Publishers' page

Publishers' page

Publishers' page

Publishers' page

Dedication Page

(optional)

Contents

1. Space-Time Adaptive Beamforming Algorithms for Airborne Radar Systems 1

1.1 Introduction 1
1.2 System and Signal Models 3
1.3 Conventional Beamforming Algorithms 4
1.4 Low-Rank Beamforming Algorithms 5
1.4.1 Eigenvalue-decomposition-based algorithms . . . 7
1.4.2 Krylov subspace-based algorithms 8
1.4.3 JIO-based algorithms 9
1.4.4 JIDF-based algorithms 10
1.5 Sparsity-aware Beamforming Algorithms 14
1.6 Knowledge-aided Beamforming Algorithms 16
1.7 Simulations 18
1.8 Concluding Remarks 18

Bibliography 19

Index 23

Chapter 1

Space-Time Adaptive Beamforming Algorithms for Airborne Radar Systems

1.1 Introduction

Space-time adaptive processing (STAP) techniques [Klemm (2002)],[Melvin (2004)] have been thoroughly investigated in the last decades as a key enabling technology for advanced airborne radar applications following the landmark publication by Brennan and Reed [Brennan and Reed (1973)]. A great deal of attention has been given to STAP algorithms and different strategies to design space-time beamformers to mitigate the effect of clutter and jamming signals [Reed *et al.* (1974)]- [Guerci (2000)]. It is fully understood that STAP techniques can improve slow-moving target detection through better mainlobe clutter suppression, provide better detection in combined clutter and jamming environments, and offer a significant increase in output signal to- interference-plus-noise-ratio (SINR). Moreover, it is also well understood that the clutter and the jamming signals often reside in a low-rank signal subspace, which is typically much lower than the number of degrees of freedom of the array and the associated space-time beamformer. Due to its large computational complexity cost by the matrix inversion operation, the optimum STAP processor is prohibitive for practical implementation. In addition, another very challenging issue that is encountered by full-rank STAP techniques is when the number of elements M in the spatio-temporal beamformer is large. It is well-known that $K \geq 2M$ independent and identically distributed (i.i.d) training samples are required for the filter to achieve the steady performance [Haykin (2002)]. Thus, in dynamic scenarios the full-rank STAP with large M usually fail or provide poor performance in tracking target signals contaminated by interference and noise.

In the recent years, a number of innovative space-time beamforming algorithms have been reported in the literature for clutter and interference mitigation in radar systems. These algorithms include reduced-rank and reduced-dimension techniques [Haimovich (1991)]-[de Lamare and Sampaio-Neto (2009)], which employ a two-stage processing framework to exploit the low-rank property of the clutter and the jamming signals. The first stage performs dimensionality reduction and is followed by second stage that employs a beamforming algorithm with a reduced dimensional filter. Another class of important space-time beamforming algorithms adopt the strategy of compressive sensing and sparsity-aware algorithms, which exploit the fact that space-time beamformers do not need all their degrees of freedom to mitigate clutter and jamming signals. These algorithms compute sparse space-time beamformers which can converge faster and are effective for STAP in radar systems. By exploiting the low-rank properties of the interference and devising sparse STAP algorithms, designers make use of prior knowledge about the clutter and the jamming signals. It has been recently shown that it is beneficial in terms of performance to also exploit prior knowledge about the environment and the data in the form of a known covariance data matrix. The class of space-time beamforming algorithms that exploit different forms of prior knowledge are called knowledge-aided STAP (KA-STAP) algorithms.

The goal of this chapter is to review the recent work and advances in the area of space-time beamforming algorithms and their application to radar systems. These systems include phased-array and multi-input multi-output (MIMO) radar systems, mono-static and bi-static radar systems and other configurations. Furthermore, this chapter also describes in detail some of the most successful space-time beamforming algorithms that exploit low-rank and sparsity properties as well as the use of prior-knowledge to improve the performance of STAP algorithms in radar systems.

The chapter is structured as follows. Section 1.2 describes the radar system under consideration and signal model used to mathematically describe. Section 1.3 formulates the problem of designing space-time beamformers and reviews conventional space-time beamforming algorithms. Section 1.4 examines low-rank space-time beamforming algorithms, whereas Section 1.5 explores the concept of sparsity-aware space-time beamforming algorithms. Section 1.6 studies knowledge-aided beamforming algorithms and discusses how these techniques can be adopted in existing radar systems. Section 1.7 is devoted to the presentation of simulation results, discussions and the comparison of a number of existing algorithms. The chapter ends with Section 1.8 which gives the concluding remarks of this chapter.

1.2 System and Signal Models

The system under consideration is a pulsed Doppler radar residing on an airborne platform. The radar antenna is a uniformly spaced linear antenna array consisting of N elements. The radar returns are collected in a coherent processing interval (CPI), which is referred to as the 3-D radar datacube shown in Fig. 1(a), where K denotes the number of samples collected to cover the range interval. The data is then processed at one range of interest, which corresponds to a slice of the CPI datacube. This slice is a $J \times N$ matrix which consists of $N \times 1$ spatial snapshots for J pulses at the range of interest. It is convenient to stack the matrix column-wise to form the $M \times 1$ vector r(i), termed the i-th range gate spacetime snapshot, where $M = JN$ and $1 < i \leq K$ [Klemm (2002)].

The objective of a radar is to ascertain whether targets are present in the data. Thus, given a space-time snapshot, radar detection is a binary hypothesis problem, where hypothesis H_0 corresponds to the absence of a target and hypothesis H_1 corresponds to the presence of a target. The radar space-time snapshot is then expressed for each of the two hypotheses in the following form

$$\begin{aligned} H_0 &: \boldsymbol{r}(i) = \boldsymbol{v}(i); \\ H1 &: \boldsymbol{r}(i) = a\boldsymbol{s} + \boldsymbol{v}(i); \end{aligned} \tag{1.1}$$

where a is a zero-mean complex Gaussian random variable with variance σ_s^2 , $\boldsymbol{v}(i) = \boldsymbol{r}_c(i) + \boldsymbol{r}_j(i) + \boldsymbol{n}(i)$ contains the input interference-plus-noise vector which consists of clutter $\boldsymbol{r}_c(i)$, jamming $\boldsymbol{r}_j(i)$ and the white noise $\boldsymbol{n}(i)$. These three components are assumed to be mutually uncorrelated. Thus, the $M \times M$ covariance matrix $\boldsymbol{R}$ of the undesired clutter-plus-jammer-plus-noise component can be modelled as

$$\boldsymbol{R} = \boldsymbol{R}_c + \boldsymbol{R}_j + \boldsymbol{R}_n \tag{1.2}$$

where $(\cdot)^H$ represents the Hermitian transpose and $E[\cdot]$ denotes expectation. The noise covariance noise matrix is given by $R_n = E[\boldsymbol{n}(i)\boldsymbol{n}^H(i)] = \sigma_n^2\boldsymbol{I}$, where σ_n^2 is the variance of the noise and $\boldsymbol{I}$ is an identity matrix. The clutter signal can be modeled as the superposition of a large number of independent clutter patches with evenly distributed in azimuth about the receiver. Thus, the clutter covariance matrix can be expressed as

$$\boldsymbol{R}_c = E[\boldsymbol{r}_c\boldsymbol{r}_c^H] = \sum_{k=1}^{N_r}\sum_{l=1}^{N_c} \xi_{k,l}^c[\boldsymbol{b}(\vartheta_{k,l}^c)\boldsymbol{b}(\vartheta_{k,l}^{c,\,H}) \otimes [\boldsymbol{a}(\varpi_{k,l}^c)\boldsymbol{a}(\varpi_{k,l}^{c,\,H}), \tag{1.3}$$

where N_r denotes the number of range ambiguities and N_c denotes the number of the clutter patches. The quantity $\xi_{k,l}^c$ is the power of reflected signal by the kl-th clutter patch. The symbol $\otimes$ denotes Kronecker product, and the quantities $\boldsymbol{b}(\vartheta_{k,l}^{c,}{}^H$ and $\boldsymbol{a}(\varpi_{k,l}^c)$ denote the spatial steering vector with the spatial frequency $\vartheta_{k,l}^c$ and the temporal steering vector with the normalized Doppler frequency $\varpi_{k,l}^c$ for the k,l-th clutter patch, respectively, which can be expressed as follows

$$\boldsymbol{b}(\vartheta_{k,l}^c) = \begin{bmatrix} 1 \\ e^{-j2\pi\vartheta} \\ e^{-j2\pi 2\vartheta} \\ \vdots \\ e^{-j2\pi(N-1)\vartheta} \end{bmatrix}, \quad \boldsymbol{a}(\varpi_{k,l}^c) = \begin{bmatrix} 1 \\ e^{-j2\pi\varpi} \\ e^{-j2\pi 2\varpi} \\ \vdots \\ e^{-j2\pi(N-1)\varpi} \end{bmatrix}, \tag{1.4}$$

where $\vartheta = \frac{d}{\lambda}cos(\phi)sin(\theta)$ and $\vartheta = f_d/f_r$, λ is the wavelength, d is the inter-element spacing which is normally set to half wavelength, and ϕ and θ are the elevation and the azimuth angles, respectively. The quantities f_d and f_r are the Doppler frequency and the pulse repetition frequency, respectively. The jamming covariance matrix $\boldsymbol{R}_j = E[\boldsymbol{r}_j(i)\boldsymbol{r}_j^H(i)]$ can be written as

$$\boldsymbol{R}_j = \sum_{q=1}^{N_j} \xi_q^j [\boldsymbol{b}(\vartheta_q^j)\boldsymbol{b}^H(\vartheta_q^j)] \otimes \boldsymbol{I}_K, \tag{1.5}$$

where ξ_q^j is the power of the q-th jammer. The vector $\boldsymbol{b}(\vartheta_q^j)$ is the spatial steering vector with the spatial frequency ϑ_q^j of the q-th jammer and N_j is the number of jamming signals. The vector $\boldsymbol{s}$, which is the $M \times 1$ normalized space-time steering vector in the space-time look-direction can be defined as

$$\boldsymbol{s} = \sqrt{\xi_t}\boldsymbol{b}(\vartheta_t)\boldsymbol{a}(\varpi_t), \tag{1.6}$$

where $\boldsymbol{a}(\varpi_t)$ is the $K \times 1$ normalized temporal steering vector at the target Doppler frequency ϖ_t and $b(\vartheta_t)$ is the $N \times 1$ normalized spatial steering vector in the direction provided by the target spatial frequency ϑ_t and ξ_t denotes the power of the target.

1.3 Conventional Beamforming Algorithms

In order to detect the presence of targets, each range bin is processed by an adaptive space-time beamformer, which is typically designed to achieve maximum output SINR, followed by a hypothesis test to determine the target presence or absence. The secondary data $\boldsymbol{r}(i)$ are taking from training samples, which should be ideally i.i.d. training samples but it is often non-heterogeneous [Klemm (2002)]. The optimum full-rank STAP that maximizes the SINR can obtained by solving the following minimum variance distortionless response (MVDR) constrained optimization given by:

$$\boldsymbol{w}_{\text{opt}} = \arg\min_{\boldsymbol{w}} \boldsymbol{w}^H \boldsymbol{R} \boldsymbol{w} \quad \text{subject to } \boldsymbol{w}^H \boldsymbol{s} = 1, \tag{1.7}$$

where the optimal space-time MVDR beamformer $\boldsymbol{w}_{\text{opt}}$ is designed to maximize the SINR and to maintain a normalized response in the target spatial-Doppler look-direction. The solution to the optimization problem above is described by:

$$\boldsymbol{w}_{\text{opt}} = \frac{\boldsymbol{R}^{-1}\boldsymbol{s}}{\boldsymbol{s}^H \boldsymbol{R}^{-1} \boldsymbol{s}}. \tag{1.8}$$

The space-time beamformer $\boldsymbol{w}_{\text{opt}}$ can be computed by using the above solution. Alternatively, the space-time beamformer can be estimated adaptive algorithms [Haykin (2002)]. These algorithms include the least mean-square (LMS), the conjugate gradient (CG) and the recursive least-squares (RLS) techniques. The computational complexity of these algorithms ranges from linear with M for the LMS to quadratic with M for the CG and RLS algorithms. A common problem with the conventional adaptive algorithms is that the laws that govern their convergence and tracking behaviors imply that they depend on M and on the eigenvalue spread of $\boldsymbol{R}$. This indicates that their performance degrades significantly when the space-time beamformer has many parameters for adaptation, which makes the computation of the parameters of the beamformer slow and costly. This problem can be addressed by some recent techniques reported in the literature, namely, low-rank, sparsity-aware and knowledged-aided algorithms.

1.4 Low-Rank Beamforming Algorithms

Reduced-rank adaptive signal processing has been considered as a key technique for dealing with large systems in the last decade. The basic idea of the reduced-rank algorithms is to reduce the number of adaptive coefficients by projecting the received vectors onto a lower dimensional subspace which consists of a set of basis vectors. The adaptation of the low-order filter within the lower dimensional subspace results in significant computational savings, faster convergence speed and better tracking performance. The first statistical reduced-rank method was based on a principal-components (PC) decomposition of the target-free covariance matrix [Haimovich (1991)]. Another class of eigen-decomposition methods was based on the cross-spectral metric (CSM) [Goldstein and Reed (1997c,a)]. Both the PC and the CSM algorithms require a high computational cost due to the eigen-decomposition. A family of the Krylov subspace methods has been investigated thoroughly in the recent years. This class of reduced-rank algorithms, including the multistage Wiener filter (MSWF) [Golstein *et al.* (1998); Guerci (2000); Gau and Reed (1998)] which projects the observation data onto a lower-dimensional Krylov subspace, and the auxiliary-vector filters (AVF) [Pados and Batalama (1999); Pados and Karystinos (2001); Pados *et al.* (2007)]. These methods are relatively complex to implement in practice and may suffer from numerical problems despite their improved convergence and tracking performance. The joint domain localized (JDL) approach, which is a beamspace reduced-dimension algorithm, was proposed by Wang and Cai [Wang and Cai (1994)] and investigated in both homogeneous and nonhomogeneous environments in [Adve *et al.* (2000a,b)], respectively. Recently, reduced-rank adaptive processing algorithms based on joint iterative optimization of adaptive filters [de Lamare and Sampaio-Neto (2007a); Fa *et al.* (2008)] and based on an adaptive diversity-combined decimation and interpolation scheme [de Lamare and Sampaio-Neto (2007b, 2009)] were proposed, respectively.

The basic idea of low-rank algorithms is to reduce the number of adaptive coefficients by projecting the received vectors onto a lower dimensional subspace as illuminated in the figure. Let $\boldsymbol{S}_D$ denote the $M \times D$ rank-reduction matrix with column vectors which form an $M \times 1$ basis for a D-dimensional subspace, where $D < M$. Thus, the received signal $\boldsymbol{r}(i)$ is transformed into its reduced-rank version $\boldsymbol{r}_D(i)$ given by

$$\boldsymbol{r}_D(i) = \boldsymbol{S}_D^H \boldsymbol{r}(i) \tag{1.9}$$

The reduced-rank signal is processed by an adaptive low-rank space-time beamformer $\boldsymbol{w}_D$ with D coefficients. Subsequently, the decision is made based on the output of the beamformer $y(i) = \boldsymbol{w}_D^H \boldsymbol{S}_D^H \boldsymbol{r}(i)$. A designer can compute the parameters of the beamformer by solving the following constrained optimization problem:

$$\boldsymbol{w}_{D,\text{opt}} = \arg\min_{\boldsymbol{w}_D} \boldsymbol{w}_D^H \boldsymbol{S}_D^H \boldsymbol{R} \boldsymbol{S}_D \boldsymbol{w}_D \quad \text{subject to } \boldsymbol{w}_D^H \boldsymbol{S}_D^H \boldsymbol{s} = 1, \tag{1.10}$$

The optimum low-rank MVDR solution for the above problem is given by

$$\boldsymbol{w}_{D,\text{opt}} = \frac{(\boldsymbol{S}_D^H \boldsymbol{R} \boldsymbol{S}_D)^{-1} \boldsymbol{S}_D^H \boldsymbol{s}}{\boldsymbol{s}^H \boldsymbol{S}_D (\boldsymbol{S}_D^H \boldsymbol{R} \boldsymbol{S}_D)^{-1} \boldsymbol{S}_D^H \boldsymbol{s}} = \frac{\boldsymbol{R}_D^{-1} \boldsymbol{s}_D}{\boldsymbol{s}_D^H \boldsymbol{R}_D^{-1} \boldsymbol{s}_D}. \tag{1.11}$$

where $\boldsymbol{R}_D = \boldsymbol{S}_D^H \boldsymbol{R} \boldsymbol{S}_D$ denotes the low-rank covariance matrix and $\boldsymbol{s}_D = \boldsymbol{S}_D^H s$ denotes the low-rank steering vector. The key challenge in the design of low-rank STAP algorithms is find a cost-effective method to compute the rank-reduction matrix $\boldsymbol{S}_D$.

1.4.1 *Eigenvalue-decomposition-based algorithms*

The eigenvalue-decomposition (EVD)-based beamforming algorithms are also known as PC-based algorithms and have been originally reported as the eigencanceler method. These PC-based algorithms refer to the beamformers constructed with a subset of the eigenvectors of the interference-only covariance matrix associated with the eigenvalues of largest magnitude. The application of this method to radar was reported in [Haimovich (1991)].

The basic idea of the EVD-based beamformer is to approximate the $M \times M$ covariance matrix $\boldsymbol{R}$ of the received data as follows:

$$\boldsymbol{R} = \sum_{d=1}^{D} \lambda_d \boldsymbol{v}_d \boldsymbol{v}_d^H, \tag{1.12}$$

where the $M \times 1$ vector $\boldsymbol{v}_d$ corresponds to the dth eigenvector of $\boldsymbol{R}$ and λ_d is the dth eigenvalue of $\boldsymbol{R}$. By assuming that the eigenvalues are obtained in decreasing order of magnitude, the EVD-based method approximates $\boldsymbol{R}$ using its D dominant eigenvectors. The rank-reduction matrix is constructed by using the D dominant eigenvectors as described by

$$\boldsymbol{S}_D = [\boldsymbol{v}_1 \boldsymbol{v}_2 \dots \boldsymbol{v}_D] \tag{1.13}$$

The low-rank MVDR solution for the above problem is given by

$$\boldsymbol{w}_D = \frac{\boldsymbol{S}_D(\boldsymbol{S}_D^H \boldsymbol{R} \boldsymbol{S}_D)^{-1} \boldsymbol{S}_D^H \boldsymbol{s}}{\boldsymbol{s}^H \boldsymbol{S}_D(\boldsymbol{S}_D^H \boldsymbol{R} \boldsymbol{S}_D)^{-1} \boldsymbol{S}_D^H \boldsymbol{s}} = \frac{(\sum_{d=1}^{D} \lambda_d^{-1} \boldsymbol{v}_d \boldsymbol{v}_d^H)\boldsymbol{s}}{\boldsymbol{s}^H (\sum_{d=1}^{D} \lambda_d^{-1} \boldsymbol{v}_d \boldsymbol{v}_d^H)\boldsymbol{s}}. \tag{1.14}$$

The EVD-based low-rank MVDR space-time beamformer described above does not take into account the target steering vector $\boldsymbol{s}$ when selecting a suitable subspace representation of the interference. Clearly, this low-rank space-time beamformer requires the computation of an EVD, which has a computational cost that is cubic with M [Golub and van Loan (2002)]. In order to reduce this computational complexity, a designer can resort to subspace tracking algorithms which bring the cost down to $O(M^2)$. Another technique associated with EVD-based beamforming that can improve the performance of low-rank MVDR space-time beamformers is the method called cross-spectral metric (CSM) [Goldstein and Reed (1997c)]. The

CSM approach chooses the set of D eigenvectors for the rank-reduction matrix which optimizes the desired criterion, namely, the maximization of the SINR, in opposition to the PC method which always chooses the dominant eigenvectors.

1.4.2 *Krylov subspace-based algorithms*

The first Krylov methods, namely, the conjugate gradient (CG) method [Hestenes and Stiefel (1952)] and the Lanczos algorithm [Lanczos (1952)] have been originally proposed for solving large systems of linear equations. These algorithms used in numerical linear algebra are mathematically identical to each other and have been derived for Hermitian and positive definite system matrices. Other techniques have been reported for solving these problems and the Arnoldi algorithm [Arnoldi (1951)] is a computationally efficient procedure for arbitrarily invertible system matrices. The multistage Wiener filter (MSWF) [Goldstein and Reed (1997c)] and the auxiliary vector filtering (AVF) [Pados and Batalama (1999)] algorithms are based on a multistage decomposition of the linear MMSE estimator. A key feature of these methods is that they do not require an EVD and have a very good performance. It turns out that Krylov subspace algorithms that are used for solving very large and sparse systems of linear equations, are highly suitable alternatives for designing low-rank space-time beamforming algorithms in radar systems. The basic idea of Krylov subspace algorithms is to construct the rank-reduction matrix $\mathbf{S}_D$ with the following structure:

$$\boldsymbol{S}_D = \left[\boldsymbol{q}\ \boldsymbol{R}\boldsymbol{q}\ \ldots\ \boldsymbol{R}^{D-1}\boldsymbol{q}\right], \tag{1.15}$$

where $\boldsymbol{q} = \frac{\boldsymbol{s}}{||\boldsymbol{s}[i]||}$ and $||\cdot||$ denotes the Euclidean norm (or the 2-norm) of a vector. In order to compute the basis vectors of the Krylov subspace (the vectors of $\boldsymbol{S}_D$), a designer can either directly employ the expression in (1.15) or resort to more sophisticated approaches such as the Arnoldi iteration [Arnoldi (1951)]. The low-rank MVDR solution for the space-time beamformer using the Krylov subspace is given by

$$\boldsymbol{w}_D = \frac{(\boldsymbol{S}_D^H\boldsymbol{R}\boldsymbol{S}_D)^{-1}\boldsymbol{S}_D^H\boldsymbol{s}}{\boldsymbol{s}^H\boldsymbol{S}_D(\boldsymbol{S}_D^H\boldsymbol{R}\boldsymbol{S}_D)^{-1}\boldsymbol{S}_D^H\boldsymbol{s}}. \tag{1.16}$$

An appealing feature of the Krylov subspace algorithms is that the required model order D does not scale with the system size. Indeed, when M goes to infinity the required D remains a finite and relatively small value. This result was established in [Xiao and Honig (2005)]. Among the disadvantages of Krylov subspace methods are the relatively high computational cost of constructing $\boldsymbol{S}_D$ ($O(DM^2)$), the numerical instability of some implementations and the lack of flexibility for imposing constraints on the design of the basis vectors.

1.4.3 *JIO-based algorithms*

The aim of this part is to introduce the reader to low-rank beamforming algorithms based on joint iterative optimization (JIO) techniques. The idea of these methods is to design the main components of a low-rank space-time beamforming scheme via a general optimization approach. The basic ideas of JIO techniques have been reported in [de Lamare and Sampaio-Neto (2007a)]. Amongst the advantages of JIO techniques are the flexibility to choose the optimisation algorithm and to impose constraints, which provides a significant advantage over eigen-based and Krylov subspace methods. One disadvantage that is shared amongst the JIO techniques, eigen-based and Krylov subspace methods are the complexity associated with the design of the matrix $\mathbf{S}_D$. For instance, if we are to design a beamforming algorithm with a very large M, we still have the problem of having to design an $M \times D$ rank-reduction matrix $\mathbf{S}_D$.

In the framework of JIO techniques, the design of the matrix $\mathbf{S}_D$ and the beamforming vector $\mathbf{w}_D$ for a fixed model order D will be dictated by the optimization problem. To this end, we will focus on a generic $\mathbf{S}_D = \left[\mathbf{s}_1\ \mathbf{s}_2\ \ldots\ \mathbf{s}_D\right]$, in which the basis vectors $\mathbf{s}_d$, $d = 1, 2, \ldots, D$ will be obtained via an optimization algorithm and iterations between the $\mathbf{S}_D$ and $\mathbf{w}_D$ will be performed. The JIO method consists of solving the following optimization problem

$$\begin{gathered}\left[\boldsymbol{S}_{D,\text{opt}}, \boldsymbol{w}_{D,\text{opt}}\right] = \arg \min_{\boldsymbol{S}_D, \boldsymbol{w}_D} \underbrace{\underbrace{\boldsymbol{w}_D^H \boldsymbol{S}_D^H \boldsymbol{R} \boldsymbol{S}_D \boldsymbol{w}_D}_{x(i)}}_{\boldsymbol{C}(\boldsymbol{S}_D, \boldsymbol{w}_D)}, \\ \text{subject to } \boldsymbol{w}_D^H \boldsymbol{S}_D^H \boldsymbol{s} = 1\end{gathered} \tag{1.17}$$

where it should be remarked that the optimization problem in (1.17) is non convex, however, the algorithms do not present convergence problems. Numerical studies with JIO methods indicate that the minima are identical and global. Proofs of global convergence have been established with different versions of JIO schemes [de Lamare and Sampaio-Neto (2007a)], which demonstrate that the LS algorithm converges to the reduced-rank Wiener filter.

In order to solve the above problem, we resort to the method of Lagrange multipliers [Haykin (2002)] and transform the constrained optimization into an unconstrained one expressed by the Lagrangian

$$\mathcal{L}(\boldsymbol{S}_D, \boldsymbol{w}_D) = \boldsymbol{w}_D^H \boldsymbol{S}_D^H \boldsymbol{R} \boldsymbol{S}_D \boldsymbol{w}_D + \lambda(\bar{\boldsymbol{w}}_D^H \mathbf{S}_D^H \boldsymbol{s} - 1), \tag{1.18}$$

where λ is a scalar Lagrange multiplier, $*$ denotes complex conjugate and the operator $\Re[\cdot]$ selects the real part of the argument. By fixing $\boldsymbol{w}_D$, minimizing (1.18) with respect to $\boldsymbol{S}_D$ and solving for λ, we obtain

$$\boldsymbol{S}_D = \frac{\boldsymbol{R}^{-1} \boldsymbol{s} \boldsymbol{w}_D^H \boldsymbol{R}_{\bar{w}}^{-1}}{\boldsymbol{w}_D^H \boldsymbol{R}_{\bar{w}}^{-1} \boldsymbol{w}_D \boldsymbol{s}^H \boldsymbol{R}^{-1} \boldsymbol{s}}, \tag{1.19}$$

where $\boldsymbol{R} = E[\boldsymbol{r}(i)\boldsymbol{r}^H(i)]$ and $\boldsymbol{R}_{\bar{w}} = E[\bar{\boldsymbol{w}}_D \bar{\boldsymbol{w}}_D^H]$. By fixing $\boldsymbol{S}_D$, minimizing (1.18) with respect to $\boldsymbol{w}_D$ and solving for λ, we arrive at the expression

$$\boldsymbol{w}_D = \frac{\bar{\boldsymbol{R}}^{-1} \boldsymbol{s}}{\boldsymbol{s}^H \boldsymbol{R}^{-1} \boldsymbol{s}}, \tag{1.20}$$

where $\boldsymbol{R}_D = E[\boldsymbol{S}_D^H \boldsymbol{r}(i)\boldsymbol{r}^H(i)\boldsymbol{S}_D] = E[\boldsymbol{r}_D(i)\boldsymbol{r}_D^H(i)]$, $\boldsymbol{s}_D = \boldsymbol{S}_D^H \boldsymbol{s}$. Note that the expressions in (1.19) and (1.20) are not closed-form solutions for $\boldsymbol{w}_D$ and $\boldsymbol{S}_D$ since (1.19) is a function of $\boldsymbol{w}_D$ and (1.20) depends on $\boldsymbol{S}_D$. Thus, it is necessary to iterate (1.19) and (1.20) with initial values to obtain a solution. Unlike the Krylov subspace-based methods [Goldstein and Reed (1997a)] and the AVF [Pados and Karystinos (2001)] methods, the JIO scheme provides an iterative exchange of information between the low-rank beamformer and the rank-reduction matrix and leads to a simpler adaptive implementation. The key strategy lies in the joint optimization of the filters. The rank D must be set by the designer to ensure appropriate performance or can be estimated via another algorithm. In terms of complexity, the JIO techniques have a computational cost that is related to the optimization algorithm. With recursive LS algorithms the complexity is quadratic with M $((O(M^2))$, whereas the complexity can be as low as linear with DM when stochastic gradient algorithms are adopted [de Lamare and Sampaio-Neto (2009)].

1.4.4 *JIDF-based algorithms*

This section is devoted to presentation of a low-rank space-time beamforming technique based on the joint interpolation, decimation and filtering (JIDF) concept [de Lamare and Sampaio-Neto (2009)]. The JIDF approach allows a designer to compute the parameters of the rank-reduction matrix and the low-rank space-time beamformer with a low complexity. The motivation for designing a rank-reduction matrix based on interpolation and decimation comes from two observations. The first is that rank reduction can be performed by constructing new samples with interpolators and eliminating (decimating) samples that are not useful in the STAP design. The second comes from the structure of the rank-reduction matrix, whose columns are a set of vectors formed by the interpolators and the decimators.

In the JIDF scheme, the number of elements for adaptive processing is substantially reduced, resulting in considerable computational savings and very fast convergence performance for the radar applications. The $M \times 1$ received vector $\boldsymbol{r}(i)$ is processed by a multiple processing branch (MPB) scheme with B branches, where each spatio-temporal processing branch contains an interpolator, a decimation unit and a low-rank space-time beamformer. In the b-th branch, the received vector $\boldsymbol{r}(i)$ is filtered by the interpolator $\boldsymbol{v}_b = [v_{b,0}\ v_{b,1}\ \ldots\ v_{b,I-1}]^T$ with I coefficients, resulting in an interpolated received vector $\boldsymbol{r}_b(i)$ with M samples, which is expressed by

$$\boldsymbol{r}_b(i) = \boldsymbol{V}_b^H \boldsymbol{r}(i), \tag{1.21}$$

where the $M \times M$ Toeplitz convolution matrix is given by

$$\boldsymbol{V}_b = \begin{bmatrix} v_{b,0} & 0 & \ldots & 0 \\ \vdots & v_{b,0} & \vdots & \vdots \\ v_{b,I-1} & \vdots & \vdots & 0 \\ 0 & v_{b,I-1} & \vdots & 0 \\ 0 & 0 & \vdots & \vdots \\ \vdots & \vdots & \vdots & 0 \\ 0 & 0 & \vdots & v_{b,0} \end{bmatrix} \tag{1.22}$$

The vector $\boldsymbol{r}_b(i)$ can be expressed in an alternative way that is useful for the design of the JIDF scheme and is described by

$$\boldsymbol{r}_b(i) = \boldsymbol{V}_b^H \boldsymbol{r}(i) = \boldsymbol{\Re}_0(i) \boldsymbol{v}_b, \tag{1.23}$$

where the $M \times I$ matrix $\boldsymbol{\Re}_o(i)$ with the samples of $\boldsymbol{r}(i)$ has a Hankel structure and is described by

$$\mathfrak{R}_o(i) = \begin{bmatrix} r_0(i) & r_1(i) & \dots r_{\mathrm{I}-1}(i) \\ r_1(i) & r_2(i) & \dots & r_{\mathrm{I}}(i) \\ \vdots & \vdots & \ddots & \vdots \\ r_{M-2}(i) & r_{M-1}(i) & \dots & 0 \\ r_{M-1}(i) & 0 & \dots & 0 \end{bmatrix} \tag{1.24}$$

The dimensionality reduction is performed by a decimation unit with $D \times M$ decimation matrices $\boldsymbol{D}_b$ that transforms $\boldsymbol{r}_I(i)$ into $D \times 1$ vectors $\boldsymbol{r}_{D,b}(i)$ with $b = 1, \dots, B$, where $D = M/L$ is the rank of the resulting system of equations that will be generated and L is the decimation factor. The $D \times 1$ vector $\boldsymbol{r}_{D,b}(i)$ for branch b is expressed by

$$\begin{aligned} \boldsymbol{r}_{D,b} &= \boldsymbol{S}_{D,b}^H \boldsymbol{r}(i) = \boldsymbol{D}_{D,b} \boldsymbol{V}_b^H \boldsymbol{r}(i) \\ &= \boldsymbol{D}_{D,b} \mathfrak{R}_o(i) \boldsymbol{v} \end{aligned} \tag{1.25}$$

where $\boldsymbol{S}_{D,b}$ is the rank-reduction matrix and the vector $\boldsymbol{r}_{D,b}(i)$ for branch b is used in the minimization of the output power for branch b. The output at the end of the JIDF scheme $y(i)$ is selected according to

$$y(i) = y_{b_s}(i) \quad \text{when} \quad b_s = \arg \min_b |y_b|^2 \tag{1.26}$$

where B is a parameter to be set by the designer. For the computation of the parameters of the JIDF scheme, it is fundamental to express the output $y_b(i)$ as a function of the interpolator $\boldsymbol{v}_b$, the decimation matrix $\boldsymbol{D}_{D,b}$ and the low-rank space-time beamformer $\boldsymbol{w}_{D,b}$ as follows:

$$\begin{aligned} y_b(i) &= \boldsymbol{w}_{D,b}^H \boldsymbol{S}_{D,b}^H \boldsymbol{r}(i) \\ &= \boldsymbol{w}_{D,b}^H \boldsymbol{D}_{D,b} \mathfrak{R}_o(i) \boldsymbol{v}, \end{aligned} \tag{1.27}$$

where the expression (1.27) indicates that the dimensionality reduction carried out by the JIDF scheme depends on finding appropriate $\boldsymbol{v}_b$, $\boldsymbol{D}_{D,b}$ and $\boldsymbol{w}_{D,b}$. Unlike the remaining low-rank beamforming techniques, the JIDF is able to substantially reduce the cost of the rank-reduction matrix.

The parameters of the JIDF scheme that performs low-rank space-time MVDR beamforming can be computed by solving the following optimization problem

$$\begin{aligned} [\boldsymbol{w}_{D,\text{opt}}, \boldsymbol{v}_{\text{opt}}, \boldsymbol{D}_{D,b_s}] = \arg \min_{\boldsymbol{w}_{D,b}, \boldsymbol{v}_b, \boldsymbol{D}_{D,b}} & \boldsymbol{w}_{D,b}^H E[\boldsymbol{D}_{D,b} \mathfrak{R}_o(i) \boldsymbol{v} \boldsymbol{v}^H \mathfrak{R}_o^H(i) \boldsymbol{D}_{D,b}^H] \boldsymbol{w}_D \\ \text{subject to } & \boldsymbol{w}_{D,b}^H \boldsymbol{D}_{D,b} \boldsymbol{S}_o \boldsymbol{v}_b = 1, \end{aligned} \tag{1.28}$$

where $\boldsymbol{S}_o$ is $M \times I$ steering matrix with a Hankel structure, which has the same form as $\boldsymbol{\Re}_o(i)$ and is given by

$$\boldsymbol{S}_o(i) = \begin{bmatrix} s_0(i) & s_1(i) & \dots & s_{\mathrm{I}-1}(i) \\ s_1(i) & s_2(i) & \dots & s_{\mathrm{I}}(i) \\ \vdots & \vdots & \ddots & \vdots \\ s_{M-2}(i) & s_{M-1}(i) & \dots & 0 \\ s_{M-1}(i) & 0 & \dots & 0 \end{bmatrix}. \tag{1.29}$$

The constrained optimization in (1.28) can be transformed into an unconstrained optimization problem by using the method of Lagrange multipliers, which results in

$$\mathcal{L}(\boldsymbol{w}_{D,b}, \boldsymbol{v}_b, \boldsymbol{D}_{D,b}) = \boldsymbol{w}_D^H E[\boldsymbol{D}_{D,b}\boldsymbol{\Re}_o(i)\boldsymbol{v}\boldsymbol{v}^H\boldsymbol{\Re}_o^H(i)\boldsymbol{D}_{D,b}^H]\boldsymbol{w}_D + \lambda(\boldsymbol{w}_D^H\boldsymbol{D}_{D,b}\boldsymbol{S}_o\boldsymbol{v} - 1), \tag{1.30}$$

where λ is a Lagrange multiplier.

The strategy to compute the parameters of the low-rank space-time beamformer based on the JIDF scheme is to minimize the cost function with respect to a set of parameters and fix the remaining parameters. By minimizing (1.30) with respect to $\boldsymbol{v}_b$, we obtain

$$\boldsymbol{v}_b = \frac{\boldsymbol{R}_{v,b}^{-1}\boldsymbol{s}_{v,b}}{\boldsymbol{s}_{v,b}^H\boldsymbol{R}_{v,b}^{-1}\boldsymbol{s}_{v,b}}, \tag{1.31}$$

where $\boldsymbol{R}_{v,b} = E[\boldsymbol{r}_{v,b}\boldsymbol{r}_{v,b}^H]$ is the $I \times I$ autocorrelation matrix, $\boldsymbol{r}_{v,b} = \boldsymbol{D}_{D,b}^H\boldsymbol{R}_o^H\boldsymbol{w}_{D,b}$, and $\boldsymbol{s}_{v,b} = \boldsymbol{D}_{D,b}^H\boldsymbol{S}_o^H\boldsymbol{w}_{D,b}$ is the $I\times$ low-rank steering vector. By minimizing (1.30) with respect to $\boldsymbol{w}_{D,b}$, we have

$$\boldsymbol{w}_{D,b} = \frac{\boldsymbol{R}_{w,b}^{-1}\boldsymbol{s}_{w,b}}{\boldsymbol{s}_{w,b}^H\boldsymbol{R}_{w,b}^{-1}\boldsymbol{s}_{w,b}}, \tag{1.32}$$

where $\boldsymbol{R}_{w,b} = E[\boldsymbol{r}_{w,b}\boldsymbol{r}_{w,b}^H]$ is the $D \times D$ autocorrelation matrix, $\boldsymbol{r}_{w,b} = \boldsymbol{D}_{D,b}\boldsymbol{R}_o\boldsymbol{v}_b$, and $\boldsymbol{s}_{w,b} = \boldsymbol{D}_{D,b}\boldsymbol{S}_o\boldsymbol{v}_b$ is the $D\times$ low-rank steering vector. In order to compute $\boldsymbol{v}_b$ and $\boldsymbol{w}_{D,b}$, a designer needs to iterate them for each processing branch b.

The decimation matrix $\boldsymbol{D}_{D,b}$ is selected to minimize the square of the output of the beamformer $y_b(i)$ obtained for all the B branches

$$\boldsymbol{D}_{D,b} = \boldsymbol{D}_{D,b_{\mathrm{s}}}[i] \ \text{ when } \ b_{\mathrm{s}} = \arg\min_{1\leq b\leq B} |y_b(i)|^2, \tag{1.33}$$

The design of the decimation matrix $\boldsymbol{D}_{D,b}$ imposes constraints on the values of the elements of the matrix such that they only take the value zero or one. Since the optimal approach for the design of $\boldsymbol{D}_{D,b}$ corresponds to an exhaustive search, we consider a suboptimal technique that employs pre-stored patterns. The decimation scheme employs a structure formed in the following way

$$\boldsymbol{S}_{D,b} = [\boldsymbol{\phi}_{b,1}\ \boldsymbol{\phi}_{b,2}\ \boldsymbol{\phi}_{b,D}], \tag{1.34}$$

where $\boldsymbol{\phi}_{b,d}$ is an $M \times 1$ vector composed of a single one and zeros as described by

$$\boldsymbol{\phi}_{b,d} = [\underbrace{0,\ \ldots,\ 0}_{z_{b,d}},\ 1,\ \underbrace{0,\ \ldots,\ 0}_{M-z_{b,d}-1}], \tag{1.35}$$

where $z_{b,d}$ is the number of zeros before the only element equal to one. We set the value of $z_{b,d}$ in a deterministic way which can be expressed as

$$z_{b,d} = \frac{M}{D} \times (d-1) + (b-1). \tag{1.36}$$

It is necessary to iterate (1.31), (1.32) and (1.33) in an alternated form (one followed by the other) with an initial value to obtain a solution. The expectations can be estimated either via time averages or by instantaneous estimates and with the help of adaptive algorithms.

1.5 Sparsity-aware Beamforming Algorithms

This section considers space-time beamforming algorithms that exploit the sparsity encountered in the data processed by radar systems. In particular, the motivation for exploiting the sparsity of data vectors observed by radar systems is given and a brief discussion on the suitability of sparsity-aware algorithms for radar applications is provided. A general approach to design space-time beamforming algorithms based on the l_1-norm regularization is described. The main principle is to employ a reduced number of weights to suppress the clutter and the jamming signals encountered in radar applications.

Recently, motivated by compressive sensing (CS) techniques used in radar, several authors have considered CS ideas for moving target indication (MTI) and STAP problems [Maria and Fuchs (2006)]-[Selesnick (2010)]. The core notion in CS is to regularize a linear inverse problem by including prior knowledge that the signal of interest is sparse [Parker and Potter (2010)]. These works on space-time beamforming techniques based on CS rely on the recovery of the clutter power in angle-Doppler plane, which is usually carried out via convex optimization tools. However, these methods are based on linear programming and have a quite high computational complexity ($O(K^3)$), where K is the dimension of the angle-Doppler plane. In this section, we describe the concept of a sparsity-aware STAP (SA-STAP) algorithm that can improve the detection capability using a small number of snapshot. To overcome the high complexity of the CS-STAP type algorithm, we design the STAP algorithm with another strategy, by imposing the sparse regularization to the minimum variance (MV) cost function. Since the interference variance has a low rank property, we assume that a number of samples of the data cube are not meaningful for processing and the optimal STAP filter weight vector is sparse, or nearly sparse. Then, we exploit this feature by using a l_1-norm regularization. With this motivation, the STAP algorithm design becomes a mixed l_1-norm and l_2-norm optimization problem.

The conventional space-time beamforming algorithms do not exploit the sparsity of the received signals. In this exposition, it is assumed that a number of samples of the data cube are not meaningful for processing and a reduced number of active weights of the space-time beamformer can effectively suppress the clutter and the jamming signals. Specifically, a sparse regularization is imposed to the MV cost function. Thus, the space-time beamformer design can be described as the following optimization problem

$$\begin{aligned} \boldsymbol{w}_{\text{opt}} &= \arg\min_{\boldsymbol{w}} \boldsymbol{w}^H \boldsymbol{R} \boldsymbol{w} \\ \text{subject to } & \boldsymbol{w}^H \boldsymbol{s} = 1 \text{ and } ||\boldsymbol{w}||_1 = 0, \end{aligned} \tag{1.37}$$

where the objective of the l_1-norm regularization is to force the components of the space-time beamformer $\boldsymbol{w}$ to zero [Angelosante *et al.* (2010)]. This problem can be solved using the method of Lagrange multipliers, which results in the following unconstrained cost function

$$\mathcal{L}(\boldsymbol{w}, \alpha, \lambda) = \boldsymbol{w}^H \boldsymbol{R} \boldsymbol{w} + \alpha(\boldsymbol{w}^H \boldsymbol{s} - 1) + \lambda(||\boldsymbol{w}||_1), \tag{1.38}$$

The unconstrained cost function above is convex, however, it is non-differentiable which makes it difficult for one to use the method of Lagrange Multipliers directly and obtain an expression for the space-time beamformer. To this end, the following approximation to the regularization term is employed

$$||\boldsymbol{w}||_1 \approx \boldsymbol{w}^H \boldsymbol{\Lambda} \boldsymbol{w}, \tag{1.39}$$

where

$$\boldsymbol{\Lambda} = \text{diag}\left(\frac{1}{|w_1| + \epsilon} \ \frac{1}{|w_2| + \epsilon} \ \cdots \ \frac{1}{|w_M| + \epsilon} \right), \tag{1.40}$$

where ϵ is a small positive constant. Simultaneously, we assume that the partial derivative of $\boldsymbol{w}^H \boldsymbol{\Lambda} \boldsymbol{w}$ with respect to $\boldsymbol{w}^*$ is given by

$$\frac{\partial \boldsymbol{w}^H \boldsymbol{\Lambda} \boldsymbol{w}}{\partial \boldsymbol{w}^*} \approx \boldsymbol{\Lambda} \boldsymbol{w}. \tag{1.41}$$

With the development above, an approximation to the unconstrained cost function can be employed as described by

$$\mathcal{L}(\boldsymbol{w}, \alpha, \lambda) \approx \boldsymbol{w}^H \boldsymbol{R} \boldsymbol{w} + \alpha(\boldsymbol{w}^H \boldsymbol{s} - 1) + \lambda \boldsymbol{w}^H \boldsymbol{\Lambda} \boldsymbol{w}, \tag{1.42}$$

By computing the gradient terms with respect to $\boldsymbol{w}^*$ and α and equating them to zero, we obtain the following expression for the space-time beamformer

$$\boldsymbol{w} = \frac{(\boldsymbol{R} + \lambda \boldsymbol{\Lambda})^{-1} \boldsymbol{s}}{\boldsymbol{s}^H (\boldsymbol{R} + \lambda \boldsymbol{\Lambda})^{-1} \boldsymbol{s}}. \tag{1.43}$$

Comparing (1.43) with the conventional optimal space-time beamformer in (1.8), we find that there is an additional term $\lambda \boldsymbol{\Lambda}$ in the inverse of the interference covariance matrix $\boldsymbol{R}$, which is due to the l_1-norm regularization. The term λ is a positive scalar which provides a trade-off between the sparsity and the output interference power. The larger the chosen λ, the more components are shrunk to zero [Zibulevsky and Elad (2010)]. It should also be remarked that the expression for the beamformer in (1.43) is not a closed-form solution since $\boldsymbol{\Lambda}$ is a function of $\boldsymbol{w}$. Thus it is necessary to develop an iterative procedure to compute the parameters of the space-time beamformer.

1.6 Knowledge-aided Beamforming Algorithms

Although STAP techniques are considered efficient tools for detection of slow targets by airborne radar systems in strong clutter environments [Klemm (2002)], due to the very large number of degrees of freedom (DoFs) conventional space-time beamformers have a slow convergence and require about twice the DoFs of the independent and identically distributed (IID) training snapshots to yield an average performance loss of roughly 3dB [Ward (1994)]. In real scenarios, it is hard to obtain so many IID training snapshots, especially in heterogeneous environments. Low-rank [Guerci (2000)]-[de Lamare and Sampaio-Neto (2009)] and sparsity-aware [Maria and Fuchs (2006)]-[Yang *et al.* (2011)] methods have been considered to counteract the slow convergence of the conventional space-time beamformers. Nevertheless, there are other alternatives to improve the training of STAP algorithms and improve their performance. Recently developed knowledge-aided (KA) STAP algorithms have received a growing interest and become a key concept for the next generation of adaptive radar systems [Wicks *et al.* (2006)]-[Fa *et al.* (2010)]. The core idea of KA-STAP is to incorporate prior knowledge, provided by digital elevation maps, land cover databases, road maps, the Global Positioning System (GPS), previous scanning data and other known features, to compute estimates of the clutter covariance matrix with high accuracy [**?**]. Prior work on KA-STAP algorithms include the exploitation of prior knowledge of the clutter ridge to form the STAP filter weights [**?**], use of prior knowledge about the terrain [Capraro *et al.* (2006)] and prior knowledge about the covariance matrix of the clutter and the jamming signals [Blunt *et al.* (2006)]-[Fa *et al.* (2010)].

In this section, we discuss a strategy to mitigate the deleterious effects of the heterogeneity in the secondary data, which makes use of a priori knowledge of the clutter covariance matrix and has recently gained significant attention in the literature [Wicks *et al.* (2006)]-[Fa *et al.* (2010)]. In KA-STAP techniques, there are two basic tasks that need to be addressed: the first one is how to obtain prior knowledge from the terrain knowledge of the clutter and how to estimate the real interference covariance matrix with the prior knowledge [Wicks *et al.* (2006)]-[Capraro *et al.* (2006)] and the second is how to apply the covariance matrix estimates in the design of the space-time beamforming algorithm [Blunt *et al.* (2006)]-[Fa *et al.* (2010)]. We first review how a designer can obtain prior knowledge of the clutter and employ this knowledge to build a known covariance matrix $\boldsymbol{R}_o$. Then, we present a method to combine this prior knowledge with commonly used estimation techniques to compute the covariance matrix of the received vector $\boldsymbol{r}(i)$, resulting in a combined covariance matrix estimate $\hat{\boldsymbol{R}}_c$ for use in the space-time beamformer that is more accurate and has an enhanced performance.

The optimal space-time beamformer employs the following expression to compute its parameters

$$\boldsymbol{w} = \frac{\hat{\boldsymbol{R}}^{-1}\boldsymbol{s}}{\boldsymbol{s}^H\hat{\boldsymbol{R}}^{-1}\boldsymbol{s}}, \tag{1.44}$$

where an estimate of the covariance matrix is typically obtained by

$$\hat{\boldsymbol{R}} = \frac{1}{K}\sum_{k=1}^{K}\boldsymbol{r}(k)\boldsymbol{r}^H(k), \tag{1.45}$$

where $\boldsymbol{r}(k)$ is taken from secondary data. The estimate $\hat{\boldsymbol{R}}$ can be sufficiently accurate when K is at least twice as great as M [Brennan and Reed (1973)] and the training samples are assumed i.i.d. However, it is by now well understood that the clutter environments are often heterogeneous and this leads to performance degradation on space-time beamforming. KA-STAP techniques can significantly help to combat the heterogeneity [Stoica *et al.* (2008)].

With KA techniques the covariance matrix $\boldsymbol{R}_c$ is estimated by combining an initial guess of the covariance matrix $\boldsymbol{R}_o$ derived from the digital terrain database or the data probed by radar in previous scans, and the sample average covariance matrix estimate in the present scan $\hat{\boldsymbol{R}}$ so that

$$\boldsymbol{R}_c = \alpha \boldsymbol{R}_o + (1-\alpha)\hat{\boldsymbol{R}}, \tag{1.46}$$

where $0 \leq \alpha \leq 1$. Alternatively, this principle can be applied to the inverse of the covariance matrix estimate

$$\boldsymbol{R}_c^{-1} = \eta \boldsymbol{R}_o^{-1} + (1-\eta)\hat{\boldsymbol{R}}^{-1}, \tag{1.47}$$

where $0 \leq \eta \leq 1$.

In order to compute the parameter η, we need to consider the optimization problem

$$dd \tag{1.48}$$

1.7 Simulations

1.8 Concluding Remarks

Bibliography

R. Klemm, Principle of space-time adaptive processing, IEE Press, Bodmin, UK, 2002.

W. L. Melvin, A STAP overview, IEEE Aero. .Elec. Syst. Mag., vol. 19, no. 1, pp. 1935, 2004.

L. E. Brennan and I. S. Reed, Theory of adaptive radar, IEEE Trans. Aero. Elec. Syst., vol. AES-9, no. 2, pp. 237252, 1973.

I. S. Reed, J. D. Mallett, and L. E. Brennan, Rapid convergence rate in adaptive arrays, IEEE Trans. Aero. Elec. Syst., vol. AES-10, no. 6, pp. 853863, 1974.

E. J. Kelly, An adaptive detection algorithm, IEEE Trans. Aero. Elec. Syst., vol. AES-22, no. 2, pp. 115127, 1986.

A. M. Haimovich and Y. Bar-Ness, An eigenanalysis interference canceler, IEEE Trans. Sig. Process., vol. 39, no. 1, pp. 7684, 1991.

F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, A CFAR adaptive matched filter detector, IEEE Trans. Aero. Elec. Syst., vol. 28, no. 1, pp. 208216, Jan 1992.

J. Ward, Space-time adaptive processing for airborne radar,, Tech. Rep. 1015, MIT Lincoln lab., Lexington, MA, Dec. 1994.

A. Haimovich, The eigencanceler: adaptive radar by eigenanalysis methods, IEEE Trans. Aero. Elec. Syst., vol. 32, no. 2, pp. 532542, 1996.

J. S. Goldstein and I. S. Reed, Reduced-rank adaptive filtering, IEEE Trans. Sig. Process., vol. 45, no. 2, pp. 492496, 1997.

J. S. Goldstein and I. S. Reed, Theory of partially adaptive radar, IEEE Trans. Aero. Elec. Syst., vol. 33, no. 4, pp. 13091325, 1997.

Y.-L. Gau and I.S. Reed, An improved reduced-rank CFAR space-time adaptive radar detection algorithm, IEEE Trans. Sig. Process., vol. 46, no. 8, pp. 21392146, Aug 1998.

I. S. Reed, Y. L. Gau, and T. K. Truong, CFAR detection and estimation for STAP radar, IEEE Trans. Aero. Elec. Syst., vol. 34, no. 3, pp. 722 735, 1998.

J. S. Goldstein, I. S. Reed, and P. A. Zulch, Multistage partially adaptive STAP CFAR detection algorithm, IEEE Trans. Aero. Elec. Syst., vol. 35, no. 2, pp. 645661, 1999.

J. R. Guerci, J. S. Goldstein, and I. S. Reed, Optimal and adaptive reduced-rank STAP, IEEE Trans. Aero. Elec. Syst., vol. 36, no. 2, pp. 647663, 2000.

S. Haykin, Adaptive Filter Theory, NJ: Prentice-Hall, 4th, ed, 2002.

J. S. Goldstein and I. S. Reed, Subspace selection for partially adaptive sensor array processing, IEEE Trans. Aero. Elec. Syst., vol. 33, no. 2, pp. 539544, 1997.

J. S. Goldstein, I. S. Reed, and L. L. Scharf, A multistage representation of the wiener filter based on orthogonal projections, IEEE Trans. Inf. Theory, vol. 44, no. 7, pp. 29432959, 1998.

D. A. Pados and S. N. Batalama, Joint space-time auxiliary-vector filtering for DS/CDMA systems with antenna arrays, IEEE Trans. Commun., vol. 47, no. 9, pp. 14061415, 1999.

D. A. Pados and G. N. Karystinos, An iterative algorithm for the computation of the MVDR filter, IEEE Trans. Sig. Process.], vol. 49, no. 2, pp. 290300, Feb 2001.

D. A. Pados, G. N. Karystinos, S. N. Batalama, and J. D. Matyjas, Short-data-record adaptive detection, 2007 IEEE Radar Conf., pp. 357 361, 17-20 April 2007.

H. Wang, and L. Cai, On adaptive spatial-temporal processing for airborne surveillance radar systems, IEEE Trans. Aero. Elec. Syst., vol. 30, no. 3, 660670, 1994.

R. S. Adve, T. B. Hale, and M. C. Wicks, Practical joint domain localised adaptive processing in homogeneous and nonhomogeneous environments. Part 1: Homogeneous environments., IEE Proceedings Radar, Sonar and Navigation, vol. 147, no. 2, 5765, 2000.

R. S. Adve, T. B. Hale, and M. C. Wicks, Practical joint domain localised adaptive processing in homogeneous and nonhomogeneous environments. Part 2: Nonhomogeneous environments., IEE Proceedings Radar, Sonar and Navigation, vol. 147, no. 2, 6674, 2000.

R. C. de Lamare and R. Sampaio-Neto, Reduced-rank adaptive filtering based on joint iterative optimization of adaptive filters, IEEE Sig. Proc. Lett., vol. 14, no. 12, pp. 980983, 2007.

R. Fa, R. C. de Lamare, and D. Zanatta-Filho, Reduced-rank STAP algorithm for adaptive radar based on joint iterative optimization of adaptive filters, in Conf. Record of the Fourty-Second Asilomar Conf. Sig. Syst. Comp., 2008.

R. C. de Lamare and R. Sampaio-Neto, Adaptive reduced-rank mmse parameter estimation based on an adaptive diversity-combined decimation and interpolation scheme, *in Proc. IEEE Int. Conf. Acous. Speech Sig. Process.*, 1520 April 2007, vol. 3, pp. III1317III1320.

R. C. de Lamare, and R. Sampaio-Neto, “Adaptive reduced-rank processing based on joint and iterative interpolation, decimation, and filtering”, *IEEE Trans. Sig. Process.*, vol.57, no.7, pp. 2503-2514, July 2009.

S. Applebaum and D. Chapman, “Adaptive arrays with main beam constraints”, *IEEE Trans. on Ant. Prop.*, vol. 24, no. 5, pp. 650662, 1976.

G. H. Golub and C. F. van Loan, Matrix Computations, Wiley, 2002.

M. R. Hestenes and E. Stiefel, Methods of Conjugate Gradients for Solving Linear

Systems, Journal of Research of the National Bureau of Standards, vol. 49, no. 6, pp. 409436, December 1952.

C. Lanczos, Solution of Systems of Linear Equations byMinimized Iterations, Journal of Research of the National Bureau of Standards, vol. 49, no. 1, pp. 33 53, July 1952.

W. E. Arnoldi, The Principle of Minimized Iterations in the Solution of the Matrix Eigenvalue Problem, Quarterly of Applied Mathematics, vol. 9, no. 1, pp. 1729, January 1951.

W. Xiao and M. L. Honig, "Large System Transient Behavior of Adaptive Least Squares Algorithms", *IEEE Transactions on Information Theory*, Vol. 51, No. 7, pp. 2447-2474, July 2005.

S. Maria and J. J. Fuchs, Application of the global matched filter to STAP data an efficient algorithmic approach, *IEEE Int. Conf. Acoust. Speech and Signal Processing*, pp. 14-19, 2006.

K. Sun, H. Zhang, G. Li, H. Meng and X. Wang, A novel STAP algorithm using sparse recovery technique, *Proc. of IGARSS*, pp.336-339, 2009.

J. T. Parker and L. C. Potter, A Bayesian perspective on sparse regularization for STAP post-processing, *IEEE Radar Conf.*, pp.1471- 1475, May 2010.

I. W. Selesnick, S. U. Pillai, K. Y. Li and B. Himed, Angle-Doppler processing using sparse regularization, *IEEE Int. Conf. Acoust. Speech and Signal Processing*, pp.2750-2753, 2010.

M. Zibulevsky and M. Elad, L1-L2 optimization in signal and image processing, *IEEE Sig. Proc. Mag.*, vol. 27, no. 3, pp. 76-88, May 2010.

D. Angelosante, J. A. Bazerque and G. B. Giannakis, Online adaptive estimation of sparse signals: where RLS meets the l1-norm, *IEEE Trans. Sig. Proc.*, vol. 58, no. 7, pp. 3436-3446, 2010.

Z. Yang, R. C. de Lamare and X. Li, "L_1-Regularized STAP Algorithms with a Generalized Sidelobe Canceler Architecture for Airborne Radar", *IEEE Transactions on Signal Processing*, vol. 60, no. 2, pp. 674- 686, February 2011.

M. C. Wicks, M. Rangaswamy, R. Adve, and T. B. Hale, "Space-time adaptive processing: a knowledge-based perspective for airborne radar," *IEEE Sig. Proc. Mag.*, vol. 23, no. 1, 2006, pp. 983-996,.

W. L. Melvin and J. R. Guerci, "Knowledge-aided signal processing: a new paradigm for radar and other advanced sensors," *IEEE Trans. Aero. Elec. Syst.*, vol. 42, no. 3, 2006, pp. 1021-1042.

W. L. Melvin and G. A. Showman, "An approach to knowledge-aided covariance estima- tion," *IEEE Trans. Aero. Elec. Syst.*, vol. 42, no. 3, 2006, pp. 1021-1042.

E. Conte, A. De Maio, A. Farina, and G. Foglia, "Design and analysis of a knowledge-aided radar detector for doppler processing," *IEEE Trans. Aero. Elec. Syst.*, vol. 42, no. 3, 2006, pp. 1058-1079.

C. T. Capraro, G. T. Capraro, I. Bradaric, D. D. Weiner, M. C. Wicks, and W. J. Baldygo, "Implementing digital terrain data in knowledge-aided space-time adaptive processing," *IEEE Trans. Aero. Elec. Syst.*, vol. 42, no. 3, 2006. pp. 1080-1099.

S. D. Blunt, K. Gerlach, and M. Rangaswamy, “STAP using knowledge-aided covariance estimation and the FRACTA algorithm,” *IEEE Trans. Aero. Elec. Syst.*, vol. 42, no. 3, 2006, pp. 1043-1057.

P. Stoica, Li Jian, Zhu Xumin, and J. R. Guerci, “On using a priori knowledge in space- time adaptive processing,” *IEEE Trans. Sig. Proc.*, vol. 56, no. 6, 2008, pp. 2598-2602.

R. Fa, R. C. de Lamare, and V. H. Nascimento, “Knowledge-Aided STAP Al- gorithm using Convex Combination of Covariance Matrix Inversions for Heterogeneous Clutter”, *Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2010)*, 2010

R. Fa and R. C. de Lamare, “ Knowledge-Aided Reduced-Rank STAP for MIMO Radar Based on Joint Iterative Optimization of Adaptive Filters, *Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2010)*, 2010.

Index

BIBTEX, 20, 23

acknowledgments, 20
appendix, 20, 39

bibliography, 20
  author-date, 20
  numbered, 21
boxed text, 20

citation, 17
  author-date, 17
  numbered, 18
    bracketed, 18
    superscript, 18

equations
  display, 6
  inline, 6

floats, 7
  figures, 9
    landscape, 12
  tables, 7
    landscape, 11
    notes to, 8
    rules, 9
footnote, 19

index, 22, 25
  multiple indexes, 26

lists, 5
  bulleted and unnumbered, 5
  numbered and lettered, 5
    Alphabetical (A, B, C...), 5
    alphabetical (a, b, c...), 5
    arabic (1, 2, 3...), 5
    Roman (I, II, III...), 5
    roman (i, ii, iii...), 5

master file, 3

packages, 1

quote, 19

running head, 19

subdocuments
  chapters, 3
  parts, 3
  sectional units, 4
    section, 4
    subsection, 4
    subsubsection, 4
    unnumbered section, 4

theorems, 14