DEFENCE R&D DÉFENSE

Analysis of the Small Sample Size Performance of Fast Fully Adaptive STAP Techniques for MTI Radar

Christoph H. Gierull

Defence R&D Canada
TECHNICAL REPORT
DREO TR 2001-079
October 2001

National Defence Défense nationale

Canada

Analysis of the Small Sample Size Performance of Fast Fully Adaptive STAP Techniques for MTI Radar

Christoph H. Gierull
Defence Research Establishment Ottawa

Defence Research Establishment Ottawa

Technical Report

DREO TR 2001-079

October 2001

Abstract

In ground surveillance from an airborne or space-based radar it is desirable to be able to detect small moving targets, such as tanks or wheeled vehicles, within severe ground clutter. For operational moving target indication (MTI) systems the clutter filter coefficients have to be updated frequently due to rapidly changing interference environment. This report examines the small sample size performance of different fast fully adaptive space-time processors (STAP) and compares it to the optimum-detector performance. These recently proposed techniques, named Matrix Transformation based Projection (MTP) and Lean Matrix Inversion (LMI), were originally developed to provide fast man-made jammer suppression in large surface phased array radars with many elements. For this application they have been proven to operate with near-optimum performance, yet with a computational expense drastically reduced from that of the optimum detector in most practical cases. The investigation herein focuses on the performance achieved when only a few data samples are available to adapt (update) the clutter filter coefficient.

In this report, the techniques are described and a number of simulations carried out. The two applications, STAP and jammer suppression, are similar; both are required to suppress an interference which is characterized by a certain number of dominant eigenvalues of the sample space-time covariance matrix. Despite the similarities the performance between the two differs due to the different shapes of their eigenvalue distribution. The LMI is shown to give the best Signal-to-Noise-plus-Clutter Ratio (SNCR) for a given computational load.

Résumé

Ce rapport examine la performance de différents processeurs spatio-temporels totalement adaptables (STAP) à haute vitesse pour les échantillons de petite taille, et il la compare à la performance du détecteur optimal. Les techniques proposées récemment, appelées projection à base de transformation matricielle (MTP) et inversion matricielle oblique (LMI), ont été mises au point à l'origine dans le but d'assurer la suppression rapide du brouillage de source artificielle dans les radars à réseau en phase présentant une grande surface et dotés de nombreux éléments. Pour cette application, il s'est avéré qu'ils fournissent une performance quasi-optimale et qu'ils exigent, dans la plupart des cas pratiques, des ressources de calcul considérablement réduites par rapport au détecteur optimal. L'étude décrite porte essentiellement sur la performance obtenue lorsque seulement quelques échantillons de données sont disponibles pour adapter (mettre à jour) le coefficient de filtrage du clutter.

Ce rapport décrit les techniques utilisées et un certain nombre de simulations effectuées. Les deux applications, STAP et la suppression de brouillage, sont semblables, et les deux sont nécessaires pour éliminer un brouillage caractérisé par un certain nombre de valeurs propres dominantes dans la matrice de covariance spatio-temporelle de l'échantillon. En dépit des similitudes, les performances des deux se distinguent par les formes différentes de leur répartition des valeurs propres. Il est démontré que la LMI procure le meilleur rapport signal sur bruit plus clutter (SNCR) pour une charge de calcul donnée.

Executive summary

Background

In ground surveillance from an airborne or space-based radar it is desirable to be able to indicate small moving targets, such as tanks or wheeled vehicles, within severe ground clutter. The clutter echoes received by the radar show a motion-induced spread of the Doppler spectrum which particularly masks out the slowly moving targets. The problem is to suppress this clutter effectively.

The optimum technique to address this problem is space-time adaptive processing (STAP). In STAP an array of antenna elements, N, as well as a number of pulses, M, is used to estimate the clutter. The optimum processor is given by the inverse of the space-time covariance matrix of the clutter. Since this matrix is of order NM, the full inversion is too computationally intensive to adapt the processor frequently. By exploiting the fact that the rank of the clutter covariance matrix is, in most cases, much smaller than the order (due to the redundancy in the space-time domain), it is possible to develop techniques requiring less computational effort.

Principal results

In this report, two different, recently developed, fast rank reduction techniques, the so called subspace or projection methods, are applied to the problem of clutter suppression. These methods were originally proposed to provide fast man-made jammer suppression in large surface phased array radars with many elements. The basic theory behind them is briefly reviewed and their performance is studied via simulations.

Significance of results

While most investigations in the literature deal only with the asymptotic performance, i.e. an infinite number of data samples, studies of the small sample size behaviour are rare. Finite sample size is the practical case for operational radar systems in which the filtering coefficients have to be adapted frequently along the range dimension. The techniques described in this report are found to produce superior Signal-to-Noise-plus-Clutter Ratios (SNCR) for a given computational complexity. The most computationally efficient algorithm reduces the requirement for high speed computation by $10 - 1000$ times in many practical cases.

Christoph H. Gierull. 2001. Analysis of the Small Sample Size Performance of Fast Fully Adaptive STAP Techniques for MTI Radar. DREO TR 2001-079. Defence Research Establishment Ottawa.

Sommaire

Contexte

Pour la surveillance air-sol, il est souhaitable de pouvoir détecter de petites cibles mobiles, comme des blindés ou des véhicules à roues, dans du clutter de sol intense. Les échos de clutter reus par l'aéronef montrent un étalement du spectre Doppler causé par le mouvement, qui masque particulièrement les cibles mobiles lentes. Le problème consiste à supprimer efficacement ce clutter. La technique optimale pour aborder ce problème est le traitement spatio-temporel adaptable (STAP). Selon cette technique, un réseau d'éléments d'antenne, N, ainsi qu'un certain nombre d'impulsions, M, sont utilisés pour évaluer le clutter. Le processeur optimal correspond à l'inverse de la matrice de covariance spatio-temporelle. Comme cette matrice est de l'ordre N M, l'inversion complète nécessite trop de calculs pour l'adaptation fréquente du processeur. En exploitant le fait que le rang de la matrice de covariance du clutter est, dans la plupart des cas, beaucoup plus faible que l'ordre (en raison de la redondance dans le domaine spatio-temporel), il est possible de mettre au point des techniques exigeant moins d'efforts de calcul.

Principaux résultats

Ce rapport décrit comment deux techniques différentes de réduction rapide du rang qui ont récemment été mises au point, soit les méthodes subspatiales ou de projection, sont appliquées au problème de suppression du clutter. Ces méthodes ont été mises au point à l'origine dans le but d'assurer la suppression rapide du brouillage de source artificielle dans les radars à réseau en phase présentant une grande surface et dotés de nombreux éléments. Les principes sous-jacents sont examinés brièvement et les performances sont étudiées à l'aide de simulations.

Signification des résultats

Bien que la plupart des études qui se retrouvent dans le documentation portent uniquement sur la performance asymptomatique, c.-à-d. sur un nombre infini d'échantillons de données, les études comportementales relatives aux échantillons de petite taille sont rares. Dans la pratique, des tailles d'échantillon finies sont utilisées pour les systèmes radar opérationnels dans lesquels les coefficients de filtrage doivent s'adapter fréquemment dans la dimension de la distance. Il s'avère que les techniques décrites dans ce rapport produisent des rapports signal sur bruit plus clutter (SNCR) supérieurs pour une complexité de calcul déterminée. L'algorithme le plus efficient du point de vue des calculs réduit de 10 à 1 000 fois la nécessité de calculs à haute vitesse, dans beaucoup de cas pratiques.

Christoph H. Gierull. 2001. Analyse de la performance de techniques spatio-temporelles totalement adaptables (STAP) à haute vitesse pour les échantillons de petite taille. DREO TR 2001-079. Centre pour la Recherche de la Défence Ottawa.

Table of contents

List of figures

Acknowledgements

The work is a result of the German-Canadian cooperation under the auspices of the Memorandum of Understanding between the Governments of Canada and Germany for the Exchange of Information in Defence Science, Annex 2 on Radar Technology. The author would like to thank Dr. Richard Klemm and Frau Gaby Gniss of the German Defence Research Establishment (FGAN/FHR) for providing the simulated space-time covariance matrices.

This page intentionally left blank.

1. Introduction

In military space-based or air-to-ground surveillance it is desirable to be able to detect small moving ground targets such as tanks or wheeled vehicles within severe ground clutter. The clutter echoes received by the aircraft show a motion-induced spread of the Doppler spectrum which particularly masks the slowly moving targets. The problem is to suppress this clutter effectively, i.e. down to the sensor noise power level. One very promising solution is the use of multiple antenna apertures or a single antenna aperture with multiple feed horns. The extra degrees of freedom offered by these structures allows one to exploit the information contained in both the temporal (Doppler) and angular (spatial) domains towards the goal of distinguishing between the clutter and the desired target signals. The different phase centers of the antenna provide the auxiliary spatial dimension which allows the use of digital signal processing to suppress the unwanted clutter energy. Good general introductions of space-time filtering of data from an array of sensors for Moving Target Indication (MTI) are given in [1], [2].

In this report, two different recently developed fast rank reduction techniques, so called subspace or projection methods, are applied to the problem of clutter suppression. They were originally proposed to provide fast man-made jammer suppression in large surface phased array radars with many elements. In this report, the basic theory behind them is reviewed and their performance is studied via simulations. While most investigations in the literature deal only with the asymptotic performance, i.e. an infinite number of data samples, studies of the small sample size behaviour are rare. Finite sample size is the practical case for an operational system for which the filtering coefficients have to be adapted frequently along the range dimension.

In chapter 2 the problem is mathematically formulated and the concept of optimum processing is briefly reviewed. The simulation scenario used for the analysis and comparison of the different methods is described. In chapter 3 the finite sample size processing scheme is introduced. This processing scheme is required to estimate the covariance matrix which is not known in practice. In section 3.2, the two SNCR-optimizing techniques SMI/LSMI are reviewed, followed by the description of the idea behind subspace or projection methods. Since the computational complexity of the best known subspace technique (the eigenvector decomposition) is enormous, a much faster method, called MTP, is proposed in the following section. The shape of the eigenvalue distribution of the covariance matrix is exploited to develop a robust and computationally efficient algorithm. It is based on a weighting of the projector and is called lean matrix inversion (LMI). The computational complexity of the various techniques is compared in chapter 4.

2. Optimum Processing

2.1 Maximization of the SNCR

Let a spatial-time-snapshot $\underline{x}$, hereafter called a snapshot, be modeled as the superposition $\underline{x}_k = \mathbf{D}\underline{b}_k + \underline{n}_k$, where $\underline{n}_k$ is the sensor noise vector with mutually independent components of power σ^2 and the time $k := t_k$. $\underline{b}_k$ is a complex normal distributed random vector with independent entries representing the amplitudes of the clutter at time k. The snapshots are sampled at such intervals that $\underline{x}_k$ and $\underline{x}_{k-1}$ are completely independent (see also [1]). In the case of MTI the snapshots are often gathered from the range bins surrounding a guard bin along the range direction. With these definitions in mind, the received data vectors can be understood as realizations of mutually independent, identically complex normal distributed random vectors $\underline{X}$. The random vectors $\underline{X}$ have mean $\underline{0}$ and the covariance matrix $\mathbf{R} = \mathrm{E}\underline{X}\underline{X}^H = \mathbf{D}\mathbf{A}\mathbf{D}^H + \sigma^2\mathbf{I}$, where E is the expectation operator and $\mathbf{I}$ denotes the identity matrix. Matrices are written bold, vectors are underlined and $(\cdot)^H$ means complex conjugate transpose.

As can be demonstrated using basic detection theory [3], the probability of detection of a signal in noise depends mainly on the SNR. Therefore, one criterion for adjusting the filter vector $\underline{w}$ could be the maximization of the signal-to-noise-plus-clutter ratio (SNCR) of the filter output when the processor is matched to frequency ϑ. The SNCR of this filter output, hereafter denoted as ρ, is given by

$$\rho(\underline{w}) := \rho(\underline{w}(\vartheta)) = |a|^2 \frac{\mathrm{E}|\underline{w}^H \underline{d}(\vartheta)|^2}{\mathrm{E}|\underline{w}^H \underline{X}|^2} = |a|^2 \frac{\underline{w}^H \underline{d}(\vartheta)\underline{d}(\vartheta)^H \underline{w}}{\underline{w}^H \mathbf{R} \underline{w}} \tag{1}$$

where a is the complex amplitude and $\underline{d}(\vartheta)$ a given signal at this frequency. It is easy to verify that the solution of (1) is given by

$$\underline{w}_{\mathrm{opt}} = \alpha \mathbf{R}^{-1} \underline{d}(\vartheta) \tag{2}$$

with

$$\rho_{\mathrm{opt}} = \underline{d}(\vartheta)^H \mathbf{R}^{-1} \underline{d}(\vartheta), \tag{3}$$

where α is an arbitrary complex factor [4, 5]. The optimum processor can be separated into two steps: a decorrelation of the clutter via multiplication with $\mathbf{R}^{-1/2}$ (pre-whitening) and a matched filtering with the adapted signal $\underline{s} = \mathbf{R}^{-1/2}\underline{d}$.

2.2 Simulation Scenario

In order to analyze and compare different suppression techniques, the SNCR can be used as a measure of the performance of the clutter filter. The SNCR is

normalized with respect to the signal-to-noise ratio in the noise-only case:

$$\begin{aligned}\rho(\vartheta) := \rho(\underline{w}(\vartheta)) &= \frac{SNCR(\vartheta)_{\text{adapted, with clutter}}}{SNR(\vartheta)_{\text{without clutter}}} \\ &= \frac{|\underline{w}(\vartheta)^H \underline{d}(\vartheta)|^2}{\underline{w}(\vartheta)^H \mathbf{R}\underline{w}(\vartheta)} \left(\frac{|\underline{d}(\vartheta)^H \underline{d}(\vartheta)|^2}{\underline{d}(\vartheta)^H \sigma^2 \mathbf{I} \underline{d}(\vartheta)} \right)^{-1} \\ &= \frac{\sigma^2}{N} \frac{|\underline{w}(\vartheta)^H \underline{d}(\vartheta)|^2}{\underline{w}(\vartheta)^H \mathbf{R}\underline{w}(\vartheta)}, \end{aligned} \tag{4}$$

i.e. the resultant curve is actually the SNCR-loss with respect to the best possible case of no interference at all.

Fig. 1 shows the normalized SNCR versus the normalized target frequency for the two special cases of, no adaptation[1] with $\underline{w}(\vartheta) = \underline{d}(\vartheta)$ and, optimum processing with $\underline{w}(\vartheta) = \mathbf{R}^{-1}\underline{d}(\vartheta)$. The simulation scenario used in this report was chosen as follows:

- Side-looking geometry
- Number of channels $N = 12$, $\lambda/2$ spacing.
- Number of Pulses $M = 12$, DPCA spacing, i.e. $(v_0/PRF = \lambda/2)$.
- Clutter-to-noise ratio $CNR = 20$dB.

The sensor noise level has been normalized to $\sigma^2 = 0$dB. One can see that almost no loss is suffered by optimum processing even for target frequencies close to the clutter notch. In contrast, the case of no adaptation equals the inverse of the filter response function and has an average loss of about -17dB over the entire frequency range. If the desired target signal perfectly overlays the clutter $\vartheta = 0$, the loss of SNCR is given as $-10 * log10(NM) - CNR \cong -40$dB.

As will be shown later, an important measure of the quality of the clutter suppression is given by the magnitudes and shape of the eigenvalue distribution of the space-time covariance matrix. Fig. 2 shows the corresponding distribution of the simulated covariance matrix. Since the DPCA-criterion is fulfilled in this side-looking case, the number of eigenvalues larger than the sensor noise power level can be determined as $N + M - 1 = 23$ (Brennan's-rule [1]). Only the first 50 eigenvalues are plotted. In order to simplify the notation throughout the report, the order of the space-time covariance matrix will henceforth be denoted as N and the number of dominant eigenvalues (larger than σ^2) as M. This notation corresponds to the jammer suppression case. Accordingly, for STAP, these numbers have to be substituted with NM and $N + M$, respectively.

[1] When no clutter is present and therefore no suppression is needed.

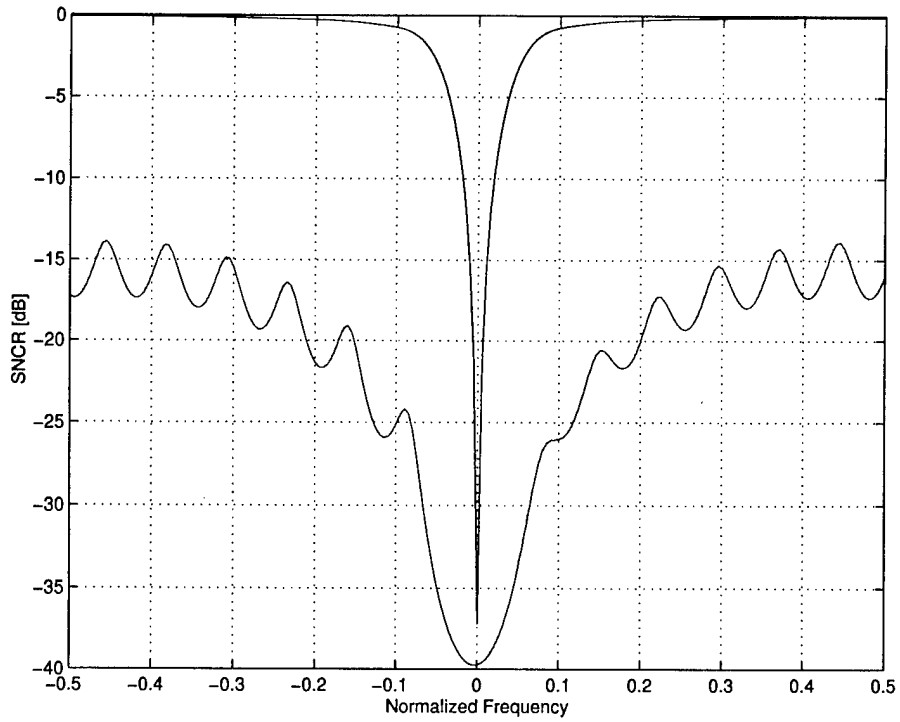

Figure 1 : *Normalized signal-to-clutter plus noise ratio versus normalized frequency for non- and optimum processing.*

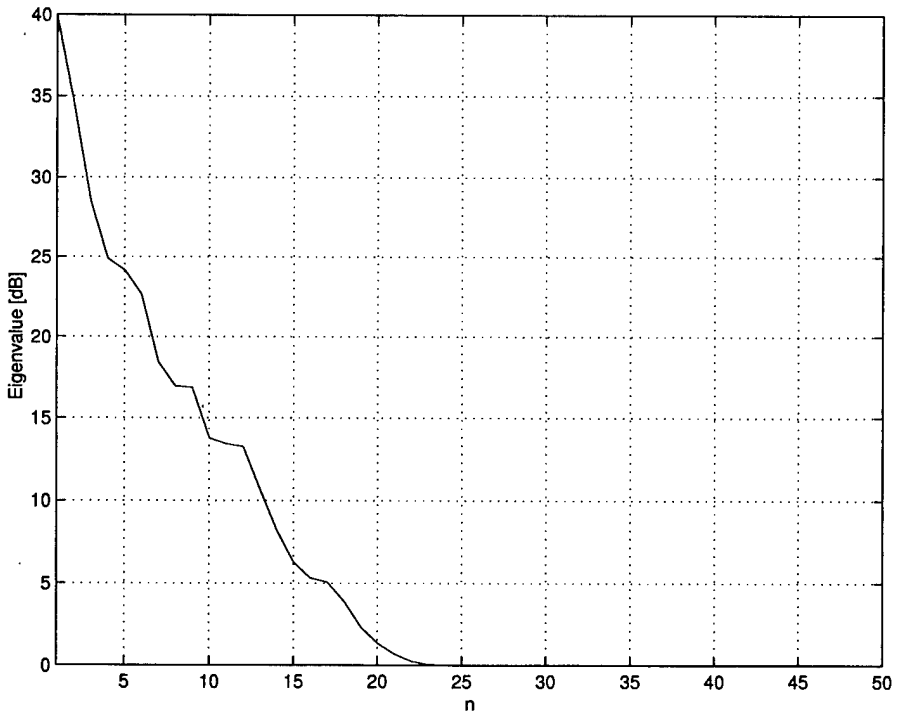

Figure 2 : *Eigenvalue distribution of the simulated space-time covariance matrix.*

3. Finite Sample Size Processing

In practice the exact (asymptotic) space-time covariance matrix is not known and must be estimated from a given number of samples. With the number of snapshots denoted by K, an often used estimation is given by the average of the dyads

$$\hat{\mathbf{R}} = \frac{1}{K} \sum_{k=1}^{K} \underline{X}_k \underline{X}_k^H. \tag{5}$$

Assuming that the vectors $\underline{X}_k$ are identically complex normal-distributed with expectation zero, $\hat{\mathbf{R}}$ is equivalent to the Maximum-Likelihood-estimator of R. That is known to to be complex Wishart-distributed with K-degrees of freedom and parameter matrix $\mathbf{R}$ [6].

A requirement for a large number of samples often proves untenable for practical reasons. For instance, the acquisition time of the samples may be too long or the extra computational effort may be wasteful if an acceptable SNCR could have been reached with fewer samples. For MTI-applications, a large number of samples are often not available since the snapshots are gathered from the range bins surrounding the guard bin along the range direction. Since the clutter is non-stationary due to its range-dependence, only a small number of surrounding range cells might be useful for estimating the clutter covariance matrix of the cell in question.

3.1 Sample Matrix Inversion (SMI)

In an approach analogous to the maximization of the SNCR with known covariance matrix it is possible to determine the filter vector by maximizing the SNCR using only a limited number of samples

$$\frac{|\underline{w}^H \underline{d}|^2}{\underline{w}^H \hat{\mathbf{R}} \underline{w}} = \max. \tag{6}$$

The optimization criterion in (6) is equivalent to the minimization of the clutter power with limited sample number under the constraint of a normalized filter vector in target direction,

$$\underline{w}^H \hat{\mathbf{R}} \underline{w} = \min \quad \text{u. c.} \quad \underline{w}^H \underline{d} = 1. \tag{7}$$

The solution of eq. (7) can be obtained as a multiple of the product of the inverse sample covariance matrix and the signal vector

$$\hat{\underline{w}} = \alpha \hat{\mathbf{R}}^{-1} \underline{d}. \tag{8}$$

This approach is well known under the name Sample Matrix Inversion SMI [7, 5, 8]. To assure the nonsingularity of the space-time covariance matrix, it is necessary that the number of snapshots at least matches the order of the matrix $K \geq N$. Substituting the adapted filter (8) into eq. (4) yields for the SNCR

$$\rho(\hat{\underline{w}}) = \frac{|\underline{d}(\vartheta)^H \hat{\mathbf{R}}^{-1} \underline{d}(\vartheta)|^2}{\underline{d}(\vartheta)^H \hat{\mathbf{R}}^{-1} \mathbf{R} \hat{\mathbf{R}}^{-1} \underline{d}(\vartheta)} \tag{9}$$

where the sensor noise level has been set to 0dB. In Fig. 3, 10 realizations of the SNCR are plotted for SMI. The sample space-time covariance matrix was estimated with $K = 150$ snapshots and shows a relatively large loss (≈ -18 dB) over the whole frequency range.

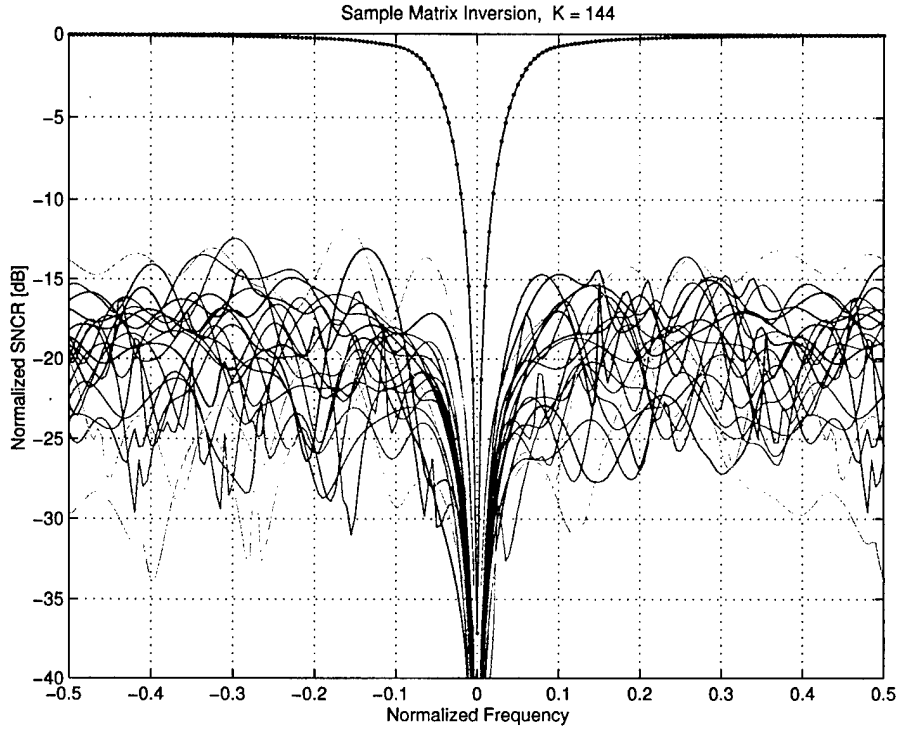

Figure 3 : *Normalized signal-to-clutter plus noise ratio versus normalized frequency for SMI.*

Here, $\rho(\hat{\underline{w}})$ is also a random variable depending on the statistics of $\hat{\mathbf{R}}^{-1}$. For the related application of man made jammer suppression, theoretical performance analysis has been done [5]. To study the performance of the SMI, Reed et. al. [5] calculated the pdf of the normalized version

$$\rho(\hat{\underline{w}}) = \rho(\hat{\underline{w}}) / \underline{d}(\vartheta)^H \mathbf{R}^{-1} \underline{d}(\vartheta).$$

The classical result they found was that the normalized SNCR for SMI is Beta-distributed with the parameters $(2(N-1), 2(K-N+2))$ and an expectation

$$E\rho = (K-N+2)/(K+1).$$

One of the main disadvantages of the SMI is the large number of samples required to adapt the weight vector and suppress the interference effectively. For example, the average number of snapshots to reach the 3 dB loss of normalized SNCR $E\rho = 0.5$ is given by $K_{3dB} = 2N - 3 \cong 2N$ (see also [7]).

3.2 Loaded Sample Matrix Inversion (LSMI)

3.2.1 Motivation

In order to improve the slow convergence of the SMI shown in Fig. 3 one can try to introduce an additional constraint to the optimization task in eq. (7). For instance, $\underline{w}^H\underline{w} = \text{const.}$ can be used; this also solves another disadvantage of the SMI, namely strong fluctuating sidelobes of the filter output. Fig. 4 shows several realizations of the adapted filter output for the SMI

$$a(\vartheta) = |\underline{w}(\vartheta_0)^H\underline{d}(\vartheta)|^2 = |\underline{d}(\vartheta_0)^H\hat{\mathbf{R}}^{-1}\underline{d}(\vartheta)|^2. \qquad (10)$$

The filter is matched to a desired signal at $\vartheta = 0.25$ which means the target is located completely outside of the clutter band. The bold curve corresponds to the case where the covariance matrix is perfectly known. Obviously, the required clutter notch does exist but the side lobes of the filter output are now strongly fluctuating and greatly increased over the entire frequency range. Low side lobes of the filter response are crucial if the probability of false alarm detections not to increase. The criterion for the minimization of the clutter power under the constraints of a normalized filter on one hand, and a preserved constant average sidelobe level of the adapted filter response on the other hand, yields

$$p(\underline{w}) = \underline{w}^H\hat{\mathbf{R}}\underline{w} = \min \quad \text{u. c.} \quad \underline{w}^H\underline{d} = 1, \quad \int_\Upsilon |\underline{w}^H\underline{d}(\kappa)|^2 d\kappa = c. \quad (11)$$

The integral in eq. (11)

$$\int_\Upsilon |\underline{w}^H\underline{d}(\kappa)|^2 d\kappa = \underline{w}^H \left(\int_\Upsilon \underline{d}(\kappa)\underline{d}(\kappa)^H d\kappa \right) \underline{w} = \underline{w}^H\mathbf{C}\underline{w}$$

is performed over the surface of the unity-sphere. For the case of plane-antennae with equidistant spacing of $\frac{\lambda}{2}$, it has been shown that $\mathbf{C}$ converges quickly towards the identity matrix. For a linear-antenna with equidistant spatial spacing $\frac{\lambda}{2}$ the identity $\mathbf{C} = \mathbf{I}$ holds exactly [9]. Using this simplification and omitting the indices, ϑ, in the following for simplification, eq. (11) can be expressed as

$$p(\underline{w}) = \underline{w}^H\hat{\mathbf{R}}\underline{w} = \min \quad \text{u. c.} \quad \underline{w}^H\underline{d} = 1, \quad \underline{w}^H\underline{w} = c \qquad (12)$$

or as the Lagrange-function

$$p(\underline{w},\mu,\gamma) = \underline{w}^H \hat{\mathbf{R}} \underline{w} + \mu(\underline{w}^H \underline{d} - 1) + \gamma(\underline{w}^H \underline{w} - c),$$

respectively.

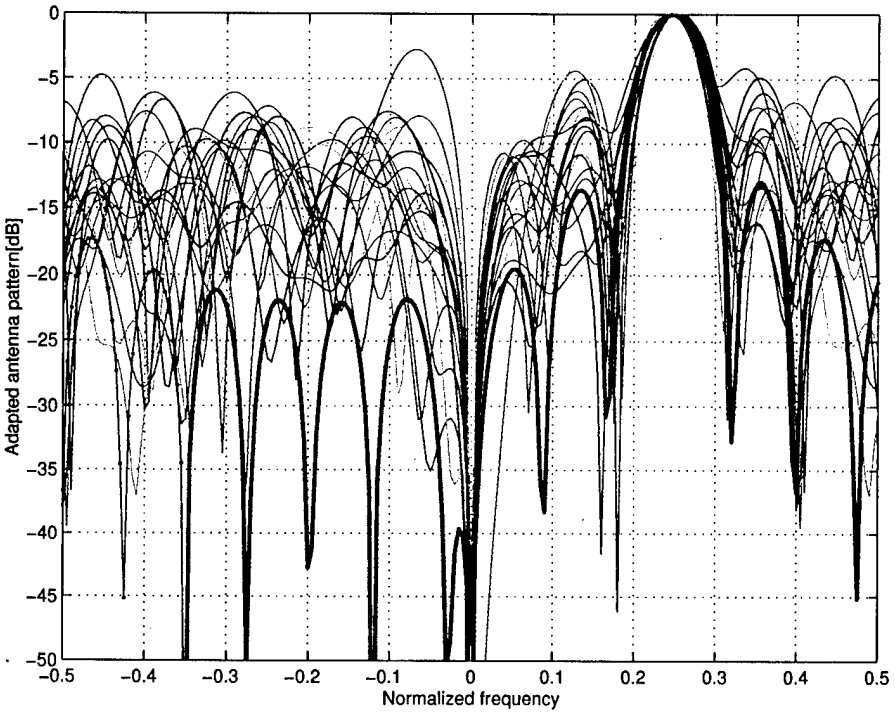

Figure 4 : *Adapted antenna pattern with SMI, $K = 150$.*

The solution results in

$$\hat{\underline{w}} = \alpha \left(\hat{\mathbf{R}} + \gamma \mathbf{I}\right)^{-1} \underline{d}, \tag{13}$$

where γ must be determined under the second constraint in eq. (12). Unfortunately, no simple analytic relation between γ and c is known. Therefore, the additive loading of the diagonal of the sample covariance matrix with a suitable chosen constant δ, i.e. $\hat{\underline{w}} = \alpha \left(\hat{\mathbf{R}} + \delta \mathbf{I}\right)^{-1} \underline{d}$, will be taken as the solution [10, 8, 11] . A definition of a reasonable range of values is given in [12, 13].

Fig. 5 shows the SNCR for the filter vector calculated with LSMI-technique. The number of snapshots was chosen to be $K = M = 24$ and the loading factor is set to three times the sensor noise power level, i.e. $\delta = 3\sigma^2$.

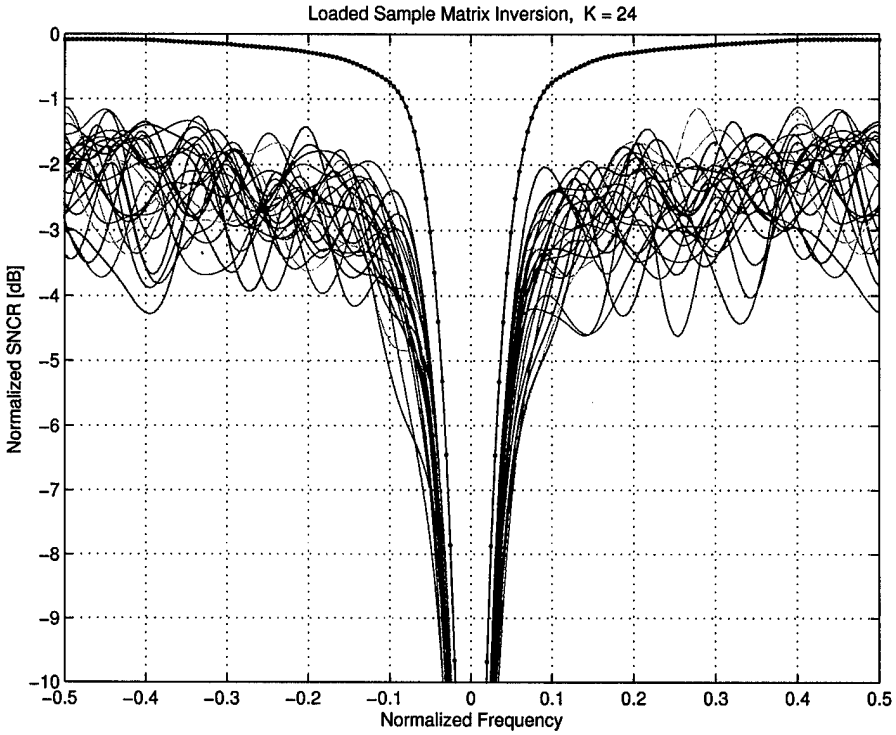

Figure 5 : *Signal-to-clutter plus noise ratio versus normalized frequency for LSMI.*

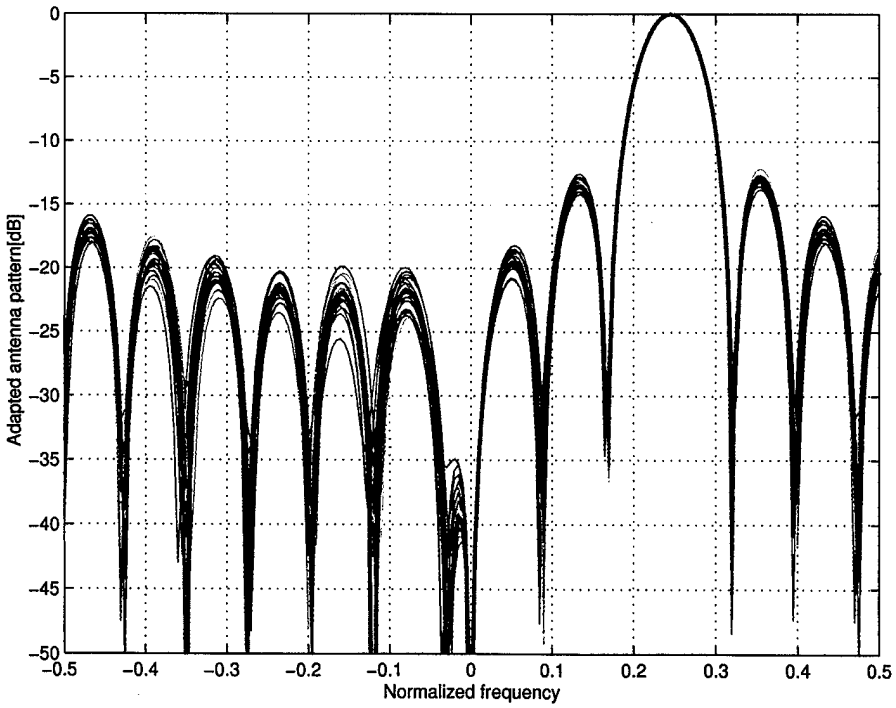

Figure 6 : *Adapted antenna pattern for LSMI,* $K = 150$, $\delta = 3$.

A significant performance improvement over that of the original SMI can be recognized over the entire frequency range. The approximate SNCR loss is only -2.5 dB, although the number of samples chosen was six times less than for Fig. 3. The improvement of the SNCR is linked to the desired reduction of fluctuation in the average sidelobe level of the adapted matched filter response, as shown in Fig. 6, [14].

3.2.2 Consequences

To study the impact of the diagonal loading on the adapted matched filter response, i.e. the gain of SNCR, let the sample covariance matrix $\hat{\mathbf{R}}$ be decomposed into its real eigenvalues λ_n (with the smallest eigenvalue $\lambda_{\min} > 0$) and corresponding orthonormal eigenvectors $\underline{u}_n$ $(n = 1,\ldots,N)$:

$$\hat{\mathbf{R}} = \sum_{n=1}^{N} \lambda_n \underline{u}_n \underline{u}_n^H \quad \text{with} \quad \sum_{n=1}^{N} \underline{u}_n \underline{u}_n^H = \mathbf{I}.$$

The inverse can be written as

$$\begin{aligned} \hat{\mathbf{R}}^{-1} &= \sum_{n=1}^{N} \frac{1}{\lambda_n} \underline{u}_n \underline{u}_n^H = \frac{1}{\lambda_{\min}} \left(\mathbf{I} - \sum_{n=1}^{N} \underline{u}_n \underline{u}_n^H \right) + \sum_{n=1}^{N} \frac{1}{\lambda_n} \underline{u}_n \underline{u}_n^H \\ &= \frac{1}{\lambda_{\min}} \left(\mathbf{I} - \sum_{n=1}^{N} \left(\frac{\lambda_n - \lambda_{\min}}{\lambda_n} \right) \underline{u}_n \underline{u}_n^H \right). \end{aligned} \tag{14}$$

Substituting eq. (14) into eq. (10) yields the adapted filter response without loading:

$$a(\vartheta) = \left| \underline{d}(\vartheta_0)^H \underline{d}(\vartheta) - \sum_{n=1}^{N} \left(\frac{\lambda_n - \lambda_{\min}}{\lambda_n} \right) \underline{d}(\vartheta_0)^H \underline{u}_n \underline{u}_n^H \underline{d}(\vartheta) \right|^2 \frac{1}{\lambda_{\min}^2}.$$

For a given target signal orthogonal to the clutter subspace, with $\underline{d}(\vartheta) \perp \{\underline{u}_1, \ldots, \underline{u}_M\}$, i. e. in the sidelobe regions of the clutter, it holds that:

$$a(\vartheta) = \left| \underline{d}(\vartheta_0)^H \underline{d}(\vartheta) - \sum_{n=M+1}^{N} \left(\frac{\lambda_n - \lambda_{\min}}{\lambda_n} \right) \underline{d}(\vartheta_0)^H \underline{u}_n \underline{u}_n^H \underline{d}(\vartheta) \right|^2 \frac{1}{\lambda_{\min}^2}.$$

When the noise eigenvalues are of different magnitude, and in particular when the sample sizes are small, the error term $\left(1 - \frac{\lambda_{\min}}{\lambda_n}\right)$ for $n \in \{M+1,\ldots,N\}$ can be large. On the basis of this notation, the effect

of the diagonal loading, i.e. $\tilde{\lambda}_n = \lambda_n + \delta$, becomes clear:

$$\tilde{a}(\vartheta) = \left| \underline{d}(\vartheta_0)^H \underline{d}(\vartheta) - \sum_{n=M+1}^{N} \left(\frac{(\lambda_n + \delta) - (\lambda_{\min} + \delta)}{\lambda_n + \delta} \right) \cdot \underline{d}(\vartheta_0)^H \underline{u}_n \underline{u}_n^H \underline{d}(\vartheta) \right|^2 \frac{1}{(\lambda_{\min} + \delta)^2}.$$

The error term after diagonal loading

$$\varepsilon = 1 - \frac{\lambda_{\min} + \delta}{\lambda_n + \delta} = \frac{\lambda_n - \lambda_{\min}}{\lambda_n + \delta}$$

becomes smaller due to the compression of the noise eigenvalues, if δ is larger than the largest noise eigenvalue λ_{M+1}. On the other hand, to preserve the clutter notch, δ must be chosen to be much smaller than the smallest dominant clutter eigenvalue, i.e. $\delta \ll \lambda_M$. For the case of adaptive jammer suppression, a less known analytical expression for the pdf of the normalized SNCR resulting from LSMI can be found in [12]. Herein it is shown that if $K > M$, and some reasonable conditions are imposed on the diagonal load γ [2], the normalized SNCR is also Beta-distributed but with parameters $(2(K-M+1), 2M)$

$$f_\rho(\rho) = \frac{\Gamma(K+1)}{\Gamma(K-M+1)\Gamma(M)} \rho^{K-M} (1-\rho)^M \quad 0 < \rho \leq 1 \qquad (15)$$

and expectation

$$E\rho = (K-M+1)/(K+1). \qquad (16)$$

It can be seen that the expected 3 dB loss of SNCR is now reached for only $K_{3dB} = 2M - 1$ snapshots, a number remarkably lower than that required with SMI. Therefore, the statistical properties are independent of the number of channels, and the number of snapshots needed to suppress the interference effectively is now proportional to M (see also [15, 13]).

To study the statistical properties of the improvement factor resulting from the application of clutter suppression in more detail, the empirical probability density function has been computed at one particular frequency. In Fig. 7 the histogram of the normalized SNCR generated from 100000 trials is plotted along with the theoretical pdf in eq. 15 for $M = 12$. The discrepancy between the two curves is due to the different shapes of the eigenvalue distributions. While the eigenvalues in

[2] $\sigma^2 < \gamma \ll \lambda_M$, where λ_M denotes the smallest interference eigenvalue

the man-made jammer case resemble step-functions (sharp edge), the eigenvalues in Fig. 2 decrease gradually down to the level of the sensor noise. This fact is discussed in more detail in the following section. It is noteworthy to point out that the variance (variation) of the SNCR for clutter suppression is remarkably smaller than in the case of jammer suppression.

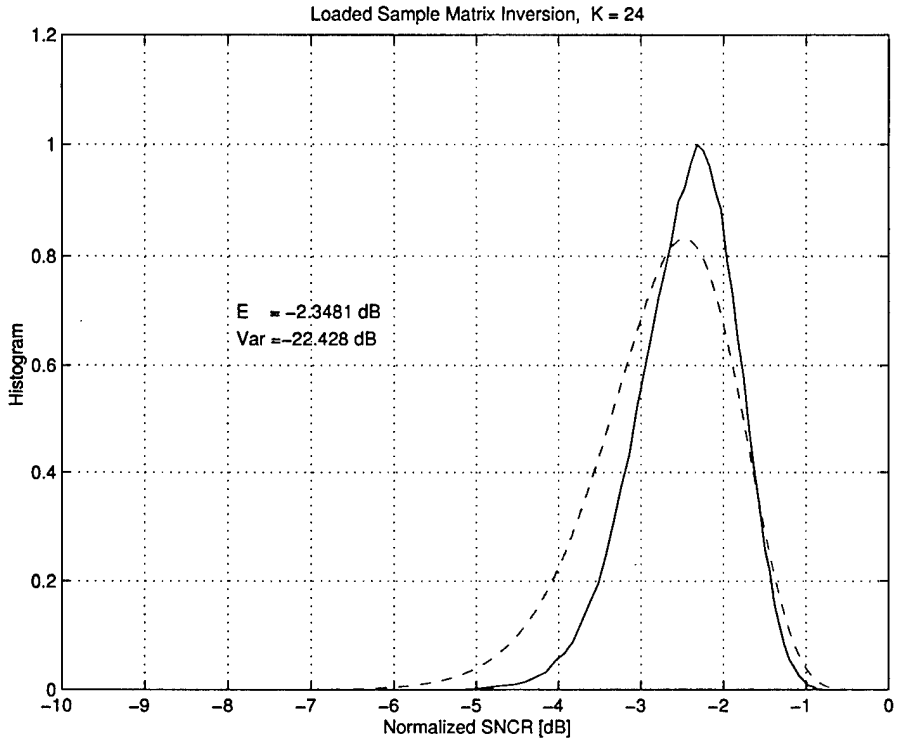

Figure 7 : *Histogram of SNCR for LSMI at normalized frequency 0.25, calculated over* 100000 *trials (dashed). Theoretical pdf (15) for* $M = 12$ *(solid).*

3.3 Subspace Techniques

Using the low rank approximation of the space-time covariance matrix, subspace techniques can be exploited for complete rejection of the clutter in terms of independency of clutter power. The basic idea of the so called projection methods is the separation of the overall space of observation into an interference and a noise part, followed by a projection of the array outputs into the clutter-free subspace to remove the interference.

The adaptive spatial interference cancellation problem is in principle closely related to that of superresolution angle estimation. The application of a projection into the space complementary to the interference or jammer subspace, often

called the (sensor) noise subspace, generates deep notches in the antenna pattern in the directions of the jammers. The detection of these notches leads directly to estimations of the unknown locations. Thus, all subspace techniques can be used for both problems, although their statistical properties can, of course, be very different. An overview of subspace methods concerning both applications can be found in [16, 17].

Starting from the model in chapter 2, a clutter data vector can be written as

$$\underline{x} = \mathbf{D}\underline{b} + \underline{n}$$

the superposition of the phase vectors $\mathbf{D} = [\underline{d}_1, \ldots, \underline{d}_M]$ weighted with the clutter amplitudes $\underline{b}$, and the sensor noise $\underline{n}$. Via a projection orthogonal to $\mathbf{D}$

$$\tilde{\underline{x}} = \mathbf{P}_{\mathbf{D}}^{\perp}\underline{x},$$

(i. e. the M-dimensional subspace of the $\mathbb{C}^N$ which is spanned by the column vectors $\underline{d}_1, \ldots, \underline{d}_{N+M}$) it is possible to suppress the clutter exactly down to zero. Then, the filtering

$$y = \underline{d}^H \tilde{\underline{x}} = \underline{d}^H \mathbf{P}_{\mathbf{D}}^{\perp}\underline{x}$$

gives the filter vector as

$$\underline{w} = \mathbf{P}_{\mathbf{D}}^{\perp}\underline{d}.$$

For the special case of an indefinite clutter-to-noise ratio CNR this orthogonal projection is equivalent to the SNCR optimization $\underline{w} = \mathbf{R}^{-1}\underline{d}$. Using the matrix inversion lemma[3], the inverse of the space-time covariance matrix $\mathbf{R} = \mathbf{D}\mathbf{B}\mathbf{D}^H + \sigma^2\mathbf{I}$ can be written as

$$\mathbf{R}^{-1} = \frac{1}{\sigma^2}\left(\mathbf{I} - \mathbf{D}\left(\mathbf{D}^H\mathbf{D} + \tilde{\mathbf{B}}^{-1}\right)^{-1}\mathbf{D}^H\right),$$

where $\tilde{\mathbf{B}} = \mathbf{B}/\sigma^2$. Assuming CNR $\to \infty$ it follows $\tilde{\mathbf{B}}^{-1} \to 0$, so that

$$\sigma^2\mathbf{R}^{-1} \to \mathbf{P}_{\mathbf{D}}^{\perp} = \mathbf{I} - \mathbf{D}\left(\mathbf{D}^H\mathbf{D}\right)^{-1}\mathbf{D}^H. \tag{17}$$

Since the phase matrix $\mathbf{D}$ is generally unknown, the projection matrix has to be estimated. This can, for instance, be achieved by a subspace fitting of the random data vectors. A M-dimensional subspace $\mathbf{S} \in \mathbb{C}^{N\times M}$ of the $\mathbb{C}^N$ must be found so as to best approximate the data vectors in a Mean Square Error MSE sense:

$$\mathrm{E}||\underline{X} - \mathbf{P}_{\mathbf{S}}\underline{X}||^2 = \min. \tag{18}$$

[3] $\left(\mathbf{A} + \mathbf{B}\mathbf{C}\mathbf{B}^H\right)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\left(\mathbf{B}^H\mathbf{A}^{-1}\mathbf{B} + \mathbf{C}^{-1}\right)^{-1}\mathbf{B}^H\mathbf{A}^{-1}$

$\mathbf{P_S}$ denotes a projection matrix into the subspace which is spanned by the columns of $\mathbf{S} = [\underline{S}_1, \ldots, \underline{S}_M]$. The minimization in eq. (18) is equivalent to the maximization of

$$\mathrm{E}||\mathbf{P_S}\underline{X}||^2 = \mathrm{E}\underline{X}^H\mathbf{P_S}\underline{X} = \mathrm{tr}\{\mathbf{P_S}\mathrm{E}\underline{X}\underline{X}^H\} = \mathrm{tr}\{\mathbf{P_S}\mathbf{R}\} = \max.$$

If the orthonormality of the vectors $\underline{S}_1, \ldots, \underline{S}_M$ is also demanded, the projection can be written as $\mathbf{P_S} = \mathbf{S}\mathbf{S}^H$. Hence, the maximization turns out to be an optimization under constraint:

$$\mathrm{tr}\{\mathbf{S}^H\mathbf{R}\mathbf{S}\} = \max \quad \text{u. c.} \quad \mathbf{S}^H\mathbf{S} = \mathbf{I}. \tag{19}$$

Substituting the eigenvalue decomposition $\mathbf{R} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$ into eq. (19) results in

$$\mathrm{tr}\{\mathbf{S}^H\mathbf{U}\mathbf{\Lambda}\mathbf{U}^H\mathbf{S}\} = \mathrm{tr}\{\mathbf{Q}^H\mathbf{\Lambda}\mathbf{Q}\} = \max \quad \text{u. c.} \quad \mathbf{Q}^H\mathbf{Q} = \mathbf{I}. \tag{20}$$

The solution of eq. (20) is shown in the appendix to be the subspace $\mathbf{Q}$ which is spanned by the eigenvectors corresponding to the M dominant eigenvalues, or an arbitrary unitary transformation of it. Denoting the eigenvectors corresponding to the M dominant eigenvalues as $\mathbf{S}$, the projection matrix can be written as

$$\mathbf{P}_{\mathbf{S}}^{\perp} = \mathbf{I} - \mathbf{S}\mathbf{S}^H$$

and the filter vector as

$$\underline{w} = \alpha\mathbf{P}_{\mathbf{S}}^{\perp}\underline{d} = \alpha\left(\underline{d} - \mathbf{S}\mathbf{S}^H\underline{d}\right). \tag{21}$$

Unlike the LSMI, the dimension of the clutter subspace M needs to be known in order to calculate $\mathbf{P_S}$ in eq. (21); it has to be estimated as well. Several different approaches exist in literature. The extended sphericity-test [18], tests the equality of the last $N - M$ smallest eigenvalues of the sample covariance matrix. The, so-called, information theory based algorithms, such as AIC- [19] and MDL-criterion [20] are different approximations of the density function of the extended sphericity-test statistic. The white noise test for active systems [21] is based mainly on the comparison of the estimated noise power level with a-priori known level. The analysis of a suitable dimension of the clutter subspace is a major task of this report.

3.3.1 Eigenvector Projection (EVP)

Since the asymptotic covariance matrix $\mathbf{R}$ is not known in practice, it is possible to use the eigenvectors of the sample covariance matrix to determine the filter vector $\underline{w}$

$$\hat{\mathbf{R}} = \hat{\mathbf{U}}\mathbf{\Lambda}\hat{\mathbf{U}}^H = \hat{\mathbf{S}}\mathbf{\Lambda}_S\hat{\mathbf{S}}^H + \hat{\mathbf{N}}\hat{\mathbf{N}}^H.$$

Using eq. (5), one can show that the desired subspace $\mathbf{S} \in \mathbb{C}^N$ is the one which best approximates the sample vectors $\underline{x}_k$ $(k = 1, \ldots, K)$ in the meaning of least square error (compare also eqs. (18) to (20)):

$$\sum_{k=1}^{K} ||\underline{x}_k - \mathbf{P}_{\mathbf{S}} \underline{x}_k||^2 = \min$$

or

$$\sum_{k=1}^{K} ||\mathbf{P}_{\mathbf{S}} \underline{x}_k||^2 = \operatorname{tr}\left\{ \mathbf{S}^H \sum_{k=1}^{K} \underline{x}_k \underline{x}_k^H \mathbf{S} \right\} = K \operatorname{tr}\left\{ \mathbf{S}^H \hat{\mathbf{R}} \mathbf{S} \right\} = \max$$
$$\text{u. c.} \quad \mathbf{S}^H \mathbf{S} = \mathbf{I}_M, \tag{22}$$

respectively. This technique is completely analogous to the MUltiple SIgnla Classification MUSIC algorithm for high resolution spectral estimation [22]. It has been shown analytically that the suppression performance of the so called EigenVector Projection (EVP) is equivalent to that of LSMI [23]. To be consistent, the eigenvectors of the sample covariance matrix are denoted by $\hat{\mathbf{S}}$ and the corresponding estimated weight vector by $\underline{\hat{w}}(\vartheta) = \left(\mathbf{I} - \hat{\mathbf{S}}\hat{\mathbf{S}}^H\right) \underline{d}(\vartheta)$. Inserting this weight into eq. (1), we get for the SNCR

$$\rho(\hat{\mathbf{S}}, \vartheta) = \frac{|\underline{d}(\vartheta)^H \mathbf{P}_{\hat{\mathbf{S}}}^{\perp} \underline{d}(\vartheta)|^2}{\underline{d}(\vartheta)^H \mathbf{P}_{\hat{\mathbf{S}}}^{\perp} \mathbf{R} \mathbf{P}_{\hat{\mathbf{S}}}^{\perp} \underline{d}(\vartheta)} \frac{\sigma^2}{N}. \tag{23}$$

The behaviour of ρ is shown in Fig. 8 and Fig. 10 for a clutter subspace dimension of 12 and 24, respectively. One can see that the average loss of SNCR in both cases is about (-2) to (-3) dB. To evaluate the statistical properties of the SNCR in more detail, identical histograms to those used with LSMI have been calculated and displayed in Fig. 9 and 11. Two remarkable results are observed.

In regards to asymptotic losses, it can be seen that even a small dimension of the clutter subspace results in near-optimum suppression. The main difference is some minor ripples; the reason for these ripples lies in the special shape of the eigenvalue distribution. The sum of all eigenvalues is identical to the total clutter power. After removing the power contained in the first 12 eigenvalues the power in the remaining eigenvalues is only a small fraction of the entire clutter power, even if they are still larger than the sensor noise floor. Further increasing the subspace dimension to 24 (Fig. 10) eliminates the remaining power and causes the ripples to disappear.

In contrast, when a limited number of samples is used to estimate the covariance matrix, surprising behaviour is observed, namely the smaller subspace dimension produces slightly better results, i.e. the loss appears to be slightly less. This impression is confirmed by the two histograms in Fig. 9 and 11, where the demonstrated expectation is slightly smaller. The reason for this behaviour is not clear but it is believed that both the magnitude of the eigenvalues and the orientation of the corresponding clutter and noise subspaces with respect to each other, play an important role (compare also section 3.3.4).

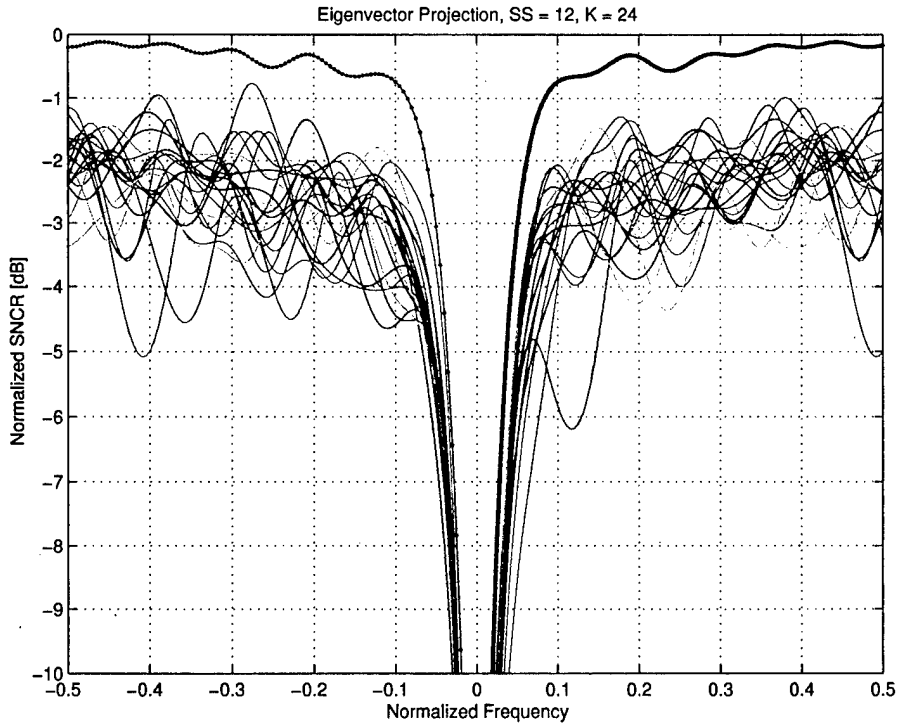

Figure 8 : *Signal-to-clutter plus noise ratio versus normalized frequency for EVP, dimension of subspace 12.*

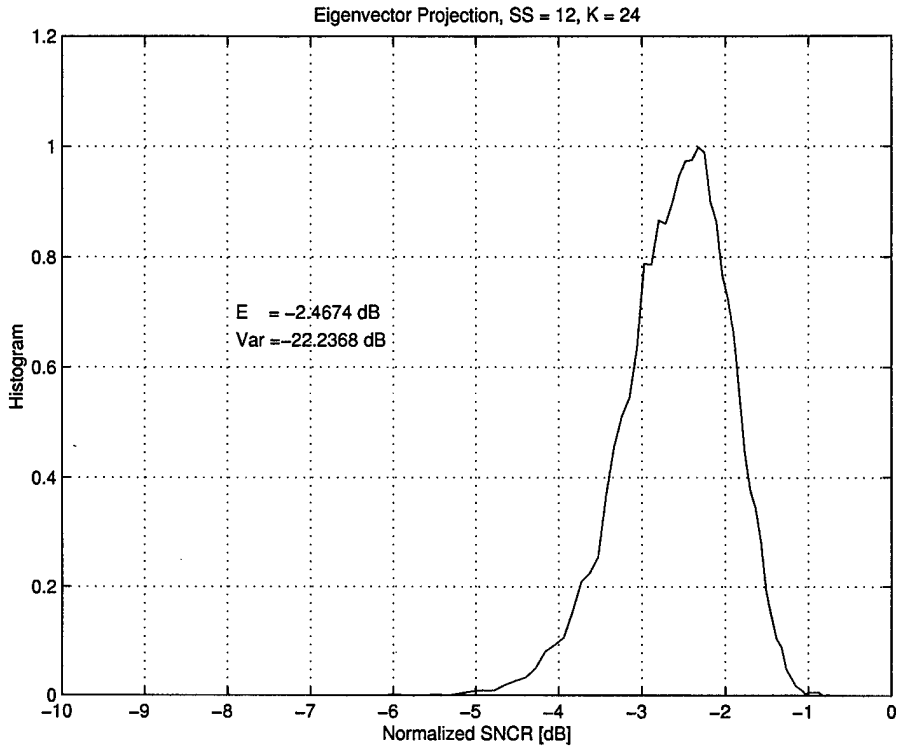

Figure 9 : *Histogram of SNCR for EVP at normalized frequency 0.25 calculated over* 10000 *trials, dimension of subspace 12.*

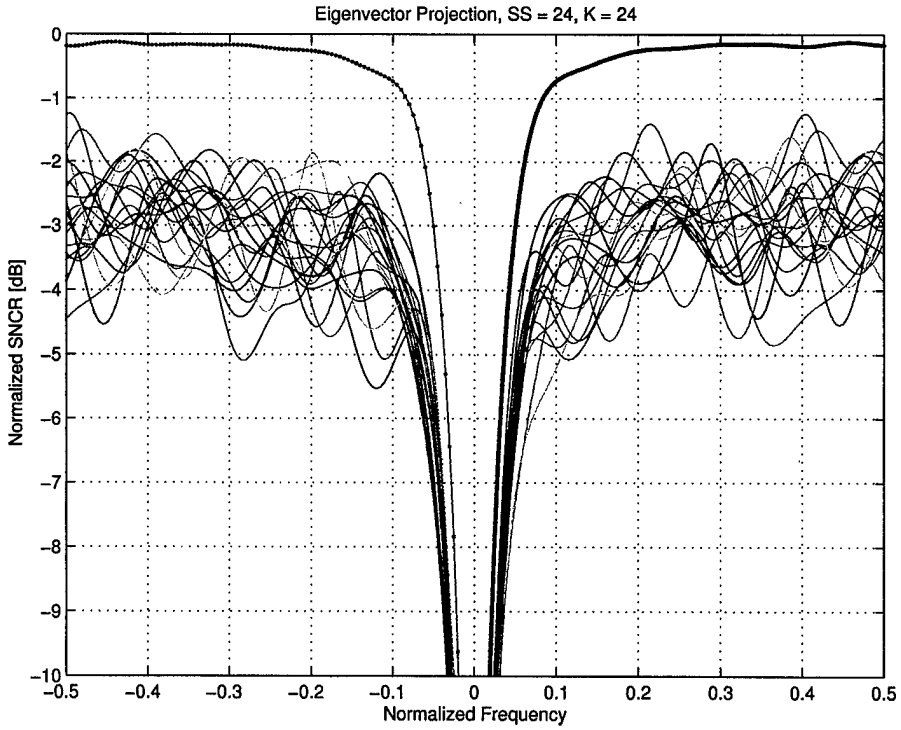

Figure 10 : *Signal-to-clutter plus noise ratio versus normalized frequency for EVP, dimension of subspace 24.*

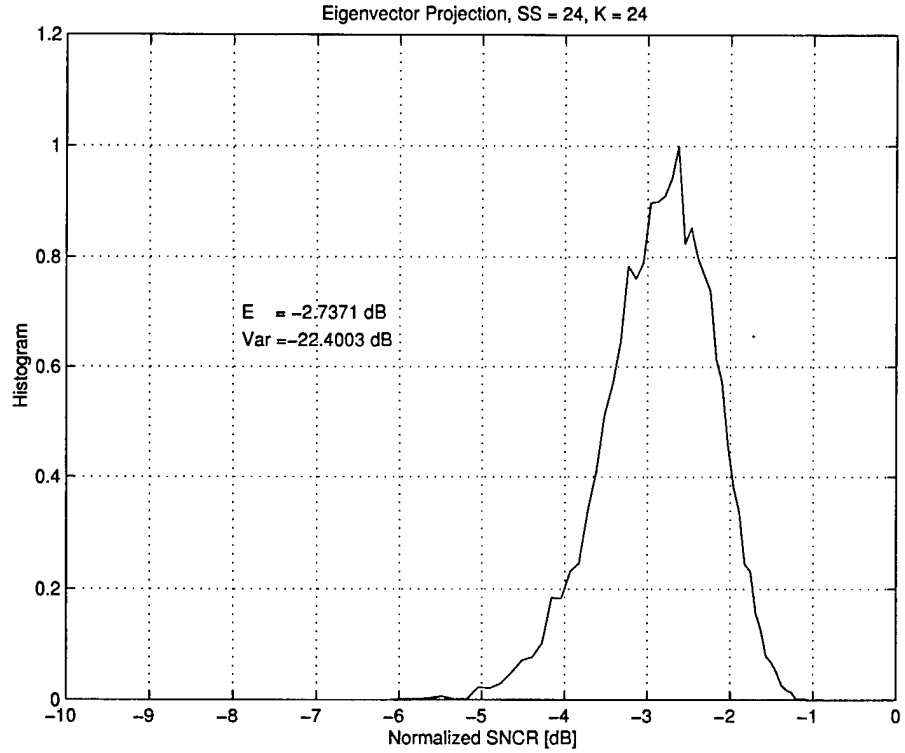

Figure 11 : *Histogram of SNCR for EVP at normalized frequency 0.25 calculated over* 10000 *trials, dimension of subspace 24.*

3.3.2 Projections without Eigenanalysis

The estimation of the complete covariance matrix and the following eigendecomposition are very intensive and time consuming computational processes. Many different techniques have been proposed in the literature, with the goal of reducing the computational complexity of subspace methods for use in real time applications in large antenna arrays, see e.g. [24, 25, 26]. These iterative methods, which employ exponential forgetting factor to partially update the eigenvectors, have the drawback of possessing a lower convergence rate than MUSIC, i.e. they need a large number of samples to perform sufficiently well. In contrast, subspace methods which work with successive short time frames (i.e. on a small block of data) show better tracking performance [17].

Since every arbitrary complete basis of the clutter subspace can be used to form the projection matrix, it is possible to develop faster subspace methods which are not based on eigenanalysis. In principle, the simplest way to estimate the subspace, as suggested by Hung and Turner [27], is to take the data vectors directly as a basis for the clutter subspace. This method, called HTP, gives only a rough estimate of the subspace and yields sufficient results only for high CNR. The theoretical statistical

analysis of the HTP, when used for man-made jammer suppression, can be found in [13].

The noneigenvector estimation of a basis of the signal subspace based on the sample space-time covariance matrix, can be conceptualized as an extension of the HTP to the case of an infinite amount of samples. One class of subspace estimation methods, called Yeh-Brandwood- Projections (YBP), is based on the idea of a sufficient data reduction preprocessing, namely the estimation of a basis of the clutter subspace based only on a part of the sample covariance matrix. Since the measured covariance matrix always contains receiver noise, the estimated subspace is disturbed.

Different techniques have been proposed in the literature to reduce the negative influence of the sensor noise e.g. [28, 29]. In [30] it was shown that the influence of the additive noise can be interpreted as a leakage of the signal power into the noise subspace. Instead of recollecting the lost jammer power, it is desirable to pre-apply an adaptive mechanism which protects the method before or during execution of the subspace estimation. A mathematical technique to achieve this property is known as the rank-revealing QR-factorization [31] of the sample covariance matrix, also called QR-decomposition with permutation (PQR) [32]. This technique was first used in this context by Hsieh et. al. [33] and Prasad et.al [34].

To avoid the time consuming recursive permutation, Reilly [35] suggests a fixed pre-selection of M columns of the covariance matrix using some a-priori knowledge about their structure. He called this method reshuffled QR decomposition, or RQR. In [36], a simple, efficient extension of the RQR, called Matrix Transformation based Projection (MTP), has been presented. It is based on a generalised representation of the class of YBP. The performance of the MTP, when applied to the closely related problem of superresolution angle estimation with small sample size, has been investigated in [37].

3.3.3 Projections Based on Covariance Matrix Transformation (MTP)

This section deals with the performance of MTP when applied to the suppression of clutter for airborne MTI. Let us first consider a general dimension reducing transformation of the covariance matrix

$$\tilde{\mathbf{R}} = \mathbf{RT} = \left(\mathbf{DAD}^H + \sigma^2\mathbf{I}\right)\mathbf{T}, \tag{24}$$

where $\mathbf{T}$ is an arbitrary transformation matrix of the $\mathbb{C}^{N\times M}$ with $\text{rank}\{\mathbf{T}\} = M$. It can be directly seen from eq. (24) that a substitution of $\mathbf{T}$ by the first M columns of the identity matrix is identical with the classical YBP. Moreover, if we replace the transformation by the permutation or preselection matrices mentioned in the previous section, the PQR and RQR methods are also included as special cases of this general expression. Using $\hat{\tilde{\mathbf{R}}} = \mathbf{X}\mathbf{X}^H\mathbf{T}$ to build up the projection matrix

$$\mathbf{P}_{\hat{\tilde{\mathbf{R}}}}^{\perp} = \left(\mathbf{I} - \hat{\tilde{\mathbf{R}}}\left(\hat{\tilde{\mathbf{R}}}^H\hat{\tilde{\mathbf{R}}}\right)^{-1}\hat{\tilde{\mathbf{R}}}^H\right)$$

and inserting the resulting weight vector $\hat{\underline{w}}(\vartheta) = \mathbf{P}_{\hat{\tilde{\mathbf{R}}}}^{\perp}\underline{d}(\vartheta)$ into eq. (1), we get for the SNCR

$$\rho(\hat{\tilde{\mathbf{R}}}, \vartheta) = \frac{|\underline{d}(\vartheta)^H\mathbf{P}_{\hat{\tilde{\mathbf{R}}}}^{\perp}\underline{d}(\vartheta)|^2}{\underline{d}(\vartheta)^H\mathbf{P}_{\hat{\tilde{\mathbf{R}}}}^{\perp}\mathbf{R}\mathbf{P}_{\hat{\tilde{\mathbf{R}}}}^{\perp}\underline{d}(\vartheta)}\,\frac{\sigma^2}{N}. \tag{25}$$

3.3.3.1 *Optimum Transformation for Known Covariance Matrices*

Eq. (24) raises the question of whether a constant (non adaptive) transformation $\mathbf{T}$ exists for which an efficient subspace estimation can be achieved, assuming the covariance structure, including the sensor noise term, is given. That means, one is looking for a transformation of the covariance matrix for which $\text{lin}\{\mathbf{RT}\} = \text{lin}\{\mathbf{D}\}$, i.e. $\mathbf{RT}$ must be a linear combination of the phase matrix $\tilde{\mathbf{R}} = \mathbf{RT} := \mathbf{DG}$ which is not at all affected by the sensor noise, with $\mathbf{G} \in \mathbb{C}^{M\times M}$ and $\text{rank}\{\mathbf{G}\} = M$. It is easy to verify that an optimum transformation in this sense is given by an arbitrary linear combination of the phase matrix itself: $\mathbf{T} = \mathbf{DF}$ with $\mathbf{F} \in \mathbb{C}^{M\times M}$ and $\text{rank}\{\mathbf{F}\} = M$. Since

$$\tilde{\mathbf{R}} = \mathbf{RT} = \mathbf{RDF} = \mathbf{DAD}^H\mathbf{DF} + \sigma^2\mathbf{DF} = \mathbf{DG},$$

where

$$\mathbf{G} = \mathbf{AD}^H\mathbf{DF} + \sigma^2\mathbf{F} \in \mathbb{C}^{M\times M}.$$

$\tilde{\mathbf{R}}$ represents a basis of the interference subspace which is not interrupted by the noise even in the case of fully correlated interference sources, i.e. when $\text{rank}\{\mathbf{A}\} < M$.

3.3.3.2 *Suboptimum Transformation for Estimated Covariance Matrices*

As the phase matrix of the clutter is supposed to be unknown, one might follow the considerations of the HTP and take M snapshots vectors $\mathbf{X}_0 = [\underline{x}_1, \ldots, \underline{x}_M]$ as a rough estimation of the interference subspace for the transformation $\hat{\tilde{\mathbf{R}}} = \hat{\mathbf{R}}\mathbf{T} = \mathbf{X}\mathbf{X}^H\mathbf{X}_0$. The result of this approach shows an unexpected effect. Although the transformation matrix $\mathbf{T} = \mathbf{X}_0$, i.e. the M snapshots are highly fluctuating and therefore the spanned subspace $\text{lin}\{\mathbf{X}_0\}$ strongly deviates from the exact one[4] $\text{lin}\{\mathbf{D}\}$, it seems that the subspace $\text{lin}\{\hat{\tilde{\mathbf{R}}}\}$ (spanned by $\tilde{\mathbf{R}} = \mathbf{D}\mathbf{A}\mathbf{D}^H\mathbf{X}_0 + \sigma^2\mathbf{X}_0$) has effectively no loss of SNCR compared to $\text{lin}\{\mathbf{D}\}$. To explain this phenomenon, we will establish a relationship between the MTP and the Power-method [32] for recursive evaluation of the eigenvectors corresponding to the M dominant eigenvalues:

$$\begin{aligned} \tilde{\mathbf{Y}}_l &= \mathbf{R}\mathbf{Y}_{l-1} = \mathbf{R}^l\mathbf{Y}_0 \\ \mathbf{Y}_l &= \text{orth}\{\tilde{\mathbf{Y}}_l\} \qquad l = 1, \ldots, L, \end{aligned} \tag{26}$$

where $\text{orth}\{\Phi\}$ denotes the orthonormalization of the columns included in Φ.

One can see in eq. (26), that depending on the initial matrix $\mathbf{Y}_0$ used in the first step of iteration $L = 1$, either a YBP with $\mathbf{Y}_0 = \tilde{\mathbf{I}}$, a PQR or a RQR will be performed. To analyse the convergence of the algorithm, we will replace the covariance matrix in (26) by its eigen-factorization. Further, let the initial matrix $\mathbf{Y}_0$ be an arbitrary linear combination of the eigenvectors

$$\mathbf{Y}_0 = \mathbf{U}\mathbf{A} = \mathbf{S}\mathbf{A}_S + \mathbf{N}\mathbf{A}_N$$

with $\mathbf{A} = \left[\mathbf{A}_S^T, \mathbf{A}_N^T\right]^T$, $\mathbf{A}_S \in \mathbb{C}^{M \times M}$ and $\mathbf{A}_N \in \mathbb{C}^{(N-M) \times M}$. Assuming linear independence of the eigenvectors and the columns of $\mathbf{Y}_0$, it directly follows that the matrix of coefficients $\mathbf{A} = \mathbf{U}^H\mathbf{Y}_0$ has the $\text{rank}\{\mathbf{A}\} = M$ and, particularly, that $\text{rank}\{\mathbf{A}_S\} = \text{rank}\{\mathbf{S}^H\mathbf{Y}_0\} = M$. Inserting these terms into the iteration (26) yields

$$\tilde{\mathbf{Y}}_l = \mathbf{R}^l\mathbf{Y}_0 = \mathbf{S}\Lambda_S^l\mathbf{A}_S + \mathbf{N}\Lambda_N^l\mathbf{A}_N$$

[4] Compare, for instance, Fig. 1 in [13]

and due to the nonsingularity of $\mathbf{A}_S$

$$\begin{aligned}\tilde{\mathbf{Y}}_l &= \left[\mathbf{S} + \mathbf{N}\boldsymbol{\Lambda}_N^l \mathbf{A}_N \mathbf{A}_S^{-1} \boldsymbol{\Lambda}_S^{-l}\right] \boldsymbol{\Lambda}_S^l \mathbf{A}_S \\ &= \left[\mathbf{S} + O\left(q^l\right)\right] \boldsymbol{\Lambda}_S^l \mathbf{A}_S \qquad l = 1, \ldots, L \end{aligned} \tag{27}$$

with

$$\begin{aligned} q = \max\{\lambda_{M+1}/\lambda_1, \ldots, \lambda_{M+1}/\lambda_M, \ldots, \lambda_N/\lambda_1, \\ \ldots, \lambda_N/\lambda_M\} = \lambda_{M+1}/\lambda_M. \end{aligned}$$

The convergence of the iteration is ensured by the successive orthonormalization of $\tilde{\mathbf{Y}}_l$. If we take into account that the smallest clutter eigenvalue λ_M is proportional to the average of the interference power of the clutter, and that λ_{M+1} is approximately equal to the sensor noise power σ^2, then the order of the error quotient q^L is proportional to $(1/\mathrm{CNR})^L$. This can be negligible even for $L = 1$, particularly in case of strong clutter. Thus, it can be concluded that the performance of the MTP and YBP depend not only on the CNR, but also on the condition of the coefficient matrix $\mathbf{A}_S$. This statement can easily be confirmed if a pure noise matrix for transformation is used. In that case, the coefficient matrix $\mathbf{A}_S = \mathbf{S}^H\mathbf{T}$ is given as a random linear combination of the clutter eigenvectors, which is generally well conditioned. Although this kind of transformation has no relation to the interference subspace, the given results shown no difference (see also [15]). Under MTP, the number of complex operations to estimate the interference subspace $\hat{\tilde{\mathbf{R}}} = \hat{\mathbf{R}}\mathbf{T} = \mathbf{X}\left(\mathbf{X}^H\mathbf{T}\right)$ is of the order $O(2KN)$. This is double that of the YBP or RQR-method which achieve greater computational efficiency through the special structure of their transformations, e.g. using the identity or any other permutation matrix for transformation requires no explicit matrix multiplication rather only a selection of corresponding columns.

3.3.3.3 *Suboptimum Transformation with Small Computational Complexity*

In this section, we use a physical interpretation of the different special cases of the YBP introduced above, to develop a transformation matrix with very small computational effort. In

other words, we are looking for a transformation matrix which is sparse but still guarantees a good condition of $\mathbf{A}_S = \mathbf{S}^H\mathbf{T}$.

The estimation of a part of the sample space-time covariance matrix corresponds to the average of the correlation of a M-element subarray output with the entire array. For example, a subarray built with the first M elements (YBP) shows a wide beamwidth due to its strongly decreased aperture. A pair of closely located clutter points which are still conventionally resolvable by the entire aperture $\Delta u \geq$ beamwidth $\sim 1/N$, now appear so close together that the corresponding subarray phase matrix $\mathbf{D}^H\mathbf{T}$ in eq. (24) will be almost singular. It now becomes clear that a fixed pre-selection scheme, like that suggested for the RQR [35], means a regular thinned array which still preserves the entire aperture. Hence, some grating lobes will appear in the subarray pattern although the condition number of the subarray phase matrix is strongly reduced. The effect of this undersampling can be interpreted as a penetration of nonsuppressable clutter power into the system through these grating lobes. From the theory of subarray design [38, 39], it is known that the use of a random subarray configuration can reduce these periodicities due to the multiplicative superposition of the ambiguities.

For a sufficiently large array, it is possible to find a transformation by randomly re-sorting the columns of the identity matrix. For example $\mathbf{E} \in \{0,1\}^{N\times N}$ is the randomly generated permutation matrix which can be divided into nonoverlapping block matrices $\mathbf{E} = [\mathbf{E}_1, \mathbf{E}_2, \ldots]$ with $\mathbf{E}_i \in \{0,1\}^{N\times M}$ and

$$\mathbf{E}_i^H\mathbf{E}_j = \begin{cases} \mathbf{I}_M & i = j \\ \mathbf{0} & \text{else.} \end{cases}$$

The transformation matrix might then be determined via a sum of a small number of these blocks $\mathbf{T} = \sum_{i=1}^{I} \mathbf{E}_i$. Figs. 12 and 14 show the resulting SNCR of the MTP for the case in which the transformation was generated as a sum of three blocks,i.e. $\mathbf{T} = \mathbf{E}_1 + \mathbf{E}_2 + \mathbf{E}_3$. It can be seen that the performance of MTP is slightly worse than EVP for the asymptotic as well as the finite sample size case. Moreover almost no difference can be recognized for different dimensions of the clutter subspace, Figs. 13 and 15. The number of operations

to estimate the interference subspace

$$\hat{\hat{\mathbf{R}}} = \mathbf{X}\left(\mathbf{X}^H\mathbf{E}_1 + \mathbf{X}^H\mathbf{E}_2 + \mathbf{X}^H\mathbf{E}_3\right)$$

is now NKM complex multiplications, therefore requiring only $2KM$ more complex additions than the YBP or RQR methods.

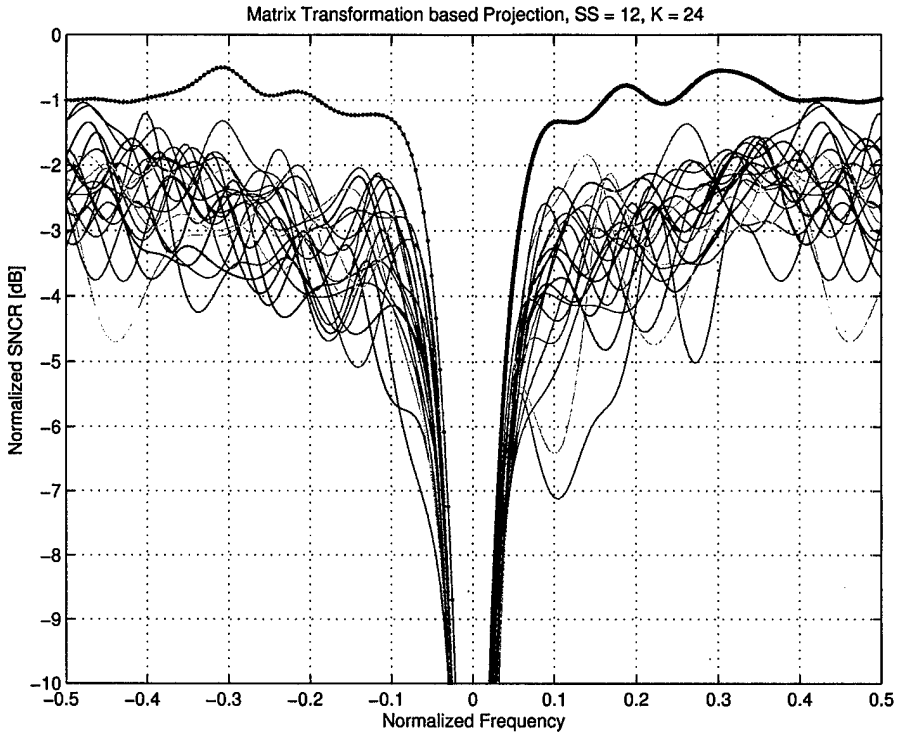

Figure 12 : *Signal-to-clutter plus noise ratio versus normalized frequency for MTP, dimension of subspace 12.*

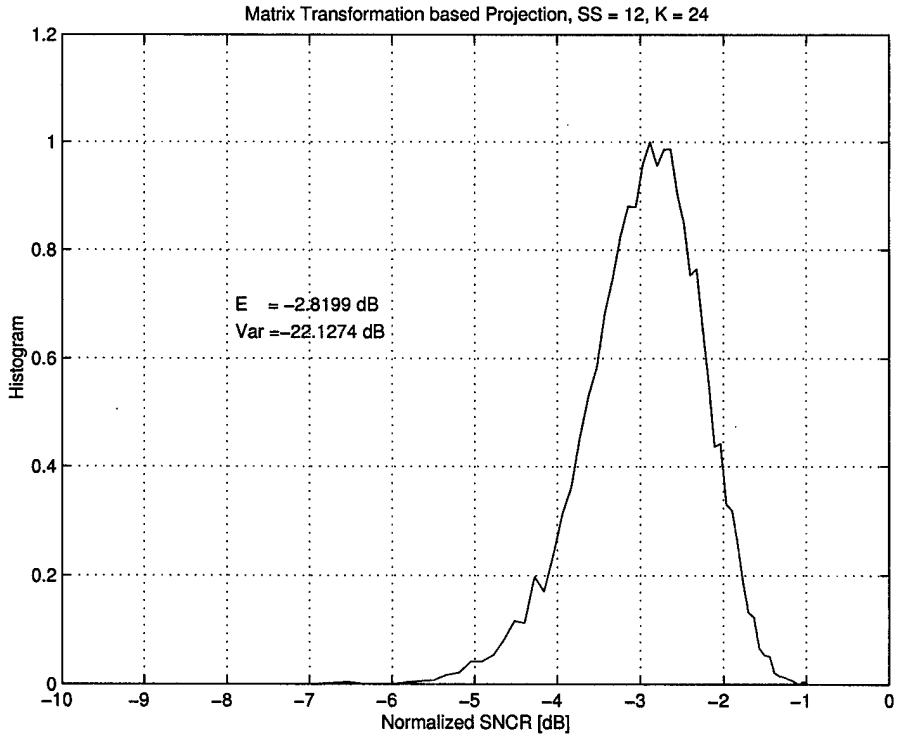

Figure 13 : *Histogram of SNCR for MTP at normalized frequency 0.25, calculated over* 10000 *trials, dimension of subspace 12.*

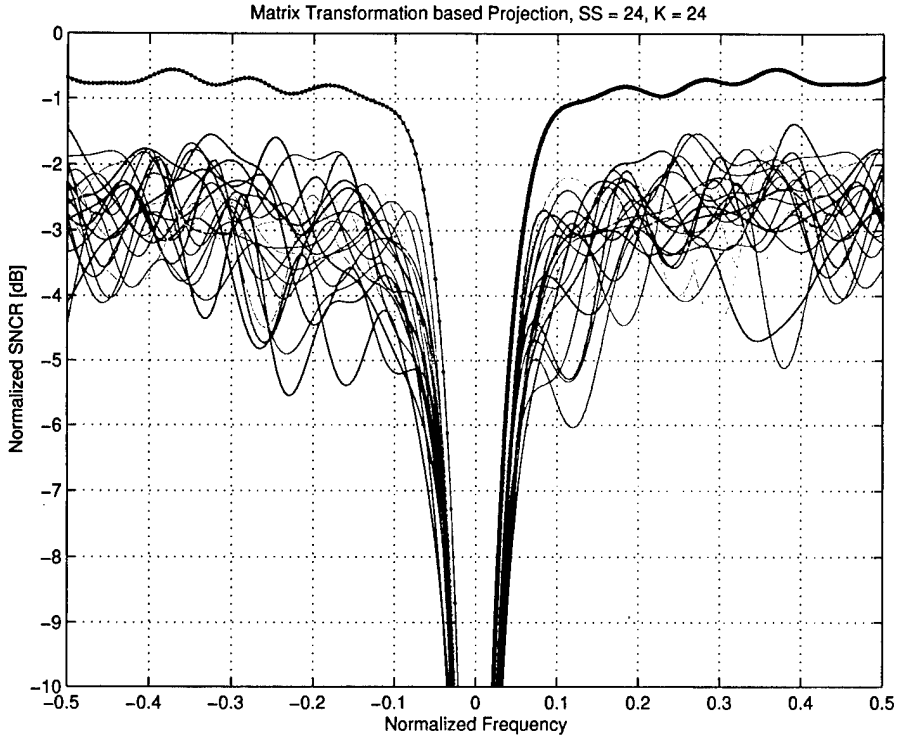

Figure 14 : *Signal-to-clutter plus noise ratio versus normalized frequency for MTP, dimension of subspace 24.*

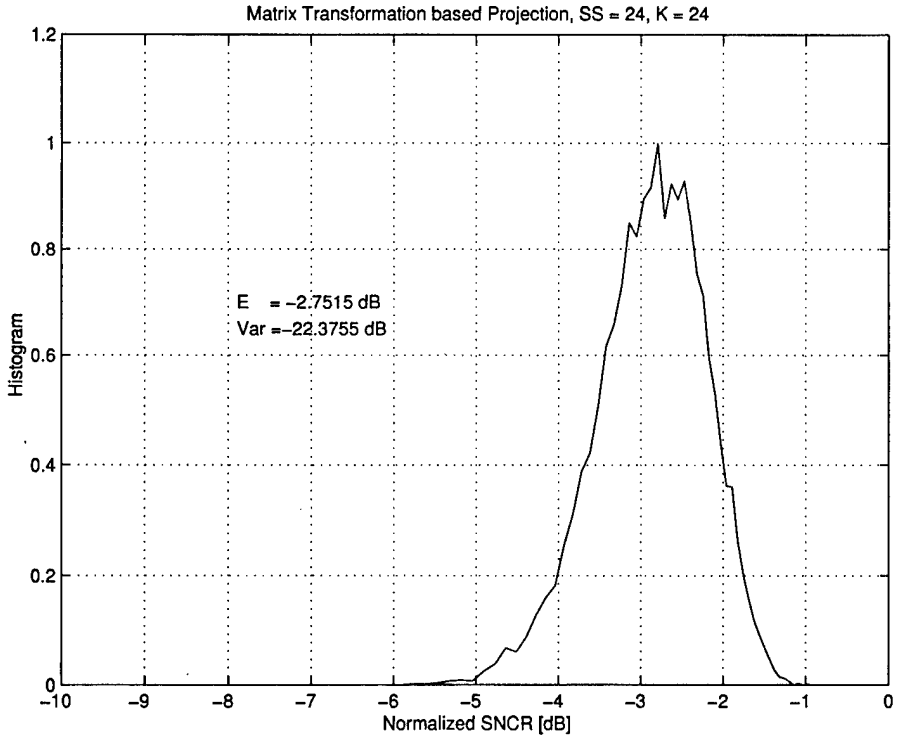

Figure 15 : *Histogram of SNCR for MTP at normalized frequency 0.25, calculated over* 10000 *trials, dimension of subspace 24.*

3.3.4 Weighted Projection - Lean Matrix Inversion (LMI)

This section describes a simple yet computationally highly efficient approach for adaptive spatial interference suppression which was originally developed as a robust method for use in the presence of dispersive channel errors [40]. The method, called lean matrix inversion (LMI), is based on a weighted projection. The corresponding interference subspace is estimated by the MTP-technique.

It is known that the dispersive channel errors included in data measured with an array of sensors will cause a decorrelation between the channels which leads to a loss of the low-rank property of the corresponding covariance matrix [41]. The resulting spread of the eigenvalue distribution can be interpreted as a leakage of interference power from the jammer subspace into the noise sub-space. The clear separation of the two subspaces disappears as a result of this leakage. Since clutter eigenvalues have a similar distribution, the application of this approach to the MTI-problem appear to hold promise.

The problem of choosing the correct dimension of the jammer subspace occurs as a result of the eigenvalue spread [41] (see also section 3.3.1). As mentioned in section 3.3.1, it is obvious that the clutter power contained in the remaining eigenvalues will not be suppressed when the dimension of the clutter subspace is chosen to be $M = 12$. As a result, a SNCR-loss occurs over the entire frequency range. It has been noted, however, that the loss

$$\rho = \frac{\underline{d}^H \mathbf{R}^{-1} \underline{d}}{N} \tag{28}$$

is strongly dependent on the clutter-to-noise ratio and probably on the orientation of the array, i.e. side-looking versus forward-looking system. The latter is the subject of future work. Let the inverse of the space-time covariance matrix in eq. (28) be substituted by its eigen-factorization $\mathbf{R}^{-1} = \mathbf{U}\Lambda^{-1}\mathbf{U}^H$, where the unitary eigenvector matrix is separated into one part $\mathbf{S} \in \mathbb{C}^{N \times M}$ corresponding to the eigenvalues $\Lambda_\mathbf{S} \in \mathbb{R}^{M \times M}$ and a part $\mathbf{N} \in \mathbb{C}^{N \times (N-M)}$ corresponding to the rest $\Lambda_\mathbf{N} \in \mathbb{R}^{(N-M)\times(N-M)}$. Then, the SNCR in the sidelobe region of the clutter is given as

$$\rho = \frac{\underline{d}^H \left(\mathbf{S}\Lambda_\mathbf{S}^{-1}\mathbf{S}^H + \mathbf{N}\Lambda_\mathbf{N}^{-1}\mathbf{N}^H \right) \underline{d}}{N} \cong \frac{\underline{d}^H \left(\mathbf{N}\Lambda_\mathbf{N}^{-1}\mathbf{N}^H \right) \underline{d}}{N}. \tag{29}$$

For $\Lambda_\mathbf{N}$ close to the identity matrix, the SNCR would be $\rho \cong 1$ outside the main clutter with $\mathbf{S}^H \underline{d} \cong 0$. With the eigenvalue distribution shown

in Fig. 2 it is evident that strong clutter with large additional eigenvalues (i.e. far beyond the noise level) results in an additional loss to the SNCR. Or, differently formulated, relatively weak CNR produces eigenvalues which are at or below the noise floor level.

The subspace dimension could be increased to ensure that all dominant eigenvalues are covered, however, this can lead to a loss in antenna gain as the remaining degrees of freedom in the array may be too small to allow detection of the moving targets. In other words, a larger clutter subspace dimension results in an unwanted suppression of target power itself. Hence one may conclude that no subspace dimension can be found for which projection methods work as well as those methods based on the Wiener solution.

The loaded sample matrix from eq. 13 can be approximated as

$$\begin{aligned}\hat{\mathbf{R}}+\gamma\mathbf{I} &= \hat{\mathbf{S}}\left(\boldsymbol{\Lambda}_{\mathbf{S}}+\gamma\mathbf{I}\right)\hat{\mathbf{S}}^{H}+\hat{\mathbf{N}}\left(\boldsymbol{\Lambda}_{\mathbf{N}}+\gamma\mathbf{I}\right)\hat{\mathbf{N}}^{H}\\ &\cong \hat{\mathbf{S}}\left(\boldsymbol{\Lambda}_{\mathbf{S}}+\gamma\mathbf{I}\right)\hat{\mathbf{S}}^{H}+\gamma\hat{\mathbf{N}}\hat{\mathbf{N}}^{H}\\ &= \hat{\mathbf{S}}\left(\boldsymbol{\Lambda}_{\mathbf{S}}+\gamma\mathbf{I}\right)\hat{\mathbf{S}}^{H}+\gamma\left(\mathbf{I}-\hat{\mathbf{S}}\hat{\mathbf{S}}^{H}\right)\\ &= \hat{\mathbf{S}}\boldsymbol{\Lambda}_{\mathbf{S}}\hat{\mathbf{S}}^{H}+\gamma\mathbf{I}. \end{aligned} \tag{30}$$

Using the matrix inversion lemma, and the fact that the eigenvectors are mutually orthonormal, the inverse of the loaded space-time covariance matrix can be written as

$$\left(\hat{\mathbf{R}}+\gamma\mathbf{I}\right)^{-1} \cong \mathbf{I}-\hat{\mathbf{S}}\mathbf{G}\hat{\mathbf{S}}^{H}. \tag{31}$$

The entries of the diagonal matrix $\mathbf{G} = \mathrm{diag}\{G_{11},\ldots,G_{MM}\} \in \mathbb{R}^{M\times M}$ are given as

$$G_{mm} = \lambda_m/(\lambda_m+\gamma) \quad \text{for} \quad m = 1,\ldots,M.$$

Eq. (31) shows that the LSMI approximates a "weighted" projection, wherein the "importance" of each subspace base vector (eigenvector) is determined by the magnitude of its corresponding eigenvalue. This approximation motivates a modification of the classical projection methods in order to make them more robust against an eigenvalue spread.

As shown in section 3.3.3 the QR-decomposition of the estimated subspace

$$\hat{\tilde{\mathbf{R}}} = \mathbf{X}\left(\mathbf{X}^{H}\mathbf{E}_1+\mathbf{X}^{H}\mathbf{E}_2+\mathbf{X}^{H}\mathbf{E}_3\right) = \mathbf{Q}\mathbf{L} = \mathbf{Q}[\mathbf{Q}_{\mathbf{S}}\mathbf{Q}_{\mathbf{N}}] \tag{32}$$

can be used to both, orthonormalize the subspace base vectors and to get an approximation of the jammer eigenvalues needed for the weighting, i.e. the diagonal elements $L_{1,1},\cdots,L_{M,M}$. Fig. 16, for instance, shows the distribution of the magnitudes of the diagonal elements of $\mathbf{L}$ in eq. (32). These magnitudes are a measure of the mutual linear independence of the columns in $\hat{\hat{\mathbf{R}}}$. The number of dominant diagonal values is the same as the eigenvalues, i.e. 24. The special shape reflects the inherent structure of the asymptotic covariance matrix and might be useful in the determination of the clutter subspace.

The weighting matrix can be determined as $\mathrm{diag}(G_{mm} = L_m/(L_m+\gamma))$. Only those diagonal elements of the QR-decomposition that are larger than a threshold α are used. The threshold is the criterion to determine the dimension of the clutter subspace. The value can be chosen in identical fashion to LSMI as $\alpha \cong 3\sigma^2$, where σ^2 describes the receiver noise power. This method was termed lean matrix inversion (LMI) method. To what extent the LMI is sensitive to the choice of the threshold α should be examined in future work.

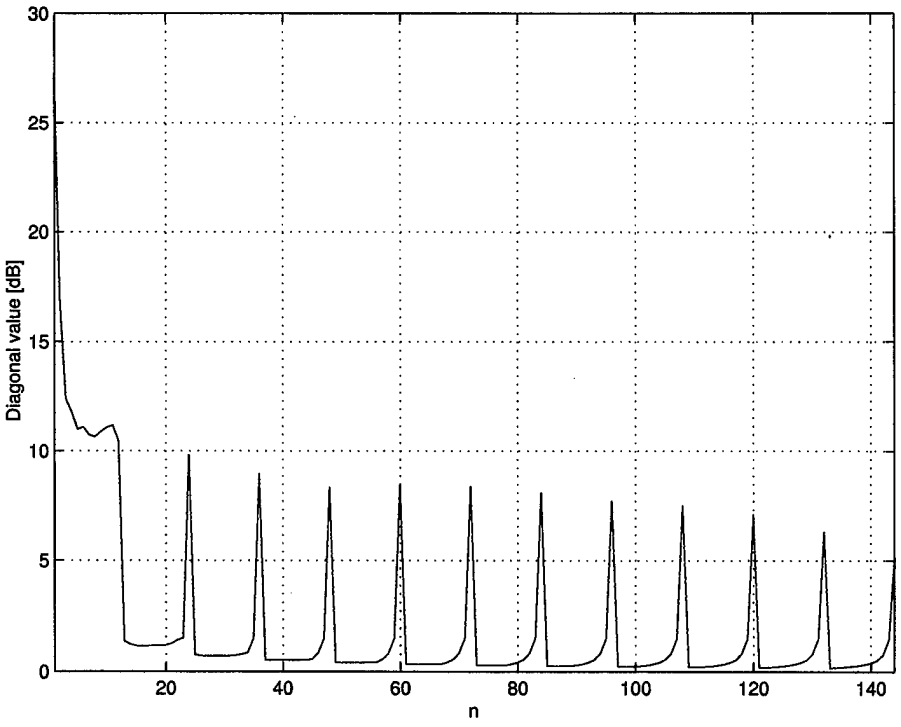

Figure 16 : *Magnitude of the diagonal elements after QR-decomposition of the space-time covariance matrix.*

Figs. 17 to 20 show the loss of SNCR for the two different clutter sub-

space dimensions, 12 and 24. It can be seen that a dimension of 12 is not large enough to show the impact of the weighting. The results in that case are similar to the MTP; increasing the subspace dimension to 24 allows the utilization of the information contained in the tail and shows superior performance compared to all other techniques. It should be pointed out that the SNCR-loss in that case is even less than with LSMI. This is remarkable since the performance of LSMI and EVP were believed to represent the optimum for the application of interference suppression. To verify that this effect is not due to variations in the estimation of ρ, the number of trials in Figs. 18 and 20 was increased to 100000, thereby obtaining a better estimate and a smoother histogram.

To investigate the performance of the LMI on the form of the transformation matrix, the three block matrices, $\mathbf{T} = \sum_{i=1}^{3} \mathbf{E}_i$, have been randomly generated in Fig. 21 for one fixed set of estimated data vectors (snapshots). For this particular set of snapshots, with an SNCR-loss of -1.57dB, the variance due to the random transformation matrix is about 10dB smaller than the fluctuation of the data itself. This result shows that the suppression performance of LMI and the MTP are not very sensitive on the choice of $\mathbf{T}$.

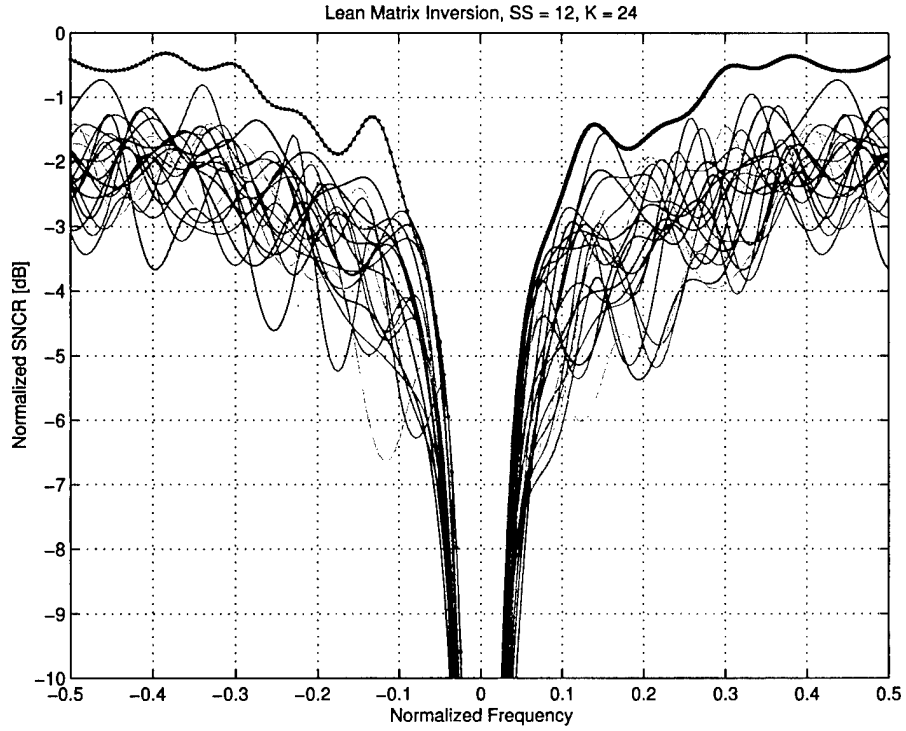

Figure 17 : *Signal-to-clutter plus noise ratio versus normalized frequency for LMI, dimension of subspace 12.*

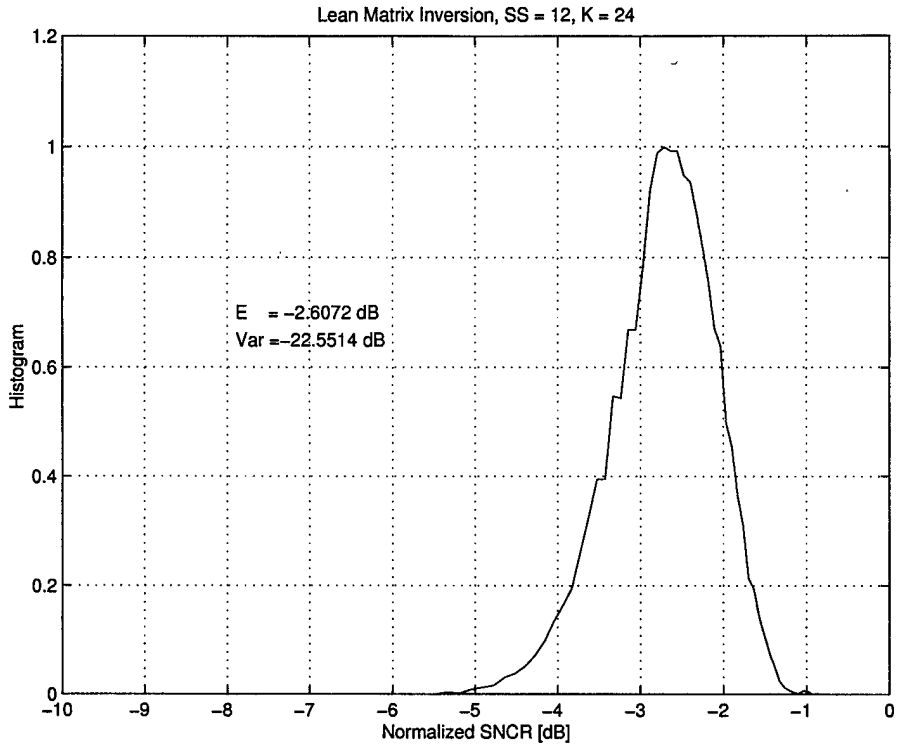

Figure 18 : *Histogram of SNCR for LMI at normalized frequency 0.25, calculated over* 10000 *trials, dimension of subspace 12.*

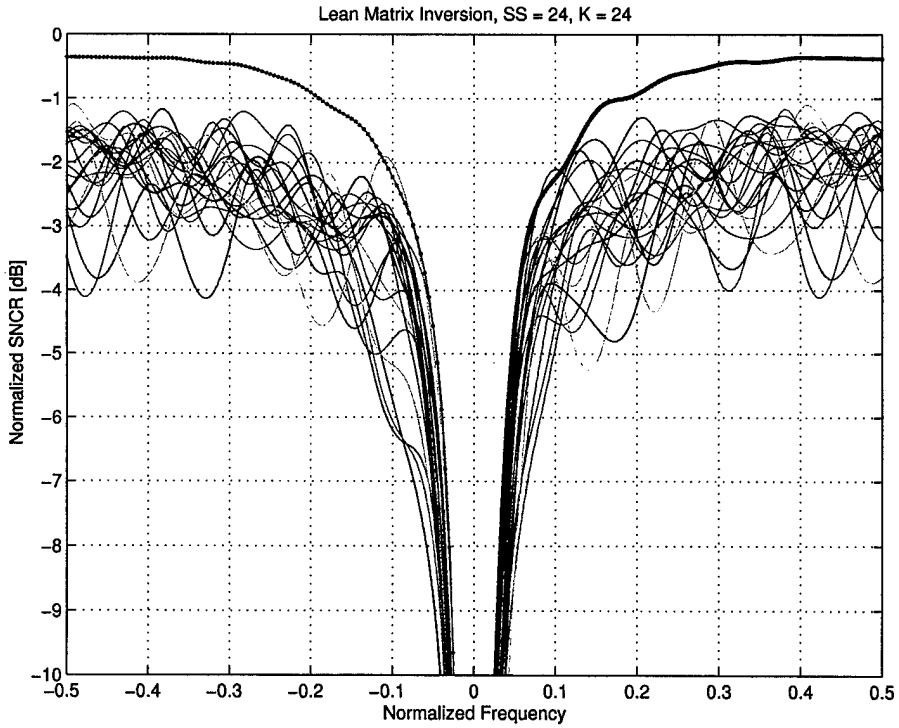

Figure 19 : *Signal-to-clutter plus noise ratio versus normalized frequency for LMI, dimension of subspace 24.*

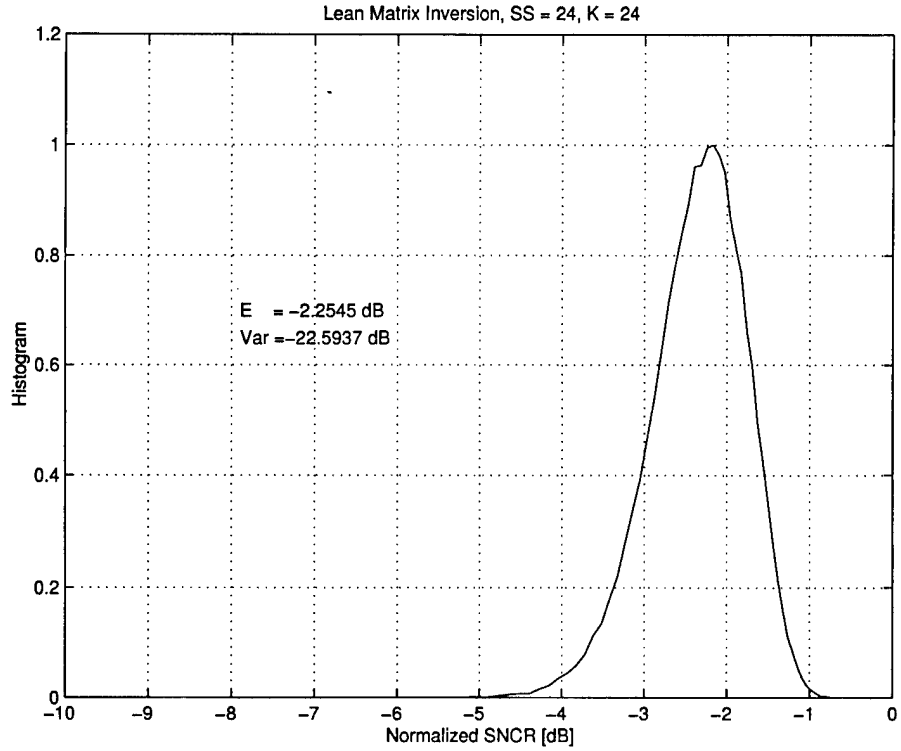

Figure 20 : *Histogram of SNCR for LMI at normalized frequency 0.25, calculated over* 100000 *trials, dimension of subspace 24.*

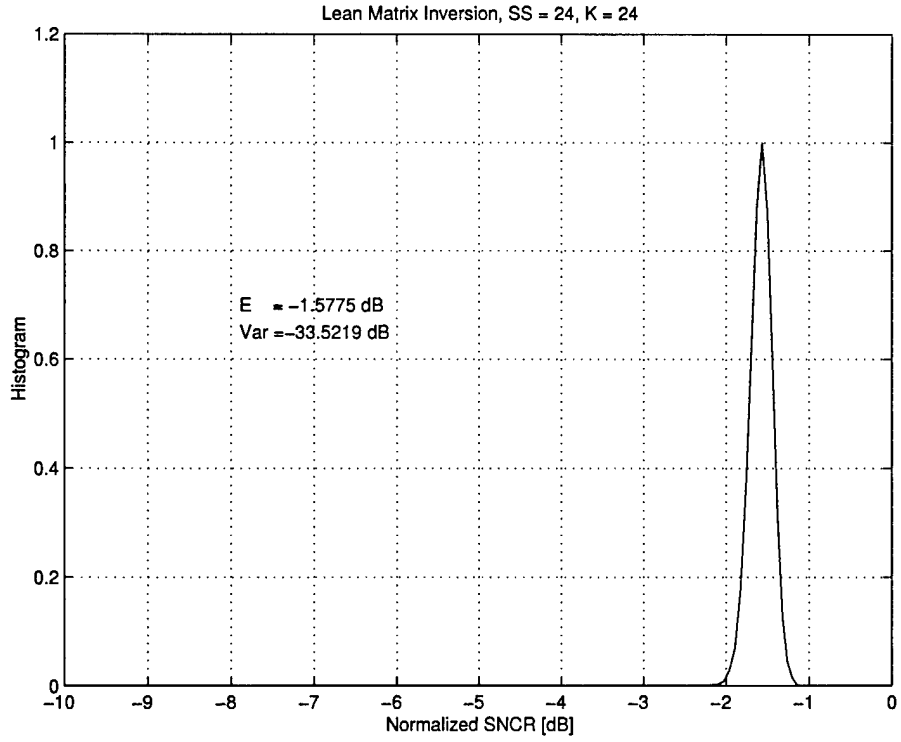

Figure 21 : *Histogram of SNCR for LMI at normalized frequency 0.25, calculated over* 10000 *trials, dimension of subspace 24. Applied on one fixed estimated covariance matrix but with varying transformation matrices.*

4. Computational Complexity

This section presents compares the computational complexity of the different techniques examined in this report. Only complex operations (multiplications and divisions) are counted for each technique as they represent the most intensive calculations. Their number is dependent on the

- number of channels: N,
- number of pulses: M,
- number of samples: K
- and the dimension of clutter subspace: C.

To determine the computational overhead of the SMI/LSMI-method, one must first consider the estimation of the space-time covariance matrix $\hat{\mathbf{R}} = \frac{1}{K}\sum_{k=1}^{K} \underline{X}_k \underline{X}_k^H + \delta\mathbf{I}$. Only the upper (or lower) triangle has to be calculated as the covariance matrix is hermitian. The number of complex multiplications is then $\frac{KNM(NM+1)}{2}$. Instead of a direct inversion of $\hat{\mathbf{R}}$ it is numerically more economical to solve the equation $\hat{\mathbf{R}}\underline{w} = \underline{d}$, e. g. via a Cholesky-decomposition. Since $\hat{\mathbf{R}}$ is hermitian positive definite, there exists a unique lower triangular $\mathbf{L}$ with positive diagonal entries such that $\hat{\mathbf{R}} = \mathbf{L}\mathbf{L}^H$. Since the matrix $\mathbf{L}$ is triangular, the solution can be found by a two-step forward-backward substitution procedure $\mathbf{L}\underline{b} = \underline{d}$ and $\mathbf{L}^H\underline{w} = \underline{b}$. The number of complex operations required to factorize the covariance matrix is $\frac{N(N^2M^2-1)}{6}$. The two substitutions give an additional $N^2M^2 - NM$ operations. Consequently, the overall number of operations is of the order $O(\frac{N^2M^2K}{2} + \frac{N^3M^3}{6})$.

A full eigenvector factorization for the EVP technique needs approximatively $O(10N^3M^3)$ complex operations [32].

The estimation of the clutter subspace $\mathbf{S}$ within the MTP or LMI technique is achieved by

$$\mathbf{S} = \mathbf{X}\left(\mathbf{X}^H\mathbf{E}_1 + \mathbf{X}^H\mathbf{E}_2 + \mathbf{X}^H\mathbf{E}_3\right). \tag{33}$$

This operation requires $NMKC$ complex multiplications since the permutation matrices, E_i, only select the appropriate columns of $\mathbf{X}^H$, they do not calculate them. The orthonormalisation of $\mathbf{S}$ then requires another NMC^2 operations.

Inserting sample values for the simulation (i.e. $N = 12, M = 12, K = 24$ and either $C = 12$ or $C = 24$) yields for the number of complex operations:

- SMI/LSMI: 746496,
- EVP: 29.8710^6,
- MTP/LMI: 41472 if $C = 12$ and 82944 if $C = 24$.

For this scenario the MTP/LMI techniques are about $10 - 20$ times faster than the LSMI, and about $375 - 750$ times faster than the EVP.

5. Conclusions

In this report, two different recently developed fast rank reduction techniques, the so called subspace or projection methods, have been applied to the problem of clutter suppression for MTI. The basic theory behind them was briefly reviewed and their small sample size performance was studied via simulations.

As a result of this study the following important conclusions can be drawn:

- Both subspace techniques, namely MTP and LMI, are capable of suppressing the clutter effectively. The loss of SNCR between the LSMI and EVP on one hand and the MTP and LMI on the other hand is only a fraction of one dB.
- It is not necessary to suppress the full clutter subspace defined by Brennan's rule, particularly in the finite sample case. It is the magnitude of the eigenvalues rather than the number larger than the noise floor which is the dominant factor, i.e. their contribution to the overall clutter power. This reduction leads to savings in computational overhead.
- The asymptotic and finite sample size behaviour of each method are noticeably different. Ripples in the shapes of the matched filter outputs for known covariance matrices, blur or even disappear in the different realizations due to the uncertainty of the estimation process.
- The LMI technique was superior to all other methods, in that it possessed both the smallest loss of SNCR and the lowest computational complexity. The first point is especially noteworthy, since LSMI and EVP were believed to represent the optimum interference suppression performance.

Since this study is based only on one particular simulation scenario, future work will have to focus on different open items, such as:

- Investigation and comparison of the proposed techniques for varying number of channels and pulses; particularly for the practical case of a small number of channels with a larger number of pulses.
- Studying the impact of a different radar orientation, such as a forward-looking system (i.e. range-dependency of the clutter), as well as the effects of inherent channel errors on the performance of MTP and LMI, especially for small sample size.
- Developing a method for automatic determination of an optimum dimension of the clutter subspace for EVP, MTP and LMI, e.g. based on the ratios of the eigenvalue magnitudes.

- Comparison of the performance versus computational cost of the fully adaptive processor concept with a so-called beamspace approach. In the latter the data are dimension-reduced before calculating the cancellation filter. This technique has also been proposed to reduce the computational effort [1].

References

1. Klemm, R. (1998). Space-Time Adaptive Processing, Stevenage, UK: IEE Press.

2. Ward, J. (1994). Space-Time Adaptive Processing for Airborne Radar. Lincoln Laboratory, MIT. (Technical Report 1015). USA.

3. Skolnik, M. I. (1970). Radar Handbook, New York: McGraw-Hill, Inc.

4. Widrow, J. D. Mantley, L. J. Griffiths, P. E. and Goode, B. B. (1967). Adaptive Antenna Systems. *Proc. IEEE*, **55**(12), 2143–2159.

5. Reed, J. D. Mallett, L. S. and Brennan, L. E. (1974). Rapid Convergence Rate in Adaptive Arrays. *IEEE Trans. Aerosp. and Electron. Systems*, **AES-10**(6), 853–863.

6. Cramér, H. (1945). Mathematical Methods of Statistics, Uppsala: Almquist & Wiksells.

7. Boroson, D. M. (1980). Sample Size Considerations for Adaptive Arrays. *IEEE Trans. Aerosp. and Electron. Systems*, **AES-16**(4), 446–451.

8. Ganz, R. L. Moses, M. W. and Wilson, S. L. (1990). Convergence of SMI and the Diagonally Loaded SMI Algorithms with Weak Interference. *IEEE Trans. Antennas and Propagation*, **AP-38**(3), 394–399.

9. Gröger, I. (1979). Nebenkeulenformung der Richtcharakteristik von Gruppenantennen. Ph.D. thesis. (in German), Technical Universiy Berlin.

10. Carlson, B. D. (1988). Covariance Matrix Estimation Errors and Diagonal Loading in Adaptive Arrays. *IEEE Trans. Aerosp. and Electron. Systems*, **AES-24**(4), 397–401.

11. Dilsavor, R. L. and Moses, R. L. (1993). Analysis of Modified SMI Method for Adaptive Array Weight Control. *IEEE Trans. Signal Proc.*, **SP-41**(2), 721–726.

12. Cheremisin, O. P. (1982). Efficiency of Adaptive Algorithms With Regulized Sample Covariance Matrix (in Russian). *Radiotechnology and Electronics (Russia)*, **2**(10), 1933–1941.

13. Gierull, C.H. (1996). Performance Analysis of Fast Projections of the Hung-Turner Type for Adaptive Beamforming. *Signal Processing*, **50**(1-2), 17–29.

14. Brookner, E. and Howell, J. M., (Eds.) (1985). Adaptive-adaptive array processing, Proceedings of the Phased Arrays 1985 Symposium.

15. Gierull, Ch. (1995). Schnelle Signalraumschätzung in Radaranwendungen, Aachen, Germany: Shaker Verlag. Ph.D. Dissertation (in German), Ruhr-University Bochum.

16. Nickel, U. (1993). Radar Target Parameter Estimation with Array Antennas. In Haykin, J. Litva, S. and Shepherd, T.J., (Eds.), *Radar Array Processing*, Vol. 25 of *Springer Series in Information Sciences*, Ch. 3, pp. 47–98. Heidelberg: Springer-Verlag Berlin.

17. Nickel, U. (1997). On the Application of Subspace Methods for Small Sample Size. *AEÜ Int. J. Electron. Commun.*, **51**(6), 279–289.

18. Bienvenu, G. and Kopp, L. (1983). Optimality of High Resolution Array Processing Using the Eigensystem Approach. *IEEE Trans. Acoust., Speech, Signal Processing*, **ASSP-31**(5), 1235–1248.

19. Childers, D. G., (Ed.) (1978). Modern Spectrum Analysis, New York: John Wiley & Sons, Inc.

20. Wax, M. and Kailath, T. (1985). Detection of Signals by Informatic Theoretic Criteria. *IEEE Trans. Acoust., Speech, Signal Processing*, **ASSP-33**, 387–392.

21. Nickel, U. (1982). Vergleich einiger Superauflösungsverfahren zur Spektralschätzung. Forschungsinstitut für Funk und Mathematik. (Technical Report 317). Wachtberg-Werthhoven.

22. Schmidt, R. O. (1981). Multiple Emitter Location and Signal Parameter Estimation. Ph.D. thesis. Stanford University.

23. Gierull, C.H. (1997). Statistical Analysis of the Eigenvector Projection Method for Adaptive Spatial Filtering of Interference. *IEE Proc. Radar, Sonar and Navig.*, **144**(2), 57–63.

24. Karhunen, J. (1984). Adaptive Algorithm for Estimating Eigenvectors of Correlation Type Matrices. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 14.6.1–14.6.4. San Diego.

25. Comon, P. and Golub, G. H. (1990). Tracking a Few Extreme Singular Values and Vectors in Signal Processing. *Proc. IEEE*, **78**(8), 1327–1343.

26. Yang, B. (1995). Projection Approximation Subspace Tracking. *IEEE Trans. Signal Processing*, **SP-43**, 95–107.

27. Hung, E. K. and Turner, R. M. (1983). A Fast Beamforming Algorithm for Large Arrays. *IEEE Trans. Aerosp. and Electron. Systems*, **AES-19**(4), 589–607.

28. Brandwood, D. H. (1987). Noise-Space Projection: MUSIC Without Eigenvectors. *IEE Proc. Pt. H*, **134**(3), 303–309.

29. Yeh, C. C. (1987). Simple Computation of Projection Matrix for Bearing Estimations. *IEE Proc. Pt. F*, **134**(2), 1347–1349.

30. Toulgoat, C.H. Gierull, M. and Turner, R. (Dec. 1993). Subspace Estimation Without Eigenvectors for Adaptive Beamforming. Defence Research Esatblishment Ottawa. (Technical Report 1194). Canada.

31. Chan, T. F. (1987). Rank Revealing QR Factorizations. *Linear Algebra and its Application*, **88**, 67–82.

32. Golub, G. H. and Van Loan, C. F. (1989). Matrix Computations, Baltimore: Johns Hopkins University Press.

33. Hsieh, K. J. R. Liu, S. F. and Yao, K. (1993). Estimation of Multiple Sinusoidal Frequencies Using Truncated Least Squares Methods. *IEEE Trans. Signal Processing*, **SP-41**(2), 990–994.

34. Prasad, S. and Chandna, B. (1991). Direction-of-Arrival Estimation Using Rank Revealing QR Factorization. *IEEE Trans. Signal Processing*, **SP-39**(5), 1224–1229.

35. Reilly, W. G. Chen, J. P. and Wong, K. M. (1988). A Fast QR-based Array-Processing Algorithm. *Proc. SPIE Advanced Algor. and Archit. for Signal Processing*, **975**, 36–47.

36. Gierull, C.H. (1997). A Fast Subspace Estimation Method for Adaptive Beamforming Based on Covariance Matrix Transformation. *AEÜ Int. J. Electr. Commun.*, **51**(4), 187–232.

37. Gierull, C.H. (1999). Angle Estimation for small sample size with fast eigenvector-free subspace method. *IEE Proc. Radar, Sonar and Navig.*, **146**(3), 126–132.

38. Nickel, U. (1989). Subarray Configurations for Interference Suppression with Phased Array Radar. In *Proc. Int. Conf. on Radar*, pp. 82–86. Paris.

39. Goffer, M. Kam , A. P. and Herczfeld, P. R. (1994). Design of Phased Arrays in Terms of Random Subarrays. *IEEE Trans. Antennas and Propagation*, **AP-42**(6), 820–826.

40. Gierull, C.H. (1998). Fast and Effective Method for Low-Rank Interference Suppression in Presence of Channel Errors. *Electronics Letters*, **El-34**(6), 518–520.

41. Nickel, U. (1992). On the Influence of Channel Errors on Array Signal Processing Methods. *Archiv f"ur Elektronik und "Ubertragungstechnik*, **AE"U-47**(4), 209–219.

Annex A
Proof of Eq. (20)

Let $\mathbf{Q} \in \mathbb{C}^{N \times M}$ be an unitary matrix and $\mathbf{\Lambda}$ a diagonal matrix with real entries $\lambda_1, \cdots, \lambda_N$, then:

$$\max \operatorname{tr}\{\mathbf{Q}^H \mathbf{\Lambda} \mathbf{Q}\} = \sum_{m=1}^{M} \lambda_m. \tag{A.1}$$

Proof:
The trace in eq. (A.1) can be re-written as

$$\operatorname{tr}\{\mathbf{Q}^H \mathbf{\Lambda} \mathbf{Q}\} = \operatorname{tr}\{\mathbf{\Lambda} \mathbf{Q} \mathbf{Q}^H\} = \sum_{m=1}^{N} \lambda_m \underline{e}_m^H \mathbf{Q} \mathbf{Q}^H \underline{e}_m = \sum_{m=1}^{N} \lambda_m \alpha_m$$

where $\underline{e}_m$ denotes the m-th column of the identity matrix $\mathbf{I}_N$. Since $\mathbf{Q}\mathbf{Q}^H$ describes a projection matrix (i.e. it is idempotent), it has M-eigenvalues equal to 1 and the rest 0. Hence it yields for the diagonal elements $\alpha_m = \underline{e}_m^H \mathbf{Q} \mathbf{Q}^H \underline{e}_m$,

$$\text{i)} \quad 0 \le \alpha_m \le 1 \quad m = 1, \ldots, N$$

$$\text{ii)} \quad \sum_{m=1}^{M} \alpha_m = \operatorname{tr}\{\mathbf{Q}\mathbf{Q}^H\} = M.$$

Proposition:

$$\sum_{m=1}^{N} \lambda_m \alpha_m = \sum_{m=1}^{M} \lambda_m \alpha_m + \sum_{m=M+1}^{N} \lambda_m \alpha_m \le \sum_{m=1}^{M} \lambda_m$$

or

$$\sum_{m=1}^{M} \lambda_m (\alpha_m - 1) + \sum_{m=M+1}^{N} \lambda_m \alpha_m \le 0,$$

respectively. Since λ_{M+1} is greater than or equal to zero, it follows

$$\sum_{m=1}^{M} (\lambda_m - \lambda_{M+1})(\alpha_m - 1) + \sum_{m=M+1}^{N} (\lambda_m - \lambda_{M+1}) \alpha_m \le 0 \tag{A.2}$$

and

$$\begin{aligned} \sum_{m=1}^{M} \lambda_m (\alpha_m - 1) + \sum_{m=M+1}^{N} \lambda_m \alpha_m &\le \lambda_{M+1} \left(\sum_{m=1}^{M} (\alpha_m - 1) + \sum_{m=M+1}^{N} \alpha_m \right) \\ &= \lambda_{M+1} \left(\sum_{m=1}^{N} \alpha_m - M \right). \end{aligned} \tag{A.3}$$

Using *ii*) the term within the brackets on the right side of inequality (A.3) equals zero.

(q.e.d)

DOCUMENT CONTROL DATA

(Security classification of title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (the name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Establishment sponsoring a contractor's report, or tasking agency, are entered in section 8.)

DEFENCE RESEARCH ESTABLISHMENT OTTAWA
DEPARTMENT OF NATIONAL DEFENCE
OTTAWA, ONTARIO, CANADA, K1A 0K2

2. SECURITY CLASSIFICATION (overall security classification of the document, including special warning terms if applicable)

UNCLASSIFIED

3. TITLE (the complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S,C or U) in parentheses after the title.)

ANALYSIS OF THE SMALL SAMPLE SIZE PERFORMANCE OF FAST FULLY ADAPTIVE STAP TECHNIQUES FOR MTI RADAR (U)

4. AUTHORS (Last name, first name, middle initial)

GIERULL, CHRISTOPH, H.

5. DATE OF PUBLICATION (month and year of publication of document)	6a. NO. OF PAGES (total containing information. Include Annexes, Appendices, etc.)	6b. NO. OF REFS (total cited in document)
October 2001	56	41

7. DESCRIPTIVE NOTES (the category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

DREO REPORT

8. SPONSORING ACTIVITY (the name of the department project office or laboratory sponsoring the research and development. Include the address.)

DEFENCE RESEARCH ESTABLISHMENT OTTAWA
DEPARTMENT OF NATIONAL DEFENCE
OTTAWA, ONTARIO, CANADA, K1A 0K2

9a. PROJECT OR GRANT NO. (if appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant)	9b. CONTRACT NO. (if appropriate, the applicable number under which the document was written)
3DB29	

10a. ORIGINATOR'S DOCUMENT NUMBER (the official document number by which the document is identified by the originating activity. This number must be unique to this document.)	10b. OTHER DOCUMENT NOS. (Any other numbers which may be assigned this document either by the originator or by the sponsor)
DREO TECHNICAL REPORT 2001-079	

11. DOCUMENT AVAILABILITY (any limitations on further dissemination of the document, other than those imposed by security classification)

(X) Unlimited distribution
( ) Distribution limited to defence departments and defence contractors; further distribution only as approved
( ) Distribution limited to defence departments and Canadian defence contractors; further distribution only as approved
( ) Distribution limited to government departments and agencies; further distribution only as approved
( ) Distribution limited to defence departments; further distribution only as approved
( ) Other (please specify):

12. DOCUMENT ANNOUNCEMENT (any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in 11) is possible, a wider announcement audience may be selected.)

13. ABSTRACT (a brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual).

(U) In ground surveillance from an airborne or space-based radar it is desirable to be able to detect small moving targets, such as tanks or wheeled vehicles, within severe ground clutter. For operational moving target indication (MTI) systems the clutter filter coefficients have to be updated frequently due to rapidly changing interference environment. This report examines the small sample size performance of different fast fully adaptive space-time processors (STAP) and compares it to the optimum-detector performance. These recently proposed techniques, named Matrix Transformation based Projection (MTP) and Lean Matrix Inversion (LMI), were originally developed to provide fast man-made jammer suppression in large surface phased array radars with many elements. For this application they have been proven to operate with near-optimum performance, yet with a computational expense drastically reduced from that of the optimum detector in most practical cases. The investigation herein focuses on the performance achieved when only a few data samples are available to adapt (update) the clutter filter coefficient.

In this report, the techniques are described and a number of simulations carried out. The two applications, STAP and jammer suppression, are similar; both are required to suppress an interference which is characterized by a certain number of dominant eigenvalues of the sample space-time covariance matrix. Despite the similarities the performance between the two differs due to the different shapes of their eigenvalue distribution. The LMI is shown to give the best Signal-to-Noise-plus-Clutter Ratio (SNCR) for a given computational load.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus. e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus-identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

SPACE-TIME ADAPTIVE PROCESSING
GROUND MOVING TARGET INDICATION
SUBSPACE TECHNIQUES

Defence R&D Canada
is the national authority for providing
Science and Technology (S&T) leadership
in the advancement and maintenance
of Canada's defence capabilities.

R et D pour la défense Canada
est responsable, au niveau national, pour
les sciences et la technologie (S et T)
au service de l'avancement et du maintien des
capacités de défense du Canada.

DEFENCE R&D DÉFENSE

www.drdc-rddc.dnd.ca