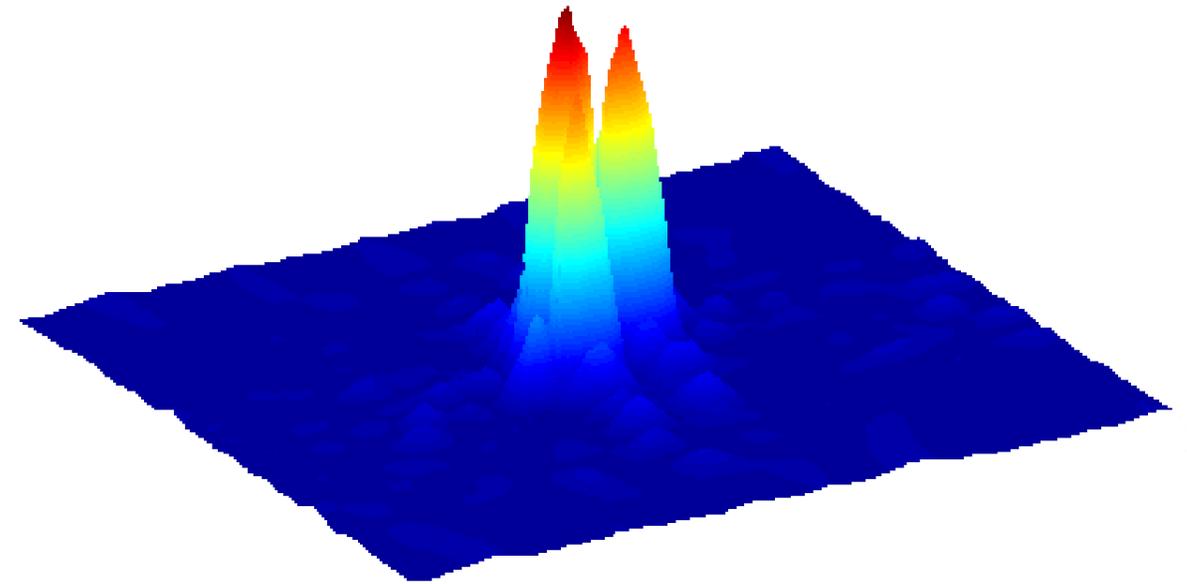


JÖRGEN ANDERTON



FOI, Swedish Defence Research Agency, is a mainly assignment-funded agency under the Ministry of Defence. The core activities are research, method and technology development, as well as studies conducted in the interests of Swedish defence and the safety and security of society. The organisation employs approximately 1000 personnel of whom about 800 are scientists. This makes FOI Sweden's largest research institute. FOI gives its customers access to leading-edge expertise in a large number of fields such as security policy studies, defence and security related analyses, the assessment of various types of threat, systems for control and management of crises, protection against and management of hazardous substances, IT security and the potential offered by new sensors.

Jürgen Anderton

Space Time Adaptive Processing for Sonar

Titel	Space Time Adaptive Processing för Sonar
Title	Space Time Adaptive Processing for Sonar
Rapportnr/Report no	FOI-R--2524--SE
Rapporttyp Report Type	Teknisk rapport Technical report
Sidor/Pages	39 p
Månad/Month	Juni/June
Utgivningsår/Year	2008
ISSN	ISSN 1650-1942
Kund/Customer	Försvarsmakten
Forskningsområde Programme area	4. Sensorer och signaturanpassning 4. Sensors and Low Observables
Delområde Subcategory	43 UV-teknik – sensorer 43 Underwater Technology - Surveillance, Target acquisition and Reconnaissance
Projektnr/Project no	E20608
Godkänd av/Approved by	Helena Bergman
FOI, Totalförsvarets Forskningsinstitut	FOI , Swedish Defence Research Agency
Avdelningen för Försvars- och säkerhetssystem	Defence & Security, Systems and Technology
164 90 Stockholm	SE-164 90 Stockholm

Sammanfattning

I praktiskt taget alla situationer där arrayer med element används, däribland radar och sonar, är estimering av bäring och hastighet av ett mål som på något sätt genererar en signal ett av huvudintressena för spaningen. Andra intressen är att filtrera signaler spatiellt och detektion av objekt. För att få ett bra estimat krävs en signal som är uppmätt på ett sätt som passar analysen i fråga. Det här examensarbetet är gjort på Totalförsvarets Forskningsinstitut, FOI. Målet är att implementera en fullt adaptiv space-time processor och genom det få insikt om svårigheterna inom området och försöka förlara dessa. Två olika metoder för att behandla interferenskovariansmatrisen beskrivs och utvärderas. Det visar sig att de adaptiva processorerna är klart bättre än den konventionella processen för analys av signaler, dock till priset av hög beräkningskomplexitet. Många förbättringar återstår att göra och huvudområdet där en större fördjupning behövs är estimeringen av kovariansmatrisen och även uppskattningen av antalet interferensskällor.

Nyckelord: Lobformning, STAP, Space-Time Adaptive Processing, Diagonal Loading

Summary

In both radar and sonar, in fact in all situations where an array of elements is used, one of the interests is to locate targets both spatially and in terms of velocity. In order to have a good estimate of a target's bearing and velocity, the signal that is analysed has to be well measured, and well suited for the kind of the analysis made. This thesis is carried out at FOI, the Swedish Defence Research Agency. The goal is to implement a fully adaptive space-time processor and identify the main difficulties and try to explain them. Two different methods for handling the noise and interference covariance matrix are described and evaluated. It turns out that the adaptive processors outperform the conventional way of analysing signals, however to a high computational cost. Still, many improvements are to be done and the main field of interest should be the estimation of the covariance matrix and also the estimation of number of interference sources.

Keywords: Beamforming, STAP, Space-Time Adaptive Processing, Diagonal loading

Contents

List of Figures	6
1 Abbreviations and symbols	7
2 Introduction	9
2.1 Background	9
2.2 Notation	10
2.3 Scope	10
2.4 Problem formulation	10
2.5 Model discussion	10
2.5.1 Assumptions	10
2.5.2 Physical limitations	11
2.5.3 Data model	11
3 Theory	17
3.1 Conventional Beamforming	17
3.2 Fully Adaptive STAP	18
3.3 Covariance Matrix Estimation	20
3.3.1 Diagonal loading	22
3.3.2 Eigenvalue Decomposition	22
4 Evaluation	25
4.1 Matlab	25
4.2 Evaluation criteria	26
4.2.1 Performance	26
4.2.2 Adapted pattern	30
5 Conclusion	35
5.1 Findings	35
5.2 Further improvements	35
Bibliography	39

List of Figures

2.1	Illustration of north-south ambiguity.	11
2.2	Illustration of scenario when a signal is emitted.	12
2.3	Illustration of the scenario when the reflected signals are recieved.	12
2.4	The motion induced Doppler-ridge.	13
2.5	M measurements over the array forms one $NM \times 1$ vector that forms the space-time snapshot.	15
3.1	Mainlobe clutter and the separate filters projected onto corresponding axis.	19
3.2	The space-time data cube.	21
4.1	The simulated model with the targets marked with white marks.	25
4.2	Targets hidden in the Doppler ridge.	26
4.3	Performance of the fully adaptive STAP with $K = 2NM$	27
4.4	Performance of the diagonally loaded STAP with $K = 2NM$	27
4.5	Performance of the fully adaptive STAP with $K = NM$	28
4.6	Performance of the diagonally loaded STAP with $K = NM$	28
4.7	Performance of the diagonally loaded STAP with $K = 40$	29
4.8	Performance of the diagonally loaded STAP with $K = 20$	29
4.9	Performance of the diagonally loaded STAP with $K = 40$. α is set to 30.	30
4.10	Adapted pattern for the fully adaptive processor with $K = 2NM$	31
4.11	Adapted pattern for the diagonally loaded STAP with $K = 2NM$	31
4.12	The adapted pattern for the fully adapted processor with $K = NM$	32
4.13	Adapted pattern for the diagonally loaded STAP with $K = NM$	32
4.14	Adapted pattern for the diagonally loaded STAP with $K = 40$	33
4.15	Adapted pattern for the diagonally loaded STAP with $K = 20$	33

1 Abbreviations and symbols

\otimes	Kronecker product
M	Number of temporal snapshots
N	Number of sensors
ω_t	Target frequency
ω_0	Carrier frequency
ϖ	Normalised Doppler frequency
d	Distance between adjacent sensors
c	Propagation speed in water, approx. 1500 m/s
λ	Wavelength
θ	Bearing
ϑ	Spatial frequency
\mathbf{v}	Steering vector
\mathbf{x}	Measured signal
$\hat{\mathbf{R}}$	Sample Covariance Matrix

2 Introduction

2.1 Background

Sonar¹ is an established technique for finding and mapping underwater objects. It is used in many different fields one of which is the estimation of a possible target's bearing and velocity. The traditional way of processing the measured signal sometimes becomes inaccurate since it does not take the properties of reverberation into consideration. Thus, it is desirable to filter the signal to limit the contribution from undesired sources. However, it is at great risk to filter more than needed if the weights of the filter are always constant. Space-Time Adaptive Processing (5), hereafter abbreviated STAP, is at help here, adapting the filter weights to become more suitable at one given moment.

The research on underwater surveillance systems has been extensive over the last decades and the evolution of digital signal processing has made it possible to implement complex algorithms in real time systems. The literature in the field of STAP is quite far-reaching and very well describes the theory (5),(12). Though, most of the research is done for STAP applied to the radar case and it is well understood by most people that the propagation difference between air and water implies a few differences in the application of the theory.

STAP is an active technique in which a signal is transmitted and properties of the environment is identified by studying the properties of the returns. Reflections that are received by the sensor may very well be reflections from the sea bottom or surface (6). These reflections are called clutter or reverberation. Although space-time clutter characteristics could be considered known through the angle-Doppler relation, eq. (2.3), adaptive techniques may be at help when non-adaptive methods fail due to e.g. perturbations in the sensor. If a model is built, the accuracy of the results depends on how exact the model is, and the more exact the model is the more sensitive to deviations in the real world it gets. Skipping the model building and instead making use of an adaptive filter is often a good way to work around the problem.

In STAP the adaptivity refers to the fact that one adjust the beamformer to the properties of the clutter, which also may be called interference. Considering that only first and second order properties the characteristics can be assembled in a covariance matrix of the interference. A lot of effort has been made trying to handle the large dimensions of the space-time covariance matrix. Fully adaptive processors though, are based on the full space-time covariance matrix. For small numbers of sensor elements and temporal samples the fully adaptive processor is quite easy to handle. However, as the number of elements and temporal samples grow, so will also the dimension of the covariance matrix, and the complexity of the processor becomes very high. Also, as the dimension gets larger, the amount of required training data grows and problems with estimating the sample covariance matrix arise. Often fully adaptive processing is used to give reference results to compare with when using suboptimum techniques.

¹Sound Navigation and Ranging

2.2 Notation

The notation used in this thesis is mainly taken from the MATLAB environment. In the matrix and vector notation $[x; y]$ means a 2×1 column vector while $[x, y]$ means a 1×2 row vector. If x and y instead are L long vectors $\mathbf{F} = [\mathbf{x}, \mathbf{y}]$ will be a $L \times 2$ matrix and $\mathbf{F} = [\mathbf{x}; \mathbf{y}]$ will correspond to a $1 \times 2L$ matrix. The \otimes is the Kronecker-product and $E\{\}$ denotes the expectation of a random quantity. When the index t is used at a variable, it denotes that the variable belongs to the desired target. When used as a variable it denotes time.

2.3 Scope

First the assumptions and model are described, after which the theory regarding covariance matrices and beamforming is ventilated. The evaluation part consists of a description of how this work is done and by which means the algorithms are evaluated. Finally some conclusions about the promises of STAP are given.

2.4 Problem formulation

The problem that is handled in this thesis comes from the fact that when an active sonar is used and a signal is emitted reflections will occur. As desired reflections of this signal will be hidden in interference emanating from the motion of the sonar platform and also by ambient noise the main task in this thesis is to implement a fully adaptive space-time processor that will detect the desired reflections against a background of interference plus noise. Figures 2.2 and 2.3 show the undesired and desired reflections. The implementation involves the problem of understanding that the main difficulty in implementing STAP in a somewhat realistic environment relates to the estimation of the interference covariance matrix. Since enough training data for the fully adaptive processor is seldom available (12), two different methods for estimation are described and compared. Besides that, a conventional beamformer is implemented for comparison reasons.

2.5 Model discussion

2.5.1 Assumptions

For this thesis, certain assumptions have been made regarding the signals. Every desired target and the interference are modelled as narrowband point sources. The sound waves are travelling in a media that is homogenous and non-dispersive. No consideration has been taken to elevation angles to targets, i.e. only azimuthal bearing is calculated. Further, the desired targets are assumed to lie in the far field, allowing for the postulation that arriving sound waves may be taken for planar. Desired source signals are considered independent and uncorrelated to the interference signals plus ambient noise. The ambient noise is regarded as normally distributed and isotropic, that is, equally strong in all directions. The sensors are omnidirectional and identical, equally spaced with $d = \frac{\lambda}{2}$ which implies that spatial aliasing is avoided (3). λ is the wavelength of the highest frequency expected to be received.

2.5.2 Physical limitations

In this work an uniform linear array, ULA, is used. The fact that the array is only used in 180° half-plane, north-south ambiguity arises, fig. 2.1. This means that two targets at opposite sides of the aperture will yield the same bearing estimate even though they are not at the same bearing at all.

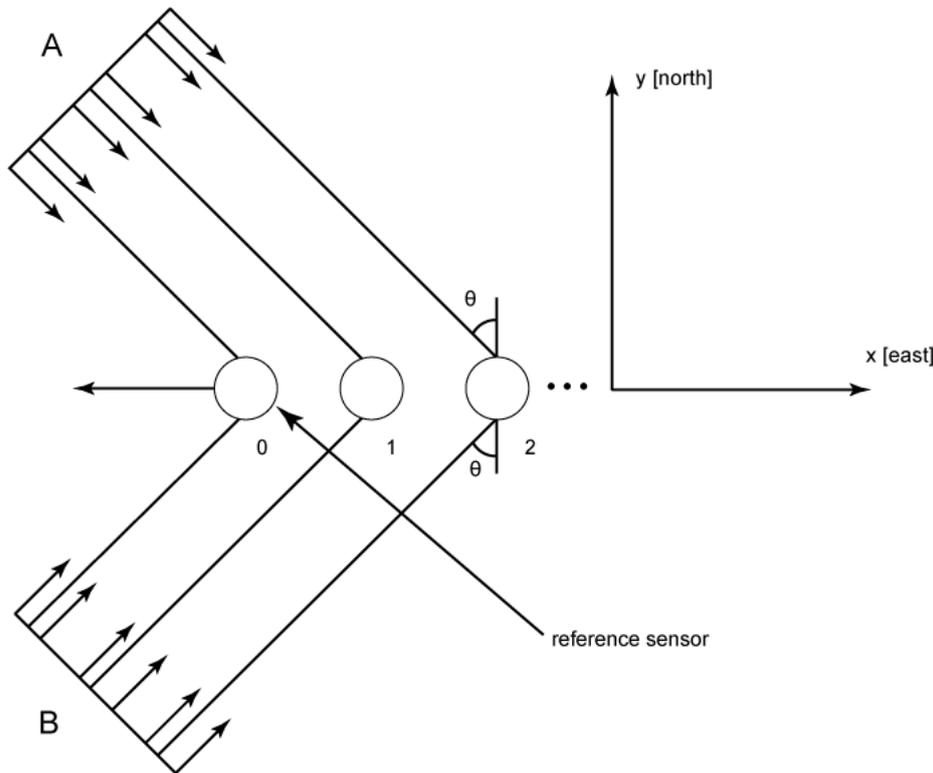


Figure 2.1: Illustration of north-south ambiguity.

2.5.3 Data model

Assume a narrow-band continuous wave signal is transmitted from the sonar at the green submarine like in fig. 2.2. The signal received by the sonar will not only be reflections from targets, like other submarines but also from the sea bottom and surface. These reflections are called reverberation and are visible in fig. 2.3 as the blue arrows. As mentioned earlier the reverberation also contains reflections from the sea bottom and surface. The green submarine is the one with sonar platform mounted on it and the red submarines are desired targets. The reverberation is in this work also called interference. The signal transmitted is a narrowband signal that, if a Hilbert transform based demodulation is applied, can be expressed as (12)

$$s_{tr}(t) = \alpha_{tr} u(t) e^{j(\omega_0 t + \phi)} \quad (2.1)$$

where α_{tr} is the random complex amplitude and ω_0 is the angular frequency and ϕ is the phase, this is usually called a pulse. $u(t)$ is the envelope of the signal and may be a window that suppresses sidelobes. The window used in this thesis though, is a rectangular one, hence $u(t) = 1$ when $0 \leq t \leq T$ and 0

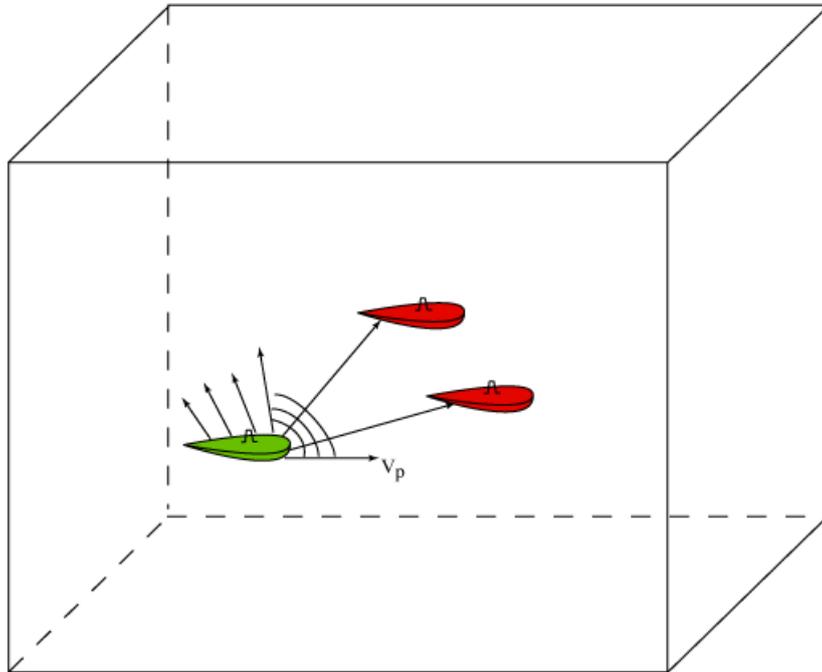


Figure 2.2: Illustration of scenario when a signal is emitted.

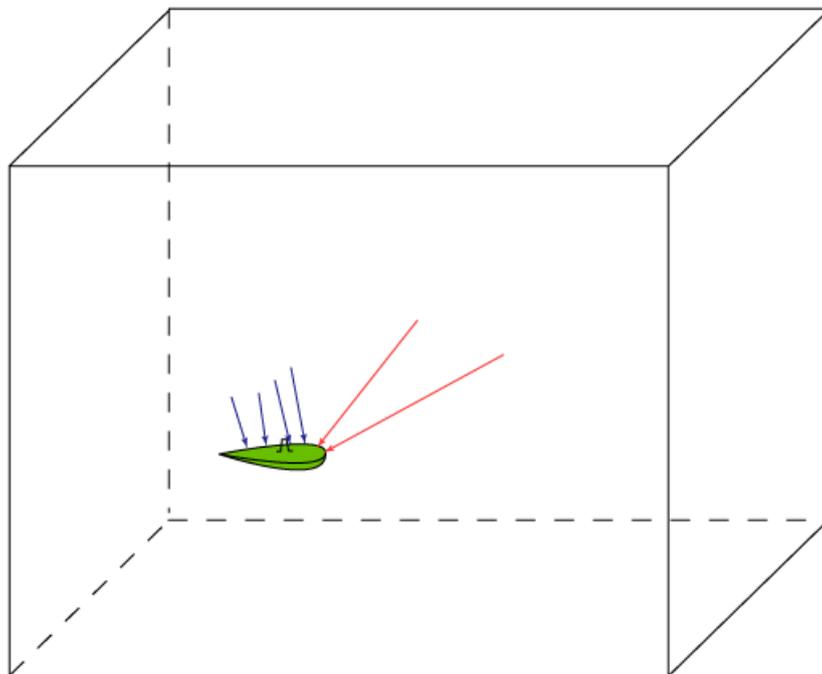


Figure 2.3: Illustration of the scenario when the reflected signals are recieved.

elsewhere. Hence, T is the pulse length. The received signal will be corrupted by a Doppler shift if the target and/or sonar platform are moving. If the sonar platform is moving then a Doppler shift will be applied to the reverberation according to (6)

$$\omega_t = \frac{2v_t}{\lambda_0} \quad (2.2)$$

where

$$v_t = V_p \cos \theta \quad (2.3)$$

which is the radial velocity of the target relative the platform. V_p is the velocity of the platform. Note that the Doppler shift is linearly dependent on $\cos \theta$, resulting in a diagonal ridge over the Doppler-Azimuth spectrum, fig. 2.4. The figure does not show a fully diagonal ridge because the bearing on the y-axis is given in degrees instead of cosine values. The Doppler effect is described by the azimuthal velocity of an object with respect to some reference point. It is this velocity that is calculated with the cosine of the angle between the object and the reference point and that is why the clutter interference is showing up like a diagonal ridge over the bearing-Doppler spectrum.

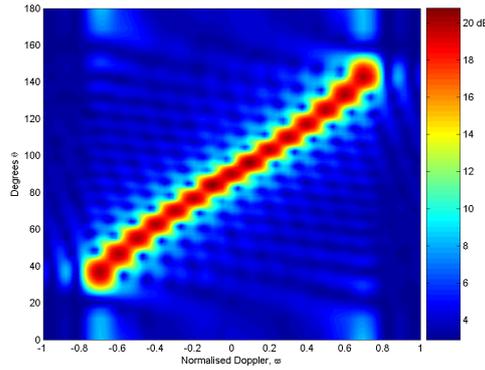


Figure 2.4: The motion induced Doppler-ridge.

The signal received by every of the N elements is expressed as

$$s_n(t) = \alpha_r u(t - \tau_n) e^{j((\omega_0 + \omega_t)(t - \tau_n) + \phi)}. \quad (2.4)$$

where α_r is the echo amplitude and ω_0 is the frequency of the transmitted signal. n denotes element $1 \dots N$. It is sometimes convenient to use the normalised Doppler frequency defined by

$$\varpi_t = \frac{\omega_t}{\omega_s} \quad (2.5)$$

where ω_s is sampling frequency. The normalisation gives that the target Doppler shift will lie in the interval $-\frac{\pi}{2}$ to $\frac{\pi}{2}$. The delay, τ_n , in (2.4) consists of two parts, the roundtrip delay, $\tau_t = \frac{2R_t}{c}$, R_t being the target range, plus the time delay between the reference sensor and n th element, which is described by

$$\tau'_n = -n \frac{d}{c} \cos \theta_t \quad (2.6)$$

when using an ULA. The *spatial frequency* is defined by

$$\vartheta_t = \frac{d}{\lambda_0} \cos \theta_t \quad (2.7)$$

and enables the phase delay to be written as

$$-\omega_0\tau'_n = n2\pi\vartheta_t. \quad (2.8)$$

If we assume a narrowband scenario then the relative delay, τ'_n , is negligible in the complex envelope term

$$\tilde{s}_n(t) = \alpha_r e^{j\phi} e^{jn2\pi\vartheta_t} u(t - \tau_t) e^{j\omega_t t} e^{j\omega_0 t} \quad (2.9)$$

where fixed phase terms have been included in the random phase ϕ . After conversion to baseband the signal is

$$s_n(t) = \tilde{s}_n(t) e^{-j\omega_0 t} = \alpha_r e^{j\phi} e^{jn2\pi\vartheta_t} u(t - \tau_t) e^{j\omega_t t}. \quad (2.10)$$

If we then consider only one specific τ_t corresponding to a specific range, the sampled target signal at snapshot k can be expressed as

$$\mathbf{x}_t = \alpha_t e^{jn2\pi\vartheta_t} e^{jm\varpi_t}, \quad \begin{matrix} n=0, \dots, N-1 \\ m=0, \dots, M-1 \end{matrix} \quad (2.11)$$

including the random complex phase term in α_t . As can be seen, one of the exponential terms depends on the spatial index n and the other on the temporal snapshot index m . These terms by themselves form spatial and temporal steering vectors, respectively. The spatial steering vector contains information about which direction the signal comes from and the temporal ditto describes the normalised Doppler shift that affects the signal when reflected by a target. One snapshot k therefore contains M samples over the array consisting of N elements stacked on top of each other, resulting in an $NM \times 1$ space-time steering vector for each angle and Doppler shift accounted for, see fig. 2.5, is formed

$$\mathbf{v}(\theta, \varpi) = \mathbf{b}(\varpi) \otimes \mathbf{a}(\theta). \quad (2.12)$$

where

$$\mathbf{a}(\theta) = \left[1, e^{j\pi\vartheta_t}, e^{j\pi2\vartheta_t}, \dots, e^{j\pi(N-1)\vartheta_t} \right]^T \quad (2.13)$$

and

$$\mathbf{b}(\varpi) = \left[1, e^{j\pi\varpi}, e^{j\pi2\varpi}, \dots, e^{j\pi(M-1)\varpi} \right]^T \quad (2.14)$$

are the spatial and temporal steering vectors respectively.

This is the response of a target at a certain bearing and normalised Doppler. Thus the target can be modelled as

$$\mathbf{x}_t = \alpha_t \mathbf{v}_t \quad (2.15)$$

where $\mathbf{v}_t = \mathbf{v}(\theta_t, \varpi_t)$ is called the target steering vector.

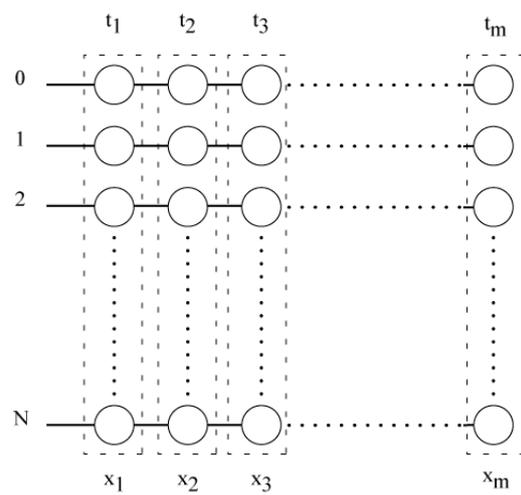


Figure 2.5: M measurements over the array forms one $NM \times 1$ vector that forms the space-time snapshot.

3 Theory

3.1 Conventional Beamforming

Conventional Beamforming is the most common method when estimating the direction of arrival. Basically, the method is to simulate a steering of the array's main beam in a certain direction. Every main beam is built up by different time delays τ_n on the sensors' output signals. The time delay will cause constructive and destructive interference and by that, it is possible to simulate the array looking in the directions of interest, and hence find the direction of arrival. Where the constructive interference is strongest, a mainlobe is formed. Constructive interference will also occur in other directions, adjacent to the direction of present interest as well, but will be much weaker. These lobes are called sidelobes. The sidelobe signals will also include contributions from targets and the sea itself (6). Those contributions are referred to as sidelobe clutter. The principle for passive and active processing is basically the same and here the theory is described for the passive processing. Passive processing means that signals emerging from possible targets are received and analysed to classify the object causing the signals. Active processing involves transmitting a signal from one's own platform which is reflected by a possible target. By analysing the received signal the target can be located and the speed of the target may be determined. With the expression given for the signal \mathbf{x}_t in (2.11) the output of the beamformer is a matched filter to \mathbf{x}_t . The information desired in \mathbf{x}_t is the complex exponential part of (2.11). Hence, matching \mathbf{x} with \mathbf{v} will show how well correlated the two signals are.

$$y = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} v^*(\theta_n, \varpi_m) x[n, m], \quad (3.1)$$

where $x[n, m]$ is the sampled version of $x(t)$ and v^* is a value in the steering vector corresponding to a certain bearing and Doppler shift. The asterisk denotes complex conjugate. (3.1) can be written as a vector multiplication

$$y = \mathbf{v}(\theta, \varpi)^H \mathbf{x}. \quad (3.2)$$

To find the direction of arrival $x(t)$ is matched with every possible θ and ϖ in \mathbf{v} . The power of the beamformer, y , is calculated by

$$P(\theta, \varpi) = E \{|y|^2\}. \quad (3.3)$$

Inserting (3.2) in (3.3) gives

$$P(\theta, \varpi) = E \{|\mathbf{v}(\theta, \varpi)^H \mathbf{x}|^2\} = E \{\mathbf{v}(\theta, \varpi)^H \mathbf{x} \mathbf{x}^H \mathbf{v}(\theta, \varpi)\} = \mathbf{v}(\theta, \varpi)^H \mathbf{R}_{xx} \mathbf{v}(\theta, \varpi) \quad (3.4)$$

since \mathbf{v} is deterministic. The power P is estimated in every direction to find the direction of arrival. The bearing that gives the maximum value of P is assumed to be the direction of arrival. Of course there may be several targets but the powers may not be the same. However, the power of a signal reflected by a target should be much stronger than the environmental noise and will therefore be detected as a target. The steering vector $\mathbf{v}(\theta, \varpi)$ is defined by the Kronecker product between the spatial and temporal steering vectors \mathbf{a} and \mathbf{b} described earlier,

$$\mathbf{v}(\theta, \varpi) = \mathbf{b}(\varpi) \otimes \mathbf{a}(\theta). \quad (3.5)$$

The conventional beamforming method is a robust method that more or less always works independently of the signal assumption. The disadvantage is that it can not make a difference between desired signals and interference. If conventional beamforming is used, the resolution depends on the number of sensors and the distance between the sensors. The minimum bearing separation between two targets is usually approximated by two different measures (11). The half power beamwidth, defined by

$$\theta_{cb} = \frac{\varepsilon\lambda}{Nd} \text{ radians} \quad (3.6)$$

and the distance to the first null from the top of the beam which is defined by

$$\theta_{cb} \approx \frac{\lambda}{Nd} \text{ radians} \quad (3.7)$$

where λ is the wavelength of the signal and Nd is the length of the aperture. ε is a constant which depends on the weights of the signal which means that if a window is used the resolution of the conventional beamformer may change, but as mentioned earlier this work assumes a rectangular window and for that case, $\varepsilon \approx 0.891$ radians. The two measures of θ_{cb} practically gives the same result and it is merely a question of which one wants to use.

3.2 Fully Adaptive STAP

The purpose of the sonar is to detect presence of targets in a noisy environment. Under some conditions the targets may be weak with respect to background noise and interference signals which makes it necessary to suppress the undesired signals to obtain appropriate detection performance. The assignment becomes exceptionally challenging if the sonar platform is moving, producing self-induced Doppler spread of reverberation. The whole idea with Space-Time Adaptive Processing is to use both spatial and temporal information simultaneously to suppress interference. Conventional processing usually processes the signal in the temporal domain at first, followed by spatial processing. The optimum spatial adaptive processor places a stopband determined by the transmit main beam which makes the sonar blind (5). Therefore, filtering in the spatio-temporal domain has two advantages over conventional processing. First, targets may be discerned. Second, as will be seen, the filtering is done better in the meaning of possibility in discovering slow moving targets. The space-time processor also makes it possible to discover weak targets that otherwise might be obscured by sidelobe clutter. As can be seen in fig. 3.1 the self-induced Doppler ridge is quite narrow if observed from the lower left corner. The space-time adaptive filter operates in that direction, applying only a narrow notch, instead of a relatively broad stopband-filter in each domain. It is well seen in the figure that slow targets which lie in the stopband of the adaptive Doppler filter may be detected when the spatio-temporal filter is applied. STAP is robust to system errors and has capability of handling non-stationary interference. However, if the interference has too much variability, i.e. undergoes severe variation between the range gate under test and the training gates, then the algorithm will deteriorate. The concept of training gates will be explained in 3.3. To calculate the angle-Doppler spectrum one wishes to apply some weight vector $\mathbf{w}(\theta, \varpi)$ to the measured signal at hand. The spectrum is then given by

$$P(\theta, \varpi) = |\mathbf{w}(\theta, \varpi)^H \mathbf{x}|^2. \quad (3.8)$$

Since the task is to mitigate interference and noise a natural design criterion

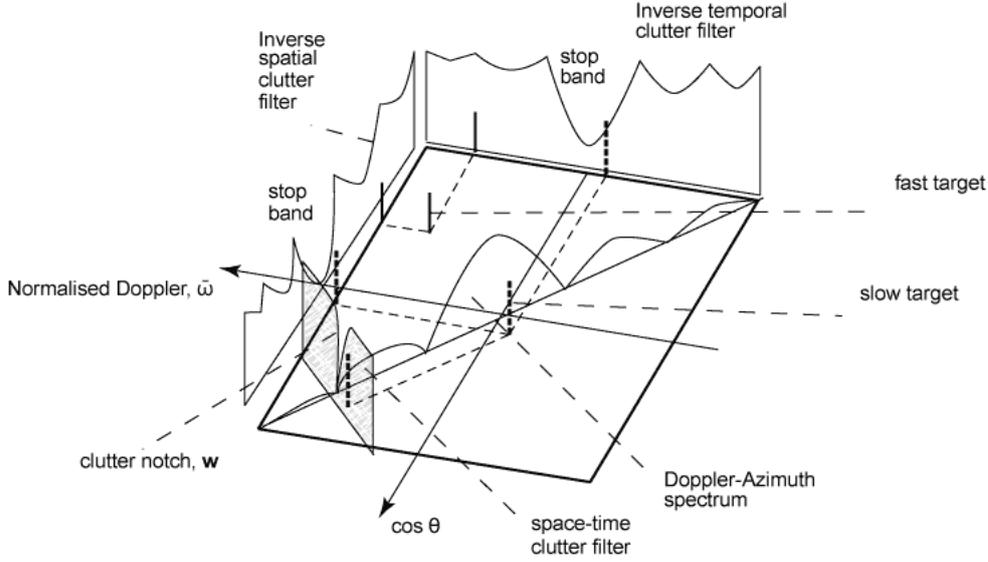


Figure 3.1: Mainlobe clutter and the separate filters projected onto corresponding axis.

for $\mathbf{w}(\theta, \varpi)$ is to maximize the signal to interference-plus-noise ratio (SINR). Under that approach the spatio-temporal beamformer is obtained as(2)

$$\mathbf{w}(\theta, \varpi) = \gamma \mathbf{R}_{nn}^{-1} \mathbf{v}(\theta, \varpi). \quad (3.9)$$

\mathbf{R}_{nn} is the interference covariance matrix, i.e. it describes the characteristics of the interference. The use of \mathbf{R}_{nn}^{-1} in 3.9 can be interpreted as that the interference is coloured and has to be whitened. Hence, a whitening filter applied to the measured signal \mathbf{x} . Note that the whitening filter is created from the interference, hence it is only the interference part of \mathbf{x} that is whitened. Performing the multiplication $\mathbf{R}_{nn}^{-1/2} \mathbf{R}_{nn} \mathbf{R}_{nn}^{-1/2}$ gives the identity matrix \mathbf{I} . Therefore, $\mathbf{R}_{nn}^{-1/2}$ is a whitening filter. $\mathbf{R}_{nn}^{-1/2}$ exists because \mathbf{R}_{nn} is positive semi-definite. The whitening of the signal \mathbf{x} implies that the steering vector \mathbf{v} also has to be whitened to match the signal, which gives

$$P(\theta, \varpi) = \left(\mathbf{R}_{nn}^{-1/2} \mathbf{v} \right)^H \mathbf{R}_{nn}^{-1/2} \mathbf{x} = |\mathbf{v}^H \mathbf{R}_{nn}^{-1} \mathbf{x}| = |\mathbf{w}^H \mathbf{x}|^2 \quad (3.10)$$

This is in the literature referred to as fully adaptive STAP (5). The term adaptive is deduced from the fact that the solution adapts to the characteristics of the interference, \mathbf{R}_{nn} . Using the Schwarz inequality it is shown that \mathbf{w} maximizes the SINR (2). γ from (3.9) is a constant and different values of γ results in different algorithms. However, the value of γ does not affect the SINR and is in this work chosen to be (9)

$$\gamma = \frac{1}{\sqrt{\mathbf{v}(\theta, \varpi)^H \mathbf{R}_{nn}^{-1} \mathbf{v}(\theta, \varpi)}} \quad (3.11)$$

yielding the complete calculation of \mathbf{w} as

$$\mathbf{w} = \frac{\mathbf{R}_{nn}^{-1} \mathbf{v}(\theta, \varpi)}{\sqrt{\mathbf{v}(\theta, \varpi)^H \mathbf{R}_{nn}^{-1} \mathbf{v}(\theta, \varpi)}}. \quad (3.12)$$

The weight vector \mathbf{w} is called the adaptive matched filter. P in (3.8) can be seen as a test variable, to be tested against two hypotheses. H_0 : Target absent and H_1 : Target present. When handling statistically estimated measures, it is always a risk that a test result like the one above proves wrong. Therefore design criteria exist such as probability of detection and probability of false alarm. The choice of γ also results in constant false alarm rate (CFAR)(9).

It is sometimes useful to apply a low-sidelobe window to the assumed target steering vector. This is usually called tapering. The tapering window is obtained by

$$\mathbf{t} = \mathbf{t}_b \otimes \mathbf{t}_a \quad (3.13)$$

where \mathbf{t}_b is the $M \times 1$ desired Doppler low-sidelobe window and \mathbf{t}_a is the desired low-sidelobe angle window. Then the low-sidelobe adapted pattern is produced by

$$\mathbf{g}_t = \mathbf{t} \odot \mathbf{v}_t. \quad (3.14)$$

If the steering vector is built in this way the term "tapered fully adaptive" is used. When the steering vector is tapered like this, not only are the sidelobes reduced but also the mainlobe is undesirably widened which leads to a trade-off.

The space-time adaptive processor has the great advantage of being able to suppress the undesired interference and noise. Though, the cost of getting this advantage is quite high, the matrix inversion is computationally heavy and also the required amount of training data for the fully adaptive processor is very high since \mathbf{R}_{nn} has to be estimated. Because of the fact that an extensive amount of training data is required other methods have evolved which may work very well even though they do not require the same amount of training data.

3.3 Covariance Matrix Estimation

As can be seen above the interference covariance matrix, (\mathbf{R}_{nn}), plays a large role in STAP. It is used to assemble information about the interference characteristics. In this thesis it is assumed that enough training data is available basically meaning that the interference plus noise covariance matrix is known. In practise this is not the case and the covariance matrix has to be estimated from adjacent range gates. The neighboring range gates to the gate under test are called guard gates and is not taken into consideration when estimating the covariance matrix since there may be leakage from the range gate under test. The range gates lying outside the guard gates are used as training data and are hence called the training gates. This reasoning is illustrated in fig. 3.2.

The training gates are assumed to be free of targets while still carrying the same interference characteristics. For natural reasons this is not the case when it comes to real data as the environment might change drastically depending on water depth etc. However, the assumption of same characteristics has to be done to enable the development of adaptive processors. The optimum performance is usually used as a reference measure when evaluating suboptimum techniques. The method used to estimate the covariance matrices is called the Sample Covariance Matrix (SCM) and is the simplest form of estimating a covariance matrix. The covariance matrix of interference plus noise is defined by

$$\mathbf{R}_{nn} = E \{ \mathbf{nn}^H \} \quad (3.15)$$

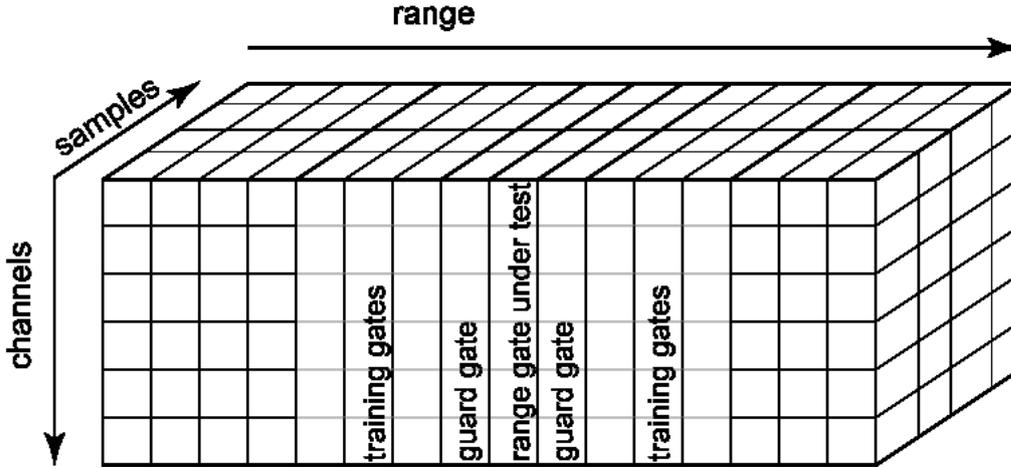


Figure 3.2: The space-time data cube.

which is estimated by the sample mean

$$\hat{\mathbf{R}}_{nn} = \frac{1}{K} \sum_K \mathbf{n}\mathbf{n}^H = \frac{1}{K} \mathbf{N}\mathbf{N}^H \quad (3.16)$$

where \mathbf{n} is a vector that contains jammers, i.e. intentional interference signals, clutter and noise respectively and \mathbf{N} is a matrix formed by several vectors (n). Jammers are disregarded in this work. The measured signal at the array sensors will be

$$\mathbf{x} = \mathbf{s} + \mathbf{n} \quad (3.17)$$

where \mathbf{s} is the desired signal component. The sample covariance matrix then is defined in the same way:

$$\mathbf{R}_{xx} = E\{\mathbf{x}\mathbf{x}^H\} \approx \frac{1}{K} \sum_K \mathbf{x}\mathbf{x}^H = \hat{\mathbf{R}}_{xx}. \quad (3.18)$$

To allow for the inversion of $\hat{\mathbf{R}}_{nn}$ and reach performance close to the optimum processor where the target steering vectors are used, the amount of training data, K , has to be sufficiently large. According to Brennan et al. (2), $K = 2NM$ to reach performance within 3 dB of optimum performance. For a realistic aperture and a realistic amount of snapshots, K becomes a very large number and gives unmanageable dimensions. Though, in a theoretical context the use of small dimensions alleviates the understanding process.

As described in sect. 3.1, the conventional beamforming method discerns both the desired signal and the jammers as targets which is unwanted. Instead a fully adaptive space time adaptive processing method was presented in section 3.2 which nulls out the interference and only discerns the desired component \mathbf{s} as a target.

Realising that the difficulty in STAP relates to the fact that it seldom exists the amount of data required by the fully adaptive approach, alternative ways of estimating the interference covariance matrix have evolved. One method to use is diagonal loading, described in 3.3.1.

3.3.1 Diagonal loading

If only the formula is considered, the method of diagonal loading does not seem too advanced. The SCM is updated with constants along its main diagonal,

$$\hat{\mathbf{R}}_{diag} = \hat{\mathbf{R}}_{nn} + \alpha \mathbf{I} \quad (3.19)$$

where \mathbf{I} is the identity matrix. The operation is also referred to as regularising of the SCM. The problem with this method is to define α , and unfortunately there is no clear theory in what way α should be designed. In (10) the empirical Bayesian approach is described. This thesis employs a strictly empirical method to design α . The diagonal loading compresses the eigenvalue spreading, which is desirable since it reduces the random noise eigenbeams. Loading the diagonal with a constant has little effect on large eigenvalues but affects originally small eigenvalues. The nulling of interference which is represented by large eigenvalues is not affected although the adaptivity against weak interference is clearly reduced. Adding the identity matrix to $\hat{\mathbf{R}}_{nn}$ is similar to adding a white noise constraint and forcing the weight vector \mathbf{w} to compensate for a higher noise level than the actual level. A too large choice of α will lead to a beamformer that becomes a scaled version of the conventional beamformer described in 3.1 since

$$\hat{\mathbf{R}}_{diag} \xrightarrow{\alpha \rightarrow \infty} \alpha \mathbf{I}. \quad (3.20)$$

A short derivation is given:

$$\mathbf{w} = \frac{(\alpha \mathbf{I})^{-1} \mathbf{v}(\theta, \varpi)}{\sqrt{\mathbf{v}(\theta, \varpi)^H (\alpha \mathbf{I})^{-1} \mathbf{v}(\theta, \varpi)}} \quad (3.21)$$

which in turn gives the expression for \mathbf{P} as

$$\mathbf{P} = \mathbf{w}^H \hat{\mathbf{R}}_{xx} \mathbf{w} = \frac{\frac{1}{\alpha^2} \mathbf{v}(\theta, \varpi)^H \hat{\mathbf{R}}_{xx} \mathbf{v}(\theta, \varpi)}{\frac{1}{\alpha} \mathbf{v}(\theta, \varpi)^H \mathbf{v}(\theta, \varpi)} \quad (3.22)$$

and since the steering vector \mathbf{v} has norm 1 the result is the same as the output of the conventional beamformer but scaled with $\frac{1}{\alpha}$. Also note that if $\alpha \rightarrow \infty$, $\mathbf{P} \rightarrow 0$. The derivation was only given to show that if α becomes large the space-time processor will approach the conventional beamformer. A large α will not necessarily be unimaginable large. During this thesis an α of 30 is large. Such an α will not cause \mathbf{P} to approach zero but will however cause the processor to approach the conventional beamformer.

The dependency of available data when choosing α is explained by the size of the eigenvalues of $\hat{\mathbf{R}}_{nn}$. When few snapshots of $\hat{\mathbf{R}}_{nn}$ are available, most eigenvalues will be zero or very close to zero. Then a very small α will still have effect on the SCM. Though, when more snapshots are used to form $\hat{\mathbf{R}}_{nn}$ but still not enough to allow for the inversion, the eigenvalues of $\hat{\mathbf{R}}_{nn}$ will be large in comparison to α and α will not have any effect on $\hat{\mathbf{R}}_{nn}$. This is also the reason to why diagonal loading is appropriate when small amount of training data exist. The regularisation will decrease the eigenvalues spread and hence improve the condition number of $\hat{\mathbf{R}}_{nn}$. The condition number is the ratio between the largest and smallest eigenvalues of $\hat{\mathbf{R}}_{nn}$. A high condition number corresponds to an ill-conditioned covariance matrix.

3.3.2 Eigenvalue Decomposition

One additional way to deal with the fact that large amounts of training data seldom exist is to use certain ways of inverting the SCM. Eigenvalue decom-

position is a method to get the significant eigenvalues and the corresponding matrices \mathbf{U} and \mathbf{V} containing the left and right eigenvectors, respectively. Any matrix can be decomposed as

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\text{H}}. \quad (3.23)$$

$\mathbf{\Sigma}$ is the matrix containing the eigenvalues or singular values. If the matrix \mathbf{A} is symmetric then $\mathbf{U} = \mathbf{V}$. Then the inversion of \mathbf{A} is carried out as

$$\mathbf{A}^{-1} = \mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{U}^{\text{H}}. \quad (3.24)$$

In (12) it is suggested that the effective rank of $\hat{\mathbf{R}}_{nn}$ is equal to the number of interference sources. Taking this into account gives opportunity to obtain the pseudo-inverse of $\hat{\mathbf{R}}_{nn}$. This means that the eigenvalues belonging to the noise subspace are not taken into consideration which in turn means that the condition number does not matter at all. In the real world it is almost impossible to know the number of interference sources that exist in data. The method of eigenvalue decomposition will therefore be very difficult to implement in a real application. Though, in theory the number of interference sources is often considered known which alleviates the use of this method. Results obtained by the use of this method will not be better than the ones obtained by regularisation. However, this method is not evaluated in this work but is brought up in section 5.2 as a possible way of handling the problem in the real world where the amount of training data is heavily limited.

4 Evaluation

4.1 Matlab

The evaluation made in the Matlab environment consisted of two targets, placed at $\mathbf{v}_1(\theta, \omega) = (85.9, -0.1)$ and $\mathbf{v}_2(\theta, \omega) = (94.5, 0.1)$, i.e. one target in direction 85.9 degrees with relative Doppler -0.1 and a second target in 94.5 with a relative Doppler of 0.1. The targets may be seen as slow moving targets if the frequency of the emitted signal and the frequency of the target reflections are close to each other. The signals are modelled as in eq. (2.10), i.e. in base band and with normalised Doppler frequency. Motion induced reverberation is modelled by 15 sources located in a diagonal ridge starting at 30 degrees up to 150 degrees and the relative Doppler shifts goes from -0.7 to 0.7. Figures 4.1 and 4.2 present the performance of the conventional beamformer. Fig. 4.1 shows the interference plus targets where the targets are marked with white asterisks. Fig. 4.2 illustrates the targets inside the motion induced Doppler ridge.

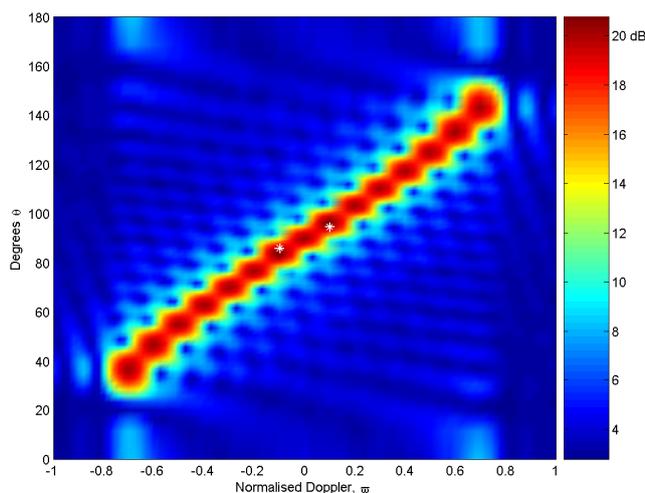


Figure 4.1: The simulated model with the targets marked with white marks.

The desired targets have power compared to the background noise giving an SNR of 7 dB and the reverberation sources are about 3 dB stronger than the targets. The signal is evaluated at 100 different angles and Doppler shifts from $0 - 180^\circ$ and -1 to 1 , respectively. The simulated uniform linear array consists of $N = 15$ equally spaced elements. The length of the continuous wave pulse is $M = 50$ samples in time which generates an overall scenario with dimensionality $NM = 750$. The model did not contain any information about the frequency of the transmitted signal since it was modelled as a received basebanded signal. The dimensionality implicates that the required amount of training data is $K = 2NM = 1500$. The evaluation is done by changing the amount of training data available to the algorithm.

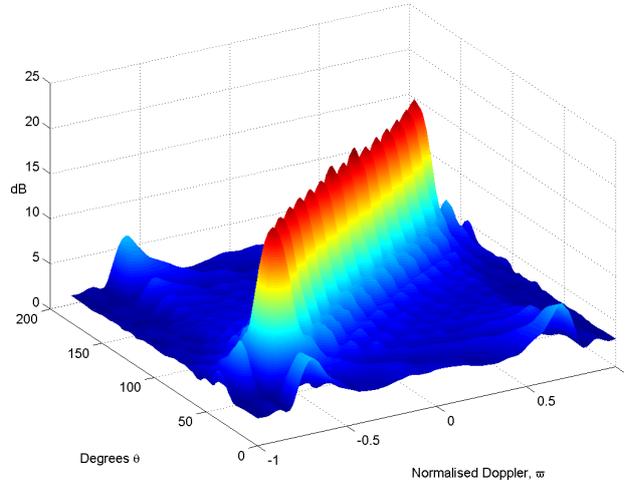


Figure 4.2: Targets hidden in the Doppler ridge.

4.2 Evaluation criteria

In this work three methods are evaluated, the conventional beamformer for comparison reasons and two methods of estimating the interference covariance matrix. There are different ways of comparing methods. One of the methods may not be better in all criteria and one has to decide which criterion is most significant. One very important criterion is if the algorithm under test actually works.

4.2.1 Performance

Fig. 4.3 speaks for itself. The reverberation from the stationary background is completely suppressed and the targets are detected at their correct bearings and Doppler shifts if $K = 2NM$. The separation of target \mathbf{v}_1 is explained by the fact that the target is placed right in the middle of the modelled interference. The processor therefore has placed a null where the interference is located which happens to be partly in the same bearing as the target. If the method of diagonal loading is used no difference is clearly visible as seen in fig. 4.4. In fact the diagonally loaded version tends to show parts of the interference due to a too large α compared to the large sample support. Fig. 4.5 shows the results when $K = NM$. This is the amount of data required to make inversion of the SCM possible at all. As we can see, no target can be distinguished from interference, due to an ill-conditioned $\hat{\mathbf{R}}_{nn}$. The diagonal loading method however, still reveals the targets as seen in 4.4. The sidelobes of the targets are less suppressed and some interference is visible. Nevertheless, the targets can be clearly discerned from interference. Reducing the amount of training data even more, to suit a more realistic situation, K is set to 40 i.e. twice the amount of modelled interference sources plus 10. The value is decided by the fact that the effective rank of the interference subspace of $\hat{\mathbf{R}}_{nn}$ is equal to the amount of interference sources, in our case 15. Brennan's rule says that the amount of training data should be twice that rank. Because of the small sample support, $\hat{\mathbf{R}}_{nn}$ becomes singular and inversion is not possible. We can still note in fig. 4.7 that the diagonal loaded approach is still able to distinguish the

targets from interference. Decreasing the sample support even more to see if one can make use of data collected in situations where the environment changes so much, that the required amount of data is not possible to collect gives the result shown in fig. 4.8. As a matter of fact, one can still see the targets even though the interference has become more visible. Decreasing the sample support even more will enhance that effect. Fig. 4.9 shows the result of the diagonal loading approach when using a too large α . Here $\alpha = 30$. As described in 3.3.1 the beamformer approaches the result of the conventional beamformer. This illustrates the importance of using a good value of α .

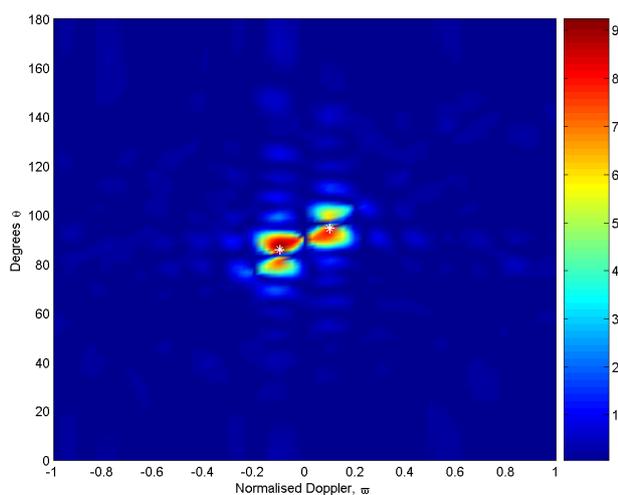


Figure 4.3: Performance of the fully adaptive STAP with $K = 2NM$.

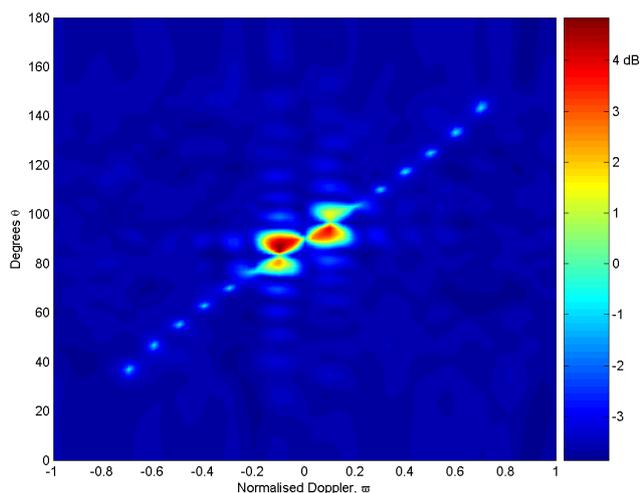


Figure 4.4: Performance of the diagonally loaded STAP with $K = 2NM$.

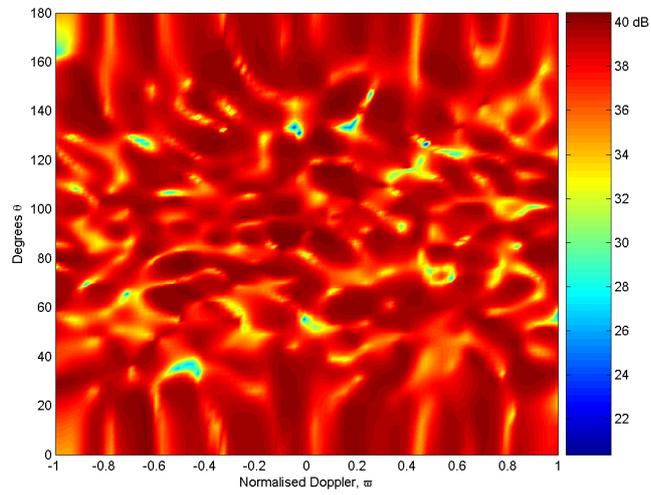


Figure 4.5: Performance of the fully adaptive STAP with $K = NM$.

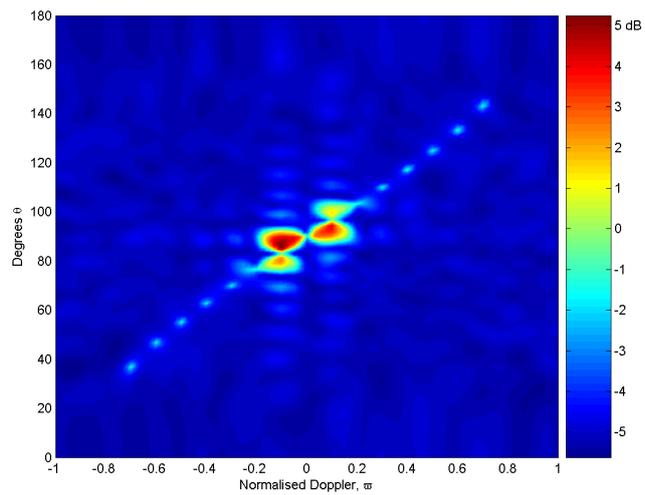


Figure 4.6: Performance of the diagonally loaded STAP with $K = NM$.

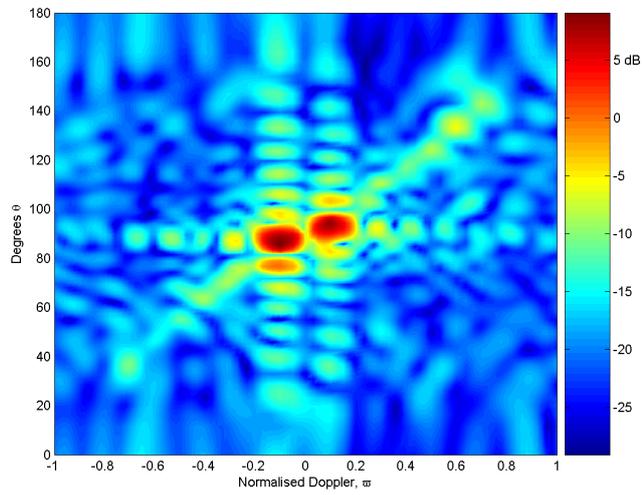


Figure 4.7: Performance of the diagonally loaded STAP with $K = 40$.

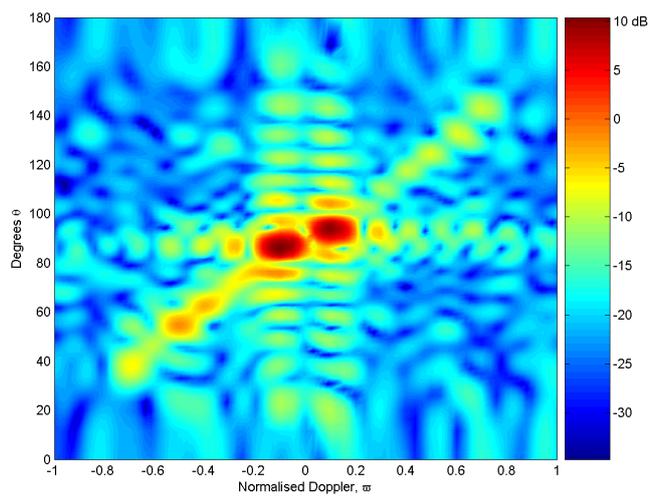


Figure 4.8: Performance of the diagonally loaded STAP with $K = 20$.

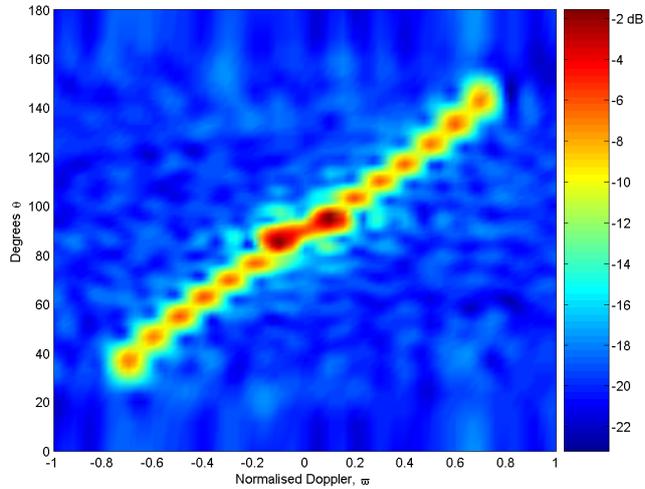


Figure 4.9: Performance of the diagonally loaded STAP with $K = 40$. α is set to 30.

4.2.2 Adapted pattern

Another thing to look at is how well the filter \mathbf{w} has adapted. This is called the adapted pattern. Given the weight vector in (3.9) the angle-Doppler response of the filter is calculated as

$$P_{\mathbf{w}}(\theta, \varpi) = |\mathbf{w}^H \mathbf{v}(\theta, \varpi)|^2. \quad (4.1)$$

This response is called the adapted pattern and shows how well the filter suppresses interference. Ideally, the adapted pattern has nulls in the directions of interference. To compare with the narrow notch shown in fig. 3.1 a cross section of the adaptive pattern is shown. Note that the values at the y-axis are $\cos \theta$ instead of just θ . This follows from that the motion induced Doppler ridge is somewhat curvy when seen from above. As described in eq.(2.3), the Doppler frequency is linearly dependent of $\cos \theta$ and not θ . Realising that these nulls may not be so deep in the case of real data, still the adapted pattern will give a hint on the promise of the algorithm under test. In fig. 4.10 it can clearly be seen that the processor places nulls in the clutter ridge and that the self induced Doppler interference is suppressed by about 17 dB. In fig. 4.11 the nulls are not as deep as in fig. 4.10 depending on the behaviour of the diagonally loaded approach. As described earlier the performance approaches the conventional beamformer, which indeed do not perform any nulling at all. The amount of suppression is naturally dependent on the strength of the interference, if the interference becomes stronger, then the adaptive filter will place a deeper null in the point of interference. However, if the strength of the reverberation is increased heavily, the covariance matrix will be ill-conditioned which impedes inversion. Though, the algorithm tends to be robust for clutter to noise ratios up to 50 dB. As expected, the adapted pattern for the fully adaptive STAP when $K = NM$, as shown in fig. 4.12, does not contain any valuable information but is only shown as a comparison to the adapted pattern of the diagonally loaded version of STAP. In the other cases where the number of samples has been heavily limited we can see that when diagonal loading is used the adapted pattern takes almost the same shape as for unlimited amounts of training data. Although, the notches get more shallow as the

amount of training data decreases. This can be seen in fig. 4.14 and 4.15. The reason for this is that when α is added to the main diagonal, the performance of the processor will approach the one of the conventional beamformer. This will lead to a not so efficient suppression of the interference.

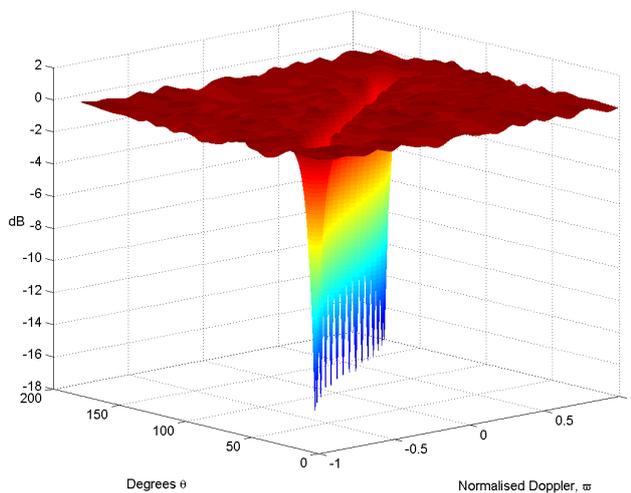


Figure 4.10: Adapted pattern for the fully adaptive processor with $K = 2NM$

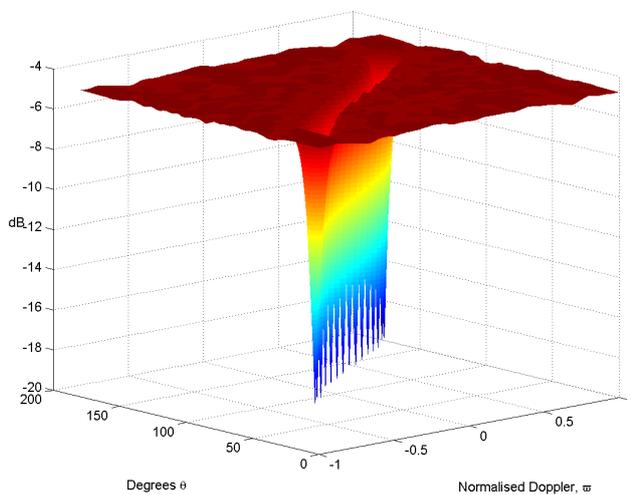


Figure 4.11: Adapted pattern for the diagonally loaded STAP with $K = 2NM$.

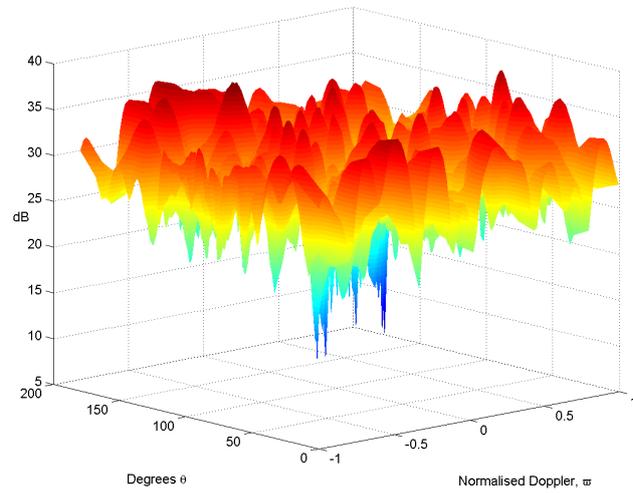


Figure 4.12: The adapted pattern for the fully adapted processor with $K = NM$.

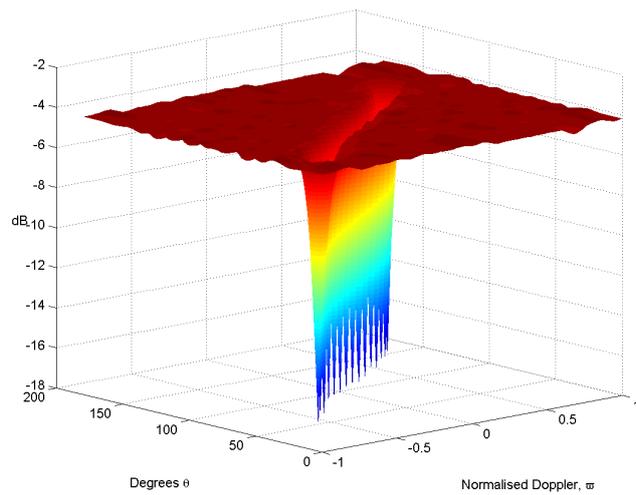


Figure 4.13: Adapted pattern for the diagonally loaded STAP with $K = NM$.

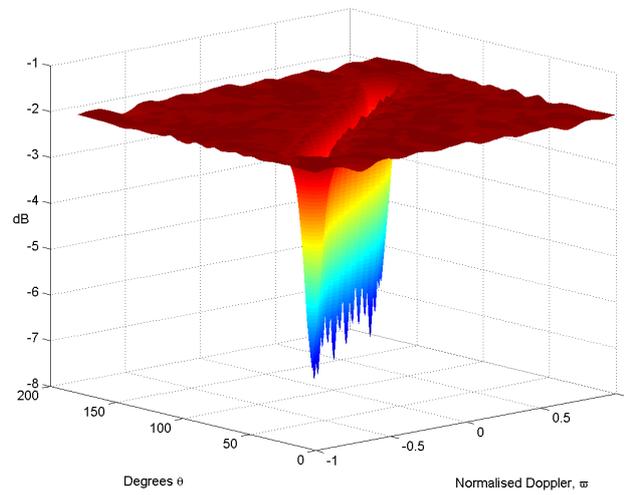


Figure 4.14: Adapted pattern for the diagonally loaded STAP with $K = 40$.

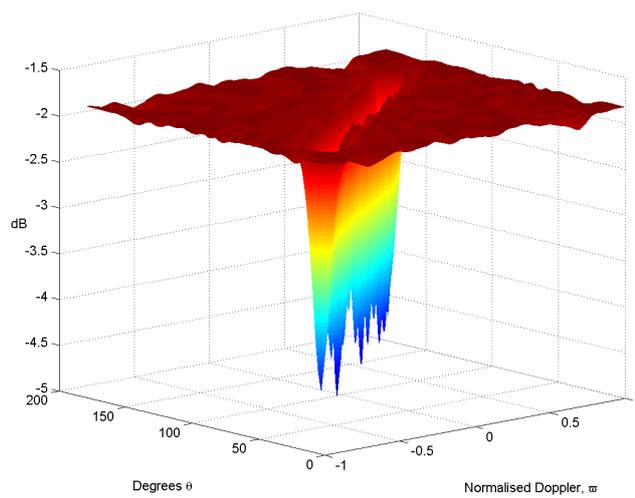


Figure 4.15: Adapted pattern for the diagonally loaded STAP with $K = 20$.

5 Conclusion

5.1 Findings

It has been shown that the fully adaptive space time processor is much better than the conventional beamformer at resolving targets and for suppressing reverberation. However, the results from the fully adaptive processor may be somewhat misleading since they depend on an unrealistic large amount of training data. The interference covariance matrix can never be completely known as is the case in this thesis. But as mentioned earlier, these results are often used as reference when evaluating suboptimum techniques. If the space time adaptive processor is to be used, other methods for estimating the covariance matrix should be employed. It is shown how the algorithm tends to fail when enough training data is not acquired and also the most common remedy to this problem. The method of diagonal loading was described in 3.3.1, which showed that STAP may be useful even in situations where small amount of training data is available. Other methods exist that address the issue of small training data sets but these are beyond the scope of this thesis to describe such methods although some references are given which enables further reading.

5.2 Further improvements

A natural step after this thesis is to explore real data and to investigate the stationarity of the reverberation in the range bins adjacent to the one under test. As mentioned in section 3.3 the real world will not be as nice as the assumptions made in this thesis. The result of the investigation will most probably show that the reverberation cannot be considered fully stationary and the following analysis then has to determine the stationarity of the reverberation in those range bins. This is of course not a simple analysis and a lot of theory on statistical analysis has to be studied. Nevertheless, the real world is of great interest and the analysis must be done. The real world also offers multipath propagation of the sound and all other intrinsic properties of the medium.

Further work may be done to explore and exploit the low rank structure of the interference covariance matrix as suggested in (8) where some promising results from the Multistage Wiener Filters (4) are shown. The method using MWF reduces the dimension of the problem in an adaptive way and it would be of great interest to implement such a processor and apply it to real data. The low rank structure is also described in (12), where a partially adaptive taxonomy is listed. One should keep in mind that even though these methods may offer good ways of estimating \mathbf{R}_{nn} , the computational complexity may not be reduced, it will rather increase. Still, that may not be the main concern since the performance characteristics of today's computers is very high. Even though Brennan's rule still applies, the order of complexity will decrease from $2NM$ to maybe $2 \times \text{rank}(\mathbf{R}_{nn})$ and the rank of \mathbf{R}_{nn} is much less than the dimension of the \mathbf{R}_{nn} . The singular value decomposition was described in 3.3.2 and is also a way of exploiting the low-rank structure of the interference. In further works that method may be well implemented. Nevertheless, the difficulty with such a method is, as mentioned before, the estimation of interference sources. Overall, implementing techniques that employ a structure that reduces the dimension

is almost a necessary step in order to evaluate real data.

Acknowledgements

This Master thesis has been carried out between August 2007 and February 2008 at the Swedish Defence Research Agency, FOI, in Kista, Sweden.

I would like to thank my supervisors at FOI, Magnus Lundberg-Nordenvaad and Elias Parastates for their support, without them it had been much harder succeeding. It has been a great pleasure doing my thesis at FOI and I thank all the people at the institution for Underwater Research for being so helpful.

Bibliography

- [1] Brennan, L.E., Mallet, J.D, Reed, I.S, (September 1976) "Adaptive Arrays in Airborne MTI Radar", *IEEE Transactions on Antennas and Propagation*, vol 24, no. 5, pp. 607-615
- [2] Brennan, L.E., Reed, I.S, (March 1973) "Theory of Adaptive Radar", *IEEE Transactions on Aerospace and Electronic Systems*, vol AES-9, no. 2, pp. 237-252
- [3] Burdic, W.S. (1984), *Underwater Acoustic System Analysis*, Prentice-Hall, ISBN 0-13-936716-0
- [4] Goldstein, J.S, Reed, I.S, Scharf, L.L, (November 1998) "A multistage representation of the wiener filter based on orthogonal projections", *IEEE Transactions on Information Theory*, vol 44, Issue 7, pp. 2943-2959.
- [5] Klemm, R. (2002), *Principles of Space-Time Adaptive Processing, IEE Radar, sonar, navigation and avionics series; no. 12*, London, The Institution of Electrical Engineers.
- [6] Klemm, R. (2004), *Applications of Space-Time Adaptive Processing, IEE Radar, sonar, navigation and avionics series; no. 14*, London, The Institution of Electrical Engineers.
- [7] Krim H., Viberg, M. (July 1996), "Two Decades of Array Signal Processing Research", *IEEE Signal Processing Magazine*, vol 13, Issue 4, pp. 67-94.
- [8] Nilsson, B, Karlsson, P-A, Karasalo, I, Lundberg, M, Parastates, E, (November 2007) "Active sonar surveillance in shallow waters", *FOI Technical report*, FOI-R-2407-SE, ISSN-1650-1942
- [9] Robey, (January 1992), "A CFAR adaptive matched filter", *IEEE Transactions on Aerospace and Electronic Systems*, vol 28, no. 1, pp. 208-216
- [10] Svensson, L, Lundberg, M, (September 2005) "On posterior distributions for signals in Gaussian noise with unknown covariance", *IEEE Transactions on Signal Processing*, vol 53. Issue 9 pp. 3554-3571.
- [11] van Trees, H.L. (2002), *Optimum array processing, part IV of detection, estimation and modulation theory* London, John Wiley & Sons Inc. 2002, ISBN 0-471-09390-4
- [12] Ward, J. (December 1994) "Space-Time Adaptive Processing for Airborne Radar", *Technical Report 1015*.