On Generalized Eigenvector Space For Target Detection in Reduced Dimensions

Gökhan M. Güvensen, Çağatay Candan, Sencer Koç and Umut Orguner

Middle East Technical University (METU)

Department of Electrical and Electronics Engineering

Ankara, Turkey.

{guvensen,ccandan,skoc,umut}@metu.edu.tr

Abstract—The detection and estimation problems with large dimensional vectors frequently appear in the phased array radar systems equipped with, possibly, several hundreds of receiving elements. For such systems, a preprocessing stage reducing the large dimensional input to a manageable dimension is required. The present work shows that the subspace spanned by the generalized eigenvectors of signal and noise covariance matrices is the optimal subspace to this aim from several different viewpoints. Numerical results on the subspace selection for the radar target detection problem is provided to examine the performance of detectors with reduced dimensions.

Index Terms—Sufficient Statistics, Mutual Information, Detection, Reduced Rank Detection, Generalized Eigenvectors.

I. INTRODUCTION

Modern radar systems, such as the phased array systems having several hundreds of receiving elements, utilize large dimensional snapshot data for decision making. The elementary operations on the snapshots such as the calculation of the sample auto-correlation matrix or its inversion (sample matrix inversion) becomes quickly infeasible with the increasing dimensions. A good way to enable the processing of large dimensional signals is the adoption of a preprocessing stage that captures the "essence" of the input at a reduced dimension.

It is well known that the minimal sufficient statistics of the detection and estimation theory reduces the dimension of the input at no loss of information. Unfortunately, it is difficult to find non-trivial sufficient statistics, yet the minimal one, in many problems and the derivation of useful statistics remains as an important challenge for such problems. In this paper, we examine the dimension reduction problem from several different viewpoints. Our goal is to find a good subspace that provides a minimal performance compromise in the subsequent statistical signal processing operations. Our application focus is the radar target detection; but the given approach can be useful in other applications such as the parameter estimation in reduced dimensions.

The subspace based detection and estimation problems are studied within different contexts. In an early work of Scharf, the generalized eigenvector space arises as the optimal subspace for the maximization of J-divergence, [1]. In the present work, we revisit the subspace problem and show that the generalized eigenvector space is also the optimal solution of several other important problems of interest. More recent work of Scharf *et al.* includes the development of reduced rank Wiener filters and their decompositions, [2]; extension of reduced dimension Wiener filtering to other applications such as detection, coding [3], [4]. A related line of work proposes the multistage Wiener filters for reduced dimensional operation, [5]. It has been shown that the multistage Wiener filtering is equivalent to the optimal Wiener filtering in a Krylov subspace [6]. Some other filtering schemes equivalent to multistage Wiener filters are also given in [7]. It should be noted that the approaches equivalent to multistage Wiener filters utilize a specially structured subspace called Krylov subspace which does not have any optimality properties. In contrast to the Krylov subspace, the generalized eigenvector space arises as the optimal of several important problems, as shown in this paper.

II. PROBLEM DESCRIPTION

We consider that the random vector $\mathbf{x}$ of interest is observed under additive, possibly colored, noise conditions, $\mathbf{y} = \mathbf{x} + \mathbf{n}$. Here the vector $\mathbf{n}$ represents the noise. Our goal is to make inference on $\mathbf{x}$ given the observation $\mathbf{y}$. The vectors $\mathbf{x}$ and $\mathbf{y}$ are of high dimensions.

It is known that the random vector $\mathbf{t}(\mathbf{y})$ ($\mathbf{t}(\cdot)$ is a vector valued function of $\mathbf{y}$) is the sufficient statistics with respect to $p_{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y})$ if the condition $p_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y}) = p_{\mathbf{X}|\mathbf{T}}(\mathbf{x}|\mathbf{t}(\mathbf{y}))$ is satisfied. Our goal is to find a "good" statistics for inference.

As a motivational example, we may consider a simple problem having a scalar variable s, which is Gaussian distributed with zero mean and variance σ_s^2; $s \sim N(0, \sigma_s^2)$, as the signal of interest. The random variable s is observed N times under additive white Gaussian noise conditions. It is assumed that the noise is independent and identically distributed with the distribution $n \sim N(0, \sigma_n^2)$. The collection of observations can be written as $\mathbf{y} = \mathbf{1}s + \mathbf{n}$, where $\mathbf{1}$ is $N \times 1$ vector of all ones. The sufficient statistics for this problem can be calculated as $t = \sum_{k=1}^{N} y_k = \mathbf{1}^{\mathbf{T}}\mathbf{y}$. With this statistics, the N dimensional observation vector y is reduced to a single scalar at no loss of information on the unknown s. Our goal is to produce similar results for arbitrary random vectors $\mathbf{x}$ and $\mathbf{n}$ having auto-correlation matrices $\mathbf{R_x}$ and $\mathbf{R_n}$. For this specific problem, we have $\mathbf{R_x} = \mathbf{1}\mathbf{1}^{\mathbf{T}}\sigma_s^2$ and $\mathbf{R_n} = \mathbf{I}\sigma_n^2$. As noted in this paper, the subspace of generalized eigenvector spanned by the dominant eigenvectors of $\mathbf{R_n^{-1}R_x}$ is a good candidate for the dimension reduction. For this specific problem, the only

generalized eigenvector with a non-zero eigenvalue is $\mathbf{1}$ and the projection of the input vector $\mathbf{y}$ to the dominant eigenvector $\mathbf{1}$ results in a statistics coinciding with the sufficient statistics.

Our main goal in this study is to explain the good performance of the generalized eigenvectors of $\mathbf{R_x}$ and $\mathbf{R_n}$, which are the eigenvectors of $\mathbf{R_n^{-1}R_x}$ for an invertible $\mathbf{R_n}$, for the choice of subspace in the reduced dimensional detection problem. As noted earlier, an optimality result for the generalized eigenvector subspace has already been noted by Scharf from the viewpoint of J-divergence in [1]. In this paper, we show that the same subspace arises as the optimal solution of several other important problems.

III. GENERALIZED EIGENVECTOR SPACE

A. Mutual Information Preserving Subspace

Figure 1 shows the data processing chain for the dimension reduction problem examined in this work. We assume that the N dimensional vectors $\mathbf{x}$ and $\mathbf{n}$ are independent, zero-mean Gaussian vectors with auto-correlation matrices $\mathbf{R_x}$ and $\mathbf{R_n}$. With the notation of [7], a D dimensional vector $\widetilde{\mathbf{y}}$ is formed by a linear transformation through $\mathbf{S}_D^{\mathbf{H}}$ matrix. Our goal is to accomplish the dimension reduction with a small loss of mutual information.
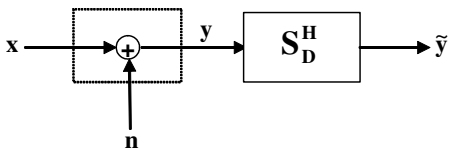

Fig. 1. Data processing chain for dimension reduction problem

The mutual information between observation $\mathbf{y}$ and $\mathbf{x}$ can be written as $I(\mathbf{x}; \mathbf{y})$, [8]. From data processing inequality, it is known that $I(\mathbf{x}; \mathbf{y}) \geq I(\mathbf{x}; \widetilde{\mathbf{y}})$. The equality is only satisfied if $\widetilde{\mathbf{y}}$ is the sufficient statistics with respect to $p_{\mathbf{X}, \mathbf{Y}}(\mathbf{x}, \mathbf{y})$.

Our goal is to find a $D \times N$ $\mathbf{S}_D^{\mathbf{H}}$ matrix so that $I(\mathbf{x}; \widetilde{\mathbf{y}})$ is as close as possible to $I(\mathbf{x}; \mathbf{y})$ under the given dimension reduction constraint. The problem can also be stated as the preservation of the mutual information with a linear transformation under a rank constraint.

The mutual information between $\mathbf{x}$ and $\widetilde{\mathbf{y}}$ can be written by finding the joint density of $\mathbf{x}$ and $\widetilde{\mathbf{y}}$. Since both $\mathbf{x}$ and $\widetilde{\mathbf{y}}$ are Gaussian vectors, the joint density can be easily expressed in terms of covariance and cross-covariance of these two vectors. In the communication literature, the mutual information calculation of Gaussian vectors $\mathbf{u}$ and $\mathbf{v} = \mathbf{Hu} + \mathbf{j}$ ($\mathbf{j}$ is a Gaussian vector independent of $\mathbf{u}$) is known as the capacity of non-fading MIMO channel and examined in depth since the work of Telatar, [9]. The formula for capacity or mutual information between $\mathbf{u}$ and $\mathbf{v}$, can be compactly written as follows, [10]:

$$C = \log\left(\frac{|\mathbf{R_j} + \mathbf{HR_uH^H}|}{|\mathbf{R_j}|}\right) = \log(|\mathbf{I} + \mathbf{R_j^{-1}HR_uH^H}|). \quad (1)$$

We can simply adapt the capacity result given in (1) to the dimension reduction problem by noting that $\widetilde{\mathbf{y}} = \mathbf{S}_D^{\mathbf{H}}\mathbf{x} + \mathbf{S}_D^{\mathbf{H}}\mathbf{n}$.

Here the $\mathbf{j}$ vector in the capacity problem is replaced by transformation dependent $\mathbf{S}_D^{\mathbf{H}}\mathbf{n}$. Substituting $\mathbf{H} \leftarrow \mathbf{S}_D^{\mathbf{H}}$, $\mathbf{R_u} \leftarrow \mathbf{R_x}$, $\mathbf{R_j} \leftarrow \mathbf{S}_D^{\mathbf{H}}\mathbf{R_n}\mathbf{S}_D$ in (1), we immediately get the mutual information between $\mathbf{x}$ and $\widetilde{\mathbf{y}}$

$$\begin{aligned} I(\mathbf{x}; \widetilde{\mathbf{y}}) &= \log(|\mathbf{I} + (\mathbf{S}_D^{\mathbf{H}}\mathbf{R_n}\mathbf{S}_D)^{-1}(\mathbf{S}_D^{\mathbf{H}}\mathbf{R_x}\mathbf{S}_D)|) \\ &= \log(\det(\mathbf{I} + \mathbf{SNR})) \end{aligned} \quad (2)$$

where the SNR matrix appearing in (2) is defined as

$$\mathbf{SNR} = (\mathbf{S}_D^{\mathbf{H}}\mathbf{R_n}\mathbf{S}_D)^{-1}(\mathbf{S}_D^{\mathbf{H}}\mathbf{R_x}\mathbf{S}_D). \quad (3)$$

It can be seen that the maximization of $\det(\mathbf{I} + \mathbf{SNR})$ is equivalent to the maximization of $I(\mathbf{x}; \widetilde{\mathbf{y}})$.

From Appendix A, it can be noted that the maximum value of the reward function $\det(\mathbf{I} + \mathbf{SNR})$ under the constraint that $\mathbf{S}_D$ is a full column rank $N \times D$ matrix is $\prod_{k=1}^{D}(1 + \lambda_k)$. This reward is achieved by the first D dominant generalized eigenvectors of $\mathbf{R_x}$ and $\mathbf{R_n}$.

The generalized eigenvectors of $\mathbf{R_x}$ and $\mathbf{R_n}$ are defined as the solutions of $\mathbf{R_x}\mathbf{e_k} = \lambda_k\mathbf{R_n}\mathbf{e_k}$. If $\mathbf{R_n}$ is invertible, the generalized eigenvectors can also be calculated by computing the eigenvectors of the matrix $\mathbf{R_n^{-1}R_x}$. Upon ordering the generalized eigenvectors in the descending eigenvalues, $\lambda_1 \geq \lambda_2 \geq \ldots > \lambda_N$, we can get the ordered generalized eigenvector set. The result of the present section claims that among full rank $N \times D$ matrices, the matrix $\mathbf{S}_D$ which is defined as

$$\mathbf{S}_D = [\mathbf{e_1}\,\mathbf{e_2}\,\ldots\,\mathbf{e}_D] \quad (4)$$

is the matrix minimizing mutual information loss due to projection onto D dimensional space.

As noted in Appendix A, replacing $\mathbf{S}_D$ with $\mathbf{S}_D\mathbf{T}$ ($\mathbf{T}$ is a $D \times D$ full rank matrix) in $\mathbf{SNR}$ matrix, does not affect the the reward function $\det(\mathbf{I} + \mathbf{SNR})$. This shows that the reward value is determined by the column space of $\mathbf{S}_D$, but not its representation according to a basis. Hence, it is always possible to orthogonalize columns of $\mathbf{S}_D$ matrix with no loss of mutual information.

B. Reconstruction Error Minimizing Subspace

Figure 2 shows the problem of data reconstruction from noisy observations after dimension reduction. A legitimate goal is the minimization of the reconstruction error according to a criterion.
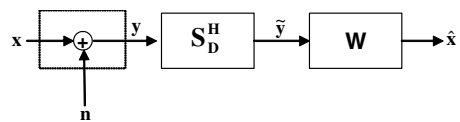

Fig. 2. Processing chain for Gaussian vector reconstruction after dimension reduction

Focusing on Gaussian vectors, the minimum mean square error (MMSE) estimator of $\mathbf{x}$ is a linear combiner that can be explicitly written as follows

$$\hat{\mathbf{x}} = \mathbf{R_x}\mathbf{S}_D\left(\mathbf{S}_D^{\mathbf{H}}(\mathbf{R_x} + \mathbf{R_n})\mathbf{S}_D\right)^{-1}\widetilde{\mathbf{y}}. \quad (5)$$

If we denote the reconstruction error with $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$, the covariance matrix of error can be written as follows

$$\mathbf{R_e} = \mathbf{R_x} - \mathbf{R_x}\mathbf{S}_D \left(\mathbf{S}_D^{\mathbf{H}}(\mathbf{R_x} + \mathbf{R_n})\mathbf{S}_D\right)^{-1}\mathbf{S}_D^{\mathbf{H}}\mathbf{R_x}. \quad (6)$$

From the error covariance matrix, the Fisher information matrix $\mathbf{R_e}^{-1}$ can be expressed via the matrix inversion lemma, [10]:

$$\mathbf{R_e}^{-1} = \mathbf{R_x}^{-1} + \mathbf{S}_D \left(\mathbf{S}_D^{\mathbf{H}}\mathbf{R_n}\mathbf{S}_D\right)^{-1}\mathbf{S}_D^{\mathbf{H}}. \quad (7)$$

By right multiplying both sides of (7) by $\mathbf{R_x}$, we get

$$\mathbf{R_e}^{-1}\mathbf{R_x} = \mathbf{I} + \mathbf{S}_D \left(\mathbf{S}_D^{\mathbf{H}}\mathbf{R_n}\mathbf{S}_D\right)^{-1}\mathbf{S}_D^{\mathbf{H}}\mathbf{R_x}. \quad (8)$$

If we take the determinant of both parts of (8), we get

$$\det(\mathbf{R_e}) = \frac{\det(\mathbf{R_x})}{\det(\mathbf{I} + \mathbf{S}_D \left(\mathbf{S}_D^{\mathbf{H}}\mathbf{R_n}\mathbf{S}_D\right)^{-1}\mathbf{S}_D^{\mathbf{H}}\mathbf{R_x})}. \quad (9)$$

It is clear from (9) that the determinant of $\mathbf{R_e}$ (error volume) is minimized with respect to $\mathbf{S}_D$ if the determinant appearing in the denominator of the ratio on the right hand side of (9) is maximized. Using the identity $\det(\mathbf{I} + \mathbf{AB}) = \det(\mathbf{I} + \mathbf{BA})$, the maximization problem can also be written as

$$\max_{\mathbf{S}_D} \det \left(\mathbf{I} + \underbrace{\left(\mathbf{S}_D^{\mathbf{H}}\mathbf{R_n}\mathbf{S}_D\right)^{-1} \left(\mathbf{S}_D^{\mathbf{H}}\mathbf{R_x}\mathbf{S}_D\right)}_{\mathbf{SNR}} \right). \quad (10)$$

The final problem is identical to the maximization of the mutual information $\det\left(\mathbf{I} + \mathbf{SNR}\right)$ where $\mathbf{SNR}$ matrix is given in (3). The $\mathbf{S}_D$ matrix maximizing $\det\left(\mathbf{I} + \mathbf{SNR}\right)$ is the matrix whose column space is formed by the first D dominant generalized eigenvectors of $\mathbf{R_x}$ and $\mathbf{R_n}$, as noted before.

An alternative criterion for the reconstruction error minimization can be the normalized mean square error (MSE). The normalized MSE (nMSE) matrix can be defined as $\mathbf{nMSE} = \mathbf{R_x}^{-1}\mathbf{R_e}$. The trace of $\mathbf{nMSE}$ can be written from (6) as follows:

$$\begin{aligned}
\text{tr}\{\mathbf{nMSE}\} &= \text{tr}\left\{\mathbf{I}_N - \mathbf{S}_D \left(\mathbf{S}_D^{\mathbf{H}}(\mathbf{R_x}+\mathbf{R_n})\mathbf{S}_D\right)^{-1}\mathbf{S}_D^{\mathbf{H}}\mathbf{R_x}\right\} \\
&= \text{tr}\left\{\mathbf{I}_D - \left(\mathbf{S}_D^{\mathbf{H}}(\mathbf{R_x}+\mathbf{R_n})\mathbf{S}_D\right)^{-1}\mathbf{S}_D^{\mathbf{H}}\mathbf{R_x}\mathbf{S}_D\right\} \\
&\quad + (N-D) \\
&= \text{tr}\left\{\mathbf{I}_D - \left(\mathbf{I}_D + \mathbf{SNR}^{-1}\right)^{-1}\right\} + (N-D) \\
&= \text{tr}\left\{(\mathbf{I}_D + \mathbf{SNR})^{-1}\right\} + (N-D) \quad (11)
\end{aligned}$$

In the second line of (11), the identity $\text{tr}\{\mathbf{AB}\} = \text{tr}\{\mathbf{BA}\}$ is used, and the last line comes from the application of matrix inversion lemma. [1] As noted in Appendix A, the cost function of $\text{tr}\{\mathbf{nMSE}\}$ is minimized with the generalized eigenvector set. We would like to note that the scalar version of the relation (11), that is $\text{MSE} = \sigma_x^2/(1 + \text{SNR})$, is utilized in the analysis of communication systems, [11], [12].

[1] The normalized MSE matrix can also be defined as $\mathbf{R_e}\mathbf{R_x}^{-1}$ (instead of $\mathbf{R_x}^{-1}\mathbf{R_e}$). Since $\text{tr}\left\{\mathbf{R_e}\mathbf{R_x}^{-1}\right\} = \text{tr}\left\{\mathbf{R_x}^{-1}\mathbf{R_e}\right\}$, the relation for the trace of the normalized MSE matrix with the SNR matrix remains the same for both definitions.

C. Detection Probability Maximizing Subspace

Figure 3 shows the processing chain specific for the radar target detection problem. Here $\mathbf{y}$ denotes the observation vector which may or may not contain target signal. The vector $\mathbf{x}$ denotes the target signal and $\mathbf{n}$ denotes the noise and jammer signal contributing to the observation vector $\mathbf{y}$. The goal is to optimize $\mathbf{S}_D$ matrix to maximize the detection probability by maximizing the SNR at the input of the detection unit.
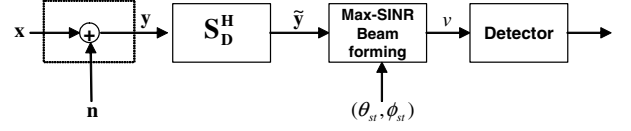

Fig. 3. Processing chain for radar target detection with dimension reduction

We model the observation vector as $\mathbf{y} = \alpha\mathbf{a}(\theta, \phi) + \mathbf{n}$. The vector $\mathbf{a}(\theta, \phi)$ represents the array manifold vector induced by an incoming wave from the elevation and azimuth angles of θ and ϕ, respectively. We consider that there exists some a-priori information about the target angular location. Hence, θ and ϕ variables are considered as random variables with a known joint probability density functions. With this modeling the vector $\mathbf{x} = \alpha\mathbf{a}(\theta, \phi)$ becomes a random vector whose covariance matrix can be calculated.

The processing chain shown in Figure 3 is utilized as follows: First, the observation vector of large dimensions is reduced to a D dimensional vector $\widetilde{\mathbf{y}}$ via the application of $\mathbf{S}_D$ matrix. Then, the signal-to-noise-and-interference ratio (SINR) maximizing beamformer (max-SINR beamformer) is constructed at reduced rank to steer the beam towards $(\theta_{\text{st}}, \phi_{\text{st}})$. Then, the beamformer output is forwarded to the detector for the declaration of target absent-present decision.

The output of max-SINR beamformer steered towards $(\theta_{\text{st}}, \phi_{\text{st}})$ can be written as

$$v(\theta_{\text{st}}, \phi_{\text{st}}) = \mathbf{a}^{\mathbf{H}}(\theta_{\text{st}}, \phi_{\text{st}})\mathbf{S}_D \left(\mathbf{S}_D^{\mathbf{H}}\mathbf{R_n}\mathbf{S}_D\right)^{-1}\mathbf{S}_D^{\mathbf{H}}\mathbf{y}. \quad (12)$$

The maximum SINR value achieved by this beamformer is

$$\begin{aligned}
\text{SINR}_{\max}(\theta_{\text{st}}, \phi_{\text{st}}) &= \\
\mathbf{a}_{(\theta_{\text{st}},\phi_{\text{st}})}^{\mathbf{H}}\mathbf{S}_D \left(\mathbf{S}_D^{\mathbf{H}}\mathbf{R_n}\mathbf{S}_D\right)^{-1}\mathbf{S}_D^{\mathbf{H}}\mathbf{a}_{(\theta_{\text{st}},\phi_{\text{st}})} &= \\
\text{tr}\left\{\left(\mathbf{S}_D^{\mathbf{H}}\mathbf{R_n}\mathbf{S}_D\right)^{-1}\left(|\alpha|^2\mathbf{S}_D^{\mathbf{H}}\mathbf{a}_{(\theta_{\text{st}},\phi_{\text{st}})}\mathbf{a}_{(\theta_{\text{st}},\phi_{\text{st}})}^{\mathbf{H}}\mathbf{S}_D\right)\right\} &\quad (13)
\end{aligned}$$

where in the last line of (13), the identity $\text{tr}\{\mathbf{AB}\} = \text{tr}\{\mathbf{BA}\}$ is used. The SINR value given on the right hand side of (13) is the maximum SNR achieved when the beam is steered towards a point, i.e. angular location of $(\theta_{\text{st}}, \phi_{\text{st}})$, in space. We may refer this SINR value as the point SINR value.

Since the target is not guaranteed to be located at a single point in space, we may suggest to use the average of point SINR values over which target can be located. This can be achieved by calculating the expected value of the point SINR

given in (13) over target angular location density as follows

$$
\begin{aligned}
\overline{\mathrm{SINR}} &= E_{(\theta,\phi)}\{\mathrm{SINR}_{\max}(\theta,\phi)\} \\
&= \mathrm{tr}\left\{\left(\mathbf{S}_D^{\mathbf{H}}\mathbf{R_n}\mathbf{S}_D\right)^{-1}\left(\mathbf{S}_D^{\mathbf{H}}E\left\{|\alpha|^2\mathbf{a}_{(\theta,\phi)}\mathbf{a}_{(\theta,\phi)}^{\mathbf{H}}\right\}\mathbf{S}_D\right)\right\} \\
&= \mathrm{tr}\left\{\left(\mathbf{S}_D^{\mathbf{H}}\mathbf{R_n}\mathbf{S}_D\right)^{-1}\left(\mathbf{S}_D^{\mathbf{H}}\mathbf{R_x}\mathbf{S}_D\right)\right\} \\
&= \mathrm{tr}\left\{\mathbf{SNR}\right\}
\end{aligned}
\tag{14}
$$

The equation (14) which shows that the average of point SINR values is the trace of the **SNR** matrix defined in (3). The $\mathbf{S}_D$ matrix maximizing the reward $\overline{\mathrm{SINR}}$ is the matrix whose column space is spanned by the first D dominant generalized eigenvectors of $\mathbf{R_x}$ and $\mathbf{R_n}$, as noted before.

IV. NUMERICAL RESULTS

We present a set of numerical results to examine the efficiency of the generalized eigenvector subspace for the reduced dimensional detection problem explained in Section III-C. Numerical evaluations are performed by using a uniform linear array (ULA) with 100 elements spaced half a wavelength apart along z-axis. We evaluate the performance metrics related to the point SINR at the output of the max-SINR beamformer given in (13), namely the Capon Beamformer, and the average output SINR in (14) over possible target elevation angle-of-arrival values. The set of elevation angle values over which target can be present is taken as $(85°, 95°)$. The target can be located at any point in this search region with equal probability. We assume three equi-powered interferers located at $\theta = 84°$, $88°$ and $97°$ for a given jammer-to-noise ratio (JNR). In this study, we compare the performance of generalized eigenvector subspace (shown to be optimal) based dimension reduction with that of the conventional subspace composed of equally spaced D beams in the search region whose steering vectors direct at $\left\{85° + \frac{10}{2D}(1+2k)\right\}_{k=0}^{D-1}$. This conventional subspace is known to be information preserving for spatially white interference case, and thus, widely used in practical beamforming applications.

In Figure 4, the average output SINR values in (14) as a function of the dimension are given for the conventional and the optimal subspaces over target search region $(85°, 95°)$. In this comparison, we have target SNR = 20 dB and JNR = 40 dB. It is seen that there is a remarkable SNR gap between the performances of reduced rank beamformers based on two subspaces especially at lower dimensions. SNR gap is defined as the output SINR difference between different methods as illustrated in Figure 4.

In Figures 5 and 6, the point SINR values given in (13) are shown for the conventional and the optimal subspaces with $D = 7$ and 9. Point SINR values are compared with that of the full dimension max-SINR beamformer ($D = N = 100$). It is seen that the optimal subspace covers the entire target search region $(85°, 95°)$ without a significant performance degradation even for $D = 7$, that is for roughly 15 fold dimension reduction. With the equally spaced conventional beams, in spite of the optimal processing after dimension reduction, the SINR performance over the region of interest
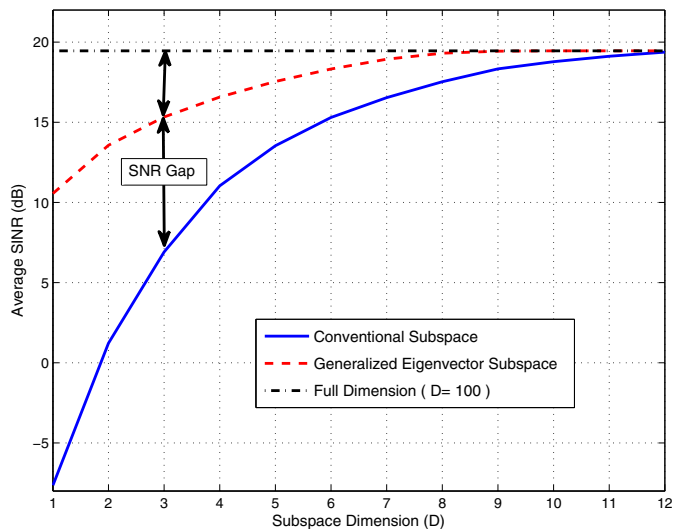

Fig. 4. Average of output point SINR values in (14) over Target Search Region $(85°, 95°)$, $SNR = 20$ dB, Interferers @ $\theta = 84°$, $88°$ and $97°$, total $JNR = 40$ dB.

is not satisfactory. There exists a significant performance degradation at points other than the jammer locations when the conventional beams are utilized for the dimension reduction.
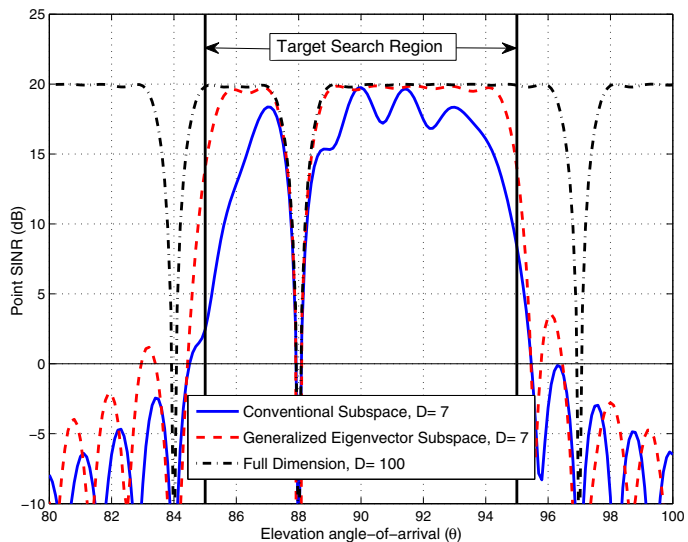

Fig. 5. Output point SINR values in (13) over Target Search Region $(85°, 95°)$, $SNR = 20$ dB, Interferers @ $\theta = 84°$, $88°$ and $97°$, total $JNR = 40$ dB, $D = 7$.

As a final comparison, the effect of JNR on the output SINR after reduced dimension processing is considered. From Figure 7, it can be observed that SNR gaps from the full dimension operation (as illustrated in Figure 4) gets large as jammer power increases for the conventional subspace. On the other hand, the optimal subspace exhibits a constant loss at all JNR values.
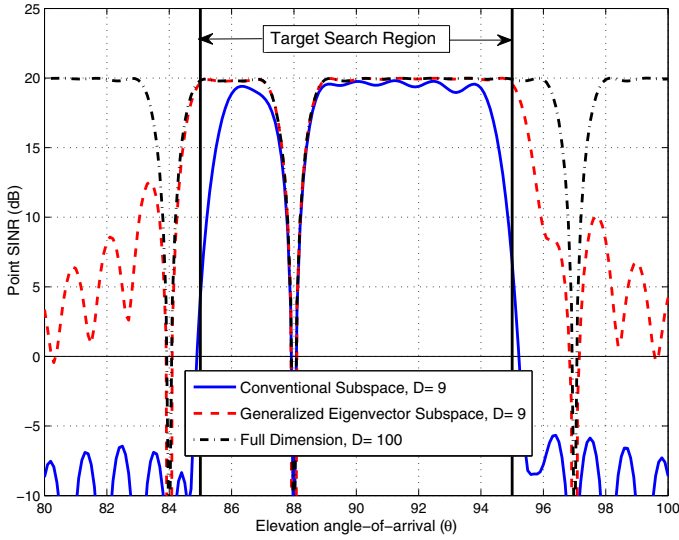
Fig. 6. Output point SINR values in (13) over Target Search Region $(85°, 95°)$, $SNR = 20$ dB, Interferers @ $\theta = 84°$, $88°$ and $97°$, total $JNR = 40$ dB, $D = 9$.
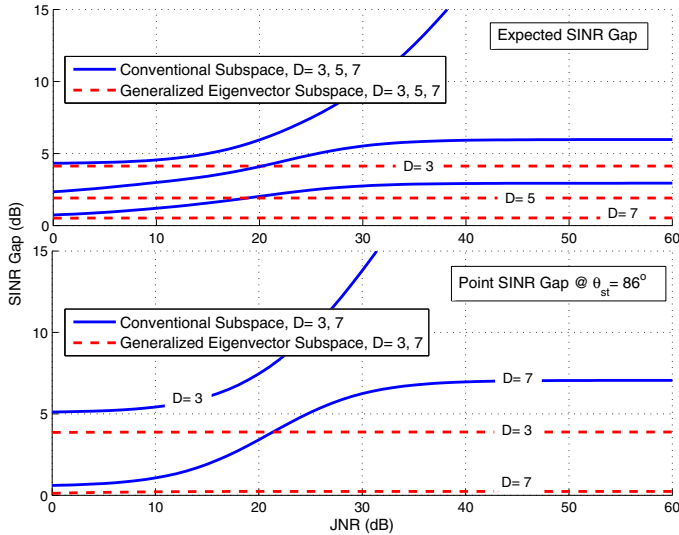

Fig. 7. Average and point SINR gap with respect to full dimension beamforming for different JNR values, $SNR = 20$ dB, Interferers @ $\theta = 84°$, $88°$ and $97°$, $D = 3, 5, 7$.

V. CONCLUSIONS

This work examines the choice of subspace for statistical signal processing in reduced dimensions. The dimension reduction is especially critical for modern radar systems equipped with several hundreds of receiving elements generating large dimensional observation vectors at the rate of millions or more times per second. The present work shows that generalized eigenvector space suggested by Scharf in 1987 [1] is indeed a suitable choice from several viewpoints. In [1], Scharf proposes the eigenvector space considering the J-divergence which is a criterion related with the asymptotics of the general Gaussian vector detection. Here we present three

other criteria, which are as important as if not more, for the optimality of the generalized eigenvector space.

APPENDIX A

We can list important properties of **SNR** matrix as follows:

1. **SNR** matrix is positive semi-definite, since $\left(\mathbf{S}_D^{\mathbf{H}}\mathbf{R_n}\mathbf{S}_D\right)$ and $\left(\mathbf{S}_D^{\mathbf{H}}\mathbf{R_x}\mathbf{S}_D\right)$ are positive semi-definite so is their inverse and their multiplication.

2. Generalized eigenvectors of $\mathbf{R_x}$ and $\mathbf{R_n}$ matrices diagonalizes the **SNR** matrix. Stated differently, if $\mathbf{R_x}\mathbf{e_k} = \mathbf{R_n}\mathbf{e_k}\lambda_k$, where λ_k and $\mathbf{e_k}$ is the k'th largest generalized eigenvalue ($\lambda_1 \geq \lambda_2 \geq \ldots > \lambda_N$) and its associated eigenvector, then a basis for N dimensional space can be written as

$$\mathbf{S}_N = [\mathbf{e_1}\, \mathbf{e_2}\, \ldots\, \mathbf{e_N}]. \qquad (15)$$

The generalized eigenvectors of symmetric matrices have the property of being $\mathbf{R_n}$ orthogonal, $\mathbf{e}_k^T\mathbf{R_n}\mathbf{e}_l = 0$, $k \neq l$, and $\mathbf{e}_k^T\mathbf{R_n}\mathbf{e}_k = 1$, and can be $\mathbf{R_n}$-orthonormalized as follows.

When $\mathbf{S}_N$ given in (15) is inserted in (3), the **SNR** reduces to a diagonal matrix with the generalized eigenvalues on its diagonal.

3. An alternative representation for the vectors of N dimensions, that is another basis for the subspace spanned by the columns of $\mathbf{S}_D$, results in a similarity transformation for **SNR** matrix. Stated differently, if $\mathbf{S}_D$ is replaced with $\mathbf{S}_D\mathbf{T}$ in (3) where $\mathbf{T}$ is an invertible matrix, **SNR** matrix becomes $\mathbf{T}^{-1}\left(\mathbf{S}_D^{\mathbf{H}}\mathbf{R_n}\mathbf{S}_D\right)^{-1}\left(\mathbf{S}_D^{\mathbf{H}}\mathbf{R_x}\mathbf{S}_D\right)\mathbf{T}$.

4. The cost/reward functions such as trace and determinant which are invariant to the basis representation remain invariant when applied to **SNR** matrix. Any other cost function depending solely on the eigenvalues of **SNR** matrix has the same property. For such functions, we may consider that basis vectors spanning the subspace are $\mathbf{R_n}$-orthonormalized without any loss of generality.

5. If we consider the trace of **SNR** matrix as the reward function, the maximum reward that can be achieved is $\sum_{k=1}^{N}\lambda_k$ where λ_k is non-negative valued generalized eigenvalues. (This is the reward value when there is no dimension reduction.) The maximal reward for $D = N - 1$ (one dimensional reduction) is $\sum_{k=1}^{N-1}\lambda_k$ and is achieved by $N - 1$ dominant generalized eigenvectors. This argument can be justified by noting that the reward of any other subspace containing $\mathbf{e_N}$ (the generalized eigenvector with the smallest eigenvalue) can be improved by replacing $\mathbf{e_N}$ with any $\mathbf{e_k}$ which is not already in the span of the subspace. (This argument, in essence, is the argument utilized to prove the mean square representation error optimality of Karhunen-Loève expansion.) Upon the repeated use of the same argument, it can be justified that the maximal reward for $D = \{1, 2, \ldots, N\}$ dimensional subspace is $\sum_{k=1}^{D}\lambda_k$ and this reward value is achieved by first D dominant generalized eigenvectors.

6. The maximum value of the reward function $\det(\mathbf{I} + \mathbf{SNR})$ is $\prod_{k=1}^{N}(1 + \lambda_k)$. (This is the reward value when there is no dimension reduction.) By noting first that each term in the product $\prod_{k=1}^{N}(1 + \lambda_k)$ is greater than 1 and repeating a similar argument to the one given in point 5; we can show that the maximal reward for $D = \{1, 2, \ldots, N\}$ dimensional subspace is $\prod_{k=1}^{D}(1 + \lambda_k)$ and this reward value is achieved by first D dominant generalized eigenvectors.

7. The minimum value of the cost function $\mathrm{tr}\left\{(\mathbf{I} + \mathbf{SNR})^{-1}\right\}$ is achieved for $D = 1$ dimensional with the dominant generalized eigenvector. The minimum value of the cost is $\frac{1}{1+\lambda_1}$. When D is increased to 2, the minimum value becomes $\frac{1}{1+\lambda_1} + \frac{1}{1+\lambda_2}$ and it is achieved by the first two dominant generalized eigenvectors. This fact can be justified by first unit $\mathbf{R_n}$-normalizing the basis vectors of candidate space and then expanding the candidate space with the basis shown in (15). It can be noted that replacing all odd indexed $\mathbf{e}_k$ vectors $\mathbf{e}_1$ and even indexed $\mathbf{e}_k$ vectors $\mathbf{e}_2$ results in an improved cost function.

ACKNOWLEDGMENT

REFERENCES

[1] L. Scharf and B. Van Veen, "Low rank detectors for Gaussian random vectors," *IEEE Trans. Acoust., Speech, and Signal Process.*, vol. 35, pp. 1579–1582, Nov. 1987.

[2] L. L. Scharf, *Statistical Signal Processing*. Addison-Wesley, 1991.

[3] L. Scharf and C. Mullis, "Canonical coordinates and the geometry of inference, rate, and capacity," *IEEE Trans. Signal Process.*, vol. 48, pp. 824–831, Mar. 2000.

[4] A. Pezeshki, L. Scharf, J. Thomas, and B. Van Veen, "Canonical coordinates are the right coordinates for low-rank Gauss-Gauss detection and estimation," *IEEE Trans. Signal Process.*, vol. 54, pp. 4817–4820, Dec. 2006.

[5] J. Goldstein, I. S. Reed, and L. Scharf, "A multistage representation of the Wiener filter based on orthogonal projections," *IEEE Trans. Information Theory*, vol. 44, pp. 2943–2959, Nov. 1998.

[6] L. Scharf, E. Chong, M. Zoltowski, J. Goldstein, and I. S. Reed, "Subspace expansion and the equivalence of conjugate direction and multistage Wiener filters," *IEEE Trans. Signal Process.*, vol. 56, pp. 5013–5019, Oct. 2008.

[7] W. Chen, U. Mitra, and P. Schniter, "On the equivalence of three reduced rank linear estimators with applications to DS-CDMA," *IEEE Trans. Information Theory*, vol. 48, pp. 2609–2614, Sep. 2002.

[8] T. Cover, *Elements of Information Theory*. Wiley-Interscience, 2006.

[9] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecomm.*, vol. 6, pp. 585–595, Oct. 1999.

[10] P. Stoica, Y. Jiang, and J. Li, "On MIMO channel capacity: an intuitive discussion," *IEEE Signal Processing Magazine*, vol. 22, pp. 83–84, May 2005.

[11] D. Palomar, J. Cioffi, and M.-A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, pp. 2381–2401, Sept. 2003.

[12] S. Shi, M. Schubert, and H. Boche, "Downlink MMSE Transceiver Optimization for Multiuser MIMO Systems: MMSE Balancing," *IEEE Trans. Signal Process.*, vol. 56, pp. 3702–3712, Aug. 2008.