STAP WITH ADAPTIVE STATE ESTIMATION IN NON-STATIONARY HETEROGENEOUS SYSTEMS

STAP WITH ADAPTIVE STATE ESTIMATION IN NON-STATIONARY HETEROGENEOUS SYSTEMS.

By

Md. Obaidul Malek, B.Sc.(Hons.), M.Sc.

2009

A Thesis

Submitted to the Department of Electrical & Computer Engineering

and the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Master of Applied Science

McMaster University

MASTER OF APPLIED SCIENCE (2009)	McMaster University
(Electrical & Computer Engineering)	Hamilton, Ontario

TITLE:	STAP with Adaptive State Estimation in Non-Stationary Heterogeneous Systems.
AUTHOR:	Md. Obaidul Malek B.Sc.(Hons.) (Applied Physics and Electronics), M.Sc. (Applied Physics and Electronics), University of Dhaka, Dhaka, Bangladesh.
SUPERVISORS:	Dr. Thia Kirubarajan
NUMBER OF PAGES:	xxii, 106

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Master of Applied Science.

Dedications

To My Beloved Late Mother

Marhooma Sufia Malek.

Abstract

In radar signal processing, the vulnerability of the desired signal to homogeneous and heterogeneous interferences increases as the communication traffic increases. The principal challenge in the radar system then is to mitigate the effects of cold (homogeneous) clutter, severe dynamic (heterogeneous) hot clutter and jamming interferences while estimating the states of targets under track. Space-Time Adaptive Processing (STAP) enhances the capability of radar systems to overcome this challenge. However, it is a sample-based system where the adaptive processing is sensitive to the underlying assumptions as well as the diversity of potential interferences. Hence, the performance of STAP deteriorates when basic assumptions are violated due to errors in receiver array elements, non-stationary nature of interferences, inadequate Independent and Identically Distributed (i.i.d.) sample data, and target like-signal in the training data set.

This thesis proposes an Adaptive State Estimation (ASE) approach to characterize STAP used simultaneously in spatial and Doppler domains for non-stationary, homogeneous and heterogeneous systems. The contributions presented here are based on the adjustment of the weight vector and the update of associated interference covariance matrix by ASE to minimize the output noise power while maximizing Signal to Interference-plus-Noise Ratio (SINR) in the Mean Squared Error (MSE) sense. The integration of STAP principle with sequential state estimation in order to decode the

target signal while rejecting the interferences due to non-stationary heterogeneous clutter and jammer effects without degrading performance is the key contribution of this paper. The Proposed STAP-ASE algorithm is shown to outperform its counterparts in terms of efficiency, IF-improvement factor, Signal to Interference-plus-Noise Ratio (SINR) convergence rate and target detection. Simulation results are presented to illustrate the performance of the proposed technique.

Acknowledgements

My first and foremost praise and gratitude is to almighty God, *Allah*, our Creator, the Most Merciful, the Most Compassionate. Without the help of *Allah*, none of the work in my life would have come to light. His kindness and help are uncountable and undefinable. I am simply very thankful to Him for allowing me to recognize His rewards.

For the grace of *Allah*, I have got the opportunity to have several great people in my life whose presence, love, cooperation and contributions have helped me to achieve my goals. I humbly acknowledge their undeniable contributions and dedications behind my success.

First of all, I would like to express my special thanks to my supervisor, Dr.Thia Kirubarajan. He has been instrumental in the development of my academic expertise, and in organizing me to establish myself as a graduate research fellow. I have benefited tremendously from his leadership, and interest, in developing the intellectual and personal growth that is generally seen among his graduate students. As a supervisor, the results in this thesis would not have been possible without his engaged and insightful discussions. He has always been a committed, supportive and flexible instructor to his students, and provided them with advice, and shown understanding during hard times. Certainly, he has left an indelible mark and there are still many things to learn from Dr.Thia Kirubarajan. In fact, his leadership, support and expertise and openness made me to complete my research successfully. I am honored to have had Dr.Kiruba as my supervisor.

I would like to extend my thanks to Dr.Bhashyam Balaji. What, I have learned about STAP is because of Dr.Balaji. His interest, support, constructive comments, appreciation and cooperation helped me to explore my knowledge and research in this area. In a sentence, perhaps, I was blessed by *Allah* that I got the opportunity to work with Dr.Balaji. I owe him a lot, since without his cooperation and thoughtful comments, I would not have been able to complete my research.

I would also like to extend my gratitude to Dr.Ratnasingham Tharmarasa, for his resourceful and constructive advice.

From the bottom of my heart, I would like to thank my wife, Salma. During my time of study, she has always shown patience, perseverance and continuous understanding. No one has been impacted by this thesis and my M.A.Sc studies more than her. I would also like to extend my gratitude to my hearts, two daughters, Sophia and Sadia. Special thanks to Sophia, for helping me with the writing, and initial proofreading of this thesis.

I would also like to sincerely thank my father, Professor Abdul Malek, father-in-law Md.Ataul Haque, brother Md.Anwarul Malek, and sister Dr.Laila Alamgir, for their continuous help, support, and trust. They have taught me many things. Most importantly, they have taught me to make my own decisions. My deepest gratitude and dedication go to my mother, Marhooma Sufia Malek. Her uninterrupted love and encouragement had tremendously inspired me to pursue higher education. Without her love and support, I wouldn't be what I am today. I would also like to thank my uncles, aunts and other family members. They have supported me all the way before and after graduation, and have always encouraged me to excel in my studies.

I have been surrounded by numerous great friends, group members and lab members who are hard to forget. I would like to thank all of them, including my lab member Dr.Wessam Mesbah, Nasser Morad, my group members Nandakumaran Nadarajah, Alikabar Gorji Daronkolaei, Dr.David Akselrod, Dan Danu, Darcy Dunne , and Xin Chen. In particular, I would like to thank Nandakumaran, Dr.Wessam, and Nasser for their significant comments and their constructive advice.

Special thanks to all of the faculty members, including Dr.Simon Haykin, Dr.T. Davidson, Dr.Reilly, Dr.Jian Kang Zhang, Dr.Drazena, and Dr.Chih-Hung (James) Chen, whose direct and indirect contributions helped me to explore my research in this area.

Finally, I am grateful to all my school teachers, university professors and teaching assistants, and to every individual who has taught me in some way.

List of Acronyms

ABF	Adaptive Beamforming
ASE	Adaptive Sequential Estimation
$ASE - STAP$	STAP-Adaptive State Estimation
BF	Beamforming
$\mathbf{c}$	Predefined Signal Vector
CPI	Coherent Processing Interval
CRLB	Cramér-Rao Lower Bound
d	Desired Signal Output
d_{non}	Desired Signal Output-nonLinear
$\mathbb{E}$	Expectation
e	Error
EKF	Extended Kalman Filter
f_d	Doppler Frequency
f_{cd}	Doppler Frequency-Clutter
f_{tarD}	Doppler Frequency-Target
g	Constraint Gain
H(.)	Measurement Matrix
h(.)	Measurement Matrix-nonLinear
I	Identity Matrix

IF	Improvement Factor
i.i.d	Independent and Identically Distributed
K	No. of Snapshot
KF	Kalman Filter
L	Samples Data Set
LSE	Least Square Estimator
LSMI	Loaded Sample Matrix Inversion
M	No. of Pulses
MMSE	Minimum Mean Square Error
MSE	Mean Squared Error
MVDR	Minimum Variance Distortionless Response
N	Number of Array Element
PCRLB	Posterior Cramér-Rao Lower Bound
PRI	Pulse Repetition Interval
P_s	Signal Power
P_n	Noise Power
Q	Interference Covariance Matrix
Q_c	Interference Covariance Matrix-Clutter
Q_j	Interference Covariance Matrix-Jammer

$\mathbf{q}_c$	Interference Vector-Clutter
q_j	Interference Vector-Jammer
$\mathbf{s}$	Signal Steering Vector
SINR	Signal Interference Noise Ratio
SMI	Sample Matrix Inversion
STAP	Space-Time Adaptive Processing
U	Whitening Block-Interference
v_p	Platform Velocity
v_{rad}	Radial Velocity
$\mathbf{w}_c$	Constraint Weight Vector
$\mathbf{w}$	Weight Vector
$\mathbf{x}$	Steering Vector-Observed Signal

List of Notations

j	the unit imaginary number $j = \sqrt{-1}$
$(\cdot)^*$	the conjugate operator
$(\cdot)^T$	the transpose of a vector or a matrix
$(\cdot)^H$	the Hermitian transpose of a vector or a matrix
$(\cdot)^{-1}$	the inversion of a matrix
$\lvert \cdot \rvert$	the determinant of a matrix
$\lVert \cdot \rVert$	the Euclidean norm of a vector
$\otimes$	the Kronecker matrix product
∇	the gradient operator
λ	the operating wavelength

Contents

List of Figures

Chapter 1

Introduction

The word "radar is an abbreviation for RAdio Detection And Ranging. It refers to the technique of using radio waves to detect the presence of the target of interest in the atmosphere. In radar communication systems, modulated waveformed and directive antennas are being used to transmit the electromagnetic energy into a specific direction in order to detect, and track the target of interest. Target within the surveillance area will reflect portions of the transmitted energy, echoes or radar returns, back to the radar. These reflected echoes from the target are then received and processed by the radar receiver in order to extract the target information such as range, range rate, acceleration, angular position and other target specific identifications. Radars are very complex electronic and electromagnetic systems. There is a great diversity in the design architecture of the radar systems based on purpose, however the fundamental operating principle and basic designed architecture is the same. They (Radars) can be classified as ground based, airborne and spaceborne radar systems. Depending on the type of waveforms radars use or by their operating frequency, radar can also be classified as Continuous Wave (CW) radar and the Pulse Radar (PR). A simplified block diagram of radar communication systems is shown in Figure 1.1.

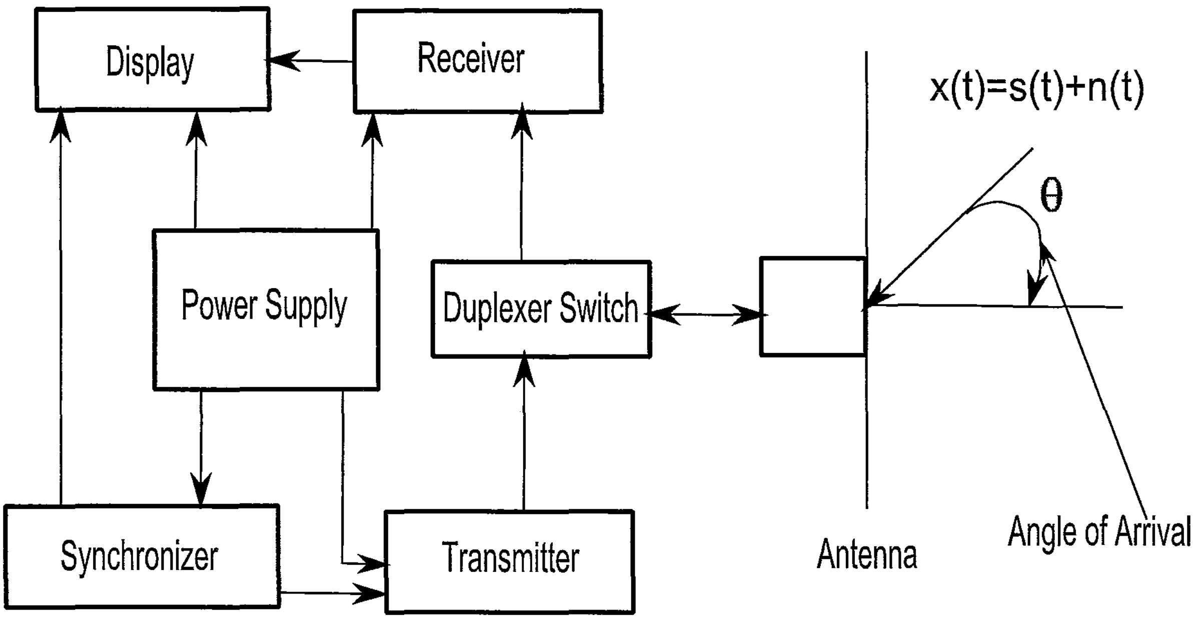


Figure 1.1: Block Diagram of Radar Communication Systems

In the radar communication systems, if the received or the observed signal $\mathbf{x}(t)$ at time t has the interference $\mathbf{n}(t)$, then the main challenges of the radar communication systems are to extract the signal of interest $\mathbf{s}(t)$ from this observed signal. This type of signal reception by the conventional radar receiver has been attractive solution for the severe problem of signal detection and estimation, but conventional signal reception technique are susceptible to degradation in Signal to Interference-plus-Noise Ratio (SINR) performance because of the constant changing of the interference environment due to nonstationary, nonlinear and heterogeneous sources. This degradation may be further aggravated as the radar communication traffic increases. As a result, adaptive signal processing is the focal area for reducing the susceptibility of signals to interferences because of its automatic sensations, hence rejection of the interferences without prior knowledge of the signal environment. Space-Time Adaptive Processing (STAP) is a multi-dimensional adaptive signal processing technique that estimates adaptive weight vectors in spatial and Doppler domains for which a target detection hypothesis is to be formed. STAP operates on the set of returns, composed of pulses, array elements, and range bins, over a period of time. Hence, STAP is a 2D processing on 3D datacube, collected from the available signals received at the radar receiver. Since, adaptation is performed in spatial and temporal domains, training is done using the range bins. Typically, it is a sample-based approach where covariance matrix may be estimated from Independent and Identically Distributed (i.i.d) sample data. The main theory of this processing is to adapt with the sample data set in order to estimate the interference covariance matrix and adjust or update the weight vector such a way that, the noise power can be minimized, and Signal to Interference-plus-Noise Ratio (SINR) can be maximized in some appropriate sense.

1.1 Fundamentals of Space-Time Adaptive Processing

STAP processing has long been considered for airborne radar in order to mitigate the target signal in strong ground clutter environment [2]. Processing in STAP is based on the adaptation of sample data, training the sample and solving a set of linear equations. In STAP, two types of data are typically processed: one is training data, which is used to estimate the interference covariance matrix and adaptive weight vector. Other is the primary data or the test data on which detection and parameter estimation are performed. In STAP, increment in the range bin within a particular Pulse Repetition Interval (PRI) is known as the fast time samples, and those across the PRIs is called slow time samples. If there are N antenna elements, and M pulses over the Coherent Processing Interval (CPI), the group of N samples associated with each pulse (over a particular CPI) are collected, or one sample is taken simultaneously from each of the N antenna element, to make one snapshot. Therefore, there are $M \times N$ snapshots for each range bin. Signal processing on STAP platform is based on the following three fundamental (core) equations.

1.) Equation for the estimation of the asymptotic interference covariance matrix for ith from the training data set [1] [Complete derivation in Appendix A]:

$$Q_i = \frac{1}{L}\sum_{l=1}^{L} \mathbf{x}_l \mathbf{x}_l^H \tag{1.1}$$

where, $L \geq MN$ is the total number of samples being used for the ith range bin, $\mathbf{x}$ is the input data vector, which may have noise only or target with noise, and $(.)^H$ represents the hermitian transposition. The dimension of $\mathbf{x}$ is same as L $\geq MN$. In this case, No. of Range Bins $\geq MN$ and i represents the range bin index.

One of the major challenges in STAP processing is to estimate the interference covariance matrix from the i.i.d sample data set, since the prior of the matrix is not known. Therefore, the Maximum Likelihood (ML) function in (1.1) needs to be utilized in order to estimate the maximum likelihood of the interference covariance matrix. However, in this approach, the input data from all the surrounding range bins except the bin under test would be considered, and it is also assumed that there is no target signals in the surrounding range bins (unsupervised or null hypothesis-Appendix D). But, there are two main possibilities for having the target signal in the secondary data: one is the perfectly matched with the target signal response, and the other is the mismatched [1]. In the proposed approach, the perfectly matched target like signal (semi-supervised) has been considered and filtered out from the secondary data set in order to estimate the maximum likelihood of the interference covariance matrix. It is also assumed that interference is Gaussian, and Independent and Identically Distributed (i.i.d).

2.) Equation for the weight vector:

$$\mathbf{w}_{opt} = Q^{-1}\mathbf{s} \tag{1.2}$$

Where, Q and $\mathbf{w}_{opt}$ are (unknown) expected asymptotic interference covariance matrix, and optimal weight vector, $\mathbf{w} = [w_1, w_2, \ldots, w_{MN}]^H$ respectively. $\mathbf{s} = [s_1, s_2, \ldots, s_{MN}]^H$ is the desired signal steering vector. The covariance (asymptotic) matrix Q can be obtained from (1.1), and hence the weight vector $\mathbf{w}$. For known value of Q, equation (1.2) is an optimal solution for the weight vector, i.e., $\mathbf{w} = \mathbf{w}_{opt}$.

3.) Equation for the scalar beamformering output:

$$\begin{aligned} y &= \mathbf{w}^H(\mathbf{s} + \Omega) \\ y &= \mathbf{w}^H\mathbf{s} + \mathbf{w}^H\Omega \end{aligned} \tag{1.3}$$

where, observed signal vector $\mathbf{x} = [x_1, x_2, \ldots, x_{MN}]^H$ with noise vector Ω.

The core concept of the STAP formulation is based on above three equations (1.1-1.3). Furthermore, a low sidelobe antenna with fixed interference does not need training, hence the adaptation. But, often there is a need to detect and estimate the target of interest in the vicinity of diverse interferences. Therefore, adaptation and adjustment are the two crucial steps in STAP processing. Adaptation involves the estimation of the interference matrix from the i.i.d training data vector and the requirement of the number of data vector for the estimation of covariance matrix increases with the increase of matrix dimension. On the other hand, adjustment of the weight vector would be required in such a way that the process can track the changes due to changes in the nature of the interferences. The main objective of the STAP is to isolate the signal from the interferences by adaptively estimating the weight vector and error covariance matrix based on the sample data (training data) received from the spatial and temporal domains. Main motivation of the STAP processing is its outstanding performance to isolate the slow and fast moving target from the interferences, even if the interference signals are stronger than the signal of interest.

Depending on how sample returns (input data) are processed, STAP can be classified into fully adaptive processing, subspace adaptive processing, and post Doppler adaptive Processing.

In the fully adaptive processing technique, the sample data is processed with full the degree of freedom as obtained by the number of array elements and the pulses [1]. In fact, this model estimates the weight vector, mitigates the target signal and rejects the interference covariance matrix on the way processor has received the sample date [1]. This technique faces challenges for the adaption with the training data set due to the insufficient i.i.d sample data set in one hand and computational complexity due to the estimation of the inverse of interference covariance matrix on the other.

In the subspace technique, the large covariance matrix is transformed into discrete matrices, so that the computational load can be reduced, and the challenge not having sufficient i.i.d sample data can also be avoided. Matrix transformation can be done on spatial or temporal domain or both. This is a suboptimal process, however, this method is widely used, since it is computationally efficient.

In post Doppler processing, Doppler filtering is performed before the STAP adaptation. This is also called frequency depended processing where scanning is performed based on the Doppler bins. It is a special form of subspace technique. However, it is suboptimal and the performance is poor compared to full and typical subspace adaptive processing.

The motivation of the STAP processing is its capability to enhance the radar receiver to mitigate the desired signal power and isolate the target even if the undesired signal is stronger compared to the target signal. One of the major challenges in STAP processing is to estimate the interference covariance matrix from the sample data set, because of insufficient i.i.d sample data set due to nonstationary, and heterogeneous clutter and jammer characteristics. More importantly, the prior of the interference covariance matrix is unknown, however, using the Maximum Likelihood Estimator (MLE) function stated in (1.1), the likelihood of interference covariance matrix may be estimated.

1.2 Space-Time Adaptive Processing with Adaptive State Estimation (STAP-ASE)

A common problem in any radar communication system is its additive noise at the receiver. One source of additive noise is from the solid state devices and resistors used in the implementation of the receiver module due to thermal effects. Another source of additive noise interference, is due to clutter and jammer effects. Therefore, the received signal at antenna elements are contaminated not only due to cold clutter interference but also nonstationary hot clutter and jammer. Hence, it is necessary to detect and isolate the effect of nonstationary and nonlinear signal interferences from the target of interest. In STAP processing, it is usually assumed that the training (secondary/auxiliary) data samples are free from target signal (unsupervised). But, in reality, there is a high possibility that the training data set is being contaminated by the target signal. Therefore, overall performance, efficiency, and accuracy of the STAP processing depends on the appropriate selection of auxiliary data set. It is crucial to have a target free secondary data for computing the covariance matrix of the range cell under consideration (test), so that the cancelation of desired target signal or self-nulling may be avoided. Therefore, the filter needs to be designed in a way that would enable to isolate the target like signal from the sample data set before starting the training process. Data sample collected by this method can be termed as semi-supervised. STAP processing not only mitigates the combined interference effect from nonstationary hot and cold clutter and jammer, but it also enhances the detection of small and slow moving target, and provides robustness in the presence of undesired signals.

The proposed STAP-ASE model is based on the minimization of the interference covariance matrix and maximization of the Signal to Interference-plus-Noise Ratio (SINR) in the MSE sense. Therefore, the processor would be able to decode or isolate the signal of interest from the nonstationary heterogeneous interferences due to clutter and jammer sources. During the weight training process, there may be no limit on the value of the weight vector [4], so that the constraint on the weight value may introduce consistency on the estimation process [4]. Therefore, a constraint was added to weight vector, which would accelerate the system convergence by limiting the weight vector within an acceptable range and may also keep the relationship required for the processing. Constrained weight vector may be considered as a key tool to cancel out the interferences due to hot clutter, cold clutter, and noise, so that the system can converge at its optimal point without affecting the original signal characteristics and at the same time maintain the rigorous reliability of the processor.

The motivation of this model is its effective suppression of clutter and jammer interferences by integrating STAP principle with sequential state estimation. In STAP-ASE, the interference covariance matrix adaptively changes over the CPI due to nonstationary clutter and jammer interferences. The interference covariance matrix is also updated within the filtering loop using previous data. Therefore, actual attainable interference suppression, and convergence rate based on worst case scenario are much higher than other techniques like Sample Matrix Inversion, Loaded Sample Matrix Inversion. In this model, the Posterior Cramér-Rao Lower Bound (PCRLB) is analyzed from the dynamic state model perspective and have used this lower bound as an achievable optimal point in order to cancel out the interferences, and mitigate the signal power. Maximum attainable SINR, performance and the Improvement Factor (IF) are also examined. The performance of Minimum Variance Distortionless Response (MVDR) is compared with the proposed STAP-ASE. ASE-SATP has shown outperform performance compared to its counterparts, SMI, LSMI, for signal decoding, Improvement Factor (IF), SINR, efficiency, consistency and convergence rate.

1.3 Contributions and Previous Work

The nonlinear model in [7] presents a robust adaptive beamforming (spatial filtering) system. The approach is to implement the robust Minimum Variance Distortionless Response (MVDR) beamformer due to mismatches in the approximation of the desired signal steering vector.

Wiener filter approach stated in [1][2], is for a linear stationary system, since Wiener filter is optimal filter under the assumption that the system is linear and time-invariant. The typical block diagram of a Wiener filter is given in Figure 1.2.

Constrained Kalman filter model stated for STAP in [3] is for a stationary linear system with unity gain, which performs well under the assumption that system environment is stationary and linear. The operational diagram for this model is stated

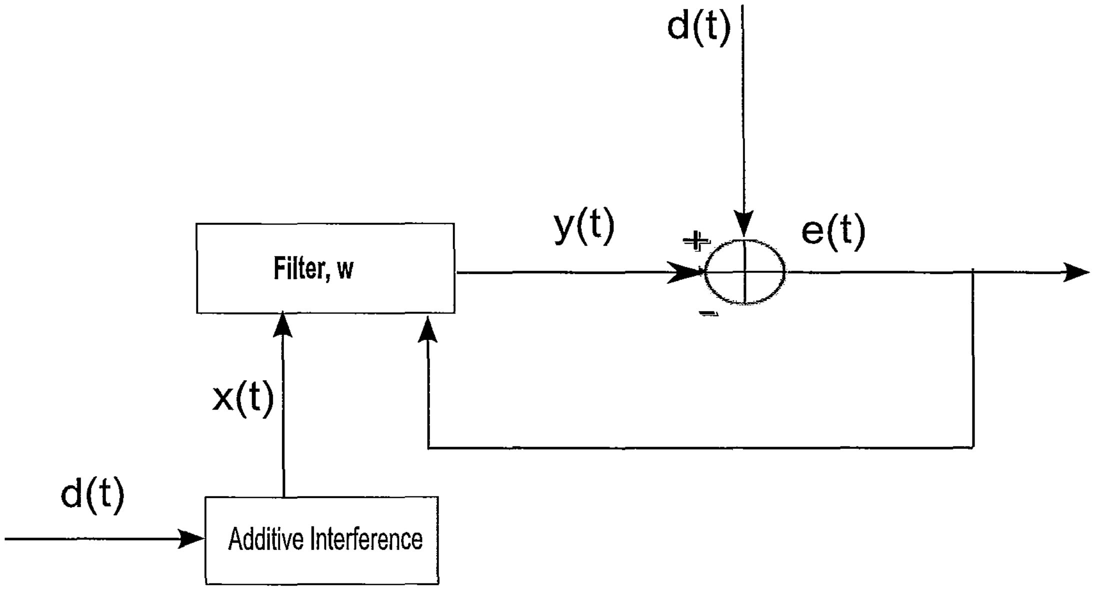


Figure 1.2: Schematic Diagram of a Wiener Filter

in Figure 1.3 [3].

In the proposed STAP-ASE model, STAP itself is integrated with sequential state estimation for rejecting the clutter and jammer effects, hence decoding the signal of interest by considering the worst case scenario.

In most of the previous works including Sample Matrix Inversion (SMI), Loaded Sample Matrix Inversion (LSMI), Adaptive Beamforming Using the Constraint Kalman Filter, and Robust Adaptive Beamforming Based on Kalman Filter, weight vector is updated within the filtering loop, and decomposed interference covariance matrix off the loop. In the proposed STAP-ASE model, weight vector, and its associated interference covariance and updated within the filtering loop. Interference covariance matrix is updated in every loop cycle using the data set from the immediate previous

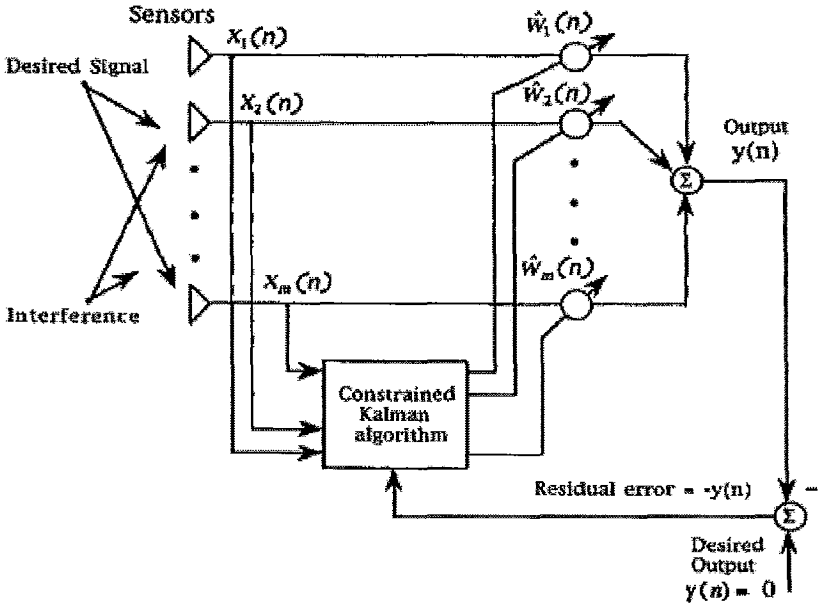


Figure 1.3: Block Diagram of Linear Constrain Kalman Beamformer (Copied From [3])

state. As a result the STAP-ASE model has higher rate of convergence as well as better rejects the interferences originated from the cold and hot clutter and jammer sources.

Furthermore, the implementation of a Kalman filter on a dynamic model (i.e., STAP), depends on the adaptation and on-line estimation of unknown parameters, and is severely biased on the nature of the received signals. More importantly, in nonstationary heterogeneous system where the signal characteristics fluctuate over time due to diverse nature of the received signal, there may be noticeable variation in interference covariance matrix over a single (same) Coherent Processing Interval

(CPI), hence the state estimation, which in turn degrades the performance and consistency of the processor with models stated in SMI, LSMI, and [3][7][18]. In the proposed model, the interference covariance matrix changes over CPI have been considered because of nonstationary and heterogeneous nature of the clutter and jammer, and has the ability to sense adaptively the presence of the nonstationary interference sources and suppress the interference while simultaneously decoding the signal with high precision.

1.4 Thesis Organization

Chapter 2 provides a brief overview of the technical background, i.e., prerequisites for STAP processing. In particular, it includes a brief summary of some of the theories about signal processing, statistical properties, and prerequisites upon which the Space-Time Adaptive Processing (STAP) has been developed. Comprehensive theory of STAP processing, interferences, possible challenges, nonstationary systems and beamformer output has also been discussed in this chapter. Finally, in this chapter, the Least Mean Squared and Wiener filter approaches are also discussed.

In Chapter 3, an overview of signals for the proposed STAP-ASE model is discussed. One of the biggest challenges of the STAP processing is to estimate the interference covariance matrix from the sample data set because of insufficient i.i.d sample data set, nonstationarity, and heterogeneous clutter and jammer characteristics. More importantly, the prior of the interference covariance matrix is unknown. Estimation of the asymptotic interference covariance matrix, STAP-signal characterization, signal and noise, clutter and jammer interference models are stated in this chapter. Furthermore, a model for constraint weight vector is also discussed. It is shown that predefined signal vector with the associated constraint weight vector, would be able to approximate the desired signal so that the signal from the interference can be isolated. Mathematical formulation for the maximum attainable Signal to Interference-plus-Noise Ratio (SINR) is also presented in this chapter. Finally, the Improvement Factor, states the performance and efficiency of the estimator, and generalized expression for Minimum Variance Distortionless Response (MVDR) are also included in this chapter.

Chapter 4 is the core one for the proposed STAP-ASE processing. In this chapter, the problem formulation, design algorithms, and design architecture for the proposed STAP-ASE model are discussed. Different aspects of finite sample size, SMI and LSMI algorithms, for STAP processing are also presented. Furthermore, the Wiener filtering approach is formulated, and it is shown that the Wiener filter is the special case of the Kalman filtering approach, since it has been developed under the assumption that the system is stationary and linear. More importantly, a novel model, which integrates STAP with the extended Kalman filter is presented. The computational burden for this model is also discussed. This chapter also presents the Posterior Cramér-Rao Lower Bound (PCRLB), and the efficiency of the proposed model is also formulated.

The simulations, results and discussions are included in Chapter 5. Finally, Chapter 6 presents the conclusions of this thesis and future work.

Chapter 2

Prerequisites for Space-Time Adaptive Processing (STAP)

2.1 Introduction

As the communication networking increases, the vulnerability of the desired signal to the homogeneous and heterogeneous interferences increases. Hence, adaptive signal processing architectures are currently the main focal point as a means for the susceptibility of the desired signal to the interferences. The main reason of widening the interest for the adaptive systems because of its automated adaptation as well as suppression of the interferences. Space-Time Adaptive Processing (STAP) is a multi-dimensional signal processing technique, which enhances the radar receiver to mitigate the desired signal power and isolates the target even if the undesired (interferences) signal is stronger compared to the signal of interest.

This chapter would provide a brief summary of some of the theories, statistical properties, and prerequisites upon which the Space-Time Adaptive Processing model has been developed. Comprehensive theory of STAP processing, challenges, characteristics of nonstationary systems and beamformering output have also been discussed. Finally, the basic concept of Least Mean Square and Wiener filter algorithms have been stated.

2.2 Statistical Properties

Statistical properties of the signal processing deals with the mean, covariance, probability, probability density function, autocorrelation and cross correlation functions. These properties are the backbone of the signal processing and detection; and parameter estimations hypothesis are being performed based on this properties. But, a fundamental limit on the performance based on this hypothesis is deviated from the optimal level due to the insufficient Independent and Identically Distributed (i.i.d) sample data set, errors in receiver elements, and the diverse nature of the interferences. As a result, the adaptive signal processing is now the subject of extensive research due to its capability of reducing the effects due to the diversification from the underlining assumption in order to work and adapt with the more real world environment. Therefore, it is very important to have a core concept of the signal properties from the statistical perspectives. Following subsections stated the fundamentals of these properties based on the works have already been done in [4][20][21].

2.2.1 Mean

Now, assumed that, $\mu_{\mathbf{X}}$ is the mean, also known as expected value, of random process $\mathbf{X}$. If $\mathbf{x}$ is vector, which contains L samples of random process $\mathbf{X}$, then it can be written:

$$\mathbf{x} = [x_1 \quad x_2 \ldots x_L]^T \tag{2.1}$$

and the expected value or the mean:

$$\mu_{\mathbf{x}} = \mathbb{E}[\mathbf{x}] = [E[x_1] \quad E[x_2] \ldots E[x_L]]^T \tag{2.2}$$

Now consider another random process Y, and y is vector of length L of that process with mean $\mu_{\mathbf{y}}$, then mean of the two random variables is the sum of their means, and can be stated as:

$$\mu_{\mathbf{x}+\mathbf{y}} = \mu_{\mathbf{x}} + \mu_{\mathbf{y}} \tag{2.3}$$

2.2.2 Covariance

The covariance can be stated as a measure of how two non-identical variables change together. If the two variables are identical, then this property is called as the variance. The covariance matrix can be defined as:

$$\begin{aligned} \mathbf{Q}_{\mathbf{xy}} &= \mathbb{E}[(\mathbf{x} - \mu_x)(\mathbf{y} - \mu_y)^H] \\ &= \mathbb{E}[\mathbf{x}\mathbf{y}^H - \mu_x\mathbf{y}^H - \mathbf{x}\mu_y + \mu_x\mu_y^H] \end{aligned} \tag{2.4}$$

In the case of zero means, covariance can be stated as:

$$\mathbf{Q}_{\mathbf{xy}} = \mathbf{R}_{\mathbf{xy}} \tag{2.5}$$

Using these relationships, variance:

$$\mathbf{Q}_{\mathbf{xx}} = \mathbf{R}_{\mathbf{xx}} \tag{2.6}$$

Where, $(.)^H \quad R_{xx} \quad and \quad R_{xy}$ are the Hermitian transpose, autocorrelation and cross-correlation respectively, and have been defined in next section.

The standard deviation can be stated as a measure of the variability or dispersion of a data set. It is represented as a square root of the variance (or covariance). Therefore standard deviation:

$$\sigma_{\mathbf{x}} = \sqrt{Q_{\mathbf{xx}}} \tag{2.7}$$

2.2.3 Cross Correlation

Let us consider two stochastic processes X and Y. Then the correlation between two stochastic processes is called the cross correlation, and can be stated as:

$$R_{\mathbf{xy}} = \mathbb{E}[\mathbf{xy}^H] \tag{2.8}$$

The autocorrelation of a stochastic process describes the correlation between the process at different points in time. Hence can be stated as:

$$R_{\mathbf{xx}} = \mathbb{E}[\mathbf{xx}^H] \tag{2.9}$$

Therefore, according to equation [2.5, 2.6, 2.8 *and* 2.9], it can be concluded that, at mean zero; variance and covariance are equal to autocorrelation and cross correlation respectively.

2.2.4 Posterior Probability Density Function

Probability theory is concerned with analysis of random phenomena. It is a possibility, and a way of expressing knowledge or belief that an event will occur or has occurred. If probability is a number that describes a set. Higher the number, the more probability of occurring event there is. It is a measure of how likely it is that some event will occur. Typically, probability can be denoted by $P[.]$. Conditional probability on the other hand is defined as the knowledge of event or occurrence of X, given the occurrence of Y. So, conditional probability can be stated as [21]:

$$P[X|Y] = \frac{P[XY]}{P[Y]} \tag{2.10}$$

Where, $P[Y] > 0$

Now, consider, sample vectors $\mathbf{x}$ and $\mathbf{y}$ of a stochastic process X and Y respectively. Then the posterior probability density function (pdf) of $\mathbf{x}$ conditioned on $\mathbf{y}$ is:

$$p[\mathbf{x}|\mathbf{y}] = \frac{p[\mathbf{y}|\mathbf{x}]p[\mathbf{x}]}{p[\mathbf{y}]} \tag{2.11}$$

This is also known as a Bayes' theorem.

If the random vector $\mathbf{x}$ follows a multivariate gaussian pdf $p_{\mathbf{x}}(\mathbf{x})$, then it can be written:

$$p_{\mathbf{x}}(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{L}{2}}|Q_{xx}|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\mu_x)^H Q_x^{-1}(\mathbf{x}-\mu_x)} \tag{2.12}$$

where $|Q_x|$ is the matrix determinant.

2.2.5 Ergodic Process

A stochastic process is termed to be ergodic if its statistical properties can be deduced from a single, sufficiently long sample of the process. In other words, it is the process when time average of the samples approaches the ensemble average. As for example, in signal processing, the process can be stated as a ergodic, if the mean of the snapshot of the process is equal to the true mean of that process for all the snapshots for $-\infty < t < \infty$.

2.2.6 Independence, Correlation, and Disjoint

The random events $\mathbf{x}$ and $\mathbf{y}$ is said to be independent if and only if their joint pdf can be written as:

$$p_{xy}(\mathbf{x}, \mathbf{y}) = p_x(\mathbf{x})p_y(\mathbf{y}) \tag{2.13}$$

if x and y have non-zero probabilities, then this implies:

$$p_{xy}(\mathbf{x}|\mathbf{y}) = p_x(\mathbf{x}) \tag{2.14}$$

and

$$p_{yx}(\mathbf{y}|\mathbf{x}) = p_y(\mathbf{y}) \tag{2.15}$$

In the case of disjoint events, it may consider:

$$p_{xy}(\mathbf{x} \cap \mathbf{y}) = 0 \tag{2.16}$$

In probability, disjoint and the independence are not exactly the same meaning. However, when any of $p_x(\mathbf{x})$ or $p_y(\mathbf{y})$ is zero, then disjoint and the independence are the same. Finally, random variables $\mathbf{x}$ and $\mathbf{y}$ are said to be uncorrelated if they satisfy the following relation:

$$\mathbb{E}[\mathbf{x}\mathbf{y}]^H = \mathbb{E}[\mathbf{x}]\mathbb{E}[\mathbf{y}]^H \tag{2.17}$$

2.3 Interferences

The principal challenge in the radar system is to mitigate the interferences due to thermal noises; and interferences because of cold (homogeneous) clutter, severe dynamic (heterogeneous) nonstationary multi-path hot clutter and jammer sources. Particularly, in the case of airborne radars, these noises and interferences play a critical role due to integrated influences of platform motion and antenna pattern. However, in the case of ground based radars, system can have a ridge at the zero Doppler.

System may consider two types of noise sources, effect the target signal; one is the internal and other is the external sources. External noise mostly is the environmental noise (thermal noise) and is received from the surrounding of the target of interest. On the other hand, internal noise (or receiver inherent noise) effect is superimposed on the desired signal received by the radar due to the radar receiver. However, with the advancement of the radar technology and high performance of the receivers, internal noise effect can be minimized at lower level compared to the other interferences.

Clutter can be stated as any object (unwanted) that may generate undesired target like echo signal for the radar professionals [24]. Typically, clutter is considered as a passive nature of interferences, since it usually appears in response of the radar signals [24]. Major sources of clutter echoes are typically returned from ground (objects), sea, rain, birds, animals, insects, chaff, decoys and atmospheric turbulence. Clutter can also be considered as a moving interference with respect to the clutter background [1]. Clutter may also be caused by the multi-path echoes from the desired targets due to the reflection and refraction [22]. To isolate the target from this mixture, may be a great challenge, since the clutter sources and the target signal may share the same resolution element. However, differing doppler shifts between the target and clutter would be very efficient technique in order to mitigate the target Signal [24].

Jammer is known as an active interference, and it is initiated by the elements outside the radar and in general, it is unrelated to the radar signals. Furthermore, a jammer may be considered as a device that intercepts signal transmissions by creating interference. It is a common interference in radar communication systems. Jammer interference from the dedicated transmitter (jammer), can cause a masking interference and deception jamming in the form of noise and simulated echoes respectively

[24]. Detection and tracking in jammer environment is very challenging task, since jamming signal usually travels only from the jammer to the radar while in the case of radar signal, it travels in two ways, hence produces signal with significantly reduced power when received by the radar receivers [22,24]. Jammer has an added effect of affecting radars along other line-of-sights, due to the radar receiver's sidelobes (Jamming). Mainlobe jamming effect can be mitigated by narrowing the mainlobe solid angle, and sidelobe jamming effect can be improved by reducing receiving sidelobes in the radar antenna design.

2.4 Space-Time Adaptive Processing (STAP)

Space-Time Adaptive Processing (STAP) is a special set of signal processing method, which enhances the radars to detect targets that might otherwise be contaminated by the interferences. This processing is a linear combination or weighted sum of the input samples. The Space in STAP means that STAP weight, signal samples at each antenna array element, at one instant of time define an antenna pattern in space, the term Adaptive in STAP refers that STAP weights are computed to reflect the actual noise and interferences, due to the clutter and jammer effect in which radar finds itself, Time in the STAP process refers that STAP weights, applied to the signal samples at one antenna element defines a system impulse, and therefore a system frequency response.

Adaptive architecture are currently the main focus point as a means for reducing the susceptibility of the desired signal to the interferences, and STAP is an extension of the adaptive antenna processing method. Adaptive antenna methods, typically adjust the phase (directional pattern) and amplitude of the received signal aperture illumination in order to mitigate the signal power. On the other hand, Space-Time

Adaptive Processing (STAP) is a multi-dimensional adaptive signal processing technique, over a spatial-doppler domains. Processing in the STAP is based on the adaptation of sample data, training the sample and solving a set of linear equations. STAP typically processing two types of data. One is training data and are used to estimate the interference covariance matrix as well as adaptive weight vector. Other is the primary data or the test data on which detection and the parameter estimation is performed. Since, adaptation is done in spatial and temporal domains, training is usually performed in the range bins. Therefore, it is typically called 2D data processing methodology on 3D datacube.

In STAP, increment in the range bin within a particular Pulse Repetition Interval (PRI) is known the fast time samples, and those across the PRIs is called as slow time samples. If there are N antenna elements, and M pulses over the Coherent Processing Interval (CPI), collect the group of N samples associated with each pulse (over a particular CPI), or taking one sample simultaneously from each of the N antenna element, to make one snapshot (one column of Figure 2.1). So, there are MN snapshots for each range bin. The typical diagram for STAP datacube is shown in Figure 2.1.

A low sidelobe antenna with fixed interference does not need training hence the adaptation. But, processor often needs to detect and estimate the target of interest in the vicinity of diverse interferences. Therefore, adaptation and adjustment are the two crucial steps in STAP processing. Adaptation involves the estimation of the interference matrix from the Independent and Identically Distributed (i.i.d) training data vector and number of data vector for the estimation of covariance (asymptotic) matrix increases with the increase of matrix dimension. On the other hand, adjustment of the weight in a way that process can track the changes due to changes in the nature of the interferences. The main objective of the STAP is to isolate the

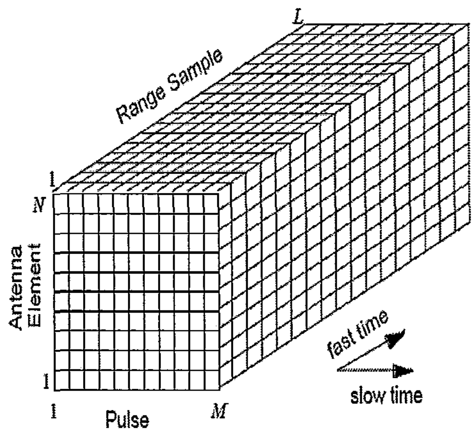


Figure 2.1: STAP datacube

signal from the interferences by adaptively estimating the weight vector and error covariance matrix, based on the sample data (training data) received from the spatial and temporal domains. Furthermore, STAP principle is also used to eliminate the effects of co-channel interference, and ISI in radar communication. It also uses multiple antennas, which typically requires to place at least half the wave length of the operating signal in order to mitigate the fading effects. The motivation of the STAP processing is its outstanding performance to isolate the slow and fast moving target from the interferences, even if the interference signals are stronger than the signal of interest.

2.4.1 Processing Procedures

Depending on how sample returns (input data) are processed, Space-Time Adaptive Processing can be classified into fully adaptive processing, subspace adaptive processing, and post Doppler adaptive processing.

In fully adaptive processing, system processes the sample data with full degree of freedom as obtained by the number of array elements and the pulses [1]. In fact, system estimates the weight vector, mitigates the target signal and rejects the interference covariance matrix on the way they have received it [1]. Fully Adaptive Processing requires to solve $M \times N$ linear equations. This is fully optimal process, however in the case of large sample data, or unavailability of sufficient number of Independent and Identically Distributed (i.i.d) sample data set, this technique faces challenges for adaption with the sample data set and computational complexity due to the estimation of the inverse of interference covariance matrix ($MN \times MN$). Typical operational principle of fully adaptive processing is given in the Figure 2.2 [70]:

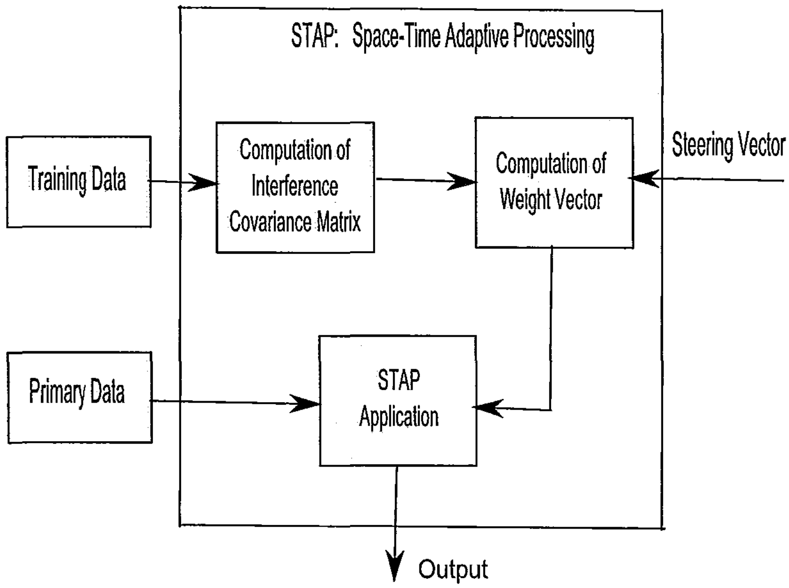


Figure 2.2: Fully Adaptive STAP Processing

In subspace technique, large covariance matrix needs to be transformed into discrete matrices so that computational load can be reduced, and also the challenges that are constituted for not having sufficient i.i.d sample data may be avoided. Matrix transformation can be done in spatial or temporal domain or both. This is suboptimal process, however this method is computationally efficient. Typical operational principle of partially adaptive processing is given in the Figure 2.3 [70]:

In post Doppler processing, Doppler filtering is performed before the STAP adaptation. This is also known as a frequency dependent processing where scanning is performed based on the Doppler bins. It is a special form of subspace technique, where $M \times N$-dimension filtering problems are transformed into M separate N-dimension problems. This technique is computationally efficient and needs fewer sample data than fully adaptive process. However, it is suboptimal and the performance is poor in comparison to full and typical subspace adaptive STAP techniques.

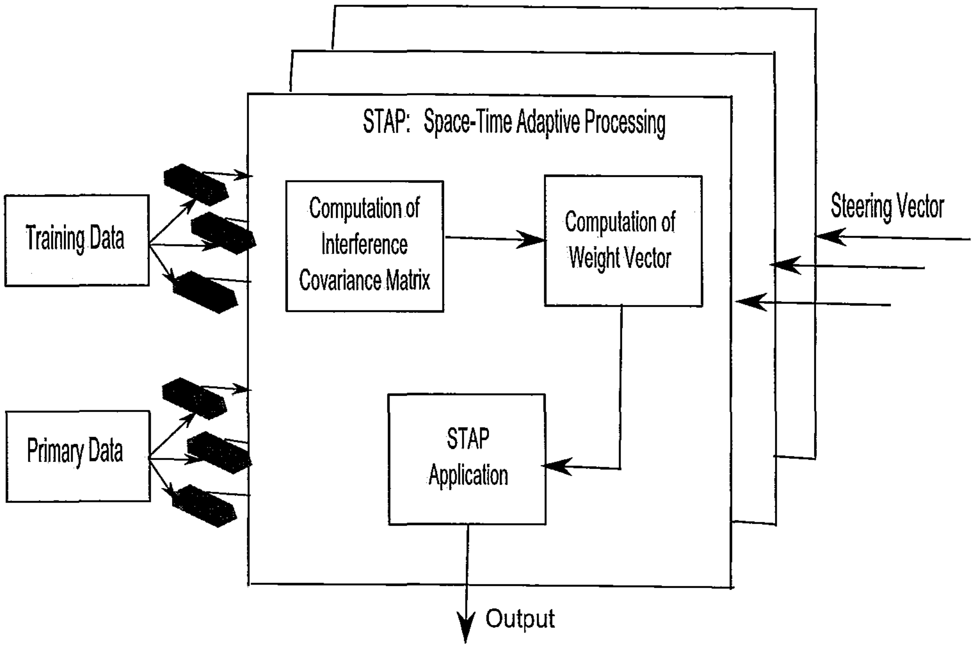


Figure 2.3: Partially Adaptive STAP Processing

2.4.2 STAP-Nonstationary Heterogeneous Systems

Nonstationary refers to the changes of properties over time, and it is usually a property of the dynamic systems. In STAP-processing, requirements of the Independent and Identically Distributed (i.i.d) training data set increase with the increase of the dimension of the covariance matrix. But, due to the nonstationary and heterogenous nature of the interferences, it might be a great challenge for having sufficient i.i.d sample data set for the estimation of interference covariance matrix as well as adjustment of the associated weight vector. Most of the time, the system model would usually consider that the radar interferences are stationary. But, in reality, spatial-temporal properties of the interferences may not be constant over the same Coherent Processing Interval (CPI). Therefore, the fluctuation in the adaptation of the spatial-temporal covariance (asymptotic) matrix over time (CPI) is highly predictable. In STAP processing, nonstationarity may be considered due to the range variation as well as variation of the direction of arrival over time i.e. over the same Coherent

Processing Interval (CPI). As a result the structure of the interference covariance matrix changes with Pulse Repetition Interval (PRI) over the same CPI. Therefore, the processor model must be able to incorporate adaption and adjustment in order to achieve effective suppression of interferences due to nonstationary heterogeneous clutter and jammer effects.

Detection of target in nonstationary heterogeneous clutter and jammer environment may involve the computation of data during each snap shot compared to the other. If the adjustment of weight vector is considered in the same time slot as the training data, there would be no impact due to time variation. Even extended clutter or jammer suppression might be possible for the diverse properties of the interferences due to nonstationary heterogeneous systems. Time varying adaptive weights by a linear extrapolation over a CPI due to moving interference may also be computed [17][18].

One of the greatest challenges in STAP processing is to estimate the asymptotic interference covariance matrix from the sample data set, because of insufficient i.i.d sample data set due to nonstationary and heterogeneous clutter and jammer characteristics, and the prior of the interference covariance matrix is unknown. More importantly, computational complexity increases with the increase of the size of the interference covariance matrix, since the accuracy of the STAP processing depends on the number of the sample; and size of the covariance matrix increases with the increase of the sample size. High computational burden is also the challenge for the future growth and expansion of the STAP method. Furthermore, the Maximum Likelihood Estimation (MLE) function stated in equation (1.1), may be used to estimate the likelihood of interference (asymptotic) covariance matrix, even though the prior of the interference covariance matrix not known. The computational bottle neck can also be overcome by using the partially adaptive (parallel processing) STAP-processing.

Finally, motivation of the STAP processing is its ability to enhance the radar receiver to mitigate the desired signal power and isolate the target even if the interference signal is stronger compared to the target of interest.

2.5 Adaptive Beamforming

Adaptive beamforming is a signal processing technique, which widely used in antenna arrays for spatial signal transmission and reception. In this beamforming technique, array of antennas is exploited to achieve maximum reception in a specified direction in order to estimate the signal arrival from a desired direction as well as obtained the optimal signal power, while rejecting the signal from other directions. The main distinction between an adaptive beamformer and a conventional beamforming systems is that, adaptive beamforming has the ability to automatically adapt with the interferences, and suppress these interferences while simultaneously enhancing and optimizing the performance of the radar reception without prior knowledge of the signal and noise environment. It is highly reliable beamforming technique, since the spatial characteristics of the signal can be automatically adjusted until the array sidelobes are reduced at the optimal level. On the other hand, conventional methods are highly vulnerable to the undesired interferences, and the performance is significantly degraded with the increase of the array sidelobes. Figure 2.4 shows the principle configuration of adaptive beamforming method:

According to the Figure 2.4, now, assume that the direction of arrival is θ, number of antenna array elements N, operating wave length λ, spacing between two elements is d. Therefore, the received signal at the Nth element of the antenna array can be stated as [1][2]:

$$x_n = e^{j2\pi(N-1)\frac{d}{\lambda}sin\theta} \tag{2.18}$$

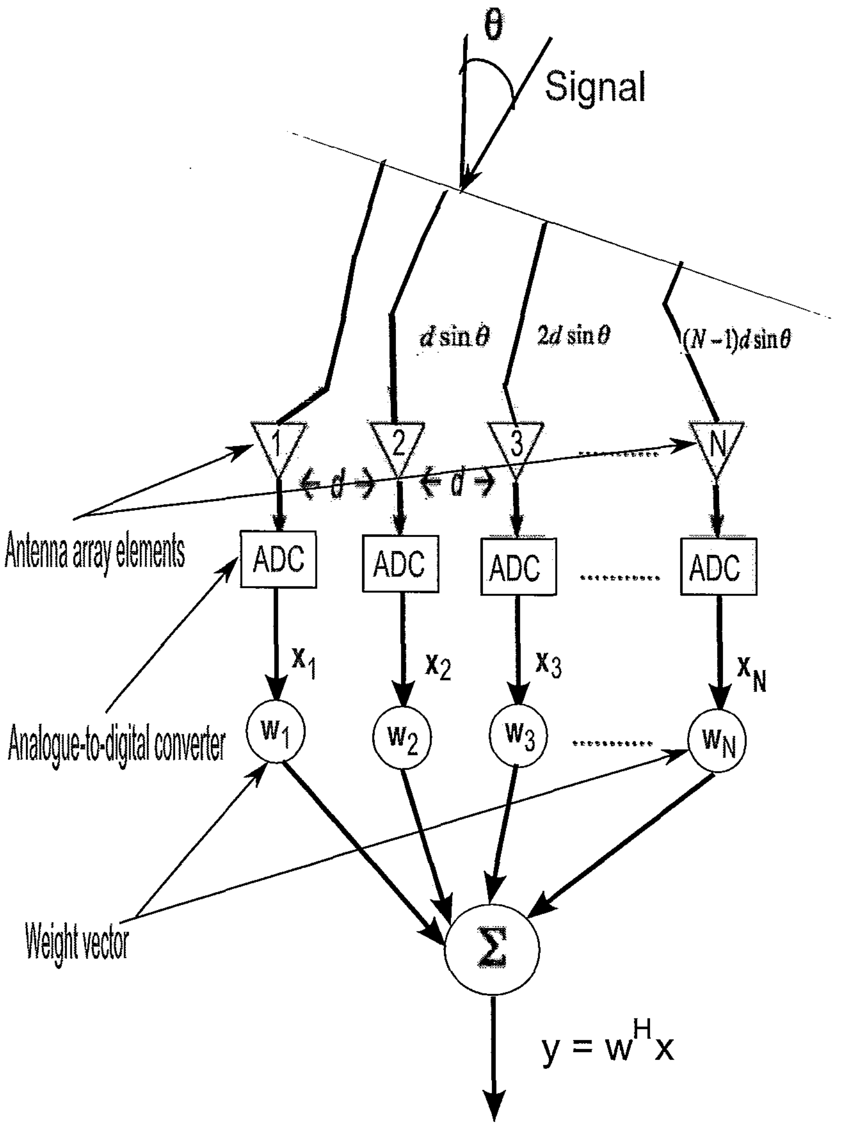


Figure 2.4: Adaptive Beamforming Systems

Now, consider the associated weighting factors (vectors) $\mathbf{w}$ for each of the array element, then the scalar output of the beamformer can be stated as [1][2]:

$$\begin{aligned} y &= \sum_{n=1}^{N} w_n^H x_n \\ &= \mathbf{w}^H \mathbf{x} \end{aligned} \tag{2.19}$$

In STAP-processing, adaptive beamforming, typically means a data dependent modification of the array elements. In this processing, the system usually considered two types of data: one is the training data and other is the primary data. Adjustment in the adaptive beamforming is achieved by varying the weights of the associated array elements, and by taking the weighted sum of the received signals at all of the array elements [25]. In the case of STAP processing, the initial weight vector would be approximated from the likelihood of the interference (asymptotic) covariance matrix, and optimal weight vector can be achieved by recursively or the iteratively training the sample data set based on underlying processing algorithms, and detection and estimation is performed with primary data set. The operation of an adaptive beamforming system can be illustrated, by considering a reference signal and comparing the signal with the estimated signal. Each time the difference (error) from this comparison is calculated, and based on this error, system needs to adjust or estimate the associated coefficient. This process is continued until it fulfils the required criteria of the processing algorithm. To overcome the challenges due to the nonlinear and nonstationary environment, sequential state estimator filtering approach, i.e., Kalman filter may be considered.

2.6 Wiener Filter and Least Mean Squared

In the statistical linear filtering solution, a simple linear filtering approach would be anticipated in order to cancel out the noise and mitigate the signal power. Wiener filter, proposed by Norbert Wiener, is a fundamental building block of optimum linear filters, which involves linear estimation of a desired signal. It is linear optimal discrete time invariant filter. The principle approach of this filtering solution, is to minimize the error due to the differences (noise) between the reference signal and the estimated signal because of the stationary input. It uses Mean Squared Error (MSE) approach in order to minimize the difference and optimize the signal of interest. Figure 2.5 is the generic block diagram of Wiener filter. The input data $\mathbf{x}(n)$ is the combination of signal and the noise, and $d(n)$ is the desired signal. The Wiener filter is inadequate for dealing with non-stationarity systems (signal and interferences), since it is optimal filter under the assumption that the operating environment is linear and stationary.

The Least Mean Squared (LMS) algorithm is an adaptive gradient based method of steepest decent, introduced by Widrow and Hoff. LMS uses an iterative procedure in order to make successive corrections to the weight vector in the direction of the negative of the gradient vector which eventually leads to the minimum mean square error. There is no cross-correlation or expectation required for least mean square. Typically, it does not consider any stochastic nature of the signal and statistical assumption, however there may be a similarity (with Wiener filter) in weight solution.

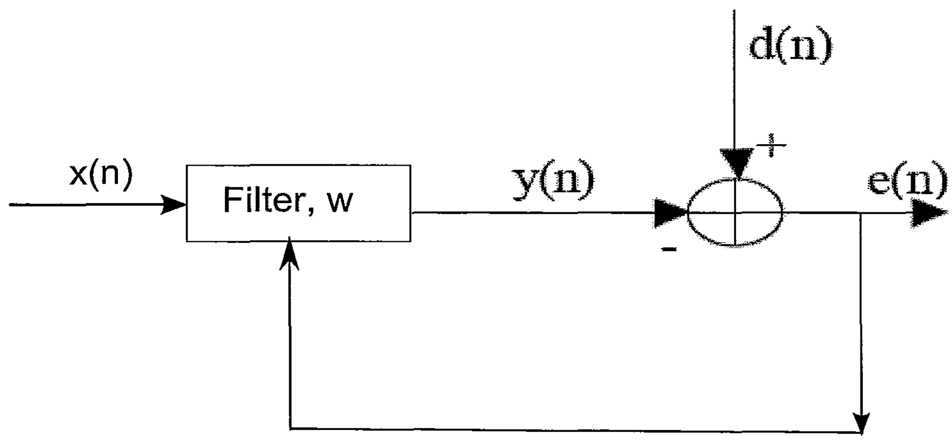


Figure 2.5: Block Diagram of a Weiner Filter

Chapter 3

Signals Overview for STAP with Adaptive State Estimation (STAP-ASE)

Adaptive signal processing is the focal area for reducing the susceptibility of the signals to the interferences because of its automatic sensation and rejection of the interferences without prior knowledge of the signal environment. Space-Time Adaptive Processing (STAP) is a multi-dimensional adaptive signal processing technique, which estimates adaptive weight vectors in spatial and Doppler domains for which a target detection hypothesis is to be performed. STAP operates on the set of returns, composed of pulses, array elements, and range bins, over a period of time. Hence, STAP is a 2D processing on 3D datacube, collected from the available signal received at the radar receiver. Since, adaptation is performed in spatial and temporal domains, training is done using the range bins. Typically, it is a sample based approach where covariance (asymptotic) matrix may be estimated from Independent and Identically Distributed (i.i.d) sample data. The main theory of this processing is to adapt with the sample data set in order to estimate thc interference covariance matrix and adjust or update the weight vector in a way that, the noise power can be minimized, and Signal to Interference-plus-Noise Ratio (SINR) can be maximized in some appropriate sense.

In this chapter, the prerequisites of problem formulation for STAP-ASE have been stated. One of the greatest challenges in STAP processing is to estimate the interference covariance matrix from the sample data set, because of insufficient i.i.d sample data set due to nonstationary, and heterogeneous clutter and jammer characteristics. More importantly, the prior of the interference covariance matrix is unknown. Therefore, estimation of the interference covariance matrix, problem formulation, STAP-signal characterization, signal and interference models are also being stated in this chapter. The improvement factor, states the performance and efficiency of the estimator, and mathematical formulation for Minimum Variance Distortionless Response (MVDR) have also been included in this chapter.

3.1 Introduction

In STAP processing, it is usually assumed that the training (secondary/auxiliary) data samples are free from target signal (unsupervised). But, in reality, there is a high probability that the training data set is being contaminated by the target signal. Therefore, overall performance, efficiency, and accuracy of the STAP processing depends on the appropriate selection of auxiliary data set. It is crucial to have a target free secondary data for computing the covariance matrix of the range cell under consideration (tcst), so that the cancelation of desired target signal or self-nulling may be avoided. Therefore, the filter needs to be designed in a way, that would be able to isolate the target like signal from the sample data set before starting the training process. The sample data obtained by this way is called as semi-supervised. STAP processing not only mitigate the combined interference effects from nonstationary hot and cold clutter and jammer, but also enhance the detection of small and slow moving target, and also provides robustness in the presence of undesired signals.

STAP involved high computational complexity for the adaptation, estimation and filtering of the signal of interest. Complexity due to adaptation and the estimation of the inverse of the interference covariance matrix increases with the increase of the sample data vector, since dimension of the interference covariance matrix is proportional to the data vector [1]. On the other hand, filtering typically involves sequentially multiplying the received data (steering vector) with the inverse of the estimated covariance (asymptotic) matrix for all the range bins for all possible Doppler frequencies.

Proposed STAP-ASE model is based on the minimization of the interference covariance matrix and maximization of the Signal to Interference-plus-Noise Ratio (SINR) in the Mean Squared Error (MSE) sense. Therefore, the system can decode or isolate the signal of interest from the interferences due to clutter and jammer sources. During the weight training process, there may be no limit on the value of the weight vector [4], so that the constraint on the weight value may introduce consistency on the estimation process [4]. Therefore, added constraint to weight vector would accelerate the system convergence by limiting the weight vector within an acceptable range and may also keep the relation stated in equation (3.13). Constraint weight vector may be considered as a key tool to cancel out the interferences (hot clutter, cold clutter, and noise), so that the system can converge to its optimal point without affecting the original signal characteristics and at the same time maintain the rigorous reliability of the systems.

3.2 Signal Characterizations

One of the greatest challenges in STAP processing is to estimate the interference covariance matrix from the i.i.d sample data set, since the prior of the matrix is not known. However, using the Maximum Likelihood Estimation (MLE) function, processor may estimate the interference covariance matrix. In this approach, the input data from all the surrounding range bins except the bin under test has been considered, and it may also be assumed that there are no target signal in the surrounding range bins (unsupervised or null hypothesis:-Appendix-D). But there are two main possibilities for having target signal in the secondary data set. One is perfect matched with the target signal response, and other is mismatched. In the proposed approach, the perfectly matched target like signal (semi-supervised) is being considered and filtered out from the secondary data set in order to estimate the maximum likelihood of the interference covariance matrix. Hence, the maximum likelihood of interference covariance matrix can be stated as [1]:

$$Q_i = \frac{1}{L}\sum_{l=1}^{L}\mathbf{x}_l\mathbf{x}_l^H \tag{3.1}$$

where,$L \geqq MN$ is the total number of samples being used for the ith range bin, $\mathbf{x}$ is the input data vector may have noise only or target with noise and $(.)^H$ represents the hermitian transposition. The dimension of $\mathbf{x}$ is $\geqq MN$.

3.2.1 Signal Models

Diverse nature of the signal properties due to the various sources, signal source may be considered as inherently random. Thermal noise, interferences due to clutter and jamming effect are also random in nature. The main concern is to process the received signal in a way that the overall detection performance of the desired signal can be improved. The adaptive methodology may exploit signal characteristics, and the nature of the signal sources in order to achieve such improvement. Therefore, it is very important to consider the signals and interferences model in different context. On the other hand, advantage for having the random signal is such that, it is plausible to assume a Gaussian random process, since the statistical properties of the Gaussian signals are more desirable because the first and second moment of the process would be able to provide a complete characterization of the signal.

Now, consider the complex desired (joint domain) steering vector $\mathbf{s} \in C^{MN}$, where N is the number of array elements and M is the number of pulses. If the maximum likelihood of interference covariance matrix is Q (its dimension depends on the dimension of input $\mathbf{x}$) with zero mean Gaussian and multivariate complex Gaussian distribution, and corresponding complex weight vector is $\mathbf{w} \in C^{MN}$, then

$$\begin{aligned} \mathbf{s} &= [s_1, s_2, \ldots, s_{MN}]^H \\ \mathbf{w} &= [w_1, w_2, \ldots, w_{MN}]^H \end{aligned} \tag{3.2}$$

The optimal weight vector for the signal with steering vector $\mathbf{s}$ and the asymptotic interference covariance matrix of the processor is Q, then the relation between them can be stated by the following well-known equation [1]:

$$\mathbf{w}_{opt} = Q^{-1}\mathbf{s} \tag{3.3}$$

Where, Q and $\mathbf{w}_{opt}$ are (unknown) expected interference covariance and weight vector respectively and depend on the adaption and training of the sample data.

Now, if $\mathbf{x} = [x_1, x_2, \ldots, x_{MN}]$ is the signal (with noise+clutter+jammer interference) from the array elements then the main challenge is to isolate the desired signal with steering vector s from the interferences by the adaptation of thc sample data, estimation of covariance matrix and adjustment of the weight vector.

Now assume that the complex weight vector w_{mn} for each receiving channel of array of N elements with equal space d (for M pulses) as shown in Figure 3.1. Now

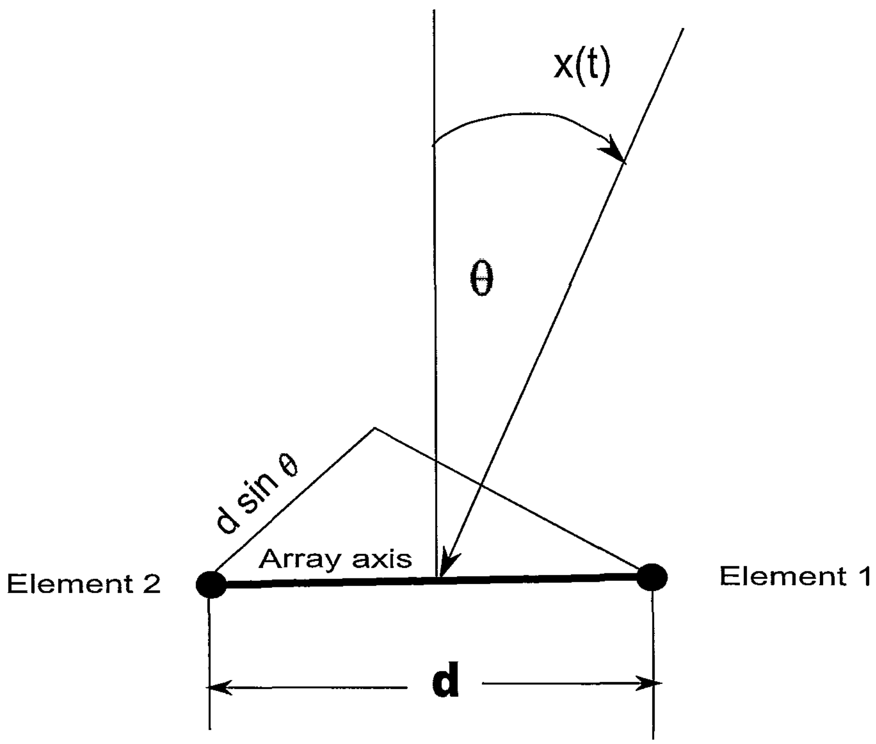


Figure 3.1: Pair of Identical Array Elements

consider the beamformer output of the process is scalar, y. Therefore, the output y can be stated as [1]:

$$\begin{aligned} y &= \sum_{nm=1}^{MN} w_{nm}^H x_{nm} \\ &= \mathbf{w}^H \mathbf{x} \end{aligned} \tag{3.4}$$

Received signal is always contaminated by the noise and the interferences due to the clutter and jammer effects. If $\mathbf{x}$ has the noise component $\Omega \in C^{MN}$, equation (3.4) may be restated as:

$$\begin{aligned} y &= \mathbf{w}^H(\mathbf{s} + \Omega) \\ y &= \mathbf{w}^H\mathbf{s} + \mathbf{w}^H\Omega \end{aligned} \tag{3.5}$$

The output power of the process [1]:

$$P_{out} = \mathbb{E}[yy^H] = \mathbb{E}[\mathbf{x}^H\mathbf{w}\mathbf{w}^H\mathbf{x}] \tag{3.6}$$

The signal power at the output [1]:

$$P_s = \mid \mathbf{w}^H\mathbf{s} \mid^2 \tag{3.7}$$

and the noise power at the output [1]:

$$\begin{aligned} P_n &= \mathbb{E}[\|\mathbf{w}^H\Omega\|^2] \\ &= \mathbf{w}^H\mathbb{E}[\Omega\Omega^H]\mathbf{w} \\ &= \mathbf{w}^H Q\mathbf{w} \end{aligned} \tag{3.8}$$

since, $Q = \mathbb{E}[\Omega\Omega^H]$
and $\|\ .\ \|$ represents the Euclidean-norm.

Signal to Interference-plus-Noise Ratio (Power),

$$SINR = \frac{P_s}{P_n} = \frac{\mid \mathbf{w}^H\mathbf{s} \mid^2}{\mathbf{w}^H Q\mathbf{w}} \tag{3.9}$$

Based on the STAP phenomenon, the processor needs to minimize the noise power while maximizing the signal response in some appropriate sense. Therefore, the system can isolate or decode the signal of interest.

The Signal to Interference-plus Noise Power ratio may be restated as:

$$SINR = \frac{P_s}{P_n} = \frac{|\mathbf{w}^H\mathbf{s}|^2}{\mathbf{w}^H Q\mathbf{w}} = \frac{|\mathbf{w}^H Q^{1/2}Q^{-1/2}\mathbf{s}|^2}{\mathbf{w}^H Q\mathbf{w}} \tag{3.10}$$

Using the Schwarz's inequality, equation (3.10) may be written as:

$$\frac{P_s}{P_n} = \frac{|\mathbf{w}^H Q^{1/2}Q^{-1/2}\mathbf{s}|^2}{\mathbf{w}^H Q\mathbf{w}} \leq \frac{(\mathbf{w}^H Q\mathbf{w})(\mathbf{s}^H Q^{-1}\mathbf{s})}{\mathbf{w}^H Q\mathbf{w}} = \mathbf{s}^H Q^{-1}\mathbf{s} \tag{3.11}$$

At the optimum condition, the upper bound can be considered, and hence:

$$(\frac{P_s}{P_n})_{opt} = \mathbf{s}^H Q^{-1}\mathbf{s} \tag{3.12}$$

3.2.2 Constraint Weight Vector

The main objective of the proposed model is to estimate the weight vector and the interference covariance matrix from the adaption of training data set in a way, that the noise power can be minimized and at the same time SINR can be maximized in the MSE sense. Therefore, the system may decode or isolate the signal of interest. Now, consider that, putting the constraint on the weight vector would allow the system to obtain this goal. Now, the equation for the constraint can be stated as:

$$\mathbf{c}^H\mathbf{w}_c = g \tag{3.13}$$

Where, $\mathbf{w}_c = [w_{c1}, w_{c2}, \ldots, w_{cMN}]^T$ is the constraint weight vector, $\mathbf{c}$ is the predefined steering vector whose elements are stated by look direction and signal Doppler frequency, and g is the constant gain due to the associated signal steering vector.

The constraint to the weight vector would also ensure that the characteristics (such as signal energy) of the signal would be unchanged, while minimizing the effects of the interference covariance matrix [1]. Now, constrain the gain of the processor at g on the prescribed direction and frequency, and null the other directional and Doppler interferences without distorting the signal of interest.

The optimal solution for the constraint weight [complete derivation in Appendix C] vector can be stated as:

$$\mathbf{w}_c = Q^{-1}\mathbf{c}(\mathbf{c}^H Q^{-1}\mathbf{c})^{-1} g \tag{3.14}$$

According to STAP principle, interference covariance matrix needs to be estimated using the sample data from the surrounding range cells except the cell under consideration (test cell). Therefore, they system needs to nullify the signal influence due to the test cell when estimating the interference covariance matrix. Now, consider a a matrix $U \in C^{MN \times MN}$, and which must also satisfy following condition,

$$\mathbf{c}^H U = 0 \tag{3.15}$$

where, 0 is $MN \times 1$ zero vector. Therefore, constraint vector $\mathbf{c}$, constraint weight vector $\mathbf{w}_c$, and null block with matrix $U(t)$ are formulated in a way, which would satisfy the above constraints.

3.3 Target Signal

The desired target signal for the proposed model depends on the weight vectors ($\mathbf{w}_c = \mathbf{w}$), and it is also apparent from equation (3.3), the weight vector depends on the asymptotic interference covariance matrix Q. However, if the system has the signal of interest $\mathbf{s}(t)$ of the moving target with radial velocity v_{rad} for the associated spatial and temporal steering vectors, then desired signal can be formulated [1] as:

$$\begin{aligned} \mathbf{v}_s &= e^{j\frac{2\pi}{\lambda}(x_i\ cos\phi_t + y_i\ sin\phi_t)cos\theta_t - z_i\ sin\theta_t} \\ \mathbf{v}_t &= e^{j\frac{2\pi}{\lambda}2v_{rad}mT\ cos\phi_t\ cos\theta_t} \\ f_{tarD} &= \frac{2v_{rad}}{\lambda}cos\phi_t\ cos\theta_t \\ \mathbf{s}(t) &= \mathbf{v}_t \otimes \mathbf{v}_s \end{aligned} \tag{3.16}$$

Where, $\mathbf{v}_t$, $\mathbf{v}_s$, and $\otimes$ are the temporal steering vector, spatial steering vector and kronecker product respectively. The terms ϕ_t, θ_t λ, f_{tarD}, T are azimuth angle, depression angle, operating wave length, target Doppler frequency, PRI respectively. Subscript t represents the position of the target at the associated angles. $m = 1, 2, \ldots, M$ and $i = 1, 2, \ldots, N$.

The tangential component of the target velocity may also be considered, however for large range and the small number of echo pulses this can be neglected. For the same reason, radial component of the target can also be assumed constant, since the azimuth angle interval passing by the desired target at the time of observation is negligible. However, in the case of long pulse sequences, radial velocity varies pulse to pulse [1].

Furthermore, if the space time signal vector $\mathbf{s}(t)$ used for the model is under the assumption of equation (3.16), then they system may consider that the model is perfectly matched to the expected signal, as a result the maximum processor performance can be anticipated in order to suppress the interferences due to clutter.

3.4 Interferences

Immunity to the multipath interference needs to be considered in order to enhance the detection performance, since interference environment always changes due to the severely diverse nature of spatial-temporal properties. Received signal in the nature is contaminated by both of white as well as the colored noises due to clutter and the jammer interferences. As the communication load increases, the detection of target due to these interferences is getting a big concern.

So, it is a paramount importance to understand the interference model in order to isolate and suppress these interferences from the target of interest. The total interferences may be stated as:

$$Q(t) = Q_n + Q_c(t) + Q_j(t) \tag{3.17}$$

where, Q_n is noise part; Q_c and Q_j are the interferences due to clutter and jammer effects respectively.

Typical configuration of omnidirectional array with two array elements where signal and interferences are arrived at certain angles is shown in Figure 3.2.

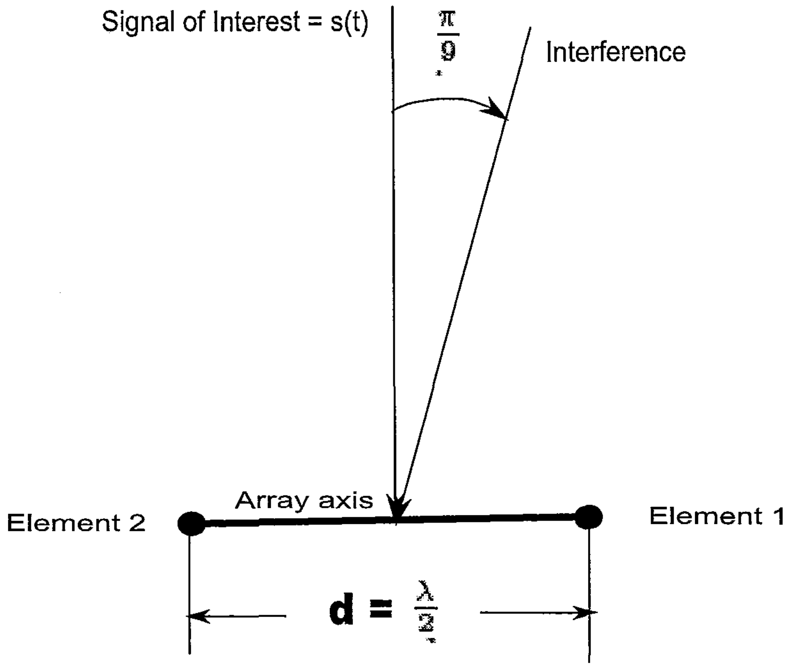


Figure 3.2: Signal and Noise mixture process

3.4.1 Noise

Noise can be defined as the variation in, and the addition of the unwanted factors to the stream of the desired target signal. It can be categorized as the white noise and the colored noise; and can also be considered as a random signal. White noise typically contains the equal power at any central frequency and within a fixed bandwidth, where power spectrum density is distributed with visible properties. This white noise may typically be considered as uncorrelated with the spatial domain (signal), and it has covariance with diagonal matrix; and may also be considered as a thermal noise received by the antenna or the receiver inherent noise. The colored noise on the other hand, is due to the presence of the interferences because of the clutter and jammer effects, and correlated with signal.

3.4.2 Clutter

Clutter may describe any source or object which can cause unwanted signal. It is a passive interference, received by the radar. The unwanted signal that comes through the antenna's main-lobe is being considered as a main-lobe clutter, and the others are being considered as a side-lobe clutter signal. The Figure 3.3 shows the typical antenna radiation pattern for side-lobes, and main-lobes [28]. Clutter and target of interest may be situated in the same radar resolution cell [27], hence angular and range selection method do not ensure sufficient suppression of clutter. Alternate technique to mitigate the target signal is to consider the difference in velocity between the target and the interference due to clutter, since it may cause different Doppler shift between the desired signal and the interferences. In space time processing, clutter rejection is being considered on a pulse-to-pulse basis, and the process is carried out for one range increment only under the assumption that, the clutter echoes are independent, identical, and asymptotically Gaussian. Furthermore, if the system considers that the clutter background is moving at the radial direction with velocity v_c, then additional clutter velocity needs to be considered with the platform velocity [1][2].

Statistical characteristics of the clutter returns mainly depended on the nature of each return, however, the clutter fluctuation can be modeled by means of Gaussian phenomenon [1]. Steering vectors due to the spatial and temporal properties for airborne radar associated with each direction because of clutter, can be stated as [1]:

$$\begin{aligned}
\mathbf{q}_{cs}(\phi) &= e^{j\frac{2\pi}{\lambda}(x_i\ cos\phi + y_i\ sin\phi)cos\theta - z_i\ sin\theta} \\
\mathbf{q}_{ct}(\phi, v_p, v_c) &= e^{j\frac{2\pi}{\lambda}(v_p\ cos\phi\ cos\theta\ +\ v_c)2mT} \\
f_{cD} &= \frac{2v_p}{\lambda} cos\phi\ cos\theta \\
\mathbf{q}_c(t) &= \mathbf{q}_{ct} \otimes \mathbf{q}_{cs} \\
Q_c &= \mathbb{E}[\mathbf{q}_c \mathbf{q}_c^H]
\end{aligned} \tag{3.18}$$

Where, θ, ϕ, v_c, v_p, f_{cD} and T are the depression angle, azimuth, radial clutter velocity, platform velocity, clutter Doppler frequency and PRI respectively.

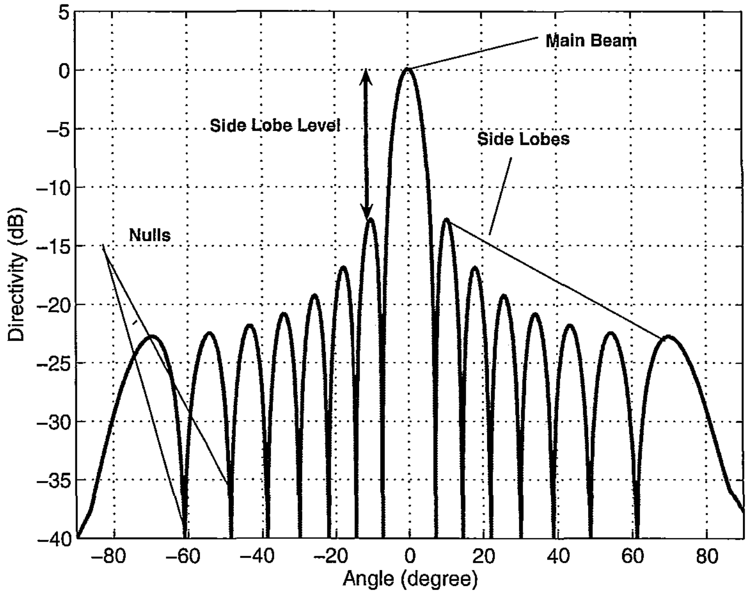


Figure 3.3: Antenna Pattern Showing Main-lobes and Side-lobes

3.4.3 Jammer

Jammer may be considered as an active interference such as masking interference and deceptive jamming [27]. Now, for jammer, system may consider only the channel mismatch due to spatial properties of the process. So, steering vector due to the spatial and temporal properties for airborne radar because of jammer, can be stated as [1]:

$$
\begin{aligned}
\mathbf{q}_{js}(\phi) &= e^{j\frac{2\pi}{\lambda}(x_i\ cos\phi_l + y_i\ sin\phi_l)cos\theta_l - z_i\ sin\theta_l} \\
\mathbf{q}_{jt} &= [1, 1_2, \ldots, 1_M]^H \\
\mathbf{q}_j(t) &= \mathbf{q}_{jt} \otimes \mathbf{q}_{js} \\
Q_j &= \mathbb{E}[\mathbf{q}_j \mathbf{q}_j^H] \qquad (3.19)
\end{aligned}
$$

where, l represents the location of the jammer at the associated angle.

Typically, the processor performed the jammer and clutter rejection in two separate steps. Since the jammer is being considered as spatial interference, first step before the transition, spatial jammer interference covariance matrix (free of clutter) in passive filter mode can be estimated. Second step is being considered after the transition, i.e, after the anti-jamming filtering. As a result, for the jammer rejection, the processor typically won't need STAP-filter, since jammer cancelation is being completed in the spatial only domain. Therefore, for STAP-processing, filter just has to cope with the clutter only colored interference if the jammer rejection is perfectly done before transmission.

3.5 Improvement Factor (IF)

In the radar signal processing, interference suppression is really a computational burden, however, performance of achieving clutter notch with some of the processors is nearly close to the optimal. The performance and the efficiency of the processor may be estimated by the Improvement Factor (IF). It is defined as the ratio of Signal to Noise Power ratios at the output and input. It can be stated as[1]:

$$\begin{aligned} IF &= \frac{\frac{P_{os}}{P_{on}}}{\frac{P_{is}}{P_{in}}} \\ &= \frac{\mathbf{w}^H \mathbf{s}\mathbf{s}^H \mathbf{w}.tr(Q)}{\mathbf{w}^H Q \mathbf{w}.\mathbf{s}^H \mathbf{s}} \end{aligned} \tag{3.20}$$

Where, P_{os}, P_{on}, P_{is} *and* P_{in} represent output signal power, output noise power, input signal power and input noise power respectively.

Equation (3.20) characterizes the efficiency or gain of any linear processor. The performance of the processor can be stated as:

$$\mathbb{P}_{IF} = \frac{IF_{op}}{IF_{ST}} \qquad (3.21)$$

where, IF_{opt} *and* IF_{ST} are the Improvement Factors (IF) for optimal processor and STAP processor (under consideration) respectively. It represents how much processor performance degraded by interference, and the expected attainable value for $\mathbb{P}_{IF}$ is unity.

3.6 Minimum Variance Distortionless Response (MVDR)

The focal point of the most adaptive techniques used today were in the Minimum Variance Distortionless Response (MVDR). According to section 3.2.1, the beamformer output:

$$\begin{aligned} y(t) &= \mathbf{w}^H(\mathbf{s}(t,\theta,f_d) + \Omega(t)) \\ y(t) &= \mathbf{w}^H\mathbf{s}(t,\theta,f_d) + \mathbf{w}^H\Omega(t) \end{aligned} \qquad (3.22)$$

The main goal is to extract $\mathbf{s}(t,\theta,f_d)$ from $y(t)$, i.e., $y(t) \approx \mathbf{s}(t)$. Therefore,

$$\begin{aligned} &\min_{\mathbf{w}} \quad \mathbf{w}^H Q_v\, \mathbf{w} \\ &Subject\ to: \quad \mathbf{w}^H\mathbf{s}(\theta,f_d) = 1 \end{aligned} \qquad (3.23)$$

where Q_v and $\mathbf{s}(t, \theta, f_d)$ are known. If variance Q_v can be replaced with the estimated variance Q from the processor, then Minimum Variance Distortionless Response (MVDR) is the variation on equation (3.23).

Using the Lagrange multiplier, the optimal MVDR weight vector with the imposed constraint stated in the above sections can be stated as:

$$\mathbf{w}_{MVDR} = \frac{Q^{-1}\mathbf{s}(\theta, f_d)}{\mathbf{s}^H(\theta, f_d)\; Q\; \mathbf{s}(\theta, f_d)} \tag{3.24}$$

and using the output from the proposed model, the form of equation (3.23) can be stated as:

$$\begin{aligned} &\min_{\mathbf{w}} \quad \mathbf{w}^H Q\, \mathbf{w} \\ \textit{Subject to}: \quad &\mathbf{w}^H \mathbf{s}(\theta, f_d) = 1 \end{aligned} \tag{3.25}$$

Hence, Minimum Variance Distortionless Response (MVDR) stated in (3.23) is the optimal solution of equation (3.25).

In summary, the Minimum Variance Distortionless Response (MVDR) is a technique, which minimizes the variance of the observed signal, while simultaneously passing the desired signal.

3.7 Conclusions

Desired target signal is always contaminated by the white and colored interferences due to the severely diverse interference sources. So it is important to have the knowledge about the signal and interference properties of the operating environment. In this chapter, the signal and interference models have been formulated. It is apparent from this chapter that, in order to improve the detection performance of the target signal, it is very important to have efficient filtering methodology. Jammer is an active interference, and the processor doesn't need the joint domains approach (STAP), since it needs to be canceled out at the spatial domain before the transmission. Clutter on the other hand, is a passive interference, however it can be considered as a moving interference with respect to the platform velocity. Typically, rejection of clutter interference is performed in space-time domain. It is also shown that the constraint to the weight vector, the predefined signal vector $\mathbf{c}$ and the blocking matrix U, play an important role for the approximation of the desired signal and pre-whiting the signal-interferences. Finally, it is also evident that the, Improvement Factor (IF) and the Minimum Variance Distortionless Response (MVDR) can also be used as a key tool for the analysis of performance as well as efficiency of the estimator and the processor.

Chapter 4

Space-Time Adaptive Processing with Adaptive State Estimation (STAP-ASE)

Due to the diverse nature of the sources, nonlinear and nonstationary interferences because of the clutter and jammer effects are highly anticipated. But ground and space surveillance from the airborne radar, it is desirable to isolate, detect, and track, small, slow and fast moving targets within this vicinity. So, filter would be able to track and adapt with the changes due to the severely diverse nature of the environment. Wiener filter is an optimal linear filter in order to cancel out the interferences for the linear estimation of a desired signal. Therefore, the conventional Wiener filter (or LMS algorithms) is inadequate for dealing with nonstationary and nonlinear systems (signal and interferences). One the other hand, sequential state estimator (i.e., Extended Kalman Filter-EKF), has the capability to adapt with the multiple stochastic constraints to achieve the suppression of diverse nature of the interferences due to the clutter and jammer effects. The approach presented here, STAP with

Aadaptive (sequential) State Estimation, incorporates and adapts with the environmental changes due to the nonlinear and nonstationary systems, while maintaining the distortionless output without degrading the signal properties.

In this chapter, problem formulation for the proposed STAP-ASE model, algorithms, and design architecture of the proposed model have been stated. The different aspects of the finite sample size STAP processing and the Wiener filtering approaches have also been included. In the middle of this chapter, the proposed STAP-ASE model, and computational complexity of this model have been included. Finally, the Posterior Cramér-Rao Lower Bound (PCRLB), and formulation of the efficiency of the proposed model have been stated.

4.1 Aspects of Finite Sample Size for STAP Processing

The biggest challenge of the STAP processing is to estimate the asymptotic interference covariance matrix from the sample data set due to the insufficient i.i.d sample data. More importantly, the exact (asymptotic) knowledge (prior) of the interference covariance matrix is unknown. As a result, the model needs to use the most popular Maximum Likelihood Estimator (MLE) function stated in following equation for number of samples L, in order to approximate the interference (asymptotic) covariance matrix:

$$Q_i = \frac{1}{L}\sum_{l=1}^{L} \mathbf{x}_l \mathbf{x}_l^H \tag{4.1}$$

Where Q_i is the equivalent to MLE of Q_{opt}, and $\mathbf{x}$ is the training data sample, which may obtain from spatial, temporal or both domains.

Requirements of the large number sample data is anticipated for the accuracy of this approximation close to the optimal value. However, large number of sample size may cause long data acquisition time as well as computational effort [30]. On the other hand, due to nonstationarity and nonlinearity, only a small number of Independent and Identically Distributed (i.i.d) data set from the surrounding may be available for the estimation of the interference covariance matrix for the range cell under test. Following subsections would include two most widely used algorithms, Simple Matrix Inversion (SMI) and Loaded Sample Matrix Inversion (LSMI), for finite sample size STAP processing.

4.1.1 Sample Matrix Inversion (SMI)

The Sample Matrix Inversion (SMI) algorithm has been widely applied in the field of adaptive processing. Most of the interference suppression due to clutter and jammer effects, are based on the inverse of the sample covariance matrix. According to the relation between asymptotic interference covariance matrix and the optimal weight vector; suppression techniques are based on the adaptation of the training data set and inverse of the interference covariance matrix. Furthermore, it is possible to estimate the Signal to Interference-plus-Noise Ratio (SINR) by using relation stated in following equation.

$$SINR = \frac{\mid \mathbf{w}^H \mathbf{s}(\theta, f_d) \mid^2}{\mathbf{w}^H Q \mathbf{w}} \tag{4.2}$$

The solution for maximizing the value of SINR can be stated under the following constraint:

$$\begin{aligned} &\min_{\mathbf{w}} \quad \mathbf{w}^H Q_v \, \mathbf{w} \\ Subject\ to: \quad &\mathbf{w}^H \mathbf{s}(\theta, f_d) = 1 \end{aligned} \tag{4.3}$$

Where, θ *and* $\mathbf{s}$ are the direction of arrival and the target signal respectively.

The solution for equation (4.3) can be obtained from the following equation:

$$\mathbf{w} = Q^{-H}\mathbf{s}(\theta, f_d) \tag{4.4}$$

This method is known as Sample Matrix Inversion (SMI) [29], and the stability of the SMI algorithm depends on the ability to invert the interference covariance matrix. If the signal and interference characteristics are known, then the interference covariance matrix and optimal solution for the weight vector, hence the SINR can be obtained from equations (4.2), (4.3) and (4.4). However, in practice, the covariance matrix is unknown and needs to be estimated from the relation stated in equation (4.1). The typical diagram for adaptive SMI beamforming is shown in Figure 4.1 [30].

SMI algorithm is based on the direct inversion of the interference covariance matrix, hence offers a faster convergence rate [30]. However, huge matrix inversions, and large sample size lead to computational complexities, which may cause the crucial challenge for implementing SMI algorithms. For example, if the number of array elements N (N=12) and pulses M (M=100) for each Coherent Processing Interval (CPI), then the required number of samples or snapshots for 3dB optimal output using SMI algorithm will be $2 \times MN$ (= 2400). On the other hand, due to the diverse nature of the sources, and inadequate number data samples in the noise domain, it is very challenging to have large number of i.i.d training data samples. As a consequence singularity of the covariance matrix occurs, which may cause the instability of the SMI algorithm.

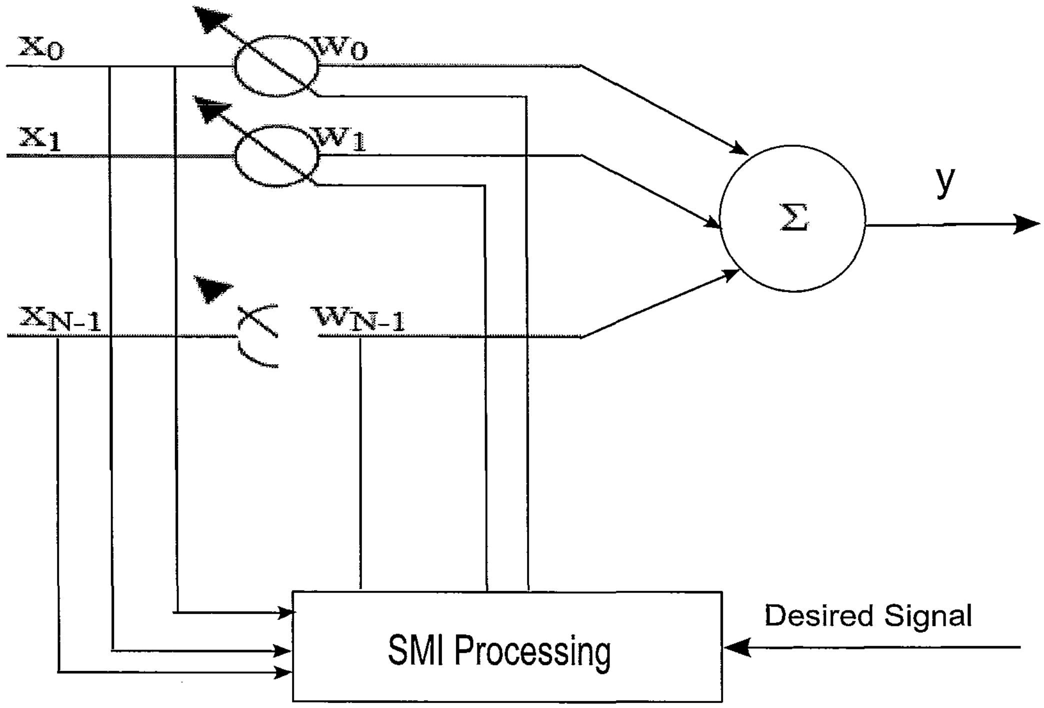


Figure 4.1: SMI- Beamforming

4.1.2 Diagonal Loading- Loaded Sample Matrix Inversion (LSMI)

The problem of the instability for SMI algorithms due to singularity of the covariance matrix for not having sufficient sample data can be circumvented by adding artificial noise, known as diagonal loading. This method is widely known as Loaded Sample Matrix Inversion (LSMI). If Q is the estimate of the interference covariance matrix, then the corresponding diagonally loaded matrix can be stated as [31]:

$$Q_{dl} = Q + I\delta \tag{4.5}$$

Where, I is the identity matrix and δ is a real constant representing the desired loading factor.

According to the EVP (eigenvalue processor) model [1], the number of the interference eigenvalues is $L = M + N - 1$, and the number of channels depend on this value. But, this might be necessary to increase the number of channels beyond L, so that system can cope with the decorrelation effect due to the signal bandwidth. Therefore, the nonlinear and the nonstationary clutter characteristics may lead to the insufficient number of the eigenvalues. Typically, interference covariance matrix has several number of eigenvalues. But with the insufficient number of required training data set, may result inadequate estimation of the interference, hence the large interference eigenvalue spread occurs. It raises the question about the singularity (non-invertibility) of the interference covariance matrix. Therefore, introducing the additional or artificial noise, diagonal loading, diagonally can be a way of overcoming this challenge. With the additional noise data, the estimation improves and as a result eigenvalues converge to the optimal value, hence the adaptive beam shape. Diagonal loading, known as Loaded Sample Matrix Inversion (LSMI), methodology improves the adaptive side lobe levels, hence enhances the shape of the mainbeam response in one hand, and overcomes the challenges due to the limited number of training data sample on the other. In the case of LSMI, it has been shown [1][2][29][31] that, for the 3dB performance, the system only needs $L = 2 \times (M + N)$ training data set, which is significantly lower than the SMI method.

4.2 Wiener Filter Estimator

Wiener filter is a linear optimal discrete time filter, based on the minimization of the cost function, commonly known as the Mean Squared Error (MSE) function. It is

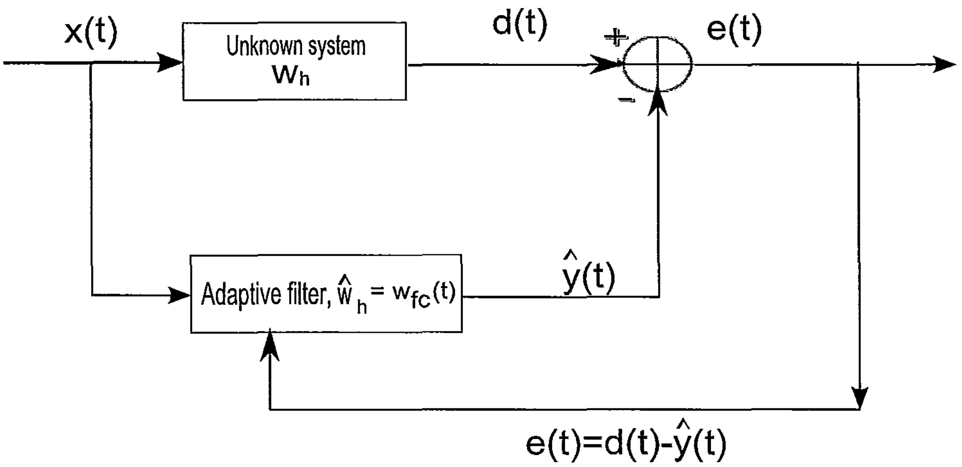


Figure 4.2: Block Diagram of a Adaptive Wiener Filter

based on the complex valued stochastic process, where filter specified in terms of its impulse response [32]. A schematic presentation of the adaptive Wiener filter is shown in Figure 4.2. Adaptive Wiener filter has the ability to adapt with the stationary, linear systems. The scalar output of the filter at time t can be stated as:

$$y(t) = \mathbf{w}_{fc}^{H}(t)\mathbf{x}(t) \tag{4.6}$$

Where, vector $\mathbf{x}$ is the observed signal, which may contain interferences, and the vector $\mathbf{w}_{fc} = [w_1, w_2, \ldots, w_N]^T$ is considered as the filter coefficients, and in this section N is considered as the number of filter coefficients.

If the system considers that the observed signal $\mathbf{x}$, has the signal part $\mathbf{s}$ and the noise part $\mathbf{n}$, then above equation can be restated as:

$$y(t) = \mathbf{w}_{fc}^H(t)\mathbf{s}(t) + \mathbf{w}_{fc}^H\mathbf{n}(t) \tag{4.7}$$

Now, consider that the filter will pass the stationary process $\mathbf{x}(t)$ in order to estimate $\hat{y}(t)$ of the desired signal $d(t)$. The model also assumed that $\mathbf{x}(t)$, $\mathbf{s}(t)$, and $\mathbf{n}(t)$ have zero mean value, and the filter coefficients $\mathbf{w}_{fc}(t)$ do not change with time. Then the estimate of the desired signal can be stated as:

$$\hat{y}(t) = \mathbf{w}_{fc}^H(t)\mathbf{x}(t) \tag{4.8}$$

Since, filter coefficients do not change with time, then it can be written:

$$\mathbf{w}_{fc}^H(t+1) = \mathbf{w}_{fc}^H(t) \tag{4.9}$$

The Mean Squared Error (MSE) is given by [complete derivation in Appendix B]:

$$\begin{aligned} MSE = \mathbb{E}[e^2(t)] &= \mathbb{E}[(d(t) - \mathbf{w}_{fc}^H(t)\mathbf{x}(t))^2] \\ &= \sigma_d^2 - 2\mathbf{w}_{fc}^H\mathbf{p}_{dx} + \mathbf{w}_{fc}^H Q_x \mathbf{w}_{fc} \end{aligned} \tag{4.10}$$

where, $\mathbf{p}_{dx} = [p_{dx}(1), p_{dx}(2), \ldots, p_{dx}(N)]^T$ = cross correlation vector between $d(t)$ and $\mathbf{x}(t)$; and σ_d^2 is the covariance of the desired signal.

Therefore the cost function can be stated as:

$$\min_{\mathbf{w}_{fc}} MSE = \sigma_d^2 - 2\mathbf{w}_{fc}^H\mathbf{p}_{dx} + \mathbf{w}_{fc}^H Q_x \mathbf{w}_{fc} \tag{4.11}$$

where, Q is $N \times N$ matrix; $\mathbf{p}_{dx}$ and $\mathbf{w}_{fc}$ are $N \times 1$ vectors respectively.

If the matrix Q is invertible, then the Wiener solution for the optimal filter coefficient $\mathbf{w}_{opt}$ can be stated as:

$$\mathbf{w}_{opt} = Q_x^{-1}\mathbf{p}_{dx} \tag{4.12}$$

where, $\mathbf{w}_{opt}$ is $N \times 1$ vector.

The main advantage of Wiener filter is its computational simplicity and ability to suppress noise in linear and stationary case. Furthermore, the optimal solution of Wiener filter involves only second order statistic, which in fact leads to a useful theory of linear filtering for many applications. The main goal of the Wiener is to filter out the noise, based on the statistical properties. Furthermore, Wiener filter is an optimal filter under the assumption that the signal and additive noise are stationary linear stochastic processes, and their spectral characteristic or auto and cross correlation are known. But, it is very likely to have time varying nonlinear system with unknown signal and noise properties. In that case, Wiener filter solution is inadequate to deal with noise suppression. Therefore, the system would need a filter that has the ability to adapt itself to nonstationary and nonlinear environment.

4.3 Model Formulations and Samples Training Using Space-Time Adaptive Processing with Adaptive State Estimation (STAP-ASE)

The challenges of the Wiener filter can be overcome by sequential state estimation (i.e., Kalman Filter/Extended Kalman Filter). This is in fact an important generalization of the Wiener filter, and has the ability to adapt with the nonstationary and nonlinear systems. In other words, Wiener filter is the special form of the Kalman

Filter when the process is stationary. Therefore, the sequential state estimation is being used for training the data samples for the proposed STAP-ASE model. In this training process, constraint is added to the weight vectors in order to filter out the signal interferences in the MSE sense. In this section, the main objective is to design a filter that would give us the estimate $\hat{y}(t)$ of the desired signal $d(t)$ by using the input data sample $\mathbf{x}(t)$. Now, consider the block $U(t)$, which satisfies equation (3.15). Therefore, the output of the block $U(t)$ has contained the sample data signal from the all range cells but the cell under test (semi-supervised). The operational block diagram of the STAP with adaptive sequential state estimation (STAP-ASE) is shown in Figure 4.3. The estimated processor output $\hat{y}(t)$(scalar) can be stated as:

$$\hat{y}(t) = \mathbf{w}^H \mathbf{x}(t) \tag{4.13}$$

According to Figure 4.3 it can be written:

$$\begin{aligned}
e(t) &= d(t) - \hat{y}(t) = d(t) - \mathbf{w}^H(t)\mathbf{x}(t) \\
MSE = \mathbb{E}[e^2(t)] &= \mathbb{E}[(d(t) - \mathbf{w}^H(t)\mathbf{x}(t))^2] \\
&= \mathbb{E}[((d(t) - \mathbf{w}^H(t)\mathbf{x}(t))(d(t) - \mathbf{w}^H(t)\mathbf{x}(t))^H] \\
&= \mathbb{E}[d^2(t) - \mathbf{w}^H(t)\mathbf{x}(t)d(t) - d(t)\mathbf{x}^H(t)\mathbf{w}(t) + \mathbf{w}^H(t)\mathbf{x}(t)\mathbf{x}^H(t)\mathbf{w}(t)] \\
&= \mathbb{E}[d^2(t)] - 2\mathbf{w}^H(t)\mathbb{E}[\mathbf{x}(t)d(t)] + \mathbf{w}^H(t)\mathbb{E}[\mathbf{x}(t)\mathbf{x}^H(t)]\mathbf{w}(t) \\
&= \sigma_d^2 - 2\mathbf{w}^H(t)\mathbf{P}_{xd} + \mathbf{w}^H(t)Q(t)\mathbf{w}(t)
\end{aligned} \tag{4.14}$$

where,

$$\begin{aligned}
\sigma_d^2 &= \textit{Variance of the desired signal.} \\
\mathbf{P}_{xd} &= \textit{Cross correlation vector between } x(t) \textit{ and } d(t)
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}[\mathbf{x}(t)d(t)] \\
&= \mathbb{E}[\mathbf{x}(t)\mathbf{x}^H(t)]\mathbf{w}_c(t) \\
\mathbf{P}_{xd} &= Q(t)[Q^{-1}\mathbf{c}(t)(\mathbf{c}^H(t)Q^{-1}(t)\mathbf{c}(t))^{-1}g] \qquad (4.15) \\
&= g\,\frac{\mathbf{c}}{\mathbf{c}^H Q^{-1}\mathbf{c}} \qquad (4.16)
\end{aligned}
$$

Therefore,

$$\min_{\mathbf{w}} \; MSE = \sigma_d^2 + \mathbf{w}^H(t)Q(t)\mathbf{w}(t) - 2\mathbf{w}^H(t)\mathbf{P}_{xd} \qquad (4.17)$$

and is also the expected outcome or objective function.

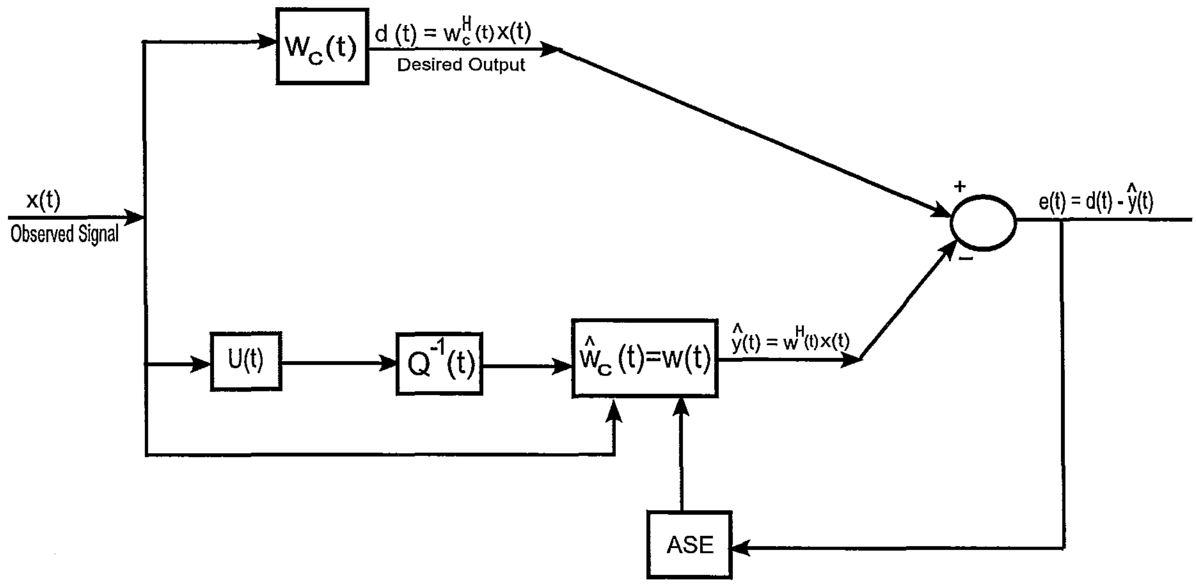


Figure 4.3: STAP with Adaptive State Estimation (STAP-ASE)

4.3.1 Adaptive Extended Kalman Filter Model for STAP Processing

The proposed model has included spatial-Doppler processing in order to cancel out the nonstationary multipath interference (hot clutter) in the MSE sense, so that it would isolate or decode the signal of interest. For the diverse nature of the hot interferences, the properties of the spatial-Doppler samples would be severely fluctuated over the CPI interval, i.e., nonstationary. Therefore, due to the nonstationary interference environment, covariance matrix, hence the weight vector associated with the array element would not be fixed over time. As a result, consideration of conventional adaptive processing won't allow us to achieve the effective suppression of hot clutter. So, a complex system needs to be modeled so that system would be able to track the changes in response to the change of the environment over time. Furthermore, STAP is a multi-dimensional signal processing problem, which would estimate the state vector by (adaptive) training the sample data. Since, adaptation is performed in spatial and temporal domains, training is done using the range bins. The signal received at the output of the processor can be stated as:

$$y(t) = \mathbf{w}^H(t)\mathbf{x}(t) = \mathbf{w}^H(t)\mathbf{s}(t) + \mathbf{w}^H(t)n(t) \tag{4.18}$$

The objective is to find out the minimum value of the Mean Squared Error (MSE), i.e., Minimum Mean Squared Error (MMSE) . Hence, the system can isolate or decode the desired signal.

Therefore, it can be written:

$$\min_{\mathbf{w}} \; MSE = \sigma_d^2 + \mathbf{w}^H(t)Q(t)\mathbf{w}(t) - 2\mathbf{w}^H(t)\mathbf{P}_{xd} \tag{4.19}$$

More importantly, the conditional mean is an optimal MMSE estimator and is computationally efficient since it is recursive so that there is no need to store the entire past data set. In many instances, the nonlinearity is benign so that the approximate Extended Kalman Filter (EKF), based on linearizing gives very good performance. It shall be seen that the EKF is sufficiently accurate in this case. Of course, in the linear case the Kalman Filter estimate of the conditional mean is optimal.

Now, consider an unknown dynamic system with state vector $\mathbf{w}_c$ and the system is driven by random noise. If the system can be modeled as filter, then the state equation can be written as [7]:

$$\mathbf{w}_c^K(t+1) = F(t+1|t)\mathbf{w}_c^K(t) + v(t) \tag{4.20}$$

Where, $F(t+1|t)$ is the transition matrix or the model parameter. $v(t)$ is processed noise, may be assumed white Gaussian and zero mean. The covariance matrix of the process noise may be assumed $R = \Psi_p^2 I$ [7], where I is identity matrix, t is time index, and K represents the snapshot of the STAP processing. For convenience, the superscript K can be ignored for rest of the cases.

If the signal environment is considered as stationary, then the state vector $\mathbf{w}_c$ would be fixed and $F(t+1|t)$ can be considered identity matrix. For nonstationary clutter and jammer environment, a more complex method for $F(t+1|t)$ needs to be developed so that model can track the change in response to the change of the environment.

Now, the measurement equation can be represented by:

$$\begin{aligned} d(t) &= \mathbf{x}(t)^H \mathbf{w}_c(t) + \rho(t) \\ &= H[t, \mathbf{w}_c(t)] + \rho(t) \end{aligned} \tag{4.21}$$

where, $\rho(t)$ is white Gaussian measurement noise with zero mean, and covariance is given by:

$$\mathbb{E}[\rho(t)\rho(i)] = \sigma_m^2(t)\delta_{ti} \tag{4.22}$$

Since the measurement model is linear (and the state model is also considered as linear), the Kalman Filter (KF) provides the MMSE estimate. However, in general, the measurement model can be nonlinear and nonstationary. Therefore, the measurement equation (4.21) due to the nonlinearity can be stated as:

$$d_{non}(t) = h[t, \mathbf{w}_c(t)] + \rho(t) \tag{4.23}$$

where, $h[.]$ is Jacobian evaluation of the measurement matrix $H[.]$ The estimate $\hat{y}(t)$ of the desired signal $d(t)$ can be stated as:

$$\hat{y}(t) = h[t, \hat{\mathbf{w}}_c(t)] \tag{4.24}$$

The MSE between the desired signal and the processor output can be defined by using equation (4.17) as follow:

$$\begin{aligned} \mathbf{e}(t) &= d(t) - \hat{y}(t) && (4.25) \\ \min_{\hat{\mathbf{w}}_c} MSE &= \mathbb{E}[|e(t)|^2] \\ &= \sigma_d^2 + \hat{\mathbf{w}}_c^H(t)Q(t)\hat{\mathbf{w}}_c(t) - 2\hat{\mathbf{w}}_c^H(t)\mathbf{P}_{xd} && (4.26) \end{aligned}$$

Initial estimate of the weight vector starts with the relation stated in equation (1.2), by using the estimated covariance (asymptotic) matrix calculated from the sample data using equation (1.1). Thus, using STAP-ASE, $\hat{\mathbf{w}}_c$ would be converged to optimal value in the MSE sense. Furthermore, the processor may update $\mathbf{x}$ for the associated weight vector using equation (1.2), hence update estimate of covariance matrix using relation stated in equation (1.1).

Now, the innovation or the measurement residue may be stated as [13]:

$$\nu(t) = d(t) - \hat{y}(t|t-1) \tag{4.27}$$

Hence, the state update equation can be stated as:

$$\hat{\mathbf{w}}_c(t+1|t+1) = \hat{\mathbf{w}}_c(t+1|t) + \mathbf{k}(t+1)\nu(t+1) \tag{4.28}$$

Where $\mathbf{k}$(t) is the filter gain, and then $\mathbf{k}$(t) can be written in terms of first order (Jacobian) EKF measurement matrix [13]:

$$\mathbf{k}(t+1) = P(t+1|t)[\frac{\partial h(t, \hat{\mathbf{w}}_c(t))}{\partial \mathbf{w}_c(t)}]\mathbf{M}^{-1}(t+1) \tag{4.29}$$

and the estimated (predicted) conditional covariance matrix [13]:

$$P(t+1|t) = F(t)P(t|t)F^H(t) + R(t) \tag{4.30}$$

Innovation covariance:

$$\mathbf{M}(t+1) = \sigma_m^2(t+1) + [\frac{\partial h(t, \hat{\mathbf{w}}_c(t))}{\partial \mathbf{w}_c(t)}]^H.P(t+1|t)[\frac{\partial h(t, \hat{\mathbf{w}}_c(t))}{\partial \mathbf{w}_c(t)}] \tag{4.31}$$

The covariance update [13]:

$$P(t+1|t+1) = P(t+1|t) - \mathbf{k}(t+1)M(t+1)\mathbf{k}(t+1)^H \tag{4.32}$$

Using the matrix Lemma [13], it can be written:

$$P(t+1|t+1)^{-1} = [(P(t+1|t))]^{-1} + \frac{\partial h(t, \hat{\mathbf{w}}_c(t))}{\partial \mathbf{w}_c(t)}[\sigma_m^2]^{-1}[\frac{\partial h(t, \hat{\mathbf{w}}_c(t))}{\partial \mathbf{w}_c(t)}]^H \tag{4.33}$$

If the first term of right hand side of equation (4.33) can be neglected for the large initial value condition, then above equation may be restated as:

$$P(t+1|t+1)^{-1} = \frac{\partial h(t, \hat{\mathbf{w}}_c(t))}{\partial \mathbf{w}_c(t)}[\sigma_m^2]^{-1}[\frac{\partial h(t, \hat{\mathbf{w}}_c(t))}{\partial \mathbf{w}_c(t)}]^H \tag{4.34}$$

This is a recursive process starting from equation (4.23), and the process would stop once the stop criteria meets.

Differences stated in equation (4.27) represent the differences between the desired output and the actual output, estimated using the weight vector, and covariance matrix. If the desired output $d(t)$ can be interpreted as an approximation of the actual desired signal using the constraint; then actual signal $a(t)$, and the approximated desired signal $d(t)$ with approximation error $\beta(t)$ may be related as follows:

$$a(t) = d(t) + \beta(t) \tag{4.35}$$

Now, if the system is considered that, $a(t), and\ \beta(t)$ are white Gaussian and zero mean, uncorrelated, then

$$\sigma^2(t) = MMSE + \mathbb{E}[\beta^2] \tag{4.36}$$

Therefore, if the measurement $d(t)$ is generated using the realistic approximation for the desired signal, then $\sigma^2(t)$ would be approximated based on the MMSE that can be achieved by an optimal array, approximation of constraint weight vector $\mathbf{w}_c(t)$ and the filtering model as stated above. The robustness of the approximation of $d(t)$ has been discussed in many papers. Using the model stated in [7], the robustness of this approximation may be analyzed based on the constraint weight vector.

Furthermore, for time varying systems, parameters $F(t+1|t)$ and covariance R of equation 4.20 would be more complex form. Therefore, these parameters of state equation must be chosen in a way that system model can track the changes due to the diversity of the environment [7]. For nonstationary environment, the state transition matrix $F(t+1|t) \neq I$, and this may make the state equation unstable, stability of the filter may be assured by the observability condition [7][68][69]. If the state covariance matrix R represents the total uncertainty due to the adapting stationary environment assumption represented by using identity state transition matrix in equations (4.20) and (4.30). The effect associated with this deviation due to nonstationary environment $[v(t) \neq 0 \textit{ and } F(t+1|t) \neq I]$ prevent the Kalman Filter (KF) gain from decaying to the values that are too small. Therefore, estimate of optimal weight vector $\hat{\mathbf{w}}_c$ able to follow the variations in the optimal weight vector due to the nonstationary heterogeneous operating environment. A typical choice of Ψ_p^2 is 10^{-4} [7], but for stationary environment the value of $\Psi_p^2 = 0$, since for time invariant systems weight vector does not change with time [7].

4.3.2 Computational Complexity

The computational complexity is an important aspect for the proposed STAP-ASE model. Starting from equation (4.23), the STAP-ASE's computational complexity for Jacobian evaluation is $\mathbf{O}(N^2)$. Complexity for equation (4.24), for estimating desired signal is $\mathbf{O}(N^2)$. Computational bottleneck for computing the inverse of updated covariance at each cycle is $\mathbf{O}(N^3)$. Complexity for updating of weight vector in equation (4.28) is $\mathbf{O}(N)$, gain in equation (4.29) is $\mathbf{O}(N^2)$, and predicted covariance in equation (4.30) is $\mathbf{O}(N^2)$. Computational cost for innovation covariance stated in equation (4.31) is $\mathbf{O}(N^2)$, covariance update in (4.32) is $\mathbf{O}(N^2)$ and for equation (4.34) is $\mathbf{O}(N^2)$.

In STAP with ASE, the immediate previous state information has been used to update current interference covariance matrix, and therefore, its computational cost is $\mathbf{O}(N^3)$ per cycle. Estimation of covariance matrix in SMI, LSMI and [3][7] are off the filtering loop, hence updated weight vector does not have any direct impact on computation of the covariance matrix. Advantage of using proposed STAP-ASE method is that, it would give the high convergence rate and high rejection of the interferences due to nonstationary heterogeneous clutter and jammer effects at the worst case scenario. In conclusion, it can be stated that, STAP with ASE is simpler and a more efficient estimator than its counterparts.

4.4 Posterior Cramér-Rao Lower Bound (PCRLB)

The estimation of the unknown state with uncertainty for the slow and fast moving target is crucial for the efficient tracking algorithms. However, the estimation of unknown parameters are always contaminated by the interferences. So, minimum achievable interference reduction level needs to be studied in order to evaluate the performance of the estimator. The Cramér-Rao Lower Bound (CRLB), inverse of the Fisher Information Matrix (FIM), provides the achievable lower bound on interference covariance matrix for the evaluation of performance of any unbiased estimator. It allows the system to establish a bound to rule out the unreachable estimators. In general, it is a benchmark against which the optimum performance of processor can be compared, regardless of the filtering algorithm employed. The CRLB has been analyzed for the dynamic state model, where state is unknown and randomly fluctuated. Therefore, for the greatest interest of deriving lower bound on covariance matrix, in this section, the Posterior Cramér-Rao Lower Bound (PCRLB) is being used.

4.4.1 Introduction

Now consider, $\hat{\mathbf{w}}(\mathbf{z})$ to be the unbiased estimate of dynamic state vector $\mathbf{w} = [w_1, w_2, \ldots, w_n]^T$ estimated from observation $\mathbf{z} = [z_1, z_2, \ldots, z_n]^T$, then the posterior pdf of the weight estimation may be stated as:

$$p(\mathbf{w} \mid \mathbf{z}) = \frac{p(\mathbf{z} \mid \mathbf{w})p(\mathbf{w})}{p(\mathbf{z})} \tag{4.37}$$

the error covariance matrix:

$$Cov(\hat{\mathbf{w}}) = \mathbb{E}[(\hat{\mathbf{w}} - \mathbf{w})(\hat{\mathbf{w}} - \mathbf{w})^T] \tag{4.38}$$

and the Fisher Information Matrix (FIM) I may be stated as:

$$I(\mathbf{w}) = -\mathbb{E}[(\frac{\partial^2 \, ln \; p(\mathbf{w} \mid \mathbf{z})}{\partial \mathbf{w} \; \partial \mathbf{w}^T}] \tag{4.39}$$

Hence, the PCRLB can be defined as:

$$Cov(\hat{\mathbf{w}}) \geq I(\mathbf{w})^{-1} \tag{4.40}$$

Covariance matrix $Cov(\hat{\mathbf{w}})$:

$$Cov(\hat{\mathbf{w}}) = \begin{bmatrix} var(\hat{\mathbf{w}}_1) & cov(\hat{\mathbf{w}}_{12}) & \ldots & cov(\hat{\mathbf{w}}_{1n}) \\ cov(\hat{\mathbf{w}}_{21}) & var(\hat{\mathbf{w}}_2) & \ldots & cov(\hat{\mathbf{w}}_{2n}) \\ \ldots & \ldots & \ldots & \\ \ldots & \ldots & \ldots & \\ cov(\hat{\mathbf{w}}_n 1) & cov(\hat{\mathbf{w}}_{n2}) & \ldots & var(\hat{\mathbf{w}}_n) \end{bmatrix} \tag{4.41}$$

4.4.2 State and Measurement Information

The dynamic state equation for the state vectors $\mathbf{w} = [w_1, w_2, \ldots, w_n]^T$ at the time cycle t may be stated as:

$$\mathbf{w}(t+1) = F(t)\mathbf{w}(t) + \nu(t) \tag{4.42}$$

where, $F(t)$ is linear fixed parameter as stated before, and $\nu(t)$ is independent white Gaussian noise with covariance Υ.

Hence, according to [14][15][16][17], it may be written:

$$J_w(t) = [\Upsilon(t-1) + F(t)J(t-1)F(t-1)^T]^{-1} \tag{4.43}$$

where, $J(t)$ is the state information matrix at time cycle t.

Measurement (nonlinear) equation may be stated as:

$$\mathbf{z}(t) = h[\mathbf{w}(t)] + \rho(t) \tag{4.44}$$

According to [15][16][17], the measurement information matrix may be stated as:

$$J_z(t) = \mathbb{E}[(\nabla_w(t)\ ln\ p(\mathbf{z}(t) \mid \mathbf{w}(t)))(\nabla_w(t)\ ln\ p(\mathbf{z}(t) \mid \mathbf{w}(t)))^T] \tag{4.45}$$

where, ∇ is the first order partial derivative and T is the matrix transpose.

If the measurement covariance is Σ for the nonlinear measurement model, then it can be written [15][16][17]:

$$J_z(t) = \mathbb{E}[h(\mathbf{w}, t)\Sigma^{-1}h^T(\mathbf{w}, t)] \tag{4.46}$$

Therefore, the posterior Fisher Information Matrix (FIM) $J(t)$, may be stated as:

$$J(t) = J_w(t) + J_z(t) \tag{4.47}$$

Now, if the prior information $\mathbf{w}$ is not known, then according to the equation (4.37), (4.43), (4.45) and (4.47), it can be stated that the maximization of posterior is equivalent to maximization of the likelihood function $p(\mathbf{z} \mid \mathbf{w})$ [4] as well as maximization of measurement information matrix. Now the likelihood function using the conditional probability with predicted covariance σ^2 and conditional mean $\bar{\mathbf{w}}(k \mid k-1)$ (Gaussian assumption) may be stated [4] as:

$$\begin{aligned} p(\mathbf{z} \mid \mathbf{w}) &= \prod_{k=1}^{n} \frac{1}{\sqrt{2\pi\sigma_k^2}}\, exp[-\frac{(\mathbf{z}^k - \bar{\mathbf{z}}(k \mid k-1))^2}{2\sigma_k^2}] \\ ln\ p(\mathbf{z} \mid \mathbf{w}) &= \sum_{k=1}^{n}[-ln\ (2\pi\sigma_k^2) - \frac{(\mathbf{z}^k - \bar{\mathbf{z}}(k \mid k-1))^2}{2\sigma_k^2}] \end{aligned} \tag{4.48}$$

Therefore, maximum likelihood cost function:

$$J(\mathbf{w}) = -ln\ p(\mathbf{z} \mid \mathbf{w}) = \sum_{k=1}^{n}[ln\ (2\pi\sigma_k^2) + \frac{(\mathbf{z}^k - \bar{\mathbf{z}}(k \mid k-1))^2}{2\sigma_k^2}] \tag{4.49}$$

The objective is to find out the Maximum Likelihood Estimator (MLE), so that the cost function can be minimized. Minimum (optimal) value for the cost function can be achieved by differentiating equation (4.49) with respect to $\mathbf{w}$ and equating it with zero.

So, it may yield:

$$\begin{aligned}
\sum_{k=1}^{n}[0+\frac{(z^k-\bar{\mathbf{z}}_{ml}(k\mid k-1)}{\sigma_k^2}] &= 0\\
\sum_{k=1}^{n}[\frac{(\mathbf{z}^k-\bar{\mathbf{z}}_{ml}(k\mid k-1)}{\sigma_k^2}] &= 0\\
\sum_{k=1}^{n}[(\mathbf{z}^k-\bar{\mathbf{z}}_{ml}(k\mid k-1)] &= 0\\
n\ \bar{\mathbf{z}}_{ml} &= \sum_{k=1}^{n}[\mathbf{z}^k]\\
\bar{\mathbf{z}}_{ml} &= \frac{1}{n}\sum_{k=1}^{n}[\mathbf{z}^k]
\end{aligned} \tag{4.50}$$

since, the system assumed that, σ^2 is independent of the weight vector $\mathbf{w}$.

Now, it may be written:

$$Var(\bar{\mathbf{z}}_{ml}) \geq\ CRLB = I(\mathbf{w})^{-1} \tag{4.51}$$

Now, an estimator is said to be optimal or will reach to the lower bound if

$$\lim_{n\to\infty} \bar{\mathbf{z}}_{ml}^{n} = \mathbf{z} \tag{4.52}$$

In that case, the estimator is said to be consistence since it provides the actual value of the parameter.

Finally, the efficiency of the estimator can be stated as:

$$Efficiency,\ \eta\ = \frac{I(\mathbf{w})^{-1}}{Var(\bar{\mathbf{z}}_{ml})} \tag{4.53}$$

Expected value for the efficiency is 1.

4.5 Conclusions

This is the core chapter for this thesis. First part of this Chapter, the aspects of the finite sample size have been stated, and the required minimal sample size for 3dB performance of SMI is $2 \times MN$, has also been formulated. It has also been shown that, using the Loaded Sample Matrix Inversion (LSMI) method, the system can achieve 3dB performance by using only $2 \times (M + N)$ samples compared to SMI $2 \times MN$ samples. Middle of this chapter, the Wiener filter algorithms has been stated and shown that Wiener filter is a special case of the sequential state estimation under the assumption that the system environment is stationary and linear. Final part of this chapter, the designed architecture of the proposed STAP-ASE model has been stated. In this model, Extended Kalman Filter (EKF) as a sequential state estimator has been used, and found that Kalman Filter is an efficient filter, and it has the ability to adapt itself to nonstationary heterogeneous environments. Finally, the Posterior Cramer-Rao Lower Bound (PCRLB) has been stated, and also formulated the efficiency for the proposed model. However for STAP-ASE model, Improvement Factor (IF) is the key tool for measuring the efficiency and the performance of the estimator.

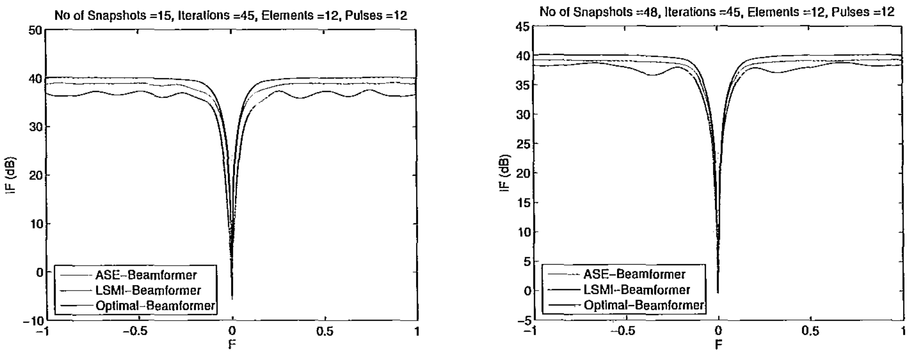
Chapter 5

Simulations and Discussions

5.1 Simulations and Discussions

In order to demonstrate the performance and efficiency of the proposed STAP-ASE estimator, the comparisons with Loaded Sample Matrix Inversion (LSMI), Sample Matrix Inversion (SMI) and optimal based beamformers have been presented.

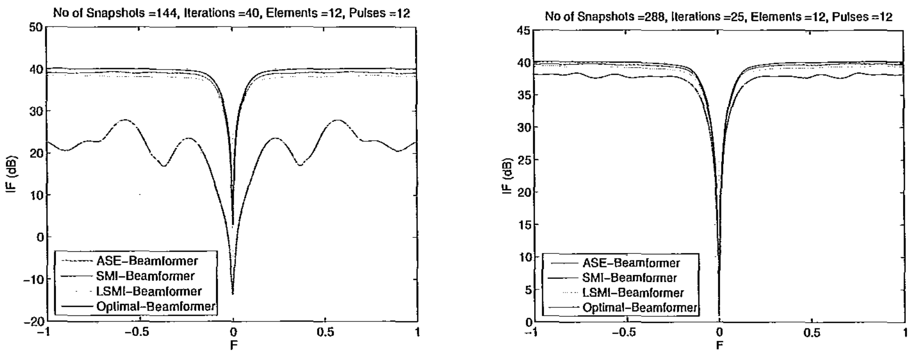
In Figure 5.1(a), Improvement Factor (IF), characterizing the efficiency and performance, among optimal, STAP-ASE and LSMI beamformers have been compared. In this simulation, 15 snapshots has been considered; and 12 no. of pulses (M) and array elements (N) respectively have also been considered. From the outcome of this simulation, it is evident that the, clutter rejection performance for the STAP-ASE for 15 snapshots is superior than LSMI-beamformer and almost close to 3dB of optimal beamformer. Furthermore, STAP-ASE processor suppresses the clutter down to the noise level compared to the optimal-processor. In Figure 5.1(b), the model has considered 48 snapshots for LSMI-3dB performance. According to the Figure, it is apparent that the clutter rejection performance of STAP-ASE beamformer is still better than LSMI beamformer and is also very close to optimal beamformer. However, LSMI beamformer is a bit better in the sense that curve in LSMI beamformer suppresses the clutter down to the noise level compared to the STAP-ASE.



(a) ASE and LSMI-IF (b) ASE and LSMI(3dB)-IF

Figure 5.1: IF-ASE and LSMI-IF –Comparison of ASE-Beamformer with Optimal and LSMI Beamformers for (a) $K = 15$ and (b) $K = 2 \times (M + N)$

In Figure 5.2(a), a comparison of Improvement Factor (IF), characterizing the efficiency and performance, among Optimal, STAP-ASE, LSMI and SMI beamformers has been presented. In this simulation, 144 snapshots; 12 no. of pulses (M) and array elements (N) respectively have been considered. It is apparent from the Figure that, clutter rejection performance is inferior for SMI beamformer. The performance of STAP-ASE beamformer is better than LSMI and is also very close to the optimal beamformer. In simulation 5.2(b), the performance of the beamformers using 288 snapshots (SMI-3dB) has been investigated. According to the outcomes of the simulation, it is apparent that, there is a noticeable improvement in the performance of SMI-beamformer. The curves between the STAP-ASE and the optimal beamformers have shown that the clutter notch as well as rejection performance caused by both processors almost the same. It also stated that the clutter rejection performance for STAP-beamformer is a bit better than LSMI-beamformer and but much superior than SMI-beamformer, however, both LSMI and SMI have shown 3dB performance. Only noticeable performance for LSMI and SMI beamformers is that, both processors suppress the clutter only down to the noise level compared to ASE-beamformer.



(a) ASE, LSMI and SMI-IF (b) ASE, LSMI and SMI(3dB)-IF

Figure 5.2: ASE, LSMI and SMI-IF –Comparison of ASE-Beamformer with LSMI, SMI and Optimal Beamformers for (a) $K = NM$ and (b) $K = 2 \times (NM)$

The performance comparisons based on formation of beampatterns (power) for ASE, SMI, LSMI, and optimal beamformers are illustrated in Figure 5.3. According to the outcomes of this simulation, it can be stated that, all of these beamformers place nulls at the direction of interferences, however in the case of SMI-beamformer, some distortions have been found. It is also apparent that the performance of the side-lobe reduction, and nulling the interferences of ASE-beamformer almost the same as the optimal beamformer.

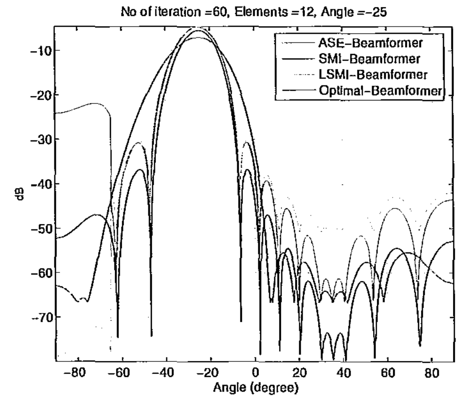


Figure 5.3: Directional Patterns –Comparison of STAP-ASE-Beamformer with SMI, LSMI and Optimal Beamformers

Figure-5.4 shows the MSE of beamformer output power versus the iteration numbers. It is apparent from the simulation that the Mean Squared Error (MSE) of STAP-ASE beamformer output is very outstanding and is always below 1 after 7 iterations.

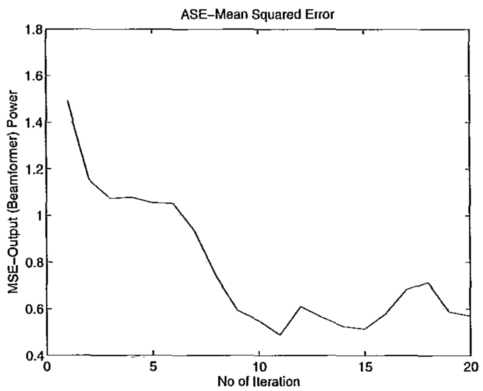


Figure 5.4: MSE-ASE –STAP Mean Squared Error of Output Power

Next simulation, stated in Figure 5.5 has been compared the Signal to Interference-plus-Noise Ratio (SINR) of the proposed STAP-ASE beamformer with SMI, LSMI and optimal beamformers versus the no. of snapshots. From the figure, it is evident that the proposed STAP-ASE has the best performance among the algorithms tested, however the performance of LSMI algorithm is almost close to the STAP-ASE model after $80th$ snapshot. More importantly, the performance of the proposed model is closed to the optimal beamformer after $40th$ snapshot.

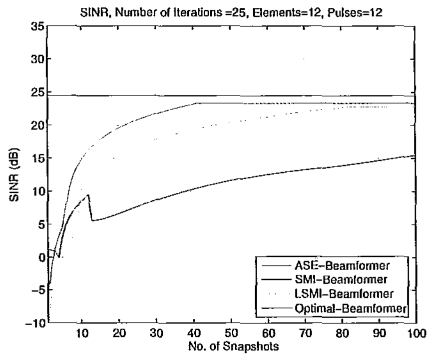


Figure 5.5: SINR –Comparison of STAP-ASE-Beamformer with SMI, LSMI and Optimal Beamformers.

Figure 5.6 has been examined and compared the MSE of the beamformer output powers among the proposed STAP-ASE, SMI and LSMI models versus the no. of snapshots. It is apparent from the simulation that the Mean Squared Error (MSE) of STAP-ASE beamformer output is outstanding compared to other two algorithms. More importantly, the MSE of the STAP-ASE model always below 1 after $19th$ snapshot.

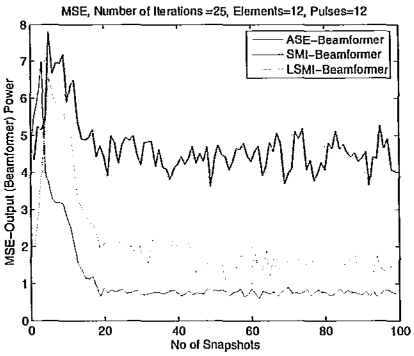


Figure 5.6: MSE– STAP Mean Squared Error of Output Power.

5.2 Conclusions

In this research, a novel STAP with Adaptive State Estimation (STAP-ASE) algorithm has been proposed. The proposed estimator is based on the integration of the sequential state estimation with the STAP processing, so that the model can detect the target signal while rejecting the interferences at worst case scenario. The performance comparison of the proposed model with the Sample Matrix Inversion (SMI), and Loaded Sample Matrix Inversion (LSMI) beamformer models has been examined. The outcomes of this algorithm is also compared with the Optimal beamformer. It is apparent from the simulation results that the STAP-ASE estimator outperforms its counterparts in clutter rejection, and target characterization. Furthermore, it has also shown better than 3dB performance for the convergence rate to the optimal Signal Interference-plus-Noise Ratio (SINR).

Chapter 6

Conclusions and Future Work

6.1 Conclusions

This thesis has proposed a model that integrates Space-Time Adaptive Processing with Adaptive (sequential) State Estimation, so that radar system can overcome the challenges due to the effects of cold (homogeneous) clutter, severe dynamic (heterogeneous) hot clutter and jamming interferences while estimating the states of targets under track. In this model, STAP-ASE is being considered as a multi-dimensional adaptive signal processing technique, which estimates adaptive weight vectors in spatial and Doppler domains for which a target detection hypothesis is to be formed. STAP operates on the set of returns, composed of pulses, array elements, and range bins, over a period of time. Hence, STAP-ASE is a 2D processing on 3D datacube, collected from the available signal received at the radar receiver. Since, adaptation is performed in spatial and temporal domains, training is done using the range bins. The main theory of this model is to adapt with the sample data set in order to estimate the interference (asymptotic) covariance matrix and adjust or update the weight vector in a way that, the noise power can be minimized, and the Signal to

Interference-plus-Noise Ratio (SINR) can be maximized in the MSE sense. Furthermore, the interferences that are considered in this proposed STAP-ASE model is due to the nonstationary and heterogeneous clutter and jammer effects.

In this model, two types of data sets have been considered: one is the training data, which is responsible to estimate the interference covariance matrix hence the weight vector, and other is the primary data or the test data on which detection and the parameter estimation is performed. The objective is achieved by considering the fully adaptive processing under the assumption that interference covariance matrix can be estimated from the sufficient number of Independent and Identically Distributed (i.i.d) semi-supervised, only the target like signal is filtered out from the training data, sample data set.

Motivation of this thesis, for its effective suppression of clutter and jammer interferences by integrating STAP principle with sequential state estimation. In STAP-ASE, interference covariance matrix adaptively changes over the CPI due to nonstationary clutter and jammer interferences, and the model also updated interference covariance matrix within the filtering loop using immediate previous data set. Therefore, actual attainable interference suppression, and convergence rate based on worst case scenario are much higher than its counterparts. In this paper, the Posterior Cramer-Rao Lower Bound (PCRLB) has been analyzed from the dynamic state model perspective and this PCRLB has been considered as an achievable optimal point in order to cancel out the interferences, and mitigate the signal power. However, in this research, the Improvement Factor (IF) has been used as a key tool to evaluate the performance and efficiency of the proposed system. Maximum attainable SINR, the performance and the Improvement Factor(IF) have been examined and compared. The proposed system has also been compared with the Minimum Variance Distortionless Response (MVDR), Sample Matrix Inversion (SMI), and Loaded Sample Matrix

Inversion (LSMI). Proposed STAP-ASE model has shown outperform compared to counterparts, SMI, and LSMI for signal decoding, Improvement Factor (IF), SINR, efficiency, consistency and convergence rate.

Furthermore, proposed STAP with Adaptive State Estimation (STAP-ASE) approach, characterizes STAP simultaneously in spatial and Doppler domains for non-stationary, homogeneous and heterogeneous systems. The contributions presented here are based on the adjustment of the weight vector and the update of associated interference covariance matrix by STAP-ASE to minimize the output noise power while maximizing Signal Interference-plus-Noise Ratio (SINR) in the MSE sense, hence improving the Improvement Factor (IF) of the estimator. Finally, the integration of STAP principle with sequential state estimation in order to decode the target signal while rejecting the interferences due to nonstationary heterogeneous clutter and jammer effects without degrading the performance is the key contribution of this thesis.

Finally, the main highlighting ideas and principles of this thesis are as follows:

- Chapter 2, has included the technical background required for the Space-Time Adaptive Processing (STAP).

- In Chapter 3, the signal and the interferences due to the nonstationary heterogeneous clutter effect has been characterized. The goal is to state the signal and interference model upon which estimation and detection of the target of interest are being performed. The inclusion of the constraint weight model is ensured the consistency of the system and also ensured that the system can minimize the effects of the interference covariance matrix by keeping the characteristics (such as signal energy) of the signal unchanged. The main task of this constraint weight vector along with the predefined signal vector $\mathbf{c}$ is to approximate the desired output and to filter out the target like signal (semi-supervised) from the secondary data set. In this chapter, the equation for optimal weight vector has been stated, and also formulated the beamforming output as well as the maximum achievable Signal Interference-plus-Noise Ratio(SINR). Finally, it is also evident that the, Improvement Factor (IF) and the Minimum Variance Distortionless Response (MVDR) can be used as a key tool for the analysis of performance as well as efficiency of the estimator and the processor.

- Chapter 4 is the core chapter for this thesis. This Chapter has stated the problem formulations and the design architecture for the proposed STAP-ASE model. First part of this chapter, Sample Matrix Inversion (SMI), and Loaded Sample Matrix Inversion (LSMI) have been included, and it also defined the finite sample size structures for STAP processing using the existence SMI and LSMI algorithms. In the second part of this chapter, the Wiener filtering approach has been included as an alternative solution for the STAP processing, and showed that the Wiener filter is an optimal filter for time invariant linear system. Middle of the chapter, Extended Kalman Filter (EKF) has been stated, and this filtering approach has included as a sequential state estimation for the proposed STAP-ASE system. It is apparent from the STAP-ASE model that, the Extended Kalman Filter (EKF) is an important generalization of the Wiener filter for nonstationary and nonlinear systems. Finally, computational complexity for the proposed STAP-ASE model has been formulated; and the Posterior Cramer-Rao Lower Bound (PCRLB), and the efficiency of the proposed model have also been discussed.

- Chapter 5 has included the simulations. The proposed STAP-ASE with SMI, LSMI and MVDR (Optimal) has been compared, and found outperform performance compared to counterparts for signal decoding, Improvement Factor (IF), SINR, efficiency, and consistency.

6.2 Future Work

Several areas of research have been opened and highlighted by this work. From the designed architecture of chapter 4, two major parts of this model are very evident. One is for the estimation of the interference (asymptotic) covariance matrix after whitening the interference signal and other is the filtering part to cancel out the interferences from the observed signal in order to extract the target of interest. In this model, Extended Kalman Filter (EKF) has been used as a sequential state estimation, however other sequential state estimation like particle filter can also be used without major changing of this model architecture. For the simulations, the interference (asymptotic) covariance matrix is estimated based on the fully adaptive STAP-processing for semi-supervised training sample data set and under the assumption that there is sufficient number of Independent and Identical Distributed (i.i.d) secondary data is available in the vicinity of the test data cell. But, it is very likely to have insufficient number of Independent, and Identically Distributed data set due to the diverse nature of the interferences, hence question about the partially adaptive STAP-processing. This model can easily adapt with the partially adaptive STAP-processing by changing only the signal pre-whitening part of the designed architecture stated in Figure 4.3 of chapter 4, while keeping the other parts of the model remain the same.

Furthermore, there are potential applications in the field of neural network, or biometric signal processing. Classical problem of STAP-ASE is to extract signal from the mixture of unknown interferences, however the system doesn't have any prior knowledge about the interferences. Since model is based on the estimation of asymptotic interference covariance matrix, hence the weight vector in order to decode or extract signal of interest. Similarly, if the system may consider an interconnected neural network consisting of neurons, and a model which is used to estimate the weight vector in order to control interconnections, so that the network model will be able to find out the signal of interest. STAP-ASE can also be implemented with the Multiple Input and Multiple Output (MIMO) radar systems. Since, due to the diversity of the signal, it is possible to achieve the improved clutter resolution for the extracted signal from the radar receiver.

Finally, proposed STAP-ASE model does not consider the signal approximation error, and hence the robustness of estimator, since the basic research model is based on the minimization of the effects of the interference covariance matrix, maximization of the Signal Interference-plus-Noise Ratio (SINR) in the MSE sense, and hence improve the Improvement Factor (IF) in order to mitigate the signal power detection. Therefore, robustness of the processor, and robustness against interference fragmentation in sample covariance matrix are also the new waiting for the further exploration of STAP-ASE model.

Appendix A

Maximum Likelihood Estimate of the Sample Covariance Matrix

Lets begin with an assumption that, there are L observations $\mathbf{x} = [x_1, x_2, x_3, \ldots, x_L]^H$.

Where $\mathbf{x}$ is assumed to be zero mean and Independent and Identically Distributed (i.i.d) Gaussian random variable, and $(.)^H$ represents the hermitian transposition. The dimension of $\mathbf{x}$ is $L \geqq MN$ and M is the number of pulses and N is the number of array elements. Following derivation is based on the work presented in [35][36][37].

Now, the distribution can be stated as [35][36]:

$$f_{\mathbf{x}}(x_1, x_2, x_3, \ldots, x_{MN}) = \frac{1}{\pi^{MN}|Q|} e^{-\mathbf{x}^H Q^{-1} \mathbf{x}} \tag{A.1}$$

Where, $Q \in C^{MN \times MN}$, is unknown positive definite interference covariance matrix, and $|.|$ represents the determinant of Q.

Now, the associated likelihood conditioned on Q can be stated as [35][37]:

$$f_{\mathbf{x}}(x_1, x_2, x_3, \ldots, x_{MN}|Q) = \prod_{l=1}^{MN} f_{x_l}(\mathbf{x}_l|Q)$$

$$\begin{aligned}
&= \frac{1}{\pi^{MN\times L}|Q|^L} e^{-\sum_{l=1}^{MN} x_l^H Q^{-1} x_l} \\
f(\mathbf{x}|Q) &= \frac{1}{\pi^{MN\times L}|Q|^L} e^{-\sum_{l=1}^{MN} x_l^H Q^{-1} x_l} \\
&= \frac{1}{\pi^{MN\times L}|Q|^L} e^{-Tr(\mathbf{x}^H Q^{-1}\mathbf{x})} \\
&= \frac{1}{\pi^{MN\times L}|Q|^L} e^{-Tr(\mathbf{x}\mathbf{x}^H Q^{-1})} \\
&= \frac{1}{\pi^{MN\times L}|Q|^L} e^{-Tr(Q^{-1}\mathbf{x}\mathbf{x}^H)} \\
&= [\frac{1}{\pi^{MN}|Q|} e^{-Tr(Q^{-1}\hat{Q})}]^L \qquad \text{(A.2)}
\end{aligned}$$

Where, $Tr|.|$ represents the trace, sum of the diagonal elements, and from the matrix cookbook [38], $Tr(XY) = Tr(YX)$, and according to [35], it can also be considered that,

$$\begin{aligned}
\hat{Q} &= \frac{1}{L}\sum_{l=1}^{L} x_l x_l^H \\
&= \frac{1}{L}\mathbf{x}\mathbf{x}^H \qquad \text{(A.3)}
\end{aligned}$$

Now, the maximum likelihood of the interference covariance matrix can be obtained by minimizing the negative log-likelihood [13][37] or maximizing the likelihood function, and can be stated as:

$$\begin{aligned}
\hat{Q}_{ML} &\triangleq arg\max_{Q}[f(\mathbf{x}|Q)] \\
&= arg\min_{Q}[-lnf(\mathbf{x}|Q)] \\
&= arg\min_{Q}[ln|Q| + Tr(Q^{-1}\hat{Q})] \\
&= arg\min_{Q}[-ln|Q^{-1}| + Tr(Q^{-1}\hat{Q})] \\
&= arg\min_{Q}[-ln|Q^{-1}\hat{Q}^{-1}\hat{Q}| + Tr(Q^{-1}\hat{Q})]
\end{aligned}$$

$$= arg\min_{Q}[ln|\hat{Q}^{-1}| - ln|Q^{-1}\hat{Q}| + Tr(Q^{-1}\hat{Q})] \quad \text{(A.4)}$$

Since the system has assumed, $\hat{Q}$ is positive definite and $L \geqq MN$.

Now according to [34][35][36][37][38][39] equation A.4 can be restated as:

$$\begin{aligned}\hat{Q}_{ML} &= arg\min_{Q}[-ln|Q^{-1}\hat{Q}^{\frac{1}{2}}\hat{Q}^{\frac{1}{2}}| + Tr(Q^{-1}\hat{Q}^{\frac{1}{2}}\hat{Q}^{\frac{1}{2}}] \\ &= arg\min_{Q}[-ln|\hat{Q}^{\frac{1}{2}}Q^{-1}\hat{Q}^{\frac{1}{2}}| + Tr(\hat{Q}^{\frac{1}{2}}Q^{-1}\hat{Q}^{\frac{1}{2}}] \quad \text{(A.5)}\end{aligned}$$

Again, if $\Omega = \hat{Q}^{\frac{1}{2}}Q^{-1}\hat{Q}^{\frac{1}{2}}$, $\Omega = U\Lambda U^H$, and $\Lambda \in R^{MN \times MN} = \text{diag } \lambda_1, \lambda_2, \ldots, \lambda_{MN}$ is a diagonal matrix with positive eigenvalues, and unitary matrix [40] $U \in C^{MN \times MN}$, then according to [35], it can be written:

$$\begin{aligned}\Omega_{ML} &= arg\min_{\Omega}[-ln|\Omega| + Tr(\Omega)] \\ &= arg\min_{U\Lambda U^H}[-ln|(U\Lambda U^H)| + Tr(U\Lambda U^H)] \\ &= arg\min_{\Lambda}[-ln|(\Lambda)| + Tr(\Lambda)] \\ &= arg\min_{\lambda_1,\ldots,\lambda_{MN}}[-\sum_{l=1}^{MN} ln\ (\lambda_l) + \sum_{l=1}^{MN}(\lambda_l)] \quad \text{(A.6)}\end{aligned}$$

Because of the convexity [35][41], first order derivatives of the above equation A.6 with respect to each eigenvalues will be canceled out. So, it can be stated [35]:

$$\frac{\partial[-\sum_{l=1}^{MN} ln\ (\lambda_l) + \sum_{l=1}^{MN}(\lambda_l)]}{\partial \lambda_l} = 0 \quad \text{(A.7)}$$

$$\begin{aligned}-\frac{1}{\lambda_l} + 1 &= 0 \\ \lambda_l &= 1 \quad \text{(A.8)}\end{aligned}$$

where, $\forall\ l : l = 1, 2, \ldots, MN$

Therefore, it be can written [35]:

$$\begin{aligned} \Lambda_{ML} &= I \\ \Omega_{ML} = \hat{Q}^{\frac{1}{2}} Q_{ML}^{-1} \hat{Q}^{\frac{1}{2}} &= I \\ Q_{ML} = \hat{Q} &= \frac{1}{L}\sum_{l=1}^{L} x_l x_l^H \\ \hat{Q} = \frac{1}{L}\sum_{l=1}^{L} x_l x_l^H &= \frac{1}{L}\mathbf{x}\mathbf{x}^H \end{aligned} \quad \text{(A.9)}$$

Finally, desired equation for the Maximum Likelihood Estimate (MLE) of (asymptotic) interference covariance matrix:

$$\hat{Q} = \frac{1}{L}\sum_{l=1}^{L} x_l x_l^H = \frac{1}{L}\mathbf{x}\mathbf{x}^H \quad \text{(A.10)}$$

Appendix B

Complete Derivations of Wiener Filter Solution for Optimal Weight Vector

According to Figure 4.2 and from the first line of equation (4.10) of chapter 4, the Mean Squared Error (MSE) can be stated as:

$$\begin{aligned}
MSE = \mathbb{E}[e^2(t)] &= \mathbb{E}[(d(t) - \mathbf{w}_{fc}^H(t)\mathbf{x}(t))^2] \\
&= \mathbb{E}[d(t) - \mathbf{w}_{fc}\mathbf{x}(t)d(t) - \mathbf{w}_{fc}\mathbf{x}(t)^H] \\
&= \mathbb{E}[d^2(t) - \mathbf{w}_{fc}^H\mathbf{x}(t)d(t) - d(t)\mathbf{x}^H(t)\mathbf{w}_{fc} + \mathbf{w}_{fc}^H\mathbf{x}\mathbf{x}^H\mathbf{w}_{fc}] \\
&= \mathbb{E}[d^2(t)] - 2\mathbf{w}_{fc}^H\mathbb{E}[d(t)\mathbf{x}(t)] + \mathbf{w}_{fc}^H\mathbb{E}[\mathbf{x}(t)\mathbf{x}^H(t)]\mathbf{w}_{fc} \\
&= \sigma_d^2 - 2\mathbf{w}_{fc}^H\mathbf{p}_{dx} + \mathbf{w}_{fc}^H Q_x \mathbf{w}_{fc} \qquad \text{(B.1)}
\end{aligned}$$

where,

$$\mathbf{p}_{dx} = [p_{dx}(1), p_{dx}(2), \ldots, p_{dx}(MN)]^T \qquad \text{(B.2)}$$

is the cross correlation vector between $d(t)$ and $\mathbf{x}(t)$. σ_d^2 = covariance of desired signal and $\mathbf{w}^H\mathbf{x}(t)$ = scaler quantity.

Where Q_x is the correlation matrix of the observed data, and can be stated as:

$$Q_x = \mathbb{E}\begin{pmatrix} x(t) \\ x(t-1) \\ \cdots \\ x(t-MN) \end{pmatrix}\begin{pmatrix} x(t), x(t-1)\ldots, x(t-MN) \end{pmatrix}$$

$$= \begin{pmatrix} r_x(1) & r_x(2) & \ldots & r_x(MN) \\ r_x(0) & r_x(1) & \ldots & r_x(MN-1) \\ \ldots r_x(-MN) & r_x(-MN+1) & \ldots & r_x(0) \end{pmatrix}$$

Furthermore, it is also symmetric matrix, since in the case of Wiener filter, the model has assumed that the system is stationary, and linear.

The Wiener solution of the optimal weight vector, can be found by taking the second order derivation of equation B.1 with respect to weight vector $\mathbf{w}_{fc}$, and equating it with zero. Therefore, solution for the optimal weight vector $\mathbf{w}_{opt}$ can be stated as:

$$\begin{aligned} \frac{\partial^2(MSE)}{\partial w_{fc}^2} &= 0 \\ Q_x\mathbf{w}_{opt} &= \mathbf{p}_{dx} \\ \mathbf{w}_{opt} &= Q_x^{-1}\mathbf{p}_{dx} \end{aligned} \quad \text{(B.3)}$$

The above equation is known as the Wiener Hopf equation.

Appendix C

Derivation of Optimal Solution for Constraint Weight Vector

Rewriting the equation of section 3.2.2 in chapter 3 for the constraint weight vector $\mathbf{w}_c$:

$$\mathbf{c}^H \mathbf{w}_c = g \tag{C.1}$$

Where, $\mathbf{w}_c = [w_{c1}, w_{c2}, \ldots, w_{cMN}]^T$ is the constraint weight vector, $\mathbf{c}$ is the predefined steering vector whose elements are stated by look direction and signal Doppler frequency, and g is the constant gain due to the associated signal steering vector.

According to the optimal weight vector equation (1.2) stated in chapter 1, it can be written:

$$\mathbf{w}_c = Q^{-1}\mathbf{c} \tag{C.2}$$

Where, Q expected interference (asymptotic) covariance matrix.

Now, constrained optimization problem can be written as [25]:

$$\min_{\mathbf{w}_c} \ \mathbf{w}_c^H Q \mathbf{w}_c$$

$$\text{subject to: } \mathbf{c}^H \mathbf{w}_c = g \tag{C.3}$$

Again, combining equations C.1 and C.2, it can be written:

$$\begin{aligned} \mathbf{w}_c^H \mathbf{c} &= g \\ (Q^{-1}c)^H \mathbf{c} &= g \\ c^H Q^{-1} \mathbf{c} &= g \end{aligned} \tag{C.4}$$

Therefore, solution for this constrained optimization problem is:

$$\mathbf{w}_c = Q^{-1}\mathbf{c}[c^H Q^{-1}\mathbf{c}]^{-1} g \tag{C.5}$$

Appendix D

Space-Time Adaptive Processing: Performance Matrices

In STAP-ASE processing, there are two possibilities, either target is present or absent in the observed signal. Based on these possibilities, two hypotheses H_0 (target is absent) and H_1 (target is present) can be formulated [2]:

$$\begin{aligned} H_0 : \mathbf{x}^k &= \mathbf{q}_c^k + \mathbf{q}_j^k + \mathbf{q}_n^k \\ H_1 : \mathbf{x}^k &= \mathbf{s} + \mathbf{x}^k(H_0) \end{aligned} \tag{D.1}$$

Where, s, $\mathbf{q}_c^k$, $\mathbf{q}_j^k$, $\mathbf{q}_n^k$ are the target steering vector, clutter interference, jammer interference, and noise interference vectors respectively and k is the snapshots. Following derivation is based on the work presented in [2].

Now, consider a weight vector $\mathbf{w}^k$ applied to the joint Space-Time snapshots [2].

Since, Space-Time processing technique linearly combined the snapshot of the elements [71], then the scalar output of the processor can be stated as [2]:

$$\begin{aligned} y^k &= [\mathbf{w}^k]^H \mathbf{x}^k \\ &= \sum_{l=1}^{MN} [w_l^k]^H [x_1^k] \end{aligned} \tag{D.2}$$

If Q is the interference (asymptotic) covariance matrix as stated before, then likelihood ratio test, under the Gaussian assumption $x^k(H_0) \sim CN(0, Q^k)$ for optimal detection statistics of equation D.2 can be stated as [1][71]:

$$| y^k | \lessgtr_{H_1}^{H_0} \upsilon_T \tag{D.3}$$

Where, υ_T is the detection threshold for two hypotheses [2][71].

The performance of the above equation D.3 can be stated as [2]:

$$\begin{aligned} P_{fa} &= e^{-\frac{\beta_T^2}{2}} \\ P_D &= \int_{\beta_T}^{\infty} u\ e^{-\frac{u^2+\alpha^2}{2}}\ I_0(\alpha u) du \\ \beta_T &= \frac{\upsilon_T}{\sqrt{(\mathbf{w}^k)^H Q^k \mathbf{w}^k}} \end{aligned} \tag{D.4}$$

Where, P_{fa}, P_D *and* β_T are the probability of false alarm, probability of detection and normalized detection threshold respectively [2]. I_0 is the modified zero-order Bessel function of the first kind [2], and α can be stated as [2]:

$$\alpha = \sqrt{2 \times SINR} \tag{D.5}$$

Where, SINR is the Signal to Interference plus Noise Ratio as stated before.

Bibliography

[1] R. Klemm, "Principles of space-time adaptive processing", The Institution of Electrical Engineers, London, UK Publishing, 2002

[2] R. Klemm, "Applications of Space-Time Adaptive Processing", The Institution of Electrical Engineers, London, UK Publishing, 2004

[3] H. Y. Chen and T. C. Chiang, "Adaptive Beamforming Using the Constraint Kalman Filter", IEEE Transactions on Antennas and Propagation, Vol. 41, No. 1, November 1993, pp. 1576-1580.

[4] S. Haykin, "Kalman Filtering and Neural Networking", Wiley-Interscience Publication, 2001.

[5] http://en.wikipedia.org/wili/cramer-Rao-lower-bound, September 15, 2008.

[6] http://biblioteca.universia.net, "Adaptive Technique Applied to Offshore Dynamic Positioning Systems", September 15, 2008

[7] A. El-Keyi, T. Kirubarajan and A. B. Gershman, "Robust Adaptive Beamforming Based on Kalman Filter", IEEE Transactions on Signal Processing, Vol. 53, No. 8, August 2005.

[8] S. Dan and C. T. Tien, "Kalman Filtering with State Equality Constraint", IEEE, October 15, 2001.

[9] J. Ward, "Maximum Likelihood Angle and Velocity Estimation with Space-time Adaptive Processing Radar", Lincoln Liboratory Technical Report, 1997, pp. 1265-1267.

[10] E. C. Barlie, L. R. Fante and A. J. Torres, "Some Limitations on the Effectiveness of Airborne Adaptive Radar", IEEE Transactions on Aerospace and Electronics Systems, Vol. AES-28, No. 1, October 1992, pp. 1015-1032.

[11] J. Ward, "Space-time Adaptive processing for Airborne Radar", Lincoln Liboratory Technical Report TR-1015, December 13, 1994.

[12] G. N. Schoenig, "Contribution to Robust Adaptive Sginal Processing with Application to Space-time Adaptive Radar", Ph.d Thesis, Faculty of Virginia Polutechnic and State University, April 12, 2007.

[13] Y. Bar-Shalom, X. R. Li and T. Kirubarajan, "Estimation with Applications to Tracking and Navigation", CR YBS, Artech House, Norwood, MA, 2001.

[14] M. Hernandez, T. Kirubarajan and Y. Bar-Shalom, "Efficient MultiSensor Resource Management using Cramer Rao Lower Bounds", IEEE Transactions on Aerospace and Electronic Systems, Vol. 40, no. 2, 2004, pp. 399-416.

[15] X. Zhang and P. Willett, "Cramér-Rao Bounds for Discrete-Time Linear Filtering with Measurement Origin Uncertainties", Proceeding of the Workshop on Estimation, Tracking and Fusion, May 2001.

[16] R. Tharmarasa, T. Lang and T. Kirubarajan, "Joint Path Planining and Sensor Subset Selection for Multistatic Sensor Networks", Tracking and Fusion, May 2001.

[17] M. Hernandez, B. Ristic, A. Farina and L. Timmoneri, "Comparison of Two Posterior Cramér-Reo Bounds for Non-Linear Filtering with Applications", Tracking and Fusion, December 4, 2002.

[18] M. I. William and D. E. Mark, "Adaptive Cancellation Method for Geometry-Induced Nonstationary Bistatic Clutter Environments", December 4, 2002.

[19] S. Haykin, "Adaptive Filtering Theory" , Communication Research Laboratory, McMaster University, Hamilton, Canada, Prentice Hall, 1996.

[20] http://en.wikipedia.org, March 3, 2009.

[21] R. D. Yates, J. David and Goodman, "Probability and Stochastic Processes", John Wiley & Sons, Inc. 2nd Edition, 2005.

[22] http://http://www.experiencefestival.com/a/Radar/Principles, March 5, 2009.

[23] Y. Bar-Shalom and X. R. Li, "Multitarget-Multisensor Tracking: Principles and Techniques", 3rd Edition, 1995.

[24] M. Cherniakov, "Bistatic Radar Principle and Practice", John Wiley & Sons, Ltd. 2007.

[25] U. Nickel, "Principles of Adaptive Array Processing", Research Institute for High-Frequency Physics and Radar Techniques, September, 2006.

[26] R. G. Lorenz, "Robust Minimum Varaince Beamforming", IEEE Transactions on Signal Processing, Vol. 53, No. 5, May, 2005.

[27] M. Cherniakov, "Bistatic Radar Principles and Practice", John Wiley & Sons, 1997.

[28] http://en.wikipedia.org/wiki/Sidelobe, March 27, 2009.

[29] C. H. Gierull, "Analysis of the Small Sample Size Performance of Fast Fully Adaptive STAP Technique for MTI Radar.", Defence R & D, October 2001.

[30] "Chapter-7: Sample Matrix Inversion", http://etd.lib.fsu.edu/theses/available/etd-04092004-143712/unrestricted, April 07, 2009.

[31] B. D. Carlson, "Covariance Matrix Estimation Errors and Diagonal Loading in Adaptive Arrays", IEEE Transactions on Aerospace and Electronics Systems, Vol. 24, No. 4, July 1988.

[32] S. Haykin, "Adaptive Filter Theory", Prentice Hall, 4th Editon, 2002.

[33] J. R. Mailoux, "Phased Array Antenna Handbook", Norwood,MA: Artech House, 1994.

[34] J. Capon, "High-Resolution, Frequency-Wavenumber Spectral Analysis", Proc. of IEEE, Vol. 57, No. 8, 1969.

[35] J. R. Guerci, "Space-Time Adaptive Processing for Radar", Artech House, Inc. 2003.

[36] R. A. Horn and C. R. Johnson, "Matrix Analysis", Cambridge, England: Cambridge University Press, 1985.

[37] W. L. Melvin, "Space-Time Adaptive Radar Performance in Heterogeneous Clutter", Transactions on Aerospace and Electronic Systems, Vol. 36, No. 2, April 2000, pp. 621-633.

[38] "The Matrix Cookbook", http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/3274/pdf/imm3274.pdf, April 15, 2009.

[39] D. Pierre, "Optimazation Theory with Applications", New York: Dover, 1986.

[40] R. Klemm, "Space-Time Adaptive Processing: Principles and Applications", London, England: IEEE Press, 1998.

[41] S. Haykin, "Communication Systems", 3rd Edition, New York: John Wiley & Sons, 1994.

[42] M. D. Boroson, "Sample Size Considerations for Adaptive Arrays", IEEE Transactions on Aerospace and Electronic Systems, AES-16(4), 1980, pp. 446-451.

[43] I. Groger, "Nebenkeulenformung der Richtcharakteristik von Gruppenantennen", Ph.D. thesis. (in German), Technical University Berlin, 1979.

[44] D. J. Reed, S. L. Mallett and E. L. Brenna, "Rapid Convergence Rate in Adaptive Antenna Arrays", Transactions on Aerospace and Electronic Systems, AES-10(6), 1974, pp. 853-863.

[45] R. Klemm, "Space-Time Adaptive Processing, Stevenage", UK: IEE Press, 1998.

[46] I. M. Skolnik, "Radar Handbook", 2nd Edition, New York: McGraw-Hill, Inc., 1992.

[47] D. B. Carlson, "Covariance Matrix Estimation Errors and Diagonal Loading in Adaptive Arrays", IEEE Transactions on Aerospace and Electronic Systems, AES-24(4), 1988, pp. 397-401.

[48] R. Klemm, "Special Issue on Space-Time Adaptive Processing and Adaptive Arrays: Special Collection of Papers", IEEE Transactions on Aerospace and Electronic Systems. Vol. 36, No. 2, April 2000, pp. 508-509.

[49] R. J. Guerci and S. J. Bergin, "Principal Components, Covariance Matrix Tapers, and the Subspace Leakage Problems", IEEE Transactions on Aerospace and Electronic Systems, Vol. 38, No. 1, January 2002.

[50] A. R. Horn and R. C. Johnson, "Matrix Analysis", Cambridge, England: Cambridge University Press, 1985.

[51] G. P. Richardson, "Analysis of the Adaptive Space-Time Processing for Airborn Radar", IEE Proc. Radar, Sonar, Navig, Vol 141, No. 4, August 1994, pp. 187-195.

[52] G. P. Richardson, "Space-Time Adaptive Processing for Manoeuvring Airborne Radar", IEE ECEJ, Special issue on STAP, Vol. 11, No. 1, February 1999, pp. 57-63.

[53] G. P. Richardson, "Effects of Manoeuvre on Space-Time Adaaptive Processing Performance", Proc. IEE Radar 97, 14-16 October 1997, Edinburgh, Scotland, pp. 285-289.

[54] D. H. Rees and D. I. Skidmore, "Adaptive Attenuation of Clutter and Jamming for Array Radar", IEE Proc. Radar, Sonar and Navigation, Vol. 145, No. 4, August, 1998, pp. 193-108.

[55] S. I. Reed, D. J. Mallett and E. L. Brennan, "Repid Convergence Rate in Adaptive Arrays", Transactions on Aerospace and Electronic Systems, Vol. AES,-10, No. 6, Noverber 1974, pp. 853-863.

[56] G. P. Richardson, "STAP Covariance Matrix Structure and Its Impact on Clutter Plus Jamming Suppression Solutions", IEE Electronics Letter, Vol. 37, No. 2, 28 January 2001, pp. 1181-119.

[57] J. D. Rabideau, "Clutter and Jammer Multipath Cancellation in Airborn Adaptive Radar", IEEE Transactions on Aerospace and Electronic Systems, Vol. 36, No. 2, April 2000, pp. 566-583.

[58] J. D. Rabideau and A. Steinhardt, "Improved Adaptive Clutter Cancellation Through Data-Adaptive Training", IEEE Transactions on Aerospace and Electronic Systems, Vol. 35, No. 3, July 1999, pp. 879-891.

[59] M. A. Movich, S. J. Ayou and S. I. Reed, "Reduced rank STAP performance Analysis", IEEE Transactions on Aerospace and Electronic Systems, Vol. 36, No. 2, April 2000, pp. 664-676.

[60] M. D. Boroson, "Sample Size Consideration for Adaptive array", IEEE Transactions on Aerospace and Electronics Systems, 16, July 4 1980.

[61] A. W. Rihaczek, "Principles of High-Resolution Radar", Mark Resources, Inc, California, McGraw-Hill, Inc., 1977.

[62] N. J. Willis, "Bastatic Radar", SciTech Publishing, U.S.A, 2005.

[63] T. TSAO, B. Himed and H. J. Michels, "Effect of Interference Rank Estimation on the Detection Performance of Rank Reduced STAP Algorithms", IEEE Radarcon'98, 11-14 May 1998, Dallas, TX, pp. 147-152.

[64] X. Sun, N. Y. Peng, G. Q. Liu and J. D. Lu, "Several Ways for Clutter Suppression in Airborn Radar", Proc. IEEE Symposium on Antennas and Propagation, Seattle, 20-24 June 1994, pp. 152-155.

[65] Y. Seliktar, B. D. Williams and J. E. Holder, "Beam-Augmented STAP for Joint Clutter and Jammer Multipath Mitigation", IEE Proc. Radar, Sonar and Navigation, Vol. 147, No. 5, October 2000, pp. 225-232.

[66] G. Q. Liu, N. Y. Peng, X. Sun and E. Z. MA, "A Space-Time Transformation Approach for Adaptive Clutter Suppression in Airborn Radar", IEEE International Symposium on Circuits, and Systems, Chicago, IL., May 3-6 1993, pp. 247-250.

[67] L. W. Melvin, "Space-Time Adaptive Radar Performance in Heterogeneous Clutter", IEEE Transactions on Aerospace and Electronic Systems, Vol. 36, No. 2, April 2000, pp. 621-633.

[68] Y. Bar-Shalom, X. R. Li and T. Kirubarajan, "Estimation with Applications to Tracking and Navigation", New York: Wiley, 2001.

[69] T. C. Chen, "Linear System Theory and Design", New York: CBS, 1984.

[70] K. Lee and A. W. Bojanczyk, "Performance Modeling and Optimization Framework for Space-Time Adaptive Processing (STAP)", School of Electrical and Computer Engineering Cornell University, NY, USA, April 26-30 2004.

[71] W. L. Melvin, "A STAP Overview", IEEE Aerospace and Electronic Systems Magazine Vol. 19, No. 1, January 2004.