A MODIFIED CONJUGATE GRADIENT ALGORITHM FOR LOW SAMPLE SUPPORT IN STAP RADAR

Hocine BELKACEMI, Sylvie MARCOS

Laboratoire des signaux et systèmes
CNRS/Supélec,
91192 Gif sur Yvette, FRANCE
{belkacemi, marcos}@lss.supelec.fr

Marc LESTURGIE

ONERA-DEMR
Chemin de la Hunière et des Joncherettes
91761 Palaiseau, FRANCE
lesturgie@onera.fr

ABSTRACT

The Conjugate Gradient (CG) has been shown recently to be equivalent to the MultiStage Wiener Filter (MSWF) [1] which is an effective tool for interference suppression in space-time adaptive processing radar. In this paper, we give further insight on the interconnection between the MSWF and the CG. We propose a modified version of the CG for low sample support where we use the forward/backward (f/b) averaging for estimating the covariance estimation. The new algorithm takes benefits of the CG for rank compression and of the (f/b) subaperture smoothing for sample support compression [2]. The effectiveness of the algorithm is demonstrated through simulations.

1. INTRODUCTION

Space **T**ime **A**daptive **P**rocessing (STAP) is a two dimensional adaptive filtering algorithm that combines signals from multiple channels N and pulses M to suppress interferences (clutter and jamming) in airborne or space borne radar [3]. In the optimum processor the weight vector which maximizes the signal to interference plus noise ratio SINR is given by $\mathbf{w}_{opt} = \kappa\mathbf{R}^{-1}\mathbf{s}$ where, $\mathbf{R}$ is the covariance matrix of the interferences and κ a constant gain. Since the covariance matrix is not known, Reed et al. [4] proposed the Sample Matrix Inversion (SMI) based on replacing $\mathbf{R}$ by the sample average estimate $\hat{\mathbf{R}}$. In general, there are two computational criteria that practical implementation should ideally possess to achieve sufficient interference suppression; rapid convergence (i.e. small sample support size) and low computationally complexity. Thus sample matrix inversion (SMI) is a poor technique for the weight computation because it converges slowly requiring a wide sense stationary (WSS) sample support of $K = 2NM$ samples to obtain an SINR performance within 3dB of the optimal one in the Gaussian case, with a computational load of $O((NM)^3)$. For high interference to noise ratio (INR) cases, the principal components (PC) approach and an appropriately diagonally loaded sample covariance estimate (LSMI) can reduce the sample support requirement to $O(2r)$, where r is the dimension of the dominant subspace [5, 6]. However, this still have a high computational cost due to matrix inversion and eigendecomposition. Recently, reduced-rank filters have attracted a considerable amount of research due to their satisfactory adaptive performance and low complexity. Goldstein et *al.* [7] proposed the Multistage Wiener Filter (MSWF) which has, in one hand, a considerable rank compression compared to the principal component based filter, and in another hand, a linear computational cost $O(NM)$ when using efficient architectures [8, 9]. In [10] Hiemstra has shown through extensive Monte Carlo simulations that the MSWF has practically the same sample support requirement as the PC method of $O(2r)$. In [9], the authors have showed that the MSWF can be identified to be the solution of the Wiener-Hopf equation in the Krylov subspace defined by the covariance matrix of the observation and the cross correlation vector of the observation and the desired signal. Recently, the authors in [1, 11] have shown the subspace equivalence of the conjugate gradient to the MSWF. In almost of these works, the applications take place in the context of communication systems. In this paper we give further insight on the interconnection between the CG and MSWF. We analyse the performance of the CG for interference suppression in STAP-radar. We also propose an iterative version of CG for low support data using the forward/backward (f/b) subaperture smoothing [2] for estimating the covariance matrix. This paper is organized as follows. In section 2, we introduce the data model. The concept of Krylov subspace projection algorithms along with the MSWF is presented in section 3. The Conjugate Gradient algorithm and the low sample support version are presented in section 4 and 5 respectively. After simulations in section 6, few concluding remarks are drawn in section 7.

2. DATA MODEL

We consider a pulsed Doppler radar mounted on an airborne platform moving at constant speed v_p. The radar antenna is a uniformly spaced linear array consisting of N elements. The radar transmits a coherent burst of M pulses at a constant pulse repetition frequency (PRF) $f_r = 1/T_r$ over a set of range directions of interest. The returned space-time snapshot may consist of the target echo and interferences such as jammer, clutter and thermal noise, and is given by [3]

$$\mathbf{x} = \alpha_t \mathbf{v}_t + \mathbf{n} \tag{1}$$

where

- $\mathbf{x} = [x_1, ..., x_{MN}]^T$ is the array output vector
- α_t and $\mathbf{v}_t \equiv \mathbf{v}(\varpi_t, \nu_t)$ are the complex target attenuation factor and target steering signal vector, respectively, associated with the spatial and Doppler parameters ϖ_t , ν_t so that [3]

  $$\mathbf{v}(\varpi_t, \nu_t) = \mathbf{b}(\varpi_t) \otimes \mathbf{a}(\nu_t) \tag{2}$$

  with

- $\mathbf{a}(\nu_t) = [1 \;\; e^{j2\pi\nu_t} \cdots \;\; e^{j2\pi(M-1)\nu_t}]^T$ is the temporal steering vector ($\nu_k = \frac{f_t}{f_r}$, f_t is the target Doppler frequency).
- $\mathbf{b}(\varpi_t) = [1 \;\; e^{j2\pi\varpi_t} \cdots \;\; e^{j2\pi(N-1)\varpi_t}]^T$ is the spatial steering vector ($\varpi_t = \frac{d}{\lambda} sin(\theta_t)$, d is the element separation distance, λ is the wavelength and θ_t is the target azimuth angle).

- $\mathbf{n} = [n_1, ..., n_{MN}]^T$

The interference vector $\mathbf{n}$ is supposed to be due to clutter $\mathbf{n}_c$, jammer $\mathbf{n}_j$ and thermal noise $\mathbf{n}_w$ supposed to be spatially and temporally white

$$\mathbf{n} = \mathbf{n}_c + \mathbf{n}_w + \mathbf{n}_j \quad (3)$$

If we suppose that these components are uncorrelated complex gaussian random variables, then the interference (clutter + jammer + thermal noise) space-time covariance matrix $\mathbf{R}$ is

$$\mathbf{R}_i = E\{\mathbf{n}\mathbf{n}^H\} = \mathbf{R_c} + \sum_{i=1}^{N_j} \mathbf{R}_j(i) + \sigma^2 \mathbf{I}_{NM} \quad (4)$$

where N_j is the number of jammers, $\mathbf{R}_j(i)$ is the covariance matrix of the i^{th} jammer, $\mathbf{R_c}$ is the clutter covariance matrix, σ^2 is the noise variance and $\mathbf{I}_{NM}$ denotes the identity matrix of dimension $NM \times NM$. The jamming plus clutter covariance matrix is generally of low rank r ($r = N + (\beta + N_j)(M-1)$)[3]. This rank deficiency is exploited [5] to derive fast STAP algorithm.

3. MULTISTAGE WIENER FILTER (MSWF)

The Multistage Wiener Filter was introduced in [7], as shown in figure 1 for rank $D = 3$, it is divided into two distinct recursions. In the first one (forward recursion or analysis stage) (see table 1) the filter decomposes the observed process, $\mathbf{x}$, by a sequence of orthogonal projections [7]. Rank reduction is accomplished by truncating this decomposition at the desired number of stages, D. The resulting transformation is given by[1]

$$\begin{aligned} \mathbf{T}_{mwf} &= \left[\begin{array}{cccccc} \mathbf{h}_1 & \mathbf{B}_1^H\mathbf{h}_2 & \mathbf{B}_1^H\mathbf{B}_2^H\mathbf{h}_3 & \cdots & \prod_{i=1}^{D-1}\mathbf{B}_i^H\mathbf{h}_D \end{array}\right] \\ &= \left[\begin{array}{ccccc} \mathbf{h}_1 & \mathbf{h}_2 & \mathbf{h}_3 & \cdots & \mathbf{h}_D \end{array}\right] \end{aligned} \quad (5)$$

where $\mathbf{h}_i = \frac{\mathbf{r}_{\mathbf{x}_{i-1}d_{i-1}}}{\|\mathbf{r}_{\mathbf{x}_{i-1}d_{i-1}}\|}$ is the normalized cross correlation between the reference signal d_{i-1} and the observed data $\mathbf{x}_{i-1}$ on each previous stage and $\mathbf{B}_i = \mathbf{I} - \mathbf{h}_i\mathbf{h}_i^H$ are the blocking matrices.

It is shown in [9, 12] that the weight vector $\mathbf{w}$ lies in the D Krylov subspace [11] $\mathcal{K}^D(\mathbf{h}_1, \mathbf{R}_{\mathbf{x}_0})$ and the following identity holds:

$$\mathbf{T}_{mwf} \equiv span\{\mathbf{h}_1, \mathbf{R}_{\mathbf{x}_0}\mathbf{h}_1, \mathbf{R}^2_{\mathbf{x}_0}\mathbf{h}_1, \ldots, \mathbf{R}^{D-1}_{\mathbf{x}_0}\mathbf{h}_1\} \equiv \mathcal{K}^D(\mathbf{h}_1, \mathbf{R}_{\mathbf{x}_0}). \quad (6)$$

In this stage the covariance matrix is transformed to a tridiagonal and is given by

$$\begin{aligned} \mathbf{R}_d &= \mathbf{T}_{mwf}\mathbf{R}_{\mathbf{x}_0}\mathbf{T}^H_{mwf} \quad (7) \\ &= \left[\begin{array}{ccccc} \sigma^2_{d_0} & \delta_1 & 0 & \ldots & 0 \\ \delta_1 & \sigma^2_{d_1} & \delta_2 & \ldots & 0 \\ 0 & \delta_2 & \sigma^2_{d_2} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \delta_{NM-1} \\ 0 & 0 & \ldots & \delta_{NM-1} & \sigma^2_{d_{NM-1}} \end{array}\right] \quad (8) \end{aligned}$$

[1] The last equality is due to the fact that the $\mathbf{h}$'s are orthogonal.

Forward Recursion (*Analysis Stage*)
for $i = 1$ to D do

$$\begin{aligned} \mathbf{r}_{x_i d_i}(n) &= E\{x_i(n)\, d_i^*(n)\} \\ \delta_i(n) &= \|\mathbf{r}_{x_i d_i}(n)\| \\ \mathbf{h}_{i+1} &= \mathbf{r}_{x_i d_i}(n) / \delta_i(n) \\ \mathbf{B}_{i+1} &= null(\mathbf{h}_{i+1}) = \mathbf{I} - \mathbf{h}^H_{i+1}\mathbf{h}_{i+1} \\ d_{i+1}(n) &= \mathbf{h}^H_{i+1}\mathbf{x}_i(n) \\ \sigma^2_{d_i} &= E\{|d_i(n)|^2\} \\ \mathbf{x}_{i+1}(n) &= \mathbf{B}_{i+1}\mathbf{x}_i(n) \end{aligned}$$

End for

Backward Recursion (*Synthesis Stage*)
for $i = D$ to 1 do

$$\begin{aligned} w_i &= \delta_i / \xi_i \\ \epsilon_{i-1}(n) &= d_{i-1}(n) - w_i\epsilon_i(n) \\ \xi_{i-1} &= E\{|\epsilon_{i-1}|^2\ \} \\ &= \sigma^2_{d_{i-1}} - \delta_i^2/\xi_i \end{aligned}$$

End for

Table 1. MSWF Recursion equations

where $\sigma^2_{d_i}$ and δ_i are defined in table 1.

After the forward recursion is completed the MSWF computes a nested chain of scalars which represent the solutions to of the transformed Wiener filter. The weight vector is then given by

$$\mathbf{w} = \mathbf{s} - \mathbf{B}_0^H\mathbf{T}_{mwf}\mathbf{w}_{mwf} \quad (9)$$

$$\mathbf{w}_{mwf} = \mathbf{R}_d^{-1}\mathbf{r}_{\mathbf{d}d_0} \quad (10)$$

where $\mathbf{B}_0 = \mathbf{I} - \mathbf{s}\mathbf{s}^H$, $\mathbf{R}_d = E[\mathbf{d}\mathbf{d}^H]$ with $\mathbf{d} = [\begin{array}{cccc} d_1 & d_2 & \cdots & d_r \end{array}]$ and $\mathbf{r}_{\mathbf{d}d_0} = E[\mathbf{d}d_0^H]$
The complete set of the MSWF recursions are given in table 1

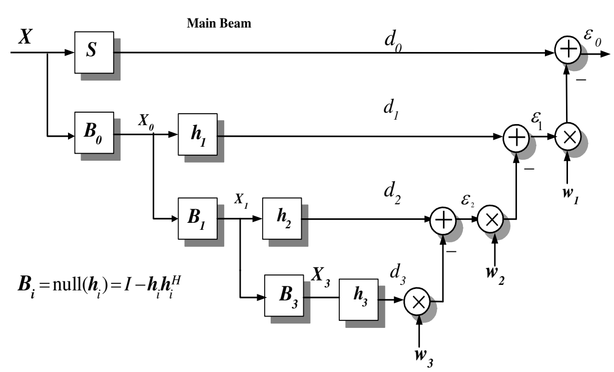

Fig. 1. Multistage Wiener Filter (Rank D=3)

4. CONJUGATE GRADIENT

The method of conjugate gradient is an iterative technique for solving symmetric positive definite linear systems. We consider the Wiener-Hopf equation

$$\mathbf{R}_{\mathbf{x}_0}\mathbf{w} = \mathbf{r}_{\mathbf{x}d} \quad (11)$$

$$
\begin{array}{l}
\mathbf{w}_0(n) = 0,\ \mathbf{p}_1(n) = \mathbf{g}_0(n) = \mathbf{r}_{\mathbf{x}_0\mathbf{d}_0} \\
\rho_0(n) = \mathbf{g}_0(n)^H \mathbf{g}_0(n) \\
\mathbf{R}(n) = \mu\mathbf{R}(n-1) + \mathbf{x}(n)\mathbf{x}^H(n) \\
\\
\text{for } i = 1 \text{ to } D \text{ do} \\
\qquad \mathbf{z}_i(n) = \mathbf{R}(n)\ \mathbf{p}_i(n) \\
\qquad \alpha_i(n) = \frac{\rho_i(n)}{\mathbf{p}_i(n)^H \mathbf{z}_i(n)} \\
\qquad \mathbf{w}_i(n) = \mathbf{w}_{i-1}(n) + \alpha_i(n)\mathbf{p}_i(n) \\
\qquad \mathbf{g}_i(n) = \mathbf{g}_{i-1}(n) - \alpha_i(n)\mathbf{z}_i(n) \\
\qquad \rho_i(n) = \mathbf{g}_i(n)^H \mathbf{g}_i(n) \\
\qquad \eta_i(n) = \frac{\rho_i(n)}{\rho_{i-1}(n)} \\
\qquad \mathbf{p}_{i+1}(n) = \mathbf{g}_i(n) + \ \eta_i(n)\mathbf{p}_i(n) \\
\text{End for}
\end{array}
$$

Table 2. Recursive Conjugate Gradient Algorithm

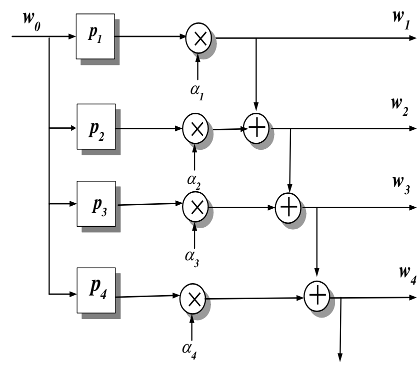

Fig. 2. Conjugate Gradient implementation

A possible approach to derive the CG solution of (11) is the minimization of the following cost function

$$\phi(\mathbf{w}) = \mathbf{w}^{\mathbf{H}}\mathbf{R}_{x_0}\mathbf{w} - \mathbf{2}Re(\mathbf{r}_{x_0 d_0}\mathbf{w}) \tag{12}$$

Table 2 depicts an iterative version of the CG [13], The $\mathbf{p}_i's$ are $\mathbf{R}_{\mathbf{x}_0}$-conjugate search directions i.e.

$$\mathbf{p}_i^H \mathbf{R}_{\mathbf{x}_0} \mathbf{p}_j = 0 \quad \forall i \neq j \tag{13}$$

Theorem 1 *After D iterations of the CG algorithm (table 2) we have the following identities*

$$
\begin{aligned}
span\{\mathbf{w}^{(1)}, \mathbf{w}^{(2)}\ \ldots, \mathbf{w}^{(D)}\} &\equiv span\{\mathbf{p}_1, \mathbf{p}_2\ \ldots, \mathbf{p}_D\} \\
&\equiv span\{\mathbf{g}_1, \mathbf{g}_2\ \ldots, \mathbf{g}_D\} \\
&\equiv \mathcal{K}^D(\mathbf{r}_{\mathbf{x}d}, \mathbf{R}_{\mathbf{x}_0})
\end{aligned} \tag{14}
$$

For the proof see [13]

Theorem 1 means that the solution to the weight vector for the Wiener-Hopf equation in (11) using the CG lies in the same subspace as the weight vector resulting from the MSWF and the resulting reduced rank transformation is given by the search directions vectors:

$$\mathbf{T}_{CG} = [\,\mathbf{p}_1, \mathbf{p}_2\ \ldots, \mathbf{p}_D] \tag{15}$$

Another connection to MSWF is that the CG also tridiagonalize the covariance matrix $\mathbf{R}_{\mathbf{x}_0}$ [14] where the elements of the transformed matrix using the MSWF $\mathbf{R}_d$. Based on the work in [15] and the equivalence of both CG and MSWF to Lanczos algorithm [1] the the relationship between the CG and MSWF coefficients can be easily found and is given by the following equations

$$\sigma^2_{d_{i+1}} = \begin{cases} \frac{1}{\alpha_i} & \text{for } i = 1 \\ \frac{1}{\alpha_i} + \frac{\eta_{i-1}}{\alpha_{i-1}} & \text{for } i > 1 \end{cases} \tag{16}$$

$$\delta_{i+1} = \frac{\sqrt{\eta_{i-1}}}{\alpha_{i-1}} \tag{17}$$

where α_i and η_i are defined in table 2.

One desirable feature of the CG is that it is a forward only recursion and the weight vector is available at each iteration as shown in figure 2, this property simplifies real time adaptation (for example, by evaluating the SINR output at each stage).

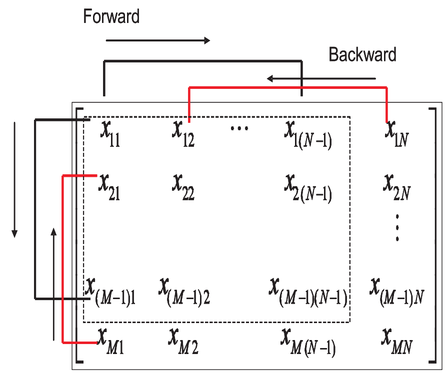

Fig. 3. Schematic of forward/backward subaperture for a single snapshot

5. A LOW SAMPLE SUPPORT CG ALGORITHM

In this section, we introduce the modified version of the CG by exploiting the space-time data structure to generate additional data vectors of reduced size for the covariance matrix estimation using the generalized forward/backward smoothing [2]. As shown in figure 3 the method consists in exploiting the symmetry property of the antenna array and temporal sampling as in the traditional forward/backward spatial smoothing. Suppose that the subspatial vector dimension is L and the subtemporal vector dimension is J, then the total vector size is JL and the effective number of snapshots is increased by a factor of $(M-J-1)(N-L-1)$, whereas the steering vector size is JL. We then propose to replace the estimated covariance matrix in the CG algorithm given in table 2 by the new covariance estimate $\mathbf{R}_{fb} \in \mathcal{C}^{JL\times JL}$ using the forward and backward data $\mathbf{x}_f$, $\mathbf{x}_b$ as described by the following set equations.

$$\mathbf{R}_f(n) = \mathbf{R}_f(n-1) + \mu\mathbf{x}_f(n)\mathbf{x}_f^H(n) \tag{18}$$

$$\mathbf{R}_b(n) = \mathbf{R}_b(n-1) + \mu\mathbf{x}_b(n)\mathbf{x}_b^H(n) \tag{19}$$

$$\mathbf{R}_{fb}(n) = 0.5(\mathbf{R}_b(n) + \mathbf{R}_f(n)) \tag{20}$$

where $\mathbf{R}_f \in \mathcal{C}^{JL \times JL}$ and $\mathbf{R}_b \in \mathcal{C}^{JL \times JL}$ define the forward and backward covariances estimates. The vectors $\mathbf{x}_f(n)$ and $\mathbf{x}_b(n)$ are defined by concatenating the rows of the forward backward submatrices, respectively (see figure 3).

6. SIMULATION RESULTS

The simulation model used is an $N = 14$ elements uniform linear array with M=16 delay taps at each element. The clutter is complex gaussian distributed with clutter to noise ratio (CNR) 40 db at each element. We assume the presence of 4 barrage jammers located at [-60° -30° 45° 60°] with respect to the flight direction and with jammer to noise ratio JNR [40 30 40 30]dB respectively. As a measure of performance we use the SINR loss as defined below

$$SINR_{Loss} = \frac{\sigma^2}{NM} \frac{\left|\mathbf{w}^H \mathbf{s}\right|^2}{\mathbf{w}^H \mathbf{R}_i \mathbf{w}} \quad (21)$$

All the simulations were realized over 100 Monte Carlo runs. Figure 4 shows the performance of the CG, MSWF, PC as a function of the rank[2] using the sample covariance matrix estimate for a sample support $K = 100$. We clearly note that the PC method is not able to achieve optimal output SINR until the rank equals the rank of the interference. Note that underestimating the rank results in degradation in SINR output whereas the MSWF and CG attain the optimal performance at very lower rank. We note that the CG has the same performance as the MSWF for low ranks values whereas for higher rank values the MSWF performance degrades considerably. This can be explained by the loss of orthogonality in the basis vectors $\mathbf{h}_i$ obtained with the MSWF. A correction to this problem has been proposed in [16]by using nested Householder transformations at the output observations.

Figure 5 demonstrates the effectiveness of the proposed method using forward/backward subaperture smoothing for accelerating the convergence in terms of the sample support. The subtemporal and subspatial vectors dimensions are $J = M - 1$ and $L = N - 1$ respectively.

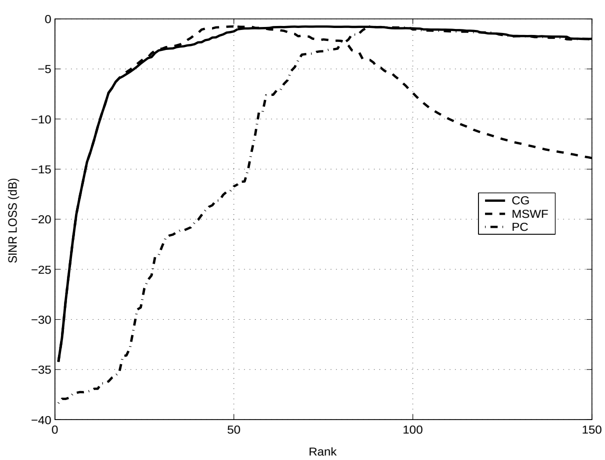

Fig. 4. SINR Loss as function of rank for PC,MSWF and CG

[2]The rank here defines the reduction transformation rank

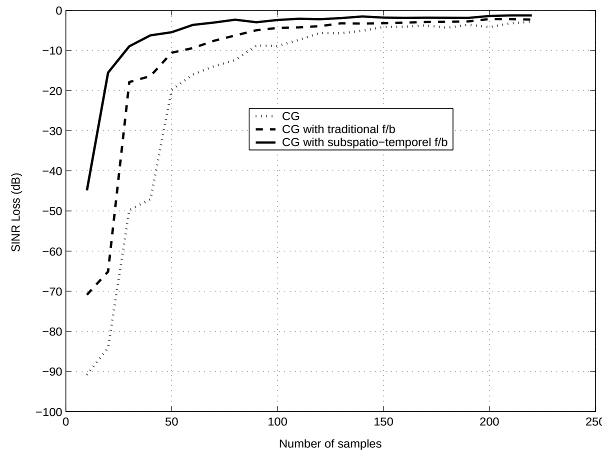

Fig. 5. SINR Loss as function of rank for CG using traditional f/b covariance estimate and CG with f/b subaperture covariance estimate

7. CONCLUSION

In this paper, the application of Reduced rank filters based on Krylov subspace projections in the STAP problem was studied. After presenting the motivation of the need of low rank processors in real world environment, the outperformance of the Krylov based subspace projection filters over eigenbased ones has been presented via simulations. We proposed a low sample support version of the CG by exploiting the space-time data structure to generate additional data vectors using the forward/backward subaperture smoothing. We demonstrated that this version provide excellent performance for short data records.

8. REFERENCES

[1] S. Burykh and K. Abed-meraim, “Reduced-rank adaptive filtering using Krylov subspace,” *EURASIP JASP*, vol. 2002, no. 12, pp. 1387–1400, Dec. 2002.

[2] S. U. Pillai, Y. L. Kim, and J. R. Guerci, “Generalized forward/backward subaperture smoothing techniques for sample starved STAP,” *IEEE Transactions on Signal Processing*, vol. 48, no. 12, pp. 3569–3574, Dec. 2000.

[3] J. Ward, “Space-time adaptive processing for airborne radar,” Lincolin Laboratory,” Technical report 1015, December 1994.

[4] L. Brennan, , and I. Reed, “Theory of adaptive radar,” *IEEE Transaction on Aerospace and Electronic Systems*, vol. AES-9, pp. 237–252, Mar. 1973.

[5] A. Haimovitch, “Eigenanalysis-based space-time adaptive radar: Performance analysis,” *IEEE Transaction on Aerospace and Electronic Systems*, vol. 33, pp. 1170–1179, October 1997.

[6] Y. Kim, S. Pillai, and J. Guerci, “Optimal loading factor for minimal support space-time adaptive processing,” pp. 2505–2508, 1998.

[7] J. S. Goldstein, I. S. Reed, and P. A. Zulch, "Multistage partially adaptive STAP CFAR detection algorithm," *IEEE Transaction on Aerospace and Electronic Systems*, vol. 35, no. 2, pp. 645–662, April 1999.

[8] D. Ricks and J. Goldstein, "Efficient architectures for implementing adaptive algorithms," *Proceedings of the Antenna Applications Symposium*, pp. 29–41, Sept. 2000.

[9] M. Honig and J. Goldstein, "Adaptive reduced-rank interference suppression based on the multistage Wiener filter," *IEEE Transactions on Communications*, vol. 50, no. 6, pp. 986–994, June 2002.

[10] J. D. Hiemstra, "Robust implementation of the Multistage Wiener filter," Ph.D. dissertation, Virginia Polytechnic Institute and State University, April 2003.

[11] M. Joham and M. D. Zoltowski, "Interpretation of the multistage nested wiener filter in the Krylov subspace framework," Munich University of Technology, Tech. Rep. TUM-LNS-TR-00-6, Nov. 2000.

[12] W. Chen, U. Mitra, and P. Schniter, "On the equivalence of three reduced rank linear estimators with applications to DS-CDMA," *IEEE Transactions on Information Theory*, vol. 48, no. 9, pp. 2609–2614, Sept. 2002.

[13] G. H. Golub and C. F. VanLoan, *Matrix computations*, 3rd ed. Johns Hopkins Univ. Press, Baltimore, 1996.

[14] M. E. Weippert, J. Hiemstra, and M. Zoltowski, "Insights from the relationship between the multistage wiener filter and the method of conjugate gradients," *Sensor Array and Multichannel Signal Processing Workshop*, pp. 388 – 392, Aug. 2002.

[15] Y. Saad, *Iterative methods for sparse linear systems*. Available at http://www-users.cs.umn.edu/ saad/.

[16] M. L. de Campos, S. Werner, and J. Apolinàrio, "On an efficient implementation of the multistage Wiener filter through housholder reflections for DS-CDMA interference suppression," *Proceedings of Globecom*, pp. 2350–2354, 2003.