Conjugate Gradient Parametric Adaptive Matched Filter

Chaoshu Jiang and Hongbin Li
Department of Electrical and Computer Engineering
Stevens Institute of Technology, Hoboken, NJ, 07030 USA
Email: cjiang@stevens.edu, hli@stevens.edu

Muralidhar Rangaswamy
Air Force Research Laboratory, Sensors Directorate
Hascom Air Force Base, MA 01731-2909 USA
Email: muralidhar.rangaswamy@hanscom.af.mil

Abstract— The parametric adaptive matched filter (PAMF) detector for space-time adaptive processing (STAP) detection is re-examined in this paper. Originally, the PAMF detector was introduced by using a multichannel autoregressive (AR) parametric model for the disturbance signal in STAP detection. While the parametric approach brings in benefits such as significantly reduced training and computational requirements as compared with fully adaptive STAP detectors, the PAMF detector as a reduced-dimensional solution remains unclear. This paper employs the conjugate-gradient (CG) algorithm to solve the linear prediction problem arising in the PAMF detector. It is shown that CG yields not only a new computationally efficient implementation of the PAMF detector, but it also offers new perspectives of PAMF as a reduced-rank subspace detector. The CG algorithm is first introduced to provide alternative implementations for the matched filter (MF) and parametric matched filter (PMF) when the covariance matrix of the disturbance signal is known. It is then extended to the adaptive case where the covariance matrix is estimated from training data. Important issues such as unknown model order and convergence rate are discussed. Performance of the proposed CG-PAMF detector is examined by using the KASSPER and other computer generated data.

I. Introduction

This paper is concerned with a multichannel signal detection problem frequently encountered in phased-array radars and many other applications. With extra spatial information provided by multiple sensors, higher performance of signal detection can be achieved (than a single-sensor system), especially in detection of signals buried in a background of directional jammers and space-time correlated clutter.

A widely explored technology for multichannel signal detection is space-time adaptive processing (STAP) [1], first proposed by Brennan, Reed and Mallett [2]. Most STAP-based methods, such as the adaptive matched filter (AMF) [3] and Kelly's generalized likelihood ratio test (GLRT) [4], need to invert a large space-time covariance matrix. These methods require not only a large number of independent, identically distributed, signal-free training data to estimate the matrix, but they also incur a high computational cost for matrix estimation and inversion.

A parametric STAP detector based on a multichannel autoregressive (AR) disturbance model has been proposed for airborne radar applications [5], [6] to reduce both the

training data requirement and computation load. This method is called the parametric adaptive matched filter (PAMF) [6]. While the PAMF detector has been found to yield exceptional performance with significantly reduced training and computational requirement when compared with fully adaptive STAP detectors, the connections between the PAMF and other reduced-dimensional or partially adaptive STAP detectors [1], which have similar benefits in training and complexity, remains unclear.

This paper aims to provide some insights into this problem by employing the conjugate-gradient (CG) method to solve the linear prediction problem underlying the PAMF detector. Our choice of the CG method is motivated by several factors. First, the CG method is a computational efficient algorithm to solve the linear prediction problem underlying the PAMF detector. In particular for airborne radar applications, due to an inherent structure of the disturbance covariance matrix, the CG algorithm can usually achieve convergence using only a few iterations, thus providing significant computational saving. Second, as we will shown, the CG iterations naturally lead to a subspace interpretation of the PAMF and offer a connection to other reduced-rank STAP detectors.

II. Data Model

Consider a received J-channel sequence $\{\boldsymbol{x}(n)|n = 1, 2, \cdots, N\}$ corrupted by a space-time correlated disturbance random process $\boldsymbol{c}(n)$. The detection problem involves the following binary hypotheses:

$$
\begin{aligned}
H_0 &: \boldsymbol{x}(n) = \boldsymbol{c}(n) \\
H_1 &: \boldsymbol{x}(n) = a\boldsymbol{s}(n) + \boldsymbol{c}(n)
\end{aligned}
\tag{1}
$$

where $\boldsymbol{s}(n)$ is a known J-channel signal and a its unknown complex amplitude. All vectors in (1) are $J \times 1$ vectors. For convenience of later discussions, define the following vectors in descending order: $\boldsymbol{s} = [\boldsymbol{s}^T(N), \boldsymbol{s}^T(N-1), \cdots, \boldsymbol{s}^T(1)]^T$, $\boldsymbol{c} = [\boldsymbol{c}^T(N), \boldsymbol{c}^T(N-1), \cdots, \boldsymbol{c}^T(1)]^T$, $\boldsymbol{x} = [\boldsymbol{x}^T(N), \boldsymbol{x}^T(N-1), \cdots, \boldsymbol{x}^T(1)]^T$. It is standard to assume that the disturbance $\boldsymbol{c}$ is a Gaussian random vector with zero-mean and space-time covariance matrix $\boldsymbol{R}_c \in \mathbb{C}^{JN \times JN}$, while the signal vector $\boldsymbol{s}(n)$ is deterministic (Swerling 0 target). Based on these assumptions, $\boldsymbol{x} \sim \mathcal{CN}(a\boldsymbol{s}, \boldsymbol{R}_c)$, where $a = 0$ under H_0 and $a \neq 0$ under H_1.

Report Documentation Page

1. REPORT DATE **MAY 2010**	2. REPORT TYPE	3. DATES COVERED **00-00-2010 to 00-00-2010**
4. TITLE AND SUBTITLE **Conjugate Gradient Parametric Adaptive Matched Filter**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Air Force Research Laboratory,Sensors Directorate,Hanscom AFB,MA,01731-2909**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM002322. Presented at the 2010 IEEE International Radar Conference (9th) Held in Arlington, Virginia on 10-14 May 2010. Sponsored in part by the Navy.

14. ABSTRACT

The parametric adaptive matched filter (PAMF) detector for space-time adaptive processing (STAP) detection is re-examined in this paper. Originally, the PAMF detector was introduced by using a multichannel autoregressive (AR) parametric model for the disturbance signal in STAP detection. While the parametric approach brings in benefits such as significantly reduced training and computational requirements as compared with fully adaptive STAP detectors, the PAMF detector as a reduced-dimensional solution remains unclear. This paper employs the conjugate-gradient (CG) algorithm to solve the linear prediction problem arising in the PAMF detector. It is shown that CG yields not only a new computationally efficient implementation of the PAMF detector, but it also offers new perspectives of PAMF as a reduced-rank subspace detector. The CG algorithm is first introduced to provide alternative implementations for the matched filter (MF) and parametric matched filter (PMF) when the covariance matrix of the disturbance signal is known. It is then extended to the adaptive case where the covariance matrix is estimated from training data. Important issues such as unknown model order and convergence rate are discussed. Performance of the proposed CG-PAMF detector is examined by using the KASSPER and other computer generated data.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**6**	

In STAP, the signal s is known as the space-time steering vector. For a side-looking uniform linear array (ULA), s is given by

$$s = s_t \otimes s_s \qquad (2)$$

where $s_t = (1/\sqrt{N})[e^{i2\pi(N-1)f_d}, \cdots, e^{i2\pi f_d}, 1]^T$ is the temporal steering vector with a normalized Doppler frequency f_d, $s_s = (1/\sqrt{J})[e^{i2\pi(N-1)f_s}, \cdots, e^{i2\pi f_s}, 1]^T$ is the spatial steering vector with a normalized spatial frequency f_s, and $\otimes$ denotes the Kronecker product. Practically, the true disturbance covariance matrix R_c is unknown, and often an estimate can be obtained from the secondary data:

$$\hat{R}_c = \frac{1}{K}\sum_{k=1}^{K} c_k c_k^H \qquad (3)$$

where $c_k, k = 1, 2\cdots, K$, denote the secondary data vectors. According to the well-known "RMB" rule [7], we need $K \geq 2JN - 3$ so that the average output signal-to-interference-plus-noise ratio (SINR) loss caused by covariance estimation error is less than 3dB. Detectors with an estimated covariance matrix are often called adaptive methods.

III. MF AND PMF

Assuming a known R_c, the matched filter (MF) is obtained by maximizing the output SINR of a linear receiver or the generalized likelihood ratio (GLR). The test is given by (e.g., [3]):

$$\frac{|s^H R_c^{-1} x|^2}{s^H R_c^{-1} s} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{\text{MF}} \qquad (4)$$

where η_{MF} is the threshold of MF. The performance of this test is regarded as a benchmark for all linear tests.

MF can also be represented by using a structure of temporal whitening cascaded with spatial whitening which comes from a block LDU decomposition of the disturbance covariance matrix [6]. This form of MF is given by

$$\frac{|(Q^{-1/2}L^{-1}s)^H(Q^{-1/2}\varepsilon)|^2}{(Q^{-1/2}L^{-1}s)^H(Q^{-1/2}L^{-1}s)} = \frac{|\tilde{s}^H \nu|^2}{\tilde{s}^H \tilde{s}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{\text{MF}} \qquad (5)$$

where $Q \in \mathbb{C}^{JN \times JN}$ is a block-diagonal matrix with Hermitian matrices $Q(n), n = 1, 2, \cdots, N$, along the main block diagonal, and $L \in \mathbb{C}^{JN \times JN}$ is a lower block-triangular matrix with $J \times J$ identity matrices along the main block diagonal. Both L and Q come from a block LDU decomposition of the disturbance covariance matrix $R_c = LQL^H$. Finally,

$$\varepsilon(n) = x(n) - \sum_{p=1}^{(n-1)} A_p^H(n)x(n-p) \qquad (6)$$

$$\nu(n) = Q^{-1/2}(n)\varepsilon(n) \qquad (7)$$

$$\tilde{s}(n) = Q^{-1/2}(n)\left[s(n) - \sum_{p=1}^{(n-1)} A_p^H(n)s(n-p)\right] \qquad (8)$$

where $A_p^H(n) \in \mathbb{C}^{J \times J}$ is a block element of L^{-1} located at the $(n-p)$th block column and the nth block row.

If the disturbance $c(n)$ is stationary in time, the MF can be simplified. A parametric matched filter (PMF) was introduced in [6] by modeling the disturbance as a stationary Pth-order multichannel autoregressive (AR) process. Specifically,

$$c(n) = \sum_{p=1}^{P} A_p^H c(n-p) + \varepsilon_P(n) \qquad (9)$$

where $A_p^H, p = 1, 2, \cdots, P$, are the matrices of the AR coefficients, and $\varepsilon_P(n)$ is the temporally white noise with a spatial covariance matrix Q_P. The PMF test is given by [6]

$$\frac{|\sum_{n=P+1}^{N} \tilde{s}_P^H(n)\nu_P(n)|^2}{\sum_{n=P+1}^{N} \tilde{s}_P^H(n)\tilde{s}_P(n)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{\text{PMF}} \qquad (10)$$

where $\nu_P(n) = Q_P^{-1/2}\varepsilon_P(n)$ and

$$\tilde{s}_P(n) = Q_P^{-1/2}\left[s(n) - \sum_{p=1}^{P} A_p^H s(n-p)\right] \qquad (11)$$

for $n = P + 1, \cdots, N$. In practice, the model order P and the AR coefficients $\{A_p\}$ are estimated from the secondary data and/or primary data. Different estimators lead to different versions of the PAMF detector [6], [8].

IV. CG-MF AND CG-PMF

In this section, we discuss alternative implementations of the MF and PMF via the CG algorithm. The resulting detectors are referred to as the CG-MF and CG-PMF detectors, respectively, for brevity. We start from the CG-MF, which also sets the basis for the CG-PMF. The latter, assuming temporally stationary disturbance, is a computationally simplified version of the CG-MF. The link between the PMF and CG as developed in the sequel reveals the PMF as a reduced-dimensional subspace detector. In this section, we assume knowledge of the covariance matrix of the disturbance signal. An adaptive version of the CG-PMF (i.e., CG-PAMF) is discussed in Section V.

A. Conjugate-Gradient MF

The MF detector, as shown in Section III, can be derived from a time-varying linear prediction process. Specifically, consider the problem of linearly predicting the nth sample $x(n)$ under H_0 from all prior received samples $x(n-1), x(n-2), \ldots, x(1)$ (cf. (9))

$$x(n) = B^H(n)y(n) + \varepsilon(n) \qquad (12)$$

where $B(n) = [A_1^H(n), A_2^H(n), \cdots, A_{n-1}^H(n)]^H = [b_1(n), b_2(n), \cdots, b_J(n)] \in \mathbb{C}^{J(n-1) \times J}$ denotes the $(n-1)$-st order time-varying multichannel linear prediction filter, and $y(n) = [y_n(1), y_n(2), \cdots, y_n(J(n-1))]^T = [x^T(n-1), x^T(n-2), \cdots, x^T(1)]^T$ contains all $n - 1$ previously received data vectors. It is noted that the above time-varying linear predictor grows in its filter order or size with n. The multichannel linear predictor is equivalent to J scalar linear predictors:

$$x_j(n) = b_j^H(n)y(n) + \varepsilon_j(n), \quad j = 1, 2, \cdots, J \qquad (13)$$

The optimum linear predictor can be obtained by solving the Wiener-Hopf equations:

$$\boldsymbol{R_{yy}}(n)\boldsymbol{b}_j(n) = \boldsymbol{r}_j(n), \quad j = 1, 2, \cdots, J \quad (14)$$

where $\boldsymbol{R_{yy}}(n) = E[\boldsymbol{y}(n)\boldsymbol{y}^H(n)] \in \mathbb{C}^{J(n-1) \times J(n-1)}$ and $\boldsymbol{r}_j(n) = E[\boldsymbol{y}(n)x_j^*(n)] \in \mathbb{C}^{J(n-1) \times 1}$. Again, note that the size of the Wiener-Hopf equation grows with n.

To obtain a temporally whitened sequence $\boldsymbol{\varepsilon}(n)$ for MF detection (cf. (6)), the above linear prediction process has to be performed multiple times, starting from $n = 2$ to $n = N$. For each n, we need solve a Wiener-Hopf equation of the form (14). While there are various solvers to the linear Wiener-Hopf equation, we consider using the conjugate gradient (CG) method, which has several properties such as fast convergence and direct link to the Krylov subspace [9]. Additional remarks on such aspects are provided shortly.

The recursive procedure involved for the determination of the linear predictors is described as follows.

for $n = 2$ to N **do**
　for $j = 1$ to J **do**
　　Initialization. Initialize the conjugate-direction vector $\boldsymbol{d}_{1,j}(n)$, gradient vector $\boldsymbol{\gamma}_{1,j}(n)$, step size $\alpha_{1,j}(n)$ and initial solution $\boldsymbol{b}_{1,j}(n)$:

$$\boldsymbol{d}_{1,j}(n) = -\boldsymbol{\gamma}_{1,j}(n) = \boldsymbol{r}_j(n) \quad (15)$$

$$\alpha_{1,j}(n) = \frac{\|\boldsymbol{\gamma}_{1,j}(n)\|^2}{\boldsymbol{d}_{1,j}^H(n)\boldsymbol{R_{yy}}(n)\boldsymbol{d}_{1,j}(n)} \quad (16)$$

$$\boldsymbol{b}_{1,j}(n) = \boldsymbol{0}. \quad (17)$$

　　for $k = 2, 3, \cdots$, till convergence ($k \leq J(n-1)$) **do**
　　Update the gradient vector $\boldsymbol{\gamma}_{k,j}$:

$$\boldsymbol{\gamma}_{k,j}(n) = \boldsymbol{\gamma}_{k-1,j}(n) + \alpha_{k-1,j}(n)\boldsymbol{R_{yy}}(n)\boldsymbol{d}_{k-1,j}(n) \quad (18)$$

　　Update the conjugate-direction vector $\boldsymbol{d}_{k,j}$:

$$\boldsymbol{d}_{k,j}(n) = \boldsymbol{d}_{k-1,j}(n)\frac{\|\boldsymbol{\gamma}_{k,j}(n)\|^2}{\|\boldsymbol{\gamma}_{k-1,j}(n)\|^2} - \boldsymbol{\gamma}_{k,j}(n) \quad (19)$$

　　Update the step size $\alpha_{k,j}$:

$$\alpha_{k,j}(n) = \alpha_{k,j}^*(n) = \frac{\|\boldsymbol{\gamma}_{k,j}(n)\|^2}{\boldsymbol{d}_{k,j}^H(n)\boldsymbol{R_{yy}}(n)\boldsymbol{d}_{k,j}(n)} \quad (20)$$

　　Update the solution $\boldsymbol{b}_{k,j}$:

$$\boldsymbol{b}_{k,j}(n) = \boldsymbol{b}_{k-1,j}(n) + \alpha_{k,j}(n)\boldsymbol{d}_{k,j}(n) \quad (21)$$

　　end for
　end for
end for

Let $\boldsymbol{B}(n)$ be the multichannel linear predictor formed from $\boldsymbol{b}_{k,j}$ after convergence. Then, $\boldsymbol{B}(n)$ can be used to whiten $\boldsymbol{x}(n)$ to produce a temporally whitened sequence $\boldsymbol{\varepsilon}(n)$. The spatial covariance matrix $\boldsymbol{Q}(n)$ of $\boldsymbol{\varepsilon}(n)$ is given by

$$\boldsymbol{Q}(n) = E[\boldsymbol{\varepsilon}(n)\boldsymbol{\varepsilon}^H(n)] = \boldsymbol{R_{xx}}(n) - \boldsymbol{B}^H(n)\boldsymbol{R_{yx}}(n)$$
$$- \boldsymbol{R_{yx}}^H(n)\boldsymbol{B}(n) + \boldsymbol{B}^H(n)\boldsymbol{R_{yy}}(n)\boldsymbol{B}(n) \quad (22)$$
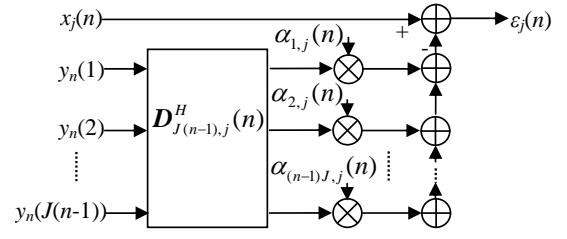
Fig. 1. Time-varying linear prediction in the conjugate-gradient MF detector.

where $\boldsymbol{R_{xx}}(n) = E[\boldsymbol{x}(n)\boldsymbol{x}^H(n)] \in \mathbb{C}^{J \times J}$, and $\boldsymbol{R_{yx}}(n) = E[\boldsymbol{y}(n)\boldsymbol{x}^H(n)] \in \mathbb{C}^{J(n-1) \times J}$, which can be used for further spatial whitening.

Fig. 1 depicts the CG-MF detector that produces the n-th sample of the temporally whitened sequence $\varepsilon_j(n)$ for the j-th channel, where $\boldsymbol{D}_{k,j}(n) = [\boldsymbol{d}_{1,j}(n), \boldsymbol{d}_{2,j}(n), \cdots, \boldsymbol{d}_{k,j}(n)]$ is the conjugate-direction matrix. CG iterations lead to a set of linearly independent vectors $\boldsymbol{d}_{1,j}(n), \ldots, \boldsymbol{d}_{k,j}(n)$ that are conjugate orthogonal, i.e., $\boldsymbol{d}_{k,j}^H(n)\boldsymbol{R_{yy}}(n)\boldsymbol{d}_{l,j}(n) = 0$ for $k \neq l$. The output of the kth iteration is given by

$$\boldsymbol{b}_{k,j}(n) = \sum_{m=1}^{k} \alpha_{m,j}(n)\boldsymbol{d}_{m,j}(n) \quad (23)$$

which is a vector in the k-dimensional vector space spanned by the conjugate-direction vectors $\{\boldsymbol{d}_{m,j}(n), m = 1, 2, \cdots, k\}$. The iterative procedure for the prediction of the n-th sample $x_j(n)$, which involves a $J(n-1)$-st order linear predictor, converges after at most $J(n-1)$ iterations. The final solution $\boldsymbol{b}_j(n)$ lies in a $J(n-1)$-dimensional vector space.

B. Conjugate-Gradient PMF

If the disturbance signal can be approximated as a temporally wide-sense stationary (WSS) multichannel AR process, the linear prediction problem of the previous subsection can be significantly simplified. Specifically, suppose the disturbance is an AR(P) process with model order P. In this case, the optimum linear predictor for the n-th sample $\boldsymbol{x}(n)$ requires only P most recently received samples (as opposed to all past samples) and the prediction filter is time-invariant with a fixed size (as opposed to time-varying with a growing size):

$$\boldsymbol{x}(n) = \boldsymbol{B}^H\boldsymbol{y}_P(n) + \boldsymbol{\varepsilon}_P(n) \quad (24)$$

where the fixed P-th order linear predictor $\boldsymbol{B} = [\boldsymbol{A}_1^H, \boldsymbol{A}_2^H, \cdots, \boldsymbol{A}_P^H]^H = [\boldsymbol{b}_1, \boldsymbol{b}_2, \cdots, \boldsymbol{b}_J] \in \mathbb{C}^{JP \times J}$ is composed of the AR coefficient matrices $\{\boldsymbol{A}_p^H\}$ (cf. (9)), $\boldsymbol{y}_P(n) = [y_n(1), y_n(2), \cdots, y_n(JP)]^T = [\boldsymbol{x}^T(n-1), \boldsymbol{x}^T(n-2), \cdots, \boldsymbol{x}^T(n-P)]^T$ denotes the regression data vector, and $n > P$. Again, it is convenient to express the above multichannel linear predictor as J scalar linear predictors:

$$x_j(n) = \boldsymbol{b}_j^H\boldsymbol{y}_P(n) + \varepsilon_{P,j}(n), \quad j = 1, 2, \cdots, J \quad (25)$$

The structure of temporally whitening via linear prediction for the PMF detector is shown in Fig. 2.
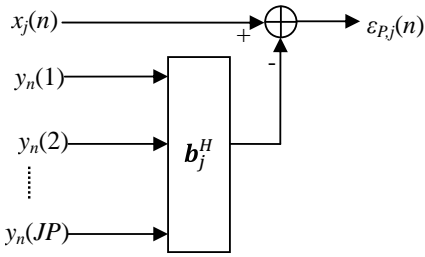
Fig. 2. Time-invariant linear prediction in the conjugate-gradient PMF detector.

The solution to the scalar linear prediction problem can be obtained by solving the following Wiener-Hopf equation

$$\boldsymbol{R_{yy}b}_j = \boldsymbol{r}_j, \quad j = 1, 2, \cdots, J \qquad (26)$$

where $\boldsymbol{R_{yy}} = E[\boldsymbol{y}_P(n)\boldsymbol{y}_P^H(n)] \in \mathbb{C}^{JP \times JP}$ and $\boldsymbol{r}_j = E[\boldsymbol{y}_P(n)x_j^*(n)] \in \mathbb{C}^{JP \times 1}$. It should be noted that unlike the MF detector, the above Wiener-Hopf is time-invariant, has a fixed size, and needs to be solved only once. The resulting solution $\boldsymbol{b}_j$ can be used to whiten the entire received signal $\boldsymbol{x}(n)$ for $n > P$. The CG algorithm can also be applied to solve (26), and the resulting detector is referred to as the CG-PMF detector. Since only one fixed-sized Wiener-Hopf equation needs to be solved, the CG-PMF detector is computationally much simpler. Specifically, the outer loop for varying n as discussed in Section IV-A varnishes, and only the conjugate-gradient processing with $n = P + 1$ is needed.

Remark: The iterative procedure of CG converges after at most JP iterations for the CG-PMF. As a result, the final solution $\boldsymbol{b}_j$ lies in a JP-dimensional vector space spanned by the conjugate direction vectors $\boldsymbol{d}_{k,j}, k = 1, 2, \ldots, JP$, or equivalently, the JP-dimensional Krylov subspace [9]: $\mathcal{K}(\boldsymbol{r}_j, \boldsymbol{R_{yy}}, JP) = \text{span}\left\{\boldsymbol{r}_j, \boldsymbol{R_{yy}}\boldsymbol{r}_j, \cdots, \boldsymbol{R_{yy}}^{JP-1}\boldsymbol{r}_j\right\}$. This shows that the PMF is a reduced JP-dimensional solution, as opposed to the full JN-dimensional MF detector. Finally, it is noted that if the disturbance signal is an AR(P) process, both the CG procedures for the MF (14) and for the PMF (26) converge to the same Wiener solution with $\boldsymbol{b}_j(n) = [\boldsymbol{b}_j^T, \boldsymbol{0}_{J(n-1-P)\times 1}^T]^T$, for $n > P$.

C. Convergence in Airborne Radar Applications

One important property of the CG algorithm is its fast convergence. In general, it takes no more than JP iterations to solve the linear equation (26) [9]. Even faster convergence is possible if the covariance matrix of the disturbance has some specific structure. In particular, if the covariance matrix is a rank-r_c correction of an identity matrix: $\boldsymbol{R_{yy}} = \boldsymbol{R}_i + \sigma_n^2 \boldsymbol{I}$, where $\boldsymbol{R}_i$ is a rank-r_c positive semi-definite matrix, then the CG algorithms converges in at most $r_c + 1$ iterations [9].

In airborne radar applications, the disturbance covariance matrix may consist of two components, namely a low-rank $\boldsymbol{R}_i$ due to the clutter and jamming and a scaled identity $\sigma_n^2 \boldsymbol{I}$ due to the white noise, where σ_n^2 denotes the noise variance. The rank r_c is typically much smaller than the joint spatio-temporal dimension JN. Specifically, if the disturbance is primarily

due to ground clutter and thermal noise, then according to Brennan's rule [2], the rank of the clutter covariance matrix for the full-dimensional MF is approximately

$$r_{c,full} \approx \lceil J + (N-1)\beta \rceil \qquad (27)$$

where $\beta = 2v_g T_r/d$, v_g is the platform velocity, T_r is the pulse repetition period, d is the antenna element spacing, and $\lceil \cdot \rceil$ rounds a real-valued number towards infinity. Likewise, the rank of the disturbance covariance matrix for the PMF detector can be approximated as

$$r_c \approx \lceil J + (P-1)\beta \rceil \qquad (28)$$

The smaller rank r_c is due to the fact that the disturbance covariance matrix is formed over P pulses, which is sufficient for the reduced-dimensional PMF detector due to the underlying AR(P) model. Meanwhile, the space-time disturbance covariance matrix for the full-dimensional MF detector is formed over N (the entire number of) pulses. As such, the PMF can benefit more from the fast convergence property of the CG algorithm.

V. RECURSIVE CONJUGATIVE-GRADIENT PAMF

The CG-PMF algorithm is now extended to the adaptive case when both the covariance matrix and the AR model order P are unknown. The resulting detector is referred to as the CG-PAMF detector. Since the true model order P is unknown, we need an upper bound $\bar{P}$. One such bound for STAP detection is considered in [6]:

$$\bar{P} = \max\left\{\left\lfloor 3\sqrt{N}/J \right\rfloor\right\} \qquad (29)$$

where $\lfloor \cdot \rfloor$ rounds a positive number towards zero. The extension of CG-PMF involves i) replacing the true covariance matrices with estimates obtained from the secondary data; and ii) integrating some model order selection criterion with conjugate-gradient iterations. In the sequel, we use the generalized Akaike information criterion (GAIC) [10], [11] for model order detection. The CG-PAMF detector is summarized next.

- **Step 1:** Estimate the disturbance covariance matrices from the training data via temporal and range averaging:

$$\hat{\boldsymbol{R}}_{yy}^{(\bar{P})} = \begin{bmatrix} \hat{\boldsymbol{R}}_{xx}(0) & \cdots & \hat{\boldsymbol{R}}_{xx}(\bar{P}-1) \\ \vdots & \ddots & \vdots \\ \hat{\boldsymbol{R}}_{xx}(1-\bar{P}) & \cdots & \hat{\boldsymbol{R}}_{xx}(0) \end{bmatrix} \qquad (30)$$

$$\hat{\boldsymbol{R}}_{yx}^{(\bar{P})} = \begin{bmatrix} \hat{\boldsymbol{R}}_{xx}(-1) \\ \vdots \\ \hat{\boldsymbol{R}}_{xx}(-\bar{P}) \end{bmatrix} \qquad (31)$$

where the sub-matrices are given by

$$\hat{\boldsymbol{R}}_{xx}^H(m) = \frac{1}{NK}\sum_{k=1}^{K}\sum_{l=1}^{N-m}\boldsymbol{x}_k(l+m)\boldsymbol{x}_k^H(l) \qquad (32)$$

K is the number of training data vectors and $\bar{P}$ is determined by (29).

- **Step 2:** Set the order of the AR process to $p = 1$ and initialize the AR coefficients $\hat{\boldsymbol{B}}_0^{(p)} = \boldsymbol{0}_J$.
- **Step 3:** Use the CG algorithm to solve $\hat{\boldsymbol{R}}_{yy}^{(p)} \hat{\boldsymbol{B}}^{(p)} = \hat{\boldsymbol{R}}_{yx}^{(p)}$ until convergence.
- **Step 4:** Compute $\hat{\boldsymbol{Q}}^{(p)}$ as the maximum-likelihood estimate of $\boldsymbol{Q}$ [8].

$$\hat{\boldsymbol{Q}}^{(p)} = \hat{\boldsymbol{R}}_{xx}(0) - \hat{\boldsymbol{B}}^{(p)\,H} \hat{\boldsymbol{R}}_{yx}^{(p)} \tag{33}$$

- **Step 5:** Compute the GAIC criterion which is given by

$$G(p) = V(p) + W(p) \tag{34}$$

where

$$V(p) = JK(N-p)\ln(e\pi) + K(N-p)\ln|\hat{\boldsymbol{Q}}^{(p)}| \tag{35}$$
$$W(p) = 2cJ^2 p\ln(\ln(K(N-p))) \tag{36}$$

where $c \geq 1$ is a parameter of user choice.

- **Step 6:** If $1 < p \leq P_{\max}$, compare $G(p)$ with $G(p-1)$. If $G(p) \geq G(p-1)$, set $\hat{P} = p - 1$ and exit; otherwise increase $p = p + 1$, set $\hat{\boldsymbol{B}}_0^{(p+1)} = [\hat{\boldsymbol{B}}^{(p)\,H}, \boldsymbol{0}_{J\times J}]^H$, and return to Step 3.

To speed up the convergence rate, the final solution from the previous iteration (in p) is used as an initial value of the current iteration. This is because the previous solution $\hat{\boldsymbol{B}}^{(p)}$ lies in a subspace of the current solution $\hat{\boldsymbol{B}}^{(p+1)}$, and therefore fewer iterations are needed for the convergence of current CG processing. Finally if the disturbance is an AR(P) process, then going beyond P iterations does not significantly change the result. For example, the solution $\hat{\boldsymbol{B}}_0^{(P+1)}$ is close to $[\hat{\boldsymbol{B}}^{(P)\,H}, \boldsymbol{0}_{J\times J}]^H$, a zero-padded version of the solution obtained in the previous iteration.

VI. NUMERICAL RESULTS

First, we examine the performance of several implementations of the PMF detector by using simulated data with AR disturbances. Specifically, the disturbance is an AR(2) process with $J = 4$ elements and $N = 64$ pulses. All PMF detectors have knowledge of the exact disturbance covariance matrix; however, they use different approaches to compute the linear predictor. Specifically, we consider the DMI-PMF, which uses direct matrix inverse to solve the Wiener-Hopf equation, and the CG-PMF as discussed in Section IV-B. In addition, two cases are considered for CG-PMF, one with knowledge of the AR model order P and the other has only an upperbound $\bar{P} = 3$. For the latter, we examine the performance of a truncated solution, and the detector is denoted as "CG-PMF (Truncated B)". Specifically, the truncated solution is obtained by running the conjugate-gradient method with 8 iterations, which yields an estimate $\boldsymbol{B}^{(2)}$ with $J\bar{P} = 12$ rows, and discarding the last 4 rows. The simulation used 20000 independent Monte Carlo data realizations, and a threshold is determined to provide a false alarm rate $P_{fa} = 0.01$. The numerical results are shown in Fig. 3. Interestingly, all 3 implementations yield nearly an identical detection performance.
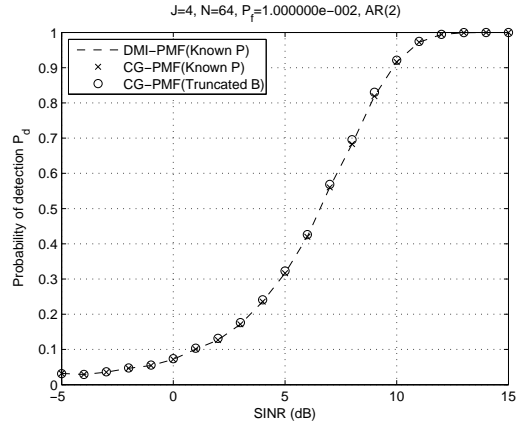

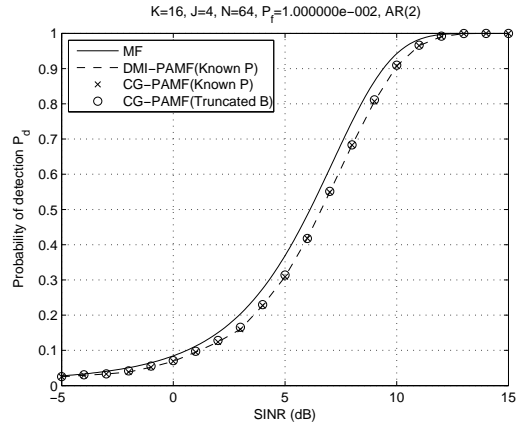Fig. 3. Probability of detection versus SINR of PMF for simulated data($J = 4$; $N = 64$; $P = 2$)

Fig. 4. Probability of detection versus SINR of PAMF for simulated data($K = 16$; $J = 4$; $N = 64$; $P = 2$; $\bar{P} = 3$)

Second, we consider the adaptive PAMF detector, for which the disturbance covariance matrix is unknown. Similar to the PMF detector, we compare three implementations of the PAMF detector, which are referred to as the DMI-PAMF, CG-PAMF, and CG-PAMF (Truncated B). For the truncated implementation, we again assume only an upperbound $\bar{P}$ is known and the truncated solution is obtained similarly as for PMF. The disturbance is also an AR(2) signal, whose disturbance covariance matrix is estimated from $K = 16$ target-free training data vectors, and the AR coefficients are estimated based on the estimated disturbance covariance matrix. The numerical results are shown in Fig. 4 for $\bar{P} = 3$. It is again observed that all 3 implementations yield nearly an identical detection performance.

Finally the performance of the recursive CG-PAMF with an unknown disturbance AR order is considered. Both the simulated data and KASSPER 2002 data set are adopted. The simulated disturbance is the same as in the previous examples. Meanwhile, the KASSPER data set is generated by considering practical airborne radar parameters and issues found in a real-world clutter environment [12]. Specifically,
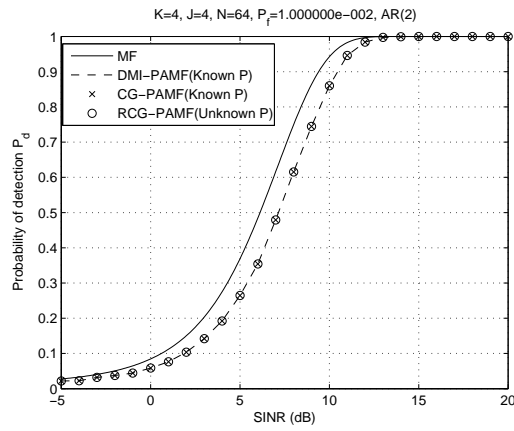
Fig. 5. Probability of detection versus SINR for Simulated data($K = 4$; $J = 4$; $N = 64$; $P = 2$; $\bar{P} = 6$)
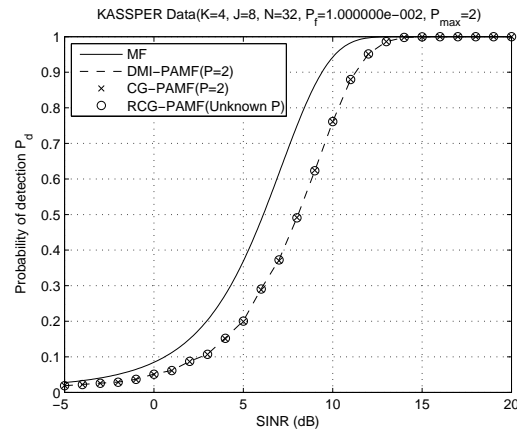

Fig. 6. Probability of detection versus SINR for KASSPER 2002 data($K = 4$; $J = 8$; $N = 32$; $P = 2$; $\bar{P} = 2$)

the simulated airborne radar platform travels at a speed of 100 m/s with a $3°$ crab angle. The radar carrier frequency is 1240 MHz. The horizontal 11 antenna elements form a ULA with a spacing of 0.1092m between adjacent elements, and the transmit array is uniformly weighted and phased to steer the mainbeam to $195°$. The pulse repetition frequency is 1984 Hz and a coherent processing interval contains 32 pulses. Here we use only the outputs of the first 8 elements for processing. The numerical results are shown as Fig. 5 and Fig. 6, in which RCG-PAMF (unknown P) represents the result of the recursive CG-PAMF under the condition of unknown P and $\boldsymbol{R}$. An upper bound $\bar{P}$ for the model order is calculated by (29). It is seen that the performance of the recursive CG-PAMF is nearly identical to that of CG-PAMF with known P (for simulated data) or a pre-selected $P = 2$ (for KASSPER data). Using the relevant parameters of the KASSPER data, we have $\beta = 2v_gT_r/d = 0.923$. It follows that for $J = 8$ elements, the maximum number of conjugate-gradient iterations of the CG-PAMF or RCG-PAMF for a given model order p is $r_{cp} + 1 = \lceil 8.077 + 0.923p \rceil$. For example, the maximum numbers of CG iterations for $p = 2$ is 10 due to the low-rank structure of the clutter, whereas without such a structure, it would require $pJ = 16$ iterations for the CG to converge.

VII. CONCLUSION

The CG algorithm was employed to solve the linear prediction problem underlying the PAMF detector. It is shown that this leads to not only new efficient implementations, but also new insights of the PAMF as a reduced-dimensional subspace detector.

REFERENCES

[1] J. Ward, "Space-time adaptive processing for airborne radar," Lincoln Laboratory, MIT, Technical Report 1015, December 1994.
[2] L. E. Brennan and I. S. Reed, "Theory of adaptive radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 9, no. 2, pp. 237–252, 1973.
[3] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 1, pp. 208–216, January 1992.
[4] E. J. Kelly, "An adaptive detection algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 22, pp. 115–127, March 1986.
[5] M. Rangaswamy and J. H. Michels, "A parametric multichannel detection algorithm for correlated non-Gaussian random processes," in *Proceedings of the 1997 IEEE National Radar Conference*, Syracuse, NY, May 1997, pp. 349–354.
[6] J. R. Román, M. Rangaswamy, D. W. Davis, Q. Zhang, B. Himed, and J. H. Michels, "Parametric adaptive matched filter for airborne radar applications," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, no. 2, pp. 677–692, April 2000.
[7] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 10, no. 6, pp. 853–863, 1974.
[8] K. J. Sohn, H. Li, and B. Himed, "Parametric Rao test for multichannel adaptive signal detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 3, pp. 920–933, July 2007.
[9] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: Johns Hopkins University Press, 1996.
[10] T. Söderström and P. Stoica, *System Identification*. London, UK: Prentice Hall International, 1989.
[11] J. Li and P. Stoica, "Efficient mixed-spectrum estimation with applications to target feature extraction," *IEEE Transactions on Signal Processing*, vol. 44, no. 2, pp. 281–295, February 1996.
[12] J. S. Bergin and P. M. Techau, "High-fidelity site-specific radar simulation: KASSPER'02 workshop datacube," Information Systems Laboratories, Inc., Vienna, VA, Technical Report ISL-SCRD-TR-02-105, May 2002.