

On the Conjugate Gradient Matched Filter

Chaoshu Jiang, Hongbin Li, and Muralidhar Rangaswamy

Abstract—The conjugate gradient (CG) algorithm is an efficient method for the calculation of the weight vector of the matched filter (MF). As an iterative algorithm, it produces a series of approximations to the MF weight vector, each of which can be used to filter the test signal and form a test statistic. This effectively leads to a family of detectors, referred to as the CG-MF detectors, which are indexed by k the number of iterations incurred. We first consider a general case involving an arbitrary covariance matrix of the disturbance (including interference, noise, etc.) and show that all CG-MF detectors attain constant false alarm rate (CFAR) and, furthermore, are optimum in the sense that the k th CG-MF detector yields the highest output signal-to-interference-and-noise ratio (SINR) among all linear detectors within the k th Krylov subspace. We then consider a structured case frequently encountered in practice, where the covariance matrix of the disturbance contains a low-rank component (rank- r) due to dominant interference sources, a scaled identity due to the presence of a white noise, and a perturbation component containing the residual interference. We show that the $(r + 1)$ st CG-MF detector achieves CFAR and an output SINR nearly identical to that of the MF detector which requires complete iterations of the CG algorithm till reaching convergence. Hence, the $(r + 1)$ st CG-MF detector can be used in place of the MF detector for significant computational saving when r is small. Numerical results are presented to verify the accuracy of our analysis for the CG-MF detectors.

Index Terms—Conjugate gradient method, Krylov subspace, matched filter, space-time adaptive processing (STAP).

I. INTRODUCTION

Detection of a multichannel signal in temporally and spatially correlated clutter and/or jamming is found in phased-array radar, sonar, and many other applications. A widely explored technique is space-time adaptive processing (STAP) [1]. Most classical STAP-based methods require to invert a large space-time covariance matrix, thus incurring a substantial amount of training signals as well as a high computational cost (e.g., [2]). Aimed at mitigating the training and computational requirements in STAP detection, reduced-rank techniques, such as eigen-canceller [3], principal-component method [4], cross-spectral metric [5], multistage Wiener filter (MWF) [6], etc., have been proposed to reduce the dimension of the data in advance of detection. Meanwhile, the conjugate gradient (CG) algorithm (e.g., [7]) is an efficient method for solving a system of linear equations and has been explored for adaptive

Manuscript received September 09, 2011; revised January 16, 2012; accepted January 23, 2012. Date of publication February 06, 2012; date of current version April 13, 2012. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Kainam Thomas Wong. The work of C. Jiang and H. Li was supported by the Air Force Office of Scientific Research (AFOSR) under Grant FA9550-09-1-0310. The work of M. Rangaswamy was supported by AFOSR under project 2311IN. This work was presented in part at the IEEE Radar Conference 2011.

C. Jiang was with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030 USA. He is now with the School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu, Sichuan, China, 610054 (e-mail: chshjiang@126.com).

H. Li is with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030 USA (e-mail: Hongbin.Li@stevens.edu).

M. Rangaswamy is with AFRL/RYP, WPAFB, OH, 45433-7132 USA (e-mail: Muralidhar.Rangaswamy@wpafb.af.mil).

Color versions of one or more of the figures in this correspondence are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSP.2012.2187200

filtering, STAP detection, and beamforming in recent studies [8]–[12]. Interestingly, there is a close connection between the CG and the MWF [9], [10].

In this correspondence, we explore the CG algorithm in the matched filter (MF) for STAP detection. The CG algorithm is employed to iteratively calculate the weight vector of the MF detector, which produces a series of progressively improved approximations to the MF weight vector. Each of the intermediate weight vectors generated by the CG can be used to form a decision variable, which, collectively, forms a family of STAP detectors referred to herein as the *CG-MF detectors*. Our goal is to examine the performance of these CG-MF detectors relative to the benchmark MF detector. We consider two cases, one involving a general covariance for the disturbance (i.e., clutter, jamming, and noise) and the other involving a structured disturbance covariance.

Specifically, for the first case, the space-time covariance matrix of the disturbance, denoted by \mathbf{R} , is arbitrary. Our analysis shows that the CG-MF detectors achieve constant false alarm rate (CFAR), irrespective of the number of iterations. The probability of detection, for a given false alarm probability, is dictated by the output signal-to-interference-and-noise ratio (SINR) of the linear filter employed by the CG-MF detector, which is non-decreasing over the CG iterations. It is found that the CG-MF detector obtained at the k th CG iteration is optimum in the sense that it yields the largest output SINR over all linear detectors within the k -dimensional Krylov subspace.

For the second structured case, we assume $\mathbf{R} = \mathbf{R}_i + \sigma_n^2 \mathbf{I} + \mathbf{\Delta}$, where \mathbf{R}_i is a rank- r matrix which contains dominant clutter scatterers or interference sources, the scaled identity $\sigma_n^2 \mathbf{I}$ is due to the presence of a white noise with variance σ_n^2 , and $\mathbf{\Delta}$ is a perturbation term with small entries compared with those in \mathbf{R}_i . The perturbation $\mathbf{\Delta}$ may be caused by insignificant, residual clutter/interference sources. In this correspondence, $\mathbf{\Delta}$ is treated as a deterministic perturbation. In the absence of the perturbation, it is well-known that the CG algorithm converges in $r + 1$ iterations [7], and if r is small, requires significant less computation than directly computing the matrix inverse needed by the MF detector. However, with $\mathbf{\Delta}$, the CG in general requires full iterations (i.e., M iterations with M being the size of \mathbf{R}) to reach convergence. Interestingly, we show that the output SINR achieved by the CG-MF detector obtained by with $r + 1$ iterations is identical to that of the MF detector within a first-order approximation. Hence, $r + 1$ CG iterations suffice and there is no need for full iterations.

Throughout the correspondence, transpose and complex conjugate transpose are denoted by $(\cdot)^T$ and $(\cdot)^H$, respectively. $\|\cdot\|$ denotes the matrix Frobenius norm. $\mathcal{CN}(\boldsymbol{\mu}, \mathbf{R})$ denotes a complex circularly symmetric Gaussian random vector with mean $\boldsymbol{\mu}$ and covariance matrix \mathbf{R} .

II. DATA MODEL

Consider detecting a known J -channel signal $\mathbf{s}(n) \in \mathbb{C}^{J \times 1}$ corrupted by a spatio-temporally correlated disturbance $\mathbf{c}(n)$:

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

where $\mathbf{x}(n)$, $n = 1, 2, \dots, N$, denotes the n th received vector and a the deterministic unknown complex amplitude of $\mathbf{s}(n)$. Let $\mathbf{s} = [\mathbf{s}^T(1), \mathbf{s}^T(2), \dots, \mathbf{s}^T(N)]^T$, $\mathbf{c} = [\mathbf{c}^T(1), \mathbf{c}^T(2), \dots, \mathbf{c}^T(N)]^T$, and $\mathbf{x} = [\mathbf{x}^T(1), \mathbf{x}^T(2), \dots, \mathbf{x}^T(N)]^T$, which are $JN \times 1$ complex vectors. In STAP, \mathbf{s} is known as the space-time steering vector. For a side-looking uniform linear array (ULA), \mathbf{s} is given by

$$\mathbf{s} = \mathbf{s}_t \otimes \mathbf{s}_s \quad (2)$$

where $\mathbf{s}_t = \left(\frac{1}{\sqrt{N}}\right) \left[1, e^{i2\pi f_d}, \dots, e^{i2\pi(N-1)f_d}\right]^T$ is the temporal steering vector with a normalized Doppler frequency f_d , $\mathbf{s}_s = \left(\frac{1}{\sqrt{J}}\right) \left[1, e^{i2\pi f_s}, \dots, e^{i2\pi(J-1)f_s}\right]^T$ is the spatial steering vector with a normalized spatial frequency f_s , and \otimes denotes the Kronecker product. A standard assumption in STAP is that the disturbance \mathbf{c} , which includes clutter and noise, is complex circularly symmetric Gaussian with zero-mean and space-time covariance matrix \mathbf{R} [1]: $\mathbf{x} \sim \mathcal{CN}(a\mathbf{s}, \mathbf{R})$, with $a = 0$ under H_0 and $a \neq 0$ under H_1 .

III. CONJUGATE GRADIENT MATCHED FILTER

The optimum detector for (1) is the matched filter (MF) (e.g., [2]):

$$t_{\text{MF}} = \frac{|\mathbf{s}^H \mathbf{R}^{-1} \mathbf{x}|^2}{\mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \quad (3)$$

where η is the threshold of MF. We will frequently denote a detector by a linear filter weight vector. The weight vector of the MF is

$$\mathbf{w}_{\text{MF}} = \mathbf{R}^{-1} \mathbf{s}. \quad (4)$$

The MF test statistic can be alternatively expressed as:

$$t_{\text{MF}} = \frac{|\mathbf{w}_{\text{MF}}^H \mathbf{x}|^2}{\mathbf{w}_{\text{MF}}^H \mathbf{R} \mathbf{w}_{\text{MF}}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \quad (5)$$

For typical STAP applications, the covariance matrix \mathbf{R} has a large dimension. As a result, direct matrix inversion is usually not recommended to compute the weight vector (4) due to its computational complexity. We consider herein an alternative way by employing the conjugate (CG) algorithm [7], which iteratively finds a sequence of linear weight vectors $\mathbf{w}_k, k = 0, 1, \dots$, that are guaranteed to converge to the MF weight vector in no more than JN iterations. Each of the weight vector \mathbf{w}_k can be used to form a detector as in (5). As such, the CG iterations yield a family of detectors, referred to as the CG-MF detectors. To introduce necessary notation, the CG algorithm is briefly summarized as follows.

Initialization. Initialize the conjugate direction vector \mathbf{d}_1 , gradient vector $\boldsymbol{\gamma}_1$ and initial solution \mathbf{w}_0 :

$$\mathbf{d}_1 = -\boldsymbol{\gamma}_1 = \mathbf{s}, \quad \text{and} \quad \mathbf{w}_0 = \mathbf{0}. \quad (6)$$

for $k = 1, 2, \dots$, till convergence ($k \leq JN$) **do**

Update the step size α_k , weight vector \mathbf{w}_k , gradient vector $\boldsymbol{\gamma}_{k+1}$, and conjugate direction vector \mathbf{d}_{k+1} as follows:

$$\alpha_k = \frac{\|\boldsymbol{\gamma}_k\|^2}{\mathbf{d}_k^H \mathbf{R} \mathbf{d}_k} \quad (7)$$

$$\mathbf{w}_k = \mathbf{w}_{k-1} + \alpha_k \mathbf{d}_k \quad (8)$$

$$\boldsymbol{\gamma}_{k+1} = \boldsymbol{\gamma}_k + \alpha_k \mathbf{R} \mathbf{d}_k \quad (9)$$

$$\mathbf{d}_{k+1} = \mathbf{d}_k \frac{\|\boldsymbol{\gamma}_{k+1}\|^2}{\|\boldsymbol{\gamma}_k\|^2} - \boldsymbol{\gamma}_{k+1}. \quad (10)$$

end for

A quick comment on the complexity is in order. Each iteration of the CG algorithm involves one matrix-vector product, requiring about $O((JN)^2)$ flops. With full JN iterations, the CG algorithm has a complexity of $O((JN)^3)$ flops, comparable with alternative linear solvers such as the QR factorization [7]. In many practical cases, the CG algorithm may require far fewer than full iterations (see Section IV-B and also [7, Ch. 10] for discussions on the convergence of the CG algorithm), leading to significant reduction in complexity.

IV. ANALYSIS

We consider the performance of the CG-MF detectors in two cases. The first involves a general covariance matrix \mathbf{R} that is positive definite but otherwise arbitrary, whereas the other deals with a structured covariance that is frequently encountered in practice.

A. General Covariance Matrix

We first represent the CG-MF detector \mathbf{w}_k by using the conjugate direction vectors $\{\mathbf{d}_k\}$. From (8),

$$\mathbf{w}_k = \mathbf{D}_k \boldsymbol{\alpha}_k \quad (11)$$

where $\boldsymbol{\alpha}_k = [\alpha_1, \alpha_2, \dots, \alpha_k]^T$ contains the stepsizes and $\mathbf{D}_k = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_k]$ consists of the first k conjugate direction directors. Note that \mathbf{D}_k diagonalizes the covariance matrix \mathbf{R} [7, p.523]:

$$\mathbf{D}_k^H \mathbf{R} \mathbf{D}_k = \boldsymbol{\Lambda}_k \quad (12)$$

where $\boldsymbol{\Lambda}_k = \text{diag}(u_1^2, u_2^2, \dots, u_k^2)$ and $u_k = (\mathbf{d}_k^H \mathbf{R} \mathbf{d}_k)^{\frac{1}{2}}$. This allows $\boldsymbol{\alpha}_k$ to be compactly expressed as

$$\boldsymbol{\alpha}_k = \boldsymbol{\Lambda}_k^{-1} \mathbf{D}_k^H \mathbf{s} \quad (13)$$

which gives the following close-form expression for \mathbf{w}_k :

$$\mathbf{w}_k = \mathbf{D}_k \boldsymbol{\Lambda}_k^{-1} \mathbf{D}_k^H \mathbf{s}. \quad (14)$$

The k th CG-MF detector using \mathbf{w}_k is given by

$$t_k = \frac{|\mathbf{w}_k^H \mathbf{x}|^2}{\mathbf{w}_k^H \mathbf{R} \mathbf{w}_k} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_k. \quad (15)$$

Theorem 1: The following are true for the k th CG-MF detector.

a) \mathbf{w}_k is a linear minimum mean square estimator that minimizes

$$\mathbf{w}_k = \arg \min_{\mathbf{w} \in \mathcal{K}(\mathbf{R}, \mathbf{s}, k)} \left\{ \mathbb{E} \|\mathbf{a} - \mathbf{w}^H \mathbf{x}\|^2 \right\} \quad (16)$$

among all linear estimators within the k -dimensional Krylov subspace

$$\mathcal{K}(\mathbf{R}, \mathbf{s}, k) \triangleq \text{span}\{\mathbf{s}, \mathbf{R}\mathbf{s}, \mathbf{R}^2\mathbf{s}, \dots, \mathbf{R}^{k-1}\mathbf{s}\}.$$

b) \mathbf{w}_k yields the largest output SINR

$$\rho_k = |a|^2 \frac{|\mathbf{w}_k^H \mathbf{s}|^2}{\mathbf{w}_k^H \mathbf{R} \mathbf{w}_k} \quad (17)$$

among all linear detectors within $\mathcal{K}(\mathbf{R}, \mathbf{s}, k)$.

c) Under the assumption $\mathbf{x} \sim \mathcal{CN}(a\mathbf{s}, \mathbf{R})$ with $a = 0$ under H_0 and $a \neq 0$ under H_1 , the probability of false alarm $P_{\text{fa},k}$ and the probability of detection $P_{\text{d},k}$ of \mathbf{w}_k are

$$P_{\text{fa},k} = \exp(-\eta_k) \quad (18)$$

$$P_{\text{d},k} = Q(\sqrt{2\rho_k}, \sqrt{2\eta_k}) \quad (19)$$

where $Q(\cdot)$ is the Marcum Q function.

Sketch of Proof: As the above results are quick extensions of standard knowledge, only a sketch of proof is provided. Result a) is due to that the Krylov subspace is also spanned by the conjugate direction vectors [7]: $\mathcal{K}(\mathbf{R}, \mathbf{s}, k) = \text{span}\{\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_k\}$. As such,

$$\begin{aligned} \mathbf{w}_k &= \arg \min_{\mathbf{w} = \mathbf{D}_k \boldsymbol{\alpha}} \left\{ \mathbb{E} \|\mathbf{a} - \mathbf{w}^H \mathbf{x}\|^2 \right\} \\ &= \mathbf{D}_k \left(\mathbf{D}_k^H \mathbf{R} \mathbf{D}_k \right)^{-1} \mathbf{D}_k^H \mathbf{s} = \mathbf{D}_k \boldsymbol{\Lambda}_k^{-1} \mathbf{D}_k^H \mathbf{s} \end{aligned} \quad (20)$$

which is identical to (14).

Result b) follows immediately from a), since minimizing the mean-square error (MSE) is equivalent to maximizing the SINR [13].

To show c), we use (12) and (11) to rewrite the test variable t_k of (15) as

$$t_k = \frac{|\boldsymbol{\alpha}_k^H \mathbf{D}_k^H \mathbf{x}|^2}{\boldsymbol{\alpha}_k^H \boldsymbol{\Lambda}_k \boldsymbol{\alpha}_k} = |y_k|^2 \quad (21)$$

where $y_k = (\boldsymbol{\alpha}_k^H \mathbf{D}_k^H \mathbf{x})(\boldsymbol{\alpha}_k^H \boldsymbol{\Lambda}_k \boldsymbol{\alpha}_k)^{-\frac{1}{2}}$. Clearly, $y_k \sim \mathcal{CN}(0, 1)$ under H_0 and $y_k \sim \mathcal{CN}\left(\left(\boldsymbol{\alpha}_k^H \mathbf{D}_k^H \mathbf{s}\right)(\boldsymbol{\alpha}_k^H \boldsymbol{\Lambda}_k \boldsymbol{\alpha}_k)^{-\frac{1}{2}}, 1\right)$ under H_1 . As such t_k is central and, respectively, noncentral Chi-square distributed under H_0 and H_1 , where the noncentrality parameter under H_1 is given by ρ_k (17). Hence, (18) and (19) follow immediately. ■

A number of remarks are in order. First, (18) implies that CG-MF detectors for all k are CFAR detectors. Their test variables t_k are all identically distributed to that of the MF test variable t_{MF} , irrespective of k the number of iterations. Second, b) implies that $\rho_k \leq \rho_{k+1}$, since $\mathcal{K}(\mathbf{R}, \mathbf{s}, k) \subseteq \mathcal{K}(\mathbf{R}, \mathbf{s}, k+1)$. Hence, the CG-MF is a family of CFAR detectors \mathbf{w}_k composed of both reduced-rank detectors ($k < JN$) and the full-rank MF detector ($k = JN$), which offers a natural way to trade complexity for performance. Specifically, the detection probability $P_{d,k}$ of the CG-MF detector \mathbf{w}_k increases with more CG iterations (i.e., a larger k), at higher computational complexity. The tradeoff and the analytical expression (19) allow one to save the computational cost by selecting an appropriate reduced-rank CG-MF detector that offers a targeted $P_{d,k}$, without going through all CG iterations.

B. CG-MF: Structured Covariance Matrix

We now examine the performance of the CG-MF detectors when the disturbance covariance matrix has a low-rank structure:

$$\mathbf{R}_i + \sigma_n^2 \mathbf{I} \quad (22)$$

where \mathbf{R}_i is a rank- r positive semi-definite matrix ($r < JN$) and \mathbf{I} an identity matrix. In such a case, the CG algorithm is known to converge using at most $r+1$ iterations [7], i.e., $\mathbf{w}_{r+1} = \mathbf{w}_{\text{MF}}$.

Many practical applications involve a disturbance covariance matrix having a structure similar to (22). For example, in airborne radars, the disturbance covariance matrix often consists of a low-rank \mathbf{R}_i due to the clutter and jamming and a scaled identity $\sigma_n^2 \mathbf{I}$ due to the thermal noise, where σ_n^2 denotes the noise variance. The rank r 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 [14], we have

$$r \approx \lceil J + (N-1)\beta \rceil \quad (23)$$

where $\beta = \frac{2v_g T_r}{d}$, 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.

Estimating the rank r can be a tricky issue since (23) may not hold for all clutter scenarios encountered in practice. The CG algorithm has an advantage of not requiring to know r a priori, since at the $(r+1)$ st iteration, the residual $\mathbf{s} - \mathbf{R}\mathbf{w}_k$, which is also the negative gradient $\boldsymbol{\gamma}_k$, vanishes. This is the stopping rule used by the CG [7]. Other STAP detectors designed to take advantage of the structure (22), such as the low rank normalized matched filter (LRNMF) [15] which employs the principal eigenvectors of the covariance matrix, requires an estimate of r and its performance is quite sensitive to the accuracy of the estimate.

While the convergence of the CG for a structured covariance matrix exactly like (22) is well known, we consider a different but related case

that \mathbf{R} is a perturbed version of (22):

$$\mathbf{R} = \mathbf{R}_i + \sigma_n^2 \mathbf{I} + \boldsymbol{\Delta} \triangleq \mathbf{R}_0 + \boldsymbol{\Delta} \quad (24)$$

where $\mathbf{R}_0 = \mathbf{R}_i + \sigma_n^2 \mathbf{I}$ as in (22) and $\boldsymbol{\Delta}$ is a Hermitian perturbation matrix assumed to be small, i.e., $\|\boldsymbol{\Delta}\| \ll \|\mathbf{R}_0\|$. Since the perturbation is small, it is of interest to examine the following questions: Can the CG algorithm reach (almost) convergence in $r+1$ iterations? How is the detection performance of the CG-MF detector \mathbf{w}_{r+1} compared to the MF detector? Before we address these questions, we note that the model (24) exists in many scenarios. For example, in airborne radar applications, the covariance matrix \mathbf{R} may not have exactly $r+1$ distinct eigenvalues as in (22). Typically, \mathbf{R} contains a few principal eigenvalues due to the dominant clutter scatterers, but the other eigenvalues are rarely identical and spread around the noise level [1]. By decomposing \mathbf{R} as (24), \mathbf{R}_i contains only the dominant clutter scatterers, and the effect of the less significant clutter scatterers can be included in $\boldsymbol{\Delta}$. The same can be extended to a general interference scenario, where \mathbf{R}_i includes the effect of a few major interference sources to be mitigated at the receiver, whereas $\boldsymbol{\Delta}$ contains the residual interference.

To answer the previous questions, we first present a result that relates the weight vectors for the two detectors.

Lemma 1: Consider the linear equation $\mathbf{R}\mathbf{w}_{\text{MF}} = \mathbf{s}$, where $\mathbf{R} = \mathbf{R}_i + \sigma_n^2 \mathbf{I} + \boldsymbol{\Delta} = \mathbf{R}_0 + \boldsymbol{\Delta}$ is a positive-definite Hermitian matrix, \mathbf{R}_i is a rank- r positive semi-definite Hermitian matrix, $\sigma_n^2 > 0$ is a constant, and $\boldsymbol{\Delta}$ is a Hermitian perturbation matrix. If the perturbation is small such that $\|\boldsymbol{\Delta}\| \ll \|\mathbf{R}_0\|$, the MF solution \mathbf{w}_{MF} can be approximated by the CG-MF solution \mathbf{w}_{r+1} , with the approximation error given by

$$\mathbf{w}_{\text{MF}} - \mathbf{w}_{r+1} = \mathbf{R}_0^{-\frac{1}{2}} \mathbf{P}_{\tilde{T}_r}^\perp \mathbf{R}_0^{-\frac{1}{2}} \mathbf{d} + o(\|\boldsymbol{\Delta}\|) \quad (25)$$

where $o(\|\boldsymbol{\Delta}\|)$ contains the second- and higher-order perturbation terms that can be neglected for small $\|\boldsymbol{\Delta}\|$,

$$\tilde{T}_r = \mathbf{R}_0^{\frac{1}{2}} T_r \quad (26)$$

$$T_r = [\mathbf{s}, \mathbf{R}_0 \mathbf{s}, \mathbf{R}_0^2 \mathbf{s}, \dots, \mathbf{R}_0^r \mathbf{s}]_{JN \times (r+1)} \quad (27)$$

$$\mathbf{P}_{\tilde{T}_r}^\perp = \mathbf{I} - \tilde{T}_r (\tilde{T}_r^H \tilde{T}_r)^{-1} \tilde{T}_r^H \quad (28)$$

$$\mathbf{d} = \left(\boldsymbol{\Delta} \mathbf{R}_0^{-1} + \mathbf{R}_0 \boldsymbol{\Phi}_{\boldsymbol{\Delta}}^{(r)} (T_r^H \mathbf{R}_0 T_r)^{-1} T_r^H \right) \mathbf{s} \in \mathbb{C}^{JN \times 1} \quad (29)$$

and $\boldsymbol{\Phi}_{\boldsymbol{\Delta}}^{(r)} \in \mathbb{C}^{JN \times (r+1)}$ is given by

$$\boldsymbol{\Phi}_{\boldsymbol{\Delta}}^{(r)} = \left[\mathbf{0}, \boldsymbol{\Delta} \mathbf{s}, \mathbf{R}_0 \boldsymbol{\Delta} \mathbf{s} + \boldsymbol{\Delta} \mathbf{R}_0 \mathbf{s}, \dots, \sum_{k=1}^r \mathbf{R}_0^{r-k} \boldsymbol{\Delta} \mathbf{R}_0^{k-1} \mathbf{s} \right]. \quad (30)$$

Proof: See the Appendix. ■

When the perturbation vanishes, it is straightforward to show from (48) of the Appendix that

$$\mathbf{w}_{\text{MF}} - \mathbf{w}_{r+1} = \mathbf{R}_0^{-\frac{1}{2}} \mathbf{P}_{\tilde{T}_r}^\perp \mathbf{R}_0^{-\frac{1}{2}} \mathbf{s} = \mathbf{0} \quad (31)$$

which offers another proof that the CG-MF detector \mathbf{w}_{r+1} converges to the MF detector \mathbf{w}_{MF} when $\boldsymbol{\Delta} = \mathbf{0}$ [7]. Interestingly, the first equality of the above equation resembles (25) except that \mathbf{d} is replaced by \mathbf{s} . It is

also noted that the matrix $\Phi_{\Delta}^{(r)}$ can be easily calculated by the following simple recursion:

$$\begin{aligned} \phi_1 &= \mathbf{0} \\ \phi_k &= \Delta \mathbf{R}_0^{k-2} \mathbf{s} + \mathbf{R}_0 \phi_{k-1}, \quad k = 2, 3, \dots, r+1. \end{aligned} \quad (32)$$

We now consider the output SINR of the CG-MF detector \mathbf{w}_{r+1} ,

$$\rho_{r+1} = |a|^2 \frac{|\mathbf{w}_{r+1}^H \mathbf{s}|^2}{\mathbf{w}_{r+1}^H \mathbf{R} \mathbf{w}_{r+1}} \quad (33)$$

and its relation to the output SINR of the MF detector. The following result addresses their relationship.

Theorem 2: Under the conditions stated in Lemma 1, the output SINRs of the MF detector \mathbf{w}_{MF} and the CG-MF detector \mathbf{w}_{r+1} are identical within a first-order approximation:

$$\delta_{\rho} = \rho_{\text{MF}} - \rho_{r+1} = o(\|\Delta\|) \quad (34)$$

where ρ_{r+1} is given by (33) and ρ_{MF} is similarly defined by replacing \mathbf{w}_{r+1} in (33) with \mathbf{w}_{MF} .

Proof: The proof goes by direct calculation and using Lemma 1. The loss of output SINR of the CG-MF relative to the MF is given by

$$\delta_{\rho} = \rho_{\text{MF}} - \rho_{r+1} = |a|^2 \left(\frac{|\mathbf{w}_{\text{MF}}^H \mathbf{s}|^2}{\mathbf{w}_{\text{MF}}^H \mathbf{R} \mathbf{w}_{\text{MF}}} - \frac{|\mathbf{w}_{r+1}^H \mathbf{s}|^2}{\mathbf{w}_{r+1}^H \mathbf{R} \mathbf{w}_{r+1}} \right). \quad (35)$$

First, we consider the difference between $\mathbf{s}^H \mathbf{w}_{\text{MF}}$ and $\mathbf{s}^H \mathbf{w}_{r+1}$. Using Lemma 1, we have

$$\mathbf{s}^H \mathbf{w}_{\text{MF}} - \mathbf{s}^H \mathbf{w}_{r+1} = \mathbf{s}^H \mathbf{R}_0^{-\frac{1}{2}} \mathbf{P}_{\mathbf{T}_r}^{\perp} \mathbf{R}_0^{-\frac{1}{2}} \mathbf{d} + o(\|\Delta\|). \quad (36)$$

Since \mathbf{s} is orthogonal to the column space of $\mathbf{R}_0^{-\frac{1}{2}} \mathbf{P}_{\mathbf{T}_r}^{\perp} \mathbf{R}_0^{-\frac{1}{2}}$ [also see (31)], we have $\mathbf{s}^H \mathbf{R}_0^{-\frac{1}{2}} \mathbf{P}_{\mathbf{T}_r}^{\perp} \mathbf{R}_0^{-\frac{1}{2}} = 0$ and (36) reduces to

$$\mathbf{s}^H \mathbf{w}_{\text{MF}} - \mathbf{s}^H \mathbf{w}_{r+1} = o(\|\Delta\|). \quad (37)$$

It follows that

$$|\mathbf{w}_{\text{MF}}^H \mathbf{s}|^2 - |\mathbf{w}_{r+1}^H \mathbf{s}|^2 = o(\|\Delta\|). \quad (38)$$

Next, we consider the difference between the denominators $\mathbf{w}_{\text{MF}}^H \mathbf{R} \mathbf{w}_{\text{MF}}$ and $\mathbf{w}_{r+1}^H \mathbf{R} \mathbf{w}_{r+1}$:

$$\begin{aligned} & \mathbf{w}_{\text{MF}}^H \mathbf{R} \mathbf{w}_{\text{MF}} - \mathbf{w}_{r+1}^H \mathbf{R} \mathbf{w}_{r+1} \\ &= 2\Re\{\mathbf{w}_{\text{MF}}^H \mathbf{R}(\mathbf{w}_{\text{MF}} - \mathbf{w}_{r+1})\} \\ & \quad - (\mathbf{w}_{\text{MF}} - \mathbf{w}_{r+1})^H \mathbf{R}(\mathbf{w}_{\text{MF}} - \mathbf{w}_{r+1}) \\ &= 2\Re\{\mathbf{s}^H(\mathbf{w}_{\text{MF}} - \mathbf{w}_{r+1})\} \\ & \quad - (\mathbf{w}_{\text{MF}} - \mathbf{w}_{r+1})^H \mathbf{R}(\mathbf{w}_{\text{MF}} - \mathbf{w}_{r+1}) \end{aligned} \quad (39)$$

where $\Re\{\cdot\}$ denotes the real part. Again using Lemma 1, we have

$$(\mathbf{w}_{\text{MF}} - \mathbf{w}_{r+1})^H \mathbf{R}(\mathbf{w}_{\text{MF}} - \mathbf{w}_{r+1}) = o(\|\Delta\|). \quad (40)$$

Substituting (37) and (40) into (39) yields

$$\mathbf{w}_{\text{MF}}^H \mathbf{R} \mathbf{w}_{\text{MF}} - \mathbf{w}_{r+1}^H \mathbf{R} \mathbf{w}_{r+1} = o(\|\Delta\|). \quad (41)$$

Finally, from (35), the output SINR loss of the CG-MF detector is given by

$$\delta_{\rho} = |a|^2 \frac{|\mathbf{w}_{\text{MF}}^H \mathbf{s}|^2 \mathbf{w}_{r+1}^H \mathbf{R} \mathbf{w}_{r+1} - |\mathbf{w}_{r+1}^H \mathbf{s}|^2 \mathbf{w}_{\text{MF}}^H \mathbf{R} \mathbf{w}_{\text{MF}}}{\mathbf{w}_{\text{MF}}^H \mathbf{R} \mathbf{w}_{\text{MF}} \mathbf{w}_{r+1}^H \mathbf{R} \mathbf{w}_{r+1}}. \quad (42)$$

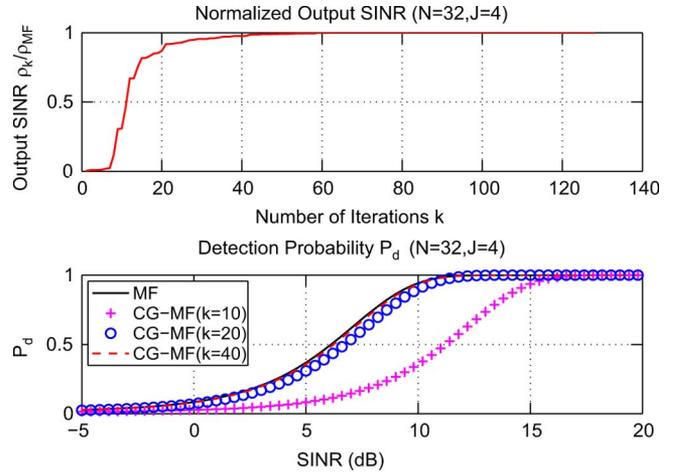


Fig. 1. General covariance matrix with $J = 4$, $N = 32$ and $P_{\text{fa}} = 0.01$. Upper: normalized output SINR versus the number of iterations for the CG-MF detector. Lower: Probability of detection versus SINR for the MF and CG-MF detectors with several different numbers of iterations.

Substituting (38) and (41) into the numerator of (42), we have

$$\begin{aligned} & |\mathbf{w}_{\text{MF}}^H \mathbf{s}|^2 \mathbf{w}_{r+1}^H \mathbf{R} \mathbf{w}_{r+1} - |\mathbf{w}_{r+1}^H \mathbf{s}|^2 \mathbf{w}_{\text{MF}}^H \mathbf{R} \mathbf{w}_{\text{MF}} \\ &= |\mathbf{w}_{\text{MF}}^H \mathbf{s}|^2 \left(\mathbf{w}_{\text{MF}}^H \mathbf{R} \mathbf{w}_{\text{MF}} - o(\|\Delta\|) \right) \\ & \quad - \left(|\mathbf{w}_{\text{MF}}^H \mathbf{s}|^2 - o(\|\Delta\|) \right) \mathbf{w}_{\text{MF}}^H \mathbf{R} \mathbf{w}_{\text{MF}} = o(\|\Delta\|) \end{aligned} \quad (43)$$

from which (34) immediately follows. \blacksquare

Remark: It is interesting to note that while Lemma 1 indicates that the difference between the weight vectors, i.e., $\mathbf{w}_{\text{MF}} - \mathbf{w}_{r+1}$, contains first-order terms of the perturbation, such first-order differences vanish in the output SINR. Theorem 2 implies that the probabilities of detection of the MF and CG-MF detectors are also identical within a first-order approximation. This has important practical implication. In particular, even though the CG algorithm using \mathbf{R} generally requires full (i.e., JN) iterations before it reaches convergence, to save computation, we can take the intermediate result \mathbf{w}_{r+1} obtained at the $(r+1)$ st iteration and obtain nearly the same detection performance as the MF detector, provided that Δ is sufficiently small.

V. NUMERICAL RESULTS

A. General Covariance Matrix

We first consider the general covariance matrix case studied in Section IV-A. We use a disturbance covariance matrix \mathbf{R} obtained from the KASSPER data set [16], which is a simulated data set that includes practical airborne radar parameters and issues found in a real-world clutter environment. The radar platform considered in this data set has 11 horizontal antenna elements. For simplicity, we use only the outputs of the first $J = 4$ channels for processing. The number of pulses is $N = 32$, and the probability of false alarm is $P_{\text{fa}} = 0.01$. We first examine the output SINR ρ_k , defined in (17), of the CG-MF detector \mathbf{w}_k . Fig. 1 (upper) shows the normalized output SINR $\frac{\rho_k}{\rho_{\text{MF}}}$, where the normalizing factor ρ_{MF} is the output SINR of the MF detector, versus the number of iterations. It is observed that ρ_k converges rapidly to ρ_{MF} as k increases.

The probability of detection for the MF detector and the CG-MF detector after $k = 10, 20$ and 40 iterations, respectively, is shown in Fig. 1 (lower) as a function of the MF output SINR, defined as $\rho_{\text{MF}} \stackrel{\Delta}{=} |a|^2 \mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}$. It is seen that with $k = 40$ iterations, the CG-MF

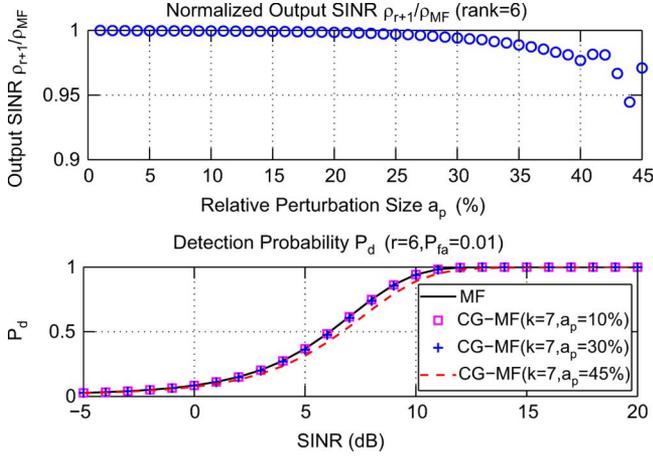


Fig. 2. Structured covariance matrix with $r = 6$, $JN = 16$, and $P_{fa} = 0.01$. Upper: Output SINR of the MF detector and CG-MF detector after $r + 1 = 7$ iterations versus the relative perturbation size. Lower: Probability of detection of the MF detector and the CG-MF detector after $r + 1 = 7$ iterations.

detector achieves nearly identical detection performance as the MF detector, which requires $JN = 128$ CG iterations.

B. Structured Covariance Matrix With Perturbation

We now consider the case examined in Section IV-B where the covariance \mathbf{R} is a perturbed version of a structured \mathbf{R}_0 . We demonstrate how the convergence of the CG-MF detector is directly affected by the size of the perturbation. We employ a *relative perturbation size*, defined as

$$a_p = \frac{\|\Delta\|}{\|\mathbf{R}_0\|}. \quad (44)$$

We tried several ways of generating the perturbation matrix Δ and obtained similar results. The ones presented here were based on the following approach. For any structured covariance matrix \mathbf{R}_0 as described in (24) and a given perturbation size a_p , 1) randomly generate \mathbf{R}' as a complex Wishhart matrix \mathbf{R}' with mean \mathbf{R}_0 and compute the difference $\Delta' = \mathbf{R}' - \mathbf{R}_0$; 2) compute Δ as: $\Delta = a_p \frac{\|\mathbf{R}_0\|}{\|\Delta'\|} \Delta'$; and 3) the perturbed covariance matrix is given by $\mathbf{R} = \mathbf{R}_0 + \Delta$. It is noted that although \mathbf{R} is generated as a random matrix, in each trial \mathbf{R} is treated as a deterministic/known matrix that is a perturbed version of \mathbf{R}_0 with perturbation size a_p .

Fig. 2 (upper) shows the output SINR of the CG-MF detector \mathbf{w}_{r+1} ($r = 6$) normalized by that of the MF detector, i.e., $\frac{\rho_{r+1}}{\rho_{MF}}$, as a function of a_p . It is seen that the output SINRs of the two detectors remain nearly identical ($\frac{\rho_{r+1}}{\rho_{MF}} > 0.99$) for a relative perturbation size as large as $a_p = 30\%$, which indicates that our perturbation analysis in Theorem 2 for the CG-MF detectors is quite accurate over a wide range of perturbation size. Fig. 2 (lower) depicts the probability of detection for the MF and CG-MF detector as a function of the MF output SINR, where several values of a_p are considered. It is seen that with a relative perturbation size as large as $a_p = 30\%$, the detection probability of the two detectors are nearly identical. At $a_p = 45\%$, a small difference is observed.

VI. CONCLUSION

The CG algorithm can be used to solve the Wiener-Hopf equation underlying the MF, which leads to a family of linear CG-MF detectors that converge to the MF in a fixed number of iterations. We have shown that the CG-MF detectors are all CFAR detectors, they can be recursively and efficiently computed via CG iterations over an expanding

Krylov subspace, and each of them is an optimum reduced-dimensional detector in the sense that it yields the maximum output SINR over all linear detectors residing the Krylov subspace. For disturbance covariances with a low-rank structure (rank- r), we have shown that the presence of a perturbation component Δ disrupting the low-rank structure has minimum effect on the convergence of the CG algorithm, in that the output SINR of the $(r + 1)$ st CG-MF detector is nearly identical to that of the MF detector. This offers significant computational saving, in particular when r is small, by using the CG-MF instead of the MF detector without incurring undue penalty in detection performance. A future topic of interest is to analyze the CG algorithm for adaptive detection when the covariance matrix \mathbf{R} is unknown and estimated from training signals.

APPENDIX PROOF OF LEMMA 1

Proof: The CG-MF solution \mathbf{w}_{r+1} obtained at the $(r + 1)$ st iteration is the \mathbf{R} -orthogonal projection of \mathbf{w}_{MF} onto the Krylov subspace $\mathcal{K}(\mathbf{R}, \mathbf{s}, r + 1)$ [7]. This means that the \mathbf{R} -norm of the approximation error is minimized over all vectors in $\mathcal{K}(\mathbf{R}, \mathbf{s}, r + 1)$, which is the column space of $\mathbf{S}_r = [\mathbf{s}, \mathbf{R}\mathbf{s}, \mathbf{R}^2\mathbf{s}, \mathbf{R}^3\mathbf{s}, \dots, \mathbf{R}^r\mathbf{s}]$ [7]. That is,

$$\|\mathbf{w}_{MF} - \mathbf{w}_{r+1}\|_{\mathbf{R}} = \min_{\mathbf{a}_k} \left\| \mathbf{w} - \sum_{k=0}^r a_k \mathbf{R}^k \mathbf{s} \right\|_{\mathbf{R}}. \quad (45)$$

Substituting $\mathbf{w} = \mathbf{R}^{-1}\mathbf{s}$ and $\|\cdot\|_{\mathbf{R}} = \|\mathbf{R}^{\frac{1}{2}}(\cdot)\|$ into (45), we have

$$\begin{aligned} \|\mathbf{w}_{MF} - \mathbf{w}_{r+1}\|_{\mathbf{R}} &= \min_{\mathbf{a}_k} \left\| \mathbf{R}^{\frac{1}{2}} (\mathbf{R}^{-1}\mathbf{s} - \sum_{k=0}^r a_k \mathbf{R}^k \mathbf{s}) \right\| \\ &= \min_{\mathbf{a}_k} \left\| \mathbf{R}^{-\frac{1}{2}} \mathbf{s} - \sum_{k=0}^r a_k \mathbf{R}^{\frac{1}{2}} \mathbf{R}^k \mathbf{s} \right\|. \end{aligned} \quad (46)$$

The minimum approximation error is achieved if and only if the vector $\sum_{k=0}^r a_k \mathbf{R}^{\frac{1}{2}} \mathbf{R}^k \mathbf{s}$ is the orthogonal projection of the vector $\mathbf{R}^{-\frac{1}{2}} \mathbf{s}$ onto the linearly transformed Krylov subspace

$$\mathbf{R}^{\frac{1}{2}} \mathcal{K}(\mathbf{R}, \mathbf{s}, r + 1) = \text{span}\{\mathbf{R}^{\frac{1}{2}} \mathbf{s}, \mathbf{R}^{\frac{1}{2}} \mathbf{R}\mathbf{s}, \dots, \mathbf{R}^{\frac{1}{2}} \mathbf{R}^r \mathbf{s}\} \quad (47)$$

or the column space of $\tilde{\mathbf{S}}_r = \mathbf{R}^{\frac{1}{2}} \mathbf{S}_r$. When the minimum of (46) is achieved, the approximation error is given by

$$\mathbf{w}_{MF} - \mathbf{w}_{r+1} = \mathbf{R}^{-\frac{1}{2}} \mathbf{P}_{\tilde{\mathbf{S}}_r}^{\perp} \mathbf{R}^{-\frac{1}{2}} \mathbf{s} \quad (48)$$

where

$$\begin{aligned} \mathbf{P}_{\tilde{\mathbf{S}}_r}^{\perp} &= \mathbf{I} - \tilde{\mathbf{S}}_r (\tilde{\mathbf{S}}_r^H \tilde{\mathbf{S}}_r)^{-1} \tilde{\mathbf{S}}_r^H \\ &= \mathbf{I} - \mathbf{R}^{\frac{1}{2}} \mathbf{S}_r (\mathbf{S}_r^H \mathbf{R} \mathbf{S}_r)^{-1} \mathbf{S}_r^H \mathbf{R}^{\frac{1}{2}} \end{aligned} \quad (49)$$

which is the orthogonal complement projection matrix of the transformed Krylov subspace $\mathbf{R}^{\frac{1}{2}} \mathcal{K}(\mathbf{R}, \mathbf{s}, r + 1)$. Substituting (49) into (48), we have

$$\mathbf{w}_{MF} - \mathbf{w}_{r+1} = \left[\mathbf{R}^{-1} - \mathbf{S}_r (\mathbf{S}_r^H \mathbf{R} \mathbf{S}_r)^{-1} \mathbf{S}_r^H \right] \mathbf{s}. \quad (50)$$

Since $\mathbf{w}_{MF} = \mathbf{R}^{-1}\mathbf{s}$, and the vector $\mathbf{S}_r (\mathbf{S}_r^H \mathbf{R} \mathbf{S}_r)^{-1} \mathbf{S}_r^H \mathbf{s} \in \mathcal{K}(\mathbf{R}, \mathbf{s}, r + 1)$, so $\mathbf{w}_r = \mathbf{S}_r (\mathbf{S}_r^H \mathbf{R} \mathbf{S}_r)^{-1} \mathbf{S}_r^H \mathbf{s}$, and

$$\mathbf{w}_{MF} - \mathbf{w}_{r+1} = \left[\mathbf{R}^{-1} - \mathbf{S}_r (\mathbf{S}_r^H \mathbf{R} \mathbf{S}_r)^{-1} \mathbf{S}_r^H \right] \mathbf{s}. \quad (51)$$

Expanding $\mathbf{R}^m = (\mathbf{R}_0 + \Delta)^m$, we have

$$\begin{aligned} \mathbf{R}^m &= \mathbf{R}_0^m + \sum_{k=1}^m \mathbf{R}_0^{m-k} \Delta \mathbf{R}_0^{k-1} + o(\|\Delta\|) \\ &\approx \mathbf{R}_0^m + \sum_{k=1}^m \mathbf{R}_0^{m-k} \Delta \mathbf{R}_0^{k-1}. \end{aligned} \quad (52)$$

If the columns of \mathbf{T}_r span the Krylov subspace $\mathcal{K}(\mathbf{R}_0, \mathbf{s}, r+1)$, then \mathbf{S}_r can be approximated by

$$\mathbf{S}_r \approx \mathbf{T}_r + \Phi_{\Delta}^{(r)} \quad (53)$$

where $\Phi_{\Delta}^{(r)}$ is defined by (30). After substituting (53) into (51), while using a first-order expansion on $(\mathbf{S}_r^H \mathbf{R} \mathbf{S}_r)^{-1}$, $(\mathbf{S}_r^H \mathbf{R} \mathbf{S}_r)^{-1}$ can be approximated as

$$\begin{aligned} (\mathbf{S}_r^H \mathbf{R} \mathbf{S}_r)^{-1} &= \left[(\mathbf{T}_r + \Phi_{\Delta}^{(r)})^H (\mathbf{R}_0 + \Delta) (\mathbf{T}_r + \Phi_{\Delta}^{(r)}) \right]^{-1} \\ &\approx (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} - (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \Psi_{\Delta}^{(r)} (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \end{aligned} \quad (54)$$

where $\Psi_{\Delta}^{(r)} = \mathbf{T}_r^H \mathbf{R}_0 \Phi_{\Delta}^{(r)} + \mathbf{T}_r^H \Delta \mathbf{T}_r + \Phi_{\Delta}^{(r)H} \mathbf{R}_0 \mathbf{T}_r$. Similarly using expansion on \mathbf{R}^{-1} , we have

$$\mathbf{R}^{-1} = (\mathbf{R}_0 + \Delta)^{-1} \approx \mathbf{R}_0^{-1} - \mathbf{R}_0^{-1} \Delta \mathbf{R}_0^{-1}. \quad (55)$$

Substituting (54) and (55) into (51), and discarding the second and higher-order terms, we have

$$\begin{aligned} \mathbf{w}_{\text{MF}} - \mathbf{w}_{r+1} &\approx \left[\mathbf{R}_0^{-1} - \mathbf{T}_r (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \mathbf{T}_r^H \right. \\ &\quad + \mathbf{T}_r (\mathbf{T}_r^H \mathbf{R} \mathbf{T}_r)^{-1} \Psi_{\Delta}^{(r)} (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \mathbf{T}_r^H \\ &\quad - \Phi_{\Delta}^{(r)} (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \mathbf{T}_r^H - \mathbf{R}_0^{-1} \Delta \mathbf{R}_0^{-1} \\ &\quad \left. - \mathbf{T}_r (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \Phi_{\Delta}^{(r)H} \right] \mathbf{s}. \end{aligned} \quad (56)$$

Since \mathbf{R}_0 is a rank- r correction matrix to $\sigma_n^2 \mathbf{I}$, the solution $\mathbf{R}_0^{-1} \mathbf{s}$ lies in the Krylov subspace $\mathcal{K}(\mathbf{R}_0, \mathbf{s}, r+1)$ [7]. Hence, $\mathbf{R}_0^{-\frac{1}{2}} \mathbf{s}$ lies in the linearly transformed Krylov subspace $\mathbf{R}_0^{\frac{1}{2}} \mathcal{K}(\mathbf{R}_0, \mathbf{s}, r+1)$ or the column space of $\tilde{\mathbf{T}}_r = \mathbf{R}_0^{\frac{1}{2}} \mathbf{T}_r$, and as such $\mathbf{P}_{\tilde{\mathbf{T}}_r}^{\perp} \mathbf{R}_0^{-\frac{1}{2}} \mathbf{s}$ is equal to zero vector, i.e.,

$$\begin{aligned} \mathbf{P}_{\tilde{\mathbf{T}}_r}^{\perp} \mathbf{R}_0^{-\frac{1}{2}} \mathbf{s} &= \left[\mathbf{I} - \mathbf{R}_0^{\frac{1}{2}} \mathbf{T}_r \left((\mathbf{R}_0^{\frac{1}{2}} \mathbf{T}_r)^H (\mathbf{R}_0^{\frac{1}{2}} \mathbf{T}_r) \right)^{-1} \mathbf{T}_r^H \mathbf{R}_0^{\frac{1}{2}} \right] \mathbf{R}_0^{-\frac{1}{2}} \mathbf{s} \\ &= (\mathbf{R}_0^{-\frac{1}{2}} - \mathbf{R}_0^{\frac{1}{2}} \mathbf{T}_r (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \mathbf{T}_r^H) \mathbf{s} = 0. \end{aligned} \quad (57)$$

Since \mathbf{R}_0 is a positive-definite Hermitian matrix, left multiplying both sides of (57) by $\mathbf{R}_0^{\frac{1}{2}}$ yields

$$\left[\mathbf{R}_0^{-1} - \mathbf{T}_r (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \mathbf{T}_r^H \right] \mathbf{s} = 0. \quad (58)$$

Substituting (58) into (56), we obtain the difference

$$\begin{aligned} \mathbf{w}_{\text{MF}} - \mathbf{w}_{r+1} &\approx \left[\mathbf{T}_r (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \left(\mathbf{T}_r^H \mathbf{R}_0 \Phi_{\Delta}^{(r)} \right. \right. \\ &\quad + \mathbf{T}_r^H \Delta \mathbf{T}_r + \Phi_{\Delta}^{(r)H} \mathbf{R}_0 \mathbf{T}_r \left. \right) (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \mathbf{T}_r^H \\ &\quad - \Phi_{\Delta}^{(r)} (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \mathbf{T}_r^H - \mathbf{R}_0^{-1} \Delta \mathbf{R}_0^{-1} \\ &\quad \left. - \mathbf{T}_r (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \Phi_{\Delta}^{(r)H} \right] \mathbf{s}. \end{aligned} \quad (59)$$

Note that

$$\begin{aligned} \mathbf{T}_r (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \mathbf{T}_r^H \Delta \mathbf{T}_r (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \mathbf{T}_r^H \mathbf{s} \\ = \mathbf{R}_0^{-\frac{1}{2}} \mathbf{P}_{\tilde{\mathbf{T}}_r} \mathbf{R}_0^{-\frac{1}{2}} \Delta \mathbf{R}_0^{-\frac{1}{2}} \mathbf{P}_{\tilde{\mathbf{T}}_r} \mathbf{R}_0^{-\frac{1}{2}} \mathbf{s} \end{aligned} \quad (60)$$

where $\mathbf{P}_{\tilde{\mathbf{T}}_r} = \tilde{\mathbf{T}}_r (\tilde{\mathbf{T}}_r^H \tilde{\mathbf{T}}_r)^{-1} \tilde{\mathbf{T}}_r^H$ is the orthogonal projection matrix onto the column space of $\tilde{\mathbf{T}}_r$. From the previous analysis, the vector $\mathbf{R}_0^{-\frac{1}{2}} \mathbf{s}$ lies in the column space of $\tilde{\mathbf{T}}_r$. Therefore,

$$\mathbf{P}_{\tilde{\mathbf{T}}_r} \mathbf{R}_0^{-\frac{1}{2}} \mathbf{s} = \mathbf{R}_0^{-\frac{1}{2}} \mathbf{s}. \quad (61)$$

Substituting (61) into (60), we have

$$\begin{aligned} \mathbf{T}_r (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \mathbf{T}_r^H \Delta \mathbf{T}_r (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \mathbf{T}_r^H \mathbf{s} \\ = \mathbf{R}_0^{-\frac{1}{2}} \mathbf{P}_{\tilde{\mathbf{T}}_r} \mathbf{R}_0^{-\frac{1}{2}} \Delta \mathbf{R}_0^{-1} \mathbf{s} \end{aligned} \quad (62)$$

and

$$\begin{aligned} \left[\mathbf{R}_0^{-1} \Delta \mathbf{R}_0^{-1} - \mathbf{T}_r (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \mathbf{T}_r^H \Delta \mathbf{T}_r (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \mathbf{T}_r^H \right] \mathbf{s} \\ = \mathbf{R}_0^{-\frac{1}{2}} \mathbf{P}_{\tilde{\mathbf{T}}_r}^{\perp} \mathbf{R}_0^{-\frac{1}{2}} \Delta \mathbf{R}_0^{-1} \mathbf{s}. \end{aligned} \quad (63)$$

Similarly, it can be proved that

$$\begin{aligned} \left[\Phi_{\Delta}^{(r)} (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \mathbf{T}_r^H \right. \\ \left. - \mathbf{T}_r (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \mathbf{T}_r^H \mathbf{R}_0 \Phi_{\Delta}^{(r)} (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \mathbf{T}_r^H \right] \mathbf{s} \\ = \mathbf{R}_0^{-\frac{1}{2}} \mathbf{P}_{\tilde{\mathbf{T}}_r}^{\perp} \mathbf{R}_0^{\frac{1}{2}} \Phi_{\Delta}^{(r)} (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \mathbf{T}_r^H \mathbf{s} \end{aligned} \quad (64)$$

and

$$\begin{aligned} \left[\mathbf{T}_r (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \Phi_{\Delta}^{(r)H} \right. \\ \left. - \mathbf{T}_r (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \Phi_{\Delta}^{(r)H} \mathbf{R}_0 \mathbf{T}_r (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \mathbf{T}_r^H \right] \mathbf{s} \\ = \mathbf{T}_r (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \Phi_{\Delta}^{(r)H} \mathbf{P}_{\tilde{\mathbf{T}}_r}^{\perp} \mathbf{R}_0^{-\frac{1}{2}} \mathbf{s} = 0. \end{aligned} \quad (65)$$

Substituting (63)–(65) into (59), we have

$$\begin{aligned} \mathbf{w}_{\text{MF}} - \mathbf{w}_{r+1} &\approx \mathbf{R}_0^{-\frac{1}{2}} \mathbf{P}_{\tilde{\mathbf{T}}_r}^{\perp} \mathbf{R}_0^{-\frac{1}{2}} \\ &\quad \times \left[\Delta \mathbf{R}_0^{-1} + \mathbf{R}_0 \Phi_{\Delta}^{(r)} (\mathbf{T}_r^H \mathbf{R}_0 \mathbf{T}_r)^{-1} \mathbf{T}_r^H \right] \mathbf{s} \\ &= \mathbf{R}_0^{-\frac{1}{2}} \mathbf{P}_{\tilde{\mathbf{T}}_r}^{\perp} \mathbf{R}_0^{-\frac{1}{2}} \mathbf{d} \end{aligned} \quad (66)$$

where \mathbf{d} is defined by (29). ■

REFERENCES

- [1] J. Ward, "Space-time adaptive processing for airborne radar," Lincoln Laboratory, Massachusetts Inst. of Technol. (MIT), Cambridge, MA, Tech. Rep. 1015, 1994.
- [2] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR adaptive matched filter detector," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 1, pp. 208–216, Jan. 1992.
- [3] A. Haimovich, "The eigencanceler: Adaptive radar by eigenanalysis methods," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 532–542, Apr. 1996.
- [4] I. P. Kirsteins and D. W. Tufts, "Adaptive detection using low rank approximation to a data matrix," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 30, no. 1, pp. 55–67, Jan. 1994.
- [5] J. S. Goldstein and I. S. Reed, "Reduced-rank adaptive filtering," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 492–496, Feb. 1997.
- [6] 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. 2943–2959, Nov. 1998.
- [7] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.
- [8] P. S. Chang and A. N. Wilson, Jr, "Analysis of conjugate gradient algorithms for adaptive filtering," *IEEE Trans. Signal Process.*, vol. 48, no. 2, pp. 409–418, Feb. 2000.
- [9] M. E. Weippert, J. D. Hiemstra, J. S. Goldstein, and M. D. Zoltowski, "Insights from the relationship between the multistage Wiener filter and the method of conjugate gradients," in *Proc. Sensor Array Multichannel Signal Process. Workshop*, Rosslyn, VA, Aug. 2002, pp. 388–392.
- [10] L. L. Scharf, E. K. P. Chong, M. D. Zoltowski, J. S. Goldstein, and I. S. Reed, "Subspace expansion and the equivalence of conjugate direction and multistage wiener filters," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5013–5019, Oct. 2008.
- [11] C. Jiang, H. Li, and M. Rangaswamy, "Conjugate gradient parametric detection of multichannel signals," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 1, Apr. 2012.
- [12] I. R. Kirsteins and H. Ge, "Performance analysis of Krylov space adaptive beamformers," in *Proc. 4th IEEE Workshop Sensor Array Multichannel Process. (SAM)*, Waltham, MA, Jul. 2006, pp. 16–20.
- [13] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE Signal Process. Mag.*, pp. 4–24, Apr. 1988.
- [14] L. E. Brennan and I. S. Reed, "Theory of adaptive radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-9, no. 2, pp. 237–252, 1973.
- [15] M. Rangaswamy, F. C. Lin, and K. R. Gerlach, "Robust adaptive signal processing methods for heterogeneous radar clutter scenarios," *Signal Process.*, vol. 84, no. 9, pp. 1477–1733, Sep. 2004.
- [16] J. S. Bergin and P. M. Techau, "High-Fidelity Site-Specific Radar Simulation: KASSPER'02 Workshop Datacube," Information Syst. Laboratories, Inc., Vienna, VA, Tech. Rep. ISL-SCRD-TR-02-105, 2002.

Inverse Methods for Reconstruction of Channel Taps in OFDM Systems

Tomasz Hrycak, Saptarshi Das, and Gerald Matz

Abstract—We describe a novel pilot-aided method for estimation of doubly selective wireless channels in OFDM systems. We compute the first few Fourier coefficients of each channel tap from the pilot information. We then estimate the BEM coefficients of the channel taps from their respective Fourier coefficients using a recently developed inverse reconstruction method. For a system with L channel taps, the proposed method uses $\mathcal{O}(L \log L)$ operations and $\mathcal{O}(L)$ memory per OFDM symbol.

We validate our method by simulating a system conforming to the IEEE 802.16e standard.

Index Terms—Basis Expansion Model (BEM), channel estimation, doubly selective, inverse reconstruction method, OFDM.

I. INTRODUCTION

A. Motivation

Orthogonal frequency-division multiplexing (OFDM) is a multicarrier modulation technique with several advantages, e.g., high spectral efficiency and robustness against multipath propagation. OFDM based communications through *rapidly varying* doubly selective wireless channels attract a great deal of scientific and commercial interest. OFDM is used in high-mobility communication systems, e.g., Mobile WiMAX (IEEE 802.16e), WAVE (IEEE 802.11p), and DVB-T (ETSI EN 300 744).

Communication systems using multicarrier modulation schemes over doubly selective channels are affected by intercarrier interference (ICI), which makes equalization more difficult. ICI is caused by the Doppler effect, and the carrier frequency offset. The Doppler effect is proportional to the receiver velocity and the carrier frequency and depends inversely on the intercarrier frequency offset.

In the case of scalable OFDM, the required bandwidth grows with the number of subcarriers. Increasing the bandwidth increases the sampling frequency, which in turn proportionally increases the number of resolvable discrete multipaths. For example, Mobile WiMAX with K subcarriers typically exhibits a discrete path delay of $\frac{K}{8}$; see [1]. A large number of channel taps makes channel estimation much harder. Such challenging regimes require an accurate channel estimation algorithm, whose complexity scales with the number of subcarriers.

B. Previous Work

Doubly selective channels, whose taps vary with time, are commonly estimated using the Basis Expansion Model (BEM); see [2] and [3]. The BEM approximates the channel taps by linear combinations of prescribed basis functions. In this approach, channel estimation reduces to estimation of the basis coefficients of the channel taps. Several bases have been proposed for modeling doubly selective channels. The BEM with complex exponentials (CE-BEM) [4] gives rise to a banded

Manuscript received August 02, 2011; revised November 18, 2011; accepted January 11, 2012. Date of publication February 10, 2012; date of current version April 13, 2012. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Lieven De Lathauwer. This paper was supported by the FWF Grant S10602-N13.

T. Hrycak and S. Das are with the Faculty of Mathematics, University of Vienna, Vienna, Austria (e-mail: saptarshi.das@univie.ac.at).

G. Matz is with Institut für Nachrichtentechnik und Hochfrequenztechnik, Vienna University of Technology, Vienna, Austria.

Digital Object Identifier 10.1109/TSP.2012.2187522