SUCCESSIVE TARGET CANCELATION FOR RADAR WAVEFORM SIDELOBE REDUCTION

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ONUR HALİLOĞLU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

SEPTEMBER 2006

Approval of the Graduate School of Natural and Applied Sciences

Prof. Dr. Canan ÖZGEN
Director Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. İsmet ERKMEN
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Asst. Prof. Dr. Ali Özgür YILMAZ
Supervisor

Examining Committee Members

Prof. Dr. Yalçın TANIK	(METU, EEE)	
Assist. Prof. Dr. Ali Özgür YILMAZ	(METU, EEE)	
Assoc. Prof. Dr. Sencer KOÇ	(METU, EEE)	
Asst. Prof. Dr. Çağtay CANDAN	(METU, EEE)	
Ülkü Çilek DOYURAN, M. Sc. in EE	(ASELSAN)	

I hearby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required, I have fully cited and referenced all material and results that are not original to this work.

Name Lastname : Onur HALİLOĞLU

Signature :

ABSTRACT

SUCCESSIVE TARGET CANCELATION FOR RADAR WAVEFORM SIDELOBE REDUCTION

HALİLOĞLU, Onur
M.Sc., Department of Electrical and Electronics Engineering
Supervisor: Asst. Prof. Dr. Ali Özgür YILMAZ

September 2006, 79 pages

Many radars suffer from masking of weaker targets by stronger ones due to range sidelobes of pulse compression codes. We propose a method to prevent this by successively detecting targets and canceling their effects. Performance of the proposed method will be investigated in various scenarios with regard to existence of noise, targets, and the Doppler effect.

Keywords: Sidelobe Suppression, Radar Signal Processing

ÖZ

RADAR DALGA BİÇİMİ YANLOB BASTIRMASI İÇİN ARDIŞIK HEDEF YOKETME

HALİLOĞLU, Onur
Yüksek Lisans, Elektrik ve Elektronik Mühendisliği Bölümü
Tez Yöneticisi: Asst. Prof. Ali Özgür YILMAZ

Eylül 2006, 79 sayfa

Bir çok radar uygulamasında, darbe sıkıştırma kodlarının ikincil kulakları (yan lobları) nedeniyle küçük hedeflerin daha büyük hedefler tarafından gölgelenmesi sorun teşkil etmektedir. Biz bu sorunu çözmek için, ardışık olarak hedefleri saptayan ve saptanan hedeflerin etkilerini eleyen bir yöntem öneriyoruz. Önerilen bu metodun performansı; gürültünün ve Doppler etkisinin dikkate alındığı çeşitli hedef kombinasyonlarından oluşan farklı senaryolar için incelenecektir.

Anahtar sözcükler: Yanlob Bastırma, Radar Sinyal İşleme

ACKNOWLEDGMENTS

I would like to thank to my advisor Assist. Prof. Dr. Ali Özgür Yılmaz for his support, contributions and valuable guidance.

I also would like to thank Fatih Bayramoğlu for his help in Latex.

I am grateful to my group coordinator Orhan Şengül for the facilities he provided in order to complete this thesis.

I wish to express my thanks to my parents Cemal and Nurten Haliloğlu, but especially to my mother Nurten Haliloğlu, who supports me and tries to understand me during my life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Introduction

Radar (Radio Detection and Ranging) is an electromagnetic system that uses radio waves to detect and to determine the distance and/or speed of objects such as aircraft, ships, people, and the natural environment. Since World War II a great effort has been put into development of radar systems and hence radar technology has immensely advanced.

A typical radar utilizes a pulse waveform. Duration of the pulse waveform is a critical issue since the properties of a radar such as resolution and range are determined by the pulse duration. High range resolution is provided when short-duration pulse is used. On the contrary, long duration pulses have lower resolution. On the other hand, long-duration pulses have some advantages over short-duration pulses. Since the width of a pulse in time domain is inversely proportional to its spectral bandwidth, the bandwidth of a short pulse is large and it is known that large bandwidth can increase system complexity. Indeed, complexity resulted from large bandwidth can be avoided by using long pulses. In addition to this, there is a peak power limitation problem for the short pulse. If it is aimed to transmit the signal to long ranges, the energy of the pulse must be high. In order to achieve the range as long as the one achieved by using a long pulse, the peak value of the short pulse must be high since equal energy is needed for transmitting the radar signal to the same range. However, the transmission line of a high peak power radar can be subject to voltage breakdown. Then, the pulse might not have sufficient energy to achieve longer ranges. This is called peak power limitation problem.

As a result, transmitting either a long pulse or a short pulse is advantageous in some aspects and disadvantageous in others. It may be desired to have the advantages of the two in a radar system. The solution is to use pulse compression in which short radar pulse is lengthened before transmission to reduce its average power level. Then at the receiver side pulse is compressed with a pulse compression filter as soon as it is received so that the range resolution of the original short pulse is restored. The pulse compression filter is a matched filter; hence, when the noise is neglected, its output envelope is the autocorrelation function of the input signal. Due to the non-ideal autocorrelation of radar signals, the envelope of a matched filter has time-sidelobes. Sometimes these sidelobes are relatively high with respect to mainlobes of other targets. This is not usually preferable because range sidelobes can mask weaker targets and additionally sidelobes can be mistaken as new targets during the radar signal processing. Therefore, sidelobe reduction has been a popular issue since the invention of pulse compression radars. There are various methods to overcome the effects of range sidelobe problem. Two of the most popular of these techniques are Least Squares (LS) method and Minimum Mean Square Error (MMSE) method.

LS method [1] estimates the received return radar samples by minimizing the squared errors between the observed received return samples and the estimated ones. LS estimation is performed by multiplying the received signal in the vector form by matrices obtained from code autocorrelation. Inversion of a $n \times n$ matrix is necessary in this method where n is the length of the received signal. Although the matrix inversion is done off-line, the multiplication operation has an order of $O(n^2)$ and thus it is prohibitively computationally intensive. Hence, the greater the number of the samples taken for the received return signal is, the more computationally complex the LS algorithm becomes.

On the other hand, in MMSE method [2] a filter is generated first by using the matched filter output to obtain target power estimates. A signal covariance matrix is obtained for each range bin by using target power estimates. Another filter is reconstructed based on the inverse of the covariance matrix and acts on the MF signal. Smaller targets are identified by checking the filtered MF signal. At each iteration of the algorithm, target powers are reestimated and filters are regenerated. The algorithm is computationally complex due to on-line matrix inversion at every iteration. Furthermore, the estimated covariance matrix is reportedly ill-conditioned and thus ad hoc methods are used to avoid numerical instability in running the algorithm.

As an alternative to these two methods, we propose a new method herein. This method is based on obtaining the target's phasor from the MF output and canceling part of the signal due to the target. This can be repeated for each detected target. Since this method successively detects a target's signal and cancels it from the overall MF output, we will refer to it as Successive Target Cancelation (STC). STC is based on basic operations such as addition and subtraction. It does not include matrix inversion or other complex operations and hence is simple and easy to handle. This algorithm known as the CLEAN algorithm in general was applied in image processing [3, 4] and astronomy [3, 5] as a deconvolution method. It is also utilized in multiuser communications [6] and called Successive Interference Cancelation. Except for a couple of works in recent years, the CLEAN algorithm has rarely been applied to detection in radar signal processing. Most recently, the CLEAN algorithm has been applied to an extended target or contiguous clutter scenario in [7]. However, performance of the CLEAN algorithm has not been evaluated in comparison to other methods.

In summary, in this study, the performance of STC is investigated in various scenarios with regard to existence of noise, targets, and the Doppler effect. Moreover, the performance of the STC method is compared to other methods proposed in literature. Differing from similar studies, we will study the effects of the Doppler shift occurring in radar signals due to target velocities.

CHAPTER 2

Review of history

2.1 Radar

2.1.1 Basic Information [8, 9]

The term radar is an acronym for Radio, Detection, And Ranging. Radar is essentially a ranging or distance measuring device. It is an electronic equipment that uses reflected electromagnetic energy for detecting the presence, direction, height, and distance of objects. Radars have their own source of energy to produce images so they are active remote sensing systems. Darkness does not affect the operating frequency of electromagnetic energy used for radar, that is radar does not require sunlight. Therefore, radar systems can determine the positions of objects that are invisible to the naked eye due to darkness, distance or weather. Of course, in order to express the positions of the detected objects radar systems need reference coordinate systems because of their limited field of view.

Radar is based on the transmission and reception of pulses in a narrow beam in electromagnetic spectrum bands. A radar system requires a transmitter to send the radar signal to the target or targets and a receiver to receive the returned signal from targets. Firstly, a powerful transmitter generates the radar signal. The reflected signal from targets, also called scattering signal, spreads in a large number of directions. Then, a sensitive receiver picks up the electrical signal which is called echo. The echoes returning from targets are then recorded and taken into consideration according to their strength, time interval, and phase.

When the reflections from targets are in the opposite direction to the incident rays, it is called backscattering. The power received by the antenna board from each radar pulse transmitted is directly related with backscattering. Backscattering depends mainly on three factors: the surface roughness, the surface humidity, and the wavelength of the radar. It is possible to declare two surface types for backscattering: rough surface and smooth surface. The former has higher backscattering compared to the latter since the dimensions of the rough surface are very small in respect of the transmitted radar signal wavelength. For instance, water surfaces with and without waves exhibit different scattering characteristics. Unlike water surfaces which are moved by wind effects, water surfaces without waves do not allow the incoming energy to be scattered back to the antenna. Surface humidity has a great influence on the electrical properties of a target. As the humidity becomes higher, backscattering increases. Since the surface roughness is dependent on the wavelength of the radar and also longer wavelengths result in higher penetration capabilities, wavelength of the radar signal is an important issue to consider with respect to backscattering.

2.2 Characteristics of Radar Systems [8, 10]

The characteristics of radar systems, which differentiate the radar systems from each other, can be categorized into three parts: measurable characteristics, descriptive characteristics, additional intercept characteristics.

2.2.1 Measurable Characteristics

Measurable characteristics are the parameters that are measured to identify a particular radar. Radio frequency, pulse repetition frequency, pulse duration, and scan rate are the certain measurable characteristics of radar systems.

2.2.1.1 Radio frequency

Generally,conventional radars operate in the microwave region approximately covering the frequencies from $1GHz$ to $30GHz$. Radio frequency abbreviated as RF is a term that refers to a portion of the electromagnetic spectrum in which an electromagnetic field suitable for wireless broadcasting and/or communications is generated when alternating current is input to an antenna. RF can be measured by either zero beat method which stands for comparison between the radar frequency and a stable frequency source or directly reading from a frequency counter. Radar RF can determine the size of the components and the radar range. For instance, high radar frequencies originating from small antennas are used at shorter ranges whereas low frequencies originating from large antennas are used at longer ranges. Usually, RF of a radar in early warning radars is lower compared to missile tracking and control radars. Therefore, RF can also be decided according to the use of radar. Moreover, power requirements and propagation characteristics of the radar systems are affected by the radio frequency.

2.2.1.2 Pulse Repetition Frequency

Pulse repetition frequency (PRF) that is used interchangeably with pulse repetition rate is the number of pulses emitted by radar in one second. The reciprocal of PRF is named as pulse repetition interval that can be defined as the time between the start of one pulse and the start of the next pulse. PRF is crucial since it determines the maximum unambiguous target range. The duration between the transmitted pulses must be arranged so as to allow echo signals belonging to one pulse to return back to the receiver antenna and to be detected before the transmission of the next pulse. Therefore, maximum unambiguous range of detection that the radar can reach can be formulated as:

$$R_{max} = \frac{c}{2 \cdot PRF}, \tag{2.1}$$

where c is the speed of light. As it can be seen from the formula, maximum unambiguous range detection R_{max} decreases with increasing PRF. The detection range that exceeds R_{max} is called ambiguous range and high PRF causes range ambiguity. On the other hand, high PRF results in more hits per sweep and higher probability of detection. Moreover, low PRF is not sufficient when high data rates are required as in fire control radars.

It has to be mentioned that there is a minimum PRF limitation that can be determined by the rotational speed of the antenna. For a rotationally constant speed system antenna, energy beam strikes a target for a short period of time. Sufficient number of pulses that have the total energy to provide detection of the signal by the radar during this limited time is required.

2.2.1.3 Pulse Duration

Pulse duration is the length of the time that transmitter is actually on. It is usually measured in milliseconds. Pulse duration determines range resolution; hence, it is an important parameter for pulse compression which will be discussed later. Range resolution is the ability of a radar to display targets separately. Targets too close together can be sensed as one target by the radar. Range resolution is solely a function of pulse duration. For high range resolutions the duration of the pulse must be small.

Pulse duration also determines theoretical minimum range. The theoretical minimum range, at which the radar receiver is unable to receive echoes during the emission of pulses from the transmitter, is the minimum distance separating the target and the radar.

Pulse duration is one of the two factors that affect total pulse energy; the other is peak value of the transmitted power. As pulse energy increases, probability of detection increases. High peak power values are not allowed because of peak power limitation of radar systems. Therefore, long pulse duration is preferred for increasing the probability of detection.

2.2.1.4 Scan Rate

Scan rate, the inverse of scan period, is the number of rotations of the antenna beam in a specified period of time. In this period of time, various functions are performed by the radar. Scan rate determines the purpose of the radar. Detection capability increases with slow antenna scan rate since more pulses return to the receiver antenna as the antenna beam scans the environment slowly. On the other hand, slow scan rate results in a longer time between observations of the target which leads to inaccurate target position prediction. Therefore there is a trade of between probability of detection and target position prediction related to scan rate.

2.2.2 Descriptive Characteristics

Descriptive characteristics of a radar are used to modify or clarify the measurable characteristics. Unlike measurable characteristics, they are non-parameter data. Modulation type and polarization are the important descriptive characteristics of the radar.

2.2.2.1 Modulation Type

Modulation type is employed according to the purpose of the radar in the transmitter part. Usually, pulse modulation types are used in radars. Pulse modulation techniques are mainly based on varying the amplitude, width, phase, and shape of the pulse. The characteristics that the pulse duration affects were mentioned earlier. The shape of the pulse determines the range accuracy, minimum range, maximum range, and target resolution. Other pulse variations will be discussed later.

2.2.2.2 Polarization

Polarization is the property of electromagnetic waves, which describes the direction of their transverse electric field. The polarization type of a radar wave is defined by the orientation of the electric field. Horizontal, vertical, circular, and elliptical are some of the polarization types. Although most radars are linearly polarized, that is, horizontally or vertically polarized,other polarization types are sometimes used to enhance the detection of targets in different environmental conditions. In both circularly and elliptically polarized antennas, electric field rotates at a rate equal to the RF frequency. However, in elliptical polarization the amplitude of the electric field varies during the rotation period.

The polarization of the radar signal can affect the radar echo from the target. Therefore, it is possible to sense the difference between a long thin rod and a sphere by using circular polarization or horizontal polarization. However, it is difficult to do this for practical targets such as aircrafts. Polarization can also be a misleading parameter because of multipath radiation from the surface of the ground or clutter, limitations in the polarization response of practical antennas, and target movements that causes distorted polarization response.

2.2.3 Additional Intercept Characteristics

Except from the measurable and descriptive characteristics, there are other parameters that differentiate radar systems from each other such as time of intercept, signal strength, and antenna beamwidth.

2.2.3.1 Time of Intercept

Time of intercept can be used in analysis of radar systems that are supported by many emitters. For operation of the radar system, functions of various emitters can be explained by the up and down times of the signal. Time of intercept is an important characteristic for collectors to understand the source of interception.

2.2.3.2 Signal Strength

Signal strength is the measure of how strongly a transmitted signal is being received, measured, or predicted at a sufficiently far away point from the transmitter antenna. Signal strength can be measured as the quantity of received signal power or signal voltage per square area. Distance between the receiver and the transmitter, propagation characteristics, relative positions of the receiving and transmitting antennas, and respective radiation patterns of the receiving and transmitting antennas are the factors that affect the signal strength measurement.

2.2.3.3 Beamwidth

Beamwidth (BW) is the angle between the half-power ($3dB$) points of the main lobe of the radiated power. It is typically measured in degrees. Beamwidth is limited by radar frequency, antenna size, and antenna shape. The formula between the beamwidth and lobe duration is given below in eqn. (2.2):

$$BW = (lobe\ \ duration \times 360) \div seconds\ \ per\ \ revolution. \tag{2.2}$$

Lobe duration is the beam illumination time of the target when a radar beam scans across the target. Beamwidth determines the accuracy of the azimuth and elevation angles. Therefore, beamwidth is a factor that determines the precision of the location of the target. The selection of the width of the beam changes according to application area of the radars. For instance, early warning radars use wide beams on the other hand fire control radars use narrow beams.

2.3 How does a radar work? [8]

Radar uses electromagnetic energy pulses in order to detect targets. The basic operation of a radar can be summarized as follows. Firstly, the RF energy is emitted to the environment. The emitted energy reflects from the objects. Actually, a small portion of the energy reflects and returns to the radar receiver. Radars use the received echo signals to determine the direction and distance of the reflecting objects. The simple block diagram of the radar is given in Figure 2.1.

Waveform generator produces the radar signal at low power. The radar signal is the input to power amplifier. The power amplification can be realized by a microwave vacuum tube or a solid state device. Microwave vacuum tubes can be a magnetron, klystron, traveling wave tube or a crossed field amplifier. Magnetron, a power oscillator, is the first invented one, it is smaller and it has a modest capability. Klystron, a power amplifier, is used when high average power is required. Also, it allows more complex waveforms provided by the waveform generator. Traveling wave tube has wider bandwidth than klystron. In some radars, crossed field amplifiers are preferred because of their wide bandwidth, modest gain, and compact size. Modulator is the

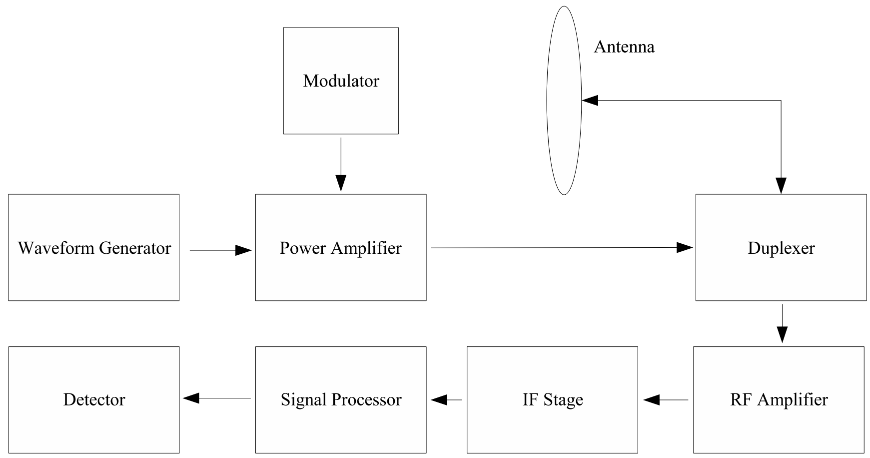

Figure 2.1: Simple radar block diagram.

required part of a radar in the transmitter side. In synchronism with the radar input pulses, the transmitter is turned on and off by the modulator. The modulator can also be used in order to generate a pulse waveform. The modulated and amplified signal is sent to the antenna. The antenna transfers the transmitter energy to signals in space with the required distribution and efficiency. Therefore, the radar transmitter produces the short duration high-power RF pulses of energy that is radiated into space by the antenna. The transmitted radar signal radiates in space and reflects from the objects. The reflected signal is picked up by the antenna. The antenna collects return echoes by focusing transmitted RF energy into directional beam. The same antenna is used for both sending and picking up the radio signal. Therefore, a device is required to allow the usage of a single antenna on a time shared basis for both transmitting and receiving. Duplexer is the microwave switch that allows the use of a single antenna for both transmit and receive cycles. It blanks the receiver in the transmit period and blanks the transmitter in the receive period.

In the receiver part of the radar, weak return RF signal is amplified first. Then the amplified RF signal is down converted to an intermediate frequency. After this process intermediate frequency (IF) stage is realized by the IF amplifier. The signal is then fed to signal processing part of the receiver. The signal processor filters the desired echo signal from clutter and receiver noise. Matched filter and Doppler filter that separates the desired moving targets from unwanted stationary clutter echoes are the examples of a signal processor. Finally, by help of the detector the target signal is detected from the overall output signal.

2.4 Matched Filter

In radar systems, the input signal is first transmitted from the transmitter, then passed through the channel, and finally received by the receiver. The received signal includes the original input signal convolved with the impulse response of the channel and additive noise as shown in

$$y(t) = x(t) * h_c(t) + n(t), \tag{2.3}$$

where $y(t)$ is the received signal, $x(t)$ is the original transmitted signal, $h_c(t)$ is the impulse response of the channel, and $n(t)$ is the additive noise. The goal is to detect the presence of targets by estimating $h_c(t)$ from the received signal. At this point, maximizing the output peak value of signal to noise power ratio of a radar receiver maximizes detection probability. Signal to noise ratio (SNR) which is defined as the ratio of the average power of the message signal to the average power of the noise is usually preferred as a criterion to describe the characteristics of a matched filter. Indeed, a matched filter (MF) is a linear filter which performs optimum detection by providing maximum attainable SNR at the output of the filter when the noise is a white Gaussian noise. Hence, practical receivers estimate the transmitted signal by using matched filtering.

Matched filtering operation is realized by convolving the unknown received signal with a time reversed version of the complex conjugate of the known transmitted signal. It is also equivalent to correlating the known transmitted signal with the unknown received signal in order to detect the presence of transmitted signal in the received signal. Mathematically, the correlation receiver and matched filter are equivalent. Frequency response of the matched filter and impulse response of MF in time domain are given as

$$H(f) = KX^*(f)e^{-j2\pi f T_d}, \tag{2.4}$$

$$h(t) = Kx^*(T_d - t), \tag{2.5}$$

where $*$ denotes the convolution operation and T_d is a constant delay needed for providing a physically realizable filter. The transfer function that is the frequency response of MF is the complex conjugate of the Fourier transform of the input signal $x(t)$ except for a delay factor. A basic matched filter system is given in Figure 2.2. The output of the matched filter is

$$m(t) = y(t) * h(t) = \int y(u)h(t-u)du = \int y(u)Kx^*(T_d - (t-u))du, \tag{2.6}$$

[11]. Autocorrelation function of the transmitted signal $x(t)$ is written as

$$r(t) = \int x(u)x^*(u-t)du. \tag{2.7}$$

Therefore, the output of the matched filter can be also described in terms of the auto-correlation function $r(t)$.

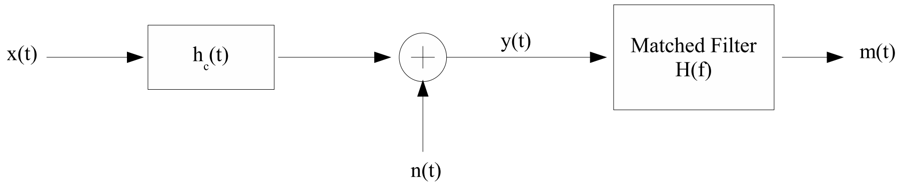

Figure 2.2: Basic Matched filter system.

The chip-rate sampled MF output signal is shown in Figure 2.3 where noise is neglected, impulse response of the channel is a delta function ($h_c(k) = \delta(k)$), and transmitted signal is a P4 code ($x(k) = e^{(\pi/N)(k-1)^2-\pi(k-1)}$).

As a result, matched filters are commonly used in radar systems and provide information from the received signal which is something similar to what was emitted from the radar transmitter. It is important to mention that pulse compression is an example of matched filtering.

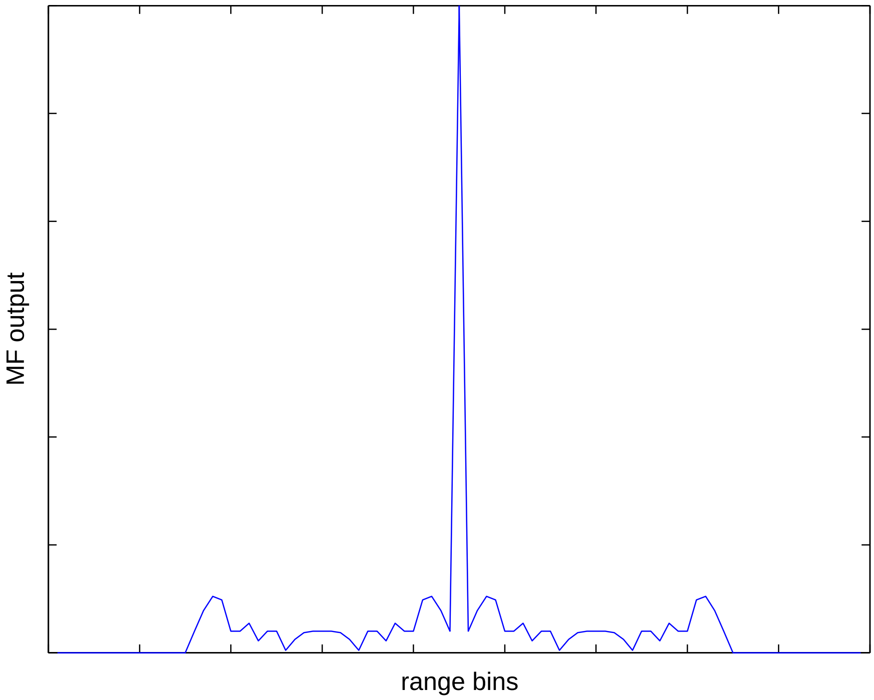

Figure 2.3: Matched filter output of a P4 code.

2.5 Doppler Effect

Doppler effect in radars is the phenomenon that is used to describe the change in the frequency of the electromagnetic signal propagating between the radar and moving target. When a moving target exists, the total phase change in the propagation path can be given as [8]

$$\phi = 2\pi \cdot \frac{2R}{\lambda} = 4\pi \cdot \frac{R}{\lambda}, \tag{2.8}$$

where λ is the wavelength of the radar and R is the line of sight range between the target and the radar. The number 2 along with R results from the two way propagation path that is the path from radar to the target and the one from target to the radar. Since the target is in motion, the range (R) between the target and the radar is changing and so is the phase. Therefore, angular velocity formed by the rate of change in the time varying phase occurs and can be formulated as

$$w_d = \frac{d\phi}{dt} = \frac{4\pi}{\lambda}\frac{dR}{dt} = \frac{4\pi v_r}{\lambda} = 2\pi f_d, \tag{2.9}$$

[8] where $v_r = dR/dt$ is the radial velocity and f_d is the Doppler frequency shift. Therefore, from equation (2.9), [8]

$$f_d = \frac{2v_r}{\lambda} = \frac{2f_t v_r}{c}, \tag{2.10}$$

where f_t is the radar frequency and equals to c/λ with c being the speed of light.

Actually, f_d is the magnitude of the frequency shift caused by the Doppler effect. Doppler frequency shift can be positive or negative ($\mp f_d$). Positive sign shows that the target is moving towards the radar while negative sign shows that the target is moving away from the radar.

Since in this thesis we are dealing with pulse radars, it is convenient to investigate the Doppler effect from the perspective of pulse radars. In pulse radars, when the transmitted signal is $A_t \sin(2\pi f_t t)$, the received signal will be $A_r \sin(2\pi f_t(t - T_r))$ similar to the transmitted signal except for the amplitude change and time shift T_r where f_t is the frequency of the signals and A_t, A_r are the amplitudes of the transmitted and received signal respectively. T_r equals to $2R/c$ and R is represented as $R_0 - v_r t$ when the target is approaching to the radar. As expected R_0 is the initial range between the target and the radar and v_r is the radial velocity of the target with respect to the radar. With the above substitutions the received signal becomes [8]

$$r(t) = A_r \sin[2\pi f_t(1 + \frac{2v_r}{c})t - \frac{4\pi f_t R_0}{c}] \tag{2.11}$$

where the factor $2f_t v_r/c$ is the Doppler frequency shift.

In pulse radars, a single pulse can be adequate to detect the Doppler frequency shift when the pulse width of the transmitted signal and Doppler frequency of the target are sufficiently large. For the realization of this detection, the condition $f_d \tau > 1$ has to be satisfied. However, this is not usually the case in pulse radars since the inverse of pulse duration $1/\tau$ is generally much greater than the Doppler frequency f_d. Thus, instead of a single pulse a pulse train consisting of many pulses is used to detect the Doppler frequency shift. Besides the Doppler effect while using a single pulse that is described earlier, another Doppler effect resulting from multiple pulses has to be taken into account. While using multiple pulses, a frequency shift occurs between the pulses different from the Doppler frequency in a single pulse since the target continues to move in time duration between pulses. How these two Doppler frequency shifts can be taken into account will be explained in Section 4.2 where we will explain our method.

2.6 Radar Types

All radars transmit high frequency signals which are reflected at targets. The echoes reflected of off targets or environment are received and assessed. Although there are many different types of radars, they can be categorized into two main groups on the basis of waveform type: continuous wave radars and pulsed radars.

2.6.1 Continuous Wave (CW) Radar

Continuous wave radar transmits a continuous high-frequency signal instead of pulses. The echo signal is permanently received and processed in CW radars. CW radars have no pulsing so maximum and minimum range definitions are meaningless while talking about CW radars. They can only detect moving targets because non-moving objects will not form a Doppler shift and the signals from these objects will be filtered out.

Due to the lack of pulses to time, CW radars can not measure distance directly. Frequency shifting methods are used to solve this problem. By comparing the frequencies of the transmitted signal and the received one, range estimation is done. Moreover, modulated CW radars where the transmitted signal is constant in the amplitude but modulated in the frequency is more successful at measuring the distance than the unmodulated CW radars in which both the amplitude and the frequency of the transmitted signal are constant. However, for long distances it is either not possible or difficult to measure the range even by using modulated CW radars. Actually, because of these limitations and difficulties, pulsed radars are widely used as alternatives to CW radars today.

2.6.2 Pulsed Radar

Pulsed radar is the most conventional radar that transmits a high frequency impulsive signal of high power. After the transmission of an impulsive signal, no signal is sent for some time. At this listening period, the echoes of the transmitted pulse are received before the transmission of the next pulse. This prevents the interference between the echoes of successive pulses. Since radar waves travel at the speed of light, range from the return can be calculated. With the help of pulsed radars, direction, distance and sometimes the altitude of the target can be determined.

Although they all contain the same basic functional components, some pulsed radars can be rather complex in their composition with additional equipment included for specific purposes. For instance, the fire control radar, one type of a tracking radar, requires additional circuitry in order to measure the target range.

The most famous of the pulsed radars is the pulse Doppler radar. It has advantages over a basic pulse radar or a CW radar. It can determine the range and detect both moving and stationary targets. At different ranges it can also differentiate two targets with the same radial velocity that is the apparent speed of a target while closing on or going away from the radar. Velocity measurements are realized by transmitting many radar pulses towards each target over a very short period of time, and measuring relative target movement between each pulse. Since received pulses are not affected by the tangential movement, these measurements are limited to measuring the component of the target velocity that is parallel to the radar beam.

Due to the power limitations of the transmitter and heating problems of the semiconductors used in the radar, it is risky to increase the transmitted power in order to enlarge the radar range. Therefore, in radar systems, it is desirable to extend the search range without increasing the transmitted power while maintaining the high range resolution. Pulse compression that will be investigated in the next section is a widely used technique for this purpose in pulsed radars [12].

CHAPTER 3

PULSE COMPRESSION

3.1 Short history of pulse compression [13]

Pulse compression relies on a family of techniques which are used to increase the bandwidth of radar pulses. These techniques have been widely employed for applications in military and aviation systems. Since 1950's, several techniques have been studied and published. However, pulse compression techniques used for distributed atmospheric targets were proposed in the 1970's. Later, a coherent radar application using a 7-bit Barker phase-coded transmit pulse and a matched-filter receiver was realized by Fetter. Moreover, pulse compression applications for incoherent scatter observations were introduced. For instance, Gray and Farley [14] studied the binary phase-coded pulse compression used for incoherent observations. During the past few years high power RF pulse compression systems have developed considerably. Today, pulse compression techniques are widely used in modern radar systems.

3.2 Why is pulse compression needed?

Pulse compression, also known as pulse coding, is a signal processing technique designed to maximize the sensitivity and resolution of radar systems. Since the invention of radar systems, high range resolution and range of detection have been important subjects for radar applications. Range resolution that is achievable in a radar system is $dR = cT/2$, where c is the speed of light and T is the pulse duration.

It is obvious that as the pulse duration becomes shorter, dR decreases and range resolution increases. Therefore, high range resolution can be obtained with a short duration pulse. Indeed, long pulses result in low range resolution. High range of detection, also called the sensitivity of the radar, depends on the energy transmitted in the radar pulses. This can be expressed in terms of the average transmitted power, that is, the peak power multiplied by the transmitter duty cycle. Transmitting longer pulses improves the radar's sensitivity by increasing the transmitted energy. The same transmitted energy can be acquired by using short pulses with high peak transmitter powers but the major drawback of the radars was the peak power limitations of the transmitters. The transmission line of a high peak power radar can be subject to voltage breakdown. Hence, in order to increase the sensitivity of the radar long pulses are preferred.

Even if high-power transmitters are used, using high power transmitters has some disadvantages: High power transmitters

- require high-voltage power supplies,
- have reliability problems,
- are bigger, heavier, and costlier.

As a result, transmitting either a long pulse or a short pulse is advantageous in some aspects and disadvantageous in others. It is desired to have the advantages of the two in a radar system. The solution is to use pulse compression in which short radar pulse is lengthened before transmission to reduce its average power level, at the receiver side pulse is compressed with a pulse compression filter as soon as it is received so that the range resolution of the original short pulse is restored.

3.3 The Pulse Compression Technique

The main aim of the pulse compression is to increase the detection range of a radar while preserving the high range resolution of a short pulse. The average energy of the pulse to be transmitted is usually fixed by the system constraints. Energy content of short and long pulses are shown in Figure 3.1. Since the average energy is fixed, the goal is to send the same energy that is

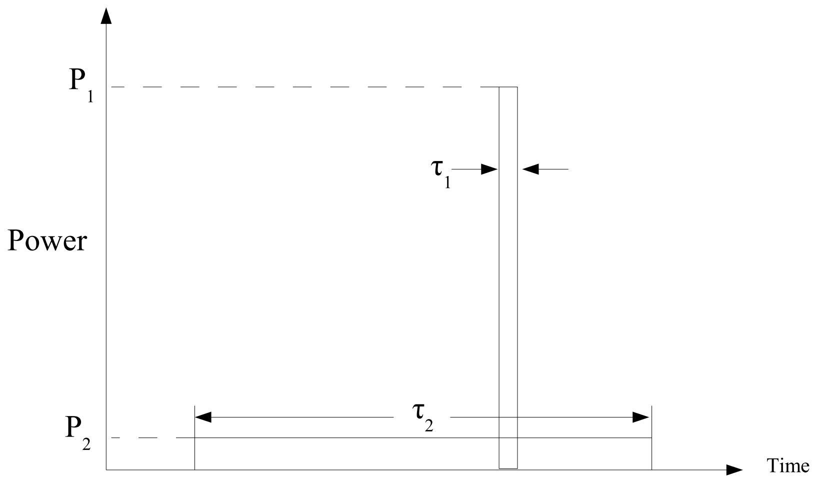

Figure 3.1: A long and a short pulse with the same energy.

$P_1\tau_1 = P_2\tau_2$.

Due to the limited peak power problem and other drawbacks in the transmission of a short pulse, the signal with long pulses is preferred. On the other hand, short pulse is needed at the receiver side for high range resolution. Therefore, a pulse compression subsystem is made of two units: expander and the compressor. The former modulates the transmit signal in order to broaden its pulse width and the latter processes the return echoes after down-conversion to IF.

The basic block diagram of the pulse compression radar (PCR) is shown in Figure 3.2. Modulation part is defined based on pulse compression type.

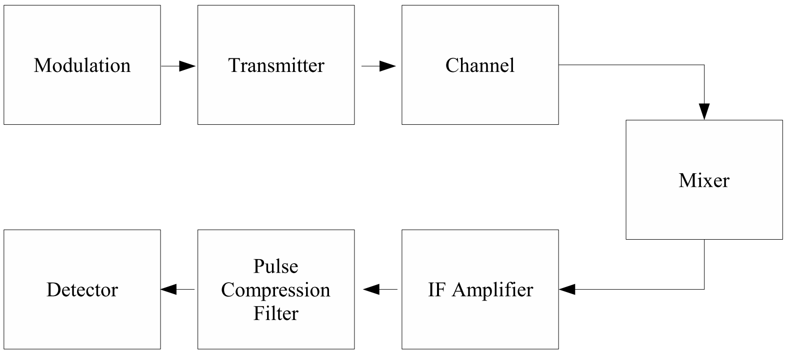

Figure 3.2: Basic PCR block diagram.

However, in all types of pulse compression, the long pulse is modulated so as to have the same spectral bandwidth of a short pulse. The modulated long pulse is sent through the transmitter. Then, in the mixer, the frequency of the signal is shifted downwards and upwards by the carrier frequency. The incoming signal is amplified in the IF amplifier stage.

On reception the modulated long pulse is passed through the pulse compression filter. Pulse compression filter, can be called as the compressor part of the system. The output envelope of the filter is the autocorrelation of the input. In the radar receiver, the received pulses are compressed in time domain, resulting in a range resolution which is finer than the uncoded pulse. Finally, a detector detects the processed return signal.

A pulse compression system is expressed by its compression ratio. Pulse compression ratio is the ratio between the pulse durations before and after the expander or the compressor. Incorporated into a practical radar system, pulse compression filters can have ratios between 10 and 10^5.

3.4 Types of Pulse Compression

Pulse compression achieves high range resolution without an incredibly high peak power. In pulse compression, a code with a peaky autocorrelation is modulated onto the pulsed waveform on transmit. On the receiver side the matched filter for the code generates peaks due to the autocorrelation at positions where targets exist. In order to achieve pulse compression, there are several techniques.

Modulation on the pulse simultaneously enhance the target range, the range resolution. Selecting a pulse compression technique is dependent on the selected waveform type, the generation method, and the processing method. Pulse modulation can be divided into two main parts namely frequency modulation and phase modulation.

3.4.1 Frequency Modulation

For a pulse modulation, frequency modulation (chirp) is a frequency change during transmission of a pulse. There are two types of frequency modulation for the pulse compression: linear frequency modulation and nonlinear frequency modulation.

3.4.1.1 Linear Frequency Modulation (LFM) pulse compression

Linear frequency modulation, also known as chirp, is the most widely used continuous phase pulse compression technique due to its ease of generation and insensitivity to Doppler shifts. A linear FM signal has a quadratic phase variation with time, so the instantaneous frequency varies linearly with time. The LFM signals are characterized so that the ratio between main-lobe amplitude and sidelobe level is large.

In LFM pulse compression, frequency modulation is realized for the modulation part in pulse Doppler radars. Usually, FM signal has low power when constructed first and then is amplified by a power amplifier. During the transmission, the modulated signal is sent as pulses with duration T and constant amplitude A. Through the pulse duration, the frequency of the signal changes linearly. If there is a linear frequency increase with time, this is called up-chirp. On the contrary, if there is a linear frequency decrease with time, this is called down-chirp. The increase and decrease in frequency can be achieved by either continuous frequency changes or discrete frequency steps. The signal that is modulated in frequency is passed through the pulse compression filter at the receiver side.

In the existence of Doppler effect, the frequency of the received signal is shifted. Shift in the received signal leads to an error in range indication which is known as range Doppler coupling. When range Doppler coupling is so large that can not be tolerated by the system, the average of the up-chirp and down-chirp FM is used for avoiding the unfavorable effects of Doppler shift.

3.4.1.2 Nonlinear Frequency Modulation (NLFM) pulse compression

Nonlinear Frequency-modulated pulse compression technique involves sweeping the carrier frequency of the transmit waveform in a nonlinear fashion. The advantage of the nonlinear FM over linear FM is that time sidelobes of the MF output are low while using a constant amplitude waveform and an ideal MF. Since the envelope of the signal is constant, power efficiency is high. Moreover, nonlinear FM does not have to deal with mismatched filters that result in SNR loss because of low time sidelobes.

Nonlinear change in frequency is similar to the amplitude shaping in LFM. For instance, reducing the amplitude of the frequency response of the system means spending less time over some part of the spectrum. If the frequency of the waveform increases during one half of the pulse and decreases during the other half of the pulse or vice versa, then this type of nonlinear FM is called symmetrical nonlinear FM. In nonsymmetrical nonlinear FM, the waveform uses only one half of the symmetrical waveform. Unlike non-symmetrical one, symmetrical nonlinear FM does not tolerate Doppler effect because it is sensitive to Doppler shifts.

One disadvantage of the non-linear FM over LFM is the system complexity. Another disadvantage is the necessity for a separate FM modulation design for each type of pulse to achieve the required sidelobe level. Finally, it is important to mention that digital processing or SAW devices can be used to process nonlinear FM.

3.4.2 Phase Modulation

Phase modulated pulse compression is realized by simply dividing the uncoded pulse length to equal subpulses (chips) and then coding these subpulses with different phases. Phase modulated pulse compressed signals have high spectral sidelobes due to sharpness of the phase transition and amplitude rise time. Phase modulated pulse compression is mainly divided into two: first one is binary phase coded pulse compression and second one is polyphase coded pulse compression.

3.4.2.1 Binary phase (Biphase) coded pulse compression

In phase coding, the phase of the transmitter RF is shifted during the pulse width. In binary phase coding, the binary bits can determine whether modulating signal is shifted to an in-phase condition or 180^0 out of phase condition with respect to the reference. The code sequence c_m is obtained as:

$$c_m = e^{j\theta_m} \tag{3.1}$$

where $\theta_m = 0$ or π so that c_m takes values $+1$ or -1.

For purposes of pulse compression, the bandwidth of a long pulse during transmission can be increased by the phase change. If the number of subpulses obtained from a long pulse by dividing it equally is N then the output of a matched filter has a peak N times greater than that of the long pulse. Also, pulse compression ratio is equal to the number of subpulses N so the duration of the compressed pulse 1/N times the uncompressed one. When the duration of the long pulse is symbolized as T, the duration of the matched filter output becomes $2T - 1$. If the phases are randomly selected, the maximum power of sidelobe of the matched filter is about $2/N$ below the compressed pulse peak.

One family of known binary codes is the Barker code. Barker binary phase codes have the lowest possible sidelobe level. If low time sidelobes are needed after pulse compression, completely random selection of phases is not suitable. Therefore, usually a selection criteria is determined for phases. For instance, one criteria of selection of subpulse phases is equalizing the time sidelobes of the compressed pulse. The benefit of this type of selection is to form all the sidelobes equally by lowering the higher sidelobes. The $0, \pi$ biphase codes that have equal time sidelobes are the Barker codes. Due to this property, Barker codes are called perfect codes. The ratio between sidelobe level and the peak value of the MF output of the Barker code is $1/N$.

3.4.2.2 Polyphase Codes

[15] Polyphase codes are obtained by dividing a long pulse into N subpulses. Phases of these subpulses are not restricted to the two levels 0 and π. The phase of each subpulse is formed as M discrete phases and $M > 2$. The code sequence $c_n = e^{j\theta_n}$ and θ_n can be written as $\theta_n = (2\pi/M)n$ for $n = 1, 2, ., M$. Compared with the binary phase codes, polyphase codes produce lower time sidelobes. Moreover, they can tolerate Doppler effect up to moderate Doppler frequencies. On the basis of the derivation technique, polyphase codes can be separated into three main groups:

- Step-frequency derived (Frank and P1 codes)
- Butler matrix derived (P2 code)
- Linear frequency derived (P3 and P4 codes)
- Randomly generated.

Frank codes is described by an $M \times M$ matrix [16] given below. In the matrix, each number is multiplied by an equal amount of phase $2\pi/M$. The pulse compression ratio of the Frank codes is the same as the number of subpulses that is $M^2 = N$. because of their structural form, Frank codes have been described as stepped phase linear FM. Frank codes have lower sidelobe levels compared to biphase codes.

0	0	0	.	.	0
0	1	2	.	.	$(M-1)$
0	2	4	.	.	$2(M-1)$
.					
.					
0	$(M-1)$	$2(M-1)$	.	.	$(M-1)^2$

Like Frank codes, P1 and P2 codes have both low sidelobes and easy implementation. Moreover, similar to Frank codes, P1 and P2 codes tolerate Doppler shifts well. The representation of N-element P1 and P2 codes can be given respectively as [15]

$$\theta_{k,l} = -(\pi/M)[M-(2l-1)][M(l-1)+(k-1)], \tag{3.2}$$

$$\theta_{k,l} = (\pi/2M)[M+(1-2k)][M+(1-2l)], \tag{3.3}$$

where k and l are integers ranging from 1 to M.

Linear frequency modulation is used to derive P3 and P4 codes. While compared to P1 and P2 codes, P3 and P4 codes have better performance in tolerating the Doppler effect. Actually, P4 code is a rearranged version of the P3 code. They have the same Doppler tolerance and they can tolerate bandwidth limitation prior to pulse compression. P3 and P4 codes are given by [17, 18]

$$\theta_k = (\pi/N)(k-1)^2, \tag{3.4}$$

$$\theta_k = (\pi/N)(k-1)^2 - \pi(k-1), \tag{3.5}$$

where k is an integer ranging from 1 to N. By converting linear frequency modulated waveform to baseband, P4 code is generated. This conversion is realized by using a local oscillator on one end of the sweep and sampling the waveform at the Nyquist rate. The generation of the P3 code is very similar to the P4 code except for the change in local oscillator frequency of the P4 code in order to solve bandwidth problem prior to pulse compression. The largest phase increments between code elements for P3 code are in the middle whereas those for P4 code are on two ends of the code. That is why P4 code has better performance in limitation of the bandwidth before the pulse compression.

3.5 Pulse Compression Effects

The peak value of the pulse compression process output depends on the energy of the input signal but not on the shape of the signal. The duration of the sidelobes of the compressed pulse is two times the duration of the uncompressed signal. The uncompressed input signal determines the sidelobe nature of the pulse compressed signal.

A major drawback to the pulse compression is appearance of range sidelobes around the main signal peak which leads a smearing of the return signals in range and introduces range ambiguities. In Figure 3.3 , the existence of a small target may not be inferred from the matched filter output when there is a small target and a large target whose power is 10dB larger than the small one. Although the small target is noticable when it is the only present target in the environment, in the existence of the large target the small target is masked by the range sidelobes of the large target.

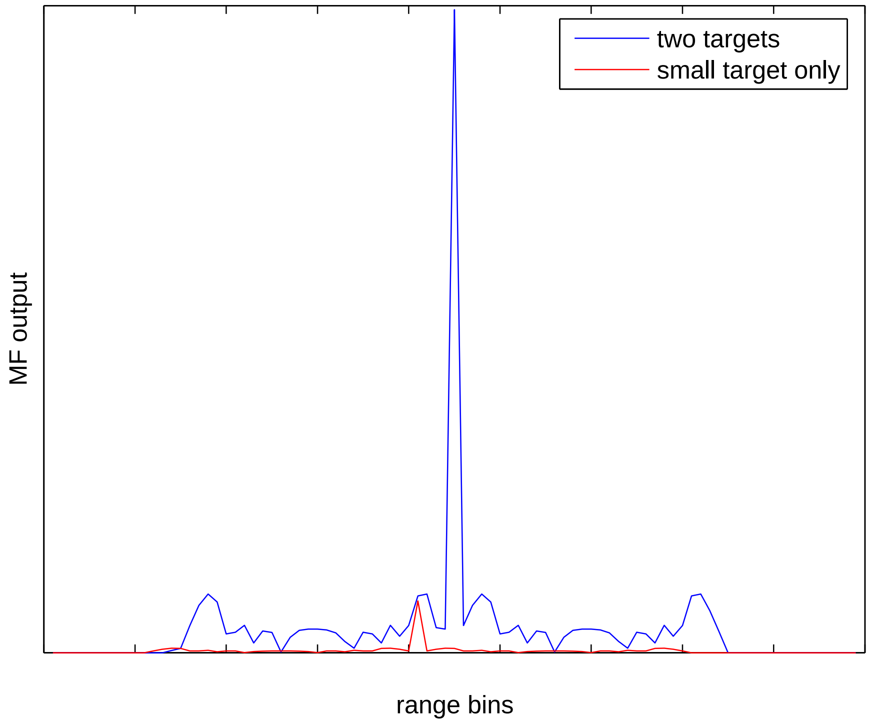

Figure 3.3: Matched filter output of received radar signal is examined when there is only a small target in the environment and when except from the small target there is also a big target whose power is 10dB greater than the small one.

It is possible that large sidelobes can result in detecting spurious targets, that is sidelobes can be mistaken as real targets. Since high sidelobes of the bigger targets can mask nearby smaller targets, suppression of range sidelobes is critical, especially in applications with multiple target systems. This effect is tried to be minimized by using carefully chosen pairs of codes or by amplitude weighting the long pulse over its duration. In general, it is not very easy to design codes with very low sidelobes. Moreover, it may not be efficient to use amplitude weighting in respect of power efficiency. For instance, if amplitude modulation is realized, it is obligatory to use linear amplifiers such as Class A amplifiers. On the other hand, because of their high power efficiency Class C amplifiers which are non-linear amplifiers are preferred in radar systems. Due to the fact that the problems resulting from large time sidelobes can not be overcome efficiently and/or sufficiently by the variations on pulse coding, it is necessary to suppress the range time sidelobes after the pulse compression. The most widely used techniques for sidelobe reduction are LS and MMSE methods which will be explained in the next two sections.

3.6 Least Squares (LS) Method

Least squares method is a mathematical optimization technique in order to estimate a value or function with the highest probability from the observations with random errors. Actually, it attempts to find the best fit function that closely approximates the given data. The least squares method minimizes the sum of square of residuals which are defined as the differences between the observation and estimated values of a function. The cost function J of the LS method for deterministic linear algebraic equations is given by [19]

$$J = \sum_{k=1}^{N} |\hat{x}_{LS}(k) - x(k)|^2, \tag{3.6}$$

where $\hat{x}_{LS}$ is the estimate, x is the observed data, and N is an integer that defines the length of the data. For radar applications where the target signals are not deterministic and noise is present in the environment, LS method tries to minimize the expected value of the square of the difference between estimated and true radar impulse response. LS estimation is optimum in the mean squared error sense for an unbiased estimation [20]. An implicit requirement for the least squares method to work is that errors in each measurement be randomly distributed.

Due to appearance of range sidelobes around the main signal peak after the pulse compression, range sidelobes of the bigger targets can mask nearby smaller targets. LS method can remedy this problem with success.

In order to understand the LS solution for range sidelobes in radar applications, it is beneficial to model the radar signals in their matrix form. The N-length transmitted radar signal is denoted by $\mathbf{s} = [s_1 s_2 ... s_N]^T$ and $\mathbf{h} = [h(0)h(1)...h(L-1)]^T$ is the L-length true radar impulse response where $(\bullet)^T$ represents the transpose matrix operation. The received return radar signal can now be expressed as [22]

$$y = S\mathbf{h} + w \tag{3.7}$$

where $w = [w(0)w(1)...w(L+N-2)]$ are the additive noise samples and $(L+N-1) \times L$ matrix S is the convolution matrix for $\mathbf{s}$ that is shown below [22]

$$S = \begin{pmatrix} s_1 & 0 & 0 & . & . & . & . & . & 0 \\ s_2 & s_1 & 0 & & & & & & 0 \\ . & s_2 & s_1 & & . & & & & . \\ . & . & s_2 & & . & & . & & . \\ s_N & . & . & & . & & . & & 0 \\ 0 & s_N & . & & & & . & & s_1 \\ 0 & 0 & s_N & & & & & & s_2 \\ . & 0 & & & . & & & & . \\ . & . & & & & & . & & . \\ 0 & 0 & . & . & . & . & . & 0 & s_N \end{pmatrix}. \tag{3.8}$$

Actually, the product $S\mathbf{h}$ stands for the discrete baseband equivalent of the convolution of the transmitted signal and the true radar impulse response. The general form of the LS solution for radar applications in order to solve the sidelobe masking problem can be given as [22]

$$\hat{h}_{LS} = (S^H S)^{-1} S^H y, \tag{3.9}$$

where $(\bullet)^H$ denotes the complex conjugate transpose or Hermitian operation.

In [21, 23, 1, 24, 25], LS based methods are proposed to solve the masking problem of weaker targets by stronger ones due to range sidelobes of pulse compression codes for radar applications. In [23], low sidelobe pulse compression method has been introduced based on a least squares solution to the deconvolution problem. Since it is the well-known optimal estimation method, it has been used in scientific works as a reference comparison method so as to measure the performance of other sidelobe suppression methods [22, 2]. That's why, we prefer to compare the performance of our method with that of LS method.

Some modified versions of the LS algorithms are also used in radar systems for sidelobe suppression [26, 27, 28]. For instance, in [26] a type of recursive least squares method that is efficient in self clutter suppression for phase-coded radar signal is analyzed. Moreover, LS method is applied for formulating parameters, calibration, and curve fitting.

Despite being the optimum unbiased estimation method in the mean square error sense, the LS method has some disadvantages. As in [1, 23], LS estimation is performed by multiplying the received signal in the vector form by matrices obtained from code auto-correlation. Inversion of an $n \times n$ matrix is necessary in this method where n is the length of the received signal. The multiplication operation has an order of $O(n^2)$ and thus it is prohibitively computationally intensive. Hence, the greater the number of samples taken for the received return signal is, the more computationally complex the LS algorithm becomes. Additionally, in the existence of Doppler effect, fast Fourier transform (FFT) has to be applied directly to the received return signal since MF operation takes place in LS. In order to determine which FFT bins to check for possible targets, it is necessary to check all FFT filters since MF is not applied prior to FFT.

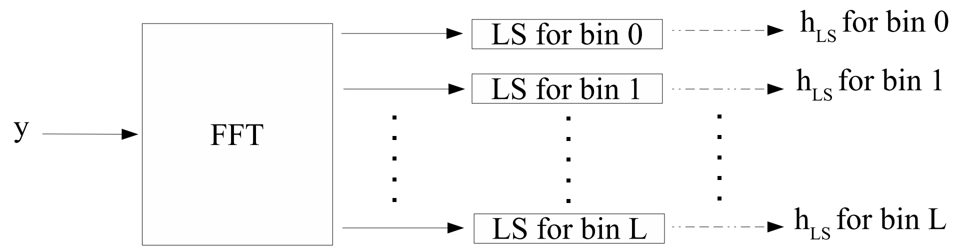

Figure 3.4: Block diagram of LS algorithm in the existence of Doppler.

3.7 Minimum Mean Squared Error (MMSE) Method

The MMSE estimate of a parameter h is an estimate which minimizes the cost function

$$J = E[|(\hat{h}_{MMSE} - h)|^2] \tag{3.10}$$

where $\hat{h}_{MMSE}$ is the estimate. The main difference between LS and MMSE algorithms is that the latter needs statistics, namely the auto-correlation functions or matrices, for the observed and estimated data. Additionally, MMSE method is optimum in the mean squared error sense for a biased estimation whereas LS method is optimum in the mean squared error sense for an unbiased estimation.

There are MMSE based filters used in radar signal processing such as Kalman adaptive filter [29] and Wiener filter [30]. Kalman adaptive filter computes the optimal linear minimum mean square errors of the finite impulse response filter coefficients. On the other hand, Wiener filter reduces the sidelobe levels.

MMSE method is also used in radar applications after pulse compression in order to improve range sidelobe problem [22, 2, 31]. In this case, h and $\hat{h}_{MMSE}$ in eqn. (3.10) become true radar impulse response and estimated true radar impulse response, respectively. The estimate of the MMSE algorithm $\hat{h}_{MMSE}$ can be assessed as a multiplication of some form of the received return samples and the complex conjugate of a matrix. Therefore, the cost function J can be rewritten as [22]

$$J = E[|(z^H(k)\tilde{y}(k) - h)|^2] \tag{3.11}$$

where [22]

$$\tilde{y}(k) = [A(k)]^T \mathbf{s} + w(k). \tag{3.12}$$

In eqn. (3.12), $\mathbf{s}$ and $w(k)$ are transmitted radar samples and additive noise defined in the previous section and $[A(k)]$ is the matrix consisting of radar impulse response coefficients given below [22]

$$[A(k)] = \begin{pmatrix} h(k) & h(k+1) & . & . & h(k+N+1) \\ h(k-1) & h(k) & & & . \\ . & & . & & . \\ . & & & . & h(k+1) \\ h(k-N+1) & . & . & h(k-1) & h(k) \end{pmatrix}. \tag{3.13}$$

By differentiating (3.11) with respect to $z^H(k)$ and then equating the result to zero, the MMSE filter $z(k)$ takes the form [22]

$$z(k) = (E[\tilde{y}(k)\tilde{y}^H(k)])^{-1}E[\tilde{y}(k)h^H(k)]. \tag{3.14}$$

It is easily deduced from (3.14) that it is needed to know the average statistics of the radar impulse response. This is actually unknown in radar applications since that information is actually what is to be obtained in a radar. In [22], these average statistics are obtained from the observed instance by

$$E[|h(k)|^2] = |\frac{m(k)}{|\mathbf{s}^2|}|^2, \tag{3.15}$$

where the matched filter output ($m(k)$) in (3.15) can be expressed as [22]

$$m(k) = \mathbf{s}^H\tilde{y}(k). \tag{3.16}$$

Although mot optimum in any sense, by substituting (3.15) and (3.16) into (3.14), an MMSE-like estimation of the radar impulse response is possible.

In [31], in order to prevent masking of weak signals, an MMSE estimator is used to suppress the range sidelobes. The MMSE estimator calculates the optimum linear weighting by using information extracted from the given observation. In [2], MMSE algorithm is run recursively. While running the MMSE algorithm in [2], a filter is generated by using the MF output first to obtain target signal power estimates to be used as average statistics. A signal covariance matrix is obtained for each range bin to be inspected from target power estimates. Another filter is constructed based on the inverse of the covariance matrix and acts on the MF signal. Smaller targets are identified by checking the filtered MF signal. At each iteration of the algorithm, target powers are reestimated and filters are regenerated. A modified version of [2] is investigated as another reiterative MMSE method in [22].

The main disadvantage of the MMSE algorithm is its complexity. For instance in [2], target power estimates and filters are newly formed at each iteration of the algorithm. Therefore, as the number of iterations increase the algorithm becomes more complex. Moreover, the algorithm is computationally complex due to on-line matrix inversion at every iteration. Furthermore, the estimate covariance matrix is reportedly ill-conditioned and thus ad hoc methods are used to avoid numerical instability in running the algorithm [2].

CHAPTER 4

Alternative Method for Sidelobe Suppression

4.1 CLEAN Method

Deconvolution methods, as the inverse operation of convolution, have been widely discussed in literature especially in image processing. In 1974, Högbom developed a nonlinear technique for image processing which was basically expressed as deconvolving the obtained image beam [32]. This technique was named as CLEAN method. After Högbom, various versions of the CLEAN method were developed and used for different applications. The CLEAN algorithm can be applied to Fourier transformed data (generally in astronomy and image processing) or MF output (mostly in radar systems). It is an effective method in eliminating the undesirable signals from the required data.

The CLEAN method is an iterative procedure widely used in image processing [3, 4] and astronomy [33, 5] as a beam removing deconvolution method. A similar method is called Successive Interference Cancelation in communication literature [6] and is usually employed for multi-user detection. Additionally, CLEAN algorithm was used for communication channel characterization problems. For instance, in [34] CLEAN algorithm processes ultra wideband measurements and provides estimates of time interval, angle of interval and waveform shape.

Except for a couple of works in recent years, the CLEAN algorithm has rarely been applied to detection in radar signal processing. Most recently, the CLEAN algorithm has been applied to an extended target or contiguous clutter scenario in [7]. In [7], two CLEAN based methods are investigated and performances of the two methods are compared. These two methods are the well-known basic CLEAN method and a modified one consisting of inner and outer loops. According to the performance analysis of the study, basic CLEAN algorithm is found suitable for discrete point targets and modified CLEAN algorithm performs well for contiguous scattering sources. In this study, it is stated that the modified CLEAN algorithm has better performance than the basic one. However, it is mentioned that computational complexity of the modified one is higher than the basic and the computational complexity of modified CLEAN algorithm depends on inner and outer loop parameters. The Doppler effect on the received return radar signal is assumed to be negligible or to be overcome before matched filtering. Hence, Doppler is not taken into account. Finally, performance of any of the CLEAN algorithms has not been evaluated in comparison to other methods.

In radar applications, the received signal y(k) is passed through the matched filter and then the CLEAN algorithm is applied to the matched filter output m(k). The representation of discrete time forms of the received signal and the MF output are respectively given below:

$$y(k) = h(k) * s(k) + w(k), \tag{4.1}$$

$$m(k) = y(k) * s^{*}(-k), \tag{4.2}$$

where $s(k)$ is the transmitted signal, $w(k)$ is the additive noise, and '$*$' denotes the convolution operation. The basic CLEAN algorithm is presented in its entirety in Table 4.1.

Table 4.1 CLEAN Algorithm

1. Obtain the MF output $m(k)$, initialize all estimates $h(k)$ to zero.

2. Find the peak in $|m(k)| > \tau$. If no such point exists, halt.

3. If the peak exceeds τ and its location is u so that $|m(u)|$ is the largest, estimate the target at $k = u$ by

$$h(u) = h(u) + \lambda \frac{m(u)}{r(0)}, (1 \geq \lambda > 0). \tag{4.3}$$

4. Cancel the effect of the most recently estimated target by

$$m(k) = m(k) - \lambda \frac{m(u)}{r(0)} r(k-u). \tag{4.4}$$

5. Go to step 2.

The threshold τ is important for detection of radar signals. When the threshold is small, weak radar signals can be detected. However, selecting a small τ allows the estimation of unwanted noise signals as a false radar signal. Therefore, the threshold has to be defined according to the noise level. For Gaussian noise, the threshold τ can be readily found as:

$$\tau = \sqrt{-N \cdot N_0 \cdot M \cdot \ln(FAR)} \tag{4.5}$$

where N is the length of the transmitted signal $s(k)$, N_0 is the noise power, M is the number of transmitted pulses, and FAR is the false alarm rate. Additionally, a small value of τ increases the process time of the algorithm since there will be more signals surpassing the threshold.

The parameter λ directly affects the computational complexity of the algorithm. If the parameter λ is very small, the number of iterations increases by a large amount because the peak value that exceeds the threshold is decreased by a small amount and has to be detected many times. On the other hand, when λ is close to 1, the estimated peak value times λ can be larger than the actual mainlobe peak of the target because of the additive effects of sidelobes of other target signals on the mainlobe of detected target.

4.2 Successive Target Cancelation (STC)

Successive target cancelation (STC), which is a CLEAN-based algorithm, can be used in sidelobe reduction as an alternative to LS and MMSE algorithms. In radar applications, it provides a solution to range sidelobe masking by obtaining the target's phasor from the MF output and canceling part of the signal due to the target. This process can be repeated for each detected target. Since this method successively detects a signal and cancels the signal from the overall MF output, we will refer to it as Successive Target Cancelation (STC).

In order to have a clear conception on the basic idea of the proposed method, it is beneficial to analyze the simplest scenario where only two targets exist both at zero Doppler frequency. In this analysis, the effect of noise will be taken into consideration whereas clutter is neglected. For simplicity, the value of λ is taken as 1. It is considered that the power difference between two targets is larger than the code sidelobes so that the weak target cannot be detected by simply checking the matched filter output.

Suppose we send a radar pulse with

$$s(t) = \sum_{p=0}^{L-1} a_p g(t - pT_c), \tag{4.6}$$

where L is the number of chips in the code, a_p's denote the code sequence, $g(\cdot)$ is taken to be the rectangular pulse in $[0, T_c]$, and T_c is the chip duration.

Assume we have two targets whose signals arrive k_1 and k_2 (both are assumed to be integers) chip durations later after we send a pulse. Thus, we have the following discrete-time impulse response observed by the receiver

$$h(k) = \alpha_1 \delta(k - k_1) + \alpha_2 \delta(k - k_2), \tag{4.7}$$

where α_1 and α_2 are the phasors of targets 1 and 2, respectively. Rather than using the continuous time notation for the signals, we will use the discrete time baseband equivalent model exemplified in (4.7) throughout the paper. The phasors are circularly symmetric zero-mean complex Gaussian random variables with high enough signal-to-noise ratios (SNR) for detection. So, target powers are quite above the noise level.

The received signal is given by

$$y(k) = h(k) * s(k) + w(k) = \alpha_1 s(k - k_1) + \alpha_2 s(k - k_2) + w(k), \tag{4.8}$$

where the front-end receiver filters and samples are constructed in a way that $w(k)$ is an additive white Gaussian noise (AWGN) with time index k. The AWGN is taken to be circularly symmetric complex Gaussian with mean 0 and variance N_0. The received signal is passed through a matched filter (MF). Neglecting the time shift for causality in MF, we obtain the MF output $m(k)$ by

$$m(k) = y(k) * s^*(-k) = \alpha_1 r(k - k_1) + \alpha_2 r(k - k_2) + n(k), \tag{4.9}$$

where $n(k) = w(k) * s^*(-k)$ and the autocorrelation of $s(k)$ is $r(k) = s(k) * s^*(-k)$.

Recalling that $|\alpha_1|^2 > |\alpha_2|^2$, the largest peak of $|m(k)|$ will be observed at $k = k_1$ with high probability since the noise power is very small compared to target powers. The MF output at $k = k_1$ is

$$m(k_1) = \alpha_1 r(0) + \alpha_2 r(k_1 - k_2) + n(k_1) \tag{4.10}$$

and there is a priori information on neither the powers nor the time shifts of targets. Then, the best estimate we can have for α_1 is

$$\hat{\alpha}_1 = \frac{m(k_1)}{r(0)} = \alpha_1 + \alpha_2 \frac{r(k_1 - k_2)}{r(0)} + \frac{n(k_1)}{r(0)}. \tag{4.11}$$

We thus obtain an estimate of α_1 which is biased. The mean of the estimate has an additive term affected by α_2, though its effect is through the sidelobes so that the code employed reduces this effect. The last term of the estimate of α_1 results from the noise in the environment and the noise effect is also reduced by the employed code.

We will now subtract the signal that is due to target 1 using its phasor estimate. This can be done both at the MF output or the video signal. Although the results are identical in the ideal case due to linearity, issues such as quantization is important in deciding which one is to be preferred in practice. We will perform cancelation at MF output in this study. Denote this new signal by $m'(k)$ such that

$$\begin{aligned} m'(k) &= m(k) - \hat{\alpha}_1 r(k - k_1) \\ &= \alpha_2 r(k - k_2) - \alpha_2 \frac{r(k_1 - k_2)}{r(0)} r(k - k_1) \\ &\quad + n(k) - \frac{n(k_1)}{r(0)} r(k - k_1). \end{aligned} \tag{4.12}$$

We now like to detect target 2 based on this updated MF signal. As seen in (4.12), the largest peak of $|m'(k)|$ will be at $k = k_2$ with high probability. At $k = k_2$,

$$m'(k_2) = \alpha_2 r(0) - \alpha_2 \frac{|r(k_1 - k_2)|^2}{r(0)} + n(k_2) - \frac{r(k_2 - k_1)}{r(0)} n(k_1) \tag{4.13}$$

so that the phasor of target 2 can be estimated as

$$\hat{\alpha}_2 = \frac{m'(k_2)}{r(0)} = \alpha_2 - \alpha_2 \frac{|r(k_1 - k_2)|^2}{r^2(0)} + \frac{1}{r(0)} n(k_2) - \frac{r(k_2 - k_1)}{r^2(0)} n(k_1). \tag{4.14}$$

There is no contribution from the first target in $\hat{\alpha}_2$ as shown in (4.14). The second term with α_2 has a squared multiplicative term and is very small compared to the first term when the maximum sidelobe levels of pulse compression codes used in practice are considered. The same argument holds for the second noise term. So, the signal-to-noise ratio (SNR) for α_2 approaches the SNR value $\frac{|\alpha_2|^2 r(0)}{N_0}$. If there was only target 2 in the system, we would have the same SNR for the estimate of its phasor. This suggests that the masking problem due to sidelobes is alleviated to a large extent with the method explained in this section.

At this point we have a good phasor estimate for target 2 and can subtract its effect from the MF output. The reason of such an iteration is the existence of a term related to α_2 in $\hat{\alpha}_1$. Though weaker compared to target 1, target 2 still has sidelobes and it affected the phasor estimate of target 1 as seen in (4.11). This overshooting may be corrected by another iteration of the algorithm. When target 2 is canceled from the MF output currently held by the algorithm, the following is obtained:

$$\begin{aligned} m''(k) &= m'(k) - \hat{\alpha}_2 r(k-k_2) \\ &= -\frac{r(k_1-k_2)}{r(0)}\alpha_2 r(k-k_1) + \frac{|r(k_1-k_2)|^2}{r^2(0)}\alpha_2 r(k-k_2) \\ &\quad + n(k) - \frac{n(k_1)}{r(0)} r(k-k_1) - \frac{n(k_2)}{r(0)} r(k-k_2) \\ &\quad + \frac{r(k_2-k_1)n(k_1)}{r^2(0)} r(k-k_2). \end{aligned} \tag{4.15}$$

Once again, $|m''(k)|$ has its peak at $k = k_1$ with high probability and we have the chance to estimate the overshoot term by $\frac{m''(k_1)}{r(0)}$. We will now add this overshoot term to the current estimate for α_1:

$$\begin{aligned} \hat{\alpha}_1' &= \hat{\alpha}_1 + \frac{m''(k_1)}{r(0)} \\ &= \alpha_1 + \frac{r(k_1 - k_2)|r(k_1 - k_2)|^2}{r^3(0)}\alpha_2 + \left[1 + \frac{|r(k_1 - k_2)|^2}{r^2(0)}\right] \cdot \frac{n(k_1)}{r(0)} \\ &\quad - \frac{r(k_1 - k_2)}{r(0)} \cdot \frac{n(k_2)}{r(0)}. \end{aligned} \tag{4.16}$$

When this new estimate is carefully examined and compared to (4.11), it can be seen that the effect of target 2 is now reduced by the $\frac{|r(k_1-k_2)|^2}{r^2(0)}$ term whereas the noise terms almost have the same variance.

The overshoot terms added at later iterations act as correction terms and enhance the estimation. The iterations can be repeated as many times as desired. However, the estimates will be very accurate after some time and the correction terms will be only due to noise. That's why we suggest only the correction terms which can pass a certain threshold be taken into consideration. This also gives way to a stopping criterion for the algorithm. In this study, we take the threshold to be equal to the constant false alarm rate (FAR) threshold τ for the noise at the MF output for a chosen FAR.

If a target has non-zero Doppler then the received signal is distorted due to frequency shift. This distortion has to be taken into consideration. In the STC method, after the MF output is obtained from the received return signal, we assume for simplicity that fast Fourier transform (FFT) is applied to the MF output for Doppler processing. The STC algorithm can be run for each FFT bin.

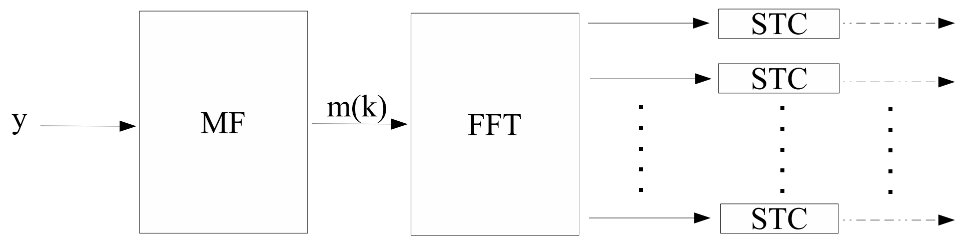

Figure 4.1: Block diagram of STC algorithm in the existence of Doppler.

By checking the maximum absolute signal in a filter, it can be determined whether at least a single target is present in that filter. If at least one target exists, then the STC algorithm is run for that filter. In the STC method, the distortion due to Doppler is taken into account by modifying the auto-correlation $c(k)$. In this case, the auto-correlation is taken as

$$c(k) = (e^{j2\pi PRF.T_c.\frac{b-1}{K}(0:K-1)} \otimes s(k)) * s^*(-k), \tag{4.17}$$

where b corresponds to the FFT bin number, PRF is the pulse repetition frequency, K is the number of pulses in a burst, and $\otimes$ denotes the Kronecker product. This modification stems from the fact that the received signal due to a target with non-zero Doppler is shifted in frequency. For the b^{th} FFT filter, it is hypothesized that the Doppler shift equals $(b-1)\frac{PRF.T_c}{K}$.

4.3 Simulation Results

In this chapter the performance of the STC algorithm is investigated with Monte Carlo simulations. We compare the performance of our method in comparison to the LS method which is optimum in the least squared sense. Moreover, the performance of STC method is examined for different values of the parameter λ. To demonstrate the performance of the STC algorithm, we investigate mainly two cases. First one is the multiple targets case. In this case, there are many targets (more than two) in the environment. The aim of this scenario is to estimate the true radar impulse response as close to the actual as possible. Therefore, we pay attention to the difference between estimated and true radar impulse response. The second case is the two targets scenario consisting of a bigger target and a smaller target that is masked by the range sidelobes of the bigger one. In this scenario, we try to demonstrate that our method can eliminate sidelobe masking of the bigger target on the smaller target and the smaller target can thus be detected in the existence of bigger target. Thus, probability of detection (P_d) of the small target versus SNR curves of STC and LS methods are compared. In addition to this, another comparison of P_d-SNR curves of the small target is done when two targets exist and when only small target exists. This comparison will clarify whether STC solves the sidelobe masking problem completely or not.

In both cases the transmitted waveform is the $P4$ code [35] of length $N = 30$ and the noise is modeled as a white circularly symmetric complex Gaussian noise with power spectral density N_0. The length of the environment response is 90 chips. In all of the simulations, the fast Fourier transform (FFT) in its plain form, i.e., no windowing, is applied to the matched filter output to account for Doppler processing. The number of transmitted pulses is 8, so the number of FFT filters is also 8. The pulse repetition frequency (PRF) is 8000Hz.

4.3.1 Multiple Targets Case

In this case, we investigate the relation between the target density and the residual error power. Residual error power (RP) is defined as the average power of the difference signal between the true radar impulse response and the estimated radar impulse response as given below:

$$RP = \frac{1}{L}\sum_{k=1}^{L} |h(k) - \hat{h}(k)|^2 \tag{4.18}$$

where L is the length of the environment. We use the target density parameter to indicate the number of targets in a given environment. Mathematical representation of the target density (TD) is given by

$$TD = \frac{N_T}{L}, \tag{4.19}$$

where N_T is the number of targets in the environment. Through this scenario, we can investigate the performance at the zero Doppler filter where many impulses due to clutter, which can be considered as targets around zero velocity, are expected. Moreover, multiple targets at other filters can be inspected within this scenario as well. When the target density is zero, there is no target in the environment. When it is 1, there exists a target signal in each range bin. When it is $\frac{N_T}{L}$, a randomly chosen set of N_T bins out of L are filled with targets. The target signals are modeled as circularly symmetric zero-mean complex Gaussian random variables with variance 1. The noise variance $N_0 = 10^{-4}$ and the threshold τ in STC is set for a false alarm rate (FAR) of 0.1. The noise and target signals are formed 1000 times randomly for each target density value. Multiple targets case is investigated for both the zero and non-zero Doppler. In the existence of Doppler, three intervals are defined for the Doppler frequencies on targets. Multiple targets case is simulated for these three different Doppler frequency intervals: -200-200Hz, -500-500Hz, and 1500-2500Hz. For the first two and no Doppler, first one of the eight FFT filters, which operates in between $-500(7500)$-500Hz, is used. For the last one, third FFT filter operating between 1500-2500Hz is used.

4.3.1.1 No Doppler on Targets

The residual error performance in this case is depicted in Figure 4.2. When there is no Doppler on targets, performance of the STC method varies according to the parameter λ. Generally, the residual error power of the STC method increases as λ decreases. On the other hand, for the target density values smaller than 0.2, higher λ leads to bigger residual errors. But the differences in the residual error powers for $\lambda < 0.2$ are very small. Therefore, it can be concluded that the residual error performance of STC is best for $\lambda = 1$ at zero Doppler. The performance of the STC method for $\lambda = 1$ leads LS method approximately up to target density of around 0.35 as seen in Figure 4.2. For the values of target density above 0.35, LS method is obviously better than the STC method. However, it is highly unlikely that more than half of the range bins at a given direction are filled with the target signals even for the zero-Doppler filter.

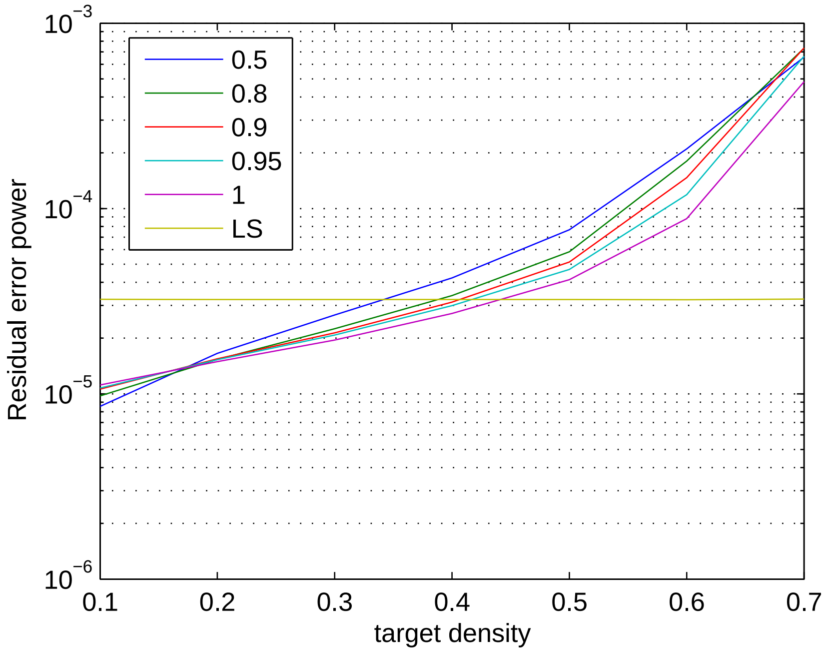

Figure 4.2: Residual error performances of LS method and STC method for 5 different λ values with respect to target density, zero Doppler, $\lambda = 0.5$, 0.8, 0.9, 0.95, and 1.

Actually, the residual error performance of any method can not be better than that of the LS method because residual error power is equivalent to the cost function of the LS method and LS method is optimum in the mean squared error sense for an unbiased estimation. However, the missing point is that our method STC obtains a biased estimate. Therefore, residual error performance of the STC method can be better than that of the LS method. Figure 4.3 demonstrates the biasedness of the STC algorithm. In this figure, STC algorithm is run 50, 250, and 10000 times for the same 9 targets in the environment and then averaged. Target signals with zero Doppler are deterministic and target locations are constant. The difference, namely the error, between the true radar impulse response and the mean of the estimated impulse response. The real and imaginary parts of the error are drawn. As seen in Figure 4.3, as the number of Monte Carlo runs used in averaging increases the random errors tend to diminish in most range bins. However, at bins where there are targets present the expected value of the error does not change. Hence, we can conclude that STC is a biased estimation algorithm.

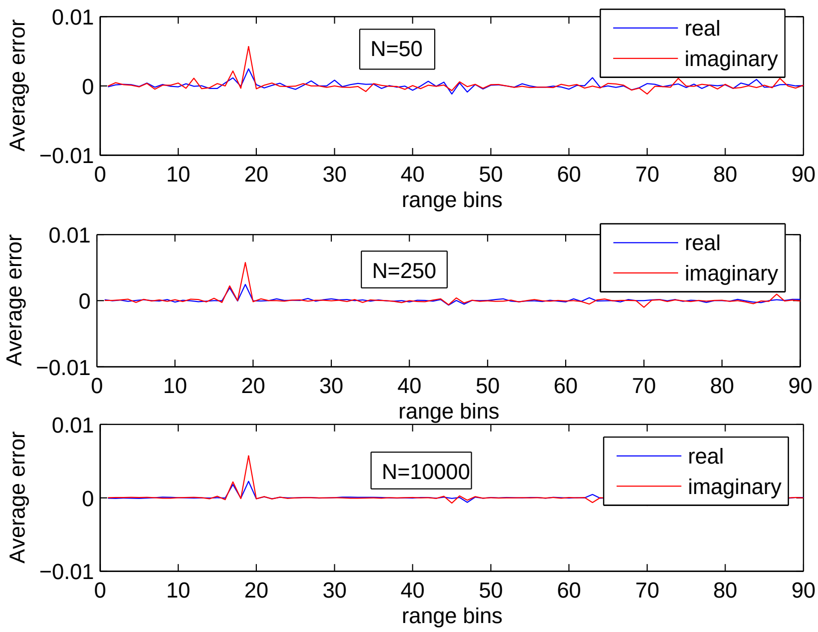

Figure 4.3: Real and imaginary parts of the average error of the STC method for $N = 50$, 250, and 10000.

In Figure 4.4, number of iterations in STC algorithm versus target density is examined for 5 different λ values. This figure gives us an idea about how the complexity of STC algorithm changes according to the parameter λ. As expected, number of iterations in STC increases when there are more targets in the environment (at high target density values). Moreover, decreasing λ results in higher number of iterations. When λ is small, comparatively small part of the found signal is canceled out from the matched filter output.

Canceling some part of the found signal instead of the whole of it can lead to false alarms at high target densities where sidelobes effect on the mainlobe of a target is more dominant. Therefore, the increase in the number of iterations for small λ values can result from detecting new targets and/or redetecting the same targets many times. We can understand the factor that leads to higher number of iteration for smaller λ values from Figures 4.5 and 4.6.

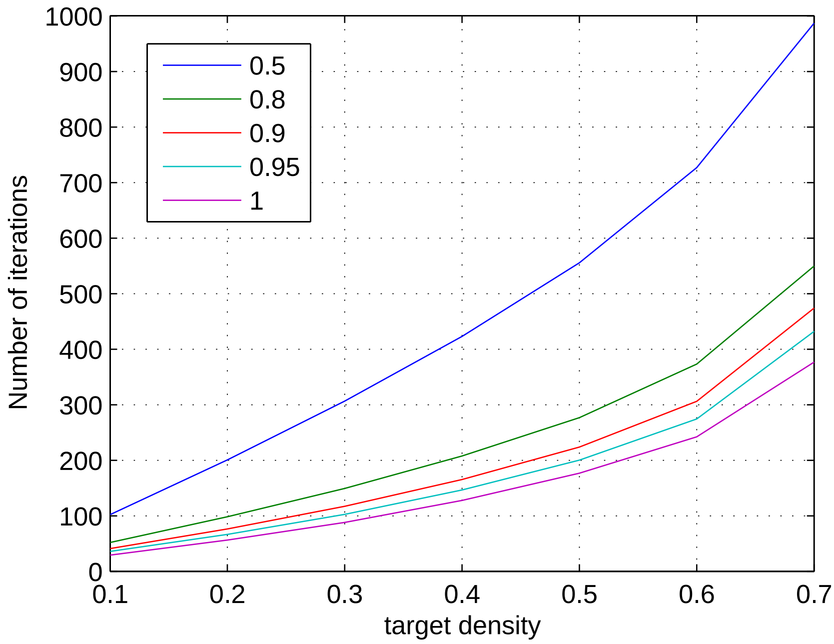

Figure 4.4: Number of iterations of STC method for $\lambda = 0.5, 0.8, 0.9, 0.95$, and 1 with respect to target density, zero Doppler.

In Figure 4.5, number of discrete targets that are detected is shown with respect to the target density. Although the number of discrete detected targets increases with decreasing λ, the results are very close for all λ's.

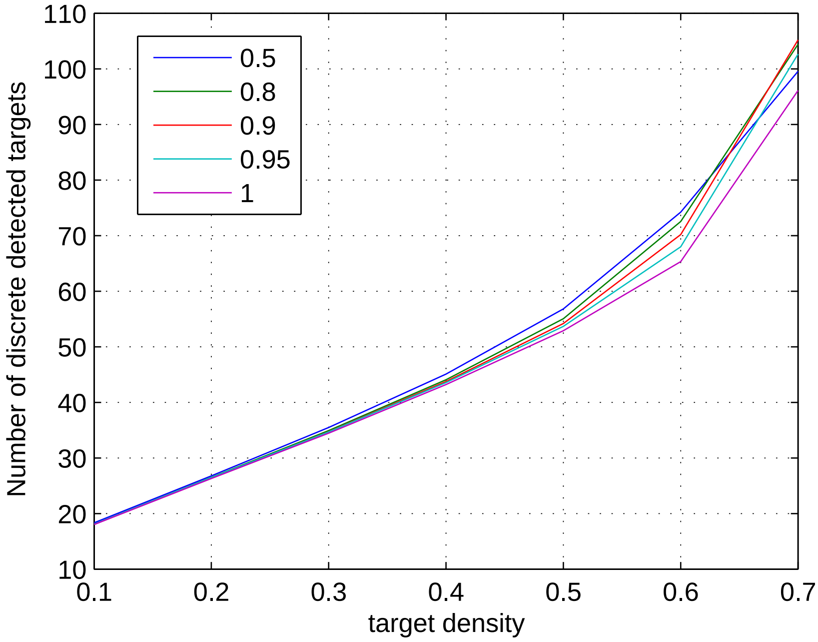

Figure 4.5: Number of discrete target detections of STC method for $\lambda = 0.5, 0.8, 0.9, 0.95$, and 1 with respect to target density, zero Doppler.

However, numbers of target redetections (detecting the same target) are considerably different as seen in Figure 4.6. As λ decreases number of target redetections increases. Therefore, increase in the number of iterations with decreasing λ mainly results from the redetected target iterations. It is crucial to investigate the detection probability of false alarms which has to be minimized especially in radar applications. At first glance, it seems that the difference between detected number of discrete targets and number of targets in the environment gives us the number of false alarms. However, it is not true when all the targets in the environment are not detected. Therefore, number of false alarms is observed in another simulation.

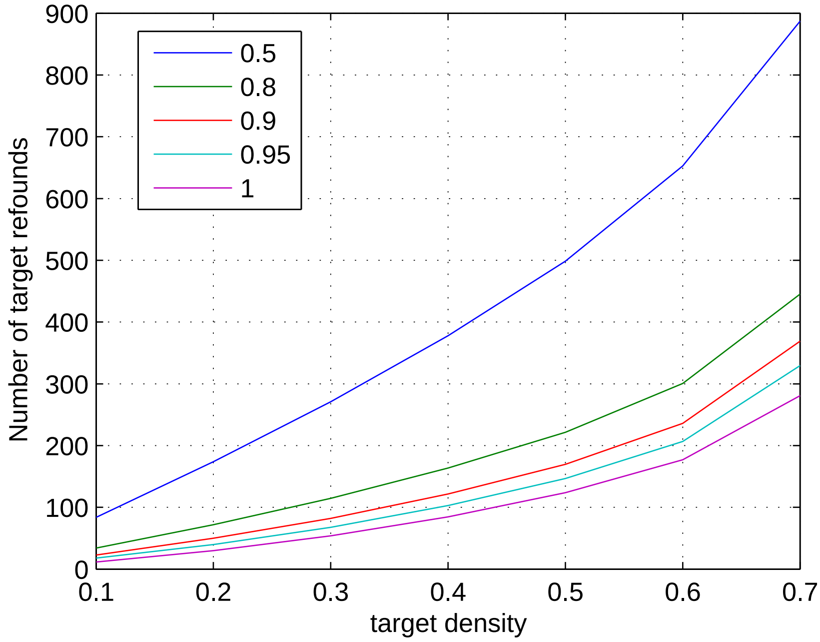

Figure 4.6: Number of target redetections of STC method for $\lambda = 0.5, 0.8, 0.9, 0.95$, and 1 with respect to target density, zero Doppler.

In Figure 4.7, the probability of false alarm (P_{FA}) performance of STC algorithm for $\lambda = 0.5, 0.8, 0.9, 0.95$, and 1 is examined. When λ gets smaller, P_{FA} increases. As it is mentioned earlier, range sidelobes leads to detect spurious targets. Therefore, due to increasing sidelobe effect, the probability of false alarm of STC increases with increasing target density.

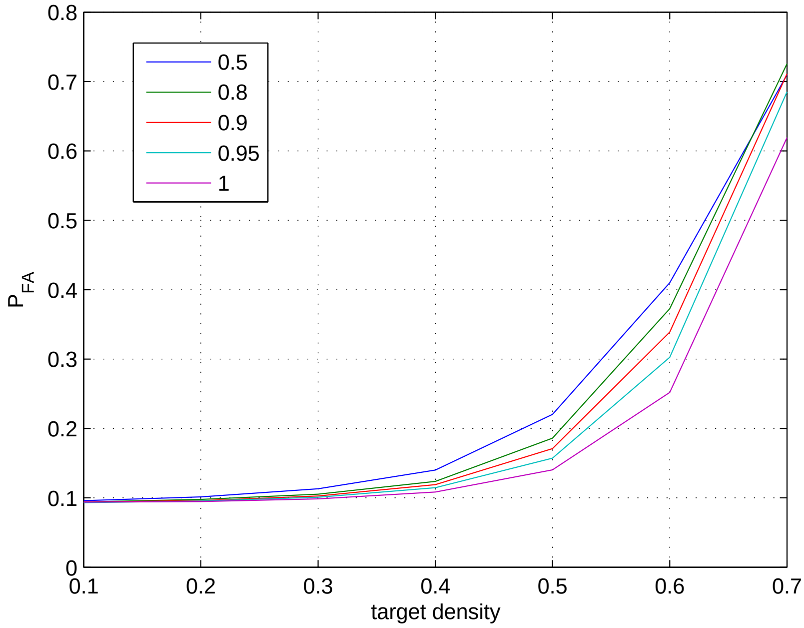

Figure 4.7: Probability of false alarm performances of STC method for $\lambda = 0.5, 0.8, 0.9, 0.95$, and 1 with respect to target density, zero Doppler.

4.3.1.2 Doppler on Targets

When there is Doppler on targets, STC performance comes even closer to LS performance. As seen in Figures 4.8 and 4.9, when the Doppler frequency increases, the performances of both methods deteriorate. However, the performance degradation in STC is less than that in LS. The degradation in LS method increases with increasing target density because as the target density increases, Doppler effect on the range bins of the radar received return samples increases. Up to some target density value, STC algorithm with smaller λ's has slightly better performance. This target density value gets bigger as the Doppler frequency increases. It can be concluded that with respect to the residual error power performance, STC can be preferred against LS when the targets do not occupy more than half of the range bins and the effect of variations in the parameter λ on the residual error performance loses its importance with improving Doppler.

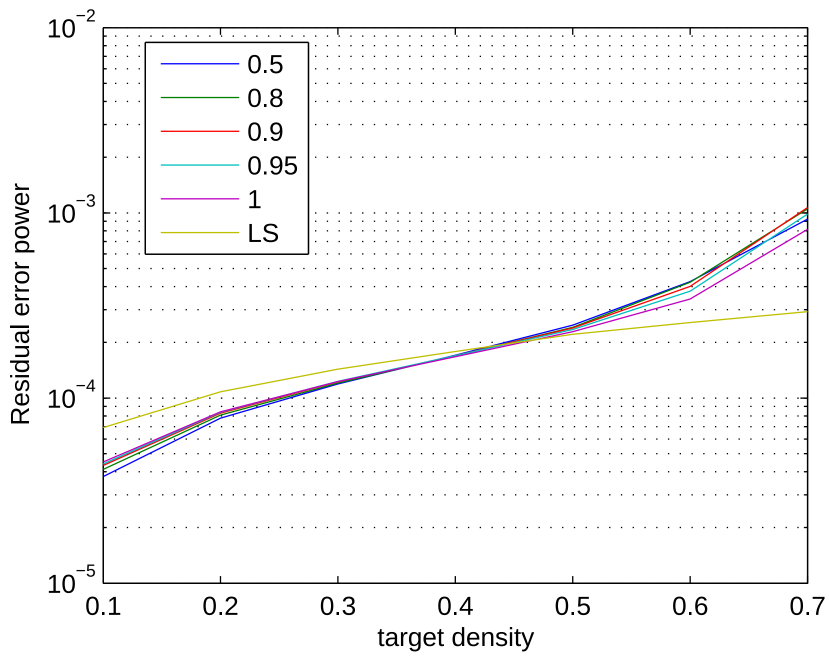

Figure 4.8: Residual error performances of LS method and STC method for $\lambda = 0.5, 0.8, 0.9, 0.95$, and 1 with respect to target density, Doppler randomly distributed with the uniform density between -200 and 200Hz.

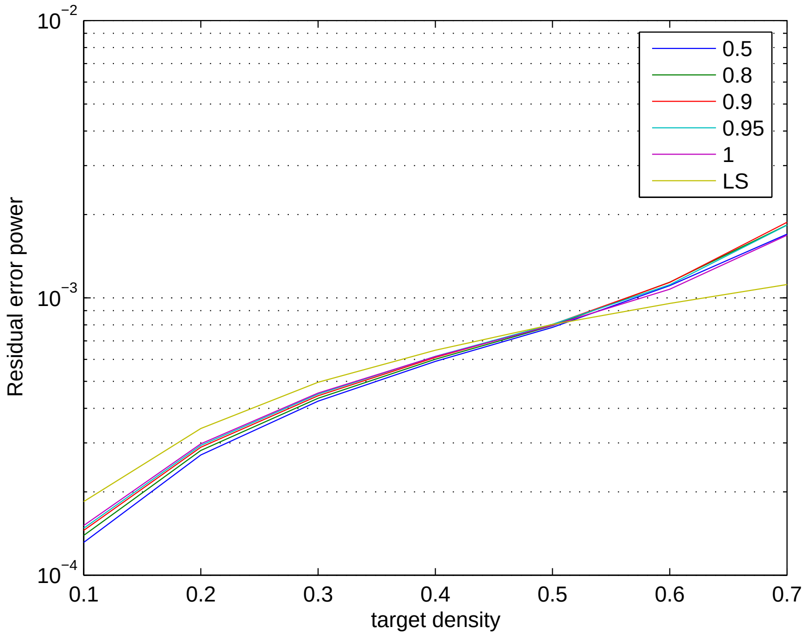

Figure 4.9: Residual error performances of LS method and STC method for $\lambda = 0.5, 0.8, 0.9, 0.95$, and 1 with respect to target density, Doppler randomly distributed with the uniform density between -500 and 500Hz.

In the existence of Doppler, the number of iterations still increases with both increasing target density and decreasing λ as seen in Figure 4.10.

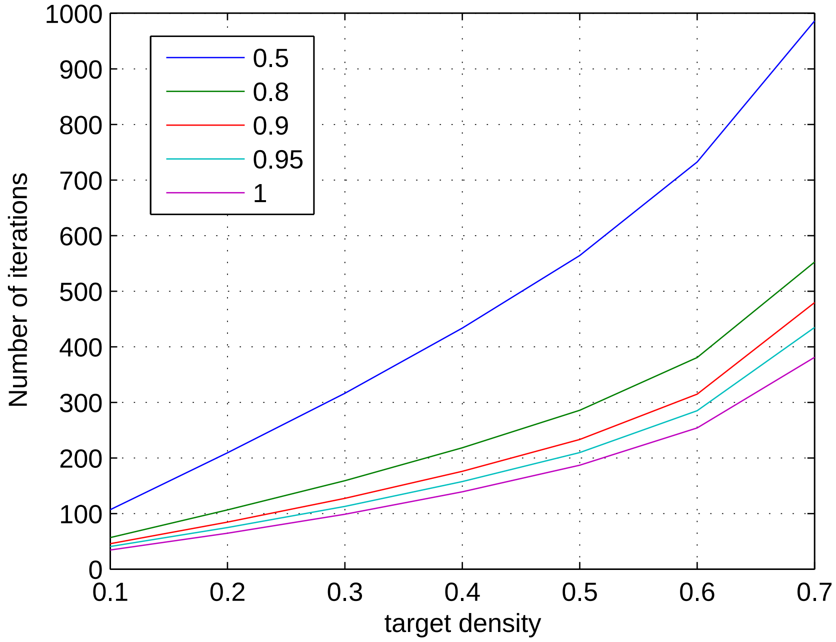

Figure 4.10: Number of iterations of STC method for $\lambda = 0.5, 0.8, 0.9, 0.95,$ and 1 with respect to target density, Doppler randomly distributed with the uniform density between -200 and 200Hz.

Moreover, for lower λ values both the number of discrete target detections (Figure 4.11) and the number of target redetections (Figure 4.12) increase when there is Doppler. As the Doppler increases, there can be seen an increase in the number of iterations for the same λ values.

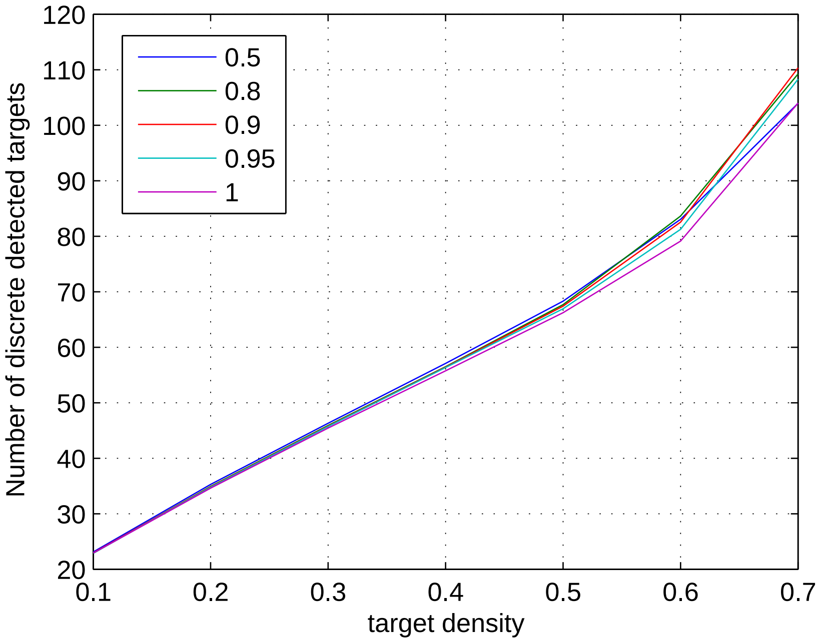

Figure 4.11: Number of discrete target detections of STC method for $\lambda = 0.5, 0.8, 0.9, 0.95,$ and 1 with respect to target density, Doppler randomly distributed with the uniform density between -200 and 200Hz.

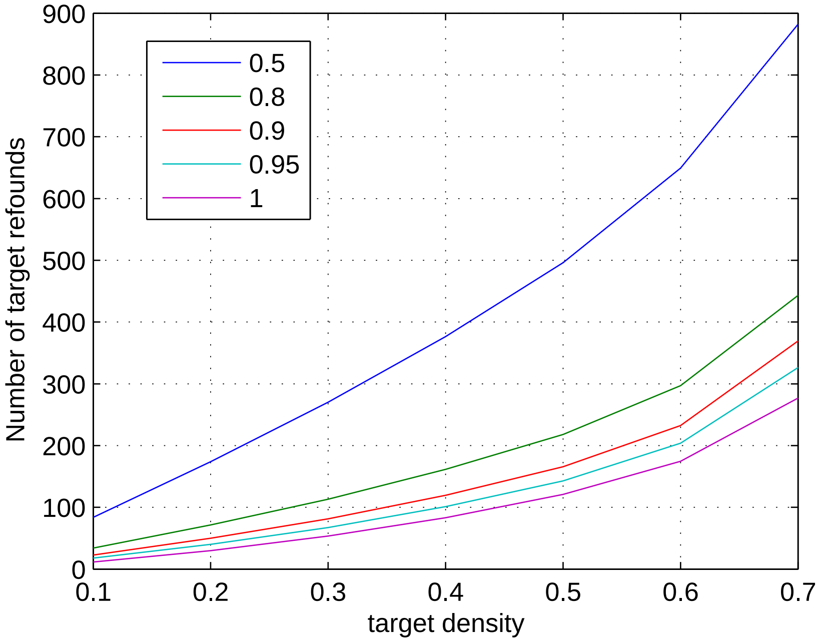

Figure 4.12: Number of target redetections of STC method for $\lambda = 0.5, 0.8, 0.9, 0.95,$ and 1 with respect to target density, Doppler randomly distributed with the uniform density between -200 and 200Hz.

When comparing the results in Figures 4.5-4.6 and 4.11-4.12, it can be said that the increase in the number of iterations due to Doppler mainly results from detecting more discrete targets. In fact, the higher the Doppler frequency, the more discrete targets are detected.

In Figures 4.13 and 4.14, it can be observed that decreasing λ increases the P_{FA} of STC in the existence of Doppler. Moreover, P_{FA} of STC algorithm is higher for high Doppler frequencies. As the target density increases, P_{FA} of STC increases when there is Doppler. When the Doppler frequencies of the targets become higher as in Figure 4.14, P_{FA} performances with respect to target density also increases.

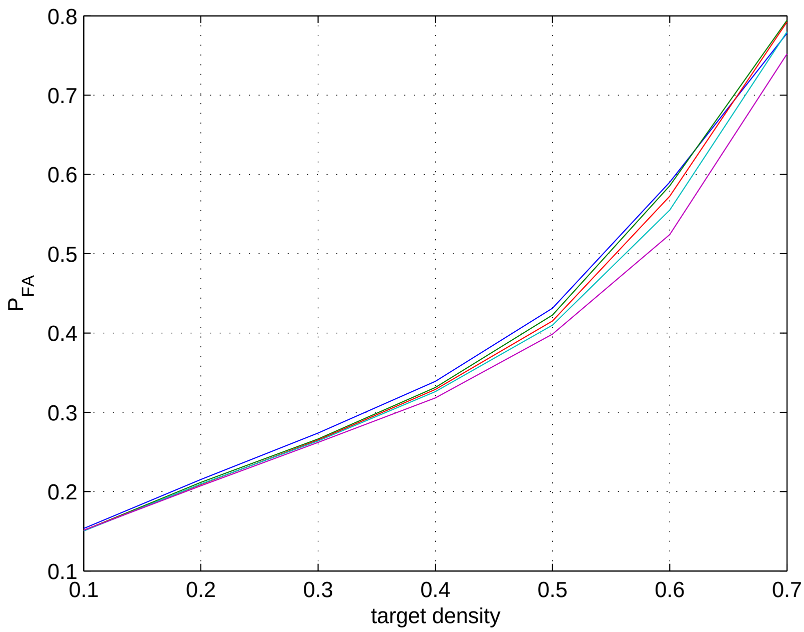

Figure 4.13: Probability of false alarm performances of STC method for $\lambda = 0.5, 0.8, 0.9, 0.95$, and 1 with respect to target density, Doppler randomly distributed with the uniform density between -200 and 200Hz.

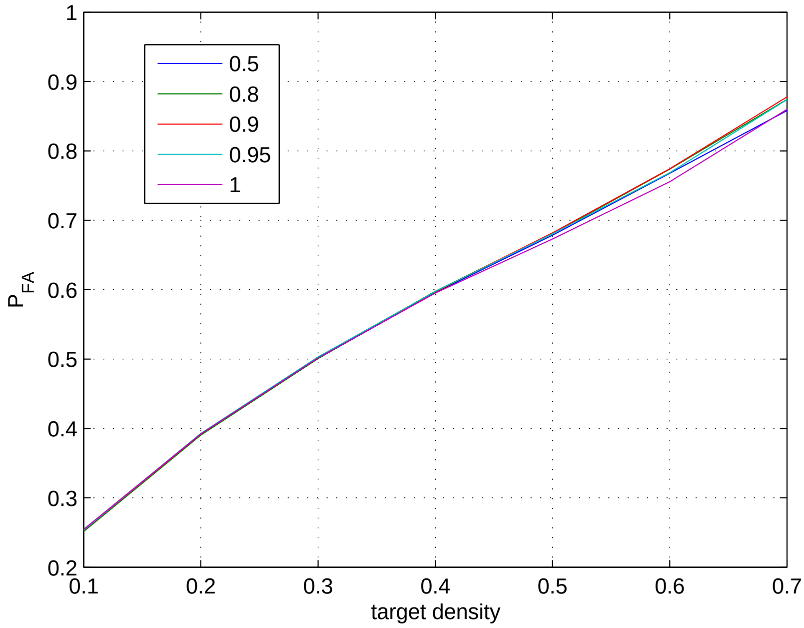

Figure 4.14: Probability of false alarm performances of STC method for $\lambda = 0.5, 0.8, 0.9, 0.95$, and 1 with respect to target density, Doppler randomly distributed with the uniform density between -500 and 500Hz.

In Figures 4.15 and 4.16, the third FFT filter is observed since the Doppler frequency of the targets is distributed between 1500 and 2500Hz. The residual error performances of Figures 4.9 and 4.15 are not very different. In addition to this, the complexity performances and false alarm detection characteristics of the two scenarios, in which the target Doppler frequencies are distributed between -500-500Hz and 1500-2500Hz, are nearly the same. When the correlation function is corrected for the center frequency of the studied filter as explained before, it can be concluded that the STC algorithm performs well even Doppler effect exists.

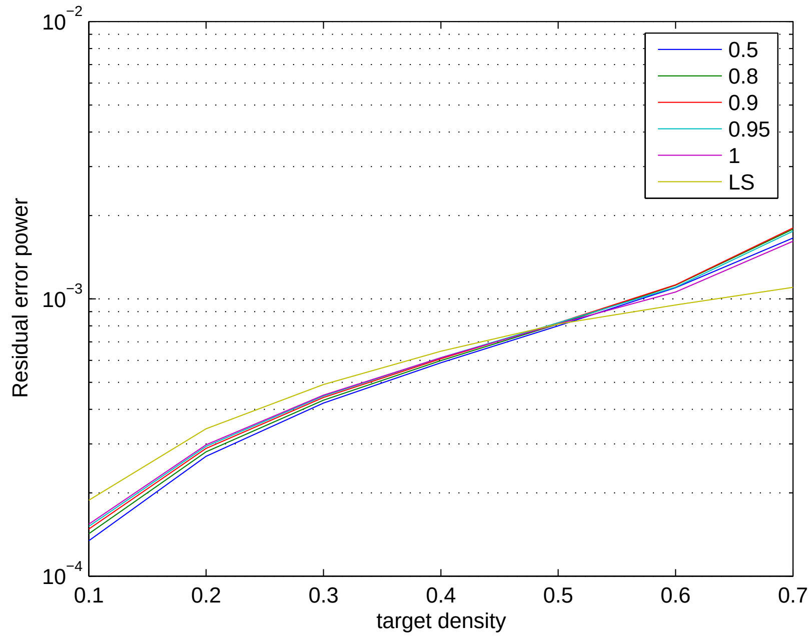

Figure 4.15: Residual error performances of LS method and STC method for $\lambda = 0.5$, 0.8, 0.9, 0.95, and 1 with respect to target density, Doppler randomly distributed with the uniform density between 1500 and 2500Hz.

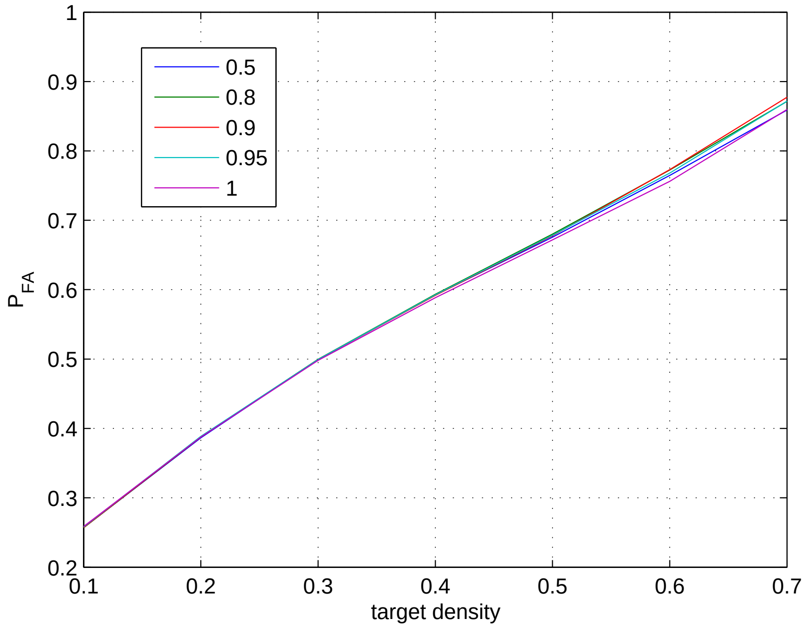

Figure 4.16: Probability of false alarm performances of STC method for $\lambda = 0.5, 0.8, 0.9, 0.95$, and 1 with respect to target density, Doppler randomly distributed with the uniform density between 1500 and 2500Hz.

4.3.2 Two Targets Case

In the second case, there is a large target and a small target in the environment. The small target is masked by the range sidelobes of the large target. We will check out the probability of detection (P_d) performance versus signal to noise ratio (SNR) of the small target. In many radar applications, targets are in motion. Therefore, Doppler effect is taken into account in this case in order to be more realistic. The third FFT filter which operates in between 1500-2500Hz is used.

STC aims to eliminate the adverse effects of sidelobes by estimating and canceling target signals. In order to observe whether STC can resolve the masking problem, we will compare the small-large target case (a) with the case that there is no masking (b), i.e., only the small target is present. If two cases have similar P_d performance, then STC will have fulfilled its promise.

In Figure 4.17, both target signals are deterministic, have fixed powers, and are 7 range bins apart from each other. In this figure, the power of the large target is 28dB higher than the power of the small target. The Doppler frequencies of the large target and the small target are 2200Hz and 1700Hz respectively. The noise power N_0 is varied so that the SNR effect on P_d can be examined. 1000 samples are taken for each N_0 value. The performance of detection based only on MF output in the single target scenario is also depicted for comparison and bears virtually no difference with STC. As seen in Figure 4.17, whether the large target is present in the environment or not, there is not much difference in P_d of the small target between LS and STC methods. Moreover, when STC method is applied, P_d of the small target does not change much according to the existence of the large target. P_d-SNR curves of the STC algorithm are different for different λ values: P_d performances of small λ values are worse.

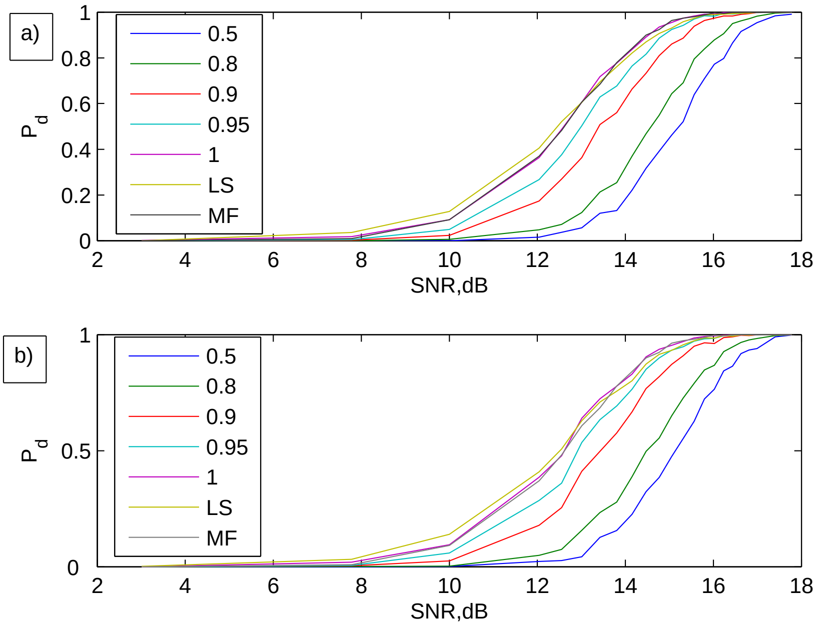

Figure 4.17: Probability of detection versus signal to noise ratio is examined for a) A deterministic small target at 1700 Hz Doppler masked by a deterministic large target at 2200 Hz Doppler whose power is 28dB higher than the small one. b) Only a deterministic small target. 0.5, 0.8, 0.9, 0.95, and 1 refers to 5 different values of λ.

In Figure 4.18, all the conditions and properties of target signals are the same as the ones in Figure 4.17 except from the Doppler effect on targets. Doppler frequency of the big target is again 2200Hz whereas Doppler frequency of the small target is 1550Hz. In this figure, it is aimed to investigate the P_d performance of the small target when the Doppler frequency of the small target is far away from the center frequency of the studied FFT filter. Although third FFT filter is used in this simulation, nearly the same amount of the target signal in the region of third FFT filter is in the region of second

FFT filter since the Doppler frequency of the small target is very close to the intersection of these two FFT filters. In Figures 4.17 and 4.18, it can be observed that the P_d performances of both LS and STC methods reduce as the Doppler frequency of the small target is moved away from the center frequency of the studied FFT filter.

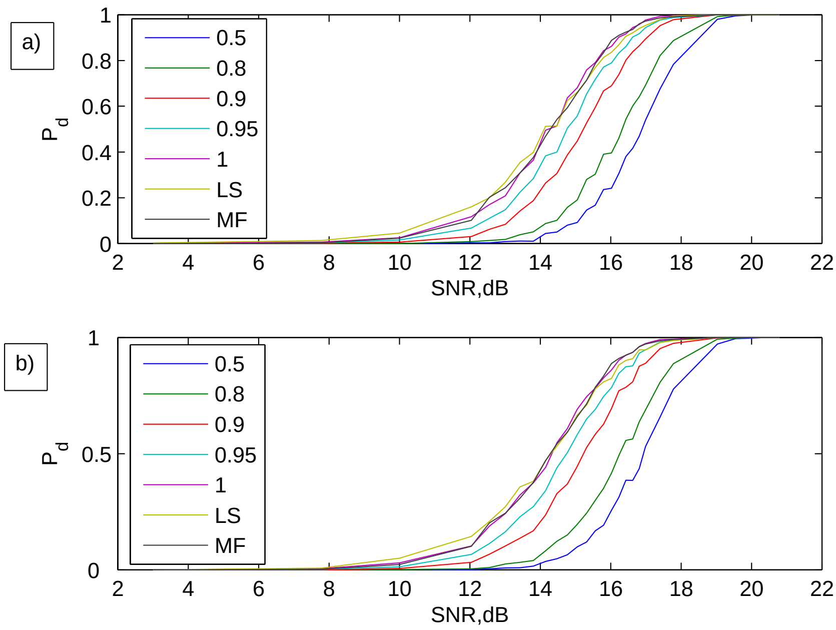

Figure 4.18: Probability of detection versus signal to noise ratio is examined for a) A deterministic small target at 1550 Hz Doppler masked by a deterministic large target at 2200 Hz Doppler whose power is 28dB higher than the small one. b) Only a deterministic small target. 0.5, 0.8, 0.9, 0.95, and 1 refers to 5 different values of λ.

In Figure 4.19, the small target signal is formed randomly with a complex Gaussian distribution for each 1000 samples while the large target is deterministic. The SNR of the big target is 50dB and N_0 of the noise is constant.

However, noise and small target signal are randomly generated. The Doppler frequency of the small target is randomly distributed with the uniform density between 1500 and 2500 Hz. In this figure, STC and LS methods give so close results that P_d-SNR curves of the methods can not easily be differentiated. P_d-SNR curve of the STC method in the existence of a large target is nearly the same as the one when the large target does not exist.

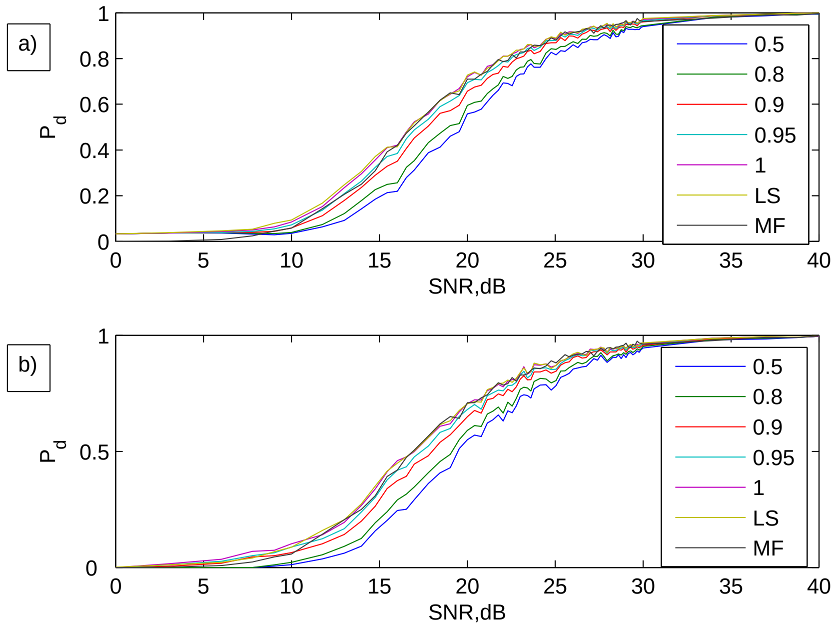

Figure 4.19: Probability of detection versus signal to noise ratio is examined for a) A randomly formed small target and a deterministic large target. b) Only a random small target. 0.5, 0.8, 0.9, 0.95, and 1 refers to 5 different values of λ.

As it is mentioned earlier, range sidelobes of big targets can be mistaken as targets during the radar signal processing. In Figure 4.19, the detection probability of the small target in the existence of the big target is greater than zero even at 0dB SNR. This unexpected result is due to the the range sidelobes of the big target.

Compared to Figure 4.17, in Figure 4.19 λ does not affect the P_d-SNR performance of STC algorithm so much because the average sidelobe effect in 4.19 is lower than that of 4.17.

In conclusion, it is obviously observed from the two figures that STC reduces most of the adverse effects of the sidelobe of the large target. Indeed, STC almost achieves optimum sidelobe reduction in the squared error sense with a reasonable complexity.

CHAPTER 5

Conclusion

Pulse compression is ubiquitously used in radar applications. The pulse compression known as standard matched filtering not only solves the limited peak power problem but also maximizes the received SNR of the target in the existence of white noise. However, this technique has the problem of sidelobe masking of smaller targets in the vicinity of large targets due to non-ideal correlation properties. To overcome this problem, the Successive Target Cancelation (STC) method which is a CLEAN algorithm is studied.

Basically, the STC method detects target signals and cancels them out from the matched filter output successively. We investigated the performance of the STC method in comparison to the LS method. Moreover, STC algorithm was performed for different loop gain parameter (λ) values. The effect of the parameter λ with respect to complexity, false alarm rate, residual error power, and probability of detection were inspected. Furthermore, the effects of the Doppler shift on the received radar signals have been examined and the STC method was correspondingly modified.

According to simulation results, STC performs well as long as the number of targets in the received signal is not excessively large. Besides this, STC performs better than the LS method for small target densities in the environment. The Doppler effect can be easily taken into account in STC without much degradation in performance. Although residual error performance of STC decreases with decreasing λ for very small target densities, both the complexity and the number of false alarms increase. Additionally, the P_d is higher when λ is high. By taking into account that the enhancement in the residual error performance is not much and this enhancement reduces with increasing Doppler, it can be concluded that performing STC with $\lambda = 1$ is preferable for sidelobe reduction in pulse compression radars. However, STC can be run for small λ's in applications where estimating the data with as small residual error as possible is important such as in image processing.

STC is based on basic operations such as addition and subtraction. It does not include matrix inversion or other complex operations and hence is simple and easy to handle. Due to inversion of a $n \times n$ matrix and the multiplication operation on the order of $O(n^2)$ where n is the length of the received signal, the computational complexity of LS method is high. Moreover, the greater the number of the samples taken for the received return signal is, the more computationally complex the LS algorithm becomes. On the other hand, while running MMSE method a signal covariance matrix is obtained for each range bin to be inspected from target power estimates and a filter is constructed based on the inverse of the covariance matrix and acts on the MF signal. At each iteration of the algorithm, target powers are reestimated and filters are regenerated. Therefore, MMSE algorithm is computationally complex due to on-line matrix inversion at every iteration and ad hoc methods used to avoid numerical instability in MMSE method resulting from ill-conditioned estimated covariance matrix.

In general, STC method almost carries out the optimum performance in the squared error sense. Moreover, it has lower complexity compared to other methods. Hence, it presents a viable solution to sidelobe problem in pulse compression radars.

As a future work, the effects of clutter suppression filters and windowed FFT operations can be studied. The performance of the proposed algorithm with constant false alarm rate detectors will be investigated.

REFERENCES

[1] Song, S.M., Kim, W.M., Park, D., and Kim, Y., "Estimation theoretic approach for radar pulse compression processing and its optimal codes", Electronics Letters, Vol. 36, No. 3, pp.250-252, Feb. 2000

[2] Blunt, S.D., and Gerlach, K., "Adaptive Pulse Compression" IEEE National Radar Conference, pp. 271-276, 2004

[3] Horiuchi, S., Kameno, S., and Ohishi, M., "Developing a Wavelet CLEAN Algorithm for Radio-Interferometer Imaging", Astronomical Data Analysis Software and Systems X ASP Conference Series, Vol. 238, 2001

[4] Skofronick-Jackson, G.M., and Gasiewski, A.J., "A nonlinear multispectral statistical CLEAN-based precipitation parameter-retrieval algorithm", Geoscience and Remote Sensing, IEEE Transactions on, Vol. 38, No. 1, pp.226-237, Jan. 2000

[5] Peichl, M., Greiner, M., and Suess, H., "DLR activities on aperture synthesis radiometry", Proceedings of the 2003 IEEE Radar Conference, pp.446-451, May 2003

[6] Verdu, S., "Multiuser Detection", Cambridge University, 1998

[7] Deng, H., "Effective CLEAN algorithms for performance-enhanced detection of binary coding radar signals" Signal Processing, IEEE Transactions on, Vol. 52, No. 1, pp.72-78, Jan. 2004

[8] Skolnik, M.I., "Introduction to Radar Systems", McGraw-Hill, 1980

[9] "Introduction to Basic Radar", Available at : "http://www.tpub.com/content/et/14089/css/14089_11.htm", August 2006

[10] Massachusetts Institute of Technology. Radar School. , "Principles of radar, by members of the staff", McGraw-Hill, 1952

[11] Proakis, J.G., "Digital Communications", Third edition, McGraw-Hill, 1995

[12] Sato, R., and Shinriki, M., "Time Sidelobe Reduction Technique for Binary Phase Coded Pulse Compression", (Translated from Denshi Joho Tsushin Gakkai Ronbunshi, Vol. J83-B-II, No. 3, March 2000, pp. 352.360, vol.84, no.12), 2001

[13] Mudukutore, A.S., Chandrasekar, V., Keeler, R.J., "Pulse Compression for Weather Radars", Geoscience and Remote Sensing, IEEE Transactions on, Vol. 36, Issue 1, pp.125-142, Jan. 1998

[14] Gray, R.W., and Farley, D. T., "Theory of incoherent-scatter measurements using compressed pulses", Radio Sci., Vol. 8, No. 2, pp.123131, Feb. 1973

[15] Aktop, O., "Optimization of Doppler Processing by Using Bank of Matched Filters", M.S. thesis, Middle East Technical University, Sept.2005

[16] Frank, S.L., "Polyphase Codes with Good Nonperiodic Correlation Properties", IEEE Transactions on Information Theory , vol. IT-9, pp.43-45, Jan. 1963

[17] Kretschmer, F.F., and Welch, L.R., "Range-time-sidelobe reduction technique for FM-derived polyphase PC codes" Aerospace and Elec-

tronic Systems, IEEE Transactions on , Vol. 29, Issue 3, pp. 834-840, July 1993

[18] Kretschmer, F.F., and Welch, L.R., "Sidelobe reduction techniques for polyphase pulse compression codes " IEEE Radar Conference, pp. 416-421, May. 2000

[19] Haykin, S., "Communication Systems", Fourth edition, John Wiley & Sons Inc., 2000

[20] Luenberger, D.G., "Optimization by Vector Space Methods", John Wiley & Sons Inc., 1969

[21] Sarkar, T.K., Park, S., Koh, J., and Schneible, R.A., "A Deterministic Least Squares Approach to Adaptive Antennas", Digital Signal Processing, A Review Journal, Vol. 6, No. 3, pp. 185-194, 1996

[22] Blunt, S.D., and Gerlach, K., "A novel pulse compression scheme based on minimum mean-square error reiteration" IEEE National Radar Conference, pp. 349-353, Sep. 2003

[23] Sarkar, T.K., and Brown, R.D., "An ultra-low sidelobe pulse compression technique for high performance radar systems" Proc. IEEE National Radar Conf., pp. 111-114, May. 1997

[24] Ackroymd, M.H., and Ghani, F., "Optimum mismatched filter for sidelobe suppression", IEEE Trans., pp. 214-218, 1973

[25] Jung, K.T., Kim, C.J., Lim, C.H., Lee, H.S., and Kwag, Y.K., "Design of Optimum Mean Square Sidelobe Suppression Filters for Barker Codes" Radar 92. International Conference, pp. 530-533, Oct. 1992

[26] Zrnic, B., Zejak, A., Petrovic, A., and Simic, I., "Range sidelobe suppression for pulse compression radars utilizing modified RLS algorithm"

Spread Spectrum Techniques and Applications, IEEE 5th International Symposium on, Vol. 3, pp. 1008-1011, Sep. 1998

[27] Zejak, A.J., Zentner, E., Rapajid, P.B., “Doppler optimized mismatched filters”, Electronics letters, vol. 21, no.7, pp. 558-560, 1991

[28] Zejak, A.J., Zatkalik, J., DukiC, M.L., “Doppler optimized compression filters”, IEEE ISSSTA ’94, pp.534-538, Finland, July 1994

[29] Gustafsson, F., Gunnarsson, F., Bergman, N., Forssell, U., Jansson, J., Karlsson, R., Nordlund, P.J., “Particle filters for positioning, navigation, and tracking”, Signal Processing, IEEE Transactions on , Vol. 50, Issue 2, pp.425-437, Feb. 2002

[30] Goldstein, J.S., Reed, I.S., Zulch, P.A., “Multistage partially adaptive STAP CFAR detection algorithm”, Aerospace and Electronic Systems, IEEE Transactions on , Vol. 35, Issue 2, pp.645-661, Apr. 1999

[31] Goodman, N., Stiles, J., “A MMSE Filter for Range Sidelobe Reduction”, Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2000. IEEE 2000 International, Vol.5 , pp.2365-2367, July 2000

[32] Schwarz, U.J., “57 Mathematical and statistical description of the Iterative Beam removing Tecnique (Method CLEAN)” Astron. Astrophys., Vol. 65, pp.345-356, 1978

[33] Starck, J., Pantin, E., and Murtagh, F., “Deconvolution in astronomy: A review”, Publications of the Astronomical Society of the Pacific, Vol. 114, pp.1051-1069, Oct. 2002

[34] Cramer, R.J., and Scholtz, R.A., “Evaluation of an Ultra-Wide-Band Propagation Channel”, IEEE Transactions on Antennas and Propagation, Vol. 50, No. 5, pp.561-570, May 2002

[35] Kretschmer, F.F., and Gerlach, K.R, "Low Sidelobe Pulse Compression Waveforms", IEEE Signals, Systems and Computers Conference, Vol. 2, pp.663-667, Nov. 1988