + |\vec{v}_\beta|[\pi - 2]\right)\cot[\phi]}{2g} \tag{4.114}$$

Utilizing equation (4.113) in conjunction with typical cruising speeds for various aircraft, a list of approximate detection distances required for various heading scenarios has been constructed and is displayed below in Table 4-4. For each of the cases, it is assumed that the initial angular position between the two aircraft is zero ($\gamma = 0$),

the avoiding aircraft initiates an instantaneous bank angle of $\phi = \pi/4$ while travelling at 30 knots, and the collision boundary distance is $r_b = 152.4$ m.

Table 4-4: Minimum required detection distances for various aircraft.

Intruder Aircraft	Intruder Speed (knots)	d_o (m) ($\beta = 180,\ \gamma = 0$)		
		$\alpha = 0$	$\alpha = 45$	$\alpha = 90$
GBS (UAV)	35	218	185	178
Cessna 185	95	551	494	483
Bell 206	115	662	600	584
Sikorsky S-92	140	795	725	711
Boeing 737	250*	1,402	1,290	1,270

* Maximum suggested airspeed from takeoff to 10,000 ft.

4.3.3 - Modified Order-Statistic Forms

The OS-CFAR and DF-CFAR detectors previously presented both use order statistics to dictate false alarm rates. The order statistic k was defined as the k^{th} largest element contained in the noise sample set $\vec{N}$ when sorted in ascending order. Although generally accepted, this approach is found to be less intuitive and less convenient for the DF-CFAR detector. This is because any modifications to N requires the value of k to inherently change in order to maintain uniformity when calculating the false alarm rate using equation (4.93). A more convenient form is to define the order statistic in terms of the descending values of $\vec{N}$ such that k does not require recalculation if N changes. To avoid confusion, the reversed order statistic will be indicated by $\bar{k}$. That is, the $\bar{k}^{th}$ element is the $\bar{k}^{th}$ largest value of $\vec{N}$ giving $\bar{k} = N - k + 1$. Thus, the false alarm rate for a single cell test using the DF-CFAR detector may be rewritten as:

$$P_{fa} = \frac{\bar{k}}{N+1} \tag{4.115}$$

which is obviously a more simplistic expression than that previously given. From the above equation, it is now evident that the minimum P_{fa} will be achieved when $\bar{k} = 1$ (max peak detector). For the case of I interfering targets, the minimum order statistic will now be given by $\bar{k}_{min} = I$.

If G number of guard cells are placed about the test cell with I interfering targets present, the false alarm rate will now be given by the following:

$$P_{fa} = \frac{\bar{k} + I}{N - G + 1} \tag{4.116}$$

4.3.4 - Harmonically Transformed Spectra

The DF-CFAR was previously presented as a method to detect unknown signals in noise of unknown statistical properties. Unlike other CFAR detectors such as the CA-CFAR and OS-CFAR, the DF-CFAR does not require noise to exhibit an exponential distribution in order to predetermine the parameters required to achieved some desired false alarm rate; the detector only requires noise components to exhibit equivalent distributions. Such a

property offers a very significant advantage since the use of any signal enhancement processors will effectively transform the underlying statistical distribution. However, the performance of the detector is not completely impervious to the application of signal enhancement processors. For such instances, modifications to the original developed form may be required to maximize operational performance. Such a case would include the use of the Harmonic Spectral Transforms (HSTs) previously presented. In brief, HSTs convert the spectra of harmonic signals such that the resultant form becomes a function of fundamental frequency. However, the transform process also generates signal components which are fractional values of the fundamental peak at various locations along the spectrum where previously only noise would have resided. Failure to exclude these values from the noise sample vector $\vec{N}$ used to determine the test statistic will thus reduce detection performance to some degree, and/or produce inaccurate false alarm rates.

Figure 4-7 displayed below provides spectral plots of the FFT magnitude and standard mean HST ($\overline{\mathrm{H}}_1$) for a signal containing 4 harmonic components. From the plots, it is evident that the HST spectrum achieves a maximum at the fundamental frequency, since all harmonic components are directly combined. It also contains smaller peaks located at frequencies which are fractions of one or more of the harmonic components. The location of all peaks (fractional and fundamental) will be contained in the set generated by:

$$f_{a,b} = \frac{af_0}{b} \tag{4.117}$$

where $a, b \in \{1,2,\ldots,R\}$, f_0 is the fundamental frequency value, and R is the total number of harmonics. The number of fractional peaks will be given by the number of unique values (excluding the fundamental) contained in the set defined by $f_{a,b}$ up to the maximum value of $2f_0$. It should be noted that the above expression provides the maximum possible number of fractional peaks that can be present. However, this is not necessarily the number that will effectively be present since peak values are influenced by spectral resolution and signal frequency values relative to this resolution. For instance, a spectrum of a 100 Hz 4 harmonic signal with a 1 Hz/bin resolution will not contain a distinct and significant fractional peak value at the 100/3 Hz position as depicted below in Figure 4-7. Table 4-5 provides the number of fractional peaks for a given number of signal harmonics.

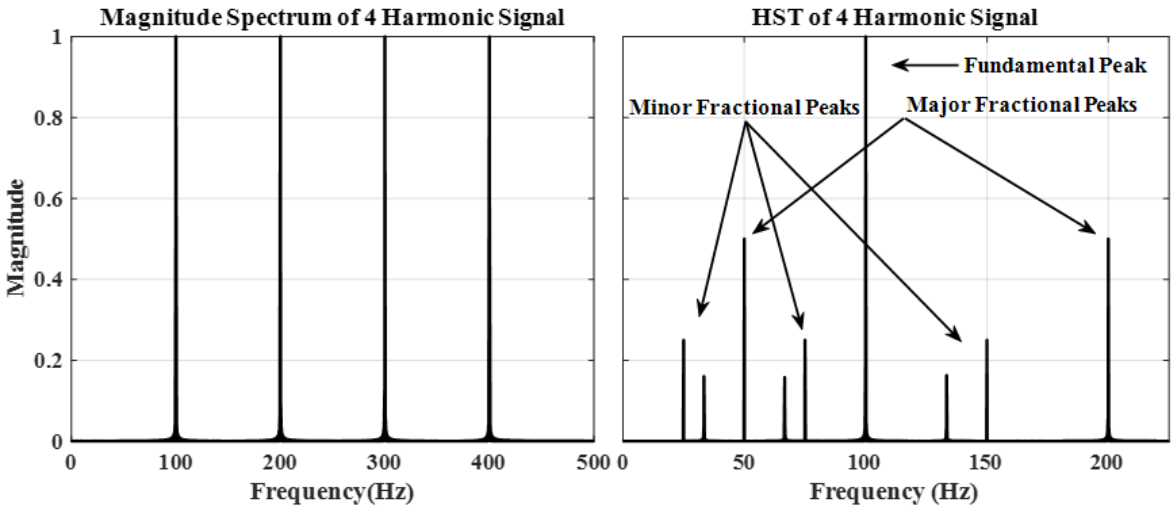

Figure 4-7: Plots of standard and harmonically transformed magnitude spectra.

Table 4-5: Maximum number of fractional peaks present for HSTs of length $2f_0$.

Harmonics (R)	Fractional Peaks (F)
2	2
3	5
4	8
5	14
6	17
7	26
8	32

Since the DF-CFAR requires all noise components to exhibit equivalent distribution types in order to maintain a constant false alarm rate, fractional peak components must be excluded from the noise estimate. One may be tempted to treat the fractional components as simply target interference, in which case the false alarm rate would be given by equation (4.116). However, doing so would produce unnecessarily high false alarm rates. Because peak locations may be calculated for a given test cell, they may instead be omitted from the noise estimate altogether. If F number of fractional peaks are omitted, the false alarm rate will now be given by:

$$P_{fa} = \frac{\bar{k} + I}{N - F - G + 1} \tag{4.118}$$

If desired, guard cells may also be placed around the fractional peak locations to minimize the effects of spectral leakage. If a total of G_F guard cells are used for each fractional peak, then the false alarm probability will now be given by:

$$P_{fa} = \frac{\bar{k} + I}{N - F(G_F + 1) - G + 1} \tag{4.119}$$

If the noise estimate is taken symmetrically about the fundamental component frequency f_0 and spans a total bandwidth of $2f_0$, then $2f_o/f_{res}$ samples will be contained in $\vec{N}$, where f_{res} is the frequency resolution of the HST spectrum. Thus, the above form may be rewritten as follows:

$$P_{fa} = \frac{\bar{k} + I}{2 f_0 / f_{res} - F(G_F + 1) - G + 1} \tag{4.120}$$

4.3.5 - Multiple Cell Testing

The following section presents a number of signal detection schemes which may be employed to reduce false alarm rates and/or increase detection probabilities. Two main scenarios are examined: 1) unconstrained independent cell testing, and 2) constrained non-independent cell testing. For each case, the issue of increased false alarm rates from testing multiple frequency locations is addressed, along with a number of methods which may be used to alleviate the problem. These include location dependent binary integration, and a more robust frequency tracking approach. The validity of all presented test schemes was confirmed through Monte Carlo simulations using trial numbers on the order of 1×10^7 runs.

4.3.5.1 - *Unconstrained Independent Events*

4.3.5.1.1 - Cumulative Probabilities

The false alarm probabilities previously presented for the CA-CFAR, OS-CFAR, and DF-CFAR detectors are only relevant for the case of one test location per trial. This is valid if the target signal frequency is known in advance; for which case we need only test the relevant frequency bin. However, if the target frequency is unknown then all possible (or expected) frequency bins must be evaluated. Such a test scheme is said to be unconstrained if the noise estimate can be taken freely about each test cell, and independent if the false alarm rate for any given cell is not affected by the result of any other. This scenario is depicted below in Figure 4-8 where the test band is defined by $\vec{B}$ and consists of B test cells, while the noise band for the b^{th} test cell is defined by $\vec{N}_b$ and consists of N_B reference cells. Since $\vec{N}_b$ contains the b^{th} test cell given by X_b, the total number of noise samples for a given test statistic will be $N_b = \left|\vec{N}_B\right| - 1$, where $|\ |$ represents the cardinality (set size) rather than Euclidean norm (vector magnitude) of the set $\vec{N}_b$.

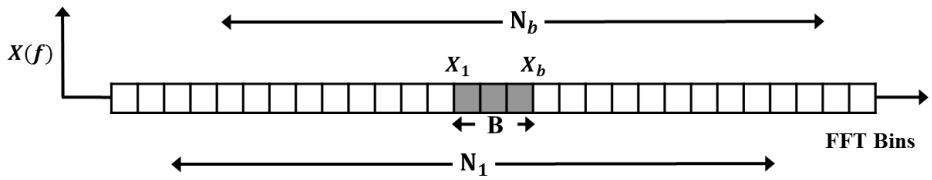

Figure 4-8: Unconstrained multiple cell testing.

If test B cells are evaluated and the tests are independent, the cumulative probability of at least one false alarm occurring will be given by:

$$P_{FA} = 1 - (1 - P_{fa})^B \tag{4.121}$$

where P_{fa} is the false alarm rate for a single cell which may be established using the CA-CFAR, OS-CFAR, or DF-CFAR detectors previously described.

If multiple realizations (windowed FFT spectra) are used to establish the test statistic, Binary Integration may be applied across windows for each tested cell to reduce false alarm rates. If T trials are performed for each cell, the total probability that at least one of the B cells will achieve a false detection on all T trials will be given by:

$$P_{FA} = 1-(1-P_{fa}^{T})^{B} \tag{4.122}$$

If D detections are required out of T trials before accepting any hypothesis, the total false alarm rate will be:

$$P_{FA} = 1-(1-P)^{B} = 1-\left(1-\sum_{r=D}^{T}\frac{\left(1-P_{fa}\right)^{T-r}P_{fa}{}^{r}T!}{(T-r)!r!}\right)^{B} \tag{4.123}$$

4.3.5.1.2 - Robust Binary Integration

Although effective in decreasing false alarm rates, one of the downsides to the above Binary Integration approach is the requirement of a stationary signal across all trials. However, many real-world signals do not adhere to these conditions, since factors such as Doppler shifting and spectral leakage may easily cause peak locations to deviate or fluctuate over time. To facilitate this deviation, a Robust Binary Integration method is proposed which can reduce false alarm rates through effectively tracking signal locations while also enabling frequency shifting.

Consider the case where a frequency band consisting of B test cells is evaluated across T trials. The probability of a false alarm for any given cell across all trials is given by:

$$P_{FA} = P_{fa}^{T} \tag{4.124}$$

which can be obtained from equation (4.97) when $D = T$.

If the location of subsequent detections are allowed to deviate by Δ cells in either direction with respect to the initial detection location, the cumulative false alarm probability for a single cell across T trials will be given by the following expression:

$$\bar{P}_{fa} = P_{fa}\left[1-\left(1-P_{fa}\right)^{2\Delta+1}\right]^{T-1} \tag{4.125}$$

Thus, the total probability that at least one false alarm will occur from testing all B test cells adhering to the above condition can be approximated as follows:

$$P_{FA} \approx 1-\left(1-\bar{P}_{fa}\right)^{B} = 1-\left(1-P_{fa}\left[1-\left(1-P_{fa}\right)^{2\Delta+1}\right]^{T-1}\right)^{B} \tag{4.126}$$

To obtain an exact expression, deviation constraints for cells near the test band edge must also be considered. Since cells outside the test band are not evaluated, deviation positions are constrained such that they remain inside the band as depicted below in Figure 4-9.

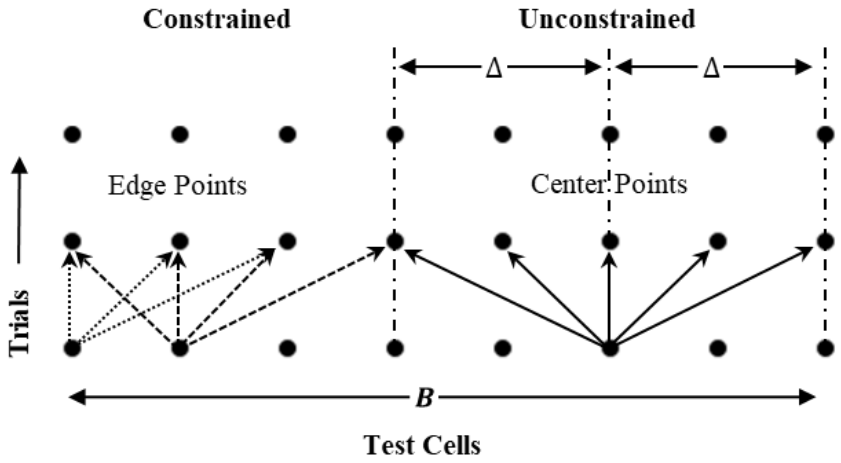

Figure 4-9: Depiction of constrained and unconstrained detection location deviation.

With respect to the above figure, it is evident that the probability of false alarm for unconstrained center points is given by:

$$P_c = P_{fa}\left[1-\left(1-P_{fa}\right)^{2\Delta+1}\right]^{T-1} \tag{4.127}$$

while the probability of false alarm for the n^{th} cell from the test band edge will be

$$P_n = P_{fa}\left[1-\left(1-P_{fa}\right)^{\Delta+n}\right]^{T-1} \tag{4.128}$$

where $n \in \{1,2,\ldots\Delta\}$. The probability that there will be at least one false detection at the $B-2\Delta$ center cells will be:

$$\bar{P}_c = 1-\left(1-P_c\right)^{B-2\Delta} = 1-\left(1-P_{fa}\left[1-\left(1-P_{fa}\right)^{2\Delta+1}\right]^{T-1}\right)^{B-2\Delta} \tag{4.129}$$

while the probability that there will be at least one false detection for the n^{th} edge cell(s) will be:

$$\bar{P}_n = 1-\prod_{n=1}^{\Delta}\left[\left(1-P_n\right)^2\right] = 1-\prod_{n=1}^{\Delta}\left[1-\left(1-P_{fa}\left[1-\left(1-P_{fa}\right)^{\Delta+n}\right]^{T-1}\right)^2\right] \tag{4.130}$$

Thus, the total probability that there will be at least one detection at either the 2Δ edge cells or $B-2\Delta$ center cells will be given by:

$$\begin{aligned} P_{FA} &= \Pr[\bar{P}_c \cup \bar{P}_1 \cup \ldots \cup \bar{P}_\Delta] \\ &= \Pr[\bar{P}_c \bigcup_{n=1}^{\Delta} \bar{P}_n] \end{aligned} \tag{4.131}$$

Assuming each of the events are independent, the above form can be expressed using DeMorgan's laws according to:

$$\Pr\left[\bar{P}_c \cup \bar{P}_1 \cup \ldots \cup \bar{P}_\Delta\right] = 1-\Pr\left[\left(\bar{P}_c \cap \bar{P}_1 \cap \ldots \cap \bar{P}_\Delta\right)^{\complement}\right] = 1-\Pr\left[\bar{P}_c^{\complement} \cap \bar{P}_1^{\complement} \cap \ldots \cap \bar{P}_\Delta^{\complement}\right] \tag{4.132}$$

where $\complement$ indicates the set complement. Thus, the final probability that there will be at least one false detection in B test cells across T trials within the deviation bounds defined by Δ is given by:

$$P_{FA} = 1 - \left(1 - \bar{P}_c\right)\left(1 - \bar{P}_n\right) \tag{4.133}$$

where $\bar{P}_c$ and $\bar{P}_n$ are given by equations (4.129) and (4.130) respectively.

4.3.5.2 - *Constrained Non-Independent Events*

One of the properties associated with the DF-CFAR detector that has not been addressed in the literature is the effect of non-independent tests on total false alarm rates when performing multiple cell testing. It was previously shown that the total probability for a given trial consisting of B test cells can be obtained via equation (4.121) for the case of independent events. In reality however, multiple cell testing using CFAR detectors rarely constitute true independent events. This is simply because testing adjacent cells requires previous and future test cells be included in the current noise estimate. For instances in which $N \gg B$, this effect can often be neglected, and the independence assumption provides an accurate approximation. However, if N and B are comparable in size this effect cannot be ignored.

Consider for example Figure 4-8 previously displayed, which depicts an unconstrained multiple cell testing scenario. It is assumed that each of the test cells in B are adjacent and evaluated in order ($X_1 \rightarrow X_B$), and the DF-CFAR order statistic is $\bar{k} = 1$ (max peak detector). Since the noise estimate is taken about the current cell, all past and upcoming test cells will also be included in this estimate provided $\vec{B}$ is a subset of $\vec{N}$ ($\vec{B} \subset \vec{N}$). If a detection is obtained at the n^{th} cell where $n \in \{1,2, \ldots B\}$, then the probability that there will be a detection at the next cell $(n+1)$ will be zero, since the current cell has already been found to be greater than all noise samples which includes all other test cells. Thus, it is evident that the outcome of each test is no longer independent of one another. An obvious solution to this problem is to simply exclude all test cells from the noise estimate. However, for instances in which source frequencies are unknown, this may result in excluding substantial portions of the signal bandwidth from the noise estimate; and since false alarm rates for the DF-CFAR detector are based on the number of noise samples used, this would greatly reduce the overall detector performance.

4.3.5.2.1 - Constrained DF-CFAR

With respect to the example outlined above, it is evident that the expressions previously developed to establish cumulative false alarm probabilities are no longer valid since tests are no longer independent. The following section thus provides an analysis for the constrained non-independent case using the DF-CFAR detection model. This model will be referred to as the Constrained Distribution Free CFAR (CDF-CFAR) detector from this point onward.

Consider the case for a constrained noise set such that $\vec{N}$ remains fixed and $\vec{B}$ is also a subset of $\vec{N}$ ($\vec{B} \subseteq \vec{N}$). Such a scenario would be inherently present if testing at the edge of a spectrum, or if all cells contained in the spectrum are used for the noise estimate as depicted below in Figure 4-10.

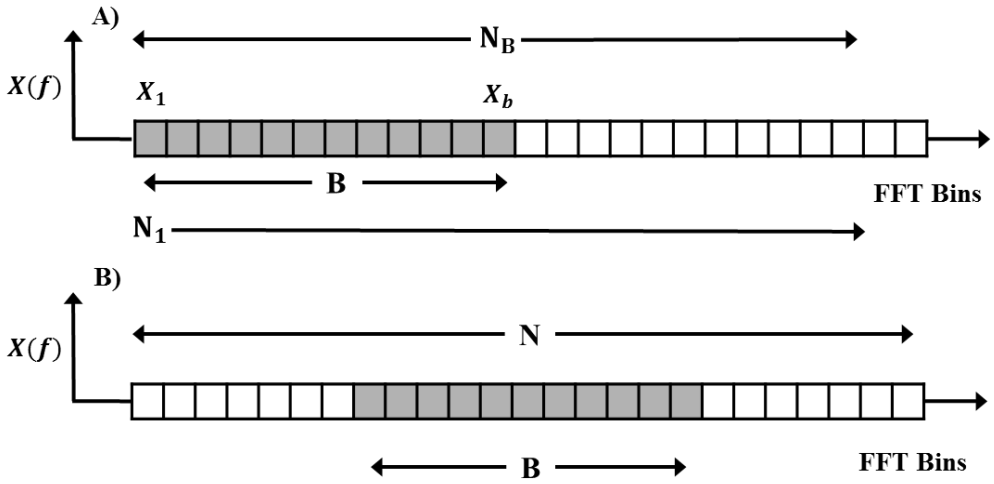

Figure 4-10: Constrained noise sets from: A) Testing at edge of spectrum, B) Using full spectrum as noise estimate.

The probability of at least one false detection for the B tests will be given by:

$$P_{FA} = \Pr[E_1 \cup E_2 \cup ... \cup E_B] = \Pr\left[\bigcup_{i=1}^{B} E_i\right] \tag{4.134}$$

where E_i is the detection event for the i^{th} test cell. Since the events are not independent, DeMorgan's laws cannot be applied. Using the principle of inclusion and exclusion instead gives:

$$\Pr\left[\bigcup_{i=1}^{B} E_i\right] = \sum_{i=1}^{B} \Pr[E_i] - \sum_{i=1}^{B-1}\sum_{j=i+1}^{B} \Pr[E_i \cap E_j] + \sum_{i=1}^{B-2}\sum_{j=i+1}^{B-1}\sum_{k=j+1}^{B} \Pr[E_i \cap E_j \cap E_k] - ..... \tag{4.135}$$

where the union of events can be obtained using the conditional probability formula:

$$\Pr[E_i \cap E_j] = \Pr[E_i]\Pr[E_j | E_i] \tag{4.136}$$

which can be extended to B number of events using the product rule:

$$\Pr\left[\bigcap_{i=1}^{B} E_i\right] = \Pr[E_1] \cdot \Pr[E_2 | E_1] \cdot \Pr[E_3 | E_1 \cap E_2] ... \Pr\left[E_B \middle| \bigcap_{i=1}^{B-1} E_i\right] \tag{4.137}$$

For the DF-CFAR detector, the probability for the first detection will be:

$$\Pr[E_1] = \frac{\bar{k}}{N+1} \tag{4.138}$$

while the probability of a detection in the second cell given one occurred in the first will be:

$$\Pr[E_2 | E_1] = \frac{\bar{k}-1}{N+1-1} = \frac{\bar{k}-1}{N} \tag{4.139}$$

If one continues the approach for $\Pr[E_3 | E_2]$, $\Pr[E_4 | E_3]$ and so on, the following general expression will be obtained:

$$\Pr\left[E_B \middle| E_n\right] = \frac{\bar{k} - n + 1}{N - n + 2} \tag{4.140}$$

It can be shown that the following relationship exists between the inclusion / exclusion functional form and the Binomial distribution:

$$\sum_{m=1}^{B}\left[(-1)^{(m+1)}\binom{B}{m}\prod_{n=1}^{m}a\right] = \sum_{i=1}^{B}a - \sum_{i=1}^{B-1}\sum_{j=i+1}^{B}a^2 + \sum_{i=1}^{B-2}\sum_{j=i+1}^{B-1}\sum_{k=j+1}^{B}a^3 - \sum_{i=1}^{B-3}\sum_{j=i+1}^{B-2}\sum_{k=j+1}^{B-1}\sum_{z=k+1}^{B}a^4 + ... \tag{4.141}$$

Substitution of equations (4.137), (4.138), and (4.140) into (4.135), noting the above relationship, and performing algebraic manipulation finally gives the following expression for the total probability of at least one false detection:

$$P_{FA} = \sum_{m=1}^{B}\left[(-1)^{(m+1)}\binom{B}{m}\prod_{n=1}^{m}\frac{\bar{k} - n + 1}{N - n + 2}\right] \tag{4.142}$$

From the above equation, it is apparent that $P_{FA} \to 1$ as $B \to N$. Evidently, the above equation may also be expressed in terms of the ordinary Hypergeometric function, which is defined according to [239] :

$${}_2F_1(a;b;c;z) = \sum_{n=0}^{\infty}\frac{(a)_n(b)_n}{(c)_n}\frac{z^n}{n!} \tag{4.143}$$

where the false alarm rate will now be given by:

$$P_{FA} = 1 - {}_2F_1(a,b;c;z) \tag{4.144}$$

with $a = -B,\ b = \bar{k},\ c = N + 1$, and $z = 1$ giving the form equivalent to (4.142).

As with the unconstrained independent case, false alarm rates may be further decreased by applying Binary Integration across multiple trials. If D detections out of T trials are required for each test cell, the total probability of at least one false detection from evaluating B cells will be thus given by the following:

$$P_{FA} = \sum_{m=1}^{B}\left[(-1)^{(m+1)}\binom{B}{m}\prod_{n=1}^{m}\left(\sum_{r=D}^{T}\frac{\left(1-\hat{\mathrm{P}}_n\right)^{T-r}\hat{\mathrm{P}}_n{}^{r}\,T!}{(T-r)!r!}\right)\right] \tag{4.145}$$

where,

$$\hat{\mathrm{P}}_n = \frac{\bar{k} - n + 1}{N - n + 2} \tag{4.146}$$

For the case of $D = T$, the above form reduces to:

$$P_{FA} = \sum_{m=1}^{B}\left[(-1)^{(m+1)}\binom{B}{m}\prod_{n=1}^{B}\left(\hat{\mathrm{P}}_n\right)^{T}\right] \tag{4.147}$$

If deviation constraints at test band edges are ignored, application of the Robust Binary Integration approach presented in the previous section will produce the following approximation:

$$P_{FA} \approx \sum_{m=1}^{B}\left[(-1)^{(m+1)}\binom{B}{m}\prod_{n=1}^{B}\left(\hat{\mathrm{P}}_n\left[1-\left(1-\hat{\mathrm{P}}_n\right)^{2\Delta+1}\right]^{T-1}\right)\right] \tag{4.148}$$

For the general case which may include G guard cells, I interfering targets, and the use of harmonically transformed spectra with F fractional peaks and G_F fractional guard cells, the base probability form will now be given by the following:

$$\hat{\mathrm{P}}_n = \frac{\bar{k}-n+I+1}{N-n-F(G_F+1)-G+2} \tag{4.149}$$

4.3.5.2.2 - Selective Cell Testing

One of the downsides to the CDF-CFAR detector is the requirement that all cells contained in $\vec{B}$ are evaluated. This may produce unacceptably high computational loads if test bands are large and low operational latency is required. To alleviate this issue, only the M largest cells contained in $\vec{B}$ may be evaluated instead. The obvious downside with this approach is the reduction in detection sensitivity for decreasing values of M. It can be shown however, that this property is only of concern if $M < \bar{k}$. This result may be concluded through a simple logical analysis of the two detection schemes (full range and M maxima).

Recall again that the variable $\bar{k}$ is defined as the reversed order statistic (largest to smallest) of the noise estimate. That is, the $\bar{k}^{th}$ element is the $\bar{k}^{th}$ largest value in $\vec{N}$. To produce a detection, the test cell value X_b must therefore be greater than or equal to the $\bar{k}^{th}$ largest value in $\vec{N}$, where the cell for the detected signal may be denoted as X_s. If $\vec{B} \subseteq \vec{N}$ and a detection is achieved by testing all B cells, then at most there can only be $\bar{k}-1$ cells larger than X_s contained in $\vec{B}$. Therefore, if the $\bar{k}$ largest cells in $\vec{B}$ are evaluated instead of all cells the detection will still be achieved. Since this constitutes the worst-case scenario in which the largest noise-only cells contained in $\vec{N}$ are also contained in $\vec{B}$, sensitivity loss will not be encountered provided $M \geq \bar{k}$. It should be noted that the total false alarm rate for a given trial will not change since the probability of at least one detection will always be governed by the largest test sample when using constrained noise sets. This approach whereby only the M largest cells in $\vec{B}$ are evaluated will be referred to as the Selective-Cell Constrained DF-CFAR (SCDF-CFAR) detector from this point fourth.

4.3.5.2.3 - Frequency Tracking CFAR

As previously indicated by equation (4.142), the DF-CFAR detector will fail ($P_{FA} = 1$) for constrained noise sets if $\vec{B} = \vec{N}$. Binary Integration may be applied to alleviate the problem and produce some level of signal/noise discrimination. However, this approach will produce unnecessarily high computational loads since all cells in $\vec{N}$ are evaluated, which will inevitably produce $\bar{k}$ false detections for each trial. Thus, a more optimal approach would be to simply track the frequency locations of the $\bar{k}$ largest cells instead. Consider a detector which only selects the maximum peak for T consecutive trials and applies Binary Integration to this single location. Given

the current location of the peak cell, the probability that the maximum will again occur at the same location for the next trial is simply $1/B$. If T consecutive trials are made, the probability that the maxima will occur at the same location for all trials is simply:

$$P_{FA} = \left(\frac{1}{B}\right)^{T-1} \tag{4.150}$$

where $T \geq 2$.

In order to better facilitate non-stationary signals, Robust Binary Integration may again be applied. If the peak location is allowed to deviate by $\triangle$ cells to either side of the initial test cell, the probability that the maxima will be within the range defined by $\triangle$ for two consecutive trials is given by:

$$\begin{aligned} P_{FA} &= 2\sum_{i=1}^{\Delta}\left[\left(\frac{1}{B}\right)\left(\frac{\Delta+i}{B}\right)\right] + \left(\frac{B-2\Delta}{B}\right)\left(\frac{2\Delta+1}{B}\right) \\ &= \frac{B+2B\Delta-\Delta(1+\Delta)}{B^2} \end{aligned} \tag{4.151}$$

where $\triangle = 0$ indicates no deviation, $\triangle = 1$ indicates the location can vary by one bin to either side of the test cell giving a total range of $2\triangle + 1 = 3$, and so on. The first term in the above equation (summation) accounts for detection points at the edge of the test band where deviations become constrained, since points outside of $\vec{B}$ will not be detected. The second term accounts for test cells which are located sufficiently far from the band edge such than that the full deviation $\triangle$ can occur without moving outside of $\vec{B}$. These two scenarios are depicted in Figure 4-9 previously displayed.

If T consecutive trials are made, the probability that all subsequent maxima are within the range specified by $\triangle$ with respect to the first maxima will be:

$$P_{FA} = \left(\frac{B+2B\Delta-\Delta(1+\Delta)}{B^2}\right)^{T-1} \tag{4.152}$$

In order to increase detection sensitivity, multiple maxima may be tracked instead of the single largest test cell. If the M largest cells are tracked, the probability of at least one false alarm across two consecutive trials will be:

$$P_{FA} = \Pr[E_1 \cup E_2 \cup ... \cup E_M] = \Pr\left[\bigcup_{i=1}^{M} E_i\right] \tag{4.153}$$

where E_i is the detection event (two consecutive trial detections) for the i^{th} largest cell. Since the detection events are not independent, the principle of inclusion / exclusion must be applied according to:

$$\Pr\left[\bigcup_{i=1}^{M} E_i\right] = \sum_{i=1}^{M}\Pr[E_i] - \sum_{i=1}^{M-1}\sum_{j=i+1}^{M}\Pr\left[E_i \cap E_j\right] + \sum_{i=1}^{M-2}\sum_{j=i+1}^{M-1}\sum_{k=j+1}^{M}\Pr\left[E_i \cap E_j \cap E_k\right] - ..... \tag{4.154}$$

where the union of events can be obtained using the conditional probability formula previously given by equation (4.136). This may be extended to M number of events using the product rule:

$$\Pr\left[\bigcap_{i=1}^{M} E_i\right] = \Pr\left[E_1\right] \cdot \Pr\left[E_2 \middle| E_1\right] \cdot \Pr\left[E_3 \middle| E_1 \cap E_2\right] \ldots \Pr\left[E_M \middle| \bigcap_{i=1}^{M-1} E_i\right] \tag{4.155}$$

The probability for the first detection event will be given by:

$$\Pr\left[E_1\right] = \frac{M}{B} \tag{4.156}$$

while the probability of the second event given the first has occurred will be:

$$\Pr\left[E_2 \middle| E_1\right] = \frac{M-1}{B-1} \tag{4.157}$$

If one continues the approach for $\Pr\left[E_3 \middle| E_2\right]$, $\Pr\left[E_4 \middle| E_3\right]$ and so on, the following general expression can be obtained:

$$\Pr\left[E_B \middle| E_n\right] = \frac{M-n+1}{B-n+1} \tag{4.158}$$

Substitution of (4.155), (4.156), and (4.158) into (4.154), noting the general relationship previously given by (4.141), and performing algebraic manipulation finally gives the following expression:

$$P_{FA} = \sum_{m=1}^{M}\left[(-1)^{(m+1)}\binom{M}{m}\prod_{n=1}^{m}\left(\frac{M-n+1}{B-n+1}\right)\right] \tag{4.159}$$

If a detection is required on all T trials where $T \geq 2$, the false alarm rate will now be given by:

$$P_{FA} = \sum_{m=1}^{M}\left[(-1)^{(m+1)}\binom{M}{m}\prod_{n=1}^{m}\left(\frac{M-n+1}{B-n+1}\right)^{(T-1)}\right] \tag{4.160}$$

If deviation constraints at test band edges are ignored, Robust Binary Integration may be applied to produce the following approximation:

$$P_{FA} \approx \sum_{m=1}^{M}\left[(-1)^{(m+1)}\binom{M}{m}\prod_{n=1}^{m}\left[\frac{(2\Delta+1)(M-n+1)}{B-n+1}\right]^{(T-1)}\right] \tag{4.161}$$

4.3.6 - Spectral Whitening

The following section addresses the issue of performing CFAR detection on signals with colored noise. More specifically, the problem of effectively utilizing the DF-CFAR with non-identically distributed noise samples is discussed. With respect to the outlined considerations, a CFAR-Enhanced Spectral Whitening method is proposed to maintain detector functionality without inhibiting detection sensitivity. The performance of the approach will also be demonstrated using simulated and experimental data.

4.3.6.1 - *Introduction*

Often for many real-world applications, signal noise does not follow a Gaussian distribution but rather exhibits some colored form that is a function of frequency. For example, the transmission of acoustic energy in a viscoelastic medium such as air results in amplitude attenuation that is proportional to the square of the component frequency as described by Stokes' Law [240]. The resulting effect is a coloring of the acoustic energy across the frequency band which includes both noise and signal components. Frequency dependent attenuation or coloring of acoustic spectra is typical of sensing performed in atmospheric conditions. In addition to coloring due to atmospheric propagation, colored broadband noise is also produced through the generation of vortices by the aircraft body and propeller as previously discussed in Section 2.3.4. An example of this can be observed by the UAV self-noise spectra displayed in Figure 2-7.

The presence of spectral coloring may greatly influence the ability to perform operations such as source detection and localization. For example, DF-CFAR methods require noise samples be Independent and Identically Distributed (IID) to maintain and predict constant false alarm rates. Other operations such as beamforming rely on the functional relationship between array steering direction and output power to approximate source locations. Thus, the attenuation of higher frequency components will reduce localization sensitivity for harmonic signals. It is therefore desired to whiten signal spectra before any enhancement, detection, or localization operations are performed to maximize the effectiveness of each process. If each frequency bin of the noise estimate has the same underlying distribution type, coloring effects may be attributed to variations in scale parameters such as mean and variance. If processing techniques can be employed to normalize these parameter values, noise samples will produce equivalent distributions and thus constitute IID variables.

4.3.6.2 - *Standard Whitening Methods*

The two most common forms of spectral whitening are inverse filtering and frequency-band gain control. The frequency-band method is a time domain approach where multiple band-pass filters are applied in parallel to section the signal into various frequency bands. Each of the filtered sections are then equalized using an active scaling approach such as Automatic Gain Control (AGC) or Linear Predictive Coding (LPC). The benefit of this method is a continuously whitened output that does not require any block-based processing such as that inherent with FFT operations. The major downside is potential phase distortions since these scaling processes are typically non-linear [76].

In contrast, inverse filtering is typically performed in the frequency domain and does not produce phase distortions. It involves dividing the spectrum of concern by the mean of its noise approximation according to the following [241]:

$$Y(f) = \frac{X(f)}{\left|\tilde{X}(f)\right|^{\gamma} + C} \tag{4.162}$$

where $\left|\tilde{X}(f)\right|$ is the approximated or smoothed magnitude spectrum of $X(f)$, γ is a scaling or degree-of-flattening factor, and C is a constant to prevent division by zero. If desired, we may exclude the division constant by simply performing the operation in the log-decibel domain instead:

$$\left|Y(f)\right| = \left|X(f)\right| - \gamma\left|\tilde{X}(f)\right| \tag{4.163}$$

To reconstruct the complex signal, the whitened spectrum is simply multiplied by the original phase response:

$$Y(f) = \left|Y(f)\right| e^{j\theta(f)} \tag{4.164}$$

where $\theta(f) = \text{Arg}[X(f)]$.

To obtain the noise approximation $\left|\tilde{X}(f)\right|$, multiple spectra are typically taken consecutively in time and averaged together. If the signal is continuously windowed and frequency transformed, a moving average function may be applied to obtain an accurate approximation. Common averaging methods include the cumulative mean, the recursive exponential mean, and the windowed mean as given by the following equations respectively:

$$\left|\tilde{X}(f,w)\right| = \frac{1}{w}\left[\left|X(f,w)\right| + (w-1)\left|\tilde{X}(f,w)\right|\right] \tag{4.165}$$

$$\left|\tilde{X}(f,w)\right| = \xi\left|X(f,w)\right| + (1-\xi)\left|\tilde{X}(f,w)\right| \tag{4.166}$$

$$\left|\tilde{X}(f,w)\right| = \frac{1}{W}\sum_{k=1}^{W}\left|X(f,w-k)\right| \tag{4.167}$$

where w is the current windowed segment number, ξ is the recursive forgetting factor ($0 < \xi < 1$), and W is the total number of windows used for the mean estimate.

Although simplistic and often effective, the major drawback with the above approach is the potential attenuation of desired signal components from a contaminated noise estimate. If target signal components are present in past windowed spectra which constitute the current noise estimate, the normalization process will act to remove them from the current whitened spectra. The obvious solution to this problem is to simply remove these components from the spectra before taking a mean estimate. However, in many instances the desired signal component(s) and location(s) are not known to facilitate removal. For such cases, the above methods are clearly not optimal in any sense.

4.3.6.3 - *CFAR Enhanced Whitening*

A proposed solution to the problem of attenuating target signal components by inclusion into the mean noise estimate is to simply remove all peak components which may constitute a potential target signal. This can be achieved through the use of a CFAR detector such as those previously presented. Using the detector, potential signals can be identified and effectively removed from the noise estimate by flooring them to some scaled value of the CFAR detection threshold. To ensure all potential components are successfully located, a very high false

alarm probability is used to maximize sensitivity. By using a value much higher than that of the final target detection stage (performed after whitening), the inability to detect a source component and subsequent inclusion into the mean noise estimate will not affect the final detection performance. It is proposed that the OS-CFAR detector be utilized since this form offers computational simplicity and superior performance in multi-target environments. However, any of the other previously presented CFAR detectors may also be used.

For the OS-CFAR detector, the following binary testing function may be constructed:

$$\mathrm{T}(f,w)=\begin{cases}1\ ,\ if\ \ |X(f,w)|\geq\eta(f,w)\\ 0\ ,\ if\ \ |X(f,w)|<\eta(f,w)\end{cases} \tag{4.168}$$

where $\eta(f,w)$ is the threshold factor given by:

$$\eta(f,w)=\alpha_{os}|X_k(f,w)| \tag{4.169}$$

where α_{os} is the order statistic scaling factor, and $|X_k(f,w)|$ is the k^{th} largest spectral component contained in the noise sample bandwidth of size N taken about the test cell $|X(f,w)|$.

Prior to calculating the mean approximation, potential signal components are effectively removed by flooring their value to some scaled fraction of the detection threshold used. This can be expressed by the following operation:

$$|X(f,w)|=\delta\eta(f,w)\mathrm{T}(f,w)+[1-\mathrm{T}(f,w)]|X(f,w)| \tag{4.170}$$

where δ is the flooring scale factor. The mean approximation is then found by substituting the above value into equations (4.165) to (4.167). Finally, the spectrally whitened form can then be obtained via equations (4.162) or (4.163) with $\gamma=0$.

To confirm the validity of the proposed whitening approach, the method is applied to experimental data taken from study TS#1 described in Chapter 6. Using a single channel recording, probability distributions were calculated from consecutive FFT spectra for the unwhitened and whitened signals and compared to that of ideal Gaussian noise. Since the purpose is to evaluate the broadband spectral noise distribution, narrowband self-noise components generated by the aircraft propulsion system were first removed via adaptive IIR notch filtering. In addition, sections of the recorded signal containing target source components were also removed leaving only broadband flight-noise for the entire data set. To calculate the probability distributions, the FFT was applied to the 1150 s duration flight recording using 0.5 s rectangular windows with a 50% overlap producing 4599 windowed points for each frequency bin. Using these observations, the PDFs for each spectral form (whitened, unwhitened, etc.) were then calculated as a function of frequency. Figure 4-11 displayed below provides the results obtained using the magnitude spectra for the original, whitened, and Gaussian noise signals. From the plots, it is evident that broadband noise in the original notch filtered signal are not IID since density values vary largely as a function of frequency. In contrast, the whitened signal PDF is nearly identical to the ideal response obtained from white Gaussian noise which follows a Rayleigh distribution. Thus, we may

conclude that the broadband noise components were effectively whitened to form a group of IID spectral components as desired.

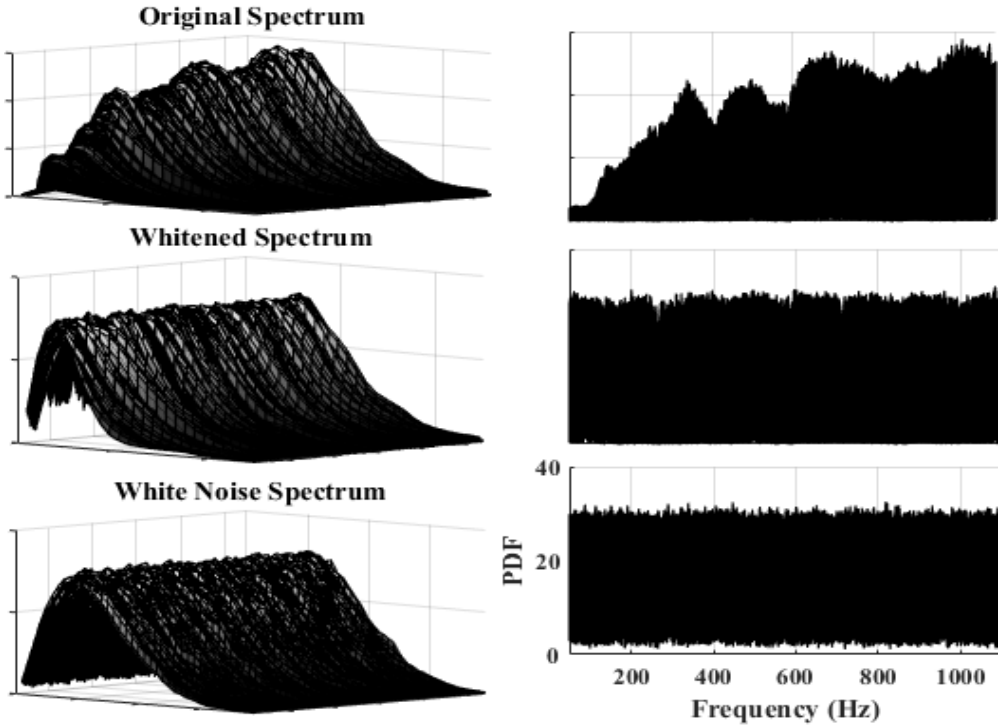

Figure 4-11: Plots illustrating spectral whitening

The performance of the proposed CFAR enhanced whitening approach is now illustrated using data from the same experiment as before, but with a continuous target source signal present. Figure 4-12 displayed below provides a spectrogram of the pre-whitened signal with a 500 Hz source component clearly visible, while Figure 4-13 displays spectrograms for the standard inverse and CFAR enhanced whitened signals respectively. The noise approximation was calculated using the recursive mean as previously given by Equation (4.166) with $\xi = 0.5$, and using a flooring scale value of $\delta = 1$. From observation of the three plots, it is evident that both methods whiten broadband noise components since power levels remain relatively constant across the frequency band. However, the standard approach also greatly attenuates the source signal to near noise-floor levels. It is evident that the proposed CFAR method does not attenuate the source component, but actually increases the SNR slightly while still maintaining an overall whitened response. This effect can be better visualized by Figure 4-14, which depicts the whitening process for a single windowed segment taken at 8.4 minutes into the flight. From these results, it can be concluded that the proposed method is an effective means of whitening colored spectra without attenuating target signal components. Quantification of the effectiveness will be established in the upcoming Simulated Studies section and later verified in Chapter 6.

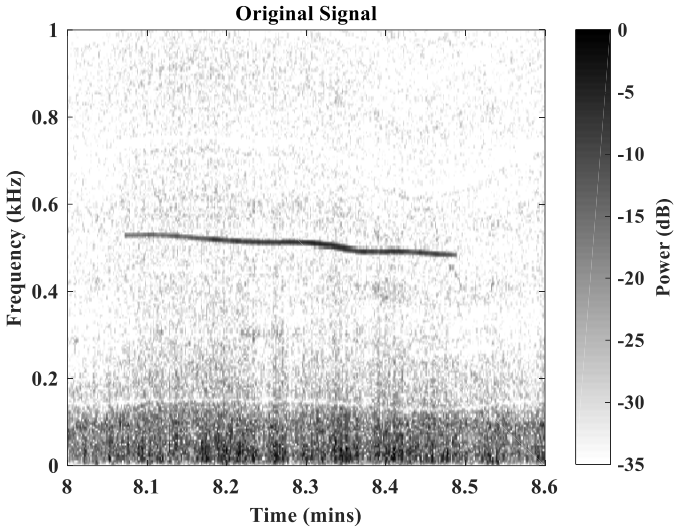

Figure 4-12: Spectrogram of original unwhitened signal.

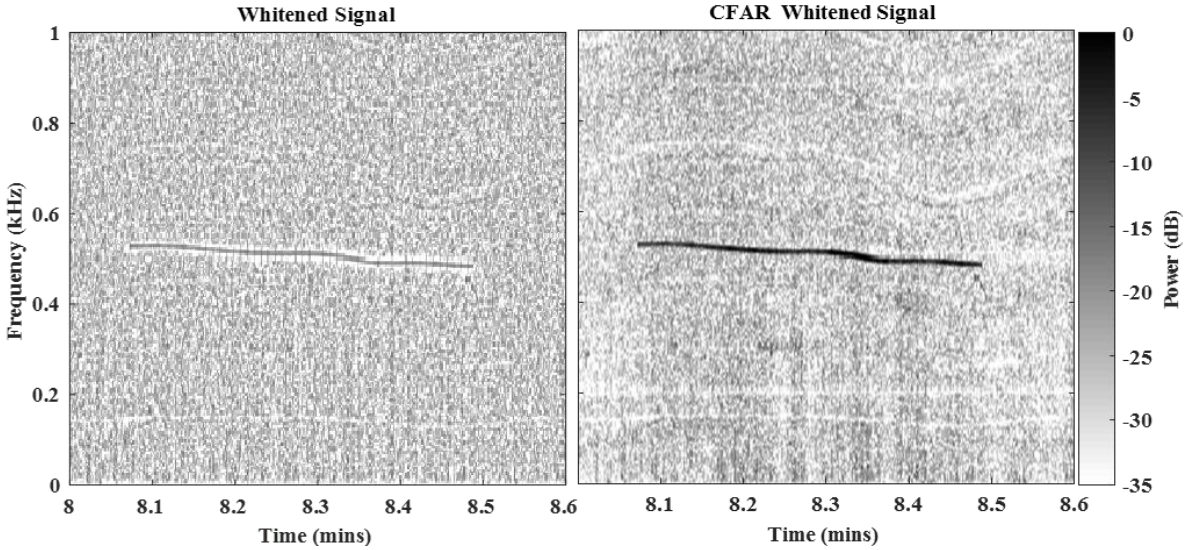

Figure 4-13: Spectrogram of whitened signal using standard and CFAR enhanced methods.

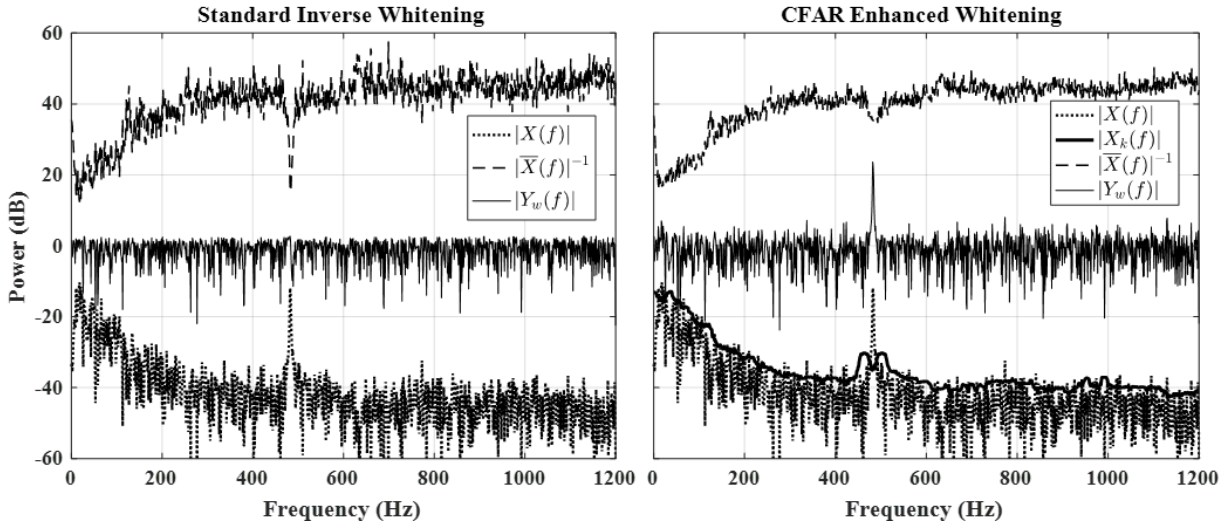

Figure 4-14: Comparison of standard and enhanced whitening methods.

4.4 - Simulation Studies

The following section provides a performance analysis of the proposed signal enhancement processors and CFAR detection schemes using computer generated signals. The purpose of the studies is to validate and identify the top-performing methods under controlled conditions, which are also reflective of those found with real-world data. An analysis of the enhancement processors is first provided, followed by an evaluation of the CFAR detection schemes, spectral whitening, and binary integration methods. Relevant results are also compared to those found in the literature using similar processing techniques (where applicable).

4.4.1 - Signal Enhancement Processors

In order to effectively compare the performance of the various enhancement processors, a numerical simulation was performed using the DF-CFAR method previously described. To simplify the analysis, it is assumed that the signal fundamental frequency is known, meaning only one test bin is required to establish the test statistic. In addition, Binary Integration is not utilized since it would only increase the performance of each processor proportionally and provide no additional insight into which approach is best suited for the application at hand. The false alarm probability was chosen somewhat arbitrarily since again the purpose of the analysis is to compare the relative performance of the processors rather than the absolute signal detectability. Table 4-6 displayed below provides a summary of the various processor forms evaluated throughout the study.

The performance of each processor is quantified in terms of the Receiver Operating Characteristics (ROC), and minimum SNR value required to achieve a detection probability of 99.5% (minimum required value as identified in Section 4.3.2.1). Lower SNR values indicate a higher detectability and thus overall better performance. The average increase in SNR values produced through application of the respective processor was not evaluated, since such values give little indication of signal detectability from a statistical sense. Such an approach will essentially give an average or expected value for the signal and noise component amplitudes without any measure of variance which largely influences detection statistics. In addition, these values are often misleading when calculated from the processor-enhanced spectra since true SNR values may only be obtained using magnitude or power values.

Table 4-6: Summary of various processor forms evaluated.

Processor Description	Symbol	Processor Description	Symbol
Incoherent Mean	$\bar{X}$	Adjusted System Acceleration Coherence (A-SAC) HST of A-SAC	Φ_{λ}^{Ψ} $\bar{\mathrm{H}}_a[\Phi_{\lambda}^{\Psi}]$
Harmogram	HAR	Wagstaff's PAC HST of PAC	$PAC(X,\phi)$ $\bar{\mathrm{H}}_a[PAC(X,\phi)]$
Harmonic Product Spectrum	HPS	Wagstaff's PAV HST of PAV	$PAV(X,\phi)$ $\bar{\mathrm{H}}_a[PAV(X,\phi)]$
Harmonic Spectrum Transform (HST)	$\bar{\mathrm{H}}_a[\ ]$	Phase-only PAC (P-PAC) HST of P-PAC	$PAC(\phi)$ $\bar{\mathrm{H}}_a[PAC(\phi)]$
HST of FFT Magnitude	$\bar{\mathrm{H}}_a[X]$	Phase-only PAV (P-PAV) HST of P-PAV	$PAV(\phi)$ $\bar{\mathrm{H}}_a[PAV(\phi)]$
Circular Convolution Enhanced (CCE) FFT Magnitude	Z	Generalized Magnitude Squared Coherence (GMSC) HST of GMSC	Γ $\bar{\mathrm{H}}_a[\Gamma]$
HST of CCE FFT Magnitude	$\bar{\mathrm{H}}_a[Z]$	Generalized Acceleration Squared Coherence (GASC) HST of GASC	Γ^{Ψ} $\bar{\mathrm{H}}_a[\Gamma^{\Psi}]$
Acceleration Vector Coherence (AVC) HST of AVC	$\vec{\Phi}$ $\bar{\mathrm{H}}_a[\vec{\Phi}]$	HST of FFT and HST of A-AVC (Product)	$\bar{\mathrm{H}}_a[X\cdot\vec{\Phi}^{\Psi}]$
Adjusted Acceleration Vector Coherence (A-AVC) HST of A-AVC	$\vec{\Phi}^{\Psi}$ $\bar{\mathrm{H}}_a[\vec{\Phi}^{\Psi}]$	HST of FFT and HST of A-SAC (Product)	$\bar{\mathrm{H}}_a[X\cdot\Phi_{\lambda}^{\Psi}]$
System Acceleration Coherence (SAC)	Φ_{λ}		
HST of SAC	$\bar{\mathrm{H}}_a[\Phi_{\lambda}]$		

4.4.1.1 - *Signal Description*

The test signal consisted of a $f_0 = 100$ Hz sine wave with $R = 6$ harmonics embedded in Gaussian noise of unity variance ($\sigma = 1$). The signal was constructed using a sampling rate of $f_s = 3000$ Hz and transformed to the frequency domain by applying the FFT on consecutive 1 s signal segments with a 0.75 s overlap. A 3000-point FFT was utilized to giving a spectral resolution of $f_{res} = 1$ Hz/bin. Two guard cells ($G = 2$) and seventeen fractional peaks ($F = 17$) were utilized to minimize effects of fractional peak values and spectral leakage. To determine the ROC and relevant detection statistics, a total of 10,000 trials were conducted for each of the SNR test points used.

A range of signals ($s = 1,2,4,6$) were used to determine the general effect of increasing the number of channels available for processing. The signal model for the s^{th} channel is given by:

$$x_s(t) = \sum_{m=1}^{M} A_m \cos\left[2\pi f_0 tm + \phi_m + \theta_s\right] + w_s(t) \tag{4.171}$$

where A_m is the amplitude of the m^{th} harmonic component, f_0 is the fundamental frequency component, ϕ_m is the initial phase of the m^{th} harmonic, θ_s is the phase shift applied to the s^{th} channel, and $w_s(t)$ is additive white Gaussian noise. Initial phase values for harmonic components where chosen randomly. Signals received by each channel were also shifted out of phase such that the coherent addition of the signals would result in a value of zero (if no noise were present). Thus, for S number of channels being utilized simultaneously the phase shift applied to each signal is given by:

$$\theta_s = \frac{2\pi}{S}(s-1) \tag{4.172}$$

for $s \in \{1,2,\ldots,S\}$. Amplitude values for subsequent harmonics were specified according to the following equation:

$$A_m = \frac{A_0}{\sqrt{m}} \tag{4.173}$$

where A_0 is the amplitude value for the fundamental harmonic component, and $1/\sqrt{m}$ is the harmonic attenuation factor. The above attenuation factor was chosen since this form was also utilized by Hinch when he first proposed the Harmogram [152, 153], and it also provides a good approximation to the harmonic attenuation properties of propeller driven aircraft [72]. The fundamental amplitude is calculated from the desired SNR of the constructed signal(s) which is given by:

$$SNR = 10\log_{10}\left[\frac{\sigma_y^2}{\sigma_n^2}\right] = 10\log_{10}\left[\frac{\sum_{m=1}^{M}\frac{A_m^2}{2}}{\sigma_n^2}\right] 10\log_{10}\left[\frac{\frac{1}{2}\sum_{m=1}^{M}\left(\frac{A_0}{\sqrt{m}}\right)^2}{\sigma_n^2}\right] \tag{4.174}$$

where σ_n^2 is noise component variance. Rearranging for A_0 gives the fundamental amplitude required to achieve some desired SNR value:

$$A_0 = \sigma_n\sqrt{2\left[\sum_{m=1}^{M}\left(\frac{1}{m}\right)\right]^{-1} 10^{\frac{SNR}{10}}} \tag{4.175}$$

Table 4-7 provides a summary of the signal and detector parameters used. Note that the false alarm rate was determined via equation (4.120) since the DF-CFAR detector was applied to harmonically transformed spectra.

Table 4-7: Simulation signal parameters.

Signal Parameters		Detector Parameters	
Source Frequency	100 Hz	Guard Cells (G)	2
Signal Harmonics	6	Fractional Peaks (F)	17
Number of Signals	1 - 6	Fractional Guard Cells (G_F)	0
Sampling Frequency	3000 Hz	Interfering Targets (I)	0
FFT Block Time	1 s	Noise Samples (N)	181
FFT Block Overlap	75%	Order Statistic ($\bar{k}$)	1
FFT Length	3000 pts	False Alarm Rate (P_{fa})	5.5E-3
Spectral Resolution	1 Hz/bin		
Number of Blocks	10,000		

4.4.1.2 - *Harmonic Spectral Transforms*

The following section compares the performance of the Harmonic Product Spectrum (HPS), Harmogram (HAR), and various Harmonic Spectral Transforms (HST). Detection values are established for a range of available signals and compared to that of the standard incoherent mean as defined below by equation (4.176). The Circular Convolution Enhancement (CCE) method is also analyzed to determine if the approach will in fact increase detectability for missing harmonic components. It should be noted that the mean form evaluated is uniform across all compounding directions for the HST operation. For example, if the standard mean was utilized across the frequency spectrum as indicated by $\overline{\mathrm{H}}_1[\,]$, the standard form would also be used across channels if $S > 1$. Thus, the MHST notation $\overline{\mathrm{H}}_{1,1}[\,]$ which indicates this operation, is simply replaced by $\overline{\mathrm{H}}_1[\,]$ for sake of simplicity in displaying results.

$$\bar{X}(f) = \frac{1}{S}\sum_{s=1}^{S}\left|X_s(f)\right| \tag{4.176}$$

Table 4-8 provides the SNR values required to achieve a detection probability of $P_D = 0.995$ for a range of processing channels, while Figure 4-15 provides an ROC plot for the case of one processing channel ($S = 1$). From the results displayed, it is evident that all harmonic transforms reduce the minimum required SNR considerably over that of the standard incoherent mean. In general, increasing the number of processing channels also increased detectability. Results obtained via the Harmogram were equal to that of the Standard Mean ($\overline{\mathrm{H}}_1$) and RMS ($\overline{\mathrm{H}}_2$) forms, which also achieved the top overall performance. In addition, the HPS was found to equal that of the Geometric Mean ($\overline{\mathrm{H}}_0$) processor. These results suggest that the specific spectral form (magnitude or power) does not affect signal detectability. The Harmonic Mean ($\overline{\mathrm{H}}_{-1}$) processor achieved the lowest detectability out of the HST forms. However, results were still higher than that of the incoherent mean (1.3 dB increase on average).

Table 4-8: Detection performance of harmonic spectral transforms.

Processor	SNR (dB)			
	S = 1	S = 2	S = 4	S = 6
$\bar{X}$	-16.4	-18.7	-21.1	-22.3
HAR	-20.2	-22.4	-24.2	-25.4
HPS	-19.4	-21.6	-23.4	-24.6
$\bar{\mathrm{H}}_{-1}[X]$	-17.8	-20.2	-22.2	-23.4
$\bar{\mathrm{H}}_{0}[X]$	-19.4	-21.7	-23.5	-24.8
$\bar{\mathrm{H}}_{1}[X]$	-20.2	-22.4	-24.2	-25.4
$\bar{\mathrm{H}}_{2}[X]$	-20.1	-22.4	-24.2	-25.4

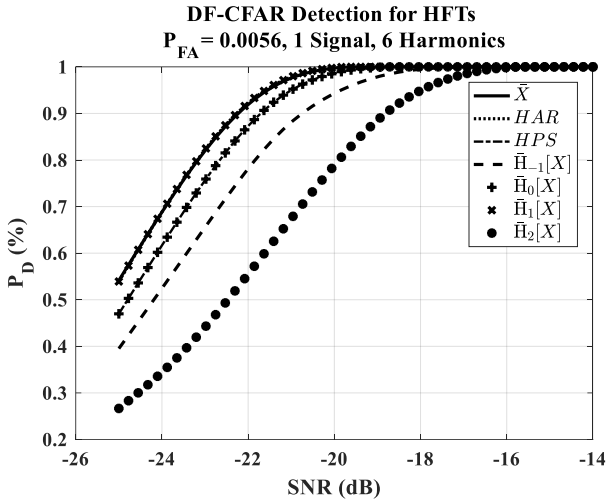

Figure 4-15: ROC curves for basic harmonic transforms.

As previously discussed, one of the downsides to product-based processors (HPS and Geometric Mean) is that degraded harmonic component(s) will greatly reduce detection performance. If a source harmonic has the same frequency as one of the narrowband self-noise components, notch filtering will completely remove this signal component. In this case the processor will fail since multiplication by zero (or a very small number) will occur during the harmonic compounding operation. Thus, it is desired to analyse the effect of utilizing the CCE method proposed by Wu [175] to enhance signal detectability for such cases.

Signals were generated according to the specifications previously outlined, but with the 4th harmonic component attenuated by 80 dB. Table 4-9 provides the results obtained with the 4th harmonic removed and with no harmonics removed. From the results, it is evident that the Geometric Mean ($\bar{\mathrm{H}}_0$) processor does in fact fail for the case of missing components as expected. Application of the CCE method did alleviate this problem and ultimately produced results comparable to that if no harmonics were missing. However, for the Standard Mean ($\bar{\mathrm{H}}_1$) processor, results were not so favourable. Application of the CCE method actually decreased performance for both cases. For the missing component case, the CCE method increased required SNR values by 3.6 dB on average, while a 4.7 dB increase was required if no components were removed. A similar result was also obtained for the Geometric Mean when no harmonics were missing; SNR values increased by 4.3 dB on average. Thus, it is evident that the CCE method does prevent the failure of product-based processors. However, detection performance is decreased if no harmonics are missing. For the case of summation-based processors, the CCE method actually lowers detectability regardless if any components are missing or not. It should be mentioned however, that these results do not necessarily conflict with those reported by Wu, since his evaluation of the technique was in terms of pitch tracking not signal detectability.

Table 4-9: Detection results from Self-Circular Convolution Enhancement.

Processor	SNR (4th Harmonic Removed)				SNR (No Harmonics Removed)			
	$S=1$	$S=2$	$S=4$	$S=6$	$S=1$	$S=2$	$S=4$	$S=6$
Z	-14.7	-15.7	-16.8	-17.3	-15.7	-16.7	-17.8	-18.3
$\bar{\mathrm{H}}_{-1}[\hat{Z}]$	-14.7	-16.1	-17.2	-17.7	-16.1	-17.3	-18.1	-18.7
$\bar{\mathrm{H}}_{0}[\hat{Z}]$	-15.7	-16.8	-17.8	-18.3	-16.6	-17.7	-18.7	-19.2
$\bar{\mathrm{H}}_{1}[\hat{Z}]$	-16.2	-17.3	-18.2	-18.8	-16.9	-18.0	-19.0	-19.5
$\bar{\mathrm{H}}_{0}[X]$	-	-	-	-	-19.4	-21.7	-23.5	-24.8
$\bar{\mathrm{H}}_{1}[X]$	-18.7	-20.7	-22.3	-23.2	-20.2	-22.4	-24.2	-25.4

Based on the results obtained from each of the scenarios investigated, it is apparent that utilizing harmonic transforms generally increases the detectability of harmonic signals. Overall, the Standard Mean processor offered the best performance as it achieved the highest detectability for all cases. In addition, the performance of the processor was not largely affected by missing harmonic components unlike that of the Geometric Mean form. The CCE method was found to salvage the operational performance of the Geometric Mean processor for missing harmonic components. However, performance was generally found to degrade for all processors if no components were actually missing. Thus, the method will be excluded from any further processing or analysis pertaining to this thesis.

4.4.1.3 - *Phase Acceleration Processors*

The performance of the presented PAPs is now analyzed using the simulation data previously outlined. These include the AVC ($\vec{\Phi}$), A-AVC ($\vec{\Phi}^{\Psi}$), SAC (Φ_{λ}), and A-SAC (Φ_{λ}^{Ψ}) processors. A beta value of $\beta = 0.9$ was used for each of the adjusted processor forms. The original and modified PAV and PAC processors are also evaluated to form a comparative basis. Results of the phase-only forms are first presented, followed by the amplitude and phase combined forms. Since the detection signal is harmonic in nature, the $\bar{\mathrm{H}}_1[\,]$ transform is also applied to each processor to achieve maximum detection capability.

Table 4-10 provides the SNR values required to achieve a detection probability of $P_D = 0.995$ for a range of processing channels, while Figure 4-15 provides ROC plots for the case of six processing channels ($S = 6$). Results obtained for the AVC processor are identical to that of the phase-only PAV processor and approximately equal to that of the SAC. The phase-only PAC was found to produce slightly better results than either of these three. The phase adjusted AVC and SAC forms attained the highest performance values with significant increases over the standard unadjusted forms. On average the adjusted forms reduced required SNR values by 3.4 and 3.8 dB for the AVC and SAC processors respectively. Overall, the adjusted SAC offered the best detection performance.

From the results obtained for the combined amplitude and phase processors, it is evident that including amplitude information significantly improved detection capabilities. From a comparison of the standard and

harmonically transformed PAV and PAC processors, it is also apparent that using the HST in conjunction with these processors significantly increased detectability. The PAV and PAC processors offered better performance over the standard AVC and SAC, but were less capable compared to the phase adjusted forms. Differences between the standard AVC and SAC were also found to increase compared to the phase-only forms, which were approximately equal. Application of the modulo 2π adjustment factor significantly increased detection performance, however the relative increase was less than that of the phase-only forms. On average the adjusted forms reduced required SNR values by 1.3 dB for both the AVC and SAC processors. The adjusted SAC processor again offered the best detection performance.

Table 4-10: Detection results for phase acceleration processors.

Phase-only Processors	SNR (dB)			Amp & Phase Processors	SNR (dB)		
	S = 2	S = 4	S = 6		S = 2	S = 4	S = 6
$\bar{H}_1[\vec{\Phi}]$	-15.3	-19.9	-21.5	$\bar{H}_1[X \cdot \vec{\Phi}]$	-21.4	-23.1	-24.2
$\bar{H}_1[\Phi_\lambda]$	-15.3	-20.1	-21.8	$\bar{H}_1[X \cdot \Phi_\lambda]$	-21.4	-24.1	-25.1
$\bar{H}_1[\vec{\Phi}^\Psi]$	-19.5	-22.9	-24.5	$\bar{H}_1[X \cdot \vec{\Phi}^\Psi]$	-22.5	-24.6	-25.6
$\bar{H}_1[\Phi_\lambda^\Psi]$	-20.1	-23.6	-25.0	$\bar{H}_1[X \cdot \Phi_\lambda^\Psi]$	-22.5	-25.4	-26.7
$\bar{H}_1[PAV(\phi)]$	-15.3	-19.9	-21.5	$\bar{H}_1[PAV(X,\phi)]$	-22.2	-24.4	-25.2
$\bar{H}_1[PAC(\phi)]$	-17.9	-20.6	-22.1	$\bar{H}_1[PAC(X,\phi)]$	-22.2	-24.4	-25.3
				$PAV(X,\phi)$	-18.7	-21.2	-22.3
				$PAC(X,\phi)$	-18.8	-21.2	-22.4

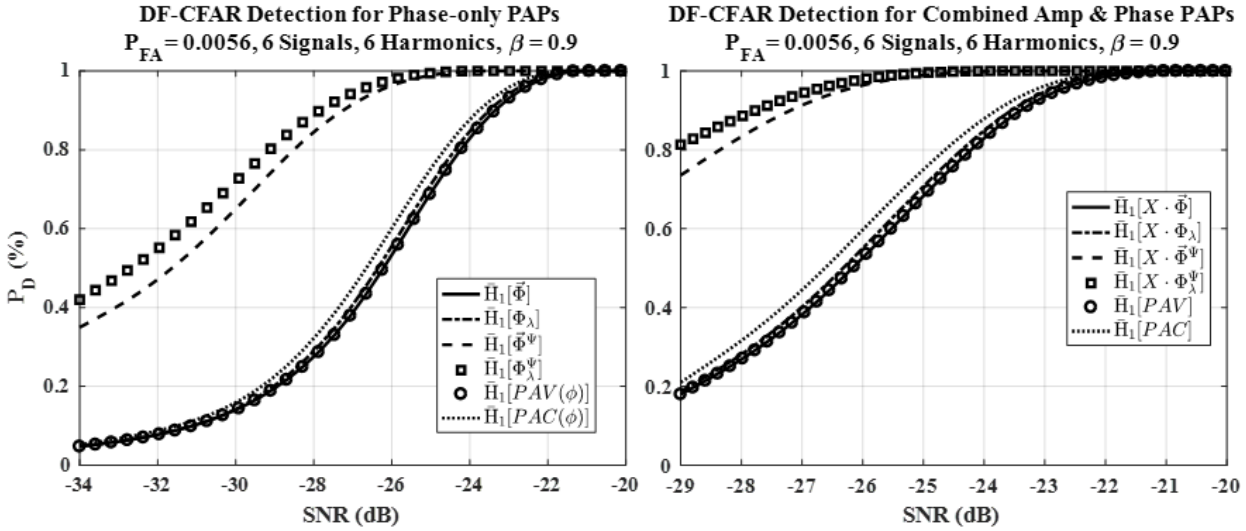

Figure 4-16: ROC plots for PAPs with six processing channels.

4.4.1.4 - *Modified Coherence Processors*

The performance of the presented coherence processors is now analyzed using the simulation data previously outlined. These include the GMSC ($\tilde{\Gamma}$), the proposed GASC ($\tilde{\Gamma}^\Psi$), and the standard mean HST ($\bar{H}_1[\,]$) taken of these forms. Table 4-11 provides the detection SNR values for a range of processing channels and windows, while Figure 4-17 provides ROC plots for two of the evaluated scenarios. Coherence values were calculated

using windows of 1 s duration with a 75% overlap. The total processing time (TPT) required using these specifications is indicated by the TPT values displayed in the table below.

From the results obtained, it is evident that the proposed GASC provides a significant performance increase over the GMSC developed by Wu. In addition, application of the HST also increased detection capabilities for both processors as expected. Differences in detection values for the GMSC and GASC are also indicated in the table below. For the standard processor form (not harmonically transformed), it is evident that a greater performance increase is achieved by the GASC for fewer processing channels. However, the harmonically transformed processors exhibit the opposite effect and shows a greater performance with increasing channel numbers. It is also apparent that increasing the number of processing windows also has a much larger effect when fewer channels are used for all processor forms.

One of the major downsides typical of using coherence based processors is the number of windowed segments usually required to achieve reasonable signal/noise discrimination. An upwards of 8 windowed segments are often required to achieve appreciable values, which is obviously problematic for non-stationary signals [75]. Indeed, this is evident by the values obtained for the GMSC. However, utilization of the phase acceleration does appear to greatly reduce this requirement. For example, required SNR values were reduced by approximately 7 dB for the case of two signals and three processing windows ($S = 2, W = 3$). Thus, it may be concluded that the application of phase acceleration should be included if utilizing coherence based processors for signal detection purposes.

Table 4-11: Detection results for coherence processors for a range of processing windows and channels.

Processor	**W = 3, TPT = 1.5 s**			**W = 4, TPT = 1.75 s**			**W = 6, TPT = 2.25 s**			**W = 8, TPT = 2.75 s**		
	S = 2	S = 4	S = 6	S = 2	S = 4	S = 6	S = 2	S = 4	S = 6	S = 2	S = 4	S = 6
Γ	-10.5	-18.4	-20.2	-13.1	-19.3	-21.1	-15.0	-20.7	-22.7	-16.4	-22.0	-23.6
Γ^{Ψ}	-17.4	-21.5	-24.1	-18.7	-22.8	-24.9	-20.4	-23.8	-25.9	-21.0	-24.7	-26.5
$\tilde{\Gamma}^{\Psi} - \tilde{\Gamma}$	-6.9	-3.2	-3.8	-5.6	-3.5	-3.8	-5.5	-3.1	-3.2	-4.6	-2.7	-2.9
$\bar{H}_1[\tilde{\Gamma}]$	-20.8	-23.7	-25.1	-21.5	-24.6	-26.0	-22.5	-25.5	-27.0	-23.1	-26.4	-27.6
$\bar{H}_1[\tilde{\Gamma}^{\Psi}]$	-23.9	-27.5	-29.4	-24.6	-28.1	-30.3	-25.6	-29.3	-31.2	-26.3	-30.2	-31.8
$\bar{H}_1[\tilde{\Gamma}^{\Psi}] - \bar{H}_1[\tilde{\Gamma}]$	-3.1	-3.7	-4.3	-3.0	-3.5	-4.3	-3.1	-3.8	-4.2	-3.2	-3.8	-4.2

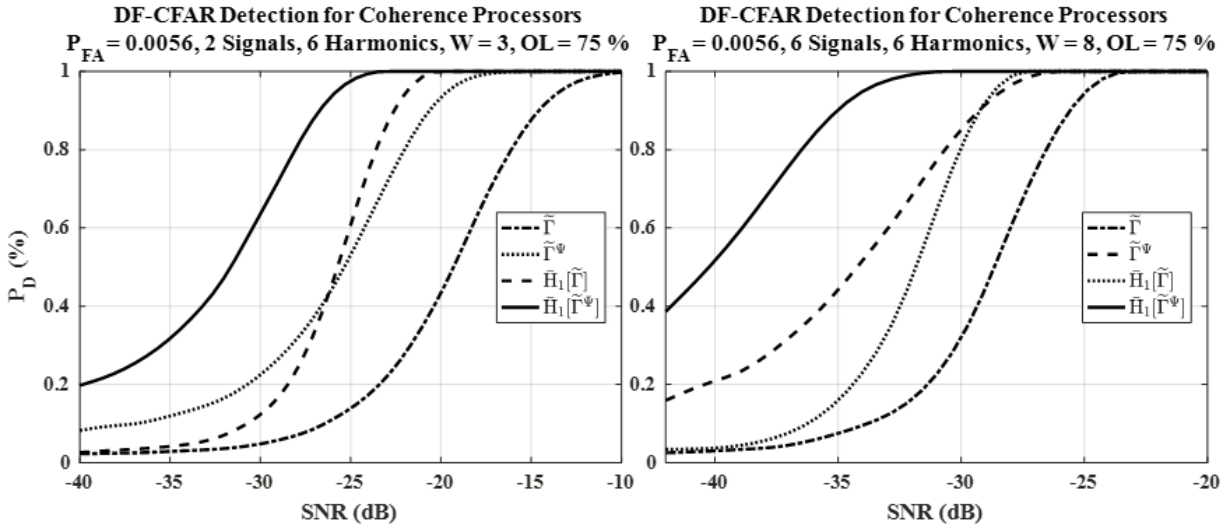

Figure 4-17: ROC plots for standard and modified coherence processors.

4.4.1.5 - *Results Summary*

The following provides a brief summary of the significant findings obtained from the enhancement processor study:

- Application of HST increased detection performance for all processors evaluated, with the Standard Mean form ($\overline{H}_1$) offering the best overall performance.
- The CCE does prevent the failure of product-based processors if harmonic components are missing. However, detection performance is actually decreased if no harmonics are missing. For the case of summation-based processors, the CCE method lowers detectability regardless if any components are missing or not.
- The modulo-2π phase adjustment factor provides a significant increase in detection performance for the AVC and SAC processors, with the harmonically transformed A-SAC providing the best performance out of all the PAPs evaluated.
- The combined HST and PAP processors were found to produce significantly better results than either of the forms applied independently. In addition, results were also found to surpass those generated by Wagstaff's PAC and PAV processors.
- The proposed GASC provided a significant performance increase over the GMSC processor for all scenarios evaluated. Results obtained for this processor were found to be significantly higher than all others evaluated including the HST and PAPs.

4.4.2 - Signal Detection

The performance of the proposed CFAR detectors for constrained non-independent tests are now demonstrated using computer-generated signals. The CDF-CFAR, SCDF-CFAR, and FT-CFAR detectors are first analysed by establishing ROC plots for the case of a stationary signal in white Gaussian noise. The SCDF-CFAR detector

is then applied to a non-stationary signal in colored noise to demonstrate the effectiveness of the proposed CFAR-Enhanced Spectral Whitening method. In addition, the ability of the Robust Binary Integration approach to increase detection rates for non-stationary signals is also demonstrated.

4.4.2.1 - *CFAR Detector Analysis*

The test signal consisted of a single-component $f_0 = 250$ Hz sine wave embedded in Gaussian noise of unity variance ($\sigma = 1$). It was constructed using a sampling rate of $f_s = 1000$ Hz and transformed to the frequency domain by applying the FFT on consecutive 1 s signal segments with no overlap. A 1000-point FFT was utilized giving a spectral resolution of $f_{res} = 1$ Hz/bin. The constructed time domain signal is given by:

$$x(t) = A\cos\left(2\pi f_0 t\right) + w(t) \tag{4.177}$$

where A is the signal amplitude, f_0 is the fundamental frequency component, and $w(t)$ is additive white Gaussian noise. To determine the ROC and relevant detection statistics, a total of 10,000 trials were conducted for each of the SNR test points used. Each of the detectors were compared by determining the SNR value required to achieve a minimum detection probability of $P_D = 0.995$ for the specified P_{FA}. The SNR value for a given amplitude can be obtained according to:

$$SNR = 10\log_{10}\left[\frac{\sigma_x^2}{\sigma_n^2}\right] = 10\log_{10}\left[\frac{A^2}{2\sigma_n^2}\right] \tag{4.178}$$

Table 4-12 displayed below provides a summary of the signal parameters used for the simulation, while

Table 4-13 provides the configuration parameters used for each detector. To verify the equivalency of the CDF-CFAR and SCDF-CFAR detectors, the number of maxima tracked (M) was determined based on the $M \geq \bar{k}$ condition as previously discussed. Parameters values for the FT-CFAR were chosen such that false alarm rates were approximately equal to that of the other two detectors in order to establish a common comparative basis.

Table 4-12: Simulation signal parameters.

Source Frequency	250 Hz
Sampling Frequency	1000 Hz
FFT Block Time	1 s
FFT Block Overlap	0%
FFT Length	1000 pts
Spectral Resolution	1 Hz/bin
Number of Blocks	10,000
SNR	-13 to -18 dB

Table 4-13: Detector parameters for ROC simulation.

	CDF-CFAR	SCDF-CFAR	FT-CFAR
Noise Sample Band ($\vec{N}$)	1-500 Hz	1-500 Hz	-
Test Band ($\vec{B}$)	200-300 Hz	200-300 Hz	1-500 Hz
Noise Samples (N)	499 pts	499 pts	-
Test Cells (B)	101 pts	101 pts	499 pts
Order Statistic ($\bar{k}$)	5	5	-
Consecutive Trials (T)	3	3	3
Consecutive Detections (D)	3	3	3
Cell Deviation (Δ)	1	1	1
Maxima Tracked (M)	-	5	3
ST (P_{FA})	0.678	0.678	1
BI (P_{FA})	1.01E-04	1.01E-04	1.08E-04
RBI (P_{FA})	8.91E-04	8.91E-04	9.76E-04

The ROC curves for each of the detectors are displayed below in Figure 4-18, while Table 4-14 provides the SNR values required to achieve 99.5% detection rate. The ROC curve for the Robust Binary Integration (RBI) scheme was not included since it is identical to that of the BI method for stationary signals. From the plots and values displayed, it is apparent that the CDF-CFAR and SCDF-CFAR produce equivalent results as expected. The FT-CFAR performed slightly less favourable with an average decrease in detectability of 0.25 dB.

In terms of detection schemes, the Single Trial (ST) test statistic produced the highest detectability results. This was expected since a positive correlation always exists between detectability and false alarm rate. The high false alarm rates produced were due to the constrained nature of the test, and relative size difference between N and B. However, the use of Binary Integration (BI) was shown to drastically reduce these values while having a relatively small effect on signal detectability.

Table 4-14: SNR value required to achieve a 99.5% detection rate.

	CDF-CFAR	SCDF-CFAR	FT-CFAR
ST	-15.2	-15.2	-14.9
BI	-14.7	-14.7	-14.5
RBI	-14.7	-14.7	-14.5

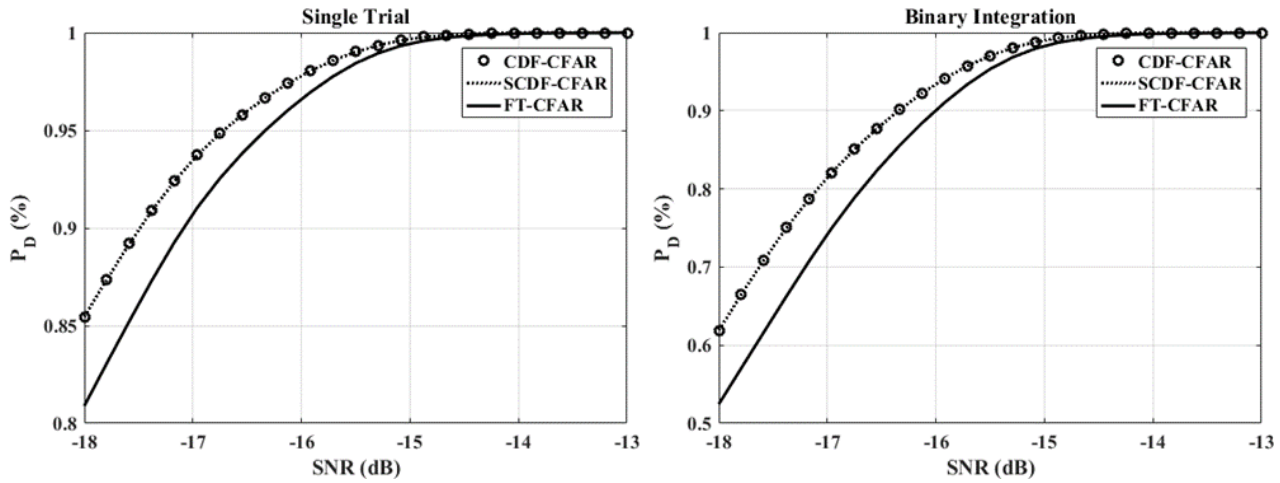

Figure 4-18: ROC curves for proposed detectors.

Based on the results obtained, it can be concluded that the SCDF-CFAR offered the best overall performance. It performed equally well to the CDF-CFAR detector but only required evaluation of 5 test cells for each trial instead of the total 101; this greatly reduced computational requirements. The FT-CFAR offers a slightly more simplistic setup, since it has only one sample set, uses no order statistics, and requires less maxima tracking points. However, this increase in simplicity and reduction in computational loads is produced at the cost of detection sensitivity. In regard to detection schemes, it is evident that the RBI is less desirable for stationary signals since false alarm rates are inherently increased with no effect on signal detectability. However, it will be shown in the next section that the RBI method is superior in detecting and tracking non-stationary signals.

4.4.2.2 - *Spectral Whitening & Binary Integration*

To illustrate the performance increase of the proposed CFAR enhanced spectral whitening technique and Robust Binary Integration, a non-stationary sinusoidal signal embedded in colored noise was utilized. The test signal consisted of a single-component quadratic chirp with a frequency value ranging from 100 to 400 Hz. The signal frequency at any given time can be obtained via the following equations:

$$f(t) = f_0(t) + \alpha t^2 \tag{4.179}$$

$$\alpha = (f_1 - f_0) / t_1^2 \tag{4.180}$$

where $f_0 = 100$ Hz, $f_1 = 400$ Hz, and $t_1 = 60$ s (total simulation duration). The signal phase can then be obtained according to:

$$\theta(t) = \frac{2\pi}{f_s} \int_0^t f(\tau) d\tau \tag{4.181}$$

Using the instantaneous phase, the signal was then constructed according to the following:

$$x(t) = A\cos\big(\theta(t)\big) + p(t) \tag{4.182}$$

where $p(t)$ is pink colored noise which has a power distribution proportional to the inverse spectral frequency value ($P(f) \propto 1/f$). Pink noise was utilized since it is a reasonable approximation of that found when

conducting acoustic sensing in atmospheric conditions (attenuation with increasing frequency). The signal was transformed via the FFT operation and whitened using the parameter values displayed below in Table 4-15. Each FFT data block was then evaluated using the SCDF-CFAR detector configured according to the values given in Table 4-16.

Table 4-15: Simulation analysis parameters.

Signal & FFT		**Spectral Whitening**	
Source Frequency	100-400 Hz	Detector Type	OS-CFAR
Sampling Frequency	1000 Hz	Forgetting Factor (λ)	0.5
FFT Block Time	0.5 s	Flooring Factor (δ)	0.001
FFT Block Overlap	50%	Noise Band ($\vec{N}$)	50 Hz
Padded FFT Length	2000 pts	Noise Samples (N)	100
Spectral Resolution	0.5 Hz/bin	Order Statistic (k_{os})	0.75 N
Simulation Duration	60 s	Guard Cell Band ($\vec{G}$)	5 Hz
		Guard Cells (G)	10

Table 4-16: CFAR detector parameters.

Noise Sample Band ($\vec{N}$)	**10-500 Hz**	**Maxima Tracked (M)**	**5**
Test Band ($\vec{B}$)	100-400 Hz	Consecutive Trials (T)	3
Noise Samples (N)	990	Consecutive Detections (D)	3
Test Cells (B)	601	Cell Deviation (Δ)	1
Order Statistic ($\bar{k}$)	5		

Figure 4-19 provides spectrogram plots of the unwhitened and CFAR-Enhanced spectrally whitened signals, while Figure 4-21 and Figure 4-22 provides spectrogram-like plots of the detection points for each testing scheme. Table 4-17 displayed below provides the results obtained from the simulation. The methods were evaluated in terms of total signal detectability, which is defined as the number of successful detections relative to the total number of possible detections (FFT data blocks).

From the detection plots and values displayed, it is apparent that the proposed spectral whitening method provides a significant increase in signal detectability with an average increase of 29%. The decreased performance for the unwhitened signal is caused by the increased power levels associated with lower frequency values. This effectively produces noise estimates that are not reflective of that found near the signal locations (unnecessarily high). This low frequency amplification is clearly visible in the power spectrum plot displayed in Figure 4-20. This trend was also visible in the experimental data spectrum previously displayed by Figure 4-14.

From the results obtained, it is evident that the RBI testing scheme provides a significant increase over the standard BI approach for both signals. An average increase in signal detectability of 20% was obtained, but at a cost of increasing the false alarm rate by a full order of magnitude. From a visual inspection of the detection plots, it is apparent that both methods perform similarly well in the low frequency region (100-150 Hz) where values are relatively stationary. However, for higher value regions it is evident that the RBI method attains a

significantly larger number of detections. This may be further increased allowing greater location deviations (Δ) between consecutive trials, provided false alarm rates are still satisfied.

Based on the results obtained, it can be concluded that the proposed spectral whitening method is very effective at increasing signal detectability for cases which involve colored noise. The RBI detection scheme was also found to increase detectability for signals producing a high degree of non-stationarity. However, this capability is attained at the expense of increased false alarm rates. Thus, depending on the degree of frequency variation, the standard BI approach may offer the better overall option.

Table 4-17: Simulation detection results.

	False Alarm Rate	**Detection Rate**	
	P_{FA}	Whitened	Unwhitened
SC	0.99	100%	100%
BI	7.7 E-05	64%	34%
RBI	6.9 E-04	83%	55%

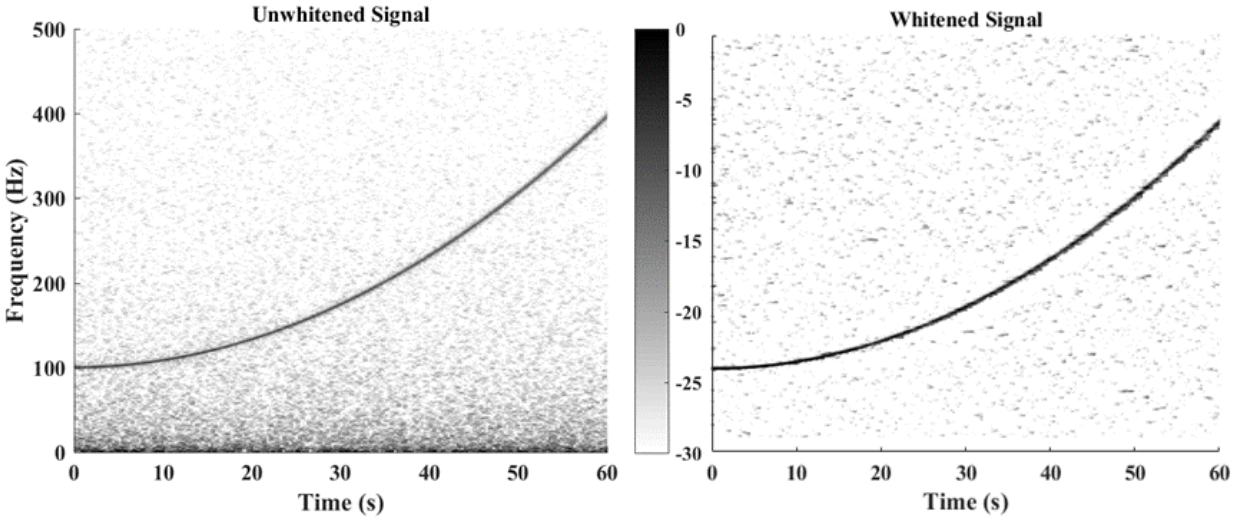

Figure 4-19: Spectrograms of standard and CFAR whitened signals.

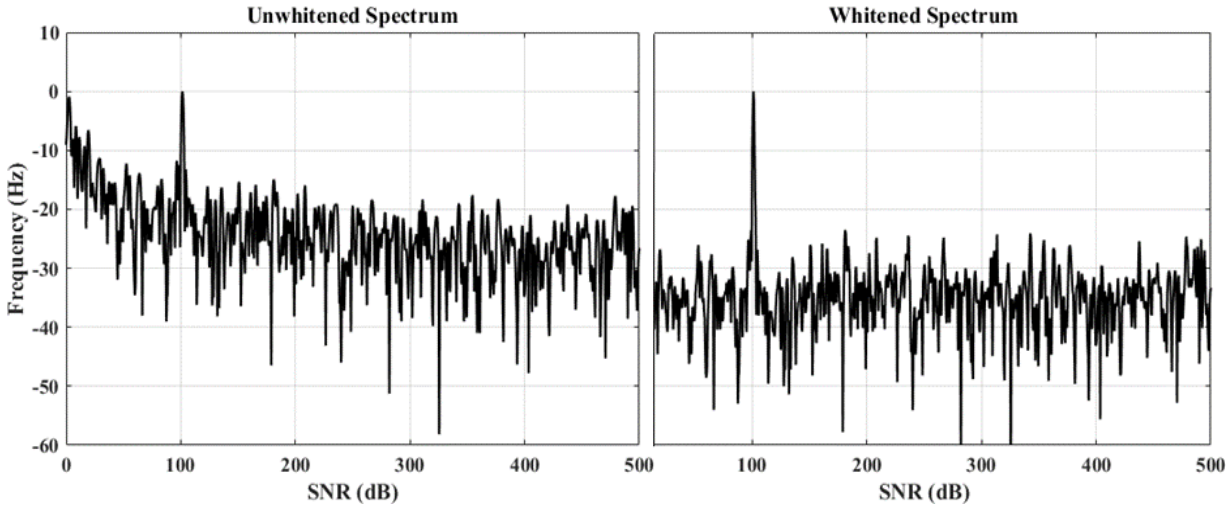

Figure 4-20: Power spectra of standard and whitened signals.

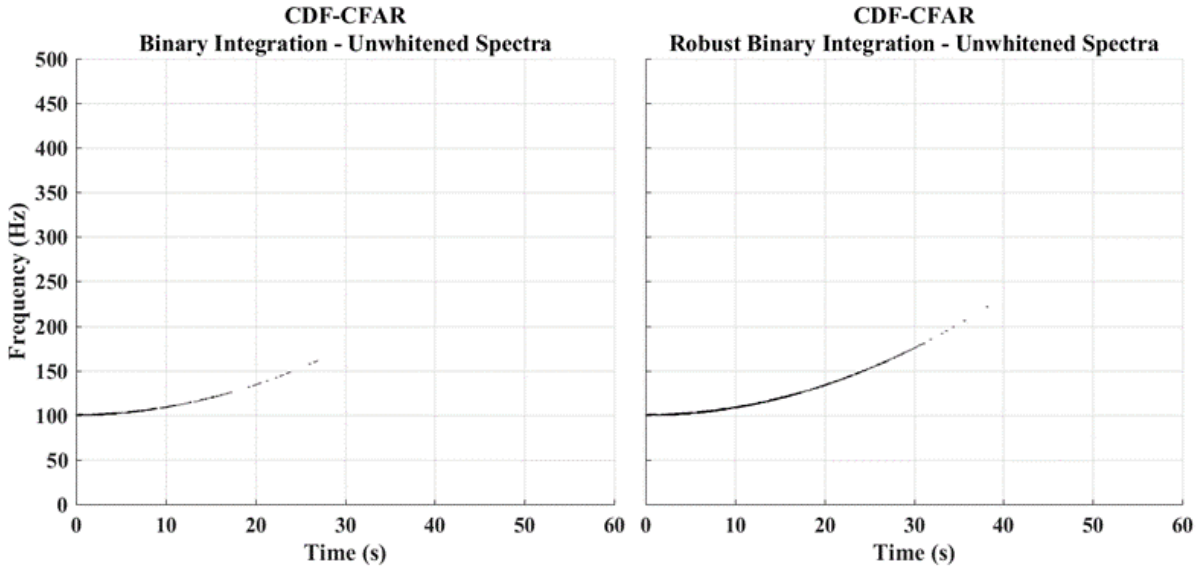

Figure 4-21: Detection plots for CDF-CFAR binary integration tests.

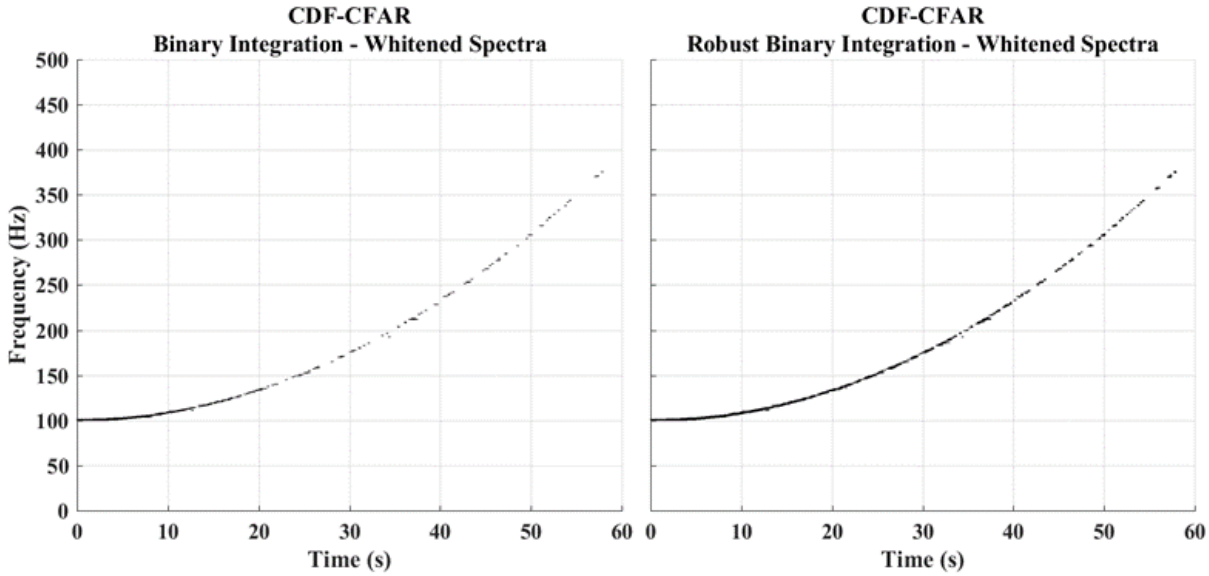

Figure 4-22: Detection plots for CDF-CFAR robust binary integration tests.

- 5 - Source Localization

The following chapter presents several array processing techniques which may be utilized to spatially locate acoustic source targets. A review of common localization methods is first provided, followed by a number of algorithms developed to reduce computational loads for the application of concern. In addition, a beamforming algorithm is proposed which exploits the properties of harmonic narrowband signals to enhance localization capabilities. A brief comparison of the proposed method to standard frequency domain beamformers is also provided.

5.1 - Introduction

Beamforming is a processing technique in which an array of sensors are used together to enhance the directional reception for a signal of interest. The array effectively acts as a spatial filter in that desired signals arriving from some location can be enhanced, while simultaneously attenuating undesired signals arriving from some other. The underlying concept behind beamforming is to utilize signal phase characteristics to constructively combine desired components while attenuating undesired components. The topic of beamforming has been extensively reported in the literature with many applications in the areas of acoustics, radar, sonar, imaging, and communications to name a few [242-244]. The subject area is extremely vast with many algorithms having been developed since the concept was first widely introduced during the early 1960s [245, 246]. Thus, only concepts and algorithms directly relevant to the application at hand will be addressed.

Consider now the localization scenario pertaining to this thesis; we wish to determine the angular position of some target source relative to the detecting aircraft. It is expected that received signals will be harmonic and continuous in time. Due to the time sensitive nature of the operation, algorithms must be capable of operating in real-time with low latency. Due to physical limitations associated with placing an array on an aircraft, it is also very unlikely that a uniform microphone spacing can be achieved; such is the case for experiments presented in this thesis. In addition, vibrations and unsteady airflow during flight operations will inherently produce motions to some degree in the array elements effectively producing positional errors. This is coupled with the fact that practical limitations during construction and installation of the array will also produce some degree of positional error with respect to the original intended design. Taking the above features into consideration, it is apparent that a robust beamforming method is required to ensure maximum operational performance for these adverse conditions. Since target detection is performed in the frequency domain, it is also desirable that localization operations be performed in this domain also. Although beamforming in the frequency domain is less common, increased computational efficiency can be achieved since target source frequencies will already have been identified in the previous detection stage.

5.2 - Background Information

5.2.1 - Conventional Beamforming

Similar to standard digital filtering, beamforming may be performed in the time or frequency domain, can be fixed or adaptive, and may be applied to narrowband or broadband signals. Narrowband beamformers focus on a narrow frequency band of interest and only filter signals across sensors in the spatial dimension. A weighting vector is applied across signals to facilitate steering of the array and promote side-lobe reduction in the directivity response. For time domain beamformers, the weighing vector typically consists of real valued numbers in combination with time delay elements, while frequency domain beamformers simply use complex numbered values [243]. For narrowband beamformers to remain maximally effective, the signal of interest should remain correlated between the closest and furthest elements of the array. That is, the time delay between the closest and furthest element should be less than the period for the signal of interest. For situations in which the above conditions are not met, the signal is instead considered broadband. For such cases, the array directivity response will vary greatly across the bandwidth for a given set of complex weights. In order to alleviate this issue and maintain a constant response, broadband beamformers apply filtering in both the time and spatial domains. Because of the variation in directivity response with respect to frequency, broadband beamformers are inherently more complex and are thus more difficult to deign with respect to physical array configurations.

The most simplistic and commonly employed beamformer is the Delay-and-Sum form depicted in Figure 5-1. Here we consider a linear array consisting of S sensors/microphones with locations given by $\vec{r}_s$ where $s \in \{1,2,\dots,S\}$. The principle behind the approach is to simply apply an appropriate time delay such that signals become aligned to produce a coherent amplified output when summed together. This can be expressed as:

$$y(\hat{k},t) = \sum_{s=1}^{S} w_s \; x\left(t + \tau_s(\hat{k})\right) \tag{5.1}$$

where w_s are weighting values applied to each channel to modify directional output characteristics. A number of methods exist to determine the optimum weighting values for a given system. Common methods include the Minimum Variance Distortionless Response (MVDR), Minimum Mean Square Error (MMSE), Maximum Signal-to-Noise Ratio (MSNR), and Minimum Power Distortionless Response (MPDR) beamformers to name a few [242]. In most real-world situations however, the performance advantage produced by optimal forms such as those previously listed are greatly lessened due to incomplete knowledge of the signal and noise spectral content [247].

The time delays τ_s are characterised by the steering direction which is given by the unit vector $\hat{k}$. The maximum output will be achieved when $\hat{k}$ aligns with the incoming wave propagation direction. For a given steering direction, the required time delay for the s^{th} microphone will be thus given by:

$$\tau_s = \frac{\hat{k} \bullet \vec{r}_s}{c} \tag{5.2}$$

where c is the speed of sound in air. For 3-dimensional space characterized by the azimuth and elevation angles ϑ and φ respectively as previously depicted in Figure 1-1, the steering and position vectors will be given by the following:

$$\hat{k} = \begin{bmatrix} \cos(\vartheta)\cos(\varphi) \\ \sin(\vartheta)\cos(\varphi) \\ \sin(\varphi) \end{bmatrix} \tag{5.3}$$

$$\vec{r}_s = [x_s, y_s, z_s]^T \tag{5.4}$$

where the spherical coordinate system was previously defined in Figure 2-9.

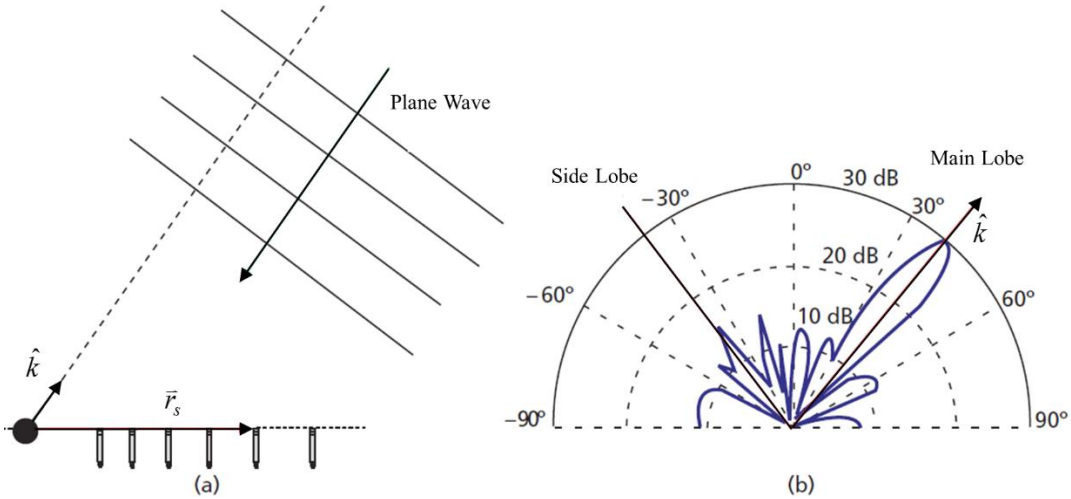

Figure 5-1: a) A microphone array with plane wave incident from the focus direction. b) A typical array directional response plot with a main lobe in the focus direction and lower side lobes in other directions [248].

Alternately, the Delay-and-Sum beamformer may also be expressed in the frequency domain instead. Applying the Fourier transform and realizing that a delay in the time domain equates to phase shift in the frequency domain yields the following:

$$Y(\hat{k}, f) = \sum_{s=1}^{S} W_s(f) \cdot X_s(f) \cdot e^{-j2\pi f \tau_s(\hat{k})} \tag{5.5}$$

Note that the above form is often referred to as the Filter-and-Sum (FAS) beamformer [249]. To avoid spatial aliasing in the array directivity response, the minimum sensor spacing must be less than half the incident wave length [249]:

$$d < \frac{\lambda_{\min}}{2} \tag{5.6}$$

where $\lambda_{\min}$ is the wavelength for the signal of interest. Spatial aliasing reduces the ability to localize a desired source through the presence of grating or side lobes and should thus be avoided whenever possible.

In order to compare various beamforming algorithms and determine overall performance, a number of measures are often used. Evaluation of the array output for all possible steering directions is referred to as the array

response plot or beam pattern as depicted in Figure 5-1. It is given by the magnitude squared of the array output and is typically expressed in decibel units:

$$B(\hat{k}, f) = \left| Y(\hat{k}, f) \right|^2 \tag{5.7}$$

The directivity gain is a measure of the maximum power output in a given steered direction compared to the average noise power in all other directions. Thus, the array gain in the direction specified by $\hat{k}_o$ is given by the following equation for the case of isotropic sound in the free field [242]:

$$D(\hat{k}_o, f) = \frac{B(\hat{k}_o, f)}{\frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} B(\hat{k}, f) \sin \vartheta d\theta d\varphi} \tag{5.8}$$

The ability of an array to localize a source with a high degree of resolution is dependent upon the beam pattern. More specifically it is dependent on the mean lobe beam width which is defined as the angular distance between the two-half power (3 dB) points on the main lobe. This is also known as the Half Power Beam Width (HPBW). Figure 5-2 displayed below illustrates the HPBW and approximate directivity gain for a 1-D linear array.

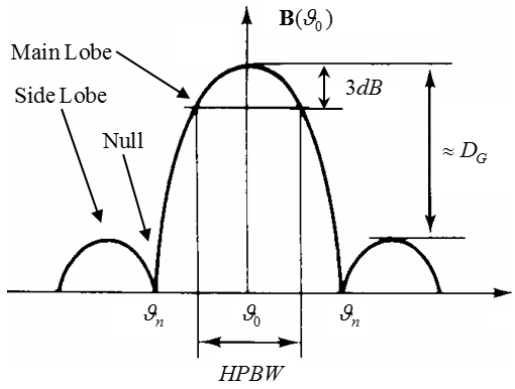

Figure 5-2: Half Power Beam Width (HPBW).

5.2.2 - Direction of Arrival Estimation

Beamformers generally have two main applications: 1) enhancing signals arriving from some location of interest, and 2) finding the spatial location for signals of interest. The process of spatially localizing signals is known as Direction of Arrival (DOA) estimation. There have been an abundance of DOA algorithms presented in the literature [245, 246, 250, 251]. However, most can be classified into the following three categories: 1) Subspace or Beam-space, 2) Time Difference of Arrival (TDOA), and 3) Steered Response Power (SRP). Each of these methods have inherent advantages and disadvantages, which ultimately dictate appropriate applications areas as will now be discussed.

5.2.2.1 - *Subspace Methods*

Subspace algorithms have arrived more recently in the field of beamforming, with algorithms generally offering superior localization accuracy over more conventional methods. In general, the technique uses an approach whereby the signal and noise are separated into separate component subspaces using an eigenvalue

decomposition of a covariance or correlation matrix. The MUSIC, Root-MUSIC, and ESPRIT are commonly reported examples of this algorithm type [251]. Although these methods have been successfully employed for a variety of array processing applications, they all possess certain restrictions that prevent their practical use for the application at hand. Some of these include: the requirement of a uniformly spaced array, high computational loads which often prevent real-time operation, and a high sensitivity to array positioning errors [245, 250, 252]. Physical limitations associated with aircraft geometry often prevent the installation of a uniformly spaced array; such is the case for experimental aircraft presented in this thesis. In addition, vibrations and unsteady airflow during flight operations will inherently produce some degree of sensor motion generating positional errors. This is coupled with the fact that practical limitations during construction and installation of the array will also produce some degree of positional error with respect to the original intended design. Although some degree of positional error can be removed via calibration procedures, those occurring during flight operations cannot be mitigated through signal processing means. Thus, with respect to the application at hand, it is apparent that such methods are not robust enough for such physically demanding operational requirements.

5.2.2.2 - *Time Difference of Arrival*

Time Difference of Arrival (TDOA) is perhaps the most commonly employed DOA method. This is largely due to its simplicity and low computational requirements which allows real-time implementation on most all digital systems [245, 247, 253-255]. Compared to other methods such the Subspace and Steered Response Power, TDOA methods offer a significant computational advantage [247, 254]. In general, the localization procedure involves the generation of hyperbolic curves which are then intersected in some optimal sense to achieve the location estimate [247]. Typically, a method such as the Least Squares approach is employed to obtain the optimal fit statistic [256]. The approach is generally a two-stage procedure whereby time delay estimates between sensor pairs are first calculated, followed by an approximation of the angular source location. Time delay estimates are typically obtained via the Generalized Cross-Correlation (GCC) function developed Knapp and Carter [257]:

$$R_{12}(\tau) = \sum_{f=0}^{L-1} \Upsilon(f) X_{12}(f)\, e^{\frac{j2\pi f \tau}{f_S}} \tag{5.9}$$

where τ is the time lag index, L is the segment length, Υ is the general weighting function, f_s is the sampling frequency, and X_{12} is the cross-power spectrum. A number of weighting functions have been presented in the literature for the purpose of enhancing TDOA estimates for various localization scenarios. The most popular of these is the PHAT processor [258]:

$$\Upsilon_{PHAT}(f) = \frac{1}{\left| X_1(f) \cdot X_2^*(f) \right|} \tag{5.10}$$

where * indicates complex conjugation. It is evident from the above equation that the PHAT weighting function is essentially a whitening filter which removes all magnitude information leaving only the phase content to determine TDOA values.

Determination of the time delay value is then given by the maximum peak of the correlation function:

$$\hat{\tau}_{12} = \arg\max\left[R_{12}(\tau)\right] \tag{5.11}$$

while angular source locations can be obtained through application the Least Squares error criterion according to the following [254]:

$$J(\vartheta,\varphi) = \sum_{i=1}^{P}\left[\hat{\tau}_i - \tau(\vartheta,\varphi)\right]^2 \tag{5.12}$$

where P is the number of sensor pairs, and τ is the true or expected phase delay as previously given by equation (5.2). The angular location estimate is finally obtained from the values which minimize the cost function:

$$\left\langle \hat{\vartheta},\hat{\varphi} \right\rangle = \arg\min\left[J(\vartheta,\varphi)\right] \tag{5.13}$$

Although TDOA methods have lower computational requirements, they often suffer from poor resolution and deteriorate extensively in the presence of multiple sources, reverberation, and low SNR environments [254]. This is essentially due to the presence of local maxima in the cross-correlation function which may obscure the true TDOA peak and subsequently produce incorrect delay estimates. The amplitudes of these erroneous maxima depend on a number of factors such as ambient noise levels and reverberation conditions. For the application at hand, this is very problematic since the SNR values of initially detected signals will be very low and multipath conditions may be present if operating at low altitudes. Thus, TDOA methods may not offer acceptable performance for the localization scenario pertaining to this thesis.

5.2.2.3 - *Steered Response Power*

The Steered Response Power (SRP) is another commonly employed localization approach. It is robust, simplistic in nature, and generally performs considerably better than TDOA methods in adverse environments [259]. In brief, the method forms a directional “beam” which is scanned over a region of interest while calculating the power output at each location. The source location is chosen based on the direction which maximizes this output power. For the case of the Filter-and-Sum beamformer, the SRP will be given by:

$$\mathrm{SRP}\left[Y(\hat{k},f)\right] = \sum_{f=f_a}^{f_b}\left|Y(\hat{k},f)\right|^2 = \sum_{f=f_a}^{f_b}\sum_{s=1}^{S} W_s(f)\cdot X_s(f)\cdot e^{-j2\pi f\tau_s(\hat{k})} \tag{5.14}$$

where S is the total number of acquired microphone signals, and $[f_a, f_b]$ defines the discrete frequency range of interest.

The angular location estimate can then be obtained by finding the direction which maximizes the output power:

$$\left\langle \hat{\vartheta}, \hat{\varphi} \right\rangle = \arg\max\left(\mathrm{SRP}\left[Y(\hat{k}, f) \right] \right) \tag{5.15}$$

Similar to the TODA method previously presented, elements of the GCC function can be included to increase performance in adverse environments. The most popular form utilizes the PHAT weighting function to whiten spectra by removing magnitude information from the SRP output [260]. Often termed the SRP-PHAT, this form has been reported extensively in the literature due to its increased performance over standard SRP methods in reverberant environments [261-264]. It is given by the sum of all steered microphone correlation pairs according to:

$$Y(\hat{k}, f) = \sum_{n=1}^{S} \sum_{s=n+1}^{S} \varphi_{PHAT}(f) \cdot X_s(f) \cdot X_n^*(f) \cdot e^{-j2\pi f \tau_{sn}(\hat{k})} \tag{5.16}$$

where $\tau_{sn} = \tau_s - \tau_n$. It should be noted however that the SRP-PHAT is only advantageous for broadband signals and/or highly reverberant conditions. For single component narrowband signals, performance tends to degrade since the maximum value attainable by the coherently aligned signal will be no greater than that possible by the random noise components.

The major benefit to the SRP method is simplicity, robustness, and the ability to utilize an unconstrained array geometry [254]. The major downside is increased computational loads since direct closed-form solutions cannot be achieved to estimate localization values. Instead, grid-search methods are employed to scan a region of interest. If the source location is completely unknown, this may result in unacceptably large computational loads. For example, if scanning the full spherical region surrounding an aircraft with a 1-degree resolution, a total of $360 \times 180 = 64{,}800$ evaluation points are required.

A number of algorithms have been proposed to reduce computational loads and facilitate real-time operations. These can be broadly categorized into three main areas: 1) Regional reduction through TDOA-based candidate location mapping [265-269], 2) Regional contraction using coarse-to-fine grid searching [261, 270-272], volumetric evaluation [262, 273], or stochastic methods [263, 264], and 3) Iterative-based search techniques [274, 275]. TDOA-based methods utilize time delays to determine a region consisting of potential candidate locations. Standard search methods are then applied to the reduced region to accurately determine source locations. Although shown to be effective, the downside to this approach is the possibility of poor initial region specifications due to inaccurate TDOA estimates. Regional contraction methods essentially perform multiple regional searches using successively smaller grid sizes until the desired accuracy is achieved. This method is simplistic in nature and a localization convergence can always be guaranteed. However, because the SRP space will generally be composed of many local maxima, there is a direct trade-off between accuracy, computational load, and convergence accuracy. Iterative-based techniques utilize methods such as the Steepest Decent and Newton-Raphson to obtain localization values. These methods offer the greatest computational savings but generally suffer from the possibility of false convergence from the presence of local maxima. This is increasingly problematic for low SNR conditions since the objective function will not have a strong global

peak. For such situations, iterative methods may produce inaccurate results which are extremely sensitive to the initial search locations [254]. Thus, with respect to the above considerations, it appears that the SRP method using a regional contraction approach may be best suited for the application at hand. It will later be shown that such methods can easily be employed using minimal computational requirements for instances in which potential source frequencies are known.

At this point one may question why use the SRP approach over phase-based methods if the signal frequency is already known. The simple answer to this question is increased accuracy since the SRP method uses both amplitude and phase information to determine the source location. In addition, it is more robust to aspects such as spectral leakage since multiple frequency bins may be efficiently utilized for the operation instead of just one. For example, if one were to account for potential leakage using phase-based methods, a phase correlation matrix (between microphone signal pairs) must be established for each frequency bin and solved using a method such as the least squares approach. The result obtained for each frequency must then be averaged together, typically using some form of weighting function. This operation is clearly more complex than simply extending the summation domain for the SRP as indicated above in equation (5.14)

5.3 - Source Localization

The following section presents a number of source localization methods for use with the SRP beamformer. Using the concept of regional reduction through coarse-to-fine grid searching, a crisscross search method is proposed which offers reduced computational loads compared to standard techniques such as that presented in [261]. In addition, a two-dimensional gradient ascent method is proposed which is also capable of achieving increased performance compared to standard grid searching methods.

5.3.1 - Regional Contraction

Regional contraction using coarse-to-fine grid searching is a commonly-employed technique to reduce computational loads when using the SRP method [261, 270-272]. In general, the method assigns a coarse grid to the spatial region of concern with SRP values being calculated at each point. The grid location producing the maximum power is then identified and subsequently used to establish a new search region which is defined by the neighbouring grid points. The grid spacing is then reduced, assigned to the new reduced region, and the maximum point is again located. The process is simply repeated until the desired localization resolution is achieved. Figure 5-3 displayed below provides a visual depiction of the process where $\times$ indicates the maximum SRP point in the region of interest.

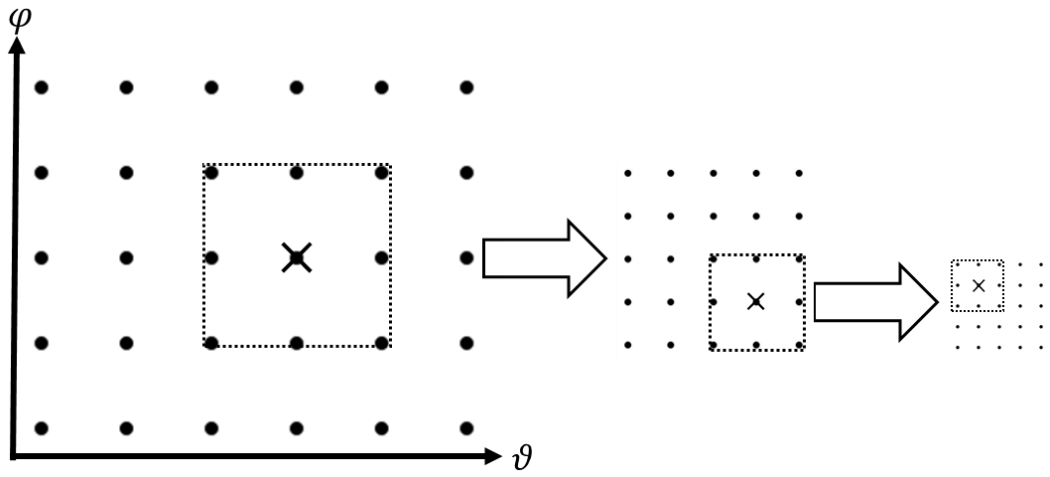

Figure 5-3: Illustration of coarse-to-fine grid search method.

As previously discussed, one of the issues associated with regional reduction and iterative-based methods is the false convergence due to the presence of local maxima and a weak global peak. To combat this issue, initial grid spacing is often chosen conservatively which consequently results in unnecessary computation. At this point, it should be emphasised that all reported methods inherently assume source frequencies are unknown. For such instances, the SRP must be calculated across a frequency band of interest or in some cases the entire signal bandwidth $[0, f_s/2]$. For narrowband signals, this will inherently reduce the effectiveness of the approach since the addition of random noise components will reduce the dynamic range of the SRP output. Such methods will also produce high computational requirements since the steered response requires calculation for every FFT bin in the frequency band; hence the reason why SRP methods are often considered computationally expensive. For the application at hand however, this is not the case since all possible source frequencies would have been identified in the previous CFAR detection stage. By exploiting this information, a much higher degree of sensitivity can be achieved in addition to great computational savings.

For a detected source signal of frequency f_o, the SRP output will now be given by:

$$\mathrm{SRP}\left[Y(\hat{k}, f)\right] = \sum_{f=f_o} \left|Y(\hat{k}, f)\right|^2 \tag{5.17}$$

which requires $f_b - f_a - 1$ less evaluation points, where $[f_a, f_b]$ defines the frequency band of interest. To establish a more robust form, the region closely surrounding f_o may also be evaluated to account for spectral leakage. Thus, if G guard cells taken equally about the detected frequency are also included, the SRP will now be given by:

$$\mathrm{SRP}\left[Y(\hat{k}, f)\right] = \sum_{f=f_o-G/2}^{f_o+G/2} \left|Y(\hat{k}, f)\right|^2 \tag{5.18}$$

By excluding essentially all noise components contained in the frequency band of interest, the sensitivity of the method will inherently increase, which is especially true for low SNR signals. This is demonstrated in Figure 5-4 and 5-5 which display plots of the SRP as a function of frequency for two SNR scenarios (-15 and 0 dB).

The signal consists of a $f_o = 500$ Hz sinusoid embedded in Gaussian noise sampled at a rate of $f_s = 2000$ Hz and transformed using a 2000-point FFT. It arrives with 45-degree angle of incidence and is acquired by a uniform linear array consisting of 4 sensors spaced 0.25 m apart. Note that -15 dB is the typical lower limit for reliable CFAR detection using the average coherent power scheme as was demonstrated in the previous chapter. The isolated component SRP (P_1) was calculated using G= 5 Hz, while the standard broadband output (P_2) was calculated for the full spectral range $[0, f_s/2]$. From the plots, it is evident that a significant increase in localization sensitivity is obtained for low SNR values. It is apparent that using an adaptive search approach would not fare well for the broadband case with low SNR values, since the presence of multiple peaks with approximately equal value will greatly increase the probably of false convergence. For the isolated component case however, localization peaks are significant and should lend well to regional reduction and/or iterative search techniques.

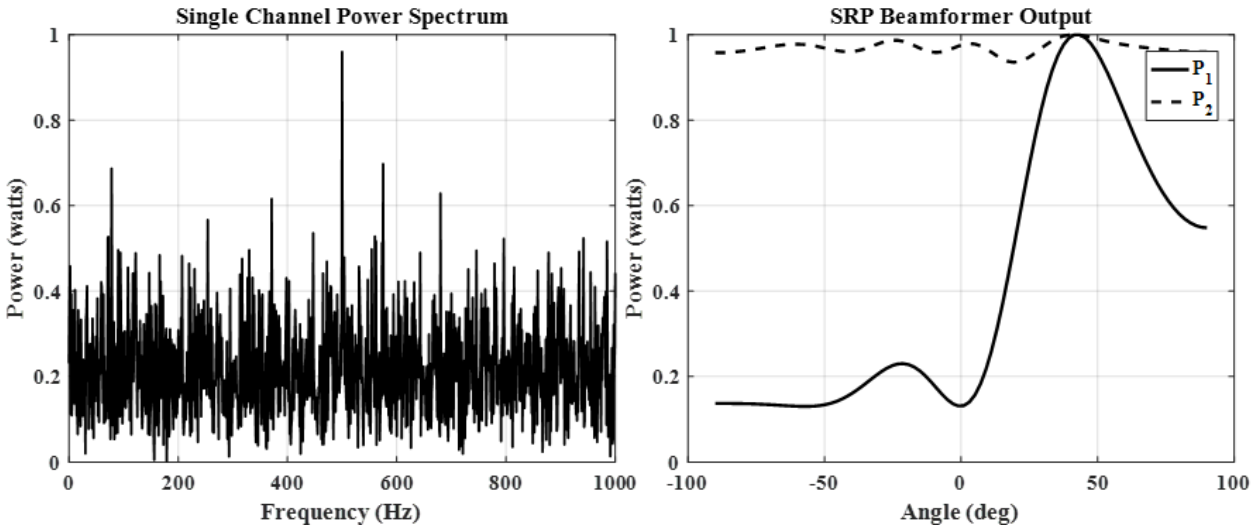

Figure 5-4: Power spectrum and SRP output for -15 dB signal, where P_1 and P_2 indicate the isolated and broadband SRP response cases respectively.

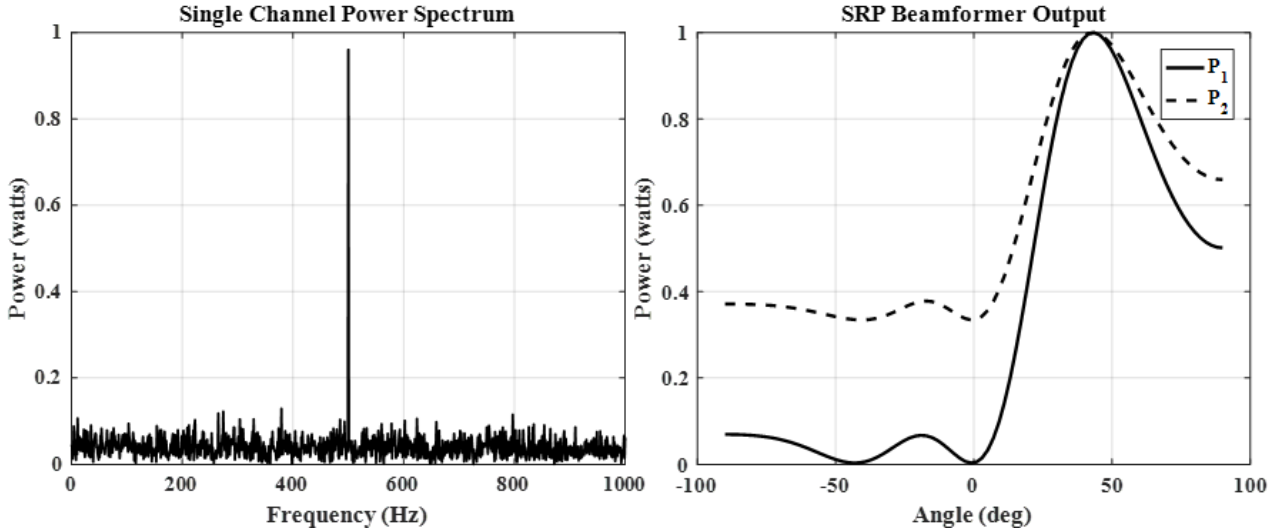

Figure 5-5: Power spectrum and SRP output for 0 dB signal, where P_1 and P_2 indicate the isolated and broadband SRP response cases respectively.

With respect to the above considerations, a regional reduction method is now proposed which utilizes a crisscross search method rather than a uniformly spaced square grid approach. The proposed algorithm is performed as follows: A very coarse uniform 2-dimensional grid is first applied to determine the general maxima region in terms of azimuth and elevation. While holding the elevation angle constant at the median point for the range of interest, SRP values are calculated for the desired range of azimuth values. The azimuth value producing the largest SRP output is then chosen and held constant while SRP values are then calculated for the desired range of elevation angles. The grid region is then chosen about this point, contracted, and the process is repeated. This is depicted below in Figure 5-5 where $\odot$ indicates the first regional max value (holding φ constant), and $\times$ indicates the final regional max. It is evident that the approach offers great computational savings since only one row and one column for each grid set is evaluated rather than all points. For example, if performing a three-level coarse-to-fine regional contraction using a 5×5 grid, a total of $5 \times 5 \times 3 = 75$ points would require evaluation. However, if using the criss-cross approach only $5 \times 5 + 2 \times 5 + 2 \times 5 = 45$ points would be required which is a significant reduction.

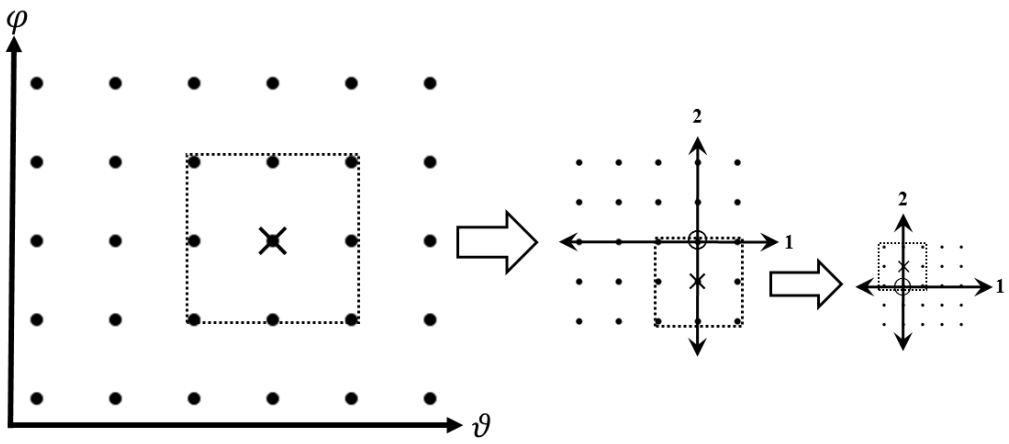

Figure 5-6: Crisscross regional reduction grid search.

It is acknowledged that the method may not be suitable for all SRP array configurations, but nevertheless it performs well for the application at hand. Consider the 2-dimensional 6-element array utilized by the Kraken multirotor as displayed by Figure 6-9. The array acquires a 100 Hz sinusoidal signal of -10 SNR dB from an azimuth and elevation of 45 and 36 degrees respectively. The SRP is given for the full detectable range which is a hemisphere in this case since cardioid directional response microphones were used. It is apparent from the plot that the SRP is highly sensitive to azimuth changes and less sensitive to elevation changes. Thus, if one were to randomly choose an elevation angle, a good approximation of the true azimuth angle can be achieved by scanning it across the region of interest. In fact, it is evident from the plot that a good approximation may be obtained for all elevation angles less than approximately 75 degrees. Thus, it is very unlikely that the proposed crisscross method would fail to converge at the true global maxima.

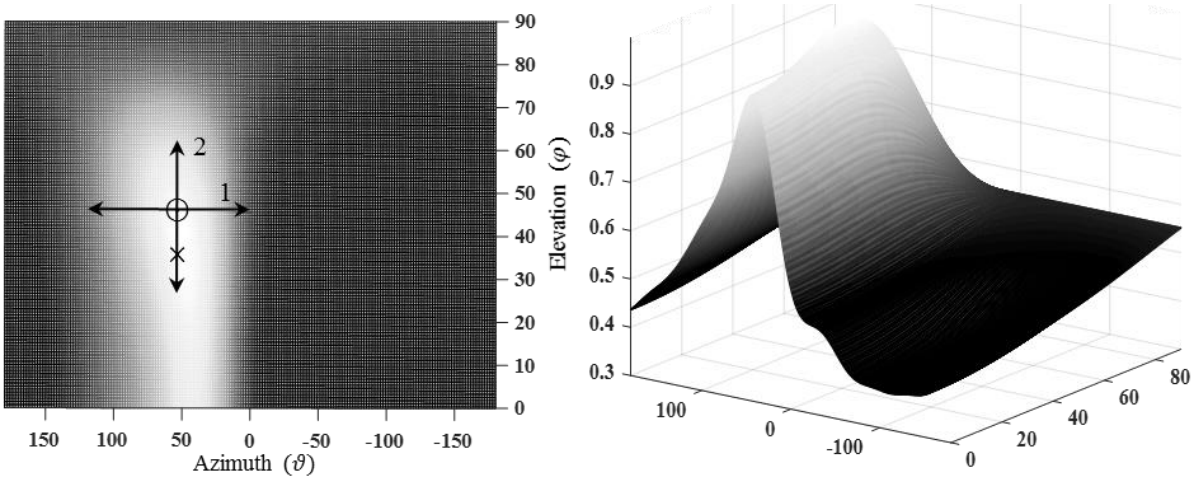

Figure 5-7: Illustration of crisscross regional convergence for 100 Hz signal acquired by Kraken array arriving with azimuth and elevation angles of 45 and 36 degrees respectively.

5.3.2 - Steepest Ascent

Direction of arrival estimation for SRP beamformers using iterative techniques are seldom used and rarely reported in the literature. Wax [275] first proposed the idea of using iterative techniques such as the Steepest Descent or Newton-Raphson method, however he did not provide any analysis or example of using such approach. Marti et al [274] used an iterative based approach in conjunction with a correlation search method to adaptively determine DOA values. Although successful, the method does not fall in the realm of typical iterative procedures such as those proposed by Wax and often found in adaptive filtering applications. Here, an adaptive approach is proposed which utilizes the steepest ascent gradient method in conjunction with the direct SRP output to approximate DOA values. Since potential source frequencies are known in advance, noise components may be omitted from the SRP output which greatly reduces the presence of local maxima and likelihood of false convergence as previously discussed. The 1-D case is first proposed since this type of array was utilized for experiments conducted using fixed-wing aircraft. The 2-D case is then presented which can be applied to any array configuration.

Consider again the general problem of DOA estimation for the SRP Beamformer: We wish to find the angle ϑ that maximizes the array output for the signal component of interest. Application of the gradient ascent method gives:

$$\vartheta_{n+1} = \vartheta_n + \mu_\vartheta \nabla J_\vartheta[n] \tag{5.19}$$

where μ_ϑ is the step size, n is the iteration index number, and J is the cost (or in this case profit) function which is given by the SRP output:

$$J(\vartheta_n) = \sum_{f=f_o} \left| Y(\vartheta_n, f) \right|^2 \tag{5.20}$$

where f_o is the fundamental frequency of the desired signal component as identified during the CFAR detection stage. The gradient of the profit function with respect to the optimization parameter may be approximated by the backwards finite difference:

$$\nabla J_{\vartheta}[n] = \frac{\partial}{\partial \vartheta} J[\vartheta_n] \approx \frac{J[\vartheta_n] - J[\vartheta_{n-1}]}{\vartheta_n - \vartheta_{n-1} + \varsigma \cdot \mathrm{sgn}[\vartheta_n - \vartheta_{n-1}]} \tag{5.21}$$

where ς is a small constant to avoid division by zero, and sgn[ ] indicates the sign of the contained expression. We may also define an error function to serve as a convergence standard for the adaptive operation:

$$\varepsilon_{\vartheta} = \left|\vartheta_n - \vartheta_{n-1}\right| \tag{5.22}$$

For each signal segment, we simply perform the adaptive iteration until the error value reaches some pre-specified threshold value. Thus, the algorithm is implemented as follows:

1) Use a standard coarse grid search to find the general global maxima region.
2) Calculate initial gradient value using the identified maximum and its closest neighbour.
3) Perform the iteration procedure using equations (5.19), (5.20), and (5.21) until the angular position error reaches the minimum desired accuracy as given by (5.22)

For the 2-dimensional case, the objective becomes determining the azimuth ϑ and elevation φ angles which maximize the SRP profit function J which is now given by:

$$J(\vartheta_n, \varphi_m) = \sum_{f=f_o} \left|Y(\vartheta_n, \varphi_m, f)\right|^2 \tag{5.23}$$

This may be achieved by applying the gradient ascent approach to each angular direction using an iterative crisscross approach similar to that described in the Regional Contraction section. For this case, the gradient ascent equation is now given by the following:

$$\left\langle \vartheta_{n+1}, \varphi_{m+1} \right\rangle = \left\langle \vartheta_n, \varphi_m \right\rangle + \mu_{\vartheta,\varphi} \left\langle \nabla J_{\vartheta}[n], \nabla J_{\varphi}[m] \right\rangle \tag{5.24}$$

Adaptive implementation may be achieved by applying the method to each angular direction separately. For the azimuth direction this may be achieved by the following equations respectively:

$$\nabla J_{\vartheta}[n] = \frac{J[n,m] - J[n-1,m]}{\vartheta_n - \vartheta_{n-1} + \varsigma \cdot \mathrm{sgn}[\vartheta_n - \vartheta_{n-1}]} \tag{5.25}$$

$$\vartheta_{n+1} = \vartheta_n + \mu_{\vartheta} \nabla J_{\vartheta}[n] \tag{5.26}$$

$$J[n+1,m] = \sum_{f=f_o} \left|Y(\vartheta_{n+1}, \varphi_m, f)\right|^2 \tag{5.27}$$

$$\varepsilon_{\vartheta} = \left|\vartheta_{n+1} - \vartheta_n\right| \tag{5.28}$$

Similarly, the elevation direction is given by the following equations:

$$\nabla J_{\varphi}[m] = \frac{J[n,m] - J[n,m-1]}{\varphi_m - \varphi_{m-1} + \varsigma \cdot \mathrm{sgn}[\vartheta_m - \vartheta_{m-1}]} \tag{5.29}$$

$$\varphi_{m+1} = \varphi_m + \mu_{\varphi} \nabla J_{\varphi}[m] \tag{5.30}$$

$$J[n,m+1] = \sum_{f=f_o} \left|Y(\vartheta_n, \varphi_{m+1}, f)\right|^2 \tag{5.31}$$

$$\varepsilon_\varphi = \left|\varphi_{m+1} - \varphi_m\right| \tag{5.32}$$

Finally, the cost function is updated for the next iteration using the latest azimuth and elevation angles $(\vartheta_{n+1}, \varphi_{m+1})$:

$$J[n+1,m+1] = \sum_{f=f_o} \left|Y(\vartheta_{n+1}, \varphi_{m+1}, f)\right|^2 \tag{5.33}$$

As with the 1-dimensional case, iterations are performed until both ε_ϑ and ε_φ reach some pre-specified value.

5.4 - Harmonic Spectral Beamformer

A beamforming method which exploits the properties of harmonic signals to spatially localize acoustic sources is now proposed. Termed the Harmonic Spectral Beamformer, the method combines the Complex Harmonic Spectral Transform with the standard Filter-and-Sum beamformer to construct a processor which produces a high degree of directional sensitivity.

Harmonic signals such as those generated by propeller-driven aircraft can be considered broadband with a narrowband decomposition structure [243]. Typically, in order to retain an optimal response for each harmonic component, the signal is decomposed into multiple segments which are then subject to their own beamforming algorithm [243]. Although effective, this method requires a much higher computational load since a number of independent algorithms are effectively required instead of just one. A potential solution to this problem is to simply transform the signal such that the harmonics are now represented by a single fundamental component. This procedure was previously described by the HSTs presented in the last chapter. However, the proposed form as given by equation (4.30) does not contain phase information which is required to facilitate beamforming operations. Thus, the phase preserving Complex Harmonic Spectral Transform (CHST) is first defined according to the following equation:

$$j\text{-}\bar{\mathrm{H}}_a[X(f)] = \left[\frac{1}{R}\sum_{r=1}^{R}\left|X(f\cdot r)\right|^a\right]^{1/a} \cdot e^{\,j\sum_{r=1}^{R}\theta(f\cdot r)} \tag{5.34}$$

where the j in the variable definition $j\text{-}\bar{\mathrm{H}}_a[]$ indicates the complex spectral transform, and the phase function θ is given by :

$$\theta(f\cdot r) = \arg[X(f\cdot r)] \tag{5.35}$$

Note that the phase transform operation is performed independently of the spectral peak operation. This is to avoid any possible reduction in the transformed spectral peaks, which will occur if harmonic components are out of phase with one another. In addition, the phase forms are summed using the scalar rather than vectoral method, to prevent cancelation of between different harmonics. It should be noted at this point that it is assumed that the harmonic signals of concern are generated acoustically by a physical means such as that found with

voiced speech, musical instruments, and aircraft propulsion systems [72, 135-137]. For such cases, the harmonic nature of the signal arises from the presence of a physical boundary which establishes the condition for standing wave generation and thus initial phase alignment between components [138].

By combining the CHST (5.34) and Filter-and-Sum (FAS) (5.5) equations, we may thus define a new Harmonic Spectral Beamformer (HSB) according to the following:

$$\mathrm{HSB}(\hat{k}, f) = \sum_{s=1}^{S}\left(W_s(f) \cdot \left[\frac{1}{R}\sum_{r=1}^{R} \left| X(f \cdot r) \right|^a \right]^{1/a} \cdot e^{j\left[\sum_{r=1}^{R} \theta_s(f \cdot r) - 2\pi \cdot f \cdot \tau_s(\hat{k}) \sum_{r=1}^{R}(r) \right]} \right) \tag{5.36}$$

where the final SRP output can be obtained by applying equation (5.18) to the above form:

$$\mathrm{SRP}\left[\mathrm{HSB}(\hat{k}, f) \right] = \sum_{f=f_o-G/2}^{f_o+G/2} \left| \sum_{s=1}^{S}\left(W_s(f) \cdot \left[\frac{1}{R}\sum_{r=1}^{R} \left| X(f \cdot r) \right|^a \right]^{1/a} \cdot e^{j\left[\sum_{r=1}^{R} \theta_s(f \cdot r) - \sum_{r=1}^{R}\left(2\pi \cdot f \cdot r \cdot \tau_s(\hat{k}) \right) \right]} \right) \right|^2 \tag{5.37}$$

The rationale behind summing phase angles across harmonic components is to ultimately provide an increased array response sensitivity. This is essentially achieved because a sight change in steering direction will now produce a much larger change in the phase function, compared to that found in the standard FAS Beamformer given by equation (5.5). This property is facilitated by the physical nature of the acoustic system whereby the initial phase of each harmonic component is equal at the point of generation. Although the phase values for each harmonic may be different at any given point thereafter, summation will still be maximized when the array is steered in the direction of wave propagation. To avoid spatial aliasing, the element spacing must now adhere to the following condition:

$$d \le \frac{\lambda_o}{2\sum_{r=1}^{R} r} = \frac{c}{2\sum_{r=1}^{R} f_o \cdot r} \tag{5.38}$$

where λ_o and f_o is the fundamental wavelength and frequency respectively. The above relationship can arrived by considering the standard condition for spatial aliasing as given by equation (5.6), and the logical operation of summing each harmonic phase in a scalar manner. Since the phase of each harmonic component is combined across time, and frequency is defined as the rate of change in phase with time; the effective frequency of the harmonically transformed wave is given by the sum of each harmonic frequency. For example, if a 100 Hz and 200 Hz signal were combined by the scalar sum of their phase functions, the resulting wave would have an effective frequency of 300 Hz.

Figure 5-8 and 5-9 provides normalized directional response plots obtained via the standard FAS and HSB beamformers for a four-element Uniform Linear Array (ULA). Response values are provided over a range of spacing values and signal harmonics for a 100 Hz sinusoidal signal. The Half Power Beam Width (HPBW) and directional gain (D_G) values are also provided in Table 5-1 and Table 5-2. The SRP using the FAS was

calculated across the entire harmonic bandwidth $[f_o, Rf_o]$ to include all signal components, while the SRP for the HSB was only evaluated at the fundamental frequency. From a comparison of the results displayed, it is evident that the HSB offers greater directionality and better overall performance compared to the standard FAS method. HPBW values for the HSB were considerably less for all harmonics and spacing combinations, while directivity gains were also considerably greater. It is also evident that the effect of spatial aliasing is greatly amplified for the HSB as predicted by equation (5.38) given above. Based on these results, one can conclude that proposed beamformer would be most effective for applications in which low frequency signals are prominent and/or array spacing is limited. The performance of the method will be demonstrated using experimental data in the next chapter.

Table 5-1: Half Power Beam Width (3 dB) in degrees for various sensor spacing values (in meters).

	Standard SRP					**HSB SRP**				
R	$d = 0.1$	$d = 0.2$	$d = 0.3$	$d = 0.4$	$d = 0.5$	$d = 0.1$	$d = 0.2$	$d = 0.3$	$d = 0.4$	$d = 0.5$
1	N/A	N/A	N/A	154	102	N/A	N/A	N/A	154	102
2	N/A	N/A	122	82	62	N/A	80	51	38	30
3	N/A	160	82	58	36	82	38	25	19	15
4	N/A	104	62	46	26	46	21	15	11	5

Table 5-2: Directivity Gain (D_G) in dB for various sensor spacing values.

	Standard SRP					**HSB SRP**				
R	$d = 0.1$	$d = 0.2$	$d = 0.3$	$d = 0.4$	$d = 0.5$	$d = 0.1$	$d = 0.2$	$d = 0.3$	$d = 0.4$	$d = 0.5$
1	1.0	1.1	1.2	1.4	1.6	1.0	1.1	1.2	1.4	1.6
2	1.1	1.2	1.5	1.9	2.3	1.2	2.0	3.3	4.2	5.0
3	1.1	1.4	1.8	2.3	2.8	2.0	4.2	6.3	7.2	5.6
4	1.2	1.6	2.1	2.7	3.2	3.7	6.6	5.6	2.8	4.7

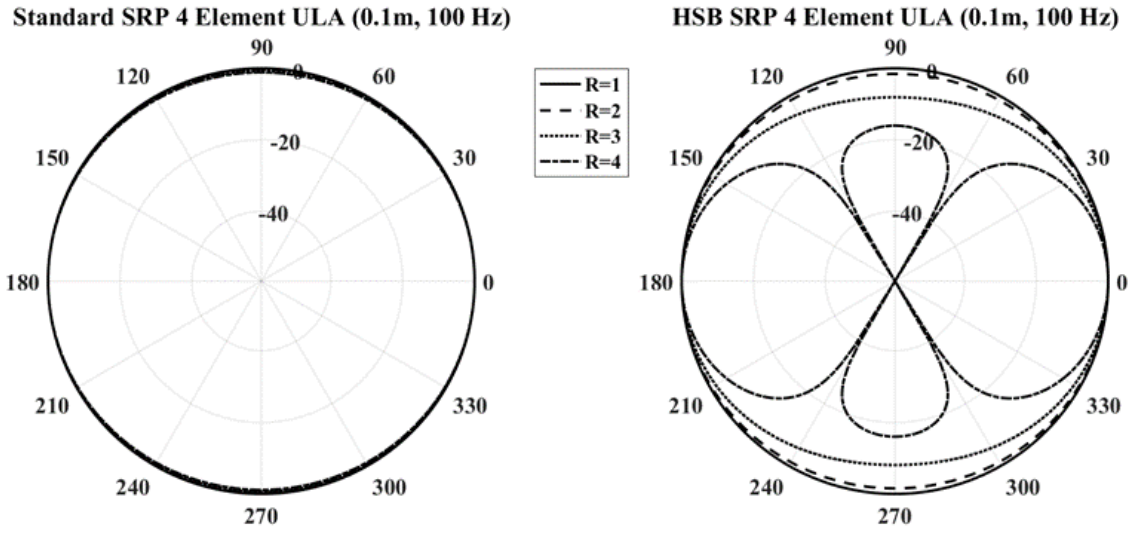

Figure 5-8: Directivity response of a 0.1 m spaced ULA acquiring a 100 Hz signal with *R* harmonic components.

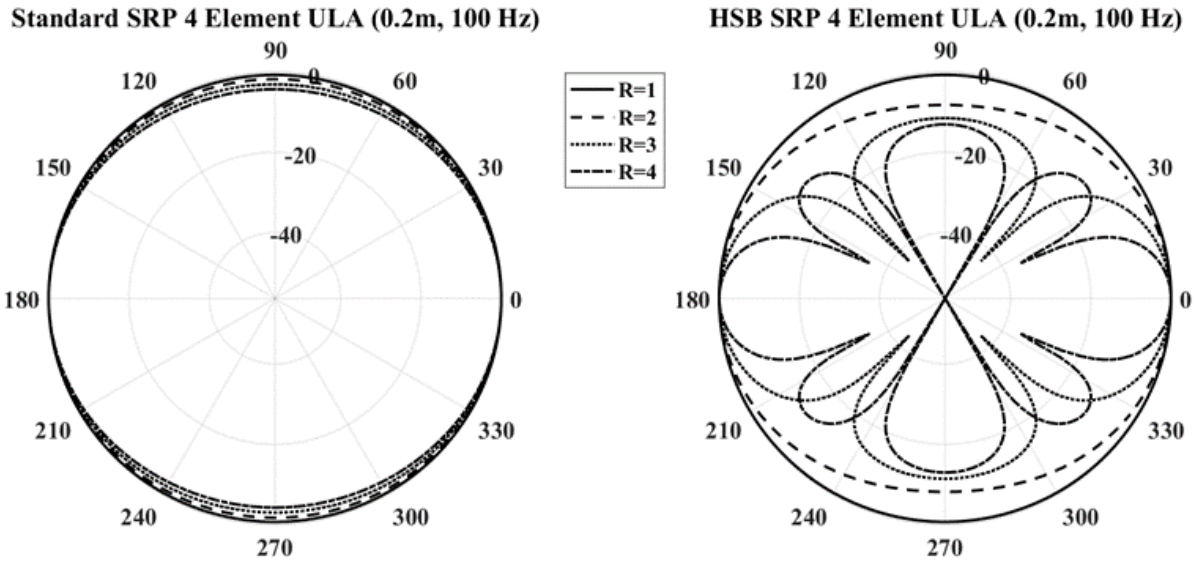

Figure 5-9: Directivity response of a 0.2 m spaced ULA acquiring a 100 Hz signal with *R* harmonic components.

- 6 - Experimental Studies

The following chapter provides the results of experiments conducted to investigate the viability of utilizing acoustic sensing to form the basis of a non-cooperative aircraft anti-collision system. In addition, the signal processing methods previously proposed in the areas of notch filtering, spectral whitening, signal enhancement, signal detection, and source localization will also be demonstrated. A description of the various aircraft and array configurations is provided, along with all relevant equipment, calibration procedures, and physical noise mitigation steps taken. Experiments are presented in the chronological order in which they were conducted, with each study increasing in technical difficulty.

Tests were conducted using both fixed-wing and multirotor aircraft using stationary ground-based and moving airborne sources. Table 6-1 displayed below provides a brief overview of the conducted experiments which are labeled according to the test set number (TS#). The basic parameters and purpose of each experiment is also given along with the various signal processing methods demonstrated.

Table 6-1: Summary of conducted experiments.

Number & Type	TS#1 - Air-to-Ground
Purpose	Preliminary study to establish the basic viability of acoustic sensing.
Operation, Detecting Aircraft, & Target Source	Detection / Fixed-wing Delta X-8 / Yorkville NX 550 Speaker
Signal Processing	SIMO IIR Notch Filtering CFAR-Enhanced Spectral Whitening Selective-cell Constrained Distribution Free CFAR Detector (SCDF-CFAR) Robust Binary Integration Detection Scheme
Number & Type	TS#2 - Air-to-Air
Purpose	Determine whether a fixed-wing UAV can be detected and localized with sufficient distance to perform an avoidance maneuver.
Operation, Detecting Aircraft, & Target Source	Detection & Localization / Fixed-wing Delta X-8 / Giant Big Stik UAV
Signal Processing	Harmonic Spectral Transforms SCDF-CFAR Detection for Harmonic Spectral Transforms Harmonic Spectral Beamformer
Number & Type	TS#3 - Air-to-Air
Purpose	Determine whether manned aircraft can be detected and localized with sufficient distance to perform an avoidance maneuver.
Operation, Detecting Aircraft, & Target Source	Detection & Localization / Kraken Octocopter / Cessna 185
Signal Processing	MIMO IIR Notch Filtering Harmonic Spectral Transforms Phase Acceleration Processors Modified Coherence Processors Harmonic Spectral Beamformer

6.1 - Equipment

6.1.1 - Data Acquisition

Two different types of capacitance microphone were used for the conducted experiments. These consisted of DPA 4053 omni-directional and RØDE M5 cardioid response microphones as displayed below in Figure 6-1. The DPA microphones were fitted with metallic laminar-flow noise cones designed to reduce self-generated noise associated with high airflow applications. These sensors were thus used for all experiments involving fixed-wing UAVs since high flow rates are inherently present. Ideally these microphones would have a directional rather than omni-directional response since this property would greatly reduce the acquired level of propeller generated self-noise. Unfortunately however, there are no high-flow directional response microphones available on the market. The RØDE M5 microphones do produce a directional response but do not offer any flow generated noise protection other than a standard foam windscreen which performs very poorly in high flow environments. Since multirotor aircraft generally operate at relatively low speeds and generate higher noise levels (from the presence of multiple lifting fans), the RØDE microphones were found to be better suited for this system. Table 6-2 displayed below provides the specifications for each microphone, while Figure 6-2 provides polar response plots. One may question why MEMS microphones are not utilized instead since they are extremely small and light weight. The simple reasoning is that these microphones have a much lower dynamic range (≈ 70 dB) and sensitivity (≈ 5 mV/Pa) and perform very poorly in low frequency regions; areas which often constitute the fundamental frequency of an aircraft [276].

Figure 6-1: DPA 4053 with and without nose cone and RØDE M5 microphone pair

Table 6-2: Microphone specifications

	DPA 4053	RØDE M5
Operating Principle	Pressure	Pressure gradient
Transducer	Prepolarized Condenser	Prepolarized Condenser
Polar Pattern	Omnidirectional	Cardioid
Diaphragm	16 mm (0.6 in)	12.70 mm (0.5 in)
Frequency Response	20 Hz - 20 kHz	20 Hz to 20 kHz
Dynamic Range	116 dB SPL	121 dB SPL
Sensitivity (ref 1 V/Pa)	30 mV/Pa, -30 dB	20 mV/ Pa, -34 dB
Maximum Input Sound Level	135 dB	140 dB
Power Requirements	48V Phantom Power	24V or 48V Phantom
Dimensions	59 x 16 mm	100 x 20 mm
Weight	22 g	80 g

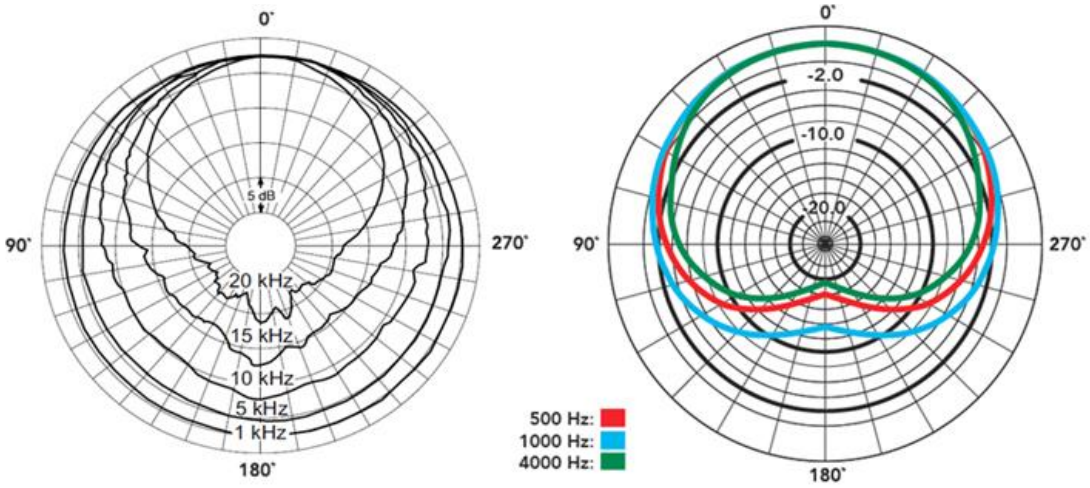

Figure 6-2: Polar response for DPA 4053 (left) RØDE M5 (right).

The recording of acoustic data was achieved using the Zoom H4 and Zoom H6 handheld recording units as displayed below in Figure 6-3. Each recording device is capable of simultaneously recording 4 and 6 channels respectively and can operate via batteries or DC power line. The H4 records at 48 kHz with a 16-bit resolution (approximately 98 dB dynamic range), while the H6 can operate at 96 kHz with a 24-bit resolution (approximately 146 dB dynamic range). The recorded data is stored in the form of .WAV files and may be processed offline after completion of the respective experiment.

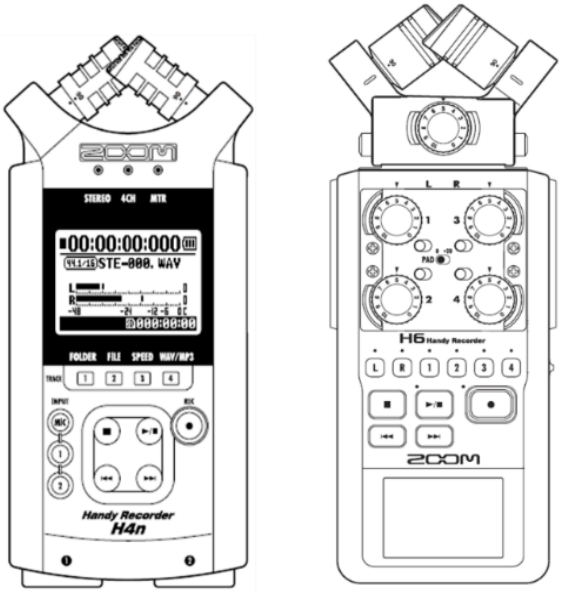

Figure 6-3: Zoom H4 (left) and H6 (right) recoding units

Flight data such as speed, GPS position, altitude, and orientation were recorded using the ArduPilot 2.5 open source autopilot system. The system can control and/or logging data for both fixed-wing and multirotor aircraft with various propulsion configurations. Data may be recorded and stored onboard, and also transmitted in real-time to a ground control station via a 2.4 GHz wireless telemetry link, with rates typically on the order of 57000 bps. Figure 6-4 displayed below provides a picture of the device while Table 6-3 lists the approximate sensor error values.

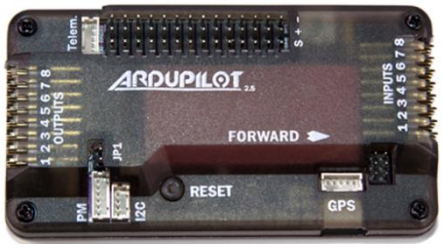

Figure 6-4: ArduPilot 2.5 autopilot system.

Table 6-3: ArduPilot sensor accuracy values.

Data Type	Sensor Accuracy
GPS Position	± 1.5 m
Compass Orientation (3D)	± 2.5 °
Altitude	± 5 m
Airspeed	± 1 m/s

6.1.2 - Sensing Platforms

A total of two different aircraft configurations were used for the experiments. These consisted of a fixed-wing pusher style (Delta Wing X-8) and a pusher-configured multirotor (Kraken). The Delta Wing X-8 is a hobby-grade aircraft powered by a single brushless DC motor. Four DPA 4053 omni-directional microphones were fitted to the aircraft via carbon fiber booms extending from the noise as depicted below in Figure 6-5. The microphones were secured using custom vibration isolation mounts constructed from Sorbothane material. Since the array is distributed in one dimension (linear form), it is only capable of localizing targets in one angular direction (azimuth). Figure 6-6 displays the geometry and directional response of the array for a range of signal frequencies. It is apparent from the plot that the array offers poor directional performance for the

frequency range of concern. This was unfortunately due to geometric and structural limitations of the aircraft which prevented the establishment of a larger microphone spacing. However, this problem can be alleviated for harmonic signals through use of the Harmonic Spectral Beamformer (HSB) previously presented in Chapter 5. Figure 6-7 provides response plots of equal scale for the standard and HSB Steered Response Power (SRP) for a 100 Hz signal with varying harmonics. It is apparent that the HSB form produces a much more desirable directivity response which has a positive correlation to the number of signal harmonics. The use of the method to increase localization accuracy will be demonstrated in the upcoming experimental results section.

Figure 6-5: Delta Wing X-8 aircraft with four DPA 4053 microphones.

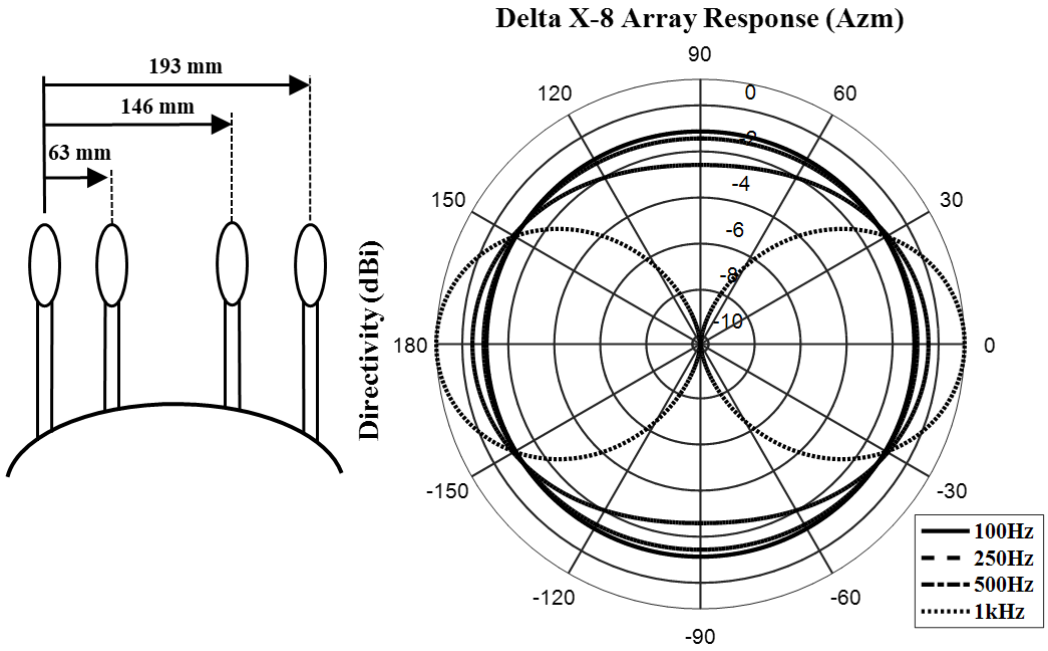

Figure 6-6: Delta X-8 array geometry and directional response.

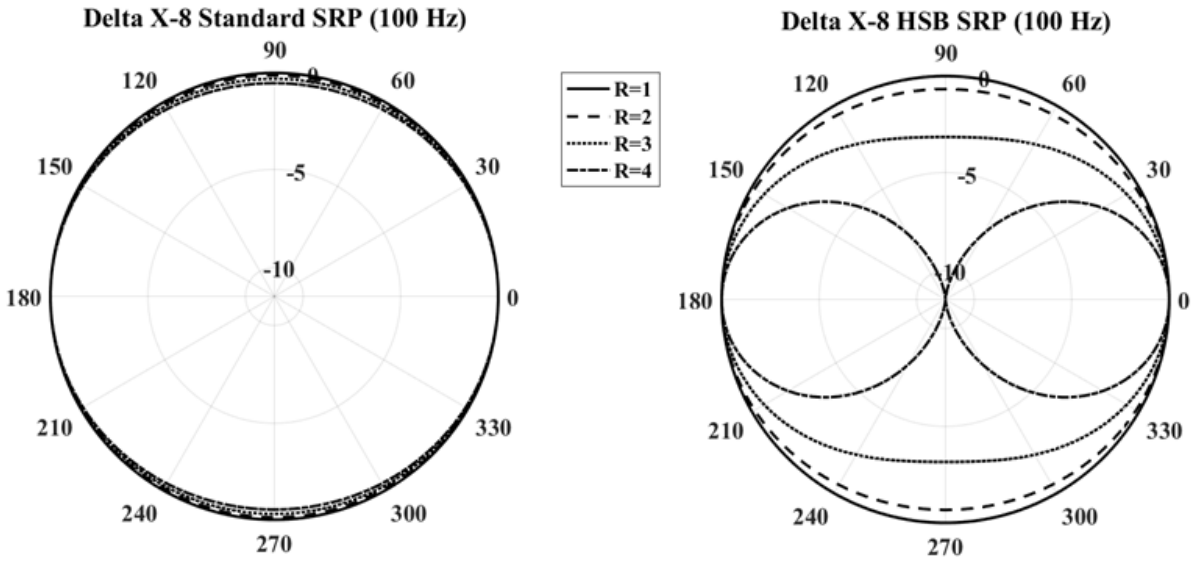

Figure 6-7: Delta X-8 array response for a 100 Hz signal with *R* harmonic components.

The Kraken multirotor is a commercial grade UAV consisting of eight brushless DC motors. The aircraft was originally designed for a conventional lifting-style configuration but was converted to a pusher-style by inverting the motors and reversing their rotation directions. By moving the propellers to the underside of the aircraft, essentially all of the topside area becomes free to mount any required sensors and instrumentation. Six RØDE M5 microphones were fitted to the aircraft via carbon fiber booms extending vertically from engine support frame(s) as depicted in Figure 6-8. Commercially available vibration isolation mounts were utilized to secure the microphones to the booms. Unlike the fixed-wing UAV, vibration induced lateral motion was found to be significant with this aircraft. This was caused by a combination of increased overall vibration levels due to the presence of multiple lifting fans, and the microphone placement location which allows flexural and torsional displacement with respect to the base mounting frame. To reduce these effects, each sensor was connected to its adjacent neighbours and base frame via a tensioned cable. It should be noted that target localization error with respect to sensor positional error is directly dependent on the received source signal wavelength. Lower wavelength signals require larger spacing values to obtain higher array directivity gains and are thus less susceptible to minor sensor displacement errors. More information on this topic may be found in [277]

Since the array is distributed in two dimensions (planar form), it is capable of localizing targets in both the azimuth and elevation angle directions. Figure 6-9 provides the array geometry, while Figure 6-10 provides response plots for the azimuth and elevation directions. From the elevation response plot, it is evident that the array does not provide any significant detection capability for targets located below the aircraft. This is expected since the microphones have a cardioid directional response and are orientated in the vertical direction.

The aircraft was controlled via a DJI NAZA-M under assisted manual operation, while data was logged via an ArduPilot 2.5 system. Because the aircraft has eight radial lifting arms and only six microphones were used,

counterweights were added to the two remaining rotor arms to maintain flight stability. Table 6-4 displayed provides a summary of the properties and sensor configuration for each of the aircraft used.

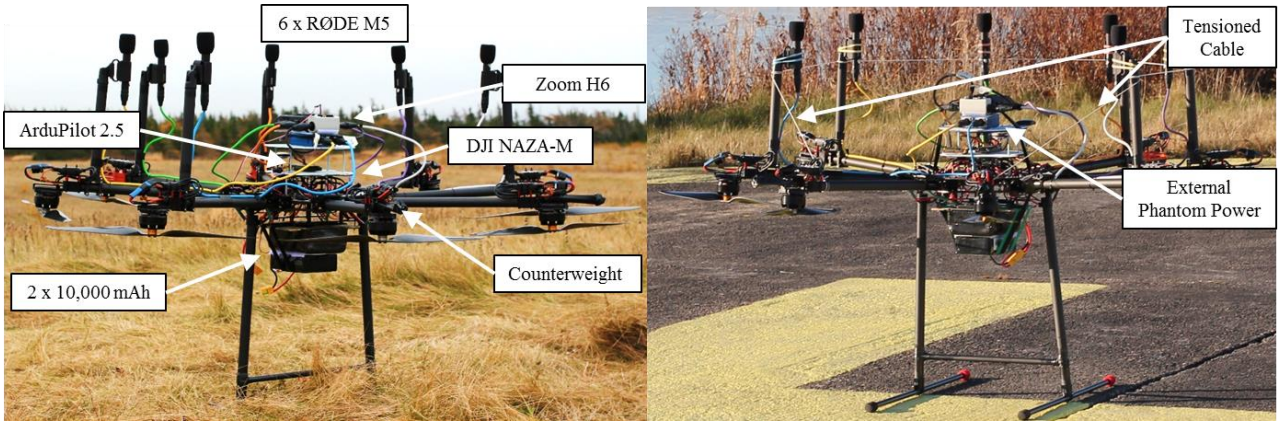

Figure 6-8: Kraken Octocopter equipped with six RØDE M5 microphones and H6 recoding unit.

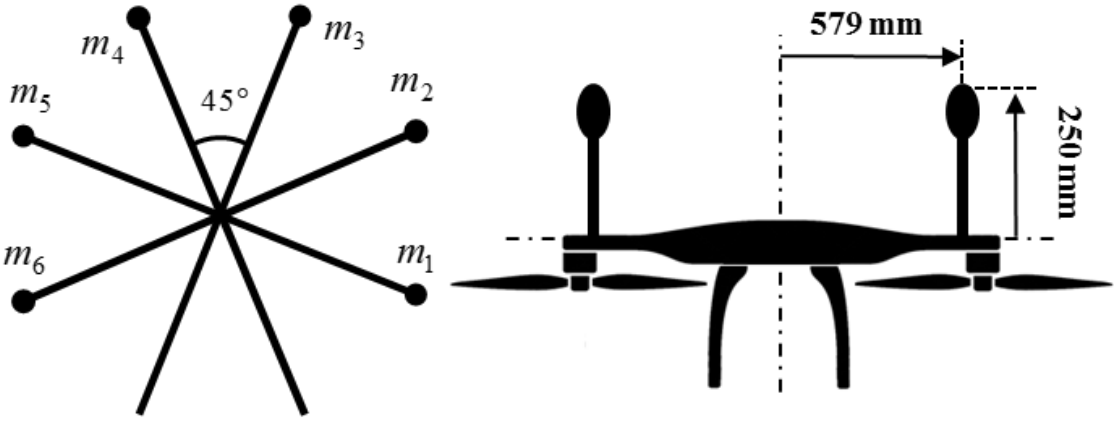

Figure 6-9: Kraken array geometry.

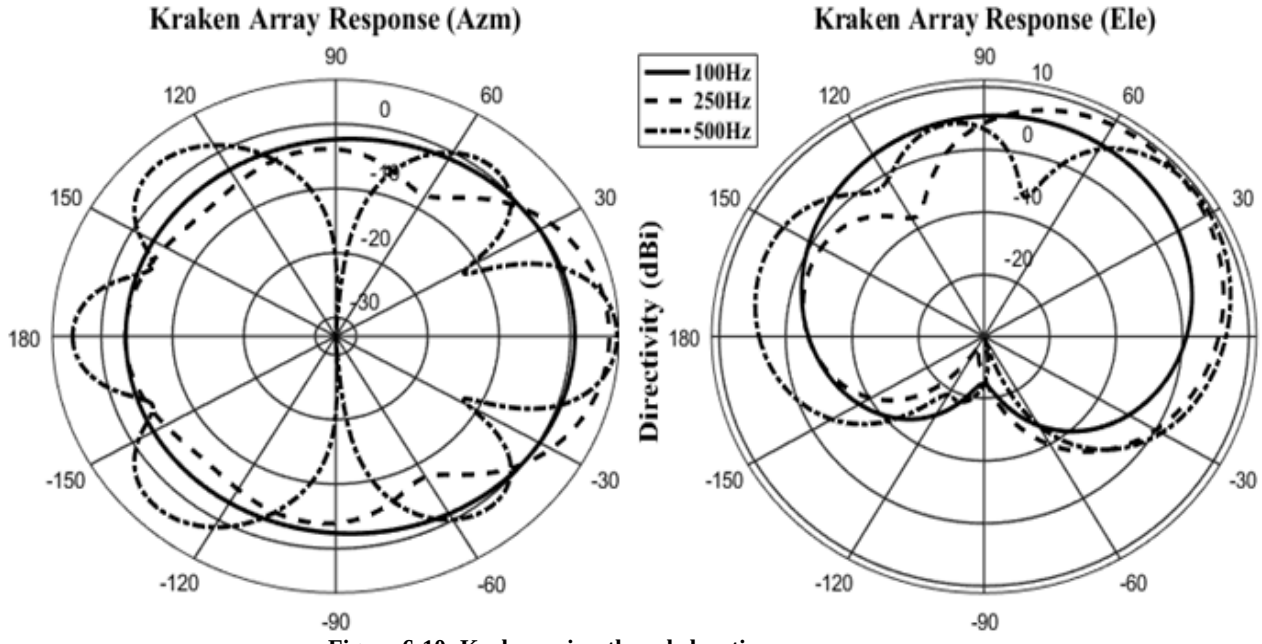

Figure 6-10: Kraken azimuth and elevation array response.

Table 6-4: Detection aircraft specifications.

	Delta X-8	Kraken
Classification	Fixed-wing Pusher	Multirotor Pusher
Cruising Speed (knots)	20-30	10-20
Max Payload Capacity (kg)	≈ 3	≈ 15
Sensor Payload (kg)	1.90	2.05
Mass (kg)	2.2	10.1
Wingspan (m)	2.12	1.79
Length (m)	0.82	1.79
Sensor Configuration	4 x DPA 4053	6 x RØDE M5
Recording Unit	Zoom H4	Zoom H6
Sampling Rate	48 kHz	96 kHz
Resolution	16 Bit	24 Bit
Flight Control	ArduPilot 2.5	DJI Naza-M
Data Logging	ArduPilot 2.5	ArduPilot 2.5

6.1.3 - Acoustic Sources

Experiments were conducted using a number of ground-based and airborne acoustic sources. The ground-based source consisted of a 500-watt Yorkville NX 550 loudspeaker, which was configured to emit a variety of continuous sinusoidal signals. Airborne sources included a gasoline powered Giant Big Stik UAV and a manned Cessna 185 airplane. Tables 6-5 and 6-6 displayed below provide the acoustic source properties and target aircraft specifications respectively, while Figure 6-11 provide images for each of the sources.

Table 6-5: Acoustic properties of sources used [278]*.

Source	SPL (dBC @ 1m)	Operating Frequency (Hz)
Yorkville NX 550 Speaker	120-125	50-20,000
Giant Big Stik	110-115	60-70
Cessna 185	125 - 130*	35-45

Table 6-6: Target aircraft specifications.

	Giant Big Stik	Cessna 185
Cruising Speed (knots)	25-35	120-140
Propeller Blades (#)	2	2
Drive Shaft RPM	3500 - 4500	2100 - 2500
Fundamental Frequency (Hz)	120-140	70-85

Figure 6-11: Acoustic source targets utilized for experimental studies.

6.1.4 - Extrapolated Detection Distances

Due to physical and regulatory constraints, acoustic detection experiments involving fixed-wing aircraft could not be physically conducted to determine the maximum detection range achievable. Instead, using simple signal analysis in conjunction with acoustic propagation laws, values for the maximum detection distance that could have been achieved for each experiment were approximated.

Determining the maximum extrapolated detection distance is a straightforward process that only requires knowledge of the closest point of approach (CPA) of the aircraft to the sound source. That is, the true SPL of the sound source is not required to determine the maximum distance at which the source should be detected. It only requires the change in power (dB) for the signal component of interest be directly proportional to the reduction in source SPL with respect to distance. Experiments were conducted where the previously described recoding setup was exposed to various acoustic sound pressure levels. The recorded signals were then transformed to the frequency domain and the power calculated in decibel units. A plot of the digital signal power with respect to the physical acoustic pressure level does in fact display a linear relationship as shown below in Figure 6-12.

The maximum expected detection distance may therefore be calculated through application of the acoustic power attenuation law previously presented by equation (2.2):

$$d_{max} = d_{ref} 10^{\frac{P_{ref} - P_{min} - A_{abs}}{20}} \tag{6.1}$$

where d_{max} is the maximum expected detection distance in meters, d_{ref} is the distance between the aircraft and acoustic source at the CPA, P_{ref} is the power of the dominant source frequency component at the CPA given in dB, P_{min} is the detection threshold for the frequency band of interest (the point at which the signal component of interest is no longer distinguishable from the surrounding noise), and A_{abs} is the atmospheric absorption factor previously given by equation (2.19). Figure 6-13 displayed below illustrates the concept. The model is of course a very simplified approximation in that losses due to environmental factors such and wind and temperature gradients are not taken into account. Nevertheless, it still provides a useful measure since it gives the maximum upper limit under ideal transmission conditions.

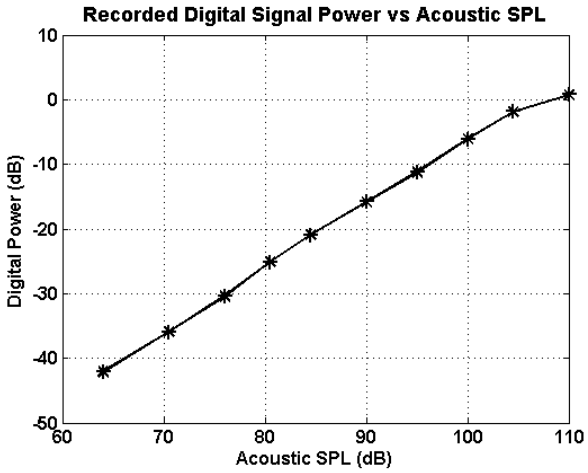

Figure 6-12: Comparison of recorded digital signal power and acoustic sound pressure levels (SPL)

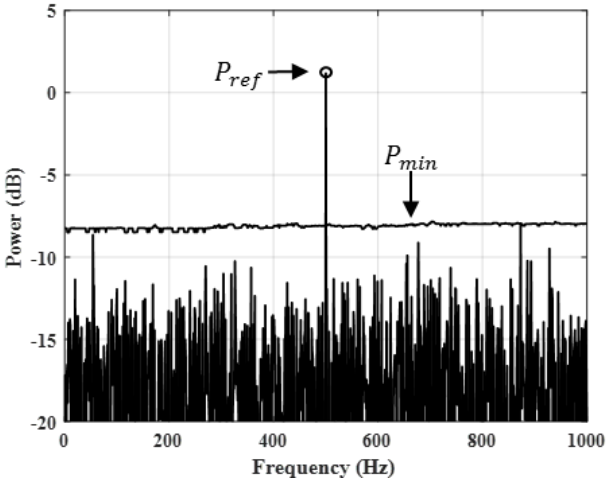

Figure 6-13: Sample spectra illustrating detection signal power range.

6.2 - Signal Processing

Figure 6-14 displayed below provides a flow diagram for overall signal processing methods used. Data for all presented experiments were processed using this general approach. Recorded signals were first decimated to reduce data processing requirements since sampling rates were on the order of 48 to 96 kHz, but only frequency information to up approximately 1000 Hz was found to be useful. The sample-reduced signals were then notch filtered to remove narrowband self-noise components using the relevant IIR notch filter type (SIMO, MIMO, etc.) as previously described in Section 3.3. The IIR filter form was used instead of the FIR or Comb filters since this method was found to perform best as previously discussed in Section 3.7. The filtered signals were then windowed, frequency transformed using the FFT, and spectrally whitened via the CFAR method proposed in Section 4.3.5. As previously discussed, this procedure was performed to transform the broadband noise components to an equivalent distribution type to facilitate use of the DF-CFAR detector. Prior to performing the detection, signals were first enhanced using the various processors described in Section 4.2.2. Finally, if a target source signal was found to be present, beamforming methods were then applied to the whitened signals

to estimate the angular source location. Details regarding the exact parameters used at each processing stage are provided in the relevant test procedure sections provided below.

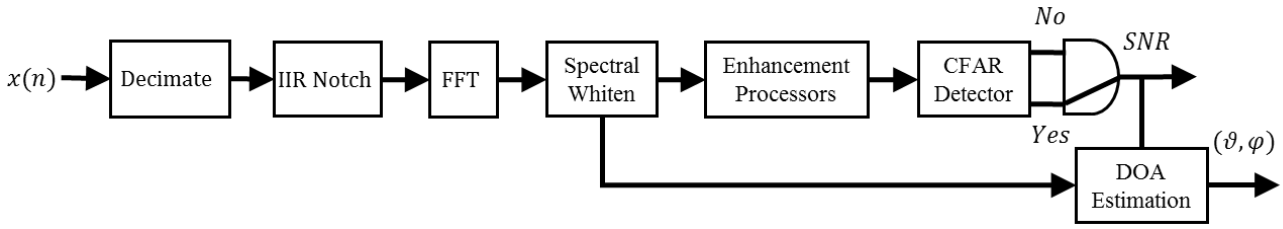

Figure 6-14: General signal processing flow diagram.

6.3 - TS#1: Fixed-wing Air-to-Ground

6.3.1 - Purpose & Procedure

The first experiment was a preliminary study to establish the basic viability of acoustic sensing. The overall goal was to determine whether a continuous pure-tone acoustic source could be detected at a relatively close proximity. The results obtained from this experiment were previously published in the Journal of Unmanned Vehicle Systems [56]. Thus, details regarding the experiment and results obtained will be only briefly discussed for the purpose of illustrating the proposed CFAR-Enhanced Spectral Whitening and SCDF-CFAR detector presented in Sections 4.3.4 and 4.3.5.

The tests involved flying the Delta X-8 UAV at various altitudes above a ground-based loudspeaker emitting various narrowband sinusoidal signals. For two of the test sets, an audio recording of a gasoline powered Giant Big Stik (GBS) UAV and a multi-frequency pure tone combination were used. Figure 6-15 provides a depiction of the experimental setup, while Table 6-7 provides the various parameters such as source frequency and passing altitudes used for the experiment. The closest point of approach (CPA) is the altitude when the aircraft is directly overhead the loudspeaker.

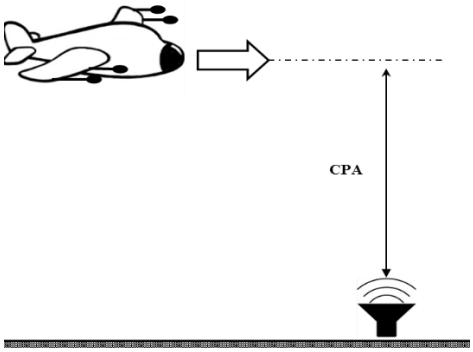

Figure 6-15: TS#1 - Depiction of experimental setup.

Table 6-7: TS#1 - Experimental parameters.

Frequency (Hz)	SPL (dB)	CPA Altitude (m)
200	127	50, 100, 150
500	122	50, 100, 150
GBS	119	50, 100, 150
$\sum_{n=1}^{10} 100n$	121	50, 100, 150

6.3.2 - Results & Discussion

6.3.2.1 - *Signal Processing*

As previously mentioned, the results of this experimental study were formerly published in a peer reviewed journal. However, neither spectral whitening nor CFAR detection were utilized for the processing of acquired acoustic signals. Instead, evaluation points were chosen manually based on a visual inspection of the signal spectrograms. Such a spectrogram is displayed in Figure 6-16 which illustrates the effectiveness of the IIR notch filtering process for a segment containing the 200 Hz source signal. Thus, results obtained via the use of these methods will now be discussed.

Tables 6-8 to 6-10 provides the notch filter, FFT, spectral whitening, and SCDF-CFAR parameters used. As previously discussed in Section 4.3.5, the proposed whitening procedure can be effectively employed without reducing the probability of detection by using a threshold value which produces a much higher false alarm rate than that used in the final detection stage. Here, thresholding values are chosen using the values displayed below such that a false alarm rate of $P_{FA} = 0.1$ is achieved. This is considerably higher than that offered by the SCDF-CFAR detector as indicated in the table below by P_{FA}^{SC}. Note that SC indicates testing a Single Cell, while ST indicates testing all B cells contained in $\vec{N}$ (Single Trial). BI and RBI indicates the use of Binary Integration and Robust Binary Integration across the T trials respectively. The proposed SCDF-CFAR detector was used instead of the CDF-CFAR form since it offers decreased computational requirements without affecting detection performance as previously discussed in Section 4.3.4.

Unfortunately, one of the recorded signals was corrupted due to a loose microphone diaphragm rendering it useless. This is evident from the spectrogram displayed in Figure 6-17 since the 200 Hz source is no longer visible. Thus, signal processing was conducted using only three of the four recorded signals.

Table 6-8: TS#1 - Signal preprocessing and filter parameters.

Sampling Frequency (f_s)	48 kHz	Number of Signals	4
Decimation Factor	8	FFT Window	0.5 s
IIR Step Size (μ)	5 E -4	Window Overlap	50%
Notch Radius (r)	0.995	Padded Length (L_{fft})	12,000 pts
Harmonics Removed (R)	8	Spectral Resolution (f_r)	0.5 Hz/bin

Table 6-9: TS#1 - CFAR-Enhanced Spectral Whitening parameters.

Detector Type	OS-CFAR	Noise Samples (N)	101
Forgetting Factor (ξ)	0.2	Order Statistic (k)	0.75 N
Flooring Factor (δ)	0.5	Guard Cell Band ($\vec{G}$)	5.5 Hz
Noise Band ($\vec{N}$)	50 Hz	Guard Cells (G)	12

Table 6-10: TS#1 - SCDF-CFAR detection parameters.

Noise Sample Band ($\vec{N}$)	1-1000 Hz	Consecutive Detections (D)	2
Test Band ($\vec{B}$)	150-550 Hz	Cell Deviation (Δ)	1
Guard Cell Band ($\vec{G}$)	10.5 Hz	Maxima Tested (M)	2
Noise Samples (N)	1998 pts	P_{FA}^{SC}	1.0E-3
Test Cells (B)	801 pts	P_{FA}^{ST}	6.5E-1
Guard Cells (G)	22 pts	P_{FA}^{BI}	8.2E-4
Order Statistic ($\bar{k}$)	2	P_{FA}^{RBI}	2.5E-3
Consecutive Trials (T)	2		

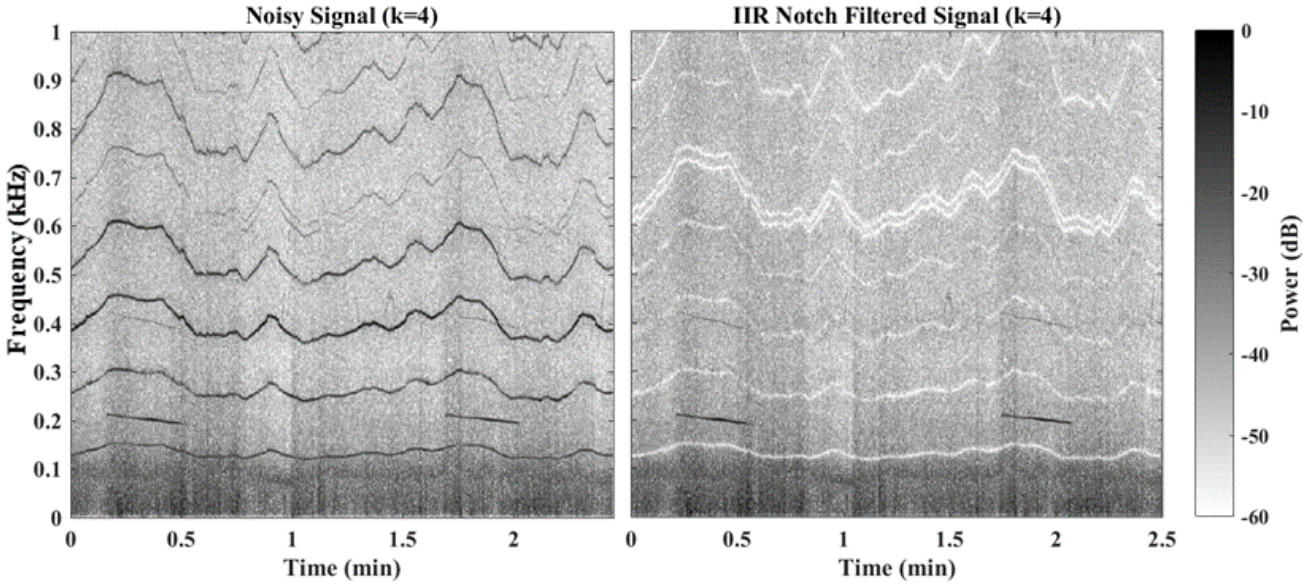

Figure 6-16: TS#1 - Spectrograms for unfiltered and filtered signal segment containing 200 Hz source.

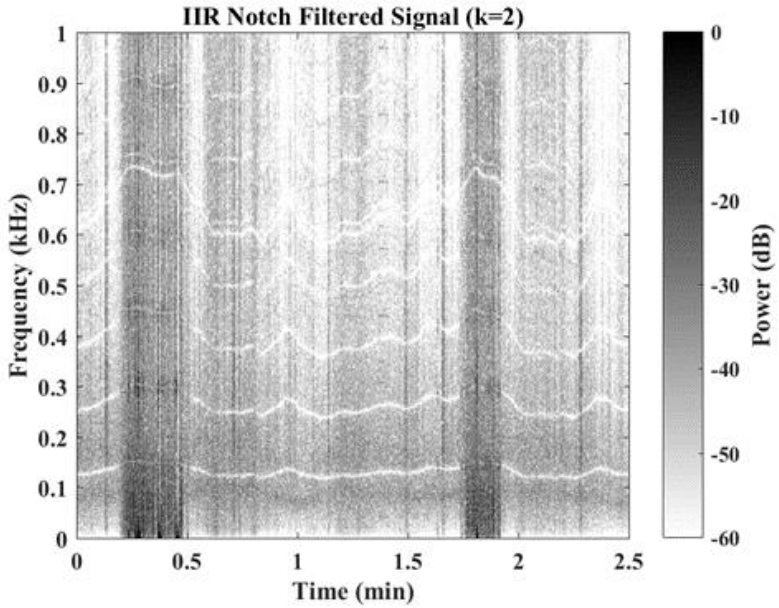

Figure 6-17: TS#1 - Spectrogram of the corrupted signal.

6.3.2.2 - *Source Detection*

Table 6-11 provides the detection results for the 200 and 500 Hz source signals with a passing altitude of 150 m. For each source frequency, results are provided for the both the whitened and unwhitened signals. From a comparison of the results obtained, it is apparent that the whitened signals produce significantly better results compared to the standard unwhitened forms; SNR values and detection rates were generally much higher. In addition to SNR values and overall detection rates, initial detection times were also found to be less for the whitened signals. This parameter is very important since it dictates whether the detecting aircraft will have enough time to perform an avoidance maneuver. It should be noted that the SNR values quoted are not calculated in the manner typical of most signal processing applications as previously given by equation (2.100) Instead, the "Effective SNR" was used which closely resembles the Spurious Free Dynamic Range as previously discussed in Section 2.6.4. This method provides a more meaningful measure since it compares the peak signal value to the point at which the signal can no longer be detected (noise floor or detection threshold), and will also be used when calculating detection results for all further experimental studies. It is depicted in the sample spectra displayed in Figure 6-19.

As previously discussed in Section 4.3.5, the decreased SNR values present for the unwhitened signals is due to an increased detection threshold caused by the non-flat power distribution of broad band noise components. This trend is clearly visible in the spectrogram and power spectrum plots displayed by Figures 6-18 and 6-19 respectively. Higher SNR values could be achieved by simply excluding lower frequency values from the threshold calculation. However, doing so would greatly increase false alarm rates since the number of noise samples N would also decrease proportionately. In general, increasing values of N and decreasing values of B will produce lower false alarm rates.

For each signal type, detection rates were calculated for the Single Trial (ST), Binary Integration (BI), and Robust Binary Integration (RBI). Figure 6-20 provides spectrogram-like plots of the detection locations obtained for the whitened and unwhitened signals. It evident that the ST detection scheme attains the highest rate for both signal types. However as previously indicated in Table 6-10, false alarm rates are much too high to facilitate a practical operating system. Application of the BI scheme was found to reduce this value substantially while causing minimal effect on signal detectability. However, when comparing the results for the 200 and 500 Hz source signals, it is evident that this decrease does become significant as the received source signal becomes more non-stationary. This non-stationarity is essentially caused by Doppler effects associated with relative motion between the source and observer, and is indicated by the detected frequency range given in the results table. For such cases, the proposed RBI scheme offers significantly increased detection rates (14-20%). However, application of the approach will also inherently increase false alarm rates to some degree.

Finally, it should be mentioned that a single harmonic component was present at the 400 Hz frequency location as indicated by the provided sample spectra. Although the emitted source signal contained only a pure 200 Hz tone, the harmonic component was generated by the presence of a reflecting boundary (ground) located directly

behind the speaker. From a visual inspection of the power spectra plots, it is evident that this component will be detected in the whitened signals but not in the unwhitened forms.

Table 6-11: TS#1 - Detection results (150 m CPA).

	f_o = 200 Hz		f_o = 500 Hz	
	Unwhitened	Whitened	Unwhitened	Whitened
Detection Rate (ST, BI, RBI)	64, 54, 55 %	100, 97, 99 %	69, 37, 51 %	100, 63, 83 %
Max SNR	28 dB	38.3 dB	26.5 dB	47.4 dB
Average SNR	12.4 dB	19.7 dB	8.2 dB	32.5 dB
Initial Detection	13.25 s	13 s	11 s	10.5 s
Second Detection	20.25 s	13.25 s	12 s	10.75 s
Observed Frequency Range	212 – 190 Hz		523 – 479 Hz	

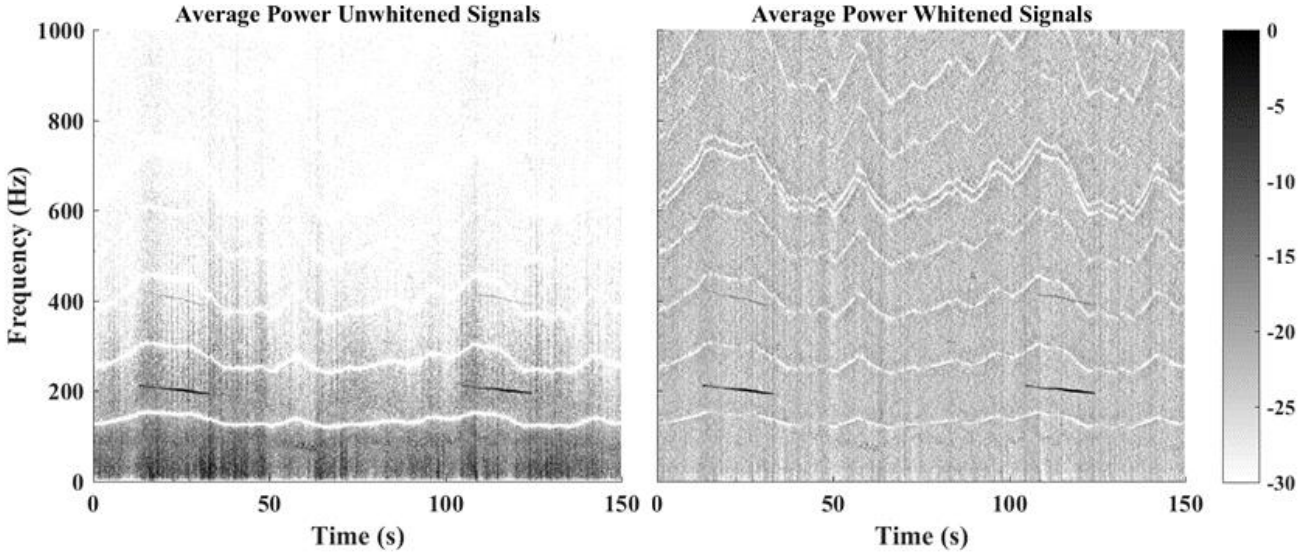

Figure 6-18: TS#1 - Spectrograms of whitened and unwhitened average power signal segments for 200 Hz source.

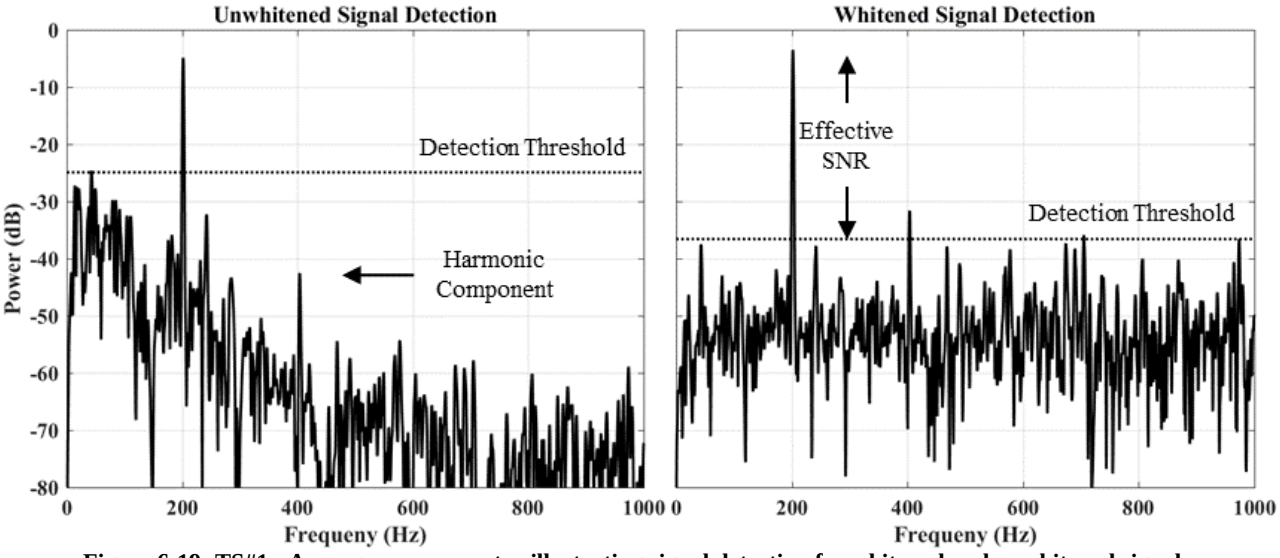

Figure 6-19: TS#1 - Average power spectra illustrating signal detection for whitened and unwhitened signals.

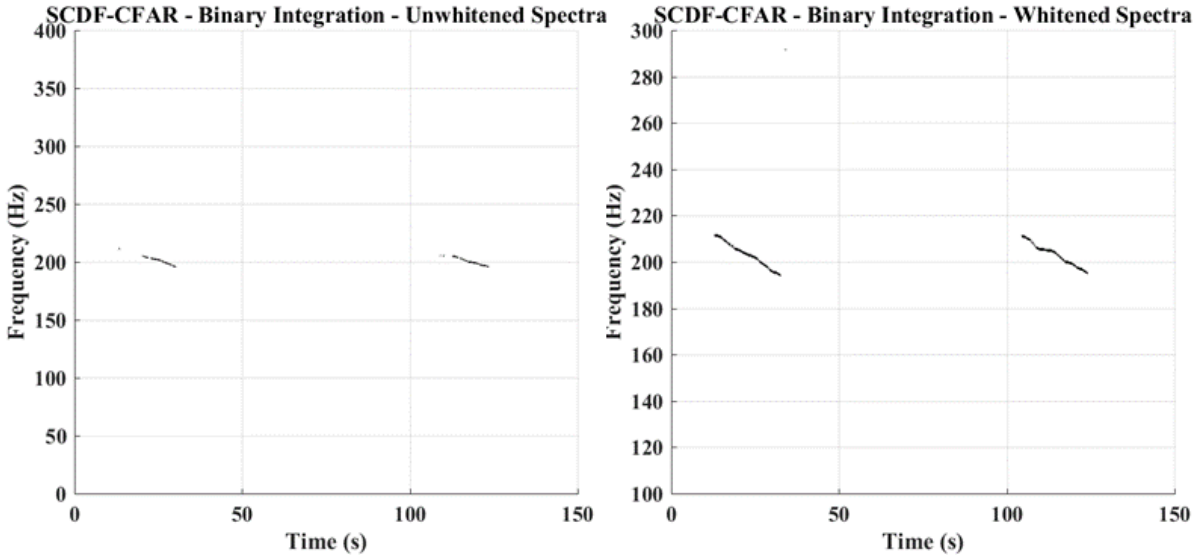

Figure 6-20: TS#1 - Spectrogram-like plots of detection locations for whitened and unwhitened signals.

6.3.3 - Conclusions

Based on the results obtained from the analysis provided, it is evident that the proposed CFAR-Enhanced Spectral Whitening method is effective at increasing the overall performance of the SCDF-CFAR detector. Results obtained for the whitened signals were significantly better in all areas evaluated. Thus, the proposed method will be employed for all further processing of experimental data using the parameter values previously listed. In addition, it was also found that the Robust Binary Integration scheme offered superior detection performance for increasingly non-stationary signals. However, the increased detectability also comes at the expense of increased false alarm rates. It is therefore concluded that the scheme should be employed for all non-stationary signals provided false alarm requirements are still met.

6.4 - TS#2: Fixed-wing Air-to-Air

6.4.1 - Purpose & Procedure

The second experiment was conducted to determine whether acoustic sensing could be used to detect and localize another moving aircraft with sufficient distance to perform an avoidance maneuver. In terms of signal processing, use of the proposed Harmonic Spectral Transforms to enhance signal detection and the Harmonic Spectral Beamformer to enhance localization will both be demonstrated.

The experiment involved flying two aircraft (sensing and intruding) in circuit formation with opposing flight paths to facilitate close mid-air encounters. The aircraft were assigned different altitudes and circuit radii to avoid any actual mid-air collisions from taking place. The intruder was assigned an altitude of 150 m with a circuit radius of approximately 500 m, while the detecting aircraft was given an altitude and circuit radius of 100 m and 200 m respectively. The sensing aircraft consisted of the Delta X-8 fitted with 4 DPA microphones, while the intruding aircraft consisted of the gasoline powered Giant Big Stik. Both aircraft were fitted with ArduPilot 2.5 systems for flight control and data logging. Figures 6-21 and 6-22 provide a depiction of the

experimental setup and a Google Earth image of the aircraft GPS tracks respectively. The total duration of the experiment was 830 s and consisted of 25 close encounters.

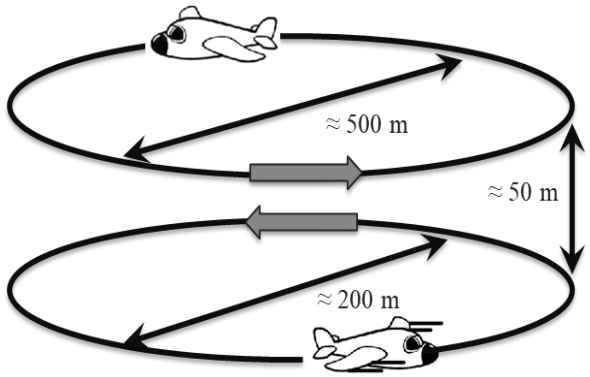

Figure 6-21: TS#2 - Depiction of aircraft flight paths utilized for the experiment.

Figure 6-22: TS#2 - Google Earth image of GPS tracks.

6.4.2 - Results & Discussion

6.4.2.1 - *Signal Processing*

The experiment was originally conducted using a total of four microphones. However, a malfunctioning external power supply corrupted two of the channels with excess noise rendering them useless. Thus, data processing was performed using only two of the recorded signals. Since both aircraft were in relatively close proximity throughout the experiment, observed source frequencies were highly non-stationary. This coupled with the fact that only two channels were available for processing, excluded the effective use of the phase-based signal enhancement processors previously presented in Section 4.2.2. The high degree of non-stationarity can be observed from Figure 6-23 which provides plots of the approximate observed Doppler frequency, corresponding phase acceleration, relative velocity, and relative acceleration. The relative velocity and acceleration were calculated directly from the GPS positional data for the two aircraft. The observed Doppler shifted frequency was approximated using equation (2.25) in conjunction with the GPS positional data, while assuming a fundamental source frequency of 135 Hz. This value was determined from a visual inspection of the recorded signal frequency spectra. The value is only an approximation since in reality the source aircraft did not maintain a constant engine speed throughout the experiment, and its actual engine speed is unknown. From the plots, it is evident that the use of phase acceleration or coherence-based enhancement processors which require a relatively stationary signal for an upwards of 4 to 6 windowed segments would not perform well. As will later be shown in upcoming experiments, those processors are best suited for long-range detection applications where the observed source frequency would change slowly over time. Instead, the signals are enhanced using the HSTs previously presented. The effectiveness of the technique is verified by comparing the results to that obtained via the standard incoherent mean as previously defined by equation (4.176).

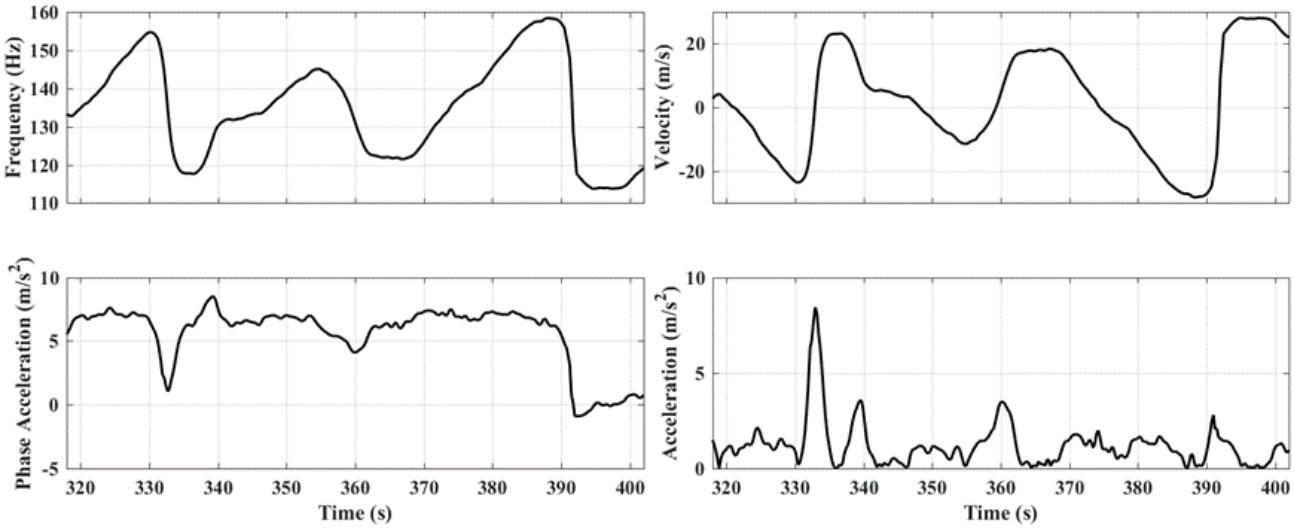

Figure 6-23: TS#2 - Plots of approximate source frequency, phase acceleration, relative velocity, and relative acceleration.

Tables 6-12 to 6-14 provide the notch filter, FFT, spectral whitening, and SCDF-CFAR parameters used. Figure 6-24 provides sample spectrogram plots, which demonstrate the effectiveness of the adaptive IIR notch filtering approach. As with the former experiment, signal detection was performed using the SCDF-CFAR detector since this method offers a decreased computational load as previously discussed. Two sets of detection parameters are provided in the table since two enhancement processors were used, which evidently have different spectral properties with regards to possible detector configurations. These include the incoherent mean (average power $\bar{X}$) and standard mean HST $\bar{\mathrm{H}}_1[X]$. A total of six ($R = 6$) harmonics were used for the HST which consisted of the summation form ($\bar{\mathrm{H}}_1$), since this transform was previously found to provide the best signal detection capabilities (see Section 4.4.1). Because the recorded signals were decimated to a final sampling rate of 4800 Hz and a total of 6 harmonics were used for the HST, the maximum noise sample band only ranged from 0 to 400 Hz for this processor. In contrast, the incoherent mean was free to utilize noise samples across the entire signal bandwidth (0 to 2400 Hz). Noise sample sizes were chosen however, such that each processor produced an equivalent false alarm rate as indicated in the table. In addition to the noise sample size, frequency test bands were also different for the two processed signal forms. The HST test band contained only the expected range of the fundamental source frequency (100 to 200 Hz). However, a visual analysis of the recorded signal spectra for the incoherent mean revealed that the second harmonic component was of much greater amplitude and thus better suited for detection. This effectively doubled the expected source frequency range (200 to 400 Hz).

Presented false alarm probabilities were calculated using the base functional form previously given by equation (4.149) since harmonically transformed signals were used. This form was also used for the incoherent mean to maintain consistency, and because the fractional peak properties of the detection method would automatically exclude surrounding harmonic components from being included in the noise sample estimate. Figure 6-25 provides sample spectrogram plots of the average power and harmonically transformed signals. It is evident from the HST plot that a large number of fractional peaks are in fact present. In theory, a total of 17 fractional peaks may be present as previously indicated by Table 4-5. In addition, it is clearly evident that the second

harmonic component has the highest peak prominence for the incoherent mean form. This is also visible from the sample detection spectra displayed below in Figure 6-26.

Table 6-12: TS#2 - Signal preprocessing and filter parameters.

Sampling Frequency (f_s)	48 kHz	Number of Signals	2
Decimation Factor	10	FFT Window	0.5 s
IIR Step Size (μ)	5 E -4	Window Overlap	50%
Notch Radius (r)	0.995	Padded Length (L_{fft})	6000 pts
Harmonics Removed (R)	8	Spectral Resolution (f_r)	0.5 Hz/bin

Table 6-13: TS#2 - CFAR-Enhanced Spectral Whitening parameters.

Detector Type	OS-CFAR	Noise Samples (N)	101
Forgetting Factor (ξ)	0.2	Order Statistic (k)	0.75 N
Flooring Factor (δ)	0.5	Guard Cell Band ($\vec{G}$)	5.5 Hz
Noise Band ($\vec{N}$)	50 Hz	Guard Cells (G)	12

Table 6-14: TS#2 - SCDF-CFAR detection parameters.

	$\bar{\mathbf{H}}_1[X]$	$\bar{X}$		$\bar{\mathbf{H}}_1[X]$	$\bar{X}$
Noise Sample Band ($\vec{N}$)	10-400 Hz	10-492.5 Hz	Order Statistic ($\bar{k}$)	3	3
Test Band ($\vec{B}$)	100-200 Hz	200-400 Hz	Consecutive Trials (T)	3	3
Guard Cell Band ($\vec{G}$)	10 Hz	10 Hz	Consecutive Detections (D)	3	3
Fractional Guard Cell Band ($\vec{G}_F$)	1 Hz	1 Hz	Cell Deviation (Δ)	1	1
Noise Samples (N)	780 pts	965 pts	Maxima Tested (M)	3	3
Test Cells (B)	201 pts	401 pts	P_{FA}^{SC}	4.2E-3	3.4E-3
Guard Cells (G)	22 pts	22 pts	P_{FA}^{ST}	0.63	0.83
Fractional Guard Cells (G_F)	2 pts	2 pts	P_{FA}^{BI}	1.5E-5	1.5E-5
Fractional Peaks (F)	17	17	P_{FA}^{RBI}	1.4E-4	1.4E-4

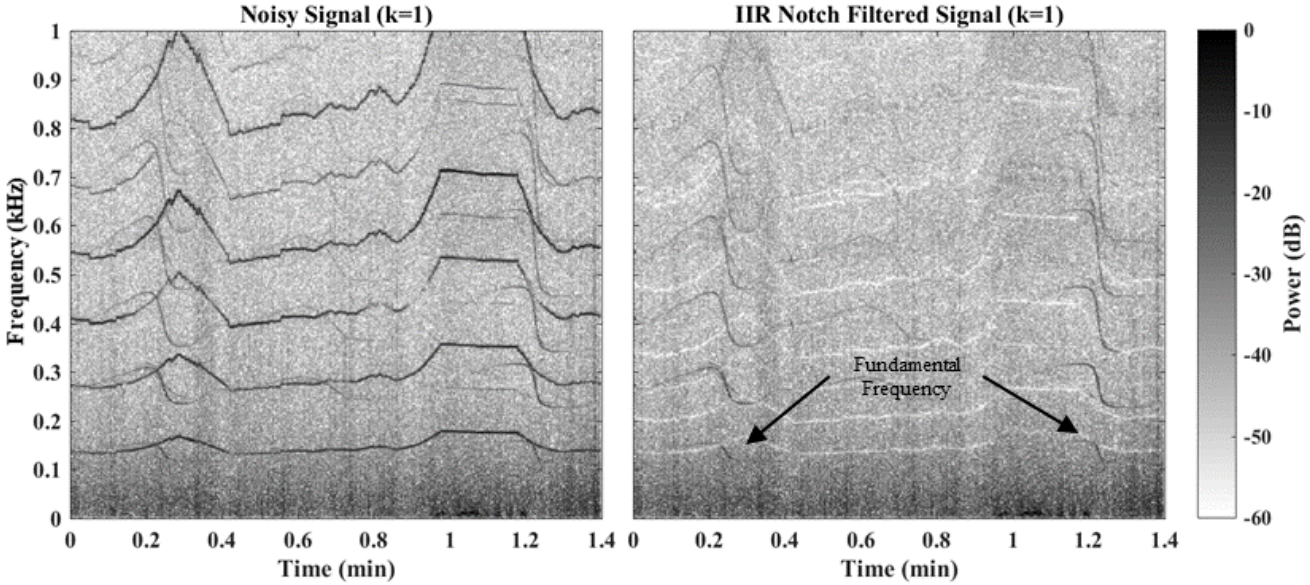

Figure 6-24: TS#2 - Spectrograms for unfiltered and filtered signal segment.

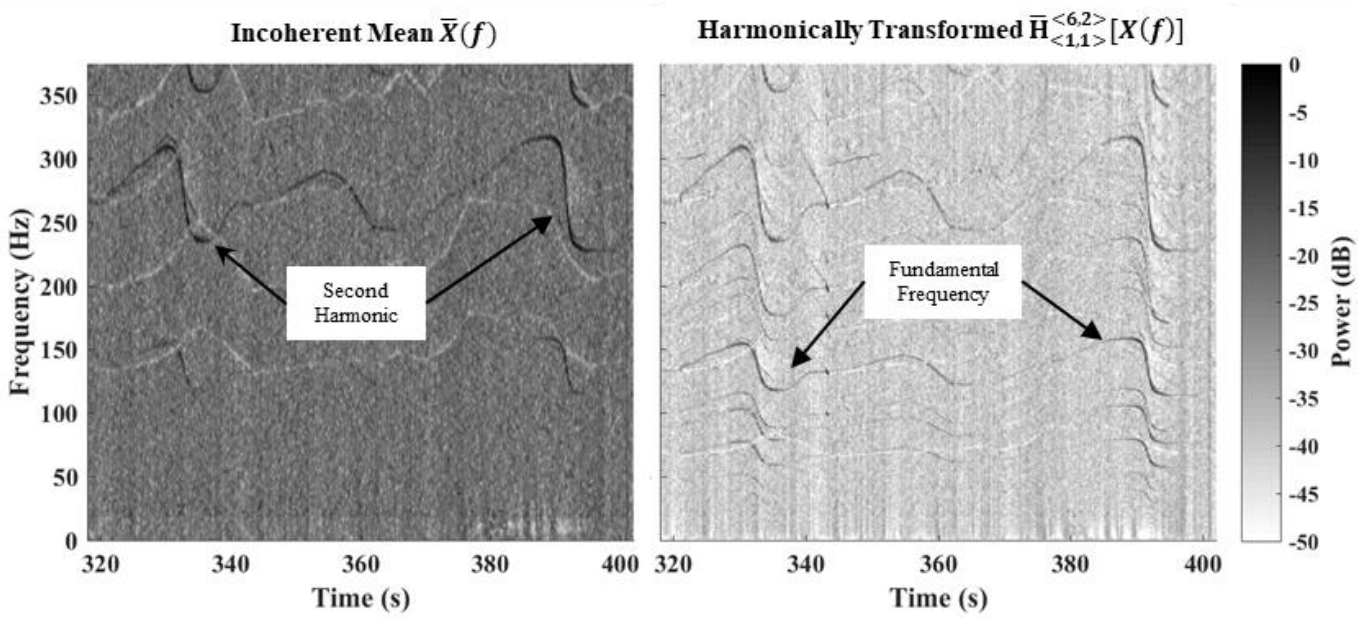

Figure 6-25: TS#2 - Sample spectrograms for the average power and harmonic sum (R=6) processors.

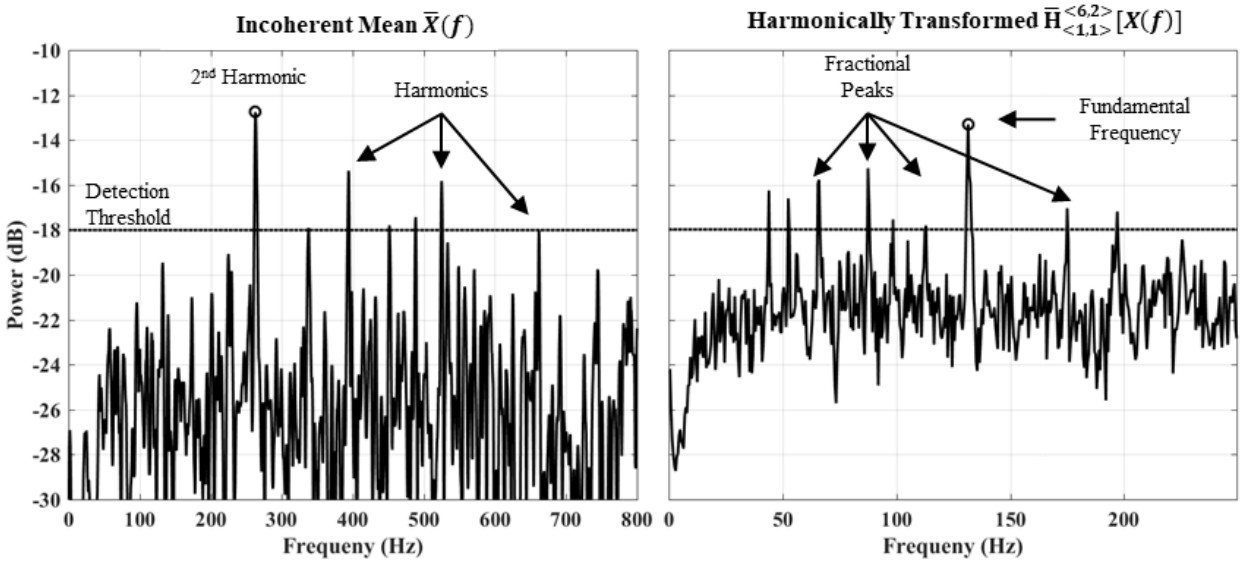

Figure 6-26: TS#2 - Sample spectra plots at a point of detection.

6.4.2.2 - *Source Detection*

Table 6-15 provides the detection results for each enhancement processor and detection scheme. From a comparison of the results obtained, it is evident that the harmonically transformed signals produce the highest detectability for all three schemes. As with the previous experiment, false alarm rates are too high for the single trial scheme to be of any practical use. Binary Integration was found to greatly reduce this value to an acceptable level, but at the expense of signal detectability. Since source signals were highly non-stationary as previously discussed, application of the Robust Binary Integration provided some alleviation to this issue. This is also evident in a more visual form via the histogram plots displayed below in Figure 6-27. From the plots, it is visible that the RBI scheme provides a significant detectability increase in the 200 to 350 m range. This result is significant since this region dictates the lower limit at which another UAV can be detected with adequate

distance to perform an avoidance maneuver as previously indicated by Table 4-4. Based on the results obtained and the requirements displayed bin the table, it is evident that the Giant Big Stik UAV was detected in time to avoid a collision.

Table 6-15: TS#2 - Detection results.

	Detections (#)		Max Distance (m)		Mean Distance (m)	
	$\bar{X}$	$\bar{\mathrm{H}}_1[X]$	$\bar{X}$	$\bar{\mathrm{H}}_1[X]$	$\bar{X}$	$\bar{\mathrm{H}}_1[X]$
Single Trial	1077	1930	672	678	270	302
Binary Integration	215	551	498	572	205	240
Robust Binary Integration	300	884	500	593	212	258

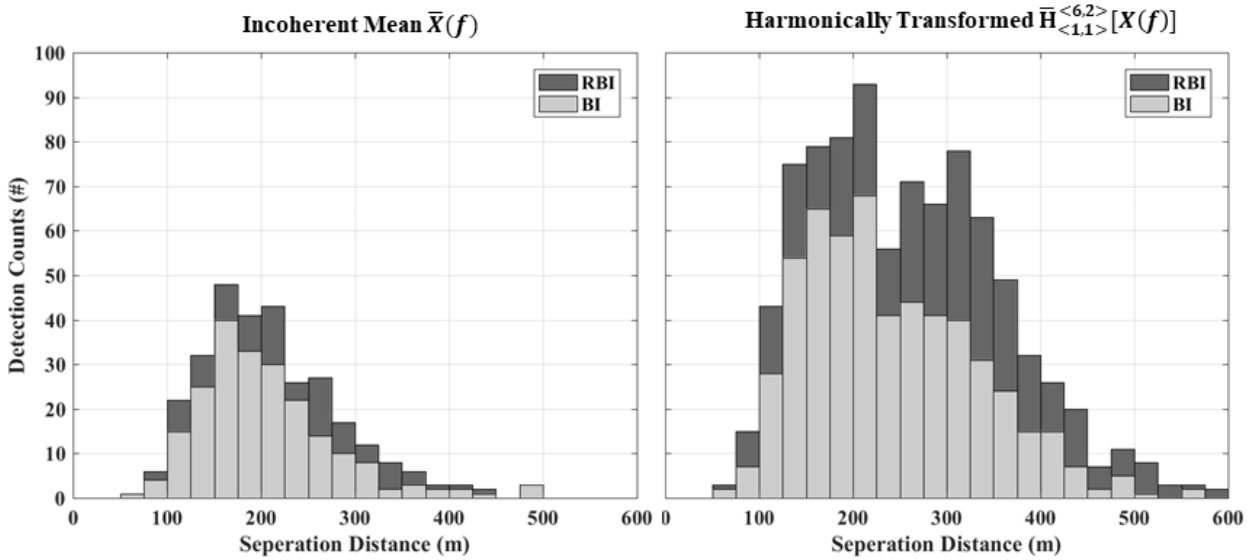

Figure 6-27: TS#2 - Histogram plot of detection counts with respect to separation distance.

Using the basic attenuation laws for acoustic propagation, a crude approximation may also be obtained for the maximum detection distances achievable for a manned aircraft such as the Cessna 185 previously described. Rewriting equation (6.1) in terms of the effective SNR as previously depicted in Figure 2-12 and ignoring the atmospheric attenuation factor gives:

$$d_{max} = d_{ref} 10^{\frac{\Delta SNR}{20}} \tag{6.2}$$

Since it was previously determined that a change in acoustic SPL (dB) is directly proportional to a change in signal power (dB), the above equation can be used to compare theoretical detection distances by simply replacing the change in SNR by the difference in SPL level for the two acoustic sources.

Table 6-16 provides the maximum expected detection results for a Cessna 185 aircraft. Δ SPL values were calculated by comparing the SPL level of the Giant Big Stik UAV used for the experiment to that of the other aircraft in accordance with Table 6-5. Since the atmospheric attenuation factor was omitted, conservative estimates using 50% of the calculated values are also provided. Note that using the attenuation factor would require an iterative calculation approach since the value is a function of separation distance. In addition, because

values are typically on the order of 0.1 dB/100 m, using the 50% approximation will still provide a very conservative estimate. Based on the extrapolated detection distances obtained in combination with the minimum required as previously given by Table 4-4, it appears that the Cessna 185 should also have been detected with sufficient range to avoid a collision; such results also agree with that previously presented in [56]. It should be noted however, that the presented values assume ideal propagation conditions which ignore effects such as wind, temperature gradients, etc., where such effects may further reduce actual detection distances. The modeling of such phenomena is complex as previously discussed in Section 2.3.2 and is outside the scope of this thesis.

Figure 6-28 displayed below provides a plot of the SNR (standard form) with respect to distance for each detection point. Since it was previously found that $SNR \propto SPL$, the observed values may be modelled using the attenuation model according to:

$$SNR = -A \cdot 20\log_{10}(x) - B \cdot x \tag{6.3}$$

where A represents the linear SPL/SNR proportionality or scaling constant, and B represents the atmospheric absorption coefficient such that $A_{abs} = Bx$. From the plot, it is evident that the observed SNR values do behave as expected, thus verifying the above extrapolation approach.

Table 6-16: TS#2 - Extrapolated detection distances for Cessna 185.

ΔSPL	d_{ref}	d_{max}	$d_{max}/2$
15 dB	258 m	1451 m	725 m

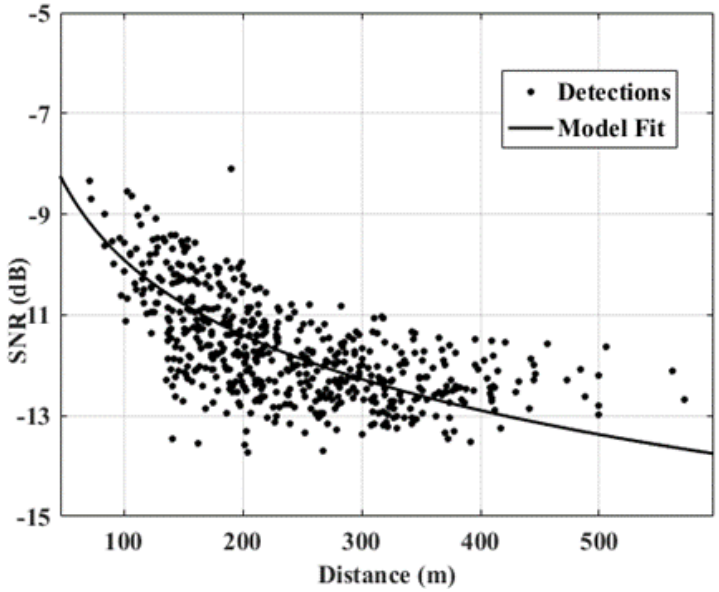

Figure 6-28: TS#2 - Relationship between detected signal SNR and source range.

6.4.2.3 - *Target Localization*

Table 6-18 displayed below provides the localization results obtained for the standard Steered Response Power (SRP), and the SRP of the proposed Harmonic Spectral Beamformer (HSB-SRP) as previously given by equations (5.14) and (5.37) respectively. These results consist of the mean and standard deviation of the angular position error at each detection point. Error values were calculated by comparing the beamforming results to aircraft heading data provided via the two autopilot systems. For each localization method, the SRP is calculated using both the entire noise sample band and that pertaining only to the detected frequency bin (including guard cells). Azimuth values were then determined via the regional contraction method previously proposed in Section 5.3. At total of six harmonics ($H = 6$) were to enhance localization sensitivity. Table 6-17 provides the parameters values used for the regional contraction search. A total of five contractions/reductions were utilized following the initial spatial scan. For subsequent detections, the location from the previous evaluation point (window) is utilized as the contraction starting point instead of re-scanning the entire domain space (360°) again. From the values displayed, it is evident that the method is much more efficient than that of the single stage approach since only 35 evaluation points are required to achieve a resolution of 0.5°. In contrast, a single stage approach would require $360/0.5 = 720$ points to achieve the same resolution.

Table 6-17: TS#2 - Regional contraction search parameters.

1st Stage (Initial Scan)		**3nd Contraction**	
Range	360°	Range	± 4°
Points	12	Points	4
Spacing	30°	Spacing	2°
1st Contraction		**4th Contraction**	
Range	± 30°	Range	± 2°
Points	6	Points	4
Spacing	10°	Spacing	1°
2nd Contraction		**5th Contraction**	
Range	± 10°	Range	± 1
Points	5	Points	4
Spacing	4°	Spacing	0.5°

Since the array effectively consisted of only two microphones (two channels were corrupted), localization of the intruding aircraft was only possible in two dimensions (azimuth). In addition, because the two microphones are also omni-directional, the array cannot discriminate between the forward or equivalent rear positions as depicted below in Figure 6-29. Such a setup is not truly practical for an aircraft collision avoidance system. Localization results are therefore examined simply to demonstrate the use of the Harmonic Spectral Beamformer.

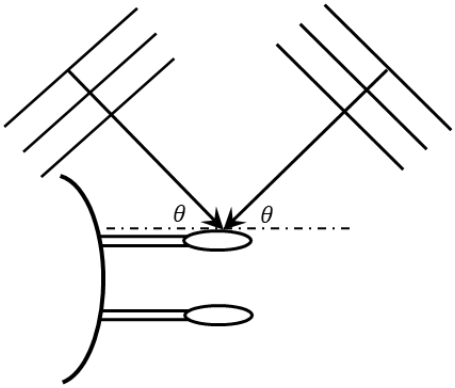

Figure 6-29: TS#2 - Depiction of spatial ambiguity for a two-element array.

Table 6-18: TS#2 - Azimuth localization results.

	SRP		HSB-SRB	
Power Band	10-492.5 Hz	f_o±10 Hz	10-400 Hz	f_o±10 Hz
Mean Error ϑ	24.5°	23.5°	13.0°	11.0°
STD Error ϑ	13.7°	13.2°	9.7°	8.8°

From the localization results displayed, it is evident that the proposed HSB provides increased localization accuracy compared to the standard SRP. This was expected since the method offers a much greater directional performance. This can be observed by the directivity response plots displayed in Figure 6-30 for the two-element array. It is also evident that evaluating only the detected frequency and surrounding guard cells (f_o±10 Hz) provides a slightly increased localization accuracy. Although the increase is only small, the result is significant since it indicates that the full frequency band of interest does not need to be evaluated, but only the region around the detected signal frequency instead. Such an approach offers great computational savings which is important since the SRP method is inherently computationally expensive compared to other techniques as previously discussed in Section 5.2.2.

Overall localization results do not appear particularly accurate, especially for the standard SRP case. Unfortunately, little data is available in the literature to form a direct comparison involving acoustic detection via UAVs. Of the relevant studies previously discussed in Section 2.2.2.6, none clearly present localization results with numerical azimuth and/or elevation error values. Ferguson [52] utilized a small UAV (Aerosonde) fitted with two microphones to detect and localize acoustic impulses from a propane cannon located on the ground. Detection distances of up to 300 m were said to be achieved with a localization bearing angle error of only 3°; although evidence for these claims was not clearly presented. Robertson [53] also conducted experiments where a ground-based propane cannon was detected and localized from a small UAV fitted with four microphones. Detection distances of up to 180 m were said to be achieved with an average localization bearing angle error of 8°. Again however, proof of these claims was not clearly demonstrated. Ohata [54] specified a minimum accuracy of 10° for the deemed-successful localization of a ground-based speaker from a

low-altitude quadcopter. However, the aircraft was fitted with 16 MEMS microphones and utilized the more complex MUSIC algorithm. In addition, SNR values were relatively high (> 0 dB). Reported values obtained from general experimental data using SRP methods are typically less than 5° [268-270]. However, localization accuracy is highly dependent on SNR values which are also much lower than that often reported in the literature. In addition, most experimental instances use much larger arrays with sizes typically on the order of 5 to 10 elements. It should also be noted that the flight data obtained from each aircraft has a heading and positional error of approximately 5° and 3 m respectively. For the HSB case, these values are significant since the heading error alone represents almost 50% of the total perceived localization error (source angular position is calculated relative to the detecting aircraft orientation). Thus, when considering these facts, it appears the results obtained are relatively good for the scenario at hand; at least for the HSB case. One would expect that an avoidance maneuver can be reliably performed under these circumstances by simply changing course such that the target source is now located at ± 90° (head perpendicularly away from target flightpath). However, determination of the minimum allowable localization error for a given kinematic setup is complex and thus outside the scope of this thesis.

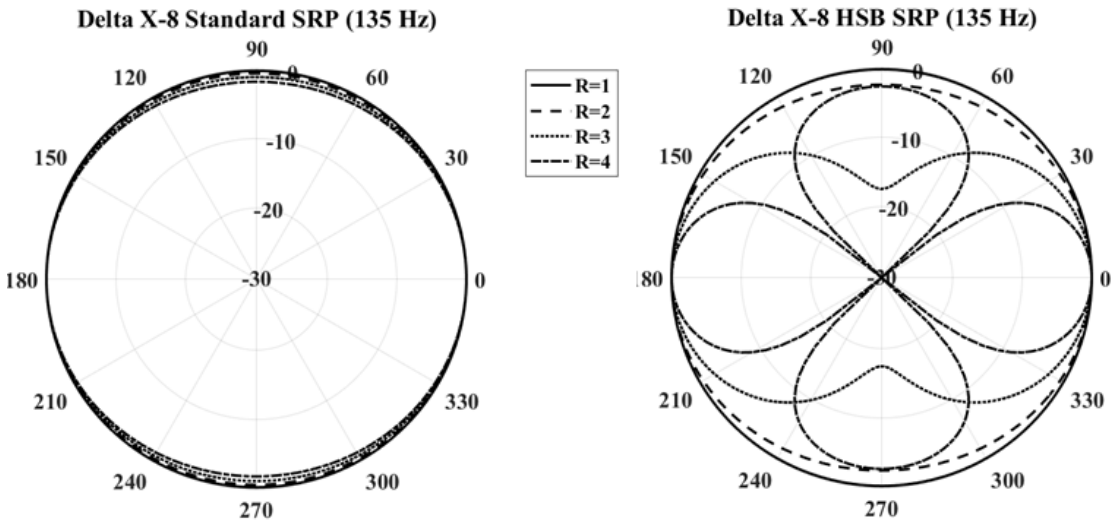

Figure 6-30: TS#2 - Directivity response for the two-element array.

6.4.3 - Conclusions

Based on the detection results obtained, it can be concluded that the Giant Big Stik UAV was detected with sufficient distance to perform an avoidance maneuver. Extrapolated detection distances also suggest that the system should also have been able to detect a Cessna 185 aircraft with sufficient distance to avoid a collision (> 550 m). The proposed Harmonic Spectral Transform was found to produce significantly increased signal detectability and overall range compared to the standard incoherent mean. Use of the proposed Robust Binary Integration scheme also offered superior detection performance since the acquired signals were highly non-stationary. Results obtained from the localization analysis indicate the proposed Harmonic Spectral

Beamformer offers superior performance to that of the standard Steered Response Power. A higher positional accuracy was achieved with less variation, and ultimately required less computational load.

6.5 - TS#3: Multirotor Air-to-Air

6.5.1 - Purpose & Procedure

The third test was conducted to determine whether acoustic sensing could be utilized with a multirotor UAV to localize another moving aircraft with sufficient distance to perform an avoidance maneuver. In terms of signal processing, use of the Harmonic Spectral Transforms (HSTs), Phase Acceleration Processors (PAPs), Modified Coherence Processors (MCPs), and combined versions of these processors to enhance signal detection will be evaluated. In addition, use of the Harmonic Spectral Beamformer to localize aircraft in 3D space will also be demonstrated.

The experiment involved flying a Cessna 185 aircraft (intruder) in circuit formation around the Kraken UAV at various distances, speeds, headings, and altitudes. Figure 6-31 provides the flight path for the intruding aircraft along with the general location of the detecting UAV. The total duration of the experiment was 834 s which consisted of 12 close encounters of various headings and distances.

Due to safety concerns over any actual collision taking place, the detecting UAV remained at a relatively low altitude (≈ 100 m) with little range movement (≈ 50 m with respect to takeoff position). In addition, the aircraft was flown under manual operation to reduce the possibility of a fly-away caused by a malfunctioning autopilot system. Both the UAV operator and aircraft pilot were also in constant communication throughout the test via an air-band radio system. As with previous experiments, flight data was logged using the ArduPilot 2.5 system.

Table 6-19: TS#3 - Kinematic parameters of intruder aircraft.

	Minimum	Maximum	Median	Mode
Speed (knots)	82	121	98	100
Altitude (m)	220	335	296	290
Range (m)	111	1500	578	650

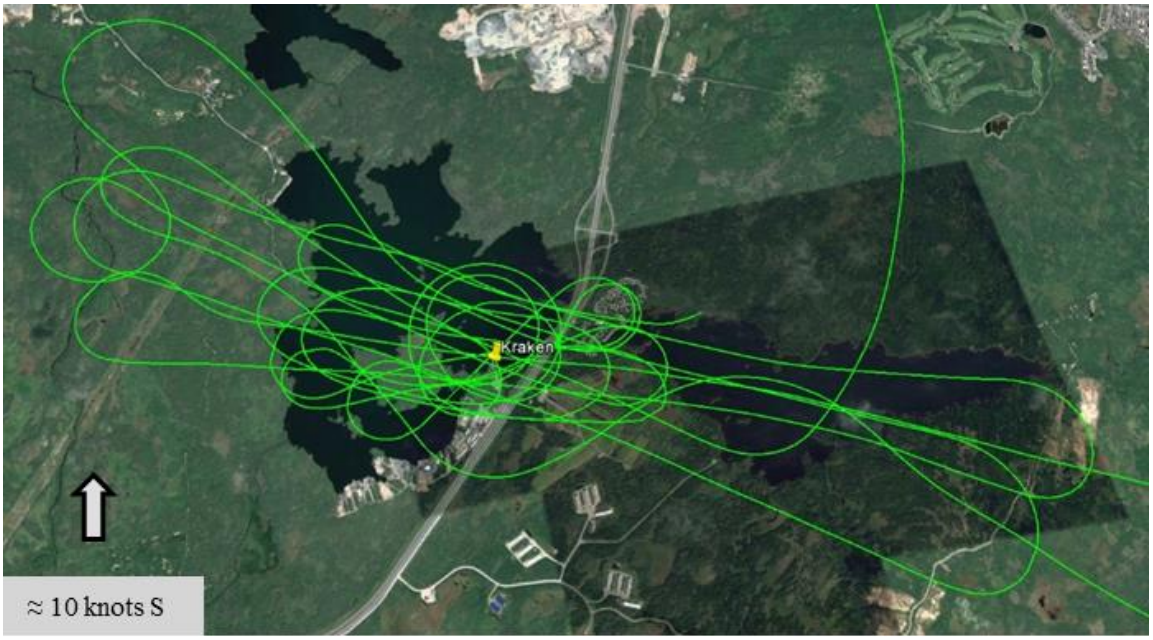

Figure 6-31: TS#3 - Cessna 185 flight path.

6.5.2 - Results & Discussion

6.5.2.1 - *Signal Processing*

Tables 6-20 to 6-22 provides the notch filter, FFT, spectral whitening, and SCDF-CFAR parameters used. Figure 6-24 provides sample spectrogram plots, which demonstrate the effectiveness of the adaptive MIMO IIR notch filter approach proposed in Section 3.3.3. A total of six harmonics ($R = 6$) and six signals ($S = 6$) were used for the HST, which again consisted of the standard mean form. To verify the results previously obtained for the enhancement processors via simulation studies, the proposed methods are applied here again using experimental data.

In addition to the proposed processors, the incoherent mean, Wagstaff's PAC [163], and the Generalized Magnitude Squared Coherence (GMSC) developed by Ramirez [197] are also applied to form a comparative basis. Table 6-23 provides a list of the enhancement processors used. For each of the processors, the HST was also applied since the source signal was known to be harmonic in nature. Note that the Generalized HST operation ($\overline{\mathrm{H}}_{<a,b,c>}^{<R,S,W>}[\,]$) is simply represented by ($\overline{\mathrm{H}}[\,]$) for sake of simplicity and neatness.

As with previous experiments, signal detection was performed using the SCDF-CFAR detector. Since the recorded signals were decimated to a final sampling rate of 6000 Hz and a total of 6 harmonics were used for the HST, the maximum noise sample band ranged from 0 to 500 Hz. Each of the processor forms also utilized the same sample noise bandwidth. Presented false alarm probabilities were again calculated using the base functional form given by equation (4.149), since all signals were harmonically transformed.

Table 6-20: TS#3 - Signal preprocessing and filter parameters.

Sampling Frequency (f_s)	48 kHz	Number of Signals	6
Decimation Factor	8	FFT Window	0.5 s
IIR Step Size ($\mu_1, \mu_1, \ldots, \mu_6$)	(10,15,20,25,30,35) E-5	Window Overlap	50%
Notch Radius (r)	0.998	Padded Length (L_{fft})	12,000
Harmonics Removed (R)	10	Spectral Resolution (f_r)	0.5 Hz/bin

Table 6-21: TS#3 - CFAR-Enhanced Spectral Whitening parameters.

Detector Type	OS-CFAR	Noise Samples (N)	102
Forgetting Factor (ξ)	0.2	Order Statistic (k)	0.75 N
Flooring Factor (δ)	0.5	Guard Cell Band ($\vec{G}$)	5.5 Hz
Noise Band ($\vec{N}$)	50 Hz	Guard Cells (G)	12

Table 6-22: TS#3 - SCDF-CFAR detection parameters.

Noise Sample Band ($\vec{N}$)	1 – 499 Hz	Order Statistic ($\bar{k}$)	2
Test Band ($\vec{B}$)	70 – 105 Hz	Consecutive Trials (T)	2
Guard Cell Band ($\vec{G}$)	10 Hz	Consecutive Detections (D)	2
Fractional Guard Cell Band ($\vec{G}_F$)	1 Hz	Cell Deviation (Δ)	1
Noise Samples (N)	995 pts	Maxima Tested (M)	2
Test Cells (B)	71 pts	P_{FA}^{SC}	2.2E-3
Guard Cells (G)	22 pts	P_{FA}^{ST}	0.15
Fractional Guard Cells (G_F)	2 pts	P_{FA}^{BI}	3.3E-4
Fractional Peaks (F)	17	P_{FA}^{RBI}	9.9E-4

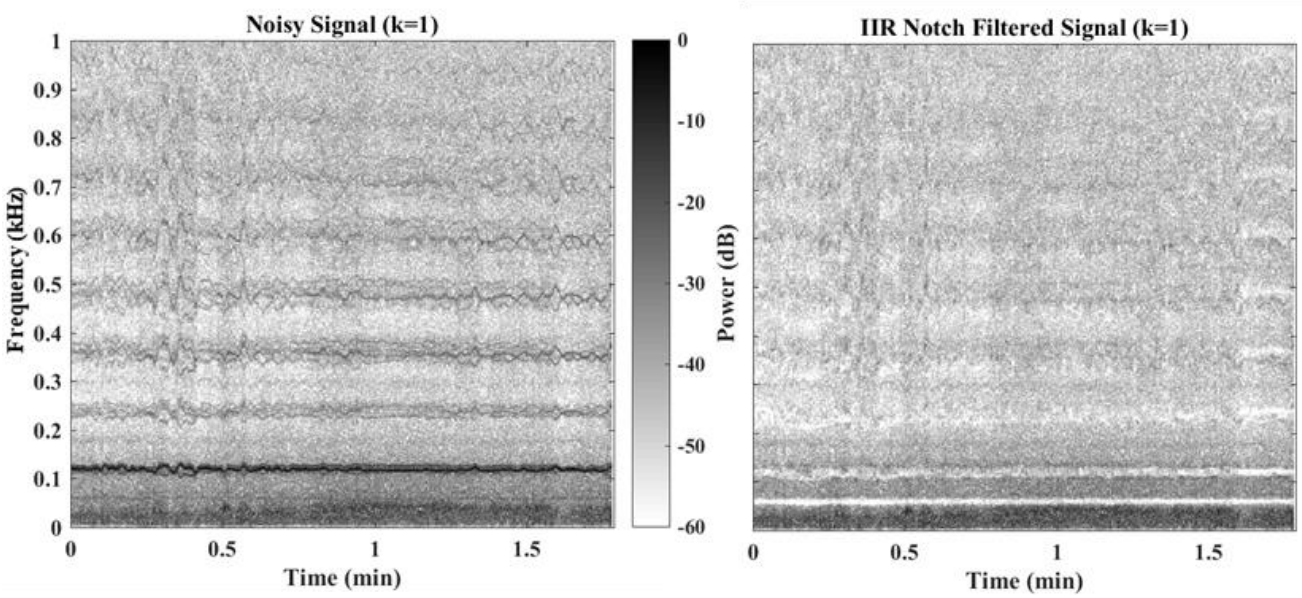

Figure 6-32: TS#2 - Spectrograms for unfiltered and filtered signal segment.

Table 6-23: TS#3 - List of enhancement processors utilized.

$\bar{X}$	Incoherent Mean of FFT Magnitude
$\bar{\mathrm{H}}[X]$	HST of FFT Magnitude (Standard HST)
$\bar{\mathrm{H}}[PAC]$	Wagstaff's PAC
$\bar{\mathrm{H}}[\vec{\Phi}]$	Acceleration Vector Coherence (AVC)
$\bar{\mathrm{H}}[\vec{\Phi}^{\Psi}]$	Adjusted Acceleration Vector Coherence (A-AVC)
$\bar{\mathrm{H}}[X \cdot \vec{\Phi}]$	Combined FFT Magnitude & AVC
$\bar{\mathrm{H}}[X \cdot \vec{\Phi}^{\Psi}]$	Combined FFT Magnitude & A-AVC
$\bar{\mathrm{H}}[\Phi_{\lambda}]$	System Acceleration Coherence (SAC)
$\bar{\mathrm{H}}[\Phi_{\lambda}^{\Psi}]$	Adjusted System Acceleration Coherence (A-SAC)
$\bar{\mathrm{H}}[X \cdot \Phi_{\lambda}^{\Psi}]$	Combined FFT Magnitude & SAC
$\bar{\mathrm{H}}[X \cdot \Phi_{\lambda}]$	Combined FFT Magnitude & A-SAC
$\bar{\mathrm{H}}[\tilde{\Gamma}]$	Generalized Magnitude Squared Coherence (GMSC)
$\bar{\mathrm{H}}[\tilde{\Gamma}^{\Psi}]$	Generalized Acceleration Squared Coherence (GASC)

6.5.2.2 - *Source Detection*

Table 6-24 provides the results for each enhancement processor and detection scheme utilized. The results are provided in terms of the number of detections with a range sufficient to avoid a head-on collision, the maximum and mean detection distance, and the minimum SNR obtained from all detections. Note that the approximate minimum required range for a Cessna 185 was previously listed in Table 4-4. The enhancement processors are sorted from largest to smallest relative to the detection range count to aid in identifying the top performing method. Distance and SNR values are only provided for the Binary Integration case since the results were nearly identical to that of the Robust Binary Integration scheme. The adjusted AVC and SAC processors were both used with a scaling value of $\beta = 0.9$, while the coherence processors were applied using four windowed segments. To form a better comparative basis between the phase acceleration and coherence based processors, results are also provided for the non-coherence forms using a total of $W = 4$ windowed segments. Referring back the Generalized HST functional form given by equation (4.35), this would be represented as $\bar{\mathrm{H}}_{<a,b,c>}^{<R,S,W>}[\,] = \bar{\mathrm{H}}_{<1,1,1>}^{<6,6,4>}[\,]$.

Based on the results displayed, it is evident that all of the processors achieve detection distances greater than the minimum required, which was approximately 550 m. Maximum distances ranged from 1381 to 1463 m, which is in agreement with the extrapolated value of 1451 m established in the previous experimental study (Table 6-16). Figure 6-33 provides plots of the separation distance at the various detection points for the best and worst enhancement processors. It is evident from the plots, that the combined FFT Mag. & A-SAC processor facilitated detection with sufficient range for each approaching run, thus maintaining the 99.5% detection requirement outlined in Section 4.3.2.1. In addition, the aircraft was also detected at the peak range locations for each of these runs. However, it is apparent that values close to the maxima when advancing or retreating are often not detected. This can be explained by the fact that the majority of the sound propagation from a propeller-driven aircraft is directed radially with respect to the propeller shaft axis [73]. Thus, when the

aircraft is advancing or retreating, the immediate sound propagation is perpendicular to the direction of flight and location of the sensing aircraft (for the head-on case). At the maximum range points, the aircraft performs a bank maneuver to reverse heading for the next approach. At this point, the sound propagation direction and sensing aircraft location are now parallel, effectively increasing transmission efficiency and thus facilitating better detection. In contrast to the proposed processor, the incoherent mean only obtained sufficient detection range for 4 of the 12 encounters. This illustrates the necessity of the proposed enhancement processors to establish an effective collision avoidance system.

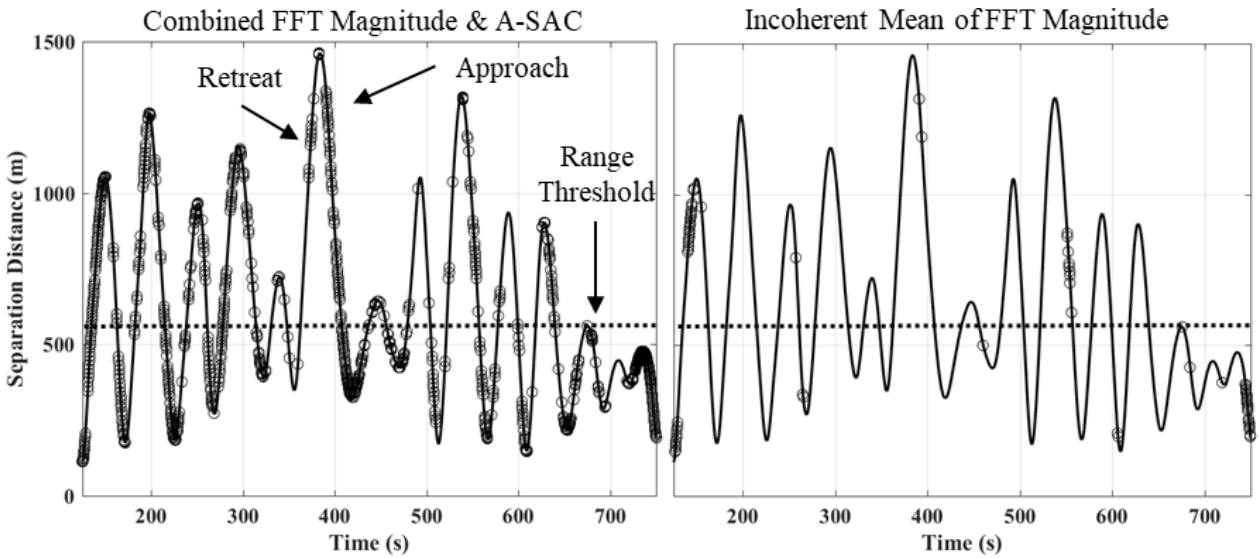

Figure 6-33: TS#3 - Separation distance at detection points for best and worst performing processors.

Table 6-24: TS#3 - Enhancement processor detection results.

$\boldsymbol{W=1}$	Detections > 550m		Max Range	Mean Range	Min SNR	$\boldsymbol{W=4}$	Detections > 550m		Max Range	Mean Range	Min SNR
	BI	RBI	BI	BI	BI		BI	RBI	BI	BI	BI
$\bar{\mathrm{H}}[\tilde{\Gamma}^{\Psi}]$	457	467	1426	623	-31.4	$\bar{\mathrm{H}}[X \cdot \Phi_{\lambda}^{\Psi}]$	694	709	1462	639	-32.3
$\bar{\mathrm{H}}[X \cdot \Phi_{\lambda}^{\Psi}]$	426	488	1463	647	-30.3	$\bar{\mathrm{H}}[X \cdot \vec{\Phi}]$	658	687	1463	634	-30.3
$\bar{\mathrm{H}}[X \cdot \vec{\Phi}]$	422	481	1463	657	-30.3	$\bar{\mathrm{H}}[X \cdot \Phi_{\lambda}]$	631	658	1461	642	-31.9
$\bar{\mathrm{H}}[X \cdot \Phi_{\lambda}]$	416	487	1462	653	-30.3	$\bar{\mathrm{H}}[PAC]$	604	620	1460	641	-31.2
$\bar{\mathrm{H}}[\tilde{\Gamma}]$	401	416	1396	616	-27.0	$\bar{\mathrm{H}}[X]$	592	620	1461	645	-29.5
$\bar{\mathrm{H}}[X]$	400	468	1463	661	-28.3	$\bar{\mathrm{H}}[\Phi_{\lambda}^{\Psi}]$	572	582	1461	663	-28.0
$\bar{\mathrm{H}}[X \cdot \vec{\Phi}^{\Psi}]$	391	429	1460	647	-30.0	$\bar{\mathrm{H}}[X \cdot \vec{\Phi}^{\Psi}]$	556	575	1455	645	-31.4
$\bar{\mathrm{H}}[PAC]$	361	424	1447	652	-30.3	$\bar{\mathrm{H}}[\vec{\Phi}]$	491	503	1462	657	-26.6
$\bar{\mathrm{H}}[\vec{\Phi}^{\Psi}]$	248	267	1377	649	-27.2	$\bar{\mathrm{H}}[\Phi_{\lambda}]$	488	502	1461	637	-26.4
$\bar{\mathrm{H}}[\vec{\Phi}]$	239	275	1377	646	-26.4	$\bar{\mathrm{H}}[\tilde{\Gamma}^{\Psi}]$	457	467	1426	623	-31.4
$\bar{\mathrm{H}}[\Phi_{\lambda}]$	233	267	1377	637	-26.4	$\bar{\mathrm{H}}[\vec{\Phi}^{\Psi}]$	417	429	1377	650	-31.2
$\bar{\mathrm{H}}[\Phi_{\lambda}^{\Psi}]$	203	223	1377	610	-26.4	$\bar{\mathrm{H}}[\tilde{\Gamma}]$	401	416	1396	616	-27.0
$\bar{X}$	180	205	1381	624	-17.6	$\bar{X}$	180	205	1463	624	-17.6

It is also evident from the results displayed that the proposed GASC obtained the highest number of detections and lowest detectable SNR. This was closely followed by the combined FFT Mag. & A-SAC, and the combined FFT Mag. & AVC processors. This result was expected since these processors also attained the highest detection performance in the simulation study previously presented in Section 4.4.2. In comparison to the GMSC developed by Ramirez, the proposed phase acceleration form (GASC) produced significantly better results with a detection increase of 14%, and a 4.4 dB decrease in detectable SNR. The proposed phase acceleration processors (PAPs) performed relatively poorly when utilized independently. However, the combined versions of these processors produced significantly better results than either of the PAPs or HST alone.

For the case of using four processing windows ($W = 4$), it is apparent that a significant increase in detectability is achieved for all the non-coherence-based processors. The combined HST-PAP processors now attained the highest performance values, with a significant increase over the previously top performing processor (GASC). A detection increase of 63%, and a 2.0 dB decrease in detectable SNR was achieved for the combined FFT Mag. & A-SAC processor. Based on these results, it appears that the coherence processors do not perform as well, if an equivalent number of processing windows had been used for the HST instead.

For both window cases, it is evident that the Robust Binary Integration scheme produced increased source detectability as expected. However, relative detection increases were not as significant as that obtained for previous experiments. This can be attributed to the fact that observed source frequencies were much more stationary with respect to time, since the total duration to perform a fly-by was significantly longer. Figure 6-34 provides detection histograms for the GASC and combined FFT Mag. & A-SAC processors. From the plots, it is evident that the RBI scheme provides increased detectability at lower separation distances, where increased variation in source signal frequencies are observed.

It should be noted that the performance of the processors is highly dependent on the evaluation parameters used. Such parameters would include: FFT length, window overlap, window type, number of processing windows, phase adjustment scale, and number of signals. Varying these values will produce different results which may indicate an alternate top performing processor. For example, Table 6-25 provides the results obtained for the top six processors when using a Hamming FFT window instead of the rectangular form previously used. From a comparison of the two data sets, it is evident that applying the window function greatly reduces the detectability performance of the processors. The relative performance between the various forms is also modified, which is evident by the fact that the GMSC now produces the third largest number of detections. In addition, the minimum detectable SNR for this processor is also decreased by 3.6 dB, while the GASC is actually increased by 3.7 dB. Modifying other parameters such as FFT length would also have a significant effect on relative and overall performance levels. However, optimization of these values for each processor type is outside the scope of this thesis. Here a basic demonstration regarding the effectiveness of the proposed methods in simply provided using values typical for the application area.

Table 6-25: TS#3 - Enhancement processor detection results (Hamming window).

$W = 1$	Detections > 550m		Max Range	Mean Range	Min SNR
	BI	RBI	BI	BI	BI
$\bar{\mathrm{H}}[\tilde{\Gamma}^{\Psi}]$	399	415	1384	620	-27.7
$\bar{\mathrm{H}}[\mathrm{X} \cdot \Phi_{\lambda}]$	394	457	1463	670	-33.4
$\bar{\mathrm{H}}[\tilde{\Gamma}]$	388	402	1377	601	-30.6
$\bar{\mathrm{H}}[X \cdot \Phi_{\lambda}^{\Psi}]$	380	438	1463	664	-32.4
$\bar{\mathrm{H}}[X \cdot \vec{\Phi}]$	377	438	1377	662	-31.6
$\bar{\mathrm{H}}[X]$	373	427	1462	669	-29.4

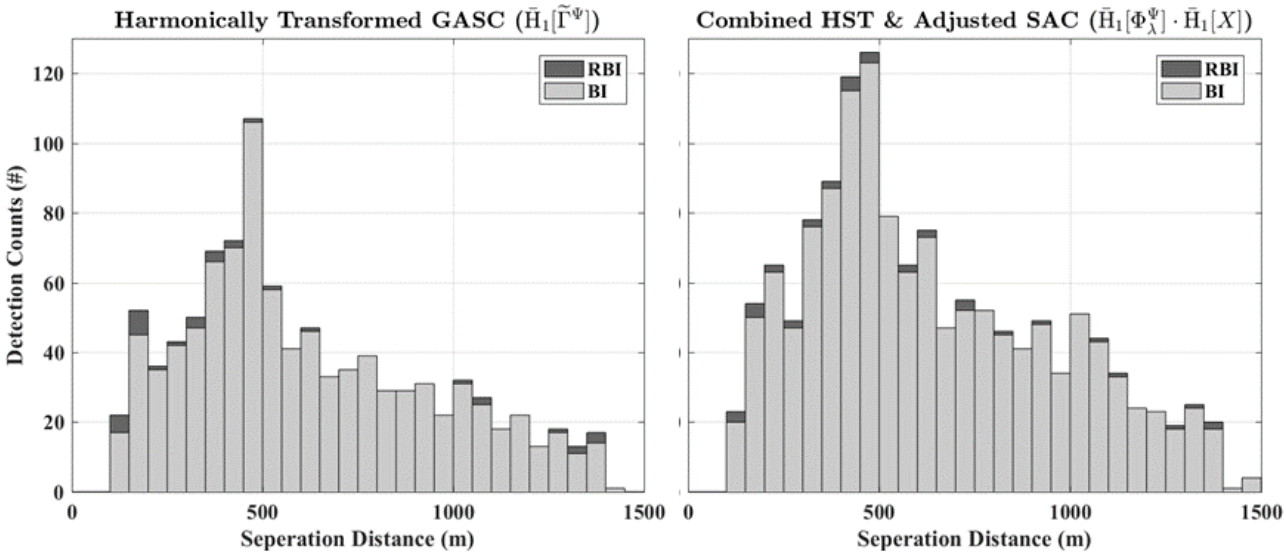

Figure 6-34: TS3# - Histogram plots of detection counts with respect to separation distance.

6.5.2.3 - *Target Localization*

Table 6-27 provides the localization results obtained for the standard SRP and proposed HSB SRP beamformer. These results consist of the mean and standard deviation of the angular position error at each detection point. Error values were calculated by comparing the beamforming results to aircraft heading data provided via the two ArduPilot systems. For each localization method the SRP is calculated using both the entire noise sample band and that pertaining only to the detected frequency bin (including guard cells). Azimuth and elevation values were then determined via the regional contraction method previously proposed in Section 5.3. Table 6-26 provides the parameter values used for the regional contraction search. A total of four elevation and five azimuth contractions were utilized following the initial spatial scan. For subsequent detections, the previous evaluation point (window) was used as the contraction starting point instead re-scanning the entire domain space (360×90°) again. From the values displayed, it is evident that the method is much more efficient than that of the single stage approach since a total of 35 + 19 = 54 evaluation points are required to achieve a resolution of 0.5° in either dimension. In contrast, a single stage approach would require 360×90/0.5 = 64,000 points to achieve the same resolution.

Table 6-26: TS#2 - Regional contraction search parameters.

	Azimuth (ϑ)	**Elevation (φ)**		**Azimuth (ϑ)**	**Elevation (φ)**
1st Stage (Initial Scan)			**3nd Contraction**		
Range	360°	90°	Range	± 4°	± 2°
Points	12	6	Points	4	4
Spacing	30°	15°	Spacing	2°	1°
1st Contraction			**4th Contraction**		
Range	± 30°	± 15°	Range	± 2°	± 1°
Points	6	6	Points	4	4
Spacing	10°	5°	Spacing	1°	0.5°
2nd Contraction			**5th Contraction**		
Range	± 10°	± 5°	Range	± 1	-
Points	5	5	Points	4	-
Spacing	4°	2°	Spacing	0.5°	-

From the localization results displayed below, it is again evident that the proposed HSB provides a significantly increased accuracy compared to the standard SRP approach. This was expected since the proposed method offers a much more localized directional response, which can be observed from the plots displayed in Figures 6-36 and 6-37. An 85 Hz signal was utilized since this is a typical fundamental frequency value for the Cessna 185 aircraft. It should also be noted that the elevation response plots are not symmetrical since the microphones have a cardioid directional response and the array is not symmetrical about the elevation axis of rotation.

Table 6-27: TS#3 - Localization results.

	SRP		**HSB-SRB**	
	1-499 Hz	f_o±5 Hz	1-499 Hz	f_o±5 Hz
Mean Error ϑ	13.5°	15.1°	6.1°	5.8°
Mean Error φ	15.5°	17.2°	13.3°	14.2°
Total Error $\sqrt{\vartheta^2 + \varphi^2}$	20.6°	22.9°	14.6°	15.4°
STD Error ϑ	16.5°	17.4°	8.6°	8.4°
STD Error φ	13.7°	13.0°	14.2°	13.9°
STD Total Error $\sqrt{\vartheta^2 + \varphi^2}$	21.4°	21.7°	14.6°	14.2°

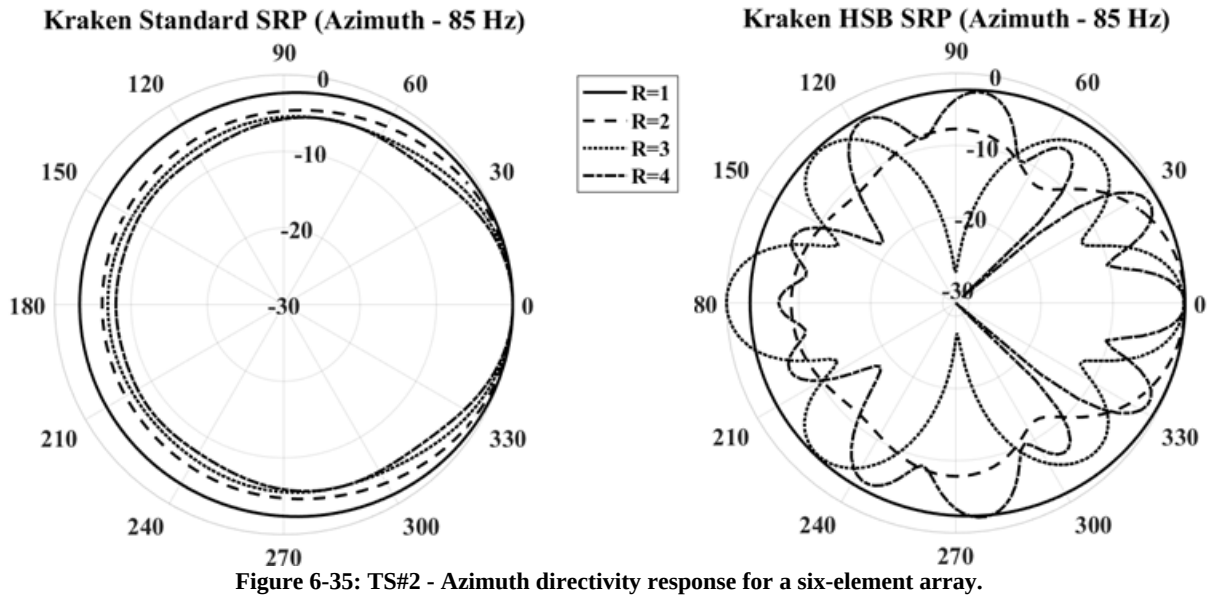

Figure 6-35: TS#2 - Azimuth directivity response for a six-element array.

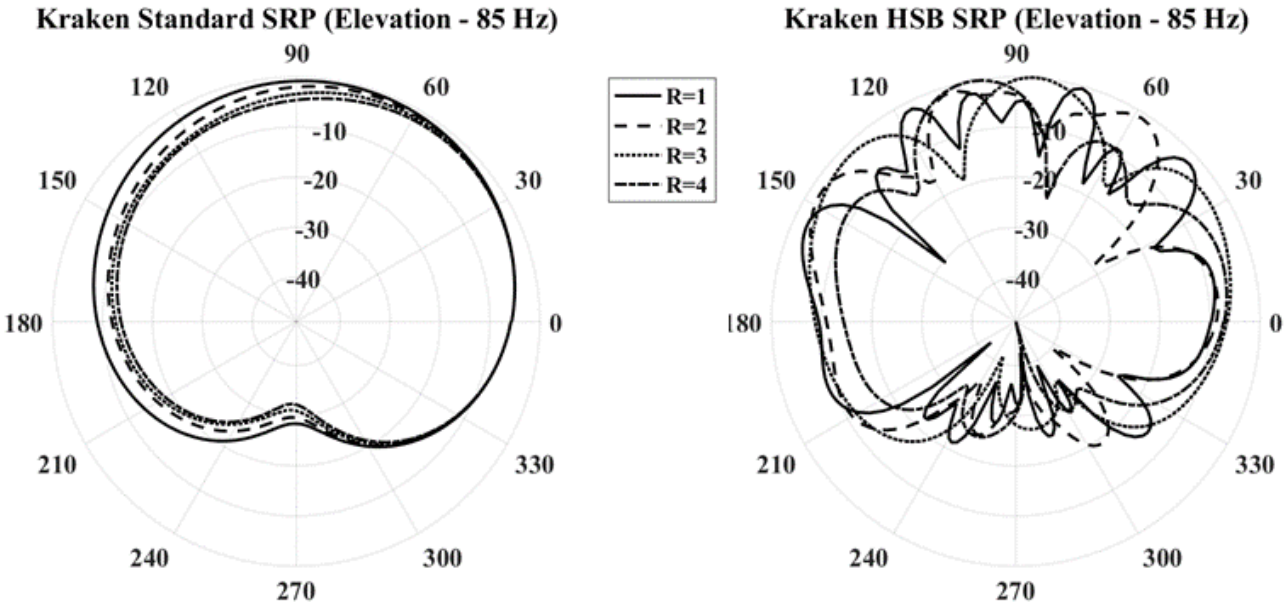

Figure 6-36: TS#2 - Elevation directivity response for a six-element array.

The narrowband evaluation method (f_o±5 Hz) produced a slightly increased azimuth and slightly decreased elevation accuracy for the HSB SRP method. In general, azimuth accuracy values were much greater than that obtained in the elevation direction for all methods. However, accuracy in this direction is generally much less important since avoidance maneuvers are mainly governed by heading modifications in the horizontal plane. This decrease in accuracy was also expected since the array offers a much better directional response in the azimuth direction. Overall, error values were considerably less than that obtained from the previous experiment, which was also expected since a larger number of array elements were used. When considering the fact that the mean angular heading error for the flight data recorder was approximately 5 °, the azimuth values obtained for the HSB SRP can be considered quite accurate. That is, the error values obtained can almost entirely be

attributed to the recorded heading/positional error associated with the device. Figure 6-37 displays plots of the recorded and predicted azimuth and elevation angles at each detection point. From the plots, it is evident that azimuth values do in fact agree with the measured/recorded values very well. Elevation angles also agree fairly well for higher valued positions. However, values in the area of $\leq 15°$ do not correlate as well. This can be explained by the fact that the array has a lower directivity gain for smaller elevation angles, which evidently occurs at larger separation distances producing lower SNR values.

In addition to the error associated with the flight data recorder(s), there is also the inherent problem associated with acoustic propagation delays. That is, the observed position of the aircraft as determined from the arriving sound wave will actually be the position when the information was originally produced. Thus, beamforming methods effectively determine where the aircraft was not where the aircraft currently is. In general, this effect will be larger and of greater concern for increasing relative angular velocities with respect to the direct line of sight. Depending on the relative heading between the two aircraft, this effect may be insignificant however. For example, the angular position error for a head-on collision will be largely unaffected by wave transit effects, since the relative angular velocity will be approximately zero. Methods have been developed to deal with this issue using both time-domain [39, 279] and frequency domain techniques [280, 281]. However, these either require knowledge of the source fundamental frequency and/or high SNR values such that changes in frequency and amplitude caused by relative motion effects can be accurately observed. In addition, all the methods assume a stationary observer and moving source. There are no models published in the literature to address the scenario of a moving source and moving observer to the author's knowledge. It should be noted however, that accurate target localization for any given instant is not completely vital provided subsequent measurements are precise and exhibit low random error. For example, if the target aircraft maintains a constant speed and heading, delayed localization values can still be utilized to determine an accurate trajectory approximation through methods such as that presented in [39].

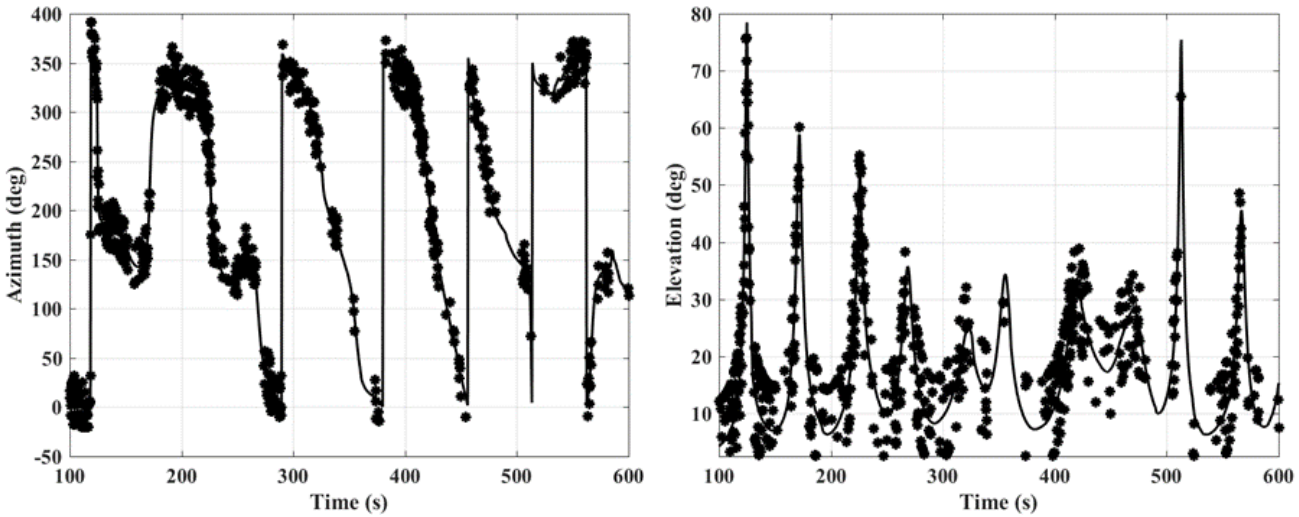

Figure 6-37: TS#3 - Sample segment illustrating localization accuracy of detection points.

6.5.3 - Conclusions

Based on the detection results obtained, it can be concluded that the Cessna 185 manned aircraft was detected and localized with sufficient range to perform an avoidance maneuver. The proposed spectral enhancement processors provided a significant increase in signal detectability and range compared to more standard methods, such as the incoherent mean and generalized coherence. However, it was also shown that the use of such methods is highly influenced by the various analysis parameters used. The Robust Binary Integration scheme was found to increase signal detectability but was only significant for short separation distances where observed signal frequencies become increasingly non-stationary. Finally, localization results suggest the proposed Harmonic Spectral Beamformer offers superior performance to that of the standard Steered Response Power; a higher positional accuracy was achieved with less variation and ultimately required less computational load. One would expect that an avoidance maneuver can be reliably performed based on the accuracy obtained. However, more work in this area is required to determined minimum allowable error values for unmanned system operations.

- 7 - Conclusions & Future Work

7.1 - Conclusions

Based on the experimental work presented in this thesis, it appears that acoustic sensing may in fact be a viable technology to establish a non-cooperative collision avoidance system for UAVs. Results obtained from experiments conducted verify that both manned and unmanned aircraft can be detected and localized with sufficient range, accuracy, and reliability to perform an avoidance maneuver. It was also shown however, that such results were only made possible using the digital signal processing developments proposed throughout this thesis. These developments are now summarized in the following section with respect to the chapter in which they were presented.

7.2 - Summary of Results & Contributions

The following provides a brief summary of the results obtained and theoretical contributions made in the areas of adaptive notch filtering, signal enhancement, source detection, and source localization.

7.2.1 - Narrowband Noise Removal

A number of techniques to adaptively filter harmonic narrowband noise without using any reference signal or producing any phase distortions was proposed. These included:

1) A distortionless zero-phase FIR notch filtering method via the use of a second-order IIR notch filter prototype.
2) A distortionless zero-phase notch filtering method via the use of FIR Comb filters.
3) Multichannel IIR notch filtering methods for SIMO and MIMO systems.

In addition, developments made to facilitate SIMO and MIMO systems were also applied to the proposed FIR Notch and FIR Comb filters. The performance of the methods was confirmed using both computer generated and experimental data.

Based on the results obtained from simulated and experimental data, it was concluded that all the proposed filtering methods provided an effective means to remove non-stationary harmonic noise without the use of any reference signal and without producing any phase distortions. The IIR notch filter offered the best performance in all filtering scenarios examined, with frequency tracking capabilities exceeding that of the proposed FIR notch and Comb filters. Modifications made to the filter to facilitate multichannel systems with partial harmonic components proved to be essential in order to effectively filter signals obtained via multirotor experiments. The proposed FIR notch filter also provided similar results to that obtained via the IIR form, proving that the method is an effective alternative for situations in which a linear-phase inherently stable filter is required. It was shown that the filter may also be effectively used for multichannel systems, although computational requirements may limit the viability for applications in which many noise sources are present. The Comb filter performed the worst out of the proposed methods. However, the filter was still effective in removing noise components for all

scenarios examined. As with the FIR form, filtering large numbers of source components may not be practical for some applications due to the processing delays needed to generate a distortionless output. However, the method does give the advantage of requiring very few computational resources compared to the other proposed forms. Thus, for applications which demand low computational loads, a zero-phase output, and inherently stable operation, this filter may offer the best solution.

7.2.2 - Signal Enhancement

A number of signal processing techniques were proposed to enhance the detection of continuous harmonic narrowband signals. These methods include:

1) A generalized spectral transform to exploit the periodic peak nature of harmonic signals in the frequency domain (Harmonic Spectral Transforms - HSTs).
2) A series of processors which exploit the phase acceleration properties of continuous periodic signals (Phase Acceleration Processors - PAPs).
3) Modifications to the generalized coherence function for multichannel systems to include phase acceleration information (Modified Coherence Processors - MCPs).

Based on the results obtained from simulated and experimental data, the following conclusions were made regarding the proposed signal enhancement processors:

- Application of HST increased detection performance for all processors evaluated, with the Standard Mean form ($\overline{H}_1$) offering the best overall performance.
- The CCE method does prevent the failure of product-based processors if harmonic components are missing. However, the detection performance is decreased if no harmonics are missing. For the case of summation-based processors, the CCE method lowers detectability regardless if any components are missing or not.
- The Modulo-2π phase adjustment factor provides a significant increase in detection performance for the AVC and SAC processors, with the harmonically transformed A-SAC providing the best performance out of all the PAPs evaluated.
- The combined HST and PAP forms were found to produce significantly better results than either of the forms applied independently. In addition, results were also found to surpass those generated by Wagstaff's PAC and PAV processors.
- The proposed GASC provided a significant performance increase over the GMSC processor for all scenarios evaluated. Results obtained for this processor were also found to be significantly higher than all other processors evaluated including the HST and PAPs.
- Use of the proposed processors produced detection distances adequate to perform an avoidance maneuver for both manned and unmanned aircraft. However, adequate detection distances were not achievable using standard methods such as the incoherent mean.

7.2.3 - Source Detection

In addition to the proposed signal enhancement processors, CFAR detection relationships for unknown signals residing in noise with fixed bandwidth regions and unknown properties were also provided. These included:

1) Establishment of distribution-free CFAR detector relationships for non-independent testing scenarios (CDF-CFAR & SCDF-CFAR).
2) Establishment of a distribution-free CFAR detector through frequency tracking of consecutive windowed spectra (FT-CFAR).
3) Development of a Robust Binary Integration scheme to better facilitate the detection of non-stationary signals.
4) Modifications for CFAR detectors to facilitate accurate usage with harmonically transformed spectra.
5) A CFAR-Enhanced Spectral Whitening technique to facilitate the accurate use of distribution-free CFAR detectors with non-identically distributed noise.

Each of the signal enhancement and detection methods was also validated using computer generated and experimental data. Based on the results obtained, it was concluded that the SCDF-CFAR offered the best overall performance. It performed equally well to the CDF-CFAR detector but required significantly less test cells, which greatly reduced computational requirements. The FT-CFAR detector offered a slightly more simplistic setup since it has only one sample set, uses no order statistics, and requires less maxima tracking points. However, the increase in simplicity and reduction in computational loads is produced at the cost of detection sensitivity. In regard to detection schemes, it was found that the RBI method proved superior in detecting and tracking non-stationary signals. However, it was also found to be less favourable compared to the standard BI method for stationary signals since false alarm rates were increased with no effect on signal detectability.

Finally, an examination of the statistical and kinematic requirements to establish a reliable UAV collision avoidance system was provided. This included a brief analysis to determine minimum required detection rates, and the development of an analytical model to approximate minimum required detection distances.

7.2.4 - Source Localization

A beamforming method was proposed (Harmonic Spectral Beamformer) to enhance the localization accuracy of harmonic continuous source signals via the Steered Response Power (SRP) method. In addition, algorithms were developed to reduce computational loads associated with the SRP localization technique. These included a Crisscross Regional Contraction method, and an adaptive approach with uses the steepest ascent gradient search.

Based on the localization results obtained from experimental data, it was concluded that the proposed Harmonic Spectral Beamformer (HSB-SRP) provided significantly increased localization accuracy compared to the standard SRP form. In addition, it was also shown that by performing signal detection prior to beamforming operations, reduced computational loads and increased localization accuracy can be obtained.

7.3 - Future Work

Although the primary purpose of the proposed research endeavor was achieved, a number of critical areas still require more work to achieve full validation in terms of commercial/industrial viability. These areas include:

1) Optimize enhancement and detection processor parameters to ensure maximum performance.
2) Incorporate trajectory mapping scheme to better facilitate avoidance maneuvering.
3) Develop a dedicated and application specific recording system to reduce payload weight and size requirements. Doing so would also reduce UAV size requirements, which in turn would reduce self-noise levels and increase source detectability.
4) Establish real-time operation using a dedicated system to verify that the outlined processing methods can be implemented in such a manner using currently available DSP hardware.
5) Establish frequency domain beamforming techniques to account for wave transient effects associated with a moving source and moving observer.
6) Conduct additional experiments under various weather conditions to determine the effect of environmental factors such as wind, rain, temperature, etc.
7) Conduct experiments using various aircraft types with different spectral signatures to verify detection capabilities.
8) Develop directional response microphones which are capable of operating under high air flow conditions.

7.4 - Final Remarks

As previously noted throughout this thesis, data analysis was performed offline using the MATLAB software suite. However, a successful collision avoidance system must be capable of operating in real-time with an extremely low latency. Graphics Processing Units (GPUs) such as the Nvidia CUDA architecture may be utilized to achieve such operational characteristics. GPUs generally offer superior computing performance over traditional CPUs for DSP applications, since many of the required operations can be performed in parallel. Indeed, such systems have been shown to offer real-time operation for many audio processing applications [282, 283].

Another area not currently addressed is the classification of target aircraft once detection is achieved. For example, it was previously illustrated in Figure 2-7 how the spectral signature for fixed-wing and multi-rotor UAVs differ vastly. Indeed, a similar situation is present for the case of fixed-wing and rotary-wing manned aircraft as illustrated in Figure 2-6 . Using this knowledge, one may be able to determine which type of aircraft is present in the local vicinity. Work has been conducted in this area for low flying aircraft using ground-based arrays [34, 284]. However, these methods attempt to classify the aircraft through direct feature extraction and pattern matching of the recorded aircraft spectra. More robust methods which utilize the power of artificial neural networks may offer much greater flexibility and accuracy in this domain [285]. In this regard,

developments made throughout this thesis may form a solid foundation for counter UAV operations. That is, the detection, localization, and classification of UAVs operating near undesired areas such as airports, which create a very high safety risk. Since UAVs are often very small and fly at relatively low altitudes, standard airport detection systems such as radar are not capable of efficiently or effectively offering detection and/or classification in this regard. In theory, a network of microphones could be located around the boundary of such a space to detect the presence of an intruding UAV. The microphones may act together essentially forming a giant array with real-time operation being facilitated by GPU processing techniques as previously discussed. Such a system would be extremely useful since it can operate in all weather conditions and would be very low cost compared to other technologies such as distributed optical or micro-radar systems.

In closing, it does appear that acoustic sensing may constitute some basis to establish a non-cooperative collision avoidance system for UAVs. Detection distances were found to be adequate to perform an avoidance maneuver for all experiments presented. However, the obvious limitation to the technology involves the sound level of the intruding aircraft. All aircraft evaluated utilized some form of internal combustion engine which is inherently loud according to human hearing standards. Although this is the case for essentially all manned aircraft, most UAVs operate using electric power instead, and are thus much quieter. For such scenarios, acoustic sensing may not produce such favorable results. It is believed that ultimately, a robust and reliable collision avoidance system will not only depend on once sensing technology; but rather multiple technologies such as those previously outlined in Section 2.2.2 acting together in tandem.

- 8 - References

[1] E. B. Carr, "Unmanned aerial vehicles: Examining the safety, security, privacy and regulatory issues of integration into U.S. airspace." National Center for Policy Analysis, 2013.

[2] C. Geyer, S. Singh and L. Chamberlain, "Avoiding collisions between aircraft: State of the art and requirements for UAVs operating in civilian airspace." Robotics Institute, Carnegie Mellon University: Pittsburgh, PA, USA, Tech. Rep. CMU-RI-TR-08-03, 2008.

[3] M. Contarino, "All weather sense and avoid system for UASs." R3 Engineering, Tech. Rep. Report to the Office of Naval Research, 2009.

[4] A. Finn and S. Franklin. Acoustic sense & avoid for UAV's. Presented at Seventh International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP). 2011.

[5] (2012). *Unmanned Aerial Vehicle (UAV) Working Group Final Report*. Available: https://www.tc.gc.ca/eng/civilaviation/standards/general-recavi-uav-2266.html.

[6] M. Huang, R. M. Narayanan, Y. Zhang and A. Feinberg, "Tracking of Noncooperative Airborne Targets Using ADS-B Signal and Radar Sensing," *International Journal of Aerospace Engineering*, vol. 2013, pp. 12, 2013.

[7] Garmin Ltd. (2014). *ADS-B Creates a New Standard of Aviation Safety*. Available: http://www8.garmin.com/aviation/adsb.html.

[8] S. B. Hottman, K. R. Hansen and M. Berry, "Literature review on detect, sense, and avoid technology for unmanned aircraft systems," Federal Aviation Administration, Washington, DC, Tech. Rep. DOT/FAA/AR-08/41, 2009.

[9] M. Shah, H. Asad and A. Basharat, "Detection and Tracking of Objects From Multiple Airborne Cameras," *The International Society of Optical Engineering*, 2006.

[10] N. Franceschini, J. M. Pichon, C. Blanes and J. M. Brady, "From insect vision to robot vision," *Philosophical Transactions: Biological Sciences*, vol. 337, pp. 283–294, 1992.

[11] J. S. Humbert and M. A. Frye, "Extracting behaviorally relevant retinal image motion cues via wide-field integration," in *American Control Conference*, 2006.

[12] J. Serres, D. Dray, F. Ruffier and N. Franceschini, "A vision-based autopilot for a miniature air vehicle: joint speed control and lateral obstacle avoidance," *Autonomous Robotics*, vol. 25, pp. 103-122, 2008.

[13] A. Beyeler, J. C. Zufferey and D. Floreano, "Optipilot: Control of take-off and landing using optic flow," in *European Conference and Competitions on Micro Air Vehicles (EMAV)*, 2009.

[14] R. K. Mehra, J. Byrne and J. Boskovic, "Flight testing of a fault-tolerant control and vision-based obstacle avoidance system for uavs," in *Association for Unmanned Vehicle Systems International (AUVSI) Conference*, 2005.

[15] J. McCandless, "Detection of aircraft in video sequences using a predictive optical flow algorithm," *Optical Engineering*, vol. 3, pp. 523-530, 1999.

[16] D. J. Lee, R. W. Beard, P. C. Merrell and P. Zhan, "See and Avoidance Behaviors for Autonomous Navigation," *SPIE Optics East, Robotics Technologies and Architectures, Mobile Robot XVII*, vol. 5609-05, pp. 25-28, 2004.

[17] T. Netter and N. Franceschini, "Neuromorphic Motion Detection for Robotic Flight Guidance," *The Neuromorphic Engineer*, 2004.

[18] K. Nordberg, P. Doherty, G. Farneback, P. Erick-Forssén, G. Granlund, A. Moe and J. Wiklund, "Vision for a UAV helicopter," in *Proceedings of IROS'02, Workshop on Aerial Robotics,* 2002.

[19] J. Keller. (2013). *Electro-optical sensor payloads for small UAVs*. Available: http://www.militaryaerospace.com/articles/print/volume-24/issue-10/technology-focus/electro-optical-sensor-payloads-for-small-uavs.html.

[20] D. Soreide, W. Tank and J. Osborne, "Development of an optical sense and avoid guidance and control system with staring infrared cameras," in *AIAA's 1st Technical Conference and Workshop on Unmanned Aerospace Vehicles,* 2002.

[21] R. Bernier, M. Bissonnette and P. Poitevin, "DSA radar – development report," in *Auvsi,* Baltimore, USA, 2005.

[22] Y. Li. Frequency-modulated continuous-wave synthetic-aperture radar: Improvements in signal processing. 2016.

[23] Sandia National Laboratories. (2005). *Synthetic Aperture Radar Applications*. Available: http://www.sandia.gov/RADAR/sarapps.html.

[24] The Society of British Aerospace Companies. (2006). *LOAM® Laser Obstacle Warning System to be Equipped on Denmark's EH101 Search and Rescue Helicopters*. Available: http://www.sbac.co.uk/community/cms/content/preview/news_item_view.asp?i=14081.

[25] D. H. Shim, H. Chung, H. J. Kim and S. Sastry, "Autonomous exploration in unknown urban environments for unmanned aerial vehicles," in *AIAA GNC Conference,* San Francisco, 2005.

[26] S. Scherer, S. Singh, L. Chamberlain and S. Saripalli, "Flying fast and low among obstacles," in *International Conference on Robotics and Automation,* 2007.

[27] A. Bachrach, R. He and N. Roy, "Autonomous flight in unstructured and unknown indoor environments," in *European Micro Air Vehicle Conference and Competititons (EMAV),* Netherlands, 2009.

[28] S. Grzonka, G. Grisetti and W. Burgard, "Towards a navigation system for autonomous indoor flying," in *IEEE International Conference on Robotics and Automation (ICRA),* Kobe, Japan, 2009.

[29] S. B. Hottman, K. R. Hansen and M. Berry, "Literature review on detect, sense, and avoid technology for unmanned aircraft systems," U.S. Department of Transportation Federal Aviation Administration, Washington, Tech. Rep. DOT/FAA/AR-08/41, 2009.

[30] S. Saripalli, J. F. Montgomery and G. S. Sukhatme, "Visually-Guided Landing of an Unmanned Aerial Vehicle," *IEEE Transactions on Robotics and Automation,* vol. 19, pp. 371-381, 2003.

[31] A. M. Zelnio, E. E. Case and B. D. Rigling. A low-cost acoustic array for detecting and tracking small RC aircraft. Presented at 2009 IEEE 13th Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop. 2009.

[32] E. E. Case, A. M. Zelnio and B. D. Rigling. Low-cost acoustic array for small UAV detection and tracking. Presented at 2008 IEEE National Aerospace and Electronics Conference. 2008.

[33] A. M. Zelnio, "Detection of Small Aircraft Using an Acoustic Array," 2009.

[34] A. Sutin, H. Salloum, A. Sedunov and N. Sedunov. Acoustic detection, tracking and classification of low flying aircraft. Presented at 2013 IEEE International Conference on Technologies for Homeland Security (HST). 2013.

[35] H. Salloum, A. Sedunov, N. Sedunov, A. Sutin and D. Masters. Acoustic system for low flying aircraft detection. Presented at 2015 IEEE International Symposium on Technologies for Homeland Security (HST). 2015.

[36] R. O. Nielsen, "Acoustic detection of low flying aircraft," in *IEEE Conference on Technologies for Homeland Security,* 2009.

[37] B. G. Ferguson. A ground-based narrow-band passive acoustic technique for estimating the altitude and speed of a propeller-driven aircraft. *J. Acoust. Soc. Am. 92(3),* pp. 1403-1407. 1992.

[38] Sadasivan, S., Gurubasavaraj, M., Sekar,S. Acoustic signature of an unmanned air vehicle exploitation for aircraft localisation and parameter estimation. *Def. Sci. J. 51(3),* 2002.

[39] J. Tong, W. Xie, Y. Hu, M. Bao, X. Li and W. He. Estimation of low-altitude moving target trajectory using single acoustic array. *J. Acoust. Soc. Am. 139(4),* pp. 1848-1858. 2016.

[40] B. G. Ferguson and K. W. Lo, "Turbo-prop and rotary-wing aircraft flight parameter estimation using both narrow-band and broadband passive acoustic signal processing methods," *J. Acoustic Society America,* vol. 108, pp. 1764, 2000.

[41] J. Wind, H. E. Bree and B. Xu, "3D sound source localization and sound mapping using a PU sensor array." in *16th AIAA/CEAS Aeroacoustics Conference,* 2010.

[42] H. E. de Bree, "Acoustic vector sensors for passive 3D trajectory monitoring of rotary wing aircraft," *European Rotorcraft Forum (ERF),* 2012.

[43] H. E. de Bree, J. Wind and S. Sadasivan, "Broad banded acoustic vector sensors for outdoor monitoring propeller driven aircraft," in *German Annual Conference on Acoustics (DAGA),* Germany, 2010.

[44] H. E. de Bree, J. Wind and P. Theije, "Detection, localization and tracking of aircraft using acoustic vector sensors," in *40th International Congress and Exposition on Noise Control Engineering (INTER-NOISE 2011),* Osaka, Japan, pp. 1, 2011.

[45] C. Reiff, T. Pham, M. Scanlon, J. Noble, A. V. Landuyt, J. Petek and J. Ratches. Acoustic detection from an aerial balloon platform. US Army Research Laboratory. 2004.

[46] K. W. Lo and B. G. Ferguson. Tactical unmanned aerial vehicle localization using ground-based acoustic sensors. Presented at Proceedings of the 2004 Intelligent Sensors, Sensor Networks, and Information Processing Conference. 2004.

[47] Jianfei Tong, Yu-Hen Hu, Ming Bao and Wei Xie. Target tracking using acoustic signatures of light-weight aircraft propeller noise. Presented at IEEE China Summit & International Conference on Signal and Information Processing (ChinaSIP). 2013.

[48] T. Pham and N. Srour. TTCP AG-6: Acoustic detection and tracking of UAVs [5417-06]. Presented at Proceedings of the International Society for Optical Engineering. 2004.

[49] R. A. Silva, "Numerical Simulation and Laboratory Testing of Time-Frequency MUSIC Beamforming for Identifying Continuous and Impulsive Ground Targets from a Mobile Aerial Platform," *Texas A&M University,* 2013.

[50] E. Tijs, D. Yntema, J. Wind and H. E. de Bree, "Acoustic vector sensors for aeroacoustics," in *CEAS Buchares,* 2009.

[51] E. H. G. Tijs, G. C. H. E. de Croon, J. W. Wind, B. Remes, C. De Wagter, H. E. de Bree and R. Ruijsink, "Hear-and-avoid for micro air vehicles," in *International Micro Air Vehicle Conference and Competitions (IMAV) ,* 2010.

[52] B. Ferguson and R. Wyber, "Detection and localization of a ground based impulsive sound source using acoustic sensors onboard a tactical unmanned aerial vehicle," in *Battlefield Acoustic Sensing for ISR Applications,* Neuilly-sur-Seine, France, pp. 16-1, 2006.

[53] D. N. Robertson, T. Pham, H. Edge, B. Porter, J. Shumaker and D. Cline, "Acoustic sensing from small-size UAVs," in *Proc. SPIE 6562 Unattended Ground, Sea, and Air Sensor Technologies and Applications IX,* Orlando, USA, pp. 656208, 2007.

[54] T. Ohata, K. Nakamura, T. Mizumoto, T. Taiki and K. Nakadai. Improvement in outdoor sound source detection using a quadrotor-embedded microphone array. Presented at IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2014.

[55] K. Okutani, T. Yoshida, K. Nakamura and K. Nakadai, "Outdoor auditory scene analysis using a moving microphone array embedded in a quadrocopter." in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).* Vilamoura, Portugal., pp. 3288, 2012.

[56] B. Harvey and S. O'Young, "Detection of continuous ground-based acoustic sources via unmanned aerial vehicles," *Journal of Unmanned Vehicle Systems*, vol. 4, pp. 83, 2015.

[57] Scientific Applications and Research Associates Inc.(SARA). (2012). *UAV Acoustic Collision-Alert System*. Available: http://www.sara.com/ISR/UAV_payloads/PANCAS.html.

[58] T. Milkie, "Passive Acoustic Non-Cooperative Collision Alert System (PANCAS) for UAV Sense and Avoid," *Unpublished White Paper by SARA, Inc.*, 2007.

[59] C. H. Hansen, B. Goelzer and G. A. Sehrndt. "Fundamentals of acoustics," in *Occupational Exposure to Noise: Evaluation, Prevention and Control.*Anonymous 2013, Available: http://www.who.int/occupational_health/publications/noise1.pdf.

[60] B. D.A. and H. C.H., *Engineering Noise Control: Theory and Practice.* New York: Spon Press, 2009.

[61] S. W. Rienstra and A. Hirschberg, "An Introduction to Acoustics," *Eindhoven University of Technology*, 2004.

[62] A. P. Dowling and J. E. Williams, *Sound and Sources of Sound.* Ellis Horwood Series in Engineering Science, 1983.

[63] W. L. Willshire and D. Chestnutt. Joint acoustic propagation experiment (JAPE-91). Presented at Workshop Proceedings of a Workshop Jointly Sponsored by the National Aeronautics and Space Administration. 1993.

[64] T. R.D., "Springer handbook of acoustics," in , T. R.D., Ed. New York: Springer New York, 2007, pp. 113.

[65] E. E. Bass, L. C. Sutherland and A. J. Zuckerwar, "Atmospheric absorption of sound: Update," *J. Acoust. Soc. Am.*, vol. 88, pp. 2019, 1990.

[66] K. W. Lo, B. G. Ferguson, Yujin Gao and A. Maguer. Aircraft flight parameter estimation using acoustic multipath delays. *IEEE Transactions on Aerospace and Electronic Systems 39(1)*, pp. 259-268. 2003.

[67] K. W. Lo, S. W. Perry and B. G. Ferguson. Aircraft flight parameter estimation using acoustical lloyd's mirror effect. *IEEE Transactions on Aerospace and Electronic Systems 38(1)*, pp. 137-151. 2002.

[68] H. Lord, W. S. Gatley and H. A. Evensen, *Noise Control for Engineers.* New York: McGraw-Hill, 1980.

[69] E. H. Brown and F. F. Hall. Advances in atmospheric acoustics. *Rev. Geophys. 16(1)*, pp. 47-110. 1978.

[70] V. Ostashev and D. W. Wilson, *Acoustics in Moving Inhomogeneous Media.* Boca Raton, FL, USA: CRC Press, 2016.

[71] G. Brooker, "Sensors and Signals," *Australian Centre for Field Robotics*, 2007.

[72] J. E. Marte and D. W. Kurtz, "A review of aerodynamic noise from propellers, rotors, and lift fans," NASA, Jet Propulsion Laboratory, California, Tech. Rep. 32-1462, 1970.

[73] U. Michel, W. Dobrzynski, W. Splettstoesser, J. Delfs, U. Isermann and F. Obermeier, "Aircraft noise," in *Handbook of Engineering Acoustics*, 1st ed., M. M. Gerhard Muller, Ed. New York: Springer Heidelberg, 2013, pp. 489.

[74] H. C. True and E. J. Rickley, "Noise characteristics of eight helicopters," U.S. Department of Transportation, Springfield, Virginia, Tech. Rep. FAA-RD-77-94, 1977.

[75] J. S. Bendat and A. G. Piersol, *Engineering Applications of Correlation and Spectral Analysis, 2nd Edition.* Wiley, 1993.

[76] R. G. Lyons, *Understanding Digital Signal Processing.* Upper Saddle River, NJ, USA: Prentice Hall, 2010.

[77] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables.* New York: Springer, 2006.

[78] J. B. Tsui, *Digital Techniques for Wideband Receivers.* SciTech Publishing, 2004.

[79] M. A. Richards, *Fundamentals of Radar Signal Processing.* McGraw-Hill, 2005.

[80] M. Nakagami, "The m-distribution: A general formula of intensity distribution of rapid fading," in *Statistical Methods in Radio Wave Propagation*, W. Hoffman, Ed. Oxford: Pergamon Press, 1960, pp. 3-36.

[81] A. Poularikas and Z. Ramadan, *Adaptive Filtering Primer with MATLAB.* CRC Press, 2006.

[82] M. Rhudy, B. Bucci, J. Vipperman, J. Allanach and B. Abraham, "Microphone array analysis methods using cross-correlations," in *ASME 2009 International Mechanical Engineering Congress and Exposition Volume 15: Sound, Vibration and Design,* Lake Buena Vista, Florida, USA, pp. 281, 2009.

[83] S. M. Boker, M. Xu, J. L. Rotondo and K. King, "Windowed cross-correlation and peak picking for the analysis of variability in the association between behavioral time series," *Psychol Methods,* vol. 7, pp. 338, 2002.

[84] P. Dodwell, "On the perceptual clarity," *Psychological Review,* pp. 275, 1971.

[85] L. Glass and E. Switkes, "Pattern recognition in humans: Correlations which cannot be perceived," *Perception,* vol. 5, pp. 67, 1976.

[86] P. Dixon and V. Di Lolo, "Beyond visible persistence: An alternative account of temporal integration and segregation in visual processing," *Cognitive Psychology,* vol. 26, pp. 33, 1994.

[87] R. M. Hennig, "Acoustic feature extraction by cross-correlation in crickets?" *J Comp Physiol A,* vol. 189, pp. 589, 2003.

[88] N. Kottege and U. R. Zimmer, "Cross-correlation tracking for maximum length sequence based acoustic localisation," in *Australasian Conference on Robotics and Automation 2008 (ACRA '08),* Canberra, ACT, Australia, pp. 1, 2008.

[89] J. M. Perez-Lorenzo, R. Viciana-Abad, P. Reche-Lopez, F. Rivas and J. Escolano, "Evaluation of generalized cross-correlation methods for direction of arrival estimation using two microphones in real environments," *Applied Acoustics,* vol. 73, pp. 698, 2012.

[90] L. Tan and J. Jiang, "Novel adaptive IIR filter for frequency estimation and tracking [DSP Tips&Tricks]," *IEEE Signal Processing Magazine,* vol. 26, pp. 186-189, 2009.

[91] L. Tan and J. Jiang, "Real-Time Frequency Tracking Using Novel Adaptive Harmonic IIR Notch Filter," *The Technology Interface Journal,* vol. 9, pp. 1, 2009.

[92] L. Tan and J. Jiang, "Simplified Gradient Adaptive Harmonic IIR Notch Filter for Frequency Estimation and Tracking," *American Journal of Signal Processing,* vol. 5, pp. 6-12, 2015.

[93] L. Tan, J. Jiang and L. Wang, "Adaptive harmonic IIR notch filters for frequency estimation and tracking," in *Adaptive Filtering,* L. Garcia, Ed. Rijeka, Croatia: InTech, 2011, pp. 313.

[94] B. Widrow, J. R. Glover, J. M. McCool, J. Kaunitz, C. S. Williams, R. H. Hearn, J. R. Zeidler, J. Eugene Dong and R. C. Goodlin, "Adaptive noise cancelling: Principles and applications," in *Proceedings of the IEEE,* pp. 1692-1716, 1975.

[95] O. L. Frost, "An algorithm for linearly constrained adaptive array processing," *Proceedings of the IEEE,* vol. 60, pp. 926-935, 1972.

[96] M. M. Dewasthale and R. D. Kharadkar, "Acoustic noise cancellation using adaptive filters: A survey," in *International Conference on Electronic Systems, Signal Processing and Computing Technologies (ICESC).* pp. 12-16, 2014.

[97] P. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation.* Springer, 2013.

[98] L. Vega and H. Rey, *A Rapid Introduction to Adaptive Filtering.* Springer, 2013.

[99] J. Treichler. Transient and convergent behavior of the adaptive line enhancer. *IEEE Transactions on Acoustics, Speech, and Signal Processing 27(1),* pp. 53-62. 1979.

[100] R. M. Ramli, A. O. Noor and S. A. Samad, "A review of adaptive line enhancers for noise cancellation," *Australian Journal of Basic and Applied Sciences,* vol. 6, pp. 337-352, 2012.

[101] I. S. Badreldin, D. S. El-Kholy and A. A. El-Wakil, "A modified adaptive noise canceler for electrocardiography with no power-line reference," in *5th Cairo International Biomedical Engineering Conference (CIBEC).* pp. 13-16, 2010.

[102] J. W. Kelly, J. L. Collinger, A. D. Degenhart, D. P. Siewiorek, A. Smailagic and W. Wang, "Frequency tracking and variable bandwidth for line noise filtering without a reference," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).* pp. 7908-7911, 2011.

[103] I. S. Badreldin, D. S. El-Kholy and A. A. Elwakil, "Harmonic adaptive noise canceler for electrocardiography with no power-line reference," in *16th IEEE Mediterranean Electrotechnical Conference (MELECON).* pp. 1017-1020, 2012.

[104] J. W. Kelly, D. P. Siewiorek, A. Smailagic and Wei Wang. An adaptive filter for the removal of drifting sinusoidal noise without a reference. *IEEE Journal of Biomedical and Health Informatics 20(1),* pp. 213-221. 2016.

[105] A. Lopez, J. C. Montano, M. Castilla, J. Gutierrez, M. D. Borras and J. C. Bravo. Power system frequency measurement under nonstationary situations. *IEEE Transactions on Power Delivery 23(2),* pp. 562-567. 2008.

[106] B. Boashash. Estimating and interpreting the instantaneous frequency of a signal I. fundamentals. *Proceedings of the IEEE 80(4),* pp. 520-538. 1992.

[107] B. Boashash. Estimating and interpreting the instantaneous frequency of a signal II. algorithms and applications. *Proceedings of the IEEE 80(4),* pp. 540-568. 1992.

[108] Y. C. Chen and H. Y. Shen, "Fundamental Frequency Analysis on A Harmonic Power Signal Using Fourier Series and Zero Crossing Algorithms," *Journal of Information Hiding and Multimedia Signal Processing,* vol. 6, pp. 924-937, 2015.

[109] D. Gerhard, "Pitch extraction and fundamental frequency: History and current techniques," Department of Computer Science, University of Regina, Tech. Rep. TR-CS 2003-06, 2003.

[110] B. Boashash, G. Jones and P. O'Shea. Instantaneous frequency of signals: Concepts, estimation techniques and applications. Presented at Proceedings of SPIE, Advanced Algorithms and Architectures for Signal Processing IV. 1989.

[111] J. Kormylo and V. Jain. Two-pass recursive digital filter with zero phase shift. *IEEE Transactions on Acoustics, Speech, and Signal Processing 22(5),* pp. 384-387. 1974.

[112] E. Arias-de-Reyna and A. I. José, "A new method for designing efficient linear phase recursive filters," *Digital Signal Processing,* vol. 14, pp. 1-17, 2004.

[113] S. R. Powell and P. M. Chau. A technique for realizing linear phase IIR filters. *IEEE Transactions on Signal Processing 39(11),* pp. 2425-2435. 1991.

[114] A. N. Willson and H. J. Orchard. An improvement to the powell and chau linear phase IIR filters. *IEEE Transactions on Signal Processing 42(10),* pp. 2842-2848. 1994.

[115] C. Tsai and A. T. Fam, "Efficient linear phase filters based on switching and time reversal," in *IEEE International Symposium on Circuits and Systems.* pp. 2161-2164, 1990.

[116] A. Mouffak and M. F. Beleachir, "Noncausal forward/backward two-pass IIR digital filters in real time," *Turkish Journal of Electrical Engineering & Computer Sciences.,* vol. 20, pp. 769-786, 2012.

[117] R. Czarnach. Recursive processing by noncausal digital filters. *IEEE Transactions on Acoustics, Speech, and Signal Processing 30(3),* pp. 363-370. 1982.

[118] B. Djokic, M. Popovic and M. Lutovac. A new improvement to the powell and chau linear phase IIR filters. *IEEE Transactions on Signal Processing 46(6),* pp. 1685-1688. 1998.

[119] S. J. Marple, "A fast least squares linear phase adaptive filter," in *IEEE International Conference on Acoustics, Speech, and Signal Processing.* pp. 219-222, 1984.

[120] P. F. Titchener, R. P. Gooch and B. Widrow, "A linear phase adaptive filter," in *Record of the Sixteenth Asilomar Conference on Circuits, Systems and Computers,* pp. 40-44, 1982.

[121] H. K. Kwan and Q. P. Li. High-speed realisation of adaptive linear phase FIR digital filters. *IEEE Proceedings of Radar and Signal Processing 140(1),* pp. 48-54. 1993.

[122] B. Friedlander and M. Morf, "Least-squares algorithms for adaptive linear-phase filtering," in *IEEE International Conference on Acoustics, Speech, and Signal Processing.* pp. 247-250, 1981.

[123] N. Kalouptsidis and G. Koyas. Efficient block LS design of FIR filters with linear phase. *IEEE Transactions on Acoustics, Speech, and Signal Processing 33(6),* pp. 1435-1444. 1985.

[124] N. Kalouptsidis and S. Theodoridis. Efficient structurally symmetric algorithms for least squares FIR filters with linear phase. *IEEE Transactions on Acoustics, Speech, and Signal Processing 36(9),* pp. 1454-1465. 1988.

[125] D. Youn and S. Prakash, "On realizations and related algorithms for adaptive linear phase filtering," in *IEEE International Conference on Acoustics, Speech, and Signal Processing.* pp. 134-137, 1984.

[126] R. Arablouei, K. Doğançay and S. Werner. On the mean-square performance of the constrained LMS algorithm. *Signal Processing 117*pp. 192-197. 2015.

[127] J. O. Smith, *Introduction to Digital Filters with Audio Applications.* W3K Publishing, 2007.

[128] P. A. Lynn and W. Fuerst, *Introductory Digital Signal Processing with Computer Applications.* New York: John Wiley & Sons, 1989.

[129] T. Yu, S. Mitra and H. Babic, "Design of linear phase FIR notch filters," *Sadhana,* vol. 15, pp. 133, 1990.

[130] S. C. Dutta Roy, B. Kumar and S. B. Jain, "FIR Notch Filter Design - A Review," *Electronics and Energetics,* vol. 14, pp. 295, 2001.

[131] S. C. Dutta Roy, S. B. Jain and B. Kumar, "Design of digital FIR notch filters from second order IIR prototype," *IETE Journal of Research,* vol. 43, pp. 275, 1997.

[132] S. Kocon and J. Piskorowski. A concept of time-varying FIR notch filter based on IIR filter prototype. Presented at 2012 17th International Conference on Methods & Models in Automation & Robotics (MMAR). 2012.

[133] R. Deshpande, B. Kumar and S. B. Jain. On the design of multi notch filters. *International Journal of Circuit Theory and Applications 40(4),* pp. 313-327. 2012.

[134] J. Williams and G. Ricker. Signal detectability performance of optimum fourier receivers. *IEEE Transactions on Audio and Electroacoustics 20(4),* pp. 264-270. 1972.

[135] K. N. Stevens, "Acoustic Properties used for the Identification of Speech Sounds," *Annals. N. Y. Acad. Sci.,* vol. 405, pp. 2-17, 1983.

[136] S. McAdams, P. Depalle and E. Clarke, "Analyzing musical sound," in *Empirical Musicology: Aims, Methods, Prospects,* E. Clarke, Ed. New York, NY: Oxford University Press, 2004, pp. 157-196.

[137] F. Farassat and K. B. Brentner, "The Acoustic Analogy and the Prediction of the Noise of Rotating Blades," *Theoret. Comput. Fluid Dynamics,* vol. 10, pp. 155-170, 1998.

[138] R. B. terMeulen, "Notes on Acoustics for Undergraduates," 2008.

[139] N. Yang, H. Ba, W. Cai, I. Demirkol and W. Heinzelman. BaNa: A noise resilient fundamental frequency detection algorithm for speech and music. *IEEE/ACM Transactions on Audio, Speech, and Language Processing 22(12),* pp. 1833-1848. 2014.

[140] Y. Doweck, A. Amar and I. Cohen. Joint model order selection and parameter estimation of chirps with harmonic components. *IEEE Transactions on Signal Processing 63(7)*, pp. 1765-1778. 2015.

[141] S. Gonzalez and M. Brookes. PEFAC - A pitch estimation algorithm robust to high levels of noise. *IEEE/ACM Transactions on Audio, Speech, and Language Processing 22(2)*, pp. 518-530. 2014.

[142] W. Chu and A. Alwan. SAFE: A statistical approach to F0 estimation under clean and noisy conditions. *IEEE Transactions on Audio, Speech, and Language Processing 20(3)*, pp. 933-944. 2012.

[143] Mingyang Wu, DeLiang Wang and G. J. Brown. A multipitch tracking algorithm for noisy speech. *IEEE Transactions on Speech and Audio Processing 11(3)*, pp. 229-241. 2003.

[144] A. Noll, "Pitch Determination of Human Speech by the Harmonic Product Spectrum, the Harmonic Sum Spectrum and a Maximum Likelihood Estimate," *Proceedings of the Symposium on Computer Processing in Communications*, vol. XIX, pp. 779-797, 1969.

[145] J. Dubnowski, R. Schafer and L. Rabiner. Real-time digital hardware pitch detector. *IEEE Transactions on Acoustics, Speech, and Signal Processing 24(1)*, pp. 2-8. 1976.

[146] J. Markel. The SIFT algorithm for fundamental frequency estimation. *IEEE Transactions on Audio and Electroacoustics 20(5)*, pp. 367-377. 1972.

[147] N. Miller. Pitch detection by data reduction. *IEEE Transactions on Acoustics, Speech, and Signal Processing 23(1)*, pp. 72-79. 1975.

[148] A. Rosenberg and M. Sambur. New techniques for automatic speaker verification. *IEEE Transactions on Acoustics, Speech, and Signal Processing 23(2)*, pp. 169-176. 1975.

[149] P. Boersma, "Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound," *Institute of Phonetic Sciences Proceedings*, vol. 17, pp. 97-110, 1993.

[150] M. Ross, H. Shaffer, A. Cohen, R. Freudberg and H. Manley. Average magnitude difference function pitch extractor. *IEEE Transactions on Acoustics, Speech, and Signal Processing 22(5)*, pp. 353-362. 1974.

[151] M. R. Schroeder, "Period Histogram and Product Spectrum: New Methods for Fundamental-Frequency Measurement," *The Journal of the Acoustical Society of America*, vol. 43, pp. 829, 1968.

[152] M. Hinich. Detecting a hidden periodic signal when its period is unknown. *IEEE Transactions on Acoustics, Speech, and Signal Processing 30(5)*, pp. 747-750. 1982.

[153] G. Planquette, C. Le Martret and G. Vezzosi. Detecting and estimating the fundamental of harmonics series when the number of harmonics is unknown. Presented at Military Communications Conference, 1995. MILCOM '95, Conference Record, IEEE. 1995.

[154] R. W. Schafer and L. W. Rabiner, "System for automatic formant analysis of voiced speech," *J. Acoust. Soc. Amer.*, vol. 47, pp. 634-648, 1970.

[155] D. G. Childers, D. P. Skinner and R. C. Kemerait. The cepstrum: A guide to processing. *Proceedings of the IEEE 65(10)*, pp. 1428-1443. 1977.

[156] X. Sun. Pitch determination and voice quality analysis using subharmonic-to-harmonic ratio. Presented at Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference On. 2002.

[157] X. Chen and R. Liu. "Multiple pitch estimation based on modified harmonic product spectrum," in *Proceedings of the 2012 International Conference on Information Technology and Software Engineering: Information Technology & Computing Intelligence*, W. Lu, G. Cai, W. Liu and W. Xing, Eds. 2013, .

[158] P. Martin. Comparison of pitch detection by cepstrum and spectral comb analysis. Presented at Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '82. 1982.

[159] A. S. Master, "Speech spectrum modeling from multiple sources," *Cambridge University Department of Engineering*, 2000.

[160] H. Ding, B. Qian, Y. Li and Z. Tang. A method combining LPC-based cepstrum and harmonic product spectrum for pitch detection. Presented at 2006 International Conference on Intelligent Information Hiding and Multimedia. 2006.

[161] C. M. Harris and M. R. Weiss, "Pitch Extraction by Computer Processing of High Resolution Fourier Analysis Data," *J. Acoust. Soc. Am.,* vol. 35, pp. 339-343, 1963.

[162] S. Venugopal, R. A. Wagstaff and J. P. Sharma, "Exploiting Phase Fluctuations to Improve Machine Performance Monitoring," *IEEE Transactions on Automation Science and Engineering,* vol. 4, pp. 153, 2007.

[163] R. A. Wagstaff, "Exploiting Phase Fluctuations to Improve Temporal Coherence," *IEEE Journal of Oceanic Engineering,* vol. 29, pp. 498, 2004.

[164] R. A. Wagstaff, "Phase coherence adaptive processor for automatic signal detection and identification," in *Proc. SPIE 6217, Detection and Remediation Technologies for Mines and Minelike Targets XI,* Orlando, USA, pp. 62171J-1, 2006.

[165] R. A. Wagstaff, "Signal processor for acoustic sensors on UAV platforms and ground vehicles," in *Proceedings of SPIE Vol. 6546, Airborne Intelligence, Surveillance, Reconnaissance (ISR) Systems and Applications IV.* pp. 654603-1, 2007.

[166] R. A. Wagstaff and H. E. Rice, "Improving temporal coherence to enhance gain and detection performance," in *Proc. of SPIE Vol. 6963, Unattended Ground, Sea, and Air Sensor Technologies and Applications X,* pp. 69630T-1, 2008.

[167] R. A. Wagstaff, "The AWSUM Filter: A 20-dB Gain Fluctuation-Based Processor," *IEEE Journal of Oceanic Engineering,* vol. 22, pp. 110, 1997.

[168] W. Q., W. C. and S. G. A coherent sinusoidal detector using phase compensation. Presented at Oceans '02 Mts/Ieee. 2002.

[169] N. Zhou. A cross-coherence method for detecting oscillations. *IEEE Transactions on Power Systems 31(1),* pp. 623-631. 2016.

[170] G. Shu and X. Liang, "Identification of complex diesel engine noise sources based on coherent power spectrum analysis," *Mechanical Systems and Signal Processing,* vol. 21, pp. 405-416, 1, 2007.

[171] W. G. Halvorsen and J. S. Bendat, "Noise Source Identification Using Coherent Output Power Spectra," *Sound and Vibration,* vol. 9, pp. 15-24, 1975.

[172] A. G. Piersol, "Use of Coherence and Phase Data Between Two Recievers in Evaluation of Noise Environments," *Journal of Sound and Vibration,* vol. 56, pp. 215-228, 1978.

[173] T. W. Parsons. Separation of speech from interfering speech by means of harmonic selection. *J. Acoust. Soc. Am. 60(4),* pp. 911-918. 1976.

[174] A. Camacho and J. G. Harris. A sawtooth waveform inspired pitch estimator for speech and music. *J. Acoust. Soc. Am. 124(3),* pp. 1638-1652. 2008.

[175] K. Wu, D. Zhang and G. Lu, "iPEEH: Improving pitch estimation by enhancing harmonics," *Expert Systems with Applications,* vol. 64, pp. 317-329, 2016.

[176] S. J.O., "Introduction to signal processing," in Anonymous Englewood Cliffs, NJ: Prentice-Hall, 1996, .

[177] R. A. Wagstaff, "The Wagstaff's integration silencing processor filter: A method for exploiting fluctuations to achieve improved sonar signal processor performance," *J. Acoust. Soc. Am.,* vol. 104, pp. 2915-2924, 1998.

[178] R. A. Wagstaff, A. E. Leybourne and J. George, "Von WISPR family of processors: Vol. 1," Naval Research Laboratory, Stennis Space Center, Tech. Rep. NRL/FR/7176- 96-9650, 1997.

[179] R. A. Wagstaff and J. George. Phase variations in a fluctuation-based processor. *Proc.SPIE 2751*pp. 132-141. 1996.

[180] P. Stoica and M. Randolph, *Spectral Analysis of Signals.* Upper Saddle River, NJ: Prentice Hall, 2005.

[181] P. Welch. The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics 15(2),* pp. 70-73. 1967.

[182] N. Yousefian, K. Kokkinakis and P. C. Loizou, "A coherence-based algorithm for noise reduction in dual-microphone applications," in *18th European Signal Processing Conference (EUSIPCO-2010),* Aalborg, Denmark, pp. 1904, 2010.

[183] A. Guerin, R. Bouquin-Jeannes and G. Faucon, "A two-sensor noise reduction system: Applications for hands-free car kit," in *EURASIP Journal on Applied Signal Processing,* pp. 1125, 2003.

[184] R. Bouquin-Jeannes, A. A. Azirani and G. Faucon, "Enhancement of Speech Degraded by Coherent and Incoherent Noise Using a Cross-Spectral Estimator," *IEEE Transactions on Audio, Speech, and Language Processing,* vol. 5, pp. 484, 1997.

[185] R. Le Bouquin and G. Faucon. Using the coherence function for noise reduction. *IEEE Proceedings I - Communications, Speech and Vision 139(3),* pp. 276. 1992.

[186] N. Yousefian and P. C. Loizou. A dual-microphone speech enhancement algorithm based on the coherence function. *IEEE Transactions on Audio, Speech, and Language Processing 20(2),* pp. 599. 2012.

[187] Shoufeng Lin, S. Nordholm, Hai Huyen Dam and Pei Chee Yong. An adaptive low-complexity coherence-based beamformer. Presented at Control, Automation and Information Sciences (ICCAIS), 2013 International Conference On. 2013.

[188] G. C. Carter. Coherence and time delay estimation. *Proceedings of the IEEE 75(2),* pp. 236-255. 1987.

[189] S. Delikaris-Manias and V. Pulkki, "Cross Pattern Coherence Algorithm for Spatial Filtering Applications Utilizing Microphone Arrays," *IEEE Transactions on Audio, Speech, and Language Processing,* vol. 21, pp. 2356, 2013.

[190] S. Delikaris-Manias and V. Pulkki. Cross spectral density based spatial filter employing maximum directivity beam patterns. Presented at Information, Intelligence, Systems and Applications, IISA 2014, the 5th International Conference On. 2014.

[191] I. A. McCowan and H. Bourlard. Microphone array post-filter based on noise field coherence. *IEEE Transactions on Speech and Audio Processing 11(6),* pp. 709. 2003.

[192] J. Y. Jong and J. H. Jones, "Roller bearing health monitoring using CPLE frequency analysis method," NASA, Tech. Rep. MSFC-329, MSFC-359, 2007.

[193] J. Y. Jong, W. D. Dorland, T. T. Fiorucci, T. Zoladz and T. Nesman, "Coherent phase line enhancer (CPLE) forRotating machinery diagnostics," in *37th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit,* pp. 1, 2001.

[194] Jen-Yi Jong, "Coherent Phase Line Enhancer Spectral Analysis Technique," US 6,408,696 Bl, 2002, 2000.

[195] S. Gade, H. Herlufsen, H. Konstantin-Hansen and N. J. Wismer, "Order tracking analysis," Brüel & Kjær, Tech. Rep. No. 2 -1995, 1995.

[196] J. Blough, "A survey of DSP methods for rotating machinery analysis, what is needed, what is available," *Journal of Sound and Vibration,* vol. 262, pp. 707-720, 2003.

[197] D. Ramirez, J. Via and I. Santamaria. A generalization of the magnitude squared coherence spectrum for more than two signals: Definition, properties and estimation. Presented at 2008 IEEE International Conference on Acoustics, Speech and Signal Processing. 2008.

[198] D. Ramirez, J. Via and I. Santamaria. Multipule-channel signal detection using the generalized coherence spectrum. Presented at 1st IAPR Workshop on Cognitive Information Processing (CIP 2008). 2008.

[199] Q. Wang and C. R. Wan. A novel CFAR tonal detector using phase compensation. *IEEE Journal of Oceanic Engineering 30(4),* pp. 900. 2005.

[200] Z. Tan and X. Zhang. Comparison of frequency domain and time domain method for single tone detection. Presented at Industrial Technology, 2008. ICIT 2008. IEEE International Conference On. 2008.

[201] H. C. So, Y. T. Chan, Q. Ma and P. C. Ching. Comparison of various periodograms for sinusoid detection and frequency estimation. *IEEE Transactions on Aerospace and Electronic Systems 35(3),* pp. 945-952. 1999.

[202] Qing Wang, Yixin Yang and Chunru Wan. Design of CFAR detector combining with frequency estimation. Presented at Fourth International Conference on Information, Communications and Signal Processing. 2003.

[203] F. M. Ahmed, K. A. Elbarbary and A. R. H. Elbardawiny. Detection of sinusoidal signals in frequency domain. Presented at 2006 CIE International Conference on Radar. 2006.

[204] S. M. Kay and J. R. Gabriel. Optimal invariant detection of a sinusoid with unknown parameters. *IEEE Transactions on Signal Processing 50(1),* pp. 27. 2002.

[205] Chun Ru Wan, Joo Thiam Goh and Hong Tat Chee. Optimal tonal detectors based on the power spectrum. *IEEE Journal of Oceanic Engineering 25(4),* pp. 540-552. 2000.

[206] D. Rife and R. Boorstyn. Single tone parameter estimation from discrete-time observations. *IEEE Transactions on Information Theory 20(5),* pp. 591-598. 1974.

[207] G. Parker and L. White. Sinusoid detection using a sequential DFT test. Presented at Information, Decision and Control, 1999. IDC 99. Proceedings. 1999. 1999.

[208] Qing Wang, Chunru Wan and Joo Thiam Goh. Theoretical performance analysis and simulation of a GLRT tonal detector. Presented at OCEANS, 2001. MTS/IEEE Conference and Exhibition. 2001.

[209] G. Tong Zhou and M. Z. Ikram. Unsupervised detection and parameter estimation of multi-component sinusoidal signals in noise. Presented at Signals, Systems and Computers, 2000. Conference Record of the Thirty-Fourth Asilomar Conference On. 2000.

[210] X. Xu, R. - Zheng, G. - Chen and E. - Blasch. Performance analysis of order statistic constant false alarm rate (CFAR) detectors in generalized rayleigh environment. Presented at Proc. of SPIE. 2007.

[211] G. V. Trunk and S. F. George. Detection of targets in non-gaussian sea clutter. *IEEE Transactions on Aerospace and Electronic Systems AES-6(5),* pp. 620-628. 1970.

[212] P. P. Gandhi and S. A. Kassam. Analysis of CFAR processors in nonhomogeneous background. *IEEE Transactions on Aerospace and Electronic Systems 24(4),* pp. 427-445. 1988.

[213] A. Jalil, H. Yousaf and M. I. Baig. Analysis of CFAR techniques. Presented at 2016 13th International Bhurban Conference on Applied Sciences and Technology (IBCAST). 2016.

[214] H. M. Finn and P. S. Johnson, "Adaptive detection mode with threshold control as a function of spatially sampled clutter estimation," *RCA Review,* vol. 29, pp. 414-464, 1968.

[215] H. Rohling. Radar CFAR thresholding in clutter and multiple target situations. *IEEE Transactions on Aerospace and Electronic Systems AES-19(4),* pp. 608-621. 1983.

[216] L. Sevgi. Hypothesis testing and decision making: Constant-false-alarm-rate detection. *IEEE Antennas and Propagation Magazine 51(3),* pp. 218-224. 2009.

[217] J. A. Ritcey. Performance analysis of the censored mean-level detector. *IEEE Transactions on Aerospace and Electronic Systems AES-22(4),* pp. 443-454. 1986.

[218] A. Di Vito and G. Morreti. Probability of false alarm in CA-CFAR device downstream from linear-law detector. *Electronics Letters 25(25),* pp. 1692-1693. 1989.

[219] A. Sarma, "Nonparametric approaches for analysis and design of incoherent adaptive CFAR detectors," 2006.

[220] E. B. El Mashade, "Performance Analysis of CFAR Detection of Fluctuating Radar Targets in Nonideal Operating Environments," *International Journal of Aerospace Sciences,* vol. 1, pp. 21-35, 2012.

[221] G. M. Dillard and C. E. Antoniak. A practical distribution-free detection procedure for multiple-range-bin radars. *IEEE Transactions on Aerospace and Electronic Systems AES-6(5),* pp. 629-635. 1970.

[222] R. Nitzberg. Constant-false-alarm-rate signal processors for several types of interference. *IEEE Transactions on Aerospace and Electronic Systems AES-8(1),* pp. 27-34. 1972.

[223] S. Nagarajan, G. Chaturvedi and S. Dhage. Modified distribution free CFAR processor for clutter edges and multi-target situations. Presented at Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '84. 1984.

[224] G. W. Zeoli and T. S. Fong. Performance of a two-sample mann-whitney nonparametric detector in a radar application. *IEEE Transactions on Aerospace and Electronic Systems AES-7(5),* pp. 951-959. 1971.

[225] S. Chikara, K. Saji, M. Sekine and T. Musha. Suppression of radar clutter via nonparametric CFAR. *Electronics and Communications in Japan (Part I: Communications) 74(6),* pp. 107-114. 1991.

[226] R. S. Raghavan. Analysis of CA-CFAR processors for linear-law detection. *IEEE Transactions on Aerospace and Electronic Systems 28(3),* pp. 661-665. 1992.

[227] J. A. Ritcey and J. R. Holm. Applications of nonlinear filtering in radar CFAR detection. Presented at Circuits and Systems, 1992. ISCAS '92. Proceedings., 1992 IEEE International Symposium On. 1992.

[228] R. Inkol, S. Wang and S. Rajan. FFT filter bank-based CFAR detection schemes. Presented at 2007 50th Midwest Symposium on Circuits and Systems. 2007.

[229] Y. Xu, C. Hou, S. Yan, J. Li and C. Hao. Fuzzy statistical normalization CFAR detector for non-rayleigh data. *IEEE Transactions on Aerospace and Electronic Systems 51(1),* pp. 383-396. 2015.

[230] M. Barkat and P. K. Varchney, "On adaptive cell-averaving CFAR radar signal detection," Rome Air Development Center Air Force Systems Command, Griffiss Air Force Base, NY, Tech. Rep. RADC-TR-17-160, 1987.

[231] M. B. El Mashade. Performance analysis of the modified versions of CFAR detectors in multiple-target and nonuniform clutter. *Radioelectronics and Communications Systems 56(8),* pp. 385-401. 2013.

[232] P. Tsakalides, F. Trinci and C. L. Nikias. Radar CFAR thresholding in heavy-tailed clutter and positive alpha-stable measurements. Presented at Signals, Systems & Computers, 1998. Conference Record of the Thirty-Second Asilomar Conference On. 1998.

[233] M. Weiss. Analysis of some modified cell-averaging CFAR processors in multiple-target situations. *IEEE Transactions on Aerospace and Electronic Systems AES-18(1),* pp. 102-114. 1982.

[234] M. Barkat, *Signal Detection and Estimation.* Norwood, MA: Artech House, 2005.

[235] F. E. Nathanson, *Radar Design Principles.* SciTech Publishing, 1999.

[236] S. Blake. OS-CFAR theory for multiple targets and nonuniform clutter. *IEEE Transactions on Aerospace and Electronic Systems 24(6),* pp. 785-790. 1988.

[237] A. Sarma and D. W. Tufts. Robust adaptive threshold for control of false alarms. *IEEE Signal Processing Letters 8(9),* pp. 261-263. 2001.

[238] J. D. Stevenson, S. O'Young and L. Rolland. Estimated levels of safety for small unmanned aerial vehicles and risk mitigation strategies. *J. Unmanned Veh. Sys. (4),* pp. 205. 2016.

[239] E. E. Kummer, "Über die hypergeometrische Reihe F(a;b;x)," *J. Reine Angew. Math.,* vol. 15, pp. 39-83, 1836.

[240] A. Farnell, *Designing Sound.* Cambridge, Massachusetts: MIT Press, 2010.

[241] M. W. Lee, "Spectral whitening in the frequency domain," United States Department of the Interior, Denver, Colorado, Tech. Rep. 86-108, 1986.

[242] H. V. Trees, *Optimum Array Processing.* New York: Wiley, 2002.

[243] B. D. Van Veen and K. M. Buckley. Beamforming: A versatile approach to spatial filtering. *IEEE ASSP Magazine 5(2),* pp. 4-24. 1988.

[244] T. E. Tuncer and B. Friedlander, *Classical and Modern Direction-of-Arrival Estimation.* Burlington, MA, USA: Elsevier, 2009.

[245] V. Krishnaveni, T. Kesavamurthy and B. Aparna, "Beamforming for Direction-of-Arrival (DOA) Estimation-A Survey," *International Journal of Computer Applications,* vol. 61, pp. 4, 2013.

[246] A. Hero, H. Messer, J. Goldberg, D. J. Thomson, M. G. Amin, G. Giannakis, A. Swami, J. K. Tugnait, A. Nehorai, A. L. Swindlehurst, J. F. Cardoso, Lang Tong and J. Krolik. Highlights of statistical signal and array processing. *IEEE Signal Processing Magazine 15(5),* pp. 21-64. 1998.

[247] M. S. Brandstein and H. F. Silverman, "A practical methodology for speech source localization with microphone arrays," *Computer Speech and Language,* vol. 11, pp. 91-126, 1997.

[248] H. K. Zaveri, "Beamforming: A technical review," Brüel & Kjær, Nærum, Denmark, Tech. Rep. 1-2004, 2004.

[249] I. A. McCowan, "Robust Speech Recognition using Microphone Arrays," 2001.

[250] L. C. Godara. Application of antenna arrays to mobile communications. II. beam-forming and direction-of-arrival considerations. *Proceedings of the IEEE 85(8),* pp. 1195-1245. 1997.

[251] H. Krim and M. Viberg. Two decades of array signal processing research: The parametric approach. *IEEE Signal Processing Magazine 13(4),* pp. 67-94. 1996.

[252] F. Li and R. J. Vaccaro. Sensitivity analysis of DOA estimation algorithms to sensor errors. *IEEE Transactions on Aerospace and Electronic Systems 28(3),* pp. 708-717. 1992.

[253] J. C. Chen, Kung Yao and R. E. Hudson. Source localization and beamforming. *IEEE Signal Processing Magazine 19(2),* pp. 30-39. 2002.

[254] J. H. DiBiase, H. F. Silverman and M. S. Brandstein, *Microphone Arrays - Signal Processing Techniques and Applications.* Springer-Verlag, 2001.

[255] J. Chen, J. Benesty and Y. A. Huang, "Time Delay Estimation in Room Acoustic Environments: An Overview," *EURASIP Journal on Applied Signal Processing,* vol. 2006, pp. 1-19, 2006.

[256] Y. Chan, R. Hattin and J. Plant. The least squares estimation of time delay and its use in signal detection. *IEEE Transactions on Acoustics, Speech, and Signal Processing 26(3),* pp. 217-222. 1978.

[257] C. Knapp and G. C. Carter. The generalized correlation method for estimation of time delay. *Acoustics, Speech and Signal Processing, IEEE Transactions On 24(4),* pp. 320-327. 1976.

[258] G. C. Carter, A. H. Nuttall and P. Cable. The smoothed coherence transform. *Proceedings of the IEEE 61(10),* pp. 1497-1498. 1973.

[259] J. Dmochowski, J. Benesty and S. Affes. Direction of arrival estimation using the parameterized spatial correlation matrix. *IEEE Transactions on Audio, Speech, and Language Processing 15(4),* pp. 1327-1339. 2007.

[260] J. H. DiBiase, "A High-Accuracy, Low-Latency Technique for Talker Localization in Reverberant Environments Using Microphone Arrays," 2000.

[261] H. Do and H. F. Silverman. A fast microphone array SRP-PHAT source location implementation using coarse-to-fine region contraction(CFRC). Presented at 2007 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. 2007.

[262] M. Cobos, A. Marti and J. J. Lopez. A modified SRP-PHAT functional for robust real-time sound source localization with scalable spatial sampling. *IEEE Signal Processing Letters 18(1),* pp. 71-74. 2011.

[263] H. Do, H. F. Silverman and Y. Yu. A real-time SRP-PHAT source location implementation using stochastic region contraction(SRC) on a large-aperture microphone array. Presented at 2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07. 2007.

[264] H. Do and H. F. Silverman. SRP-PHAT methods of locating simultaneous multiple talkers using a frame of microphone array data. Presented at 2010 IEEE International Conference on Acoustics, Speech and Signal Processing. 2010.

[265] G. Lathoud and M. Magimai-Doss. A sector-based, frequency-domain approach to detection and localization of multiple speakers. Presented at Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005. 2005.

[266] D. Salvati, C. Drioli and G. L. Foresti. Exploiting a geometrically sampled grid in the steered response power algorithm for localization improvement. *J. Acoust. Soc. Am. 141(1),* pp. 586-601. 2017.

[267] J. Dmochowski, J. Benesty and S. Affes. Fast steered response power source localization using inverse mapping of relative delays. Presented at 2008 IEEE International Conference on Acoustics, Speech and Signal Processing. 2008.

[268] Y. Oualil, F. Faubel and D. Klakow. A fast cumulative steered response power for multiple speaker detection and localization. Presented at 21st European Signal Processing Conference (EUSIPCO 2013). 2013.

[269] J. P. Dmochowski, J. Benesty and S. Affes. A generalized steered response power method for computationally viable source localization. *IEEE Transactions on Audio, Speech, and Language Processing 15(8),* pp. 2510-2526. 2007.

[270] D. N. Zotkin and R. Duraiswami. Accelerated speech source localization via a hierarchical search of steered response power. *IEEE Transactions on Speech and Audio Processing 12(5),* pp. 499-508. 2004.

[271] L. O. Nunes, W. A. Martins, M. V. S. Lima, L. W. P. Biscainho, M. V. M. Costa, F. M. Gonçalves, A. Said and B. Lee. A steered-response power algorithm employing hierarchical search for acoustic source localization using microphone arrays. *IEEE Transactions on Signal Processing 62(19),* pp. 5171-5183. 2014.

[272] J. M. Peterson and C. Kyriakakis. Analysis of fast localization algorithms for acoustical environments. Presented at Conference Record of the Thirty-Ninth Asilomar Conference onSignals, Systems and Computers, 2005. 2005.

[273] M. V. S. Lima, W. A. Martins, L. O. Nunes, L. W. P. Biscainho, T. N. Ferreira, M. V. M. Costa and B. Lee. A volumetric SRP with refinement step for sound source localization. *IEEE Signal Processing Letters 22(8),* pp. 1098-1102. 2015.

[274] A. Marti, M. Cobos, J. J. Lopez and J. Escolano, "A steered response power iterative method for high-accuracy acoustic source localization," *J Acoust Soc Am.,* vol. 134, pp. 2627, 2013.

[275] M. Wax and T. Kailath. Optimum localization of multiple sources by passive arrays. *IEEE Transactions on Acoustics, Speech, and Signal Processing 31(5),* pp. 1210-1217. 1983.

[276] J. Lewis. (May 2012). *Understanding Microphone Sensitivity*. Available: http://www.analog.com/en/analog-dialogue/articles/understanding-microphone-sensitivity.html.

[277] M. S. Brandstein, J. E. Adcock and H. F. Silverman. Microphone-array localization error estimation with application to sensor placement. *J. Acoust. Soc. Am. 99(6),* pp. 3807-3816. 1996. Available: https://doi.org/10.1121/1.414998.

[278] A. M. Đuranec, D. Miljković and T. Bucak, "Community noise analysis of GA aircraft - local airports case study," in *5th Congress of Alps-Adria Acoustics Association,* Petrčane, Croatia, pp. AR-02, 2012.

[279] J. Ahrens and S. Spors. Reproduction of moving virtual sound sources with special attention to the doppler effect. Presented at Audio Engineering Society Convention 124. 2008.

[280] H. Camargo and R. Burdisso. "A frequency domain technique to de-dopplerize the acoustic signal from a moving source of sound," in *17th AIAA/CEAS Aeroacoustics Conference (32nd AIAA Aeroacoustics Conference)*Anonymous 2011, .

[281] H. E. Camargo, "A Frequency Domain Beamforming Method to Locate Moving Sound Sources," 2010.

[282] J. A. Belloch, A. Gonzalez, F. J. Martínez-Zaldívar and A. M. Vidal, "Real-time massive convolution for audio applications on GPU," *The Journal of Supercomputing*, vol. 8, pp. 449–457, 2011.

[283] N. Jillings and Y. Wang. CUDA accelerated audio digital signal processing for real-time algorithms. pp. 697-710. 2014.

[284] A. Yakubovskiy, H. Salloum and A. Sutin, "Feature extraction for acoustic classification of small aircraft," in *2015 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, NY, USA, USA, pp. 18-21, 2015.

[285] E. Çak&imath, r, G. Parascandolo, T. Heittola, H. Huttunen and T. Virtanen. Convolutional recurrent neural networks for polyphonic sound event detection. *IEEE/ACM Transactions on Audio, Speech, and Language Processing 25(6)*, pp. 1291-1303. 2017.