Conjugate Gradient Adaptive Matched Filter

ZHU CHEN, Member, IEEE
HONGBIN LI, Senior Member, IEEE
Stevens Institute of Technology
Hoboken, NJ, USA

MURALIDHAR RANGASWAMY, Fellow, IEEE
Air Force Research Laboratory
Wright Patterson Air Force Base, OH, USA

We consider an adaptive reduced-rank detector, referred to as the CG-AMF detector, which is obtained by using the conjugate gradient (CG) algorithm to solve for the weight vector of the adaptive matched filter (AMF). The CG is a computationally efficient iterative algorithm, which finds the projection of the AMF weight vector to the Krylov subspace with a dimension growing with the CG iterations. This effectively leads to a family of reduced-rank detectors indexed by the number of CG iterations. The main purpose of this paper is to examine the output signal-to-interference-and-noise ratio (SINR) of the CG-AMF detector in the presence of strong clutter/interference. Specifically, by exploiting a connection between the CG algorithm and the Lanczos algorithm, we show the output SINR can be asymptotically expressed in a simple form involving a Ritz vector of the sample covariance matrix. The probability density function (pdf) and expected value of the output SINR are then obtained based on this approximation. Our theoretical analysis of the CG-AMF detector is verified by computer simulation. Numerical comparisons are also made with several popular reduced-rank detectors using either data-independent or data-dependent rank reduction approaches. Our results show that for a fixed training size, the CG-AMF detector often reaches its peak output SINR with a lower rank compared with the other reduced-rank detectors, which implies that the CG-AMF detector has lower computational complexity and less training requirement.

I. INTRODUCTION

The problem of detecting a multichannel signal from temporally and spatially correlated disturbance is encountered in a variety of applications including radar, sonar, wireless communications, and others [1–3]. The problem has been extensively studied under the framework of space-time adaptive processing (STAP) in phased-array radar, which employs multielement antennas and multipulse waveforms to probe and observe the radar scene. Numerous STAP detectors have been proposed [1, 2, 4–14]. Among them, the covariance matrix based detectors, which need knowledge of the space-time covariance matrix of the disturbance signal to suppress the interferences, are the most widely used multichannel signal detectors. Examples of such detectors include the Reed, Mallett, and Brennan detector [4], Kelly's generalized likelihood ratio test (GLRT) [5], the adaptive matched filter (AMF) detector [6, 7], the adaptive coherence estimator detector [8], among others. All of them involve estimating and inverting a space-time covariance matrix of the disturbance signal for each test cell using target-free training data, which may impose excessive training and computational burdens when the joint space-time dimension is large.

Aiming at mitigating the training and computational requirements of the full-rank covariance matrix based detectors, reduced-rank techniques have been proposed to reduce the dimension of the input signal in advance of detection [1, 2]. Specifically, reduced-rank methods employ rank reduction mechanisms to decrease the degrees of freedom (DoFs) of the detectors so as to alleviate the estimation burden. For a fixed training size, reduced-rank methods often offer a better estimation accuracy than their full-rank counterparts and, despite the loss of DoFs, may actually attain a better detection performance.

The rank reduction of the input signal space can be achieved by a linear transformation matrix **T**. There are two broad families of reduced-rank transformations, namely, data-independent transformations and data-dependent transformations. The former include using submatrices formed from the discrete Fourier transform (DFT), discrete cosine transform (DCT), or other data-independent linear matrices. Examples of data-dependent reduced-rank transformations include the eigencanceler (EIG) [9], the cross-spectral metric (CSM) [10], and others. The main advantage of data-independent transformations is their low computational cost. Data-dependent schemes are normally computationally more involved. For example, the EIG and CSM require the eigendecomposition of a covariance or transformed covariance matrix. Despite the higher complexity, better performance is often obtained by using data-dependent transformations.

The conjugate gradient (CG) algorithm is a computationally efficient method for solving a linear system. The CG algorithm is guaranteed to converge in a

fixed number of iterations, and able to provide a series of approximations of the solution in an expanding Krylov subspace. Due to these properties, the CG algorithm has been investigated for estimation and detection in a number of recent studies [15–19]. Specifically, [15] considered the use of CG for adaptive filtering. Krylov subspace beamforming was studied in [16]. The CG algorithm was employed to provide efficient order recursive implementation of the post-Doppler STAP detector in [17]. Parametric adaptive detection exploiting the computational efficiency of the CG algorithm for linear prediction was proposed in [18]. The CG algorithm was used to approximate the nonadaptive matched filter in [19] and a number of convergence properties were established.

We consider the CG-AMF detector for adaptive reduced-rank detection. Specifically, the computationally efficient CG algorithm is employed to iteratively calculate the weight vector of the AMF detector. This in turn produces a group of CG-AMF detectors, each having a different rank determined by the number of CG iterations. The CG-AMF detector is closely related to the multistage Wiener filter (MWF), which is a popular reduced-rank detector [20, 21]. In particular, it has been found that under certain conditions (namely, CG starts the iteration from a zero initial condition and the MWF employs an orthogonal transfer matrix), the two detectors are identical in the sense that they produce the same weight vector at each iteration [20].

The main purpose of this paper is to analyze the output signal-to-interference-and-noise ratio (SINR) of the CG-AMF detectors for different iterations numbers. It is natural to ask which CG-AMF detector yields the highest output SINR on the average. It should be noted that due to a trade-off between DoFs and estimation accuracy when the training size is fixed, the output SINR does not always grow with the rank or the number of CG iterations. An optimal rank is usually achieved well before the CG reaches its full iterations (at which point it yields a full-rank solution). Any additional CG iterations beyond the optimal rank are therefore not only computationally wasteful, but also reduce the detection performance. To this end, we examine the statistical behavior of the output SINR of the CG-AMF detector. Under the condition that the signal is contaminated by strong low-rank interference, an asymptotic expression of the probability density function (pdf) and expected value of the output SINR are obtained by using a connection between the CG and the Lanczos algorithms. Computer simulations are then used to verify the analysis and comparisons are made to several other reduced-rank detectors.

The rest of the paper is organized as follows. Section II presents the signal detection problem and an overview of reduced-rank detection. The CG-AMF detector is introduced in Section III. Section IV presents the asymptotic performance analysis of the CG-AMF detector. Section V contains numerical simulations to verify our analysis and compare the CG-AMF detector

with other reduced-rank detectors. Finally, conclusions are given in Section VI.

Notation: Vectors and matrices are denoted by boldface lowercase and uppercase letters, respectively. Transpose, complex conjugate, and complex conjugate transpose are, respectively, represented by $(.)^T$, $(.)^*$, and $(.)^H$. $\mathbb{C}$ and $\mathbb{R}$ denote the complex and real number fields. $\mathcal{CN}(\boldsymbol{\mu}, \mathbf{R})$ denotes the multivariate complex Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{R}$, $\Re\{.\}$ denotes the real part of a complex variable.

II. DATA MODEL AND REDUCED-RANK DETECTION

Consider the following signal detection problem [1, 2, 4–10, 18, 19, 21]:

$$\begin{aligned} H_0 : & \quad \mathbf{x} = \mathbf{d} \\ H_1 : & \quad \mathbf{x} = \alpha \mathbf{s} + \mathbf{d} \end{aligned} \tag{1}$$

where $\mathbf{x} \in \mathbb{C}^{M \times 1}$ denotes the observation, $\mathbf{s}$ denotes a deterministic signal of an unknown complex amplitude α, and $\mathbf{d}$ denotes a disturbance signal that is assumed to be spatially and temporally correlated. The disturbance $\mathbf{d}$ is often modeled as a Gaussian random vector with zero-mean and space-time covariance matrix $\mathbf{R} \in \mathbb{C}^{M \times M}$. As a result, $\mathbf{x} \sim \mathcal{CN}(\alpha \mathbf{s}, \mathbf{R})$, where $\alpha = 0$ under H_0 and $\alpha \neq 0$ under H_1.

The above problem is referred to as the STAP problem in array radar literature. Specifically, the $M \times 1$ vector $\mathbf{x}$ contains the samples obtained with an array of J elements and over N temporal pulses, $M = JN$ denotes the joint space-time dimension, $\mathbf{s}$ is the space-time steering vector, and $\mathbf{d}$ contains clutter and noise. For a side-looking uniform linear array, the space-time steering vector $\mathbf{s}$ is given by $\mathbf{s} = \mathbf{s}_t \otimes \mathbf{s}_s$ where $\mathbf{s}_t = (1/\sqrt{N})[1 \ e^{i2\pi f_d} \cdots e^{i2\pi(N-1)f_d}]^T$ is the spatial steering vector with a normalized Doppler frequency f_d, $\mathbf{s}_s = (1/\sqrt{J})[1 \ e^{i2\pi f_s} \cdots e^{i2\pi(J-1)f_s}]^T$ is the spatial steering vector with a normalized spatial frequency f_s, and $\otimes$ denotes the Kronecker product.

The adaptive matched filter (AMF) test [7] is a popular solution to the detection problem (1), which is given by

$$\frac{\left|\hat{\mathbf{w}}^H \mathbf{x}\right|^2}{\hat{\mathbf{w}}^H \mathbf{s}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{\text{AMF}} \tag{2}$$

where η_{AMF} denotes the test threshold, and the linear weight vector of the AMF detector is

$$\hat{\mathbf{w}} = \hat{R}^{-1} \mathbf{s} \tag{3}$$

where $\widehat{\mathbf{R}}$ denotes the sample covariance matrix which is an estimate of $\mathbf{R}$ obtained from L training data $\{\mathbf{x}_l\}_{l=1}^L$, which are assumed to be independent and identically distributed (IID) with the same covariance matrix $\mathbf{R}$:

$$\hat{\mathbf{R}} = \frac{1}{L} \sum_{l=1}^L \mathbf{x}_l \mathbf{x}_l^H. \tag{4}$$

The AMF is a full-rank detector that uses all M DoFs for detection. In general, a STAP detector is not

recommended in training-limited scenarios, since the performance loss due to estimation error can be substantial when L is small. Reduced-rank detection offers an alternative approach when training data is limited. There are two broad families of reduced-rank detectors: data-independent methods and data-dependent methods. The former methods apply a data-independent transformation matrix $\mathbf{T} \in \mathbb{C}^{k \times M}, k < M$ to the received signal $\mathbf{x}$ followed by its computational efficiency standard detection. For example, a reduced-rank version of the AMF has a weight vector given by

$$\hat{\mathbf{w}} = \mathbf{T}^H (\mathbf{T}\hat{\mathbf{R}}\mathbf{T}^H)^{-1} \mathbf{T}\mathbf{s}. \qquad (5)$$

Among others, submatrices of the DFT and DCT matrices are popular choices for data-independent reduced-rank processing due to their computational efficiency. Many pre-Doppler and post-Doppler reduced-rank techniques (e.g., [1]) also belong to the data-independent categories.

While a data-independent reduced-rank detector uses a deterministic and fixed transformation matrix $\mathbf{T}$, a data-dependent reduced-rank detector employs some rank reduction approach that is derived from the observed signal. For example, the EIG [9] utilizes the noise eigenvectors $\hat{\mathbf{E}}_n \in \mathbb{C}^{M \times k}$ of the sample covariance matrix $\hat{\mathbf{R}}$, where k is any integer between 1 and the maximum number of noise eigenvectors of $\hat{\mathbf{R}}$. $\hat{\mathbf{E}}_n$ spans a noise subspace that is orthogonal to the dominant interference contained in the observed signal. It also plays the role of rank reduction like the $\mathbf{T}$ matrix for data-independent reduced-rank methods. The weight vector of the EIG can be expressed as

$$\hat{\mathbf{w}}_E = \hat{\mathbf{E}}_n \hat{\mathbf{E}}_n^H \mathbf{s}. \qquad (6)$$

Another well-known example of data-dependent reduced-rank detector is the CSM-based detector [10], whose weight vector is given by

$$\hat{\mathbf{w}}_{CSM} = [\mathbf{I}_M - \mathbf{A}^H \mathbf{U}(\mathbf{U}^H \mathbf{A}\hat{\mathbf{R}}\mathbf{A}^H \mathbf{U})^{-1}\mathbf{U}^H \mathbf{A}\hat{\mathbf{R}}]\mathbf{s} \qquad (7)$$

where $\mathbf{A}$ is a signal blocking matrix satisfying $\mathbf{As} = \mathbf{0}$, and $\mathbf{U}$ is formed from the k eigenvectors of $\mathbf{A}\hat{\mathbf{R}}\mathbf{A}^H$ that maximize the quantity $|\mathbf{q}_i^H \mathbf{A}\hat{\mathbf{R}}\mathbf{s}|^2/\lambda_i$, with $\mathbf{q}_i$ and λ_i denoting the eigenvectors and eigenvalues of $\mathbf{A}\hat{\mathbf{R}}\mathbf{A}^H$, respectively.

In general, data-dependent reduced-rank AMF detectors such as the EIG and CSM outperform the data-independent reduced-rank AMF detectors, especially in training limited cases (L is small). However, these detectors require eigenvalue decomposition and are computationally intensive. Meanwhile, the CG algorithm offers a computationally efficient approach to reduced-rank detection. A recent study of the CG algorithm used with the nonadaptive matched filter (MF) for reduced-rank detection leads to not only computational complexity reduction but also interesting insights of CG for reduced-rank detection [19]. The problem of interest to this work is to examine the performance of the CG for adaptive reduced-rank detection relative to other data-dependent reduced-rank detectors.

III. CG-AMF DETECTOR

The CG algorithm was employed in [19] along with the nonadaptive MF for reduced-rank detection. It is straightforward to extend it for adaptive reduced-rank detection. In particular, the CG algorithm can iteratively provide a sequence of approximations $\hat{\mathbf{w}}_k, k = 1, 2, \ldots$, to the AMF weight vector $\hat{\mathbf{w}}$. Each $\hat{\mathbf{w}}_k$ can be used to form a detector as in (2). A summary of the iterative algorithm that can be used to compute the CG-AMF detector is in Table I. A few comments on the algorithm are in order. At the k-th CG iteration, the goal is to find the best approximation to the AMF weight vector (3) over the k-dimensional Krylov subspace:

$$\mathcal{K}(\hat{\mathbf{R}}, \mathbf{s}, \mathbf{k}) \triangleq \text{span}\{\mathbf{s}, \hat{\mathbf{R}}\mathbf{s}, \hat{\mathbf{R}}^2\mathbf{s}, \cdots, \hat{\mathbf{R}}^{k-1}\mathbf{s}\}. \qquad (8)$$

The Krylov subspace is also spanned by a set of basis vectors $\hat{\mathbf{d}}_1, \ldots, \hat{\mathbf{d}}_k$ referred to as the conjugate direction vectors which are $\hat{\mathbf{R}}-$conjugate, i.e., $\hat{\mathbf{d}}_i^H \hat{\mathbf{R}}\hat{\mathbf{d}}_j = 0$ for $i \neq j$. The $\hat{\mathbf{R}}-$conjugate property leads to decoupled coefficients $\hat{\alpha}_k$. As a result, after the k-th conjugate direction vector $\hat{\mathbf{d}}_k$ becomes available, the weight vector only needs to be updated along the direction of $\hat{\mathbf{d}}_k$, by using (10), and the k-th coefficient $\hat{\alpha}_k$ can be thought of as a step-size corresponding to the amount of change along the direction of $\hat{\mathbf{d}}_k$. The gradient vector $\hat{\gamma}_{k+1}$ computed by (11) is effectively the residual $\mathbf{s} - \hat{\mathbf{R}}\hat{\mathbf{w}}_{k+1}$. Given the gradient, the next conjugate direction vector $\hat{\mathbf{d}}_{k+1}$ can be efficiently computed by (12). More details about the CG iterations can be found in [22].

The CG algorithm converges to the full-rank AMF solution in no more than M iterations [22]. Even faster convergence is possible if the covariance matrix is structured. For example, it is known that if the covariance matrix $\hat{\mathbf{R}}$ contains a rank-r component plus an identity matrix, then the CG algorithm converges in no more than $r + 1$ iterations [22].

In a training-limited environment, however, convergence to the full-rank adaptive solution is often not the desired objective, since the full-rank solution may suffer considerable estimation error induced performance loss. Meanwhile, a reduced-rank solution may yield better detection performance. Since the CG iteration yields a family of reduced-rank CG-AMF detectors $\hat{\mathbf{w}}_k$, it is naturally of interest to investigate the statistical behavior of these CG-AMF detectors and determine how the detection performance changes as a function of the number of CG iterations.

IV. PERFORMANCE ANALYSIS

In this section, we consider the performance of the CG-AMF detector in terms of its output SINR. Since in adaptive detection, the output SINR is a random variable,

TABLE I
Conjugate-Gradient AMF Detector

Input:
 $M \times M$ sample covariance matrix $\hat{\mathbf{R}}$
 $M \times 1$ signal vector s
Output:
 $M \times 1$ CG-AMF weight vectors $\hat{\mathbf{w}}_k$, $k = 1, 2, \ldots$
Algorithm:
 1) **Initialization:**
  Initial conjugate-direction vector: $\hat{\mathbf{d}}_1 = \mathbf{s}$
  Initial gradient vector: $\hat{\gamma}_1 = -\mathbf{s}$
  Initial weight vector: $\hat{\mathbf{w}}_0 = \mathbf{0}$
 2) **Iterations:**
  for $k = 1, 2, \ldots$, till convergence ($k \leq M$)
  do
   a) Update the step size $\hat{\alpha}_k$:
$$\hat{\alpha}_k = \frac{\|\hat{\gamma}_k\|^2}{\hat{\mathbf{d}}_k^H \hat{\mathbf{R}} \hat{\mathbf{d}}_k} \qquad (9)$$
   b) Update the solution $\hat{\mathbf{w}}_k$:
$$\hat{\mathbf{w}}_k = \hat{\mathbf{w}}_{k-1} + \hat{\alpha}_k \hat{\mathbf{d}}_k \qquad (10)$$
   c) Update the gradient vector $\hat{\gamma}_{k+1}$:
$$\hat{\gamma}_{k+1} = \hat{\gamma}_k + \hat{\alpha}_k \hat{\mathbf{R}} \hat{\mathbf{d}}_k \qquad (11)$$
   d) Update the conjugate-direction vector $\hat{\mathbf{d}}_{k+1}$:
$$\hat{\mathbf{d}}_{k+1} = \hat{\mathbf{d}}_k \frac{\|\hat{\gamma}_{k+1}\|^2}{\|\hat{\gamma}_k\|^2} - \hat{\gamma}_{k+1} \qquad (12)$$
 end for

it is necessary to consider the statistical distribution of the output SINR.

A. Output SINR of the CG-AMF Detector

The CG-AMF weight vector $\hat{\mathbf{w}}_k$ after k iterations is in the k-dimensional Krylov subspace $\mathcal{K}(\hat{\mathbf{R}}, \mathbf{s}, k)$. The normalized gradient vectors $\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2, \cdots, \hat{\mathbf{q}}_k$, where $\hat{\mathbf{q}}_i = \frac{\hat{\gamma}_i}{\|\hat{\gamma}_i\|_2}$, $1 \leq i \leq k$, also span the same k-dimensional Krylov subspace:

$$\mathcal{K}(\hat{\mathbf{R}}, \mathbf{s}, k) = \text{span}\{\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2, \cdots, \hat{\mathbf{q}}_k\}. \qquad (13)$$

Therefore, by defining the matrix of residuals $\hat{\mathbf{Q}}_k \in \mathbb{C}^{M \times k}$ as

$$\hat{\mathbf{Q}}_k = \begin{bmatrix} \hat{\mathbf{q}}_1 & \hat{\mathbf{q}}_2 & \cdots & \hat{\mathbf{q}}_k \end{bmatrix} \qquad (14)$$

$\hat{\mathbf{w}}_k$ can be compactly expressed as

$$\hat{\mathbf{w}}_k = \hat{\mathbf{Q}}_k \hat{\mathbf{a}} \qquad (15)$$

where $\hat{\mathbf{a}} = [\hat{a}_1 \ \hat{a}_2 \cdots \hat{a}_k]^T$ contains the coefficients which can be determined as follows.

Specifically, $\hat{\mathbf{w}}_k$ can be considered as the $\hat{\mathbf{R}}$−orthogonal projection of $\hat{\mathbf{w}} = \hat{\mathbf{R}}^{-1}\mathbf{s}$ onto the Krylov subspace $\mathcal{K}(\hat{\mathbf{R}}, \mathbf{s}, k)$ [22], which means that the $\hat{\mathbf{R}}$−norm of the approximation error is minimized over all vectors in the Krylov subspace or, equivalently, the column space of $\hat{\mathbf{Q}}_k$. That is,

$$\|\hat{\mathbf{w}} - \hat{\mathbf{w}}_k\|_{\hat{\mathbf{R}}} = \min_{\hat{\mathbf{a}}} \|\hat{\mathbf{w}} - \hat{\mathbf{w}}_k\|_{\hat{\mathbf{R}}}$$
$$= \min_{\hat{\mathbf{a}}} \left\| \hat{\mathbf{R}}^{\frac{1}{2}} \hat{\mathbf{w}} - \hat{\mathbf{R}}^{\frac{1}{2}} \hat{\mathbf{w}}_k \right\| \qquad (16)$$

where the $\hat{\mathbf{R}}$−norm is defined as $\| \cdot \|_{\hat{\mathbf{R}}} = \|\hat{\mathbf{R}}^{\frac{1}{2}}(\cdot)\|$. Denote by ϵ the approximation error

$$\boldsymbol{\epsilon} \triangleq \hat{\mathbf{R}}^{\frac{1}{2}}\hat{\mathbf{w}} - \hat{\mathbf{R}}^{\frac{1}{2}}\hat{\mathbf{w}}_k = \hat{\mathbf{R}}^{-\frac{1}{2}}\mathbf{s} - \hat{\mathbf{R}}^{\frac{1}{2}}\hat{\mathbf{Q}}_k\hat{\mathbf{a}}. \qquad (17)$$

Since $\hat{\mathbf{R}}^{\frac{1}{2}}\hat{\mathbf{w}}_k$ is the orthogonal projection of the vector $\hat{\mathbf{R}}^{\frac{1}{2}}\hat{\mathbf{w}}$ onto $\hat{\mathbf{R}}^{\frac{1}{2}}\mathcal{K}(\hat{\mathbf{R}}, \mathbf{s}, k)$, we have $\varepsilon \perp \hat{\mathbf{R}}^{\frac{1}{2}}\hat{\mathbf{Q}}_k$, namely,

$$\left(\hat{\mathbf{R}}^{-\frac{1}{2}}\mathbf{s} - \hat{\mathbf{R}}^{\frac{1}{2}}\hat{\mathbf{Q}}_k\hat{\mathbf{a}} \right)^H \hat{\mathbf{R}}^{\frac{1}{2}}\hat{\mathbf{Q}}_k = \mathbf{0} \qquad (18)$$

from which we obtain

$$\hat{\mathbf{a}} = \left(\hat{\mathbf{Q}}_k^H \hat{\mathbf{R}} \hat{\mathbf{Q}}_k \right)^{-1} \hat{\mathbf{Q}}_k^H \mathbf{s}. \qquad (19)$$

Therefore, the CG-AMF weight vector can be written as

$$\hat{\mathbf{w}}_k = \hat{\mathbf{Q}}_k \left(\hat{\mathbf{Q}}_k^H \hat{\mathbf{R}} \hat{\mathbf{Q}}_k \right)^{-1} \hat{\mathbf{Q}}_k^H \mathbf{s} \qquad (20)$$

where $\hat{\mathbf{Q}}_k$ plays the same role as the data-dependent reduced-rank transform matrix $\mathbf{T}$ in Section II.

For the full-rank AMF detector, the output SINR is given by

$$\rho \triangleq \frac{|\alpha|^2 |\hat{\mathbf{w}}^H \mathbf{s}|^2}{\hat{\mathbf{w}}^H \mathbf{R} \hat{\mathbf{w}}}. \qquad (21)$$

Replacing $\hat{\mathbf{w}}$ in (21) by $\hat{\mathbf{w}}_k$ of (20), we have the output SINR of the CG-AMF detector with k CG iterations as

$$\rho_k = \frac{|\alpha|^2 \left| \mathbf{s}^H \hat{\mathbf{Q}}_k (\hat{\mathbf{Q}}_k^H \hat{\mathbf{R}} \hat{\mathbf{Q}}_k)^{-1} \hat{\mathbf{Q}}_k^H \mathbf{s} \right|^2}{\mathbf{s}^H \hat{\mathbf{Q}}_k (\hat{\mathbf{Q}}_k^H \hat{\mathbf{R}} \hat{\mathbf{Q}}_k)^{-1} \hat{\mathbf{Q}}_k^H \mathbf{R} \hat{\mathbf{Q}}_k (\hat{\mathbf{Q}}_k^H \hat{\mathbf{R}} \hat{\mathbf{Q}}_k)^{-1} \hat{\mathbf{Q}}_k^H \mathbf{s}}. \qquad (22)$$

B. Low-Rank Approximation

Henceforth, we consider the case when the disturbance covariance matrix has some low-rank structure. Specifically, we assume that $\mathbf{R}$ has the following structure:

$$\mathbf{R} = \mathbf{R}_i + \sigma_n^2 \mathbf{I} \qquad (23)$$

where $\mathbf{R}_i$ is a rank-r ($r < M$) positive semidefinite matrix and $\mathbf{I}$ is an identify matrix. The above structure of the covariance matrix is frequently encountered in practice. For example, in airborne radar applications, $\mathbf{R}$ may consist of two components, namely a low-rank $\mathbf{R}_i$ due to the presence of clutter and jamming and a scaled identity $\sigma_n^2 \mathbf{I}$ due to the presence of a thermal noise with variance σ_n^2. The rank r is typically much smaller than the joint spatial-temporal dimension $M = JN$ where J is the number of array elements and N is the number of pulses. Specifically, if the disturbance is primarily due to ground clutter and thermal noise, then according to Brennan's rule [23], the rank of the covariance matrix for the full-dimensional MF is approximately

$$r \approx \lceil J + (N - 1)\beta \rceil \qquad (24)$$

where $\beta = 2 v_g T_r / d$, 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.

In airborne radar detection, the clutter is often the dominating factor compared with the noise. Let $\lambda(\mathbf{R}) = \{\lambda_1, \ldots, \lambda_M\}$ denote the spectrum of $\mathbf{R}$ with the eigenvalues in descending order: $\lambda_1 \geq \cdots \geq \lambda_r \gg \lambda_{r+1} = \cdots = \lambda_M = \sigma_n^2$, where the first r dominant eigenvalues are due to the clutter while the rest are due to the noise. Similarly, let $\lambda(\hat{\mathbf{R}}) = \{\hat{\lambda}_1, \cdots, \hat{\lambda}_M\}$ denote the spectrum of the sample covariance matrix $\hat{\mathbf{R}}$ with descending order: $\hat{\lambda}_1 \geq \cdots \geq \hat{\lambda}_r \gg \hat{\lambda}_{r+1} \geq \cdots \geq \hat{\lambda}_M$. For a sufficiently large L, the number of training signals [cf. (4)], the eigenvalues of $\hat{\mathbf{R}}$, like $\mathbf{R}$, are clustered with r dominant ones and the rest close to the noise variance σ_n^2.

Our goal here is to obtain a useful approximation of the output SINR (22) of the CG-AFM detector for analysis, when the covariance matrix is low rank and the clutter is dominating. Our approach is based on a connection of the CG iterations and the Lanczos tridiagonalization. Specifically, the normalized gradient vectors $\hat{\mathbf{q}}_i$ in $\hat{\mathbf{Q}}_k$ are also called Lanczos vectors which tridiagonalize $\hat{\mathbf{R}}$ [22, 24–26],

$$\hat{\mathbf{Q}}_k^H \hat{\mathbf{R}} \hat{\mathbf{Q}}_k = \hat{\mathbf{T}}_k \qquad (25)$$

where $\hat{\mathbf{T}}_k$ is a tridiagonal matrix. Consider the eigenvalue decomposition of $\hat{\mathbf{T}}_k \in \mathbb{C}^{k \times k}$

$$\hat{\mathbf{T}}_k = \hat{\mathbf{U}} \hat{\mathbf{\Lambda}} \hat{\mathbf{U}}^H \qquad (26)$$

where the diagonal matrix $\hat{\mathbf{\Lambda}}$ consists of the k eigenvalues of $\hat{\mathbf{T}}_k$ in descending order: $\hat{\theta}_1 \geq \hat{\theta}_2 \geq \cdots \geq \hat{\theta}_k$, and the unitary matrix $\hat{\mathbf{U}} = [\hat{\mathbf{u}}_1 \ \hat{\mathbf{u}}_2 \cdots \hat{\mathbf{u}}_k]$ contains the corresponding eigenvectors. Based on (26) and the Lanczos decomposition (25), we can diagonalize $\hat{\mathbf{R}}$ as

$$\hat{\mathbf{P}}^H \hat{\mathbf{R}} \hat{\mathbf{P}} = \hat{\mathbf{\Lambda}} \qquad (27)$$

in which the column orthogonal matrix $\hat{\mathbf{P}} = \hat{\mathbf{Q}}_k \hat{\mathbf{U}} = [\hat{\mathbf{p}}_1 \ \hat{\mathbf{p}}_2 \cdots \hat{\mathbf{p}}_k]$ contains the Ritz vectors of $\hat{\mathbf{R}}$, and $\hat{\theta}_i$ are the Ritz values of $\hat{\mathbf{R}}$ or the "Lanczos estimates" of the eigenvalues of $\hat{\mathbf{R}}$. The Ritz values are known to converge rapidly (with respect to k) to the extremal eigenvalues, i.e., eigenvalues at the edges of the spectrum $\lambda(\hat{\mathbf{R}})$ [22, 27–29]. For the considered case, the spectrum $\lambda(\hat{\mathbf{R}})$ has two clustered groups. One contains the $M - r$ noise eigenvalues, which are uniformly spread around the noise variance σ_n^2, and the one contains the r clutter eigenvalues, which are significantly larger and may be considered as "outliers" compared with the noise eigenvalues. For such a case, the Ritz values are known to converge to the outlier eigenvalues first [30]. In particular, under the condition $k \leq r + 1$, where k denotes the number of CG iterations, $\hat{\theta}_k$ converges to the smallest (i.e., extremal) eigenvalue $\hat{\lambda}_M$, $\hat{\theta}_1$ converges to the largest (i.e., extremal) $\hat{\lambda}_1$, whereas each of the $k - 2$ Ritz values $\{\hat{\theta}_2, \cdots, \hat{\theta}_{k-1}\}$ converges to one of the outlier eigenvalues in $\{\hat{\lambda}_2, \cdots, \hat{\lambda}_r\}$. The rapid convergence of the Ritz values implies that with a few iterations, we have

$\hat{\theta}_1 \geq \hat{\theta}_2 \geq \cdots \geq \hat{\theta}_{k-1} \gg \hat{\theta}_k$. Therefore, we have

$$(\hat{Q}_k^H \hat{\mathbf{R}} \hat{Q}_k)^{-1} = \hat{\mathbf{T}}_k^{-1} = \hat{\mathbf{U}} \hat{\mathbf{\Lambda}}^{-1} \hat{\mathbf{U}}^H$$
$$= \sum_{i=1}^{k} \hat{\mathbf{u}}_i \hat{\theta}_i^{-1} \hat{\mathbf{u}}_i^H \approx \hat{\mathbf{u}}_k \hat{\theta}_k^{-1} \hat{\mathbf{u}}_k^H. \qquad (28)$$

It follows that the output SINR (22) of the CG-AMF detector can be simplified as

$$\rho_k \approx \frac{|\alpha|^2 \left| \mathbf{s}^H \hat{\mathbf{Q}}_k \hat{\mathbf{u}}_k \hat{\theta}_k^{-1} \hat{\mathbf{u}}_k^H \hat{\mathbf{Q}}_k^H \mathbf{s} \right|^2}{\mathbf{s}^H \hat{\mathbf{Q}}_k \hat{\mathbf{u}}_k \hat{\theta}_k^{-1} \hat{\mathbf{u}}_k^H \hat{\mathbf{Q}}_k^H \mathbf{R} \hat{\mathbf{Q}}_k \hat{\mathbf{u}}_k \hat{\theta}_k^{-1} \hat{\mathbf{u}}_k^H \hat{\mathbf{Q}}_k^H \mathbf{s}}$$
$$= \frac{|\alpha|^2 \hat{\mathbf{p}}_k^H \mathbf{s} \mathbf{s}^H \hat{\mathbf{p}}_k}{\hat{\mathbf{p}}_k^H \mathbf{R} \hat{\mathbf{p}}_k} \qquad (29)$$

where $\hat{\mathbf{p}}_k = \hat{\mathbf{Q}}_k \hat{\mathbf{u}}_k$ is the Ritz vector corresponding to the minimum Ritz value of $\hat{\mathbf{R}}$.

C. Statistical Analysis

The output SINR ρ_k is a random variable. To understand its statistical behavior, we derive its pdf by using (29). Our approach is similar to standard subspace perturbation analysis, e.g., as in [31], but we have to cater to the particular structure of (29) which involves CG iterations and Lanczos transform. To begin with, we consider the following eigendecompositions:

$$\mathbf{T}_k = \mathbf{Q}_k^H \mathbf{R} \mathbf{Q}_k = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H \qquad (30)$$

$$\hat{\mathbf{T}}_k = \hat{\mathbf{Q}}_k^H \hat{\mathbf{R}} \hat{\mathbf{Q}}_k = \hat{\mathbf{U}} \hat{\mathbf{\Lambda}} \hat{\mathbf{U}}^H \qquad (31)$$

and

$$\tilde{\mathbf{T}}_k = \mathbf{Q}_k^H \hat{\mathbf{R}} \mathbf{Q}_k = \tilde{\mathbf{U}} \tilde{\mathbf{\Lambda}} \tilde{\mathbf{U}}^H \qquad (32)$$

where $\mathbf{Q}_k$ is the Lanczos transform matrix obtained from the true covariance matrix $\mathbf{R}$, the diagonal matrices $\mathbf{\Lambda} = \text{diag}[\theta_1 \ \theta_2 \cdots \theta_k]$, $\hat{\mathbf{\Lambda}} = \text{diag}[\hat{\theta}_1 \ \hat{\theta}_2 \cdots \hat{\theta}_k]$ and $\tilde{\mathbf{\Lambda}} = \text{diag}[\tilde{\theta}_1 \ \tilde{\theta}_2 \cdots \tilde{\theta}_k]$ contain the k eigenvalues of $\mathbf{T}_k$, $\hat{\mathbf{T}}_k$, and $\tilde{\mathbf{T}}_k$ in descending order, and finally the unitary matrices $\mathbf{U}$, $\hat{\mathbf{U}}$, and $\tilde{\mathbf{U}}$ contain the corresponding eigenvectors of $\mathbf{T}_k$, $\hat{\mathbf{T}}_k$, and $\tilde{\mathbf{T}}_k$, respectively.

Note that both the numerator and denominator of (29) involve a quadratic form of the Ritz vector $\hat{\mathbf{p}}_k$, which consists of a data-dependent Lanczos transform matrix $\hat{\mathbf{Q}}_k$ and a data-dependent eigenvector $\hat{\mathbf{u}}_k$ and, therefore, is rather complicated. To simplify the problem, we consider the approximation $\hat{\mathbf{p}}_k \approx \mathbf{Q}_k \tilde{\mathbf{u}}_k$ with $\tilde{\mathbf{u}}_k$ denoting the eigenvector corresponding to the minimum eigenvalue of $\tilde{\mathbf{T}}_k$. This leads to

$$\rho_k \approx \frac{|\alpha|^2 \tilde{\mathbf{u}}_k^H \mathbf{Q}_k^H \mathbf{s} \mathbf{s}^H \mathbf{Q}_k \tilde{\mathbf{u}}_k}{\tilde{\mathbf{u}}_k^H \mathbf{Q}_k^H \mathbf{R} \mathbf{Q}_k \tilde{\mathbf{u}}_k} = \frac{|\alpha|^2 \tilde{\mathbf{u}}_k^H \mathbf{s}_1 \mathbf{s}_1^H \tilde{\mathbf{u}}_k}{\tilde{\mathbf{u}}_k^H \mathbf{T}_k \tilde{\mathbf{u}}_k} \qquad (33)$$

where $\mathbf{s}_1 = \mathbf{Q}_k^H \mathbf{s}$ is the Lanczos transformed steering vector.

Next, we represent $\tilde{\mathbf{u}}_k$ in terms of the subspace of $\mathbf{T}_k$. Define the $k \times k$ matrix

$$\bar{\mathbf{\Lambda}} \triangleq \mathbf{U}^H \tilde{\mathbf{T}}_k \mathbf{U} \qquad (34)$$

which is in general complex-valued and nondiagonal. It can be considered a perturbed version of the diagonal $\mathbf{\Lambda}$. Let us define the perturbation matrix

$$\mathbf{H} \triangleq \sqrt{L}(\bar{\mathbf{\Lambda}} - \mathbf{\Lambda}). \tag{35}$$

From (32) and (34), we have

$$\bar{\mathbf{\Lambda}} = \mathbf{U}^H \widetilde{\mathbf{U}} \bar{\mathbf{\Lambda}} \widetilde{\mathbf{U}}^H \mathbf{U} = \mathbf{Y} \tilde{\mathbf{\Lambda}} \mathbf{Y}^H \tag{36}$$

where the unitary matrix $\mathbf{Y} \triangleq \mathbf{U}^H \widetilde{\mathbf{U}}$ is the product of two unitary matrices: $\mathbf{U}^H$ and $\widetilde{\mathbf{U}}$. Let $\mathbf{Z} \triangleq \sqrt{L}\,(\mathbf{Y} - \mathbf{I}_k)$ and then

$$\mathbf{Y} = \mathbf{I}_k + \frac{1}{\sqrt{L}}\mathbf{Z}. \tag{37}$$

Substituting (37) back into (36), we have

$$\begin{aligned}
\bar{\mathbf{\Lambda}} &= \left(\mathbf{I}_k + \frac{1}{\sqrt{L}}\mathbf{Z}\right) \tilde{\mathbf{\Lambda}} \left(\mathbf{I}_k + \frac{1}{\sqrt{L}}\mathbf{Z}\right)^H \\
&= \left(\mathbf{I}_k + \frac{1}{\sqrt{L}}\mathbf{Z}\right)\left(\mathbf{\Lambda} + \frac{1}{\sqrt{L}}\mathbf{D}\right)\left(\mathbf{I}_k + \frac{1}{\sqrt{L}}\mathbf{Z}\right)^H \\
&= \mathbf{\Lambda} + \frac{1}{\sqrt{L}}\left(\mathbf{Z}\mathbf{\Lambda} + \mathbf{D} + \mathbf{\Lambda}\mathbf{Z}^H\right) + \mathbf{M}
\end{aligned} \tag{38}$$

where $\mathbf{D} \triangleq \sqrt{L}\left(\tilde{\mathbf{\Lambda}} - \mathbf{\Lambda}\right)$ and $\mathbf{M} \triangleq \frac{1}{L}\left(\mathbf{Z}\mathbf{D} + \mathbf{D}\mathbf{Z}^H + \mathbf{Z}\mathbf{\Lambda}\mathbf{Z}^H\right) + \frac{1}{\sqrt{L^3}}\mathbf{Z}\mathbf{D}\mathbf{Z}^H$. Since $\mathbf{M} \sim O(1/L)$ which contains quantities that are in the order of $1/L$ and higher order terms, it can be neglected for large L. Thus we have

$$\bar{\mathbf{\Lambda}} \approx \mathbf{\Lambda} + \frac{1}{\sqrt{L}}\left(\mathbf{Z}\mathbf{\Lambda} + \mathbf{D} + \mathbf{\Lambda}\mathbf{Z}^H\right). \tag{39}$$

It follows from (39) and (35) that

$$\mathbf{H} \approx \mathbf{Z}\mathbf{\Lambda} + \mathbf{D} + \mathbf{\Lambda}\mathbf{Z}^H. \tag{40}$$

Henceforth, we assume $k > 1$. The case of $k = 1$ is of little interest, since the detector reduces to a nonadaptive detector whose weight vector is $\mathbf{s}$. For $k > 1$, we partition the $k \times k$ matrix of eigenvalues $\mathbf{\Lambda}$ as

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \theta_k \end{bmatrix} \tag{41}$$

where θ_k is the minimum eigenvalue. Similarly, $\mathbf{Z}$ can be partitioned as

$$\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_{1,1} & \mathbf{z}_{1,2} \\ \mathbf{z}_{2,1} & z_{2,2} \end{bmatrix} \tag{42}$$

where $\mathbf{Z}_{1,1} \in \mathbb{C}^{(k-1)\times(k-1)}$, $\mathbf{z}_{1,2} \in \mathbb{C}^{(k-1)\times 1}$, $\mathbf{z}_{2,1} \in \mathbb{C}^{1\times(k-1)}$, and $z_{2,2}$ is a scalar. Let the matrix $\mathbf{H}$ be partitioned in a similar way. It is easy to see from (40)–(42) and the fact that $\mathbf{D} = \sqrt{L}\left(\tilde{\mathbf{\Lambda}} - \mathbf{\Lambda}\right)$ is a diagonal matrix that

$$\mathbf{h}_{1,2} = \theta_k \mathbf{z}_{1,2} + \mathbf{\Lambda}_1 \mathbf{z}_{2,1}^H. \tag{43}$$

The above expression can be further simplified by using the following result.

LEMMA 1 *Within the approximation of O(1/L),*

$$\mathbf{z}_{1,2} \approx -\mathbf{z}_{2,1}^H. \tag{44}$$

PROOF See Appendix A.

Applying the relationship (44) to (43), we obtain that

$$\mathbf{z}_{1,2} = (\theta_k \mathbf{I}_{k-1} - \mathbf{\Lambda}_1)^{-1}\mathbf{h}_{1,2}. \tag{45}$$

The next result establishes the asymptotic distribution of $\mathbf{z}_{1,2}$.

LEMMA 2 *The limiting distribution of the $(k-1) \times 1$ vector $\mathbf{z}_{1,2} = [z_{1,k}\ z_{2,k}\ \cdots\ z_{k-1,k}]^T$ is normal with zero mean and*

$$E[z_{m,k}z_{n,k}^*] = \begin{cases} \dfrac{\theta_k \theta_m}{(\theta_k - \theta_m)^2}, & m = n \\ 0, & m \neq n \end{cases} \tag{46}$$

where $m, n = 1, 2, \cdots, k-1$.

PROOF See Appendix B.

Partition the eigenvector matrix of $\mathbf{T}_k$ as $\mathbf{U} = [\mathbf{U}_1\ \mathbf{u}_k]$, where $\mathbf{u}_k \in \mathbb{C}^{k\times 1}$ is the eigenvector corresponding to the minimum eigenvalue θ_k of $\mathbf{T}_k$, and $\mathbf{U}_1 \in \mathbb{C}^{k\times(k-1)}$ contains the other eigenvectors. Note that $\mathbf{U}_1$ and $\mathbf{u}_k$ are orthogonal, i.e., $\mathbf{U}_1^H \mathbf{u}_k = 0$. Similarly, let $\widetilde{\mathbf{U}} = [\widetilde{\mathbf{U}}_1\ \tilde{\mathbf{u}}_k]$. Recall $\widetilde{\mathbf{U}} = \mathbf{U}\mathbf{Y}$ and (37). It follows that

$$\begin{aligned}
\tilde{\mathbf{u}}_k &= \mathbf{U}_1 \mathbf{y}_{1,2} + \mathbf{u}_k y_{2,2} \\
&= \frac{1}{\sqrt{L}}\mathbf{U}_1 \mathbf{z}_{1,2} + \mathbf{u}_k\left(1 + \frac{z_{2,2}}{\sqrt{L}}\right)
\end{aligned} \tag{47}$$

where $\mathbf{y}_{1,2}$ and $y_{2,2}$ are the partitioning components of $\mathbf{Y}$.

We now consider the numerator of the output SINR (33). Expanding $\tilde{\mathbf{u}}_k$ using (47), we can write the numerator as

$$\begin{aligned}
|\alpha|^2 \tilde{\mathbf{u}}_k^H \mathbf{s}_1 \mathbf{s}_1^H \tilde{\mathbf{u}}_k = |\alpha|^2 \mathbf{s}_1^H \Big(&\frac{2}{\sqrt{L}}\Re\left\{\mathbf{U}_1 \mathbf{z}_{1,2}\mathbf{u}_k^H\right\} \\
&+ \mathbf{u}_k \mathbf{u}_k^H\Big)\mathbf{s}_1 + O(1/L).
\end{aligned} \tag{48}$$

Suppose that the interference is mainly located in the sidelobe directions (i.e., no major mainlobe interference). Then, the projection of the transformed steering vector $\mathbf{s}_1$ to the interference subspace spanned by $\mathbf{U}_1$ is small compared with the projection to the noise subspace spanned by $\mathbf{u}_k$. That is, $|\mathbf{u}_k^H \mathbf{s}_1|^2 \gg \mathbf{s}_1^H \mathbf{U}_1 \mathbf{z}_{1,2}\mathbf{u}_k^H \mathbf{s}_1$. This leads to the following approximation for the numerator of ρ_k:

$$|\alpha|^2 \tilde{\mathbf{u}}_k^H \mathbf{s}_1 \mathbf{s}_1^H \tilde{\mathbf{u}}_k \approx |\alpha|^2 \mathbf{u}_k^H \mathbf{s}_1 \mathbf{s}_1^H \mathbf{u}_k. \tag{49}$$

Similarly, the denominator of ρ_k can be simplified by using (47):

$$\begin{aligned}
\tilde{\mathbf{u}}_k^H \mathbf{T}_k \tilde{\mathbf{u}}_k &= \left[\frac{1}{\sqrt{L}}\mathbf{U}_1 \mathbf{z}_{1,2} + \mathbf{u}_k\left(1 + \frac{z_{2,2}}{\sqrt{L}}\right)\right]^H \mathbf{T}_k \\
&\quad \left[\frac{1}{\sqrt{L}}\mathbf{U}_1 \mathbf{z}_{1,2} + \mathbf{u}_k\left(1 + \frac{z_{2,2}}{\sqrt{L}}\right)\right] \\
&= \frac{1}{L}z_{1,2}^H \mathbf{\Lambda}_1 \mathbf{z}_{1,2} + \theta_k\left(1 + \frac{z_{2,2}}{\sqrt{L}}\right)^*\left(1 + \frac{z_{2,2}}{\sqrt{L}}\right)
\end{aligned} \tag{50}$$

where the second equality is obtained by using the eigenpair relations: $\mathbf{U}_1^H \mathbf{T}_k \mathbf{U}_1 = \mathbf{\Lambda}_1$, $\mathbf{u}_k^H \mathbf{T}_k \mathbf{u}_k = \theta_k$ and the orthogonality relation: $\mathbf{U}_1^H \mathbf{u}_k = 0$. Since $(1 + z_{2,2}/\sqrt{L})^*(1 + z_{2,2}/\sqrt{L}) \approx 1$ [see (64) in Appendix A], the denominator reduces to

$$\tilde{\mathbf{u}}_k^H \mathbf{T}_k \tilde{\mathbf{u}}_k \approx \frac{1}{L} \mathbf{z}_{1,2}^H \mathbf{\Lambda}_1 \mathbf{z}_{1,2} + \theta_k. \qquad (51)$$

Finally, the output SINR after applying (49) and (51) is given by

$$\rho_k \approx \frac{L |\alpha|^2 \mathbf{u}_k^H \mathbf{s}_1 \mathbf{s}_1^H \mathbf{u}_k}{\mathbf{z}_{1,2}^H \mathbf{\Lambda}_1 \mathbf{z}_{1,2} + L\theta_k}. \qquad (52)$$

Define

$$\mu \stackrel{\Delta}{=} \mathbf{z}_{1,2}^H \mathbf{\Lambda}_1 \mathbf{z}_{1,2} = \sum_{m=1}^{k-1} \theta_m z_{m,k}^* z_{m,k}. \qquad (53)$$

From lemma 2, $\{z_{m,k}\}$ are independent zero-mean Gaussian random variables. Therefore, μ is a weighted complex central Chi-square random variable. The output SINR ρ_k is a one-to-one function of μ with a distribution summarized in the following result.

THEOREM 1 Under the condition that the covariance matrix $\mathbf{R}$ has a rank-r component as specified in (23), the asymptotic pdf of the output SINR ρ_k of the CG-AMF detector for $1 < k \leq r + 1$ is given by

$$f(\rho_k) = \frac{\kappa \left(\frac{\kappa}{\rho_k \theta_k} - L \right)^{k-2} \exp\left(L - \frac{\kappa}{\rho_k \theta_k} \right)}{\rho_k^2 \theta_k \Gamma(k-1)},$$
$$0 < \rho_k \leq \frac{\kappa}{\theta_k L} \qquad (54)$$

where $\kappa = L |\alpha|^2 |u_{1,k}|^2 \mathbf{s}^2$ with $u_{1,k}$ denoting the first element of the eigenvector $\boldsymbol{u}_k$.

PROOF   See Appendix C.

Following theorem 1, it is shown in Appendix D that the statistical mean of the output SINR ρ_k is given by

$$E[\rho_k] = \begin{cases} \frac{\kappa e^L E_1(L)}{\theta_k}, & k = 2 \\ \sum_{j_1=0}^{k_1-1} \sum_{j_2=0}^{k_1-j_1-1} \frac{\kappa(-1)^{j_1} L^{j_1+j_2}}{\theta_k(k_1-j_1)j_1! j_2!} & \\ \quad + \frac{\kappa e^L (-L)^{k_1} E_1(L)}{\theta_k k_1!}, & k > 2. \end{cases} \qquad (55)$$

where $k_1 = k - 2$ and the exponential integral $E_1(L)$ is defined as

$$E_1(L) \stackrel{\Delta}{=} \int_L^\infty v^{-1} e^{-v} dv. \qquad (56)$$

Some comments on the calculation of $E_1(L)$ are in order. The series representation of the exponential integral is given by [32]

$$E_1(L) = -a - \ln L + \sum_{i_1=0}^\infty \frac{(-1)^{i_1+1} L^{i_1}}{i_1 i_1!} \qquad (57)$$

where a is the Euler constant. A truncated version of this formula can be used to approximate $E_1(L)$ with good accuracy for small L. For large L, a better approximation is [33]

$$E_1(L) \approx \frac{e^{-L}}{L} \sum_{i_2=0}^{I_2-1} \frac{i_2!}{(-L)^{i_2}} \qquad (58)$$

where I_2 is an integer that is used to control the precision of $E_1(L)$. It has an error of order $O(I_2! L^{-I_2})$ and is only valid for large values of L. Our simulation results in Section V use $I_2 = 20$ for training size $L > 64$, which yields good accuracy.

It would be useful to use (55) to find an estimate of the optimum iteration number k that attains the peak output SINR. Note that (55) cannot be computed in practice since it depends on θ_k, namely the smallest eigenvlaue of the transformed true covariance matrix $\mathbf{T}_k = \mathbf{Q}_k^H \mathbf{R} \mathbf{Q}_k$ as defined in (30), which is generally unknown in adaptive detection. A reasonable approximation is to replace θ_k by $\hat{\theta}_k$ which is computed from the sample covariance matrix, i.e., $\hat{\theta}_k$ is the smallest eigenvalue of $\hat{\mathbf{T}}_k = \hat{\mathbf{Q}}_k^H \hat{\mathbf{R}} \hat{\mathbf{Q}}_k$. With this substitution, we can use (55) to find out approximately how the output SINR ρ_k varies with the number of CG iterations k, and identify a suitable k.

V. NUMERICAL RESULTS

Computer simulation is employed to verify the analytical results presented in the previous section. In addition, we compare the proposed CG-AMF detector and several conventional reduced-rank STAP detectors including the EIG [9], CSM [10], and DCT based detectors. The simulation uses a normalized spatial frequency $f_s = 0.3$ and a normalized Doppler frequency $f_d = 0.3$ for the target. The performance is assessed at different interference/noise levels specified by the reference SINR defined by

$$\text{SINR} = |\alpha|^2 \mathbf{s}^H \mathbf{R}^{-1} \mathbf{s} \qquad (59)$$

which coincides with the output SINR obtained by the clairvoyant MF, which requires knowledge of $\mathbf{R}$. For testing we use the KASSPER data set [34], 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 $N = 16$.

For the CG-AMF detector, we verify our analysis in Section IV with computer simulation. Note that our analysis is based on the assumption that the covariance matrix $\mathbf{R}$ has a rank-r component as in (23). In many real-world scenarios including the KASSPER data, although $\mathbf{R}$ usually does not have the exact low-rank structure, (23) is a useful approximation of $\mathbf{R}$, in which the rank-r component $\mathbf{R}_i$ contains the effect of dominant interference or clutter sources that have to be effectively
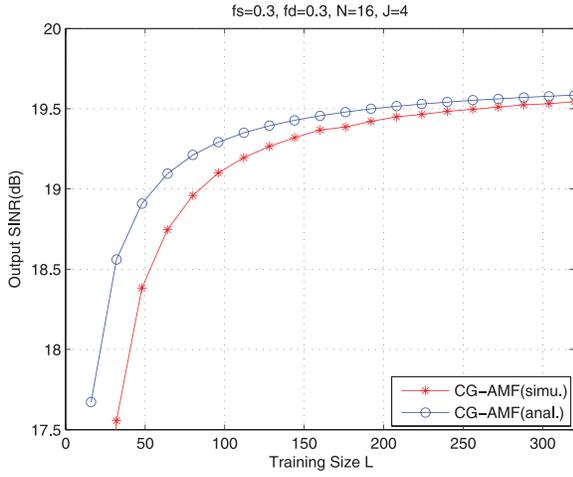
Fig. 1. Output SINR of CG-AMF with rank k = 11 versus training size L.
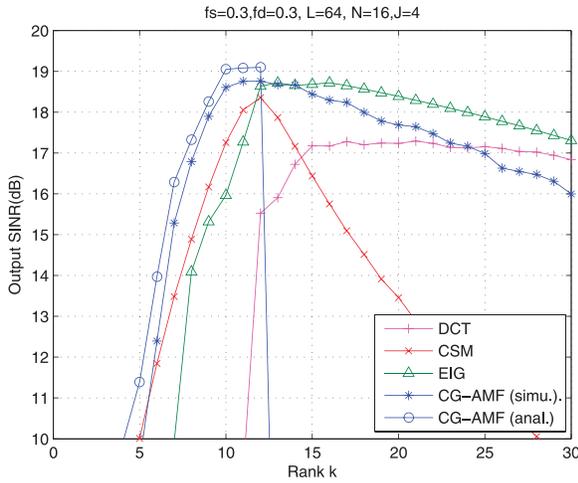
Fig. 2. Output SINR versus detection rank k with $L = 64$ training signals.

Fig. 3. Probability of detection for clairvoyant MF, fully adaptive AMF, and reduced-rank CG-AMF detectors with $L = 64$ and $P_f = 10^{-2}$.

mitigated for detection. We assess the performance of the CG-AMF detector and verify our analysis in such cases.

We first examine the accuracy of our analysis. Fig. 1 shows the mean of the output SINR versus the training size for the CG-AMF detector with rank $k = 11$, and reference SINR = 20 dB. We compare the asymptotic mean output SINR (55) with the output SINR computed by Monte Carlo simulation. When $L > 64$, we use (55) along with (58) to calculate the mean output SINR, with I_2 being set as 20. The variable-precision arithmetic (VPA) function in MATLAB is employed to ensure the calculation accuracy in this case. It can be seen that as the training size increased, the theoretical results converged rapidly to the numerical results. For instance, when the training size is 128, the gap between asymptotic analysis and simulation is less than 0.15 dB.

The size of the training data used in Figs. 2 to 4 is $L = JN = 64$. Fig. 2 depicts the output SINR of the various reduced-rank detectors versus the rank used in detection when the reference SINR = 20 dB. The theoretical results

are computed based on (55). The CG iteration yields an estimate of the rank of the $\mathbf{R}_i$ component in (23) to be $r = 11$. Note that our analysis is valid up to $k = r + 1$, which is why the analytical result is shown only up to $k = 12$. Meanwhile, It is seen that our analysis is able to predict the optimal rank of the CG-AMF detector, which yields the highest output SINR. The gap between the analysis and the simulation is due to the relative small L, which does not meet the asymptotic condition assumed in our analysis. Meanwhile, it should be noted that the DCT-based reduced-rank detector exhibits significant performance degradation compared with the other three data-dependent reduced-rank detectors at low rank. As for the data-dependent methods, they all have a similar maximum SINR, and the CG-AMF detector reaches its maximum output SINR at $k = 11$. Since both the EIG and CSM require a full eigendecomposition which has a complexity of $O(J^3N^3)$, while the complexity of each CG iteration is just $O(J^2N^2)$, the CG-AMF scheme is computationally more efficient than the EIG and CSM.

The probability of detection of the MF, AMF, and CG-AMF detectors with 7, 11, 24, and 64 CG iterations, respectively, are shown in Fig. 3 as a function of the reference SINR. The probability of false alarm is set as $P_f = 0.01$. It can be seen that the optimal CG-AMF detector is obtained with 11 CG iterations as predicted in Fig. 2. Additional iterations should not be pursued since it leads to deteriorated detection performance and higher complexity. Normally, with full iterations $k = 64$, the CG-AMF should converge to the AMF detector; there is a gap between the two, since with $L = JN = 64$ as considered here, the sample covariance matrix $\hat{\mathbf{R}}$ is poorly conditioned, which causes some numerical errors.

Fig. 4 presents the performance comparisons among the CG, EIG, and CSM based reduced-rank detectors with different ranks. It shows that by setting the optimal rank for each detector, namely, $k = 11$ for CG-AMF, $k = 13$ for EIG, and $k = 12$ for CSM, respectively, these detectors yield similar detection performance. When we reduce the
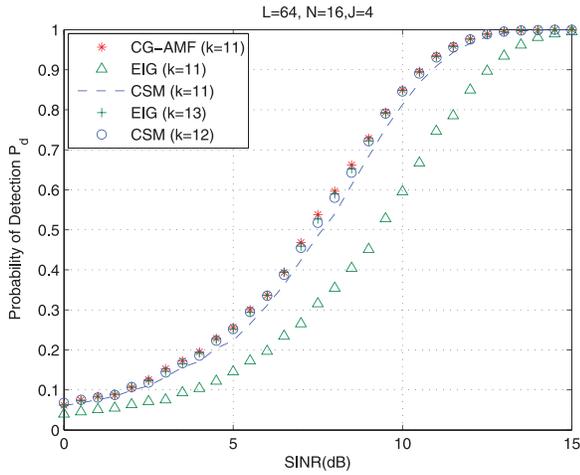
Fig. 4. Probability of detection for reduced-rank EIG, CSM, and CG-AMF detectors with $L = 64$ and $P_f = 10^{-2}$.

Fig. 5. Output SINR versus detection rank k with $L = 128$ training signals.

Fig. 6. Probability of detection for reduced-rank EIG, CSM, and CG-AMF detectors with $L = 128$ and $P_f = 10^{-2}$.

Fig. 7. Angle-Doppler spectrum of KASSPER data at range bin 503.

rank for EIG and CSM to $k = 11$, their performance degrades significantly.

In Fig. 5, the training size is increased to $L = 2\,NJ = 128$, which leads to improvements on the output SINR of all four reduced-rank methods, and the difference between the analysis and the simulation result of the CG-AMF detector decreases. Interestingly, the optimal rank of the computationally efficient CG-AMF detector, which yields the maximum output SINR, is smaller than that of the CSM and EIG based detectors.

We also increase the size of training data to $L = 2\,JN$ for Fig. 6. It shows that the CG-AMF detector with the optimal rank $k = 12$ offers the best detection performance. Note that the performance of the CG-AMF detector with full rank $k = 64$ is identical to that of the AMF detector.

Finally, we examine the performance of the CG-AMF detector with varying normalized target Doppler frequency f_d. Fig. 7 shows the angle-Doppler power spectrum of the KASSPER data at range bin 503, which shows a clutter hot spot around 0.1 Hz of the Doppler

frequency. Detection performance degradation is expected when the target is within the hot-spot region, although some detectors may have a stronger capability to detect endo-clutter targets than others. To see this, we consider a target located at 190 deg in azimuth angle but with varying f_d. Fig. 8 depicts the SINR loss of various detectors measured relative to the peak output SINR of the optimum MF detector. It is seen that all detectors, including the MF, experience a higher SINR loss as the target is in the hot-spot region. However, the CG-AMF detector with $k = 11$, which is the optimum rank from earlier results, yields overall the best performance and is closest to the MF detector among all adaptive reduced-rank detectors considered.

VI. CONCLUSIONS

We considered an adaptive reduced-rank approach by employing the iterative CG algorithm and the AMF. The resulting reduced-rank CG-AMF detector is a projection of the full-rank AMF to a Krylov subspace with a dimension determined by the number of CG iterations. Asymptotic analysis of the output SINR of the CG-AMF
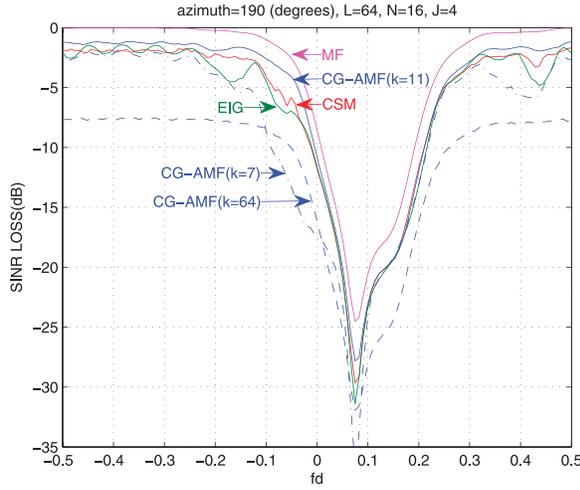
Fig. 8. SINR loss versus normalized Doppler frequency with $L = 64$ training signals.

detector was carried out by exploiting a relation between the CG algorithm and the Lanczos method. Our results show that, in the training limited cases, not only is the CG-AMF detector computationally more efficient (thanks to the efficiency of the CG), but it also often enjoys the benefit of reaching the peak output SINR with a lower rank, when compared with several other popular reduced-rank solutions.

APPENDIX A. PROOF OF LEMMA 1

Partition $\mathbf{Y}$ similarly as $\mathbf{Z}$ in (42):

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_{1,1} & \mathbf{y}_{1,2} \\ \mathbf{y}_{2,1} & y_{2,2} \end{bmatrix}. \tag{60}$$

It follows from (37) that

$$\mathbf{Y} = \begin{bmatrix} \mathbf{I}_{k-1} + \frac{1}{\sqrt{L}}\mathbf{Z}_{1,1} & \frac{1}{\sqrt{L}}\mathbf{z}_{1,2} \\ \frac{1}{\sqrt{L}}\mathbf{z}_{2,1} & 1 + \frac{z_{2,2}}{\sqrt{L}} \end{bmatrix}. \tag{61}$$

Since $\mathbf{Y}$ is unitary, i.e., $\mathbf{Y}^H\mathbf{Y} = \mathbf{I}_k$, we have

$$\begin{aligned} 1 &= \mathbf{y}_{2,1}\mathbf{y}_{2,1}^H + y_{2,2}y_{2,2}^H \\ &= \frac{1}{L}\mathbf{z}_{2,1}\mathbf{z}_{1,2}^H + \left(1 + \frac{z_{2,2}}{\sqrt{L}}\right)\left(1 + \frac{z_{2,2}}{\sqrt{L}}\right)^* \end{aligned} \tag{62}$$

and

$$\begin{aligned} \mathbf{0} &= \mathbf{Y}_{1,1}\mathbf{y}_{2,1}^H + \mathbf{y}_{1,2}y_{2,2}^H \\ &= \frac{1}{\sqrt{L}}\left(\mathbf{z}_{1,2} + \mathbf{z}_{2,1}^H\right) + \frac{1}{L}\left(\mathbf{Z}_{1,1}\mathbf{z}_{2,1}^H + \mathbf{z}_{1,2}z_{2,2}^H\right). \end{aligned} \tag{63}$$

By neglecting the terms in the order of $O(1/L)$, we have

$$\left(1 + \frac{z_{2,2}}{\sqrt{L}}\right)\left(1 + \frac{z_{2,2}}{\sqrt{L}}\right)^* \approx 1 \tag{64}$$

and

$$\frac{1}{\sqrt{L}}\left(\mathbf{z}_{1,2} + \mathbf{z}_{2,1}^H\right) \approx 0. \tag{65}$$

Then, (44) follows immediately from (65).

APPENDIX B. PROOF OF LEMMA 2

Let $\mathbf{h}_{1,2} = [h_{1,k} \ h_{2,k} \ \cdots \ h_{k-1,k}]^T$. From (45), we have

$$z_{m,k} = \frac{h_{m,k}}{\theta_k - \theta_m} \quad m = 1, 2, \cdots, k - 1. \tag{66}$$

Thus, we have

$$E[z_{m,k}z_{n,k}^*] = \frac{E[h_{m,k}h_{n,k}^*]}{(\theta_k - \theta_m)^2}. \tag{67}$$

Using (30), (32), (34), and (35), we can rewrite the eigenvalue perturbation matrix as

$$\begin{aligned} \mathbf{H} &= \sqrt{L}(\bar{\mathbf{\Lambda}} - \mathbf{\Lambda}) = \sqrt{L}(\mathbf{U}^H\widetilde{\mathbf{T}}_k\mathbf{U} - \mathbf{U}^H\mathbf{T}_k\mathbf{U}) \\ &= \left(\mathbf{U}\mathbf{Q}_k\right)^H \mathbf{B}\mathbf{Q}_k\mathbf{U} \triangleq \mathbf{C}^H\mathbf{B}\mathbf{C} \end{aligned} \tag{68}$$

where $\mathbf{B} = \sqrt{L}(\hat{\mathbf{R}} - \mathbf{R})$. It was shown in [35] and [31] that the asymptotic distribution of $\mathbf{B}$ is Gaussian with zero mean and covariance:

$$E[b_{i,j}b_{m,n}^*] = r_{i,m}r_{j,n}^* \tag{69}$$

where b_{ij} and $r_{i,m}$ are the elements of $\mathbf{B}$ and $\mathbf{R}$, respectively.

Since $\mathbf{H}$ is a linear transform of $\mathbf{B}$, it follows that $\mathbf{H}$ is also asymptotically Gaussian with zero mean. Furthermore, it is clear from (66) that $z_{m,k}$ is a Gaussian random variable with zero mean. Let $\mathbf{b} \triangleq \text{vec}(\mathbf{B}) \in \mathbb{C}^{M^2 \times 1}$, where vec(.) denotes the operation of stacking the columns of a matrix on top of one another. An equivalent expression of (69) is given as

$$E[\mathbf{b}\mathbf{b}^H] = \mathbf{R}^* \otimes \mathbf{R}. \tag{70}$$

Applying the matrix Kronecker product result [36]

$$\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}(\mathbf{B}) \tag{71}$$

we have

$$\mathbf{h} \triangleq \text{vec}(\mathbf{H}) = \left(\mathbf{C}^T \otimes \mathbf{C}^H\right)\mathbf{b}. \tag{72}$$

Therefore,

$$\begin{aligned} E[\mathbf{h}\mathbf{h}^H] &= \left(\mathbf{C}^T \otimes \mathbf{C}^H\right) E[\mathbf{b}\mathbf{b}^H]\left(\mathbf{C}^T \otimes \mathbf{C}^H\right)^H \\ &= \left(\mathbf{C}^T \otimes \mathbf{C}^H\right)\left(\mathbf{R}^* \otimes \mathbf{R}\right)\left(\mathbf{C}^T \otimes \mathbf{C}^H\right)^H \\ &= \left(\mathbf{C}^H\mathbf{R}\mathbf{C}\right)^* \otimes \left(\mathbf{C}^H\mathbf{R}\mathbf{C}\right) = \mathbf{\Lambda}^* \otimes \mathbf{\Lambda} \end{aligned} \tag{73}$$

which indicates that $E[h_{i,j}h_{m,n}^*] = \beta_{i,m}\beta_{j,n}^*$, where $\beta_{i,m}$ are the elements of $\mathbf{\Lambda}$. Moreover, as $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues of $\mathbf{T}_k$, we have $\beta_{i,m} = \theta_i\delta_{i,m}$ and $\beta_{j,n} = \theta_j\delta_{j,n}$, where $\delta_{i,n}$ denotes the Kronecker delta. Therefore,

$$E[h_{m,k}h_{n,k}^*] = \beta_{m,n}\beta_{k,k}^* = \theta_k\theta_m\delta_{m,n}. \tag{74}$$

Substituting (74) in (67) leads to (46). This completes the proof.

APPENDIX C.   PROOF OF THEOREM 1

Since $\theta_k \ll \theta_m$, (46) can be approximated as

$$E\left[z_{m,k}z_{m,k}^*\right] \approx \frac{\theta_k}{\theta_m}. \tag{75}$$

Let $v_m = \sqrt{\frac{\theta_m}{\theta_k}}z_{m,k}$, $m = 1, \cdots, k-1$. These new variables are independent, Gaussian random variables with zero mean and unit variance. Then we can rewrite μ as defined in (53) in terms of v_m as

$$\mu = \theta_k \sum_{m=1}^{k-1} v_m v_m^*. \tag{76}$$

Thus, $\zeta = \mu / \theta_k$ is a complex central Chi-squared random variable with $k-1$ DoFs [37], i.e., $\zeta \sim \chi_{k-1}^2$, and

$$f_\zeta(\zeta) = \frac{\zeta^{k-2}e^{-\zeta}}{\Gamma(k-1)}, \quad \zeta \geq 0. \tag{77}$$

The output SINR in (52) can be expressed in terms of ζ by

$$\rho_k = g(\zeta) = \frac{\kappa}{(\zeta + L)\theta_k}, \quad \zeta \geq 0 \tag{78}$$

where $\kappa = L|\alpha|^2 \mathbf{u}_k^H \mathbf{s}_1 \mathbf{s}_1^H \mathbf{u}_k$. Recall that $\mathbf{s}_1 = \mathbf{Q}_k^H \mathbf{s}$ and the first column of $\mathbf{Q}_k$ is a normalized $\mathbf{s}$ with unit norm (i.e., normalized $\hat{\gamma}_1$ in Table I), $\mathbf{s}_1$ is a $k \times 1$ vector with the first element given by $\|\mathbf{s}\|^2$, while all other k-1 elements are zeros. Therefore, κ can be further simplified as $\kappa = L|\alpha|^2 |u_{1,k}|^2 \|\mathbf{s}\|^2$.

Since the function $g(\zeta)$ is monotonic with the inverse function:

$$g^{-1}(\rho_k) = \zeta = \frac{\kappa}{\rho_k \theta_k} - L, \quad 0 < \rho_k \leq \frac{\kappa}{\theta_k L}. \tag{79}$$

The pdf of ρ_k can be calculated in terms of the following equation:

$$f_{\rho_k}(\rho_k) = \left|\frac{dg^{-1}(\rho_k)}{d\rho_k}\right| f_\zeta\left(g^{-1}(\rho_k)\right), \quad 0 < \rho_k \leq \frac{\kappa}{\theta_k L}. \tag{80}$$

It follows that the pdf of the SINR is given by

$$f(\rho_k) = \frac{\kappa\left(\frac{\kappa}{\rho_k \theta_k} - L\right)^{k-2}\exp\left(L - \frac{\kappa}{\rho_k \theta_k}\right)}{\rho_k^2 \theta_k \Gamma(k-1)},$$
$$0 < \rho_k \leq \frac{\kappa}{\theta_k L}. \tag{81}$$

APPENDIX D.   MEAN OF THE OUTPUT SINR

The mean of the output SINR can be computed by using (77) and (78):

$$E[\rho_k] = \int_0^\infty g(\zeta) f_\zeta(\zeta) d\zeta = \frac{\kappa e^L}{\theta_k k_1!}\int_L^\infty \frac{(v-L)^{k_1}e^{-v}}{v}dv \tag{82}$$

where $k_1 = k-2$. For $k = 2$, (82) reduces to

$$E[\rho_2] = \frac{\kappa e^L}{\theta_2}E_1(L) \tag{83}$$

where $E_1(L)$ is the exponential integral defined in (56). For $k > 2$, (82) can be computed as follows:

$$E[\rho_k] = \frac{\kappa e^L}{\theta_k k_1!}\int_L^\infty \sum_{j_1=0}^{k_1}\binom{k_1}{j_1}v^{k_1-1-j_1}(-L)^{j_1}e^{-v}dv$$

$$= \frac{\kappa e^L}{\theta_k k_1!}\sum_{j_1=0}^{k_1-1}(-L)^{j_1}\binom{k_1}{j_1}\Gamma(k_1-j_1, L)$$

$$+ \frac{\kappa e^L}{\theta_k k_1!}(-L)^{k_1}E_1(L) \tag{84}$$

where $\Gamma(k_1-j_1, L)$ denotes the incomplete gamma function and is defined as $\Gamma(k_1-j_1, L) \triangleq \int_L^\infty v^{k_1-j_1-1}e^{-v}dv$. Alternatively, $\Gamma(k_1-j_1, L)$ can be expressed as

$$\Gamma(k_1-j_1, L) = (k_1-j_1-1)!e^{-L}\sum_{j_2=0}^{k_1-j_1-1}\frac{L^{j_2}}{j_2!}. \tag{85}$$

Substituting (85) into (82) yields the closed-form solution in (55).

REFERENCES

[1]   Ward, J.
       Space-time adaptive processing for airborne radar.
       Lincoln Laboratory, MIT, Technical Report 1015, Dec. 1994.

[2]   Klemm, R.
       Principles of Space-Time Adaptive Processing. London: The Institute of Electrical Engineers, 2002.

[3]   Paulraj, A., and Papadias, C. B.
       Space-time processing for wireless communications.
       IEEE Signal Processing Magazine, **14** (Nov. 1997), 49–83.

[4]   Reed, I. S., Mallett, J. D., and Brennan, L. E.
       Rapid convergence rate in adaptive arrays.
       IEEE Transactions on Aerospace and Electronic Systems, **AES-10**, 6 (1974), 853–863.

[5]   Kelly, E. J.
       An adaptive detection algorithm.
       IEEE Transactions on Aerospace and Electronic Systems, **AES-22**, 1 (Mar. 1986), 115–127.

[6]   Chen, W., and Reed, I. S.
       A new CFAR detection test for radar.
       Digital Signal Processing, **1**, 4 (1991), 198–214.

[7]   Robey, F. C., Fuhrmann, D. R., Kelly, E. J., and Nitzberg, R.
       A CFAR adaptive matched filter detector.
       IEEE Transactions on Aerospace and Electronic Systems, **28**, 1 (Jan. 1992), 208–216.

[8]   Kraut, S., and Scharf, L. L.
       The CFAR adaptive subspace detector is a scale-invariant GLRT.
       IEEE Transactions on Signal Processing, **47**, 9 (Sept. 1999), 2538–2541.

[9]   Haimovich, A.
       The eigencanceler: Adaptive radar by eigenanalysis methods.
       IEEE Transactions on Aerospace and Electronic Systems, **32**, 2 (Apr. 1996), 532–542.

[10] Goldstein, J. S., and Reed, I. S.
       Reduced-rank adaptive filtering.
       IEEE Transactions on Signal Processing, **45**, 2 (Feb. 1997), 492–496.

[11] Sohn, K. J., Li, H., and Himed, B.
       Parametric Rao test for multichannel adaptive signal detection.

IEEE Transactions on Aerospace and Electronic Systems, **43**, 3 (July 2007), 920–933.

[12] Sohn, K. J., Li, H., and Himed, B.
Parametric GLRT for multichannel adaptive signal detection.
IEEE Transactions on Signal Processing, **55**, 11 (Nov. 2007), 5351–5360.

[13] Wang, P., Li, H., and Himed, B.
A new parametric GLRT for multichannel adaptive signal detection.
IEEE Transactions on Signal Processing, **58**, 1 (Jan. 2010), 317–325.

[14] Wang, P., Li, H., and Himed, B.
A Bayesian parametric test for multichannel adaptive signal detection in nonhomogeneous environments.
IEEE Signal Processing Letters, **17**, 4 (Apr. 2010), 351–354.

[15] Chang, P., and Wilson, A., Jr.
Analysis of conjugate gradient algorithms for adaptive filtering.
IEEE Transactions on Signal Processing, **48**, 2 (Feb. 2000), 409–418.

[16] Kirsteins, I. R., and Ge, H.
Performance analysis of Krylov space adaptive beamformers.
Proceedings of the Fourth IEEE Workshop on Sensor Array and Multichannel Processing (SAM), Waltham, MA, July 2006, pp. 16–20.

[17] Candan, C., and Erol, Y. B.
Conjugate directions based order recursive implementation of post-Doppler adaptive target detectors.
IET Radar, Sonar Navigation, **6** (2012), 577–586.

[18] Jiang, C., Li, H., and Rangaswamy, M.
Conjugate gradient parametric detection of multichannel signals.
IEEE Transactions on Aerospace and Electronic Systems, **48**, 2 (Apr. 2012), 1521–1536.

[19] Jiang, C., Li, H., and Rangaswamy, M.
On the conjugate gradient matched filter.
IEEE Transactions on Signal Processing, **60**, 5 (May 2012), 2660–2666.

[20] Weippert, M. E., Hiemstra, J. D., Goldstein, J. S., and Zoltowski, M. D.
Insights from the relationship between the multistage Wiener filter and the method of conjugate gradients.
Sensor Array and Multichannel Signal Processing Workshop Proceedings, 2002, Rosslyn, VA, Aug. 2002, pp. 388–392.

[21] Scharf, L. L., Chong, E. K. P., Zoltowski, M. D., Goldstein, J. S., and Reed, I. S.
Subspace expansion and the equivalence of conjugate direction and multistage wiener filters.
IEEE Transactions on Signal Processing, **56**, 10 (Oct. 2008), 5013–5019.

[22] Golub, G. H., and Van Loan, C. F.
Matrix Computations (3rd ed.). Baltimore, MD: Johns Hopkins University Press, 1996.

[23] Brennan, L. E., and Reed, I. S.
Theory of adaptive radar.

IEEE Transactions on Aerospace and Electronic Systems, **AES-9**, 2 (1973), 237–252.

[24] Axelsson, O.
Iterative Solution Methods. New York: Cambridge University Press, 1994.

[25] Parlett, B. N.
The Symmetric Eigenvalue Problem. Englewood Cliffs, NJ: Prentice-Hall, 1980.

[26] Cullum, J., and Willoughby, R. A.
Lanczos Algorithms for Large Symmetric Eigenvalue Computations, Vol. I, Theory. Philadelphia, PA: SIAM, 2002.

[27] Kaniel, S.
Estimates for some computational techniques in linear algebra.
Mathematics of Computation, **20** (1966), 369–378.

[28] Paige, C. C.
The computation of eigenvalues and eigenvectors of very large sparse matrices.
Ph.D thesis, London University, 1971.

[29] Saad, Y.
On the rates of convergence of the Lanczos and the block Lanczos methods.
SIAM Journal on Numerical Analysis, **17** (1980), 687–706.

[30] Trefethen, L. N., and Bau, D., III, *Numerical Linear Algebra.* Philadelphia, PA: Society for Industrial and Applied Mathematics, 1997.

[31] Haimovich, A. M.
Asymptotic distribution of the conditional signal-to-noise ratio in an eigenanalysis-based adaptive array.
IEEE Transactions on Aerospace and Electronic Systems, **33**, 3 (July 1997), 988–997.

[32] Sharma, R., and Bahman, Z.
A general method for an accurate evaluation of exponential integrals $E_1(x)$, $x > 0$.
Journal of Computational Physics, **25** (1977), 199–204.

[33] Milgram, M.
The generalized integro-exponential function.
Mathematics of Computation, **44** (1985), 443–458.

[34] Bergin, J. S., and Techau, P. M.
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.

[35] Gupta, R.
Asymptotic theory for principal component analysis in the complex case.
Journal of the Indian Statistical Association, **3** (1965), 97–106.

[36] Graham, A.
Kronecker Products and Matrix Calculus with Applications. Chichester, England: Ellis Horwood, 1981.

[37] Kelly, E. J., and Forsythe, K.
Adaptive detection and parameter estimation for multidimensional signal models.
Lincoln Laboratory, MIT, Technical Report 848, 1989.

Zhu Chen (S'05-M'12) received the B.S. degree in automation from South-Central University for Nationalities, Wuhan, China, in 2006, the M.S. degree in electrical engineering from Chonbuk National University, Jeonju, South Korea, in 2008, and the Ph.D. degree in electrical engineering from Stevens Institute of Technology, Hoboken, NJ, USA, in 2013.

He was a research intern at the Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA, in summer 2012. Since September 2013, he has been with Broadcom Corporation, Matawan, NJ. His current research interests include statistical signal processing, wireless communication and machine learning.

Dr. Chen received the Francis T. Boesch Award and the Outstanding Dissertation Award in 2014 from Stevens Institute of Technology.

Hongbin Li (M'99–SM'08) received the B.S. and M.S. degrees from the University of Electronic Science and Technology of China, in 1991 and 1994, respectively, and the Ph.D. degree from the University of Florida, Gainesville, in 1999, all in electrical engineering.

From July 1996 to May 1999, he was a Research Assistant in the Department of Electrical and Computer Engineering at the University of Florida. Since July 1999, he has been with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ, where he is a Professor. He was a Summer Visiting Faculty Member at the Air Force Research Laboratory in the summers of 2003, 2004, and 2009. His general research interests include statistical signal processing, wireless communications, and radars.

Dr. Li received the IEEE Jack Neubauer Memorial Award in 2013 for the best systems paper published in the *IEEE Transactions on Vehicular Technology*, the Outstanding Paper Award from the IEEE AFICON Conference in 2011, the Harvey N. Davis Teaching Award in 2003 and the Jess H. Davis Memorial Award for excellence in research in 2001 from Stevens Institute of Technology, and the Sigma Xi Graduate Research Award from the University of Florida in 1999. He has been a member of the IEEE SPS Signal Processing Theory and Methods (2011 to now) Technical Committee (TC) and the IEEE SPS Sensor Array and Multichannel TC (2006–2012). He is an Associate Editor for *Signal Processing* (Elsevier), and served on the editorial boards for *IEEE Transactions on Wireless Communications*, *IEEE Signal Processing Letters*, and *IEEE Transactions on Signal Processing*. He was a Guest Editor for *EURASIP Journal on Applied Signal Processing*. He has been involved in various conference organization activities, including serving as a General Co-Chair for the 7th IEEE Sensor Array and Multichannel Signal Processing (SAM) Workshop, Hoboken, NJ, June 17-20, 2012. Dr. Li is a member of Tau Beta Pi and Phi Kappa Phi.

Muralidhar Rangaswamy (S'89–M'93–SM'98–F'06) received the B.E. degree in electronics engineering from Bangalore University, Bangalore, India in 1985 and the M.S. and Ph.D. degrees in electrical engineering from Syracuse University, Syracuse, NY, in 1992.

He is presently employed as the Senior Advisor for Radar Research at the RF Exploitation Branch within the Sensors Directorate of the Air Force Research Laboratory (AFRL). Prior to this he has held industrial and academic appointments. His research interests include radar signal processing, spectrum estimation, modeling non-Gaussian interference phenomena, and statistical communication theory.

Dr. Rangaswamy has co-authored more than 190 refereed journal and conference record papers in the areas of his research interests. Additionally, he is a contributor to 8 books and is a co-inventor on 3 U.S. patents. He is the Technical Editor (Associate Editor-in-Chief) for Radar Systems in the *IEEE Transactions on Aerospace and Electronic Systems* (IEEE-TAES). He served as the Co-Editor-in-Chief for the *Digital Signal Processing* journal between 2005 and 2011. He serves on the Senior Editorial Board of the *IEEE Journal of Selected Topics in Signal Processing* (Jan 2012-Dec 2014). He was a 2-term elected member of the sensor array and multichannel processing technical committee (SAM-TC) of the IEEE Signal Processing Society between January 2005 and December 2010 and serves as a member of the Radar Systems Panel (RSP) in the IEEE-AES Society. He was the General Chairman for the 4th IEEE Workshop on Sensor Array and Multichannel Processing (SAM-2006), Waltham, MA, July 2006. Dr. Rangaswamy has served on the Technical Committee of the IEEE Radar Conference series in a myriad of roles (Track Chair, Session Chair, Special Session Organizer and Chair, Paper Selection Committee Member, Tutorial Lecturer). He served as the Publicity Chair for the First IEEE International Conference on Waveform Diversity and Design, Edinburgh, U.K., Nov. 2004. He presently serves on the conference subcommittee of the RSP. He is the Technical Program Chairman for the 2014 IEEE Radar Conference.

Dr. Rangaswamy received the IEEE Warren White Radar Award in 2013, the 2013 Affiliate Societies Council Dayton (ASC-D) Outstanding Scientist and Engineer Award, the 2007 IEEE Region 1 Award, the 2006 IEEE Boston Section Distinguished Member Award, and the 2005 IEEE-AESS Fred Nathanson memorial outstanding young radar engineer award. He received the 2012 and 2005 Charles Ryan basic research award from the Sensors Directorate of AFRL, in addition to more than 40 scientific achievement awards.