[5] Hiamovich, A., and Bar-Ness, Y. (1991)
An eigenanalysis interference canceler.
IEEE Transactions on Acoustics, Speech, and Signal Processing, **39**, 1 (Jan. 1991), 76–84.

[6] Moffet, T. A. (1968)
Minimum redundancy linear arrays.
IEEE Transactions on Antennas and Propagation, **AP-16**, 2 (Mar. 1968).

[7] Leech, J. (1956)
On the representation of $1, 2, \ldots, n$ by difference.
Journal of the London Math Society, **31** (1956), 160–169.

[8] Pilai, S. U., Bar-Ness, Y., and Haber, F. (1985)
A new approach to array geometry for the improved spatial spectrum estimation.
IEEE Proceedings, **73**, 10 (Oct. 1985), 1522–1524.

[9] Panayirci, E., Bar-Ness, Y., and Chen, W.-L. (1992)
Conventional interference cancellation for minimum redundancy array structure.
Presented at the International Conference on Acoustics, Speech, and Signal Processing, ICASSP-92, San Francisco, 1992.

[10] Bar-Ness, Y., Panayirci, E., and Chen, W.-L. (1995)
Minimum redundancy array structure for interference cancellation.
Signal Processing, **42** (1995), 319–334.

[11] Bar-Ness, Y., Panayirci, E., and Chen, W.-L. (1992)
Eigenanalysis for interference cancellation with minimum redundancy array structure.
Presented at the IEEE 6th SP Workshop on Statistical Signal and Array Processing, Victoria, Canada, Oct. 1992.

[12] Frost, O. L. (1972)
An algorithm for linearly constrained adaptive array processing.
Proceedings of the IEEE, **60** (1972), 926–935.

[13] Golub, G. H., and Van Loan, C. F. (1983)
Matrix Computations.
Baltimore, MD: The John Hopkins University Press, 1983.

Asymptotic Distribution of the Conditional Signal-to-Noise Ratio in an Eigenanalysis-Based Adaptive Array

The statistical characterization of the conditioned signal-to-noise ratio (SNR) of the sample matrix inversion (SMI) method has been known for some time. An eigenanalysis-based detection method, referred to as the eigencanceler, has been shown to be a useful alternative to SMI, when the interference has low rank. In this work, the density function of the conditioned SNR is developed for the eigencanceler. The development is based on the asymptotic expansion of the distribution of the principal components of the covariance matrix. It is shown that, unlike the SMI method, the eigencanceler yields a conditional SNR distribution that is dependent on the covariance matrix. It is further shown that simpler, covariance matrix-independent approximations can be found for the large interference-to-noise case. The new distribution is shown to be in good agreement with the numerical data obtained from simulations.

Manuscript received June 27, 1995; revised June 17 and August 22, 1996.

IEEE Log No. T-AES/33/3/04764.

This work was supported in part by Air Force Rome Laboratory Grant F30602-94-1-0012.

0018-9251/97/$10.00 © 1997 IEEE

I. INTRODUCTION

Reed, et al. suggested the use of the maximum likelihood estimate of the covariance matrix in lieu of the true covariance matrix, to detect a known signal vector in unknown colored noise [1]. This method is commonly referred to as sample matrix inversion (SMI). While not optimal in any sense, this procedure has the advantage of being implemented as a simple linear combination of the array inputs.

The loss incurred by the estimation of the unknown colored noise can be evaluated from the conditioned signal-to-noise ratio (CSNR). The CSNR is defined as the ratio of the SNR achieved by the adaptive filter derived from the available data, to the optimal SNR (when the true data covariance matrix is available). As the data has a statistical model, the CSNR is a random variable. Reed, et al. [1] determined its distribution, and used it to analyze the performance of the SMI method. This distribution turned out to have the remarkable property of being independent of the actual covariance matrix. In particular, they showed that if the signal vector dimension is N, the number of samples required to achieve performance within 3 dB of the optimal (i.e., when the true covariance matrix is known) is approximately $K = 2N$. Other authors provided alternative proofs or extended the SMI analysis [2–4].

When the colored noise can be modeled as the aggregate of an interference with a low-rank covariance matrix and white noise, eigenanalysis can be exploited to design a detector with faster convergence than SMI. The detector is derived from the interference subspace and has been referred to as the *eigencanceler* [5]. The eigencanceler is formulated as a modified minimum variance beamformer. A related approach is the PCI method derived from a low-rank approximation to the data matrix [6]. The minimum variance beamformer minimizes the array output subject to a set of linear constraints [7]. The eigencanceler produces the minimum norm weight vector meeting the set of linear constraints, and the additional constraint of orthogonality to the interference subspace [8]. In the special case of a single linear constraint, the eigencanceler and PCI provide the same solution.

An expression for the probability density of the CSNR for the PCI method has been derived in [9].

It is shown therein that this probability density is the same as the probability density for the SMI method, except that the parameter $(r + 1)$, where r is the interference subspace rank, is substituted for the signal dimensionality. It results that the method converges within 3 dB of optimal for $K \simeq 2r$.

Here, we suggest a new approach resulting in a different expression for the probability density of the CSNR for the eigenanalysis-based detector. The new probability density is derived from the asymptotic properties of the eigenvectors of the estimated covariance matrix. It is shown that, unlike the SMI method and the result in [9], the probability density depends on the true covariance matrix. Two simpler approximations, independent of the covariance matrix, are derived for a large interference-to-noise ratio. While not similar in form, these approximations produce numerical results close to those obtained using the probability density in [9]. It is hoped that the asymptotic analysis presented here provides new insights into the properties of reduced-rank methods.

Section II in the sequel presents the signal model. The CSNR probability density is developed in Section III. Section IV contains numerical simulations to support the theory developed here. Conclusions are provided in Section V.

II. SIGNAL MODEL

Consider an N-dimensional signal space spanned by signals received by a spatial (possibly space-time) array. Under hypothesis $\mathbf{H}_0$, the received signal $\mathbf{x}$ consists only of interference and noise contributions

$$\mathbf{x} = \mathbf{c} + \mathbf{v} \tag{1}$$

where $\mathbf{x}$ is assumed to be a zero-mean, circularly symmetric N-dimensional Gaussian random vector with covariance matrix $\mathbf{R}$. Under hypothesis $\mathbf{H}_1$, $\mathbf{x}$ is given by

$$\mathbf{x} = a\mathbf{s} + \mathbf{c} + \mathbf{v} \tag{2}$$

where a is a zero-mean, circularly symmetric Gaussian random variable with variance σ_s^2, and $\mathbf{s}$ is the target vector. Consequently, under both hypothesis models, $\mathbf{x}$ is a zero-mean, $N \times 1$ complex-valued normal random vector with the $N \times N$ covariance matrix $\mathbf{R}_i$, $i = 0, 1$, where $\mathbf{R}_0 = \mathbf{R}$ and $\mathbf{R}_1 = \sigma_s^2 ss^H + \mathbf{R}$, and the superscript stands for complex transposition. A distribution so defined is denoted $\mathbf{x} \sim \mathcal{N}(0, \mathbf{R}_i)$. The matrix $\mathbf{R}$ is further assumed to obey the model

$$\mathbf{R} = \mathbf{Q}_1 \mathbf{\Lambda}_1 \mathbf{Q}_1^H + \sigma_v^2 \mathbf{Q}_2 \mathbf{Q}_2^H \tag{3}$$

where $\mathbf{Q}_1$ is the $N \times r$ matrix of principal eigenvectors, $\mathbf{\Lambda}_1$ is the $r \times r$ diagonal matrix of principal eigenvalues, σ_v^2 is the variance of the white noise, and $\mathbf{Q}_2$ is the $N \times p$ matrix of noise eigenvectors, where $p = N - r$. The principal and noise eigenvectors are said to span the *interference* and *noise subspaces*, respectively. The eigencanceler weight vector for the case of a single steering vector linear constraint is given by [5, 8]

$$\begin{aligned} \mathbf{w} &= (\mathbf{I}_N - \hat{\mathbf{Q}}_1 \hat{\mathbf{Q}}_1^H)\mathbf{s} \\ &= \hat{\mathbf{Q}}_2 \hat{\mathbf{Q}}_2^H \mathbf{s} \end{aligned} \tag{4}$$

where $\mathbf{I}_N$ is the $N \times N$ unit matrix, $\hat{\mathbf{Q}}_1$ and $\hat{\mathbf{Q}}_2$ are the respective interference and noise subspaces in the spectral decomposition of the estimated covariance matrix, $\hat{\mathbf{R}} = \hat{\mathbf{Q}}_1 \hat{\mathbf{\Lambda}}_1 \hat{\mathbf{Q}}_1^H + \hat{\mathbf{Q}}_2 \hat{\mathbf{\Lambda}}_2 \hat{\mathbf{Q}}_2^H$. The estimate $\hat{\mathbf{R}}$ is computed from a set of K independent interference $N \times 1$ snapshot vectors $\mathbf{x}_k$, $k = 1, \ldots, K$, sometimes referred to as secondary data [1]:

$$\hat{\mathbf{R}} = \frac{1}{K} \sum_{k=1}^{K} \mathbf{x}_k \mathbf{x}_k^H. \tag{5}$$

In this work we exploit some asymptotic properties of $\hat{\mathbf{R}}$. Statisticians often make use of asymptotic analysis [10, 12], but this approach has found applications in signal processing as well [11, 13]. In reference [12], Gupta establishes that the asymptotic distribution of $\mathbf{B} = \sqrt{K}[\hat{\mathbf{R}} - \mathbf{R}]$ is normal as a direct result of the Central Limit Theorem. We use the normality of $\mathbf{B}$ to prove the following result.

LEMMA 1 *Let $\hat{\mathbf{R}} = (1/K)\sum_{k=1}^{K} \mathbf{x}_k \mathbf{x}_k^H$, where $\mathbf{x}_k$ are assumed to be zero-mean, circularly symmetric N-dimensional random vectors, independent identically distributed (IID) with $\mathcal{N}(\mathbf{0}, \mathbf{R})$ distribution. Then the limiting distribution of $\mathbf{B} = \sqrt{K}[\hat{\mathbf{R}} - \mathbf{R}]$, as $K \to \infty$, has zero-means and covariances*

$$E[b_{ij} b_{lm}^*] = \sigma_{il} \sigma_{jm}^* \tag{6}$$

where σ_{im} are the elements of the matrix $\mathbf{R}$, and the Hermitian property of $\mathbf{R}$ implies that $\sigma_{il} = \sigma_{li}^*$.

PROOF Let $x_i(k)$ be the ith component, $i = 1, \ldots, N$, of the snapshot vector $\mathbf{x}_k$. From the definition of σ_{ij}, $E[x_i(k) x_j^*(k)] = \sigma_{ij}$. The expected value of the product of four zero-mean Gaussian random variables is given by (for example [13, p. 381]):

$$E[x_i(k) x_j^*(k) x_l^*(k) x_m(k)] = \sigma_{ij} \sigma_{lm}^* + \sigma_{il} \sigma_{jm}^*. \tag{7}$$

It results that

$$\begin{aligned} E[b_{ij} b_{lm}^*] &= KE\left[\left(\frac{1}{K}\sum_{k=1}^{K} x_i(k) x_j^*(k) - \sigma_{ij}\right) \times \left(\frac{1}{K}\sum_{k=1}^{K} x_l^*(k) x_m(k) - \sigma_{lm}^*\right)\right] \\ &= \sigma_{il} \sigma_{jm}^*. \end{aligned} \tag{8}$$

III. DISTRIBUTION OF CSNR

In this section, the probability density of the CSNR is developed. The CSNR is the SNR obtained by the application of a specified weight vector $\mathbf{w}$, normalized by the SNR of the optimal case [1]:

$$\rho = \frac{|\mathbf{w}^H\mathbf{s}|^2}{\mathbf{w}^H\mathbf{R}\mathbf{w}}\frac{1}{\mathbf{s}^H\mathbf{R}^{-1}\mathbf{s}}. \tag{9}$$

Clearly, the CSNR is bounded, $0 \le \rho \le 1$. Substituting (4) in (9) we obtain

$$\rho = \frac{(\mathbf{s}^H\hat{\mathbf{Q}}_2\hat{\mathbf{Q}}_2^H\mathbf{s})^2}{\mathbf{s}^H\hat{\mathbf{Q}}_2\hat{\mathbf{Q}}_2^H\mathbf{R}\hat{\mathbf{Q}}_2\hat{\mathbf{Q}}_2^H\mathbf{s}}\frac{1}{\mathbf{s}^H\mathbf{R}^{-1}\mathbf{s}} \tag{10}$$

for the eigencanceler. Our goal is to characterize the CSNR in (10) statistically. The perturbation analysis of the noise subspace $\hat{\mathbf{Q}}_2$ is required before the characterization of ρ can proceed.

Perturbation Analysis of Noise Subspace

The objective is to develop an expression for the estimated noise subspace $\hat{\mathbf{Q}}_2$, in terms of the interference and noise subspaces of the true covariance matrix. The spectral decompositions of the true and estimated covariance matrices are

$$\mathbf{R} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H \tag{11}$$

and

$$\hat{\mathbf{R}} = \hat{\mathbf{Q}}\hat{\mathbf{\Lambda}}\hat{\mathbf{Q}}^H \tag{12}$$

respectively. Let the notation $\mathbf{Q} = \{q_{ij}\}$ denote the matrix $\mathbf{Q}$ with elements q_{ij}. For uniqueness, $\mathbf{Q} = \{q_{ij}\}$ and $\hat{\mathbf{Q}} = \{\hat{q}_{ij}\}$ are selected such that $q_{ii}, \hat{q}_{ii} > 0$. Assume that the covariance matrix is estimated using relation (5).

Define the $N \times N$ matrix

$$\tilde{\mathbf{\Lambda}} = \mathbf{Q}^H\hat{\mathbf{R}}\mathbf{Q}. \tag{13}$$

This matrix represents perturbed eigenvalues and is in general complex-valued. The perturbation from the true eigenvalues is given by

$$\mathbf{U} = \sqrt{K}(\tilde{\mathbf{\Lambda}} - \mathbf{\Lambda}). \tag{14}$$

The following theorem is formulated for the asymptotic distribution of the elements of $\mathbf{U}$.

THEOREM 1 *The limiting distribution of* $\mathbf{U}$ *is normal with zero-mean and covariance*

$$E[u_{ij}u_{lm}^*] = \lambda_i\lambda_j\delta_{il}\delta_{jm}, \tag{15}$$

where λ_i *is an element of* $\mathbf{\Lambda}$, *and* δ_{im} *is the Kronecker delta.*

PROOF The proof follows from Lemma 1 applied to the vectors $\mathbf{z}_k = \mathbf{Q}^H\mathbf{x}_k$ and the normality of the limiting distribution of the matrix $\mathbf{B}$. From (13) and (14) and with $\mathbf{B} = \sqrt{K}[\hat{\mathbf{R}} - \mathbf{R}]$, the matrix $\mathbf{U}$ can be expressed as

$$\mathbf{U} = \mathbf{Q}^H\mathbf{B}\mathbf{Q}. \tag{16}$$

The normality of the limiting distribution of $\mathbf{B}$ is maintained through the linear transformation $\mathbf{Q}$, hence the limiting distribution of $\mathbf{U}$ is normal as well. Applying Lemma 1, we have $E[u_{ij}u_{lm}^*] = \mu_{il}\mu_{jm}^*$, where μ_{im} is an element of $\mathbf{Q}^H\mathbf{R}\mathbf{Q} = \mathbf{\Lambda}$. But $\mathbf{\Lambda}$ is a nonnegative diagonal matrix of the eigenvalues of $\mathbf{R}$, hence we have $\mu_{il} = \lambda_i\delta_{il}$ and $\mu_{jm}^* = \lambda_j\delta_{jm}$. Q.E.D.

Define the $N \times N$ matrix $\mathbf{Y} = \mathbf{Q}^H\hat{\mathbf{Q}}$. Since both $\mathbf{Q}$ and $\hat{\mathbf{Q}}$ are unitary, so is $\mathbf{Y}$. The perturbed eigenvectors can be expressed $\hat{\mathbf{Q}} = \mathbf{Q}\mathbf{Y}$. From the definition of $\mathbf{Y}$ and (13) we have $\tilde{\mathbf{\Lambda}} = \mathbf{Q}^H\hat{\mathbf{R}}\mathbf{Q} = \mathbf{Q}^H\hat{\mathbf{Q}}\hat{\mathbf{\Lambda}}\hat{\mathbf{Q}}^H\mathbf{Q} = \mathbf{Y}\hat{\mathbf{\Lambda}}\mathbf{Y}^H$. Define

$$\mathbf{V} = \sqrt{K}(\mathbf{Y} - \mathbf{I}_N) \tag{17}$$

then $\mathbf{Y} = \mathbf{I}_N + (1/\sqrt{K})\mathbf{V}$. From $\tilde{\mathbf{\Lambda}} = \mathbf{Y}\hat{\mathbf{\Lambda}}\mathbf{Y}^H$, the previous relation and (14), $\tilde{\mathbf{\Lambda}}$ can be written

$$\begin{aligned}
\tilde{\mathbf{\Lambda}} &= \mathbf{\Lambda} + \frac{1}{\sqrt{K}}\mathbf{U} \\
&= \left(\mathbf{I}_N + \frac{1}{\sqrt{K}}\mathbf{V}\right)\hat{\mathbf{\Lambda}}\left(\mathbf{I}_N + \frac{1}{\sqrt{K}}\mathbf{V}\right)^H \\
&= \left(\mathbf{I}_N + \frac{1}{\sqrt{K}}\mathbf{V}\right)\left(\mathbf{\Lambda} + \frac{1}{\sqrt{K}}\mathbf{D}\right)\left(\mathbf{I}_N + \frac{1}{\sqrt{K}}\mathbf{V}\right)^H \\
&= \mathbf{\Lambda} + \frac{1}{\sqrt{K}}(\mathbf{V}\mathbf{\Lambda} + \mathbf{D} + \mathbf{\Lambda}\mathbf{V}^H) + \mathbf{M}
\end{aligned} \tag{18}$$

where $\mathbf{D} = \sqrt{K}(\hat{\mathbf{\Lambda}} - \mathbf{\Lambda})$ and $\mathbf{M} = 1/K(\mathbf{V}\mathbf{D} + \mathbf{D}\mathbf{V}^H + \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H) + (1/K^{3/2})\mathbf{V}\mathbf{D}\mathbf{V}^H$. In the following we make use of the notation $o(1/K)$. The notation $\mathbf{M} = o(1/K)$ is defined as $\lim_{K\to\infty}\Pr[\mathbf{M} \le (1/K)\mathbf{H}(K)] = 1$, where the inequality is component-wise and $\mathbf{H}(K)$ is bounded as $K \to \infty$. Note that according to this definition $\mathbf{V}$, $\mathbf{\Lambda}$, and $\mathbf{D}$ are $o(1)$, i.e., bounded as $K \to \infty$. Also note that the calculus of $o(1/K)$ implies that if $\mathbf{A}$ and $\mathbf{B}$ are $o(1/K)$, so is $\mathbf{A} + \mathbf{B}$. Using the $o(1/K)$ notion, we can define the symbol $\cong$ to denote equality to the $1/K$ order, i.e., $\mathbf{A} \cong \mathbf{B}$ implies that $\mathbf{A} - \mathbf{B} = o(1/K)$. The matrix $\mathbf{M}$ in (18) is $o(1/K)$. If we neglect $\mathbf{M}$ in the equation, the resulting equality holds to order $1/K$ accuracy. A term-by-term comparison of the first and last lines in (18) yields

$$\mathbf{U} \cong \mathbf{V}\mathbf{\Lambda} + \mathbf{D} + \mathbf{\Lambda}\mathbf{V}^H \tag{19}$$

where the symbol $\cong$ is used for asymptotic approximation.

According to the model assumed in (3), the $N \times N$ matrix of eigenvalues $\mathbf{\Lambda}$ can be partitioned

$$\mathbf{\Lambda} = \begin{pmatrix} \mathbf{\Lambda}_1 & \mathbf{0} \\ \mathbf{0} & \sigma_v^2\mathbf{I}_p \end{pmatrix}. \tag{20}$$

Similarly, the $N \times N$ matrix $\mathbf{V}$ is partitioned into the following matrices: $r \times r$ $\mathbf{V}_{11}$, $r \times p$ $\mathbf{V}_{12}$, $p \times r$ $\mathbf{V}_{21}$,

$p \times p$ $\mathbf{V}_{22}$:

$$\mathbf{V} = \begin{pmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{pmatrix}. \tag{21}$$

Thus, if the elements of $\mathbf{V}$ are denoted, v_{ij}, i.e., $\mathbf{V} = \{v_{ij}\}_{i=1,\ldots,N,j=1,\ldots,N}$, then $V_{12} = \{v_{ij}\}_{i=1,\ldots,r,j=r+1,\ldots,N}$. Let the matrix $\mathbf{U}$ be partitioned in a similar way. We have from (19), (20), and (21), and noting that $\mathbf{D} = \sqrt{\mathbf{K}}(\hat{\mathbf{\Lambda}} - \mathbf{\Lambda})$ is a diagonal matrix,

$$\mathbf{U}_{12} = \sigma_v^2 \mathbf{V}_{12} + \mathbf{\Lambda}_1 \mathbf{V}_{21}^H. \tag{22}$$

To proceed, we show that $\mathbf{V}_{12} + \mathbf{V}_{21}^H \cong \mathbf{0}$. Partition $\mathbf{Y}$ in a fashion similar to $\mathbf{V}$, i.e.,

$$\mathbf{Y} = \begin{pmatrix} \mathbf{Y}_{11} & \mathbf{Y}_{12} \\ \mathbf{Y}_{21} & \mathbf{Y}_{22} \end{pmatrix}$$

with $\mathbf{Y}_{11}$ $r \times r$, $\mathbf{Y}_{12}$ $r \times p$, etc. It follows from relation (17) that

$$\mathbf{Y} = \begin{pmatrix} \mathbf{I}_r + \frac{1}{\sqrt{K}}\mathbf{V}_{11} & \frac{1}{\sqrt{K}}\mathbf{V}_{12} \\ \frac{1}{\sqrt{K}}\mathbf{V}_{21} & \mathbf{I}_p + \frac{1}{\sqrt{K}}\mathbf{V}_{22} \end{pmatrix}. \tag{23}$$

Using the unitary property of the matrix $\mathbf{Y}$ and the relation above, we have

$$\begin{aligned} \mathbf{I}_p &= \mathbf{Y}_{21}\mathbf{Y}_{21}^H + \mathbf{Y}_{22}\mathbf{Y}_{22}^H \\ &= \frac{1}{K}\mathbf{V}_{21}\mathbf{V}_{12}^H + \left(\mathbf{I}_p + \frac{1}{\sqrt{K}}\mathbf{V}_{22}\right)\left(\mathbf{I}_p + \frac{1}{\sqrt{K}}\mathbf{V}_{22}\right)^H. \end{aligned} \tag{24}$$

From the last relation it follows that

$$\begin{aligned} \left(\mathbf{I}_p + \frac{1}{\sqrt{K}}\mathbf{V}_{22}\right)\left(\mathbf{I}_p + \frac{1}{\sqrt{K}}\mathbf{V}_{22}\right)^H &= \mathbf{I}_p - \frac{1}{K}\mathbf{V}_{21}\mathbf{V}_{12}^H \\ &\cong \mathbf{I}_p. \end{aligned} \tag{25}$$

From the unitary property of $\mathbf{Y}$ we also have

$$\begin{aligned} \mathbf{0} &= \mathbf{Y}_{11}\mathbf{Y}_{21}^H + \mathbf{Y}_{12}\mathbf{Y}_{22}^H \\ &= \frac{1}{\sqrt{K}}(\mathbf{V}_{12} + \mathbf{V}_{21}^H) + \frac{1}{K}(\mathbf{V}_{11}\mathbf{V}_{21}^H + \mathbf{V}_{12}\mathbf{V}_{22}^H) \\ &\cong \frac{1}{\sqrt{K}}(\mathbf{V}_{12} + \mathbf{V}_{21}^H) \end{aligned} \tag{26}$$

where $o(1/K)$ terms were neglected to obtain the last relation. Using $\mathbf{V}_{12} \cong -\mathbf{V}_{21}^H$ in (22) we obtain

$$\mathbf{V}_{12} = -(\mathbf{\Lambda}_1 - \sigma_v^2 \mathbf{I}_r)^{-1}\mathbf{U}_{12}. \tag{27}$$

The next theorem establishes the asymptotic distribution of the components of $\mathbf{V}_{12}$.

THEOREM 2 *The limiting distribution of the $r \times p$ matrix $\mathbf{V}_{12}$ is normal with zero-mean and*

$$E[v_{ij}v_{lm}^*] = \frac{\lambda_i \sigma_v^2}{(\lambda_i - \sigma_v^2)^2}\delta_{il}\delta_{jm} \tag{28}$$

where $\mathbf{V} = \{v_{ij}\}_{i=1,\ldots,N,j=1,\ldots,N}$, *and* $\mathbf{V}_{12} = \{v_{ij}\}_{i=1,\ldots,r,j=r+1,\ldots,N}$.

PROOF A term-by-term listing of (27) yields

$$v_{ij} = -\frac{u_{ij}}{\lambda_i - \sigma_v^2} \tag{29}$$

for $i = 1,\ldots,r$ and $j = r+1,\ldots,N$. Therefore,

$$\begin{aligned} E[v_{ij}v_{lm}^*] &= \frac{E[u_{ij}u_{lm}^*]}{(\lambda_i - \sigma_v^2)^2} \\ &= \frac{\lambda_i \sigma_v^2}{(\lambda_i - \sigma_v^2)^2}\delta_{il}\delta_{jm}. \end{aligned} \tag{30}$$

The last relation was obtained applying Theorem 1. Q.E.D.

A consequence of Theorem 2 is that the elements of $\mathbf{V}_{12}$ are asymptotically independent.

The columns of the matrices $\mathbf{Q}$ and $\hat{\mathbf{Q}}$ are the eigenvectors of $\mathbf{R}$ and $\hat{\mathbf{R}}$, respectively. Define the $N \times r$ matrix $\mathbf{Q}_1$ and the $N \times p$, $(r + p = N)$ matrix $\mathbf{Q}_2$ such that $\mathbf{Q} = [\mathbf{Q}_1 \mid \mathbf{Q}_2]$. Note that $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are orthogonal, i.e., $\mathbf{Q}_1^H\mathbf{Q}_2 = 0$. The matrix $\hat{\mathbf{Q}}$ is partitioned similarly, $\hat{\mathbf{Q}} = [\hat{\mathbf{Q}}_1 \mid \hat{\mathbf{Q}}_2]$. From the relations above, and from $\hat{\mathbf{Q}} = \mathbf{Q}\mathbf{Y}$ and (17), we find that

$$\begin{aligned} \hat{\mathbf{Q}}_2 &= \mathbf{Q}_1\mathbf{Y}_{12} + \mathbf{Q}_2\mathbf{Y}_{22} \\ &= \mathbf{Q}_2\left(\mathbf{I}_p + \frac{1}{\sqrt{K}}\mathbf{V}_{22}\right) + \frac{1}{\sqrt{K}}\mathbf{Q}_1\mathbf{V}_{12}. \end{aligned} \tag{31}$$

Note that $\mathbf{Q}_2(\mathbf{I}_p + (1/\sqrt{K})\mathbf{V}_{22})$ and $\mathbf{Q}_1$ are orthogonal, i.e., $(\mathbf{I}_p + (1/\sqrt{K})\mathbf{V}_{22})^H\mathbf{Q}_2^H\mathbf{Q}_1 = \mathbf{0}$. The last expression provides the perturbation analysis of the estimated noise subspace $\hat{\mathbf{Q}}_2$ in terms of the true noise and interference subspaces, $\mathbf{Q}_2$ and $\mathbf{Q}_1$, respectively. This relation is used to derive the probability density of the conditioned SNR.

Computation of Probability Density Function

Define the transformed noise subspace, the $N \times p$ matrix $\mathbf{C} = \mathbf{R}^{1/2}\hat{\mathbf{Q}}_2$, where the square root is defined such that $\mathbf{R}^{1/2}\mathbf{R}^{1/2} = \mathbf{R}$. Starting from the definition of the CSNR ρ, substituting $\mathbf{R}^{-1/2}\mathbf{C} = \hat{\mathbf{Q}}_2$, and defining the transformed steering vector $\mathbf{a} = \mathbf{R}^{-1/2}\mathbf{s}$, we get

$$\begin{aligned} \rho &= \frac{(\mathbf{s}^H\hat{\mathbf{Q}}_2\hat{\mathbf{Q}}_2^H\mathbf{s})^2}{\mathbf{s}^H\hat{\mathbf{Q}}_2\hat{\mathbf{Q}}_2^H\mathbf{R}\hat{\mathbf{Q}}_2\hat{Q}_2^H\mathbf{s}}\frac{1}{\mathbf{s}^H\mathbf{R}^{-1}\mathbf{s}} \\ &= \frac{(\mathbf{s}^H\mathbf{R}^{-1/2}\mathbf{C}\mathbf{C}^H\mathbf{R}^{-1/2}\mathbf{s})^2}{\mathbf{s}^H\mathbf{R}^{-1/2}\mathbf{C}\mathbf{C}^H\mathbf{C}\mathbf{C}^H\mathbf{R}^{-1/2}\mathbf{s}}\frac{1}{\mathbf{s}^H\mathbf{R}^{-1}\mathbf{s}} \\ &= \frac{(\mathbf{a}^H\mathbf{C}\mathbf{C}^H\mathbf{a})^2}{\mathbf{a}^H\mathbf{C}\mathbf{C}^H\mathbf{C}\mathbf{C}^H\mathbf{a}}\frac{1}{\mathbf{a}^H\mathbf{a}}. \end{aligned} \tag{32}$$

Substituting the unit magnitude vector $\mathbf{b} = \mathbf{a}/|\mathbf{a}|$ for $\mathbf{a}$ above and letting $\mathbf{z} = \mathbf{C}^H\mathbf{b}$ yields

$$\rho = \frac{(\mathbf{z}^H\mathbf{z})^2}{\mathbf{z}^H\mathbf{C}^H\mathbf{C}\mathbf{z}}. \tag{33}$$

Using the relation $\mathbf{R}^{1/2} = \mathbf{Q}_1\boldsymbol{\Lambda}_1^{1/2}\mathbf{Q}_1^H + \sigma_v\mathbf{Q}_2\mathbf{Q}_2^H$, (31), and the orthogonality between $\mathbf{Q}_1$ and $\mathbf{Q}_2$, we get

$$\mathbf{C} = \sigma_v\mathbf{Q}_2\left(\mathbf{I}_p + \frac{1}{\sqrt{K}}\mathbf{V}_{22}\right) + \frac{1}{\sqrt{K}}\mathbf{Q}_1\boldsymbol{\Lambda}_1^{1/2}\mathbf{V}_{12}. \tag{34}$$

Expand the term $\mathbf{z}$ using (34) and the orthogonality relations $\mathbf{Q}_2^H\mathbf{Q}_2 = \mathbf{I}_p$ and $\mathbf{Q}_2^H\mathbf{Q}_1 = 0$:

$$\mathbf{z} = \sigma_v\left(\mathbf{I}_p + \frac{1}{\sqrt{K}}\mathbf{V}_{22}\right)^H\mathbf{Q}_2^H\mathbf{b} + \frac{1}{\sqrt{K}}\mathbf{V}_{12}^H\boldsymbol{\Lambda}_1^{1/2}\mathbf{Q}_1^H\mathbf{b}. \tag{35}$$

We make the assumption that the projection of the steering vector on the true interference subspace is negligible with respect to its projection on the true noise subspace, i.e., $|\mathbf{Q}_1^H\mathbf{s}| \ll |\mathbf{Q}_2^H\mathbf{s}|$. This assumption is reasonable and is just an expression of the requirement that the interference be received in the sidelobe region. Mainlobe interference is not addressed here. From the definition of vectors $\mathbf{a}$ and $\mathbf{b}$ we have: $|\mathbf{Q}_1^H\mathbf{s}| = |\mathbf{Q}_1^H\mathbf{R}^{1/2}\mathbf{a}| = |\mathbf{a}|\,|\boldsymbol{\Lambda}_1^{1/2}\mathbf{Q}_1^H\mathbf{b}| \geq |\mathbf{a}|\,|\sigma_v\mathbf{Q}_1^H\mathbf{b}|$, where the last inequality follows from the observation that σ_v is smaller than any element of $\boldsymbol{\Lambda}_1^{1/2}$. Likewise, $|\mathbf{Q}_2^H\mathbf{s}| = |\mathbf{a}|\,|\sigma_v\mathbf{Q}_2^H\mathbf{b}|$. Since, $|\mathbf{a}|\,|\sigma_v\mathbf{Q}_1^H\mathbf{b}| \leq |\mathbf{Q}_1^H\mathbf{s}| \ll |\mathbf{Q}_2^H\mathbf{s}| = |\mathbf{a}|\,|\sigma_v\mathbf{Q}_2^H\mathbf{b}|$, it follows that $|\mathbf{Q}_1^H\mathbf{b}| \ll |\mathbf{Q}_2^H\mathbf{b}|$. The vector $\mathbf{b}$ has unity length, hence $1 = \mathbf{b}^H\mathbf{b} = \mathbf{b}^H(\mathbf{Q}_1\mathbf{Q}_1^H + \mathbf{Q}_2\mathbf{Q}_2^H)\mathbf{b}$ implies that

$$\mathbf{b}^H\mathbf{Q}_2\mathbf{Q}_2^H\mathbf{b} \cong 1. \tag{36}$$

With this result, and using (25), we have $\mathbf{z}^H\mathbf{z} \cong \sigma_v^2\mathbf{b}^H\mathbf{Q}_2(\mathbf{I}_p + (1/\sqrt{K})\mathbf{V}_{22})(\mathbf{I}_p + (1/\sqrt{K})\mathbf{V}_{22})^H\mathbf{Q}_2^H\mathbf{b} \cong \sigma_v^2$, and the numerator of (33) becomes

$$(\mathbf{z}^H\mathbf{z})^2 \cong \sigma_v^4. \tag{37}$$

To evaluate the denominator, first compute the $p \times p$ matrix $\mathbf{C}^H\mathbf{C}$. From (34) and (25):

$$\begin{aligned}\mathbf{C}^H\mathbf{C} &= \sigma_v^2\left(\left(\mathbf{I}_p + \frac{1}{\sqrt{K}}\mathbf{V}_{22}\right)^H\mathbf{Q}_2^H\mathbf{Q}_2\left(\mathbf{I}_p + \frac{1}{\sqrt{K}}\mathbf{V}_{22}\right)\right.\\ &\qquad \left. + \frac{1}{K}\sigma_v^{-2}\mathbf{V}_{12}^H\boldsymbol{\Lambda}_1\mathbf{V}_{12}\right)\\ &\cong \sigma_v^2\left(\mathbf{I}_p + \frac{1}{K}\sigma_v^{-2}\mathbf{V}_{12}^H\boldsymbol{\Lambda}_1\mathbf{V}_{12}\right).\end{aligned} \tag{38}$$

As a result, the denominator of (33) is evaluated:

$$\begin{aligned}\mathbf{z}^H\mathbf{C}^H\mathbf{C}\mathbf{z} &\cong \sigma_v^2\mathbf{z}^H\left(\mathbf{I}_p + \frac{1}{K}\sigma_v^{-2}\mathbf{V}_{12}^H\boldsymbol{\Lambda}_1\mathbf{V}_{12}\right)\mathbf{z}\\ &\cong \sigma_v^2\mathbf{z}^H\mathbf{z} + \frac{1}{K}\mathbf{z}^H\mathbf{V}_{12}^H\boldsymbol{\Lambda}_1\mathbf{V}_{12}\mathbf{z}\\ &\cong \sigma_v^4 + \frac{1}{K}\mathbf{z}^H\mathbf{V}_{12}^H\boldsymbol{\Lambda}_1\mathbf{V}_{12}\mathbf{z}\end{aligned} \tag{39}$$

where we made use of the relation $\mathbf{z}^H\mathbf{z} \cong \sigma_v^2$ (from (37)). Using (35) in (39) and neglecting terms $o(1/K^{3/2})$ and higher, the term $(1/K)\mathbf{z}^H\mathbf{V}_{12}^H\boldsymbol{\Lambda}_1\mathbf{V}_{12}\mathbf{z}$ is asymptotically equal to $(1/K)\sigma_v^2\mathbf{b}^H\mathbf{Q}_2\mathbf{V}_{12}^H\boldsymbol{\Lambda}_1\mathbf{V}_{12}\mathbf{Q}_2^H\mathbf{b}$. Thus we have

$$\mathbf{z}^H\mathbf{C}^H\mathbf{C}\mathbf{z} \cong \sigma_v^4 + \frac{1}{K}\mathbf{t}^H\mathbf{V}_{12}^H\boldsymbol{\Lambda}_1\mathbf{V}_{12}\mathbf{t} \tag{40}$$

where $\mathbf{t} = \sigma_v\mathbf{Q}_2^H\mathbf{b}$ and $\mathbf{t}^H\mathbf{t} \cong \sigma_v^2$ (see (36)). Write the $r \times p$ matrix $\mathbf{V}_{12}$ in terms of its rows:

$$\mathbf{V}_{12} = \{v_{ij}\}_{i=1,\ldots,r,j=r+1,\ldots,N} = \begin{pmatrix}\mathbf{v}_1^H\\ \vdots\\ \mathbf{v}_r^H\end{pmatrix}.$$

Then the following expression is obtained:

$$\boldsymbol{\Lambda}_1\mathbf{V}_{12}\mathbf{t} = \begin{pmatrix}\lambda_1\mathbf{v}_1^H\mathbf{t}\\ \vdots\\ \lambda_r\mathbf{v}_r^H\mathbf{t}\end{pmatrix}. \tag{41}$$

It follows that the CSNR can be written as

$$\begin{aligned}\rho &= \frac{1}{1 + \sigma_v^{-4}\dfrac{1}{K}\mathbf{t}^H\mathbf{V}_{12}^H\boldsymbol{\Lambda}_1\mathbf{V}_{12}\mathbf{t}}\\ &= \frac{1}{1 + \dfrac{1}{K}\zeta}\end{aligned} \tag{42}$$

where from (41), $\zeta = \sigma_v^{-4}\mathbf{t}^H\mathbf{V}_{12}^H\boldsymbol{\Lambda}_1\mathbf{V}_{12}\mathbf{t} = \sigma_v^{-4}\sum_{i=1}^r \lambda_i|\mathbf{v}_i^H\mathbf{t}|^2$. From Theorem 2, the terms $\mathbf{v}_i^H\mathbf{t}$ in the previous relation are complex Gaussian random variables, with $E[\mathbf{v}_i^H\mathbf{t}] = 0$ and variance $E[|\mathbf{v}_i^H\mathbf{t}|^2] = \mathbf{t}^H E[\mathbf{v}_i\mathbf{v}_i^H]\mathbf{t}$. To find an expression for the variance, we first evaluate the $p \times p$ covariance matrix $E[\mathbf{v}_i\mathbf{v}_i^H]$:

$$\begin{aligned}E[\mathbf{v}_i\mathbf{v}_i^H] &= E\begin{pmatrix}v_{i,r+1}v_{i,r+1}^* & \cdots & v_{i,r+1}v_{i,N}^*\\ \vdots & & \\ v_{i,N}v_{i,r+1}^* & & v_{i,N}v_{i,N}^*\end{pmatrix}\\ &= \frac{\lambda_i\sigma_v^2}{(\lambda_i - \sigma_v^2)^2}\mathbf{I}_p\end{aligned} \tag{43}$$

where the last relation is a direct consequence of Theorem 2. Using this relation we obtain

$$\begin{aligned}E[|\mathbf{v}_i^H\mathbf{t}|^2] &= \frac{\lambda_i\sigma_v^2}{(\lambda_i - \sigma_v^2)^2}\mathbf{t}^H\mathbf{t}\\ &\cong \sigma_v^4\frac{\lambda_i}{(\lambda_i - \sigma_v^2)^2}\end{aligned} \tag{44}$$

where we used $\mathbf{t}^H\mathbf{t} \cong \sigma_v^2$. Next we obtain the probability density of the Hermitian form $\mu_i = \mathbf{v}_i^H\mathbf{t}\mathbf{t}^H\mathbf{v}_i$. According to a lemma by Goodman [14], the characteristic function (CF) of the variate μ_i with respect to the density of $\mathbf{v}_i$ is given by

$$\Phi_{\mu_i}(\theta) = \frac{1}{\det(\mathbf{I}_p - j\theta\boldsymbol{\Sigma}_i\mathbf{t}\mathbf{t}^H)}, \qquad i = 1,\ldots,r \tag{45}$$

where $\mathbf{\Sigma}_i = E[\mathbf{v}_i\mathbf{v}_i^H]$. From (44), $\mathbf{\Sigma}_i = \lambda_i\sigma_v^2/(\lambda_i - \sigma_v^2)^2\mathbf{I}_p$. The determinant in the relation above can be computed by recognizing that the $p \times p$ matrix $(\mathbf{I}_p - j\theta\mathbf{\Sigma}_i\mathbf{t}\mathbf{t}^H)$ is the sum of an identity matrix and the rank-one matrix $j\theta\mathbf{\Sigma}_i\mathbf{t}\mathbf{t}^H = j\theta\lambda_i\sigma_v^2/(\lambda_i - \sigma_v^2)^2\mathbf{t}\mathbf{t}^H$. The only non-zero eigenvalue of the rank-one matrix is $j\theta\lambda_i\sigma_v^2/(\lambda_i - \sigma_v^2)^2\mathbf{t}^H\mathbf{t} = j\theta\lambda_i\sigma_v^4/(\lambda_i - \sigma_v^2)^2$. Consequently, the matrix $(\mathbf{I}_p - j\theta\mathbf{\Sigma}_i\mathbf{t}\mathbf{t}^H)$ has an eigenvalue equal to $(1 - j\theta\lambda_i\sigma_v^4/(\lambda_i - \sigma_v^2)^2)$ and $(p-1)$ unit eigenvalues. It follows that $\det(\mathbf{I}_p - j\theta\mathbf{\Sigma}_i\mathbf{t}\mathbf{t}^H) = 1 - j\theta\lambda_i\sigma_v^4/(\lambda_i - \sigma_v^2)^2$. From the last relation and (45) it follows that the CF of μ_i is given by

$$\Phi_{\mu_i}(\theta) = (1 - j\bar{\mu}_i\theta)^{-1} \tag{46}$$

where $\bar{\mu}_i = \lambda_i\sigma_v^4/(\lambda_i - \sigma_v^2)^2$. This CF is that of an exponential random variable (also a nonnormed chi-square variate with 2 degrees of freedom) and with mean $E[\mu_i] = \bar{\mu}_i$. The variate $\nu_i = \sigma_v^{-4}\lambda_i\mu_i$ then has an exponential distribution with mean

$$\bar{\nu}_i = E[\nu_i] = \lambda_i^2/(\lambda_i - \sigma_v^2)^2. \tag{47}$$

The random variable ζ in (42) can now be written

$$\zeta = \sum_{i=1}^{r} \nu_i. \tag{48}$$

The statistical independence of the ν_i variates follows from Theorem 2 and the independence of the vectors $\mathbf{v}_i$. Consequently, the CF of ζ is given by

$$\Phi_\zeta(\theta) = \prod_{i=1}^{r} \Phi_{\nu_i}(\theta) = \prod_{i=1}^{r}(1 - j\bar{\nu}_i\theta)^{-1}. \tag{49}$$

The product appearing in the expression of the CF can be converted to a sum by applying the following partial fraction expansion:

$$\prod_{i=1}^{r} \frac{1}{1 - j\bar{\nu}_i\theta} = \sum_{i=1}^{r} \frac{\pi_i}{1 - j\bar{\nu}_i\theta} \tag{50}$$

where $\pi_i = \prod_{j=1,j\neq i}^{r} \bar{\nu}_i/(\bar{\nu}_j - \bar{\nu}_i)$, [15]. The inverse Fourier transform of the CF $\Phi_\zeta(\theta)$ yields the density of the random variable ζ in the form of a sum of exponential densities weighted by the factors π_i:

$$f(\zeta) = \sum_{i=1}^{r} \frac{\pi_i}{\bar{\nu}_i} e^{-\zeta/\bar{\nu}_i}, \qquad \zeta \geq 0. \tag{51}$$

It follows that the density of the CSNR $\rho = (1 + (1/K)\zeta)^{-1}$ is given by the expression

$$f(\rho) = K\rho^{-2}\sum_{i=1}^{r} \frac{\pi_i}{\bar{\nu}_i} \exp\left(\frac{-K\left(\frac{1}{\rho} - 1\right)}{\bar{\nu}_i}\right), \qquad 0 \leq \rho \leq 1. \tag{52}$$

Discussion

The expression for the CSNR probability density in (52) merits further consideration. It is observed that the probability density depends on the number of samples K, the number of dominant modes r, and the eigenvalues of the true covariance matrix (through the quantities $\bar{\nu}_i$. Since the sum of the eigenvalues equals the total power in the received signal (interference + noise), it follows that the probability density is parameterized by the interference-to-noise ratio. This result is different in form, as well as in substance, from the distribution of the SMI method in [1]. That expression, provided below for reference, is that of a beta distribution and is independent of the covariance matrix:

$$f(\rho) = \frac{\Gamma(K+1)}{\Gamma(N-1)\Gamma(K+2-N)}(1-\rho)^{N-2}\rho^{K+1-N} \tag{53}$$

where $\Gamma(K+1) = K!$ is the gamma function. Another interesting comparison is with the density of the PCI method in [9]. As mentioned in the introduction, the eigencanceler and the PCI method of reference [6] provide the same weight vector in the case of a single linear constraint. The density of PCI's CSNR is similar in form to that of the SMI method, with the difference that in (53) the signal dimensionality N is replaced by the quantity $(r+1)$, [9], i.e.,

$$f(\rho) = \frac{\Gamma(K+1)}{\Gamma(r)\Gamma(K+1-r)}(1-\rho)^{r-1}\rho^{K-r}. \tag{54}$$

Like SMIs density, this density is independent of the covariance matrix and the CNR. An additional difference between the densities of (52) and (53) is the asymptotic nature of the former. A closer look at expressions where the asymptotic approximation is applied (such as (18)) reveals that the nature of the approximation is to neglect $o(1/\sqrt{K})$ terms relative to $o(1)$ terms. However, this expression is later squared (42), thus the asymptotic approximation implies neglecting terms $o(1/K)$ terms relative to $o(1)$. To achieve an order of magnitude ratio between the terms, the number of samples in data set needs to meet $K \geq 10$. Thus a modest number of samples is sufficient to satisfy the asymptotic approximation. Further analysis of the newly developed probability density, as well as comparisons with other densities and an assessment of the effect of the asymptotic approximations, are provided in the numerical results section.

Approximation for Large Interference-to-Noise Ratio

An approximation to the expression in (52) can be obtained for the case of a large interference-to-noise ratio, $\lambda_i \gg \sigma_v^2$. Earlier in this section it was shown that the quantity μ_i is an exponential variate (see (46)). For $\lambda_i \gg \sigma_v^2$, we have $\overline{\mu}_i = \lambda_i\sigma_v^4/(\lambda_i - \sigma_v^2)^2 \simeq \sigma_v^4/\lambda_i$. It follows that $\overline{v}_i = \sigma_v^{-4}\lambda_i\overline{\mu}_i = 1$. Then from (48), ζ is the sum of r IID, unit-mean, exponential variates; i.e., ζ has a gamma distribution with parameters r and 1. The density of ζ is then given by

$$f(\zeta) = \frac{1}{\Gamma(r)}\zeta^{r-1}e^{-\zeta}, \qquad \zeta \geq 0. \tag{55}$$

Using this density in conjunction with the relation between the variates ρ and ζ in (42) yields the density for the CSNR in the case of large interference-to-noise ratio:

$$f(\rho) = \frac{K^r e^K}{\Gamma(r)}e^{-K/\rho}(1-\rho)^{r-1}\rho^{-(r+1)}, \qquad 0 \leq \rho \leq 1. \tag{56}$$

It should be noted that as $\lambda_i/\sigma_v^2 \to \infty$, (52) $\to$ (56).

A further simplification can be obtained by keeping only the first two terms of the series $(1+(1/K)\zeta)^{-1} = \sum_{i=0}^{\infty}(-\zeta/K)^i$:

$$\rho = \frac{1}{1+\dfrac{1}{K}\zeta} \cong 1 - \frac{1}{K}\zeta. \tag{57}$$

The last formulation does not strictly guarantee that $0 \leq \rho$. However, since $E[\zeta] = r$ and $\mathrm{var}[\zeta] = (2r)^2$, then for $K \gg r$, ρ is almost certain to be nonnegative. The density of the CSNR is obtained from (55) and (57):

$$f(\rho) = \frac{K^r}{\Gamma(r)}e^{-K(1-\rho)}(1-\rho)^{r-1}, \qquad 0 \leq \rho \leq 1. \tag{58}$$

It should be noted that with the large interference-to-noise ratio approximation, the dependency on interference-to-noise ratio is not present in either (56) or (58). Finally, we note that if we define the CSNR loss, $\gamma = 1 - \rho$, then it follows that the probability density of the loss is given by the gamma function with parameters r and K:

$$f(\gamma) = \Gamma(r,K) = \frac{K^r}{\Gamma(r)}e^{-K\gamma}\gamma^{r-1}, \qquad 0 \leq \gamma \leq 1. \tag{59}$$

The CSNR loss is thus distributed as the incomplete gamma function.

IV. NUMERICAL RESULTS

Computer simulations were conducted to support the theory presented in the previous section. The simulation scenario consisted of distributed clutter, white Gaussian noise, and a space-time array. The clutter consisted of 120 point sources randomly distributed in the angular sector 0–20 deg with respect to the array boresight. The array boresight

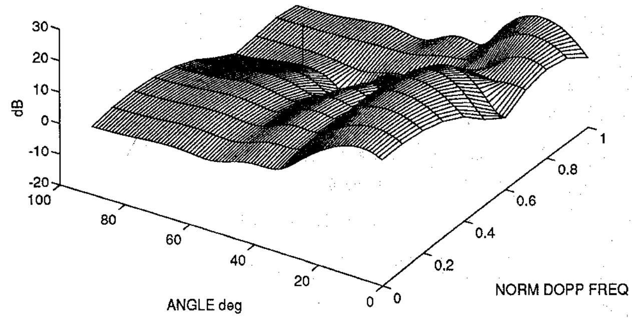

Fig. 1. Simulation clutter map. Vertical line indicates steering vector.

was assumed perpendicular to the platform motion. The clutter signals were summed noncoherently to form the signal received at the array. The space-time processor was fed by an $N_s = 8$ element linear array with $N_t = 4$ tap delay lines at each element, resulting in a signal space dimensionality of $N = N_sN_t = 32$. The steering vector was pointed at 50 deg and a normalized Doppler frequency of 0.4. The clutter map in angle-Doppler coordinates is shown in Fig. 1. For this scenario, the interference subspace was found to have rank $r = 4$. The space-time covariance matrix was estimated from a specified number of snapshots of those signals (see (5)), and weight vectors were derived for SMI and eigencanceler processors (see (4) for the latter). The CSNR was then evaluated from (9). Maintaining the same scenario, the process was then repeated numerous times to yield independent realizations of the CSNR variate. The numerical results were compared with the theory as described below.

Fig. 2 plots the mean CSNR for different clutter-to-noise ratios (CNR). The covariance matrix was estimated from $K = 2N = 2N_sN_t = 64$ snapshots. For each CNR, the mean was computed from 100 runs. The simulation results are shown together with values of $E[\rho]$ computed using numerical integration of the following densities: for the SMI, (53), and for the eigencanceler, (52), (56), and (58). Also shown are results obtained using the PCI density, (54). For this case (single linear constraint embodied by the steering vector), the PCI and eigencanceler methods produce the same weight vectors, and thus the same CSNR ρ. The objective is to evaluate the accuracy of the four probability density expressions for eigenanalysis-based processing (see (52), (56), (58), and the PCI density). From the figure it can be seen that for CNR $\geq$ 20 dB, all four expressions provide a good representation of the data. The approximation in (58) and the PCI method provide almost indistinguishable mean values.

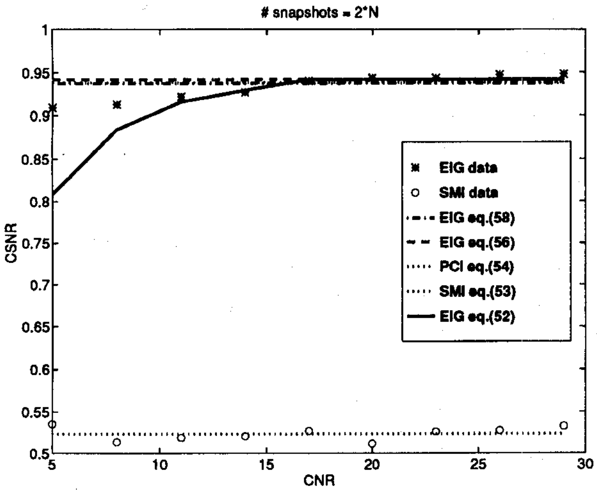

Fig. 2. Mean conditioned SNR as function of CNR for covariance matrix estimates using $K = 2N_sN_t$ samples.

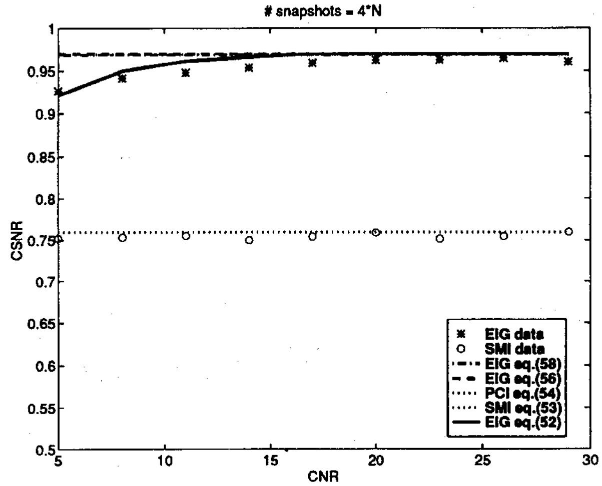

Fig. 4. Mean conditioned SNR as function of CNR for covariance matrix estimates using $K = 4N_sN_t$ samples.

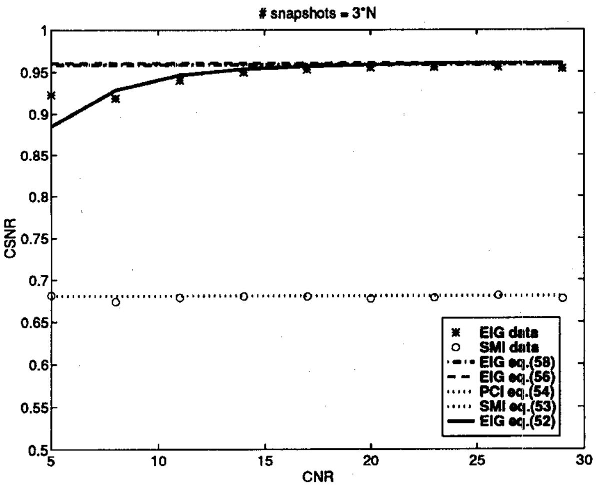

Fig. 3. Mean conditioned SNR as function of CNR for covariance matrix estimates using $K = 3N_sN_t$ samples.

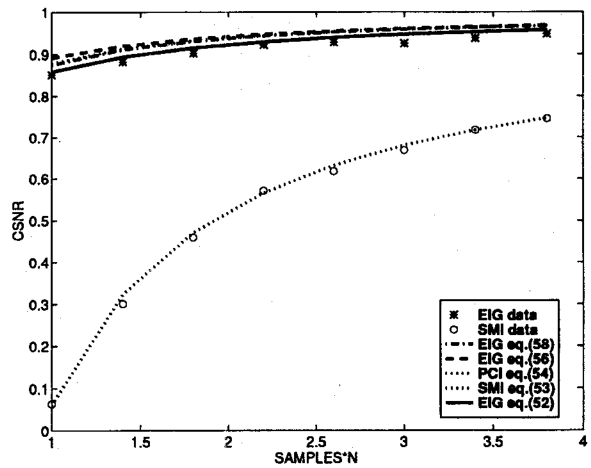

Fig. 5. Mean conditioned SNR as function of number of snapshots for CNR = 10 dB.

For large CNRs, the expectations yielded by (52) and (56) are also indistinguishable, as they should be. However, the most interesting observation is that for CNR $\leq$ 15 dB, the approximations in (56) and (58) as well as the PCI expression, fail to represent the data. Expression (52), however, provides a better fit to the data, at least for 10 dB $\leq$ CNR $\leq$ 15 dB. This observation supports the assertion that, unlike SMI processing, for eigenanalysis-based processing the CSNR is *not* independent of the covariance matrix. It is clearly evident in the figure that the CSNR of SMI-processed data is independent of the CNR, while the CSNR is CNR-dependent for the eigencanceler. It is also clear from the figure that the eigencanceler theory provides a lesser fit to the data for CNR $\leq$ 10 dB. That is due to the asymptotic nature of the expressions. Indeed, when the number of snapshots was increased to $K = 3N = 3N_sN_t$, Fig. 3 shows that the expectation using the new probability density in (52) represents the data down to CNR = 8 dB. Fig. 4 shows the result of a further increase in the number of samples to $K = 4N = 4N_sN_t$. The theory now provides a good fit to the data down to CNR = 5 dB.

As noted above, the asymptotic approximations seem to have little adverse effects at CNR $\geq$ 10. Further support to this can be found in Figs. 5 and 6. Each of the data points in the figure is an average of 100 Monte-Carlo runs. In Fig. 5 the mean CSNR is shown as a function of the number of snapshots K, at CNR = 10 dB. The new probability density provides an accurate fit to the data, down to a sample size of $K = N$ samples. The two approximations to the probability density, as well as the PCI density, show a slight bias. The SMI density provides a good fit as well. For $K < N$ the SMI cannot be applied, since the estimated covariance matrix $\hat{\mathbf{R}}$ becomes singular. Fig. 6 shows the mean CSNR of the eigencanceler only, as a function of the number of samples K, down to $K = N/16$. Again there is good agreement between the new probability density and the data. The

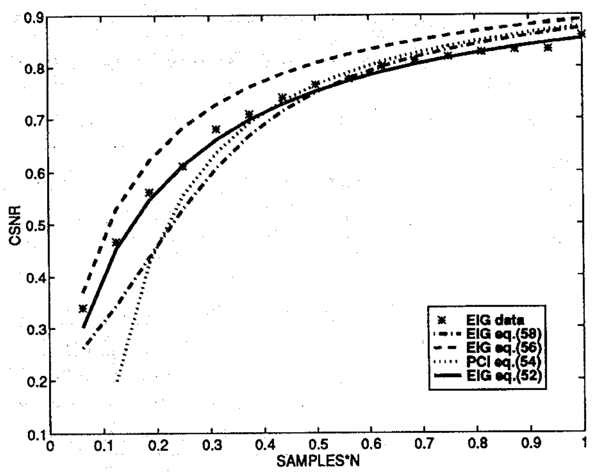

Fig. 6. Mean conditioned SNR as function of number of snapshots for CNR = 10 dB for low number of samples.

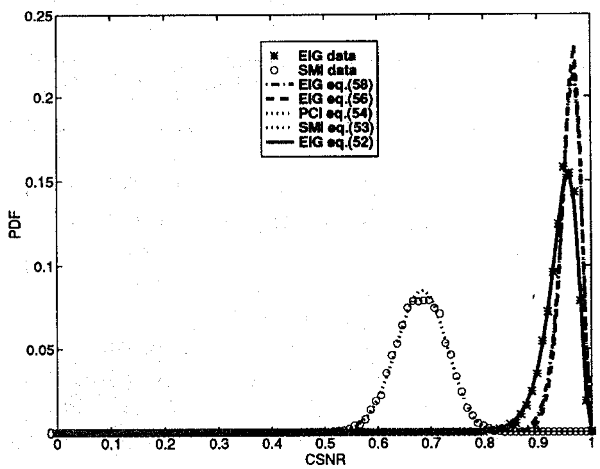

Fig. 7. Density functions of CSNR = 10 dB and $K = 4N$ snapshots.

approximation given by (56) maintains only a slight bias. The bias of approximation (58) is slightly larger, while the PCI density is more significantly off at very low number of samples. We can conclude from these plots that at CNR ≥ 10, (52) provides a very accurate description of the data down to a very low number of samples.

Fig. 7 provides the histograms of ρ for the SMI and the eigencanceler produced by 10,000 runs at CNR = 10 dB. The number of samples used to estimate each covariance matrix was $K = 4N$. A good match is observed between the probability density in (52) and the data. Likewise, the SMI density also provides an accurate representation of the data. The approximations to the eigencanceler's probability density as well as the PCI's provide a reasonable, though markedly less accurate, match to the data.

V. CONCLUSIONS

A new expression has been developed for the probability density function of the CSNR of eigenanalysis-based array processing. The new expression is derived from the asymptotic theory of the principal components of the estimated covariance matrix. It is shown that, unlike the case for the SMI method, the eigencanceler's CSNR probability density is *not* independent of the covariance matrix, for a wide range of interference-to-noise powers. Two simpler approximations are derived for the case of large interference-to-noise ratio. These approximations are shown to be independent of the covariance matrix and to provide a good fit to the data for interference-to-noise ratio ≥ 15 dB. The asymptotic nature of the expressions seems to have little effect for interference-to-noise ratio ≥ 15 dB. However, by increasing the number of samples to $K = 4N = 128$, the new density produces a good match to the data down to interference-to-noise ratio = 5 dB.

A. M. HAIMOVICH
Dept. of Electrical and Computer Engineering
New Jersey Institute of Technology
University Heights
Newark, NJ 07102

REFERENCES

[1] Reed, I. S., et al. (1974)
Rapid convergence rates in adaptive arrays.
IEEE Transactions on Aerospace and Electronic Systems, **AES-10** (Nov. 1974), 853–863.

[2] Boroson, D. M. (1980)
Sample size considerations for adaptive arrays.
IEEE Transactions on Aerospace and Electronic Systems, **AES-16** (July 1980), 446–451.

[3] Hanumara, R. C. (1986)
An alternate derivation of the distribution of the conditioned signal-to-noise ratio.
IEEE Transactions on Antennas Propagation, **AP-34** (Mar. 1986), 463–464.

[4] Khatri, C. G. (1987)
Effects of estimated noise covariance matrix in optimal signal detection.
IEEE Transactions on Acoustics, Speech, and Signal Processing, **ASSP-35** (May 1987), 671–679.

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

[6] Kirsteins, I. P., and Tufts, D. W. (1994)
Adaptive detection using low rank approximation to a data matrix.
IEEE Transactions on Aerospace and Electronic Systems, **30** (Jan. 1994), 55–67.

[7] Monzingo, R., and Miller, T. (1980)
An Introduction to Adaptive Arrays.
New York: Wiley, 1980.

[8] Haimovich, A. M., and Bar-Ness, Y. (1991)
An eigenanalysis interference canceler.
IEEE Transactions on Acoustics, Speech, and Signal Processing, **39** (Jan. 1991), 76–84.

[9] Kirsteins, I. P., and Tufts, D. W. (1991)
Rapidly adaptive nulling of interference.
In M. Bouvet and G. Bienvenu (Eds.), *High Resolution Methods in Underwater Acoustics.*
New York: Springer-Verlag, 1991, ch. 6, 220–249.

[10] Anderson, T. (1963)
Asymptotic theory for principal component analysis.
Annals of Mathematical Statistics, **34** (1963), 122–148.

[11] Jeffries, D. J. (1985)
Asymptotic results for eigenvector methods.
IEE Proceedings, Pt. F, **132** (Dec. 1985), 589–594.

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

[13] Orfanidis, S. (1993)
Optimum Signal Processing: An Introduction (2nd ed.).
New York: McGraw-Hill, 1993.

[14] Goodman, N. (1963)
Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction).
Annals of Mathematical Statistics, **34** (1963), 152–177.

[15] Proakis, J. (1995)
Digital Communications (3rd ed.).
New York: McGraw-Hill, 1995.

Blind Identification and Separation of Convolutively Mixed Independent Sources

A trispectra method for solving the m-input n-output ($n \geq m$) wideband blind identification and signal separation problem with unknown number of sources m is presented. The method is universal in the sense that it does not impose any restriction on the probability distribution of the input signals provided that they are non-Gaussian. A criterion, which states a sufficient condition for identification and separation, has been proved. An algorithm is also developed based on the criterion, whose efficiency is verified by the simulations.

I. INTRODUCTION

During recent years, there has been much interest focused on blind identification and blind source separation. The driving force behind this trend is that in the diverse fields from data communication to array signal processing, observations are made of the outputs of an unknown multiple-input multiple-output (MIMO) system. For example, in the environment with multiple speakers, multiple microphone measurements will typically have components from each speaker. The objective of blind identification is to identify the system and recover the source signals simultaneously, only utilizing the system outputs.

Manuscript received August 20, 1996.

IEEE Log No. T-AES/33/3/04765.

Many approaches have been proposed mainly for the special and simple case in which the MIMO system is linear memoryless. Most of them are based on the second- and/or fourth-order statistics of the sources [2–4, 9, 10]. Another kind of possible solutions is the neural network approaches [1, 11, 12]. However, in a variety of applications in which sources are convolutively mixed (known as wideband problem), e.g., speech dereverberation, speech enhancement in the presence of background noise and competing speakers separation, these approaches cannot be applied. The few works on this wideband aspect are restricted in the two-input two-output case with two channels assumed identity systems [5–7]. Reference [5] used a maximum likelihood (ML) estimator to separate a Gaussian autoregressive (AR) process and an uncorrelated white Gaussian noise. Reference [6] proposed a decorrelation criterion for uncorrelated source separation by imposing the uncorrelated condition on the reconstructed signals. However, the criterion doesn't specify a unique solution for the reconstruction system and is insufficient to solve the problem [6]. A symmetric adaptive decorrelation algorithm was presented in [7] with its stability, convergence, and solution uniqueness analyzed.

We propose a new type of trispectra method for the m-input n-output ($n \geq m$) wideband blind identification and signal separation problem with unknown number of sources m. The method is universal in the sense that it does not impose any restriction on the probability distribution of the input signals provided that they are non-Gaussian. A sufficient condition for identification and separation is proved and an algorithm for implementing the criterion is developed.

II. PROBLEM FORMULATION

Let m denote the number of sources and n that of sensors ($n \geq m$) with m unknown. The wideband formulation of the problem is written as

$$x_i(t) = \sum_{j=1}^{m} \sum_{k=0}^{K_i} a_{ij}^{(k)} s_j(t-k) + e_i(t) \qquad i = 1,\ldots,n$$

where $x_i(t)$ is the observation at the ith sensor, $a_{ij}^{(k)}$ is the kth coefficient of the unit sample response of the finite-duration impulse response (FIR) filter which characterizes the channel between the jth source and ith sensor. Here without loss of generality, we suppose $a_{1j}^{(0)} = 1$, $j = 1,\ldots,m$; $s_j(t)$ is the jth source, K_i is the assumed order of FIR filters associated with the ith sensor and $e_i(t)$ is measurement noise at the ith sensor.

Assumptions.

A1) $s_i(t)$ is a zero-mean wide-sense stationary process. For each t, $\{s_i(t),\ i = 1,\ldots,m\}$ are mutually independent.