SUBSPACE LEAKAGE ANALYSIS OF SAMPLE DATA COVARIANCE MATRIX

Mahdi Shaghaghi *

University of Alberta
Electrical and Computer Engineering
Edmonton, AB T6G 2V4, Canada
mahdi.shaghaghi@ualberta.ca

Sergiy A. Vorobyov

Aalto University
Dept. Signal Process. & Acoustics
Espoo, FI-00076 AALTO, Finland
svor@ieee.org

ABSTRACT

Subspace based methods provide a good compromise between performance and complexity. However, these methods are exposed to performance breakdown at the low SNR and/or small sample size region. It has been known for a long time that a major reason for such performance breakdown is the subspace swap phenomenon. However, in some scenarios such as the case of closely spaced sources, the breakdown happens before the subspace swap occurs. The reason is identified to be the intersubspace leakage where some portion of the true signal subspace resides in the estimated noise subspace. In this paper, we formally define the notion of subspace leakage which can be used as a measure for performance analysis and comparison of different methods used for estimating the signal and noise subspaces. We further study the statistical properties of the subspace leakage for the case of sample data covariance matrix.

***Index Terms*—** Eigenvalue decomposition, data covariance matrix, subspace leakage.

1. INTRODUCTION

Subspace methods have applications in array signal processing [1], [2], channel estimation [3], multiuser detection in communications [4], [5] and so on. Such methods are based on estimating the signal and noise subspaces from the singular value (or eigenvalue) decomposition of the data matrix (or the data covariance matrix).

The fidelity of the estimated subspaces to the true subspaces plays a critical role in a successful parameter estimation. At the low signal-to-noise ratio (SNR) and/or small sample size region, the estimated subspaces can largely deviate from the true ones, which can lead to performance breakdown. In most studies, the cause of such breakdown is considered to be the so-called *subspace swap* where the "measured data is better approximated by some components of the noise subspace than by the components of the signal subspace" [6].

It is shown in [7] that in some scenarios (e.g. closely spaced sources), as the SNR and/or the number of data samples is reduced, the performance breakdown can occur before the subspace swap happens. In [7], numerical examples are used to show that even a small portion of the estimated signal subspace energy residing in the true noise subspace (denoted by inter-subspace leakage) can result in performance breakdown. The notion of leakage comes originally from the performance assessment based on perturbed subspace estimation [8–12]. The results in [8] are asymptotic in the number of samples, and the results in [9–12] are asymptotic in the effective SNR. Here, we formally define the subspace leakage notion as a Frobenius norm of the perturbation matrix. Such a measure, can be used for performance analysis and comparison of different methods which are used for estimating the subspaces from measured data [13], [14].

The subspace leakage is by definition a random variable, and the analysis of its statistical properties can provide insights into the dynamics of the subspace estimation methods. In this paper, the expected value and the variance of the subspace leakage are studied. We consider the classical method of estimating the subspaces using the eigenvalue decomposition of the sample data covariance matrix, and we derive the corresponding subspace leakage and its mean and variance.

2. SYSTEM MODEL

Consider the system model where the vector of measured data $\boldsymbol{x}(t) \in \mathbb{C}^{M\times 1}$ at time instant $t \in \mathbb{N}$ is given by

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{n}(t) \tag{1}$$

where $\boldsymbol{A} \in \mathbb{C}^{M\times K}$ $(M > K)$ is a structured matrix with unknown parameters, $\boldsymbol{s}(t) \in \mathbb{C}^{K\times 1}$ is a vector containing unknown signal amplitudes and $\boldsymbol{n}(t) \in \mathbb{C}^{M\times 1}$ is a noise vector at time t.

We consider the noise vector $\boldsymbol{n}(t)$ to be independent from the signal vector at any time instant and the noise vector at other time instances. The noise vector also has the circularly-symmetric complex jointly-Gaussian distribution $\mathcal{N}_C(0, \sigma_{\mathrm{n}}^2 \boldsymbol{I}_M)$ where $\boldsymbol{I}_M$ is the identity matrix of size M.

*Supported in part by the Natural Science and Engineering Research Council (NSERC) of Canada.

Considering the system model (1), the data covariance matrix $\boldsymbol{R} \in \mathbb{C}^{M \times M}$ is given by

$$\boldsymbol{R} \triangleq E\left\{\boldsymbol{x}(t)\boldsymbol{x}^H(t)\right\} = \boldsymbol{ASA}^H + \sigma_{\mathrm{n}}^2 \boldsymbol{I}_M \tag{2}$$

where $\boldsymbol{S} = E\left\{\boldsymbol{s}(t)\boldsymbol{s}^H(t)\right\} \in \mathbb{C}^{K \times K}$ is the signal covariance matrix (assumed to be full rank) and $(\cdot)^H$ and $E\{\cdot\}$ stand for the Hermitian transposition and the expectation operators, respectively.

Let N number of snapshots (samples) be available. The basic method for estimating the data covariance matrix from the samples $\boldsymbol{x}(t)$ $(1 \leq t \leq N)$ is the following sample estimate

$$\widehat{\boldsymbol{R}} \triangleq \frac{1}{N}\sum_{t=1}^{N} \boldsymbol{x}(t)\boldsymbol{x}^H(t) \tag{3}$$

where $\widehat{\boldsymbol{R}} \in \mathbb{C}^{M \times M}$ is called the sample data covariance matrix.

Let $\hat{\lambda}_1 \leq \hat{\lambda}_2 \leq \cdots \leq \hat{\lambda}_M$ be the eigenvalues of $\widehat{\boldsymbol{R}}$ arranged in nondecreasing order, and let $\hat{\boldsymbol{g}}_1, \hat{\boldsymbol{g}}_2, \cdots, \hat{\boldsymbol{g}}_{M-K}$ be the orthonormal noise eigenvectors associated with $\hat{\lambda}_1, \hat{\lambda}_2, \cdots, \hat{\lambda}_{M-K}$ and $\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2, \cdots, \hat{\boldsymbol{e}}_K$ be the orthonormal signal eigenvectors corresponding to $\hat{\lambda}_{M-K+1}, \hat{\lambda}_{M-K+2}, \cdots, \hat{\lambda}_M$. Let also $\widehat{\boldsymbol{G}} \in \mathbb{C}^{M \times (M-K)}$ and $\widehat{\boldsymbol{E}} \in \mathbb{C}^{M \times K}$ be defined as $\widehat{\boldsymbol{G}} \triangleq \left[\hat{\boldsymbol{g}}_1, \hat{\boldsymbol{g}}_2, \cdots, \hat{\boldsymbol{g}}_{M-K}\right]$ and $\widehat{\boldsymbol{E}} \triangleq [\hat{\boldsymbol{e}}_1, \hat{\boldsymbol{e}}_2, \cdots, \hat{\boldsymbol{e}}_K]$. The range spaces of $\widehat{\boldsymbol{G}}$ and $\widehat{\boldsymbol{E}}$ represent the estimations of the noise and signal subspaces, respectively.

3. SUBSPACE LEAKAGE

Consider the eigendecomposition of the data covariance matrix $\boldsymbol{R}$. Form $\boldsymbol{G} \in \mathbb{C}^{M \times (M-K)}$ and $\boldsymbol{E} \in \mathbb{C}^{M \times K}$ by placing the noise and signal eigenvectors as the columns of $\boldsymbol{G}$ and $\boldsymbol{E}$, respectively. The range spaces of $\boldsymbol{G}$ and $\boldsymbol{E}$ represent the true noise and signal subspaces. Note that the matrix of the eigenvectors $\boldsymbol{Q}_R = [\boldsymbol{G}\ \boldsymbol{E}] \in \mathbb{C}^{M \times M}$ is a unitary matrix $\left(\boldsymbol{Q}_R\boldsymbol{Q}_R^H = \boldsymbol{I}_M\right)$. Therefore, $\boldsymbol{GG}^H + \boldsymbol{EE}^H = \boldsymbol{I}_M$ or

$$\boldsymbol{P}^{\perp} + \boldsymbol{P} = \boldsymbol{I}_M \tag{4}$$

where $\boldsymbol{P}^{\perp} \triangleq \boldsymbol{GG}^H$ and $\boldsymbol{P} \triangleq \boldsymbol{EE}^H$ are the true projection matrices into the noise and signal subspaces, respectively.

Ideally, the estimation of each signal eigenvector $\hat{\boldsymbol{e}}_k$ $(1 \leq k \leq K)$ would perfectly fall in the true signal subspace. In practice, however, the energy of the projection of $\hat{\boldsymbol{e}}_k$ into the noise subspace $\left(\|\boldsymbol{P}^{\perp}\hat{\boldsymbol{e}}_k\|_2^2\right)$ is almost surely nonzero, which can be viewed as the leakage of $\hat{\boldsymbol{e}}_k$ into the true noise subspace. Here, we define the subspace leakage as the average value of the energy of the estimated signal eigenvectors leaked into the true noise subspace, i.e.,

$$\rho \triangleq \frac{1}{K}\sum_{k=1}^{K} \|\boldsymbol{P}^{\perp}\hat{\boldsymbol{e}}_k\|_2^2. \tag{5}$$

Note that $\boldsymbol{P}^{\perp}$ is a Hermitian matrix and $\left(\boldsymbol{P}^{\perp}\right)^2 = \boldsymbol{P}^{\perp}$. Therefore, ρ can be written as

$$\rho = \frac{1}{K}\sum_{k=1}^{K} \hat{\boldsymbol{e}}_k^H \boldsymbol{P}^{\perp} \hat{\boldsymbol{e}}_k. \tag{6}$$

Using (4), the expression (6) can be further simplified as

$$\begin{aligned}
\rho &= \frac{1}{K}\sum_{k=1}^{K} \hat{\boldsymbol{e}}_k^H \left(\boldsymbol{I}_M - \boldsymbol{P}\right)\hat{\boldsymbol{e}}_k = \frac{1}{K}\left(K - \sum_{k=1}^{K}\hat{\boldsymbol{e}}_k^H \boldsymbol{P}\hat{\boldsymbol{e}}_k\right) \\
&= 1 - \frac{1}{K}\sum_{k=1}^{K} \mathrm{Tr}\left\{\hat{\boldsymbol{e}}_k\hat{\boldsymbol{e}}_k^H\boldsymbol{P}\right\} \\
&= 1 - \frac{1}{K}\mathrm{Tr}\left\{\widehat{\boldsymbol{E}}\widehat{\boldsymbol{E}}^H\boldsymbol{P}\right\} = 1 - \frac{1}{K}\mathrm{Tr}\left\{\widehat{\boldsymbol{P}}\boldsymbol{P}\right\}
\end{aligned} \tag{7}$$

where $\widehat{\boldsymbol{P}} \triangleq \widehat{\boldsymbol{E}}\widehat{\boldsymbol{E}}^H$ is the estimated signal projection matrix.

Let $\Delta\boldsymbol{P} \triangleq \widehat{\boldsymbol{P}} - \boldsymbol{P}$ be the estimation error of the signal projection matrix. Then, using the properties that $\boldsymbol{P}^2 = \boldsymbol{P}$ and $\mathrm{Tr}\{\boldsymbol{P}\} = K$, the expression (7) can be rewritten as

$$\rho = 1 - \frac{1}{K}\mathrm{Tr}\left\{\left(\boldsymbol{P} + \Delta\boldsymbol{P}\right)\boldsymbol{P}\right\} = -\frac{1}{K}\mathrm{Tr}\left\{\Delta\boldsymbol{P}\boldsymbol{P}\right\}. \tag{8}$$

Now, let $\Delta\boldsymbol{R} \triangleq \widehat{\boldsymbol{R}} - \boldsymbol{R}$ be the estimation error of the covariance matrix. Define also the following matrix

$$\boldsymbol{V} \triangleq \boldsymbol{R} - \sigma_{\mathrm{n}}^2\boldsymbol{I}_M = \boldsymbol{ASA}^H = \sum_{k=1}^{K}\left(\lambda_{M-K+k} - \sigma_{\mathrm{n}}^2\right)\boldsymbol{e}_k\boldsymbol{e}_k^H \tag{9}$$

and let $\boldsymbol{V}^{\dagger} \in \mathbb{C}^{M \times M}$ denote the pseudo-inverse of $\boldsymbol{V}$ given by

$$\boldsymbol{V}^{\dagger} = \sum_{k=1}^{K}\frac{1}{\lambda_{M-K+k} - \sigma_{\mathrm{n}}^2}\boldsymbol{e}_k\boldsymbol{e}_k^H \tag{10}$$

where $\lambda_{M-K+1} \leq \lambda_{M-K+2} \leq \cdots \leq \lambda_M$ are the K largest eigenvalues of $\boldsymbol{R}$ and $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_K$ are their corresponding eigenvectors. It is shown in [15] that the series expansion of $\widehat{\boldsymbol{P}}$ based on $\Delta\boldsymbol{R}$ is given by

$$\widehat{\boldsymbol{P}} = \boldsymbol{P} + \delta\boldsymbol{P} + \cdots + \delta^n\boldsymbol{P} + \cdots \tag{11}$$

where

$$\delta\boldsymbol{P} = \boldsymbol{P}^{\perp}\Delta\boldsymbol{R}\boldsymbol{V}^{\dagger} + \boldsymbol{V}^{\dagger}\Delta\boldsymbol{R}\boldsymbol{P}^{\perp} \tag{12}$$

and the rest of the terms are related by the following recurrence

$$\begin{aligned}
\delta^n\boldsymbol{P} = &-\boldsymbol{P}^{\perp}\left(\delta^{n-1}\boldsymbol{P}\right)\Delta\boldsymbol{R}\boldsymbol{V}^{\dagger} + \boldsymbol{P}^{\perp}\Delta\boldsymbol{R}\left(\delta^{n-1}\boldsymbol{P}\right)\boldsymbol{V}^{\dagger} \\
&-\boldsymbol{V}^{\dagger}\Delta\boldsymbol{R}\left(\delta^{n-1}\boldsymbol{P}\right)\boldsymbol{P}^{\perp} + \boldsymbol{V}^{\dagger}\left(\delta^{n-1}\boldsymbol{P}\right)\Delta\boldsymbol{R}\boldsymbol{P}^{\perp} \\
&-\sum_{i=1}^{n-1}\boldsymbol{P}\left(\delta^{i}\boldsymbol{P}\right)\left(\delta^{n-i}\boldsymbol{P}\right)\boldsymbol{P} \\
&+\sum_{i=1}^{n-1}\boldsymbol{P}^{\perp}\left(\delta^{i}\boldsymbol{P}\right)\left(\delta^{n-i}\boldsymbol{P}\right)\boldsymbol{P}^{\perp}.
\end{aligned} \tag{13}$$

The following lemma regarding the columns of $\boldsymbol{V}^{\dagger}$ is in order.

Lemma 1. *The columns of $\boldsymbol{V}^\dagger$ belong to the signal subspace, i.e., $\boldsymbol{P}\boldsymbol{V}^\dagger = \boldsymbol{V}^\dagger$.*

Proof. The following train of equalities is valid

$$\begin{aligned}\boldsymbol{P}\boldsymbol{V}^\dagger &= \boldsymbol{E}\boldsymbol{E}^H \sum_{k=1}^{K} \frac{1}{\lambda_{M-K+k} - \sigma_{\text{n}}^2} \boldsymbol{e}_k \boldsymbol{e}_k^H \\ &= \sum_{i=1}^{K} \boldsymbol{e}_i \boldsymbol{e}_i^H \sum_{k=1}^{K} \frac{1}{\lambda_{M-K+k} - \sigma_{\text{n}}^2} \boldsymbol{e}_k \boldsymbol{e}_k^H \\ &= \sum_{k=1}^{K} \frac{1}{\lambda_{M-K+k} - \sigma_{\text{n}}^2} \boldsymbol{e}_k \boldsymbol{e}_k^H = \boldsymbol{V}^\dagger. \qquad (14)\end{aligned}$$

In the last step, we used the fact that $\boldsymbol{e}_i^H \boldsymbol{e}_k$ is equal to 1 for $i = k$ and it equals zero otherwise. □

Using (8) together with the series expansion of $\widehat{\boldsymbol{P}}$ in (11), (12), and (13) up to the $\delta^2 \boldsymbol{P}$ term, the facts that $\boldsymbol{P}\boldsymbol{P}^\perp = \boldsymbol{P}^\perp \boldsymbol{P} = \boldsymbol{0}$, $\left(\boldsymbol{P}^\perp\right)^2 = \boldsymbol{P}^\perp$, $(\boldsymbol{P})^2 = \boldsymbol{P}$, and Lemma 1, we can compute ρ as

$$\begin{aligned}\rho &\approx -\frac{1}{K}\text{Tr}\left\{-\boldsymbol{P}\left(\delta\boldsymbol{P}\right)\left(\delta\boldsymbol{P}\right)\right\} \\ &= \frac{1}{K}\text{Tr}\Big\{\boldsymbol{P}\left(\boldsymbol{P}^\perp \Delta\boldsymbol{R}\boldsymbol{V}^\dagger + \boldsymbol{V}^\dagger \Delta\boldsymbol{R}\boldsymbol{P}^\perp\right) \\ &\quad \times \left(\boldsymbol{P}^\perp \Delta\boldsymbol{R}\boldsymbol{V}^\dagger + \boldsymbol{V}^\dagger \Delta\boldsymbol{R}\boldsymbol{P}^\perp\right)\Big\} \\ &= \frac{1}{K}\text{Tr}\left\{\boldsymbol{P}\boldsymbol{V}^\dagger \Delta\boldsymbol{R}\boldsymbol{P}^\perp \boldsymbol{P}^\perp \Delta\boldsymbol{R}\boldsymbol{V}^\dagger\right\} \\ &= \frac{1}{K}\text{Tr}\left\{\boldsymbol{V}^\dagger \Delta\boldsymbol{R}\boldsymbol{P}^\perp \Delta\boldsymbol{R}\boldsymbol{V}^\dagger\right\}. \qquad (15)\end{aligned}$$

3.1. Expected value of subspace leakage

Computation of the expected value of the subspace leakage requires considering the statistical properties of $\Delta\boldsymbol{R}$. We use the following two properties in our derivations [15].

Lemma 2. *For all matrices $\boldsymbol{A}_1,\ \boldsymbol{A}_2 \in \mathbb{C}^{M\times M}$, we have*

$$E\left\{\Delta\boldsymbol{R}\boldsymbol{A}_1\Delta\boldsymbol{R}\right\} = \frac{1}{N}\mathit{Tr}\left\{\boldsymbol{R}\boldsymbol{A}_1\right\}\boldsymbol{R} \qquad (16)$$

and

$$E\left\{\mathit{Tr}\left\{\Delta\boldsymbol{R}\boldsymbol{A}_1\right\}\mathit{Tr}\left\{\Delta\boldsymbol{R}\boldsymbol{A}_2\right\}\right\} = \frac{1}{N}\mathit{Tr}\left\{\boldsymbol{R}\boldsymbol{A}_1\boldsymbol{R}\boldsymbol{A}_2\right\}. \qquad (17)$$

Using (15) and (16), $E\left\{\rho\right\}$ can be computed as

$$\begin{aligned}E\left\{\rho\right\} &\approx \frac{1}{K}\text{Tr}\left\{\boldsymbol{V}^\dagger E\left\{\Delta\boldsymbol{R}\boldsymbol{P}^\perp\Delta\boldsymbol{R}\right\}\boldsymbol{V}^\dagger\right\} \\ &= \frac{1}{K}\text{Tr}\left\{\boldsymbol{V}^\dagger \frac{1}{N}\text{Tr}\left\{\boldsymbol{R}\boldsymbol{P}^\perp\right\}\boldsymbol{R}\boldsymbol{V}^\dagger\right\} \\ &= \frac{1}{NK}\text{Tr}\left\{\boldsymbol{P}^\perp\boldsymbol{R}\right\}\text{Tr}\left\{\boldsymbol{V}^\dagger\boldsymbol{V}^\dagger\boldsymbol{R}\right\}. \qquad (18)\end{aligned}$$

Since the range space of the matrix $\boldsymbol{A}$ is the same as the signal subspace, we have $\boldsymbol{P}^\perp\boldsymbol{A} = \boldsymbol{0}$. As a result, $\text{Tr}\left\{\boldsymbol{P}^\perp\boldsymbol{R}\right\}$ can be simplified as

$$\begin{aligned}\text{Tr}\left\{\boldsymbol{P}^\perp\boldsymbol{R}\right\} &= \text{Tr}\left\{\boldsymbol{P}^\perp\left(\boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^H + \sigma_{\text{n}}^2\boldsymbol{I}_M\right)\right\} = \text{Tr}\left\{\sigma_{\text{n}}^2\boldsymbol{P}^\perp\right\} \\ &= \sigma_{\text{n}}^2\text{Tr}\left\{\boldsymbol{I}_M - \boldsymbol{P}\right\} = \sigma_{\text{n}}^2\left(M-K\right). \qquad (19)\end{aligned}$$

Furthermore, using (10) and the fact that the eigenvectors of $\boldsymbol{R}$ are orthonormal, $\boldsymbol{V}^\dagger\boldsymbol{V}^\dagger\boldsymbol{R}$ can be written as

$$\boldsymbol{V}^\dagger\boldsymbol{V}^\dagger\boldsymbol{R} = \sum_{k=1}^{K}\frac{\lambda_{M-K+k}}{\left(\lambda_{M-K+k} - \sigma_{\text{n}}^2\right)^2}\boldsymbol{e}_k\boldsymbol{e}_k^H \qquad (20)$$

which results in

$$\text{Tr}\left\{\boldsymbol{V}^\dagger\boldsymbol{V}^\dagger\boldsymbol{R}\right\} = \sum_{k=1}^{K}\frac{\lambda_{M-K+k}}{\left(\lambda_{M-K+k} - \sigma_{\text{n}}^2\right)^2}. \qquad (21)$$

Finally, $E\left\{\rho\right\}$ is obtained by substituting (19) and (21) in (18) as

$$E\left\{\rho\right\} \approx \frac{\sigma_{\text{n}}^2\left(M-K\right)}{NK}C \qquad (22)$$

where

$$C = \sum_{k=1}^{K}\frac{\lambda_{M-K+k}}{\left(\lambda_{M-K+k} - \sigma_{\text{n}}^2\right)^2}. \qquad (23)$$

Remark. We used the recurrence (13) up to the $\delta^2\boldsymbol{P}$ term while computing ρ in (15). Higher order terms can also be used to estimate ρ with higher accuracy. However, this unnecessary complication would only give a more accurate estimate of the coefficient C in (22). It is of much higher significance however that $E\left\{\rho\right\}$ is inversely proportional to N and proportional to σ_{n}^2.

3.2. Variance of subspace leakage

The variance of the subspace leakage is given by

$$\text{Var}(\rho) = E\left\{\rho^2\right\} - \left[E\left\{\rho\right\}\right]^2. \qquad (24)$$

Here, we show that $\text{Var}(\rho)$ is in the order of $1/N^2$.

Using (15) and $\Delta\boldsymbol{R} = \widehat{\boldsymbol{R}} - \boldsymbol{R}$, $E\left\{\rho^2\right\}$ can be computed as

$$\begin{aligned}E\left\{\rho^2\right\} &= \frac{1}{K^2} \\ &\quad \times E\left\{\left[\text{Tr}\left\{\left(\widehat{\boldsymbol{R}} - \boldsymbol{R}\right)\boldsymbol{P}^\perp\left(\widehat{\boldsymbol{R}} - \boldsymbol{R}\right)\boldsymbol{V}^\dagger\boldsymbol{V}^\dagger\right\}\right]^2\right\}. \qquad (25)\end{aligned}$$

It was shown in (19) that $\boldsymbol{P}^\perp\boldsymbol{R} = \sigma_{\text{n}}^2\boldsymbol{P}^\perp$ which is also equal to $\boldsymbol{R}\boldsymbol{P}^\perp$. By using this fact and expanding the terms in (25), $E\left\{\rho^2\right\}$ can be written as

$$\begin{aligned}E\left\{\rho^2\right\} &= \frac{1}{K^2} \\ &\times E\Bigg\{\Bigg[\text{Tr}\left\{\widehat{\boldsymbol{R}}\boldsymbol{P}^\perp\widehat{\boldsymbol{R}}\boldsymbol{V}^\dagger\boldsymbol{V}^\dagger\right\} - \sigma_{\text{n}}^2\text{Tr}\left\{\widehat{\boldsymbol{R}}\boldsymbol{P}^\perp\boldsymbol{V}^\dagger\boldsymbol{V}^\dagger\right\} \\ &\quad -\sigma_{\text{n}}^2\text{Tr}\left\{\boldsymbol{P}^\perp\widehat{\boldsymbol{R}}\boldsymbol{V}^\dagger\boldsymbol{V}^\dagger\right\} + \sigma_{\text{n}}^4\text{Tr}\left\{\boldsymbol{P}^\perp\boldsymbol{V}^\dagger\boldsymbol{V}^\dagger\right\}\Bigg]^2\Bigg\}. \qquad (26)\end{aligned}$$

From Lemma 1, we know that the columns of $\boldsymbol{V}^{\dagger}$ belong to the signal subspace and therefore $\boldsymbol{P}^{\perp}\boldsymbol{V}^{\dagger} = \boldsymbol{V}^{\dagger}\boldsymbol{P}^{\perp} = \boldsymbol{0}$. As a result, all the terms in (26) except for the first term are equal to zero. Therefore, $E\left\{\rho^2\right\}$ is given by

$$E\left\{\rho^2\right\} = \frac{1}{K^2}E\left\{\left[\mathrm{Tr}\left\{\widehat{\boldsymbol{R}}\boldsymbol{P}^{\perp}\widehat{\boldsymbol{R}}\boldsymbol{V}^{\dagger}\boldsymbol{V}^{\dagger}\right\}\right]^2\right\}. \tag{27}$$

The following lemma is used to proceed with the computation of $E\left\{\rho^2\right\}$. For details about this property see [16].

Lemma 3. *For all matrices $\boldsymbol{A}_1,\ \boldsymbol{A}_2 \in \mathbb{C}^{M\times M}$, we have*

$$\begin{aligned}
&E\left\{\left[Tr\left\{\widehat{\boldsymbol{R}}\boldsymbol{A}_1\widehat{\boldsymbol{R}}\boldsymbol{A}_2\right\}\right]^2\right\} = \\
&\quad\left[Tr\left\{\boldsymbol{R}\boldsymbol{A}_1\boldsymbol{R}\boldsymbol{A}_2\right\}\right]^2 + \frac{2}{N}\Big\{Tr\left\{\boldsymbol{R}\boldsymbol{A}_1\boldsymbol{R}\boldsymbol{A}_2\boldsymbol{R}\boldsymbol{A}_1\boldsymbol{R}\boldsymbol{A}_2\right\} \\
&\quad+Tr\left\{\boldsymbol{R}\boldsymbol{A}_1\boldsymbol{R}\boldsymbol{A}_1\boldsymbol{R}\boldsymbol{A}_2\boldsymbol{R}\boldsymbol{A}_2\right\} + Tr\left\{\boldsymbol{R}\boldsymbol{A}_1\boldsymbol{R}\boldsymbol{A}_2\right\} \\
&\quad\times Tr\left\{\boldsymbol{R}\boldsymbol{A}_1\right\} Tr\left\{\boldsymbol{R}\boldsymbol{A}_2\right\}\Big\} + \mathcal{O}\left(\frac{1}{N^2}\right).
\end{aligned} \tag{28}$$

Using Lemma 3, $E\left\{\rho^2\right\}$ can be further computed as

$$\begin{aligned}
E\left\{\rho^2\right\} = \frac{1}{K^2}\Bigg\{&\left[\mathrm{Tr}\left\{\boldsymbol{R}\boldsymbol{P}^{\perp}\boldsymbol{R}\boldsymbol{V}^{\dagger}\boldsymbol{V}^{\dagger}\right\}\right]^2 \\
&+\frac{2}{N}\Big\{\mathrm{Tr}\left\{\boldsymbol{R}\boldsymbol{P}^{\perp}\boldsymbol{R}\boldsymbol{V}^{\dagger}\boldsymbol{V}^{\dagger}\boldsymbol{R}\boldsymbol{P}^{\perp}\boldsymbol{R}\boldsymbol{V}^{\dagger}\boldsymbol{V}^{\dagger}\right\} \\
&\quad+\mathrm{Tr}\left\{\boldsymbol{R}\boldsymbol{P}^{\perp}\boldsymbol{R}\boldsymbol{P}^{\perp}\boldsymbol{R}\boldsymbol{V}^{\dagger}\boldsymbol{V}^{\dagger}\boldsymbol{R}\boldsymbol{V}^{\dagger}\boldsymbol{V}^{\dagger}\right\} \\
&\quad+\mathrm{Tr}\left\{\boldsymbol{R}\boldsymbol{P}^{\perp}\boldsymbol{R}\boldsymbol{V}^{\dagger}\boldsymbol{V}^{\dagger}\right\} \\
&\quad\times\mathrm{Tr}\left\{\boldsymbol{R}\boldsymbol{P}^{\perp}\right\}\mathrm{Tr}\left\{\boldsymbol{R}\boldsymbol{V}^{\dagger}\boldsymbol{V}^{\dagger}\right\} + \mathcal{O}\left(\frac{1}{N^2}\right)\Big\}.
\end{aligned} \tag{29}$$

Finally, using the facts that $\boldsymbol{R}\boldsymbol{P}^{\perp} = \boldsymbol{P}^{\perp}\boldsymbol{R} = \sigma_{\mathrm{n}}^2\boldsymbol{P}^{\perp}$ and $\boldsymbol{P}^{\perp}\boldsymbol{V}^{\dagger} = \boldsymbol{V}^{\dagger}\boldsymbol{P}^{\perp} = \boldsymbol{0}$, it is concluded that

$$E\left\{\rho^2\right\} = \mathcal{O}\left(\frac{1}{N^2}\right) \tag{30}$$

and from (22) and (24), $\mathrm{Var}(\rho)$ is given by

$$\mathrm{Var}(\rho) = \mathcal{O}\left(\frac{1}{N^2}\right) - \frac{\sigma_{\mathrm{n}}^4\left(M-K\right)^2}{N^2K^2}C^2 = \mathcal{O}\left(\frac{1}{N^2}\right). \tag{31}$$

4. NUMERICAL EXAMPLE

Consider the example of direction-of-arrival (DOA) estimation using a subspace based method. Let $K = 2$ sources be impinging on a uniform linear array (ULA) with $M = 10$ antenna elements from directions $\theta_1 = 35^\circ$ and $\theta_2 = 37^\circ$. The interelement spacing is set to half a wavelength and the number of snapshots is $N = 10$. The sources $\boldsymbol{s}(t)$ are considered to be independent to each other in time and to have the circularly-symmetric complex jointly-Gaussian distribution $\mathcal{N}_C(0, \sigma_{\mathrm{s}}^2\boldsymbol{I}_2)$ where σ_{s}^2 is the signal power. The SNR is defined as $\mathrm{SNR} \triangleq 10\log_{10}\left(\sigma_{\mathrm{s}}^2/\sigma_{\mathrm{n}}^2\right)$.

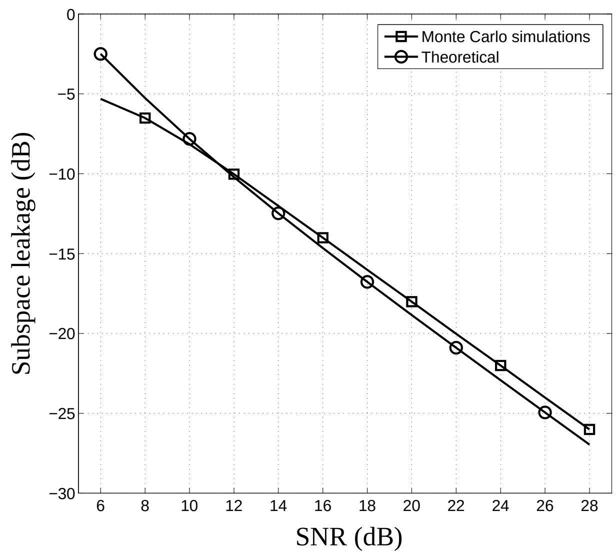

Fig. 1. Expected value of subspace leakage versus SNR.

The expected value of the subspace leakage is estimated using (7) and the Monte Carlo simulations with 10^5 number of trials. The theoretical value for the expected subspace leakage is also obtained from (22) and (23). The results are shown in Fig. 1. It can be seen that the curves obtained from the simulations are very close to those obtained from our theoretical derivations. The difference between the curves is due to the fact that in our derivations, we used up to the second order term of the series expansion of the signal projection matrix. For very low SNR values, the subspace swap phenomenon occurs where one or more of the noise eigenvectors are mistakenly taken for the signal eigenvectors and used for estimating the subspace leakage. Consequently, the the curve obtained by simulations is deviated from the curve obtained by theoretical derivations.

Monte Carlo simulations for the variance of the subspace leakage have been presented in [17]. The results could not be included here due to the lack of space.

5. CONCLUSION

In this paper, the subspace leakage as a cause for performance breakdown of the subspace based methods has been introduced. In the classical case of estimating the signal and noise subspaces from the sample data covariance matrix, we have computed the closed-form expression for the subspace leakage. The expected value and the variance of the subspace leakage has been also studied, which presented the dependance of the subspace leakage on the SNR, sample size, and other parameters. Finally, a numerical example has been given to compare the theoretical derivations with Monte Carlo simulations.

6. REFERENCES

[1] R. O. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Trans. Antennas Propagat.*, vol. AP-34, no. 3, pp. 276–280, Mar. 1986.

[2] R. Roy and T. Kailath, “ESPRIT–estimation of signal parameters via rotational invariance techniques,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 7, pp. 984–995, Jul. 1989.

[3] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, “Subspace methods for the blind identification of multichannel FIR filters,” *IEEE Trans. Signal Process.*, vol. 43, no. 2, pp. 516–525, Feb. 1995.

[4] S. Bensley and B. Aazhang, “Subspace-based channel estimation for code division multiple access communication systems,” *IEEE Trans. Commun.*, vol. 44, no. 8, pp. 1009–1020, Aug. 1996.

[5] X. Wang and H. V. Poor, “Blind multiuser detection: A subspace approach,” *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 677–690, Nov. 1998.

[6] J. Thomas, L. Scharf, and D. Tufts, “The probability of a subspace swap in the SVD,” *IEEE Trans. Signal Process.*, vol. 43, no. 3, pp. 730–736, Mar. 1995.

[7] B. A. Johnson, Y. I. Abramovich, and X. Mestre, “MUSIC, G-MUSIC, and maximum-likelihood performance breakdown,” *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3944–3958, Aug. 2008.

[8] B. D. Rao and K. V. S. Hari, “Performance analysis of ESPRIT and TAM in determining the direction of arrival of plane waves in noise,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 12, pp. 1990–1995, Dec. 1989.

[9] F. Li, H. Liu, and R. J. Vaccaro, “Performance analysis for DOA estimation algorithms: Unification, simplification, and observations,” *IEEE Trans. Aerosp., Electron. Syst.*, vol. 29, no. 4, pp. 1170–1184, Oct. 1993.

[10] Z. Xu, “Perturbation analysis for subspace decomposition with applications in subspace-based algorithms,” *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2820–2830, Nov. 2002.

[11] J. Liu, X. Liu, and X. Ma, “First-order perturbation analysis of singular vectors in singular value decomposition,” *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3044–3049, Jul. 2008.

[12] J. Steinwandt, F. Roemer, M. Haardt, and G. D. Galdo, “R-dimensional ESPRIT-type algorithms for strictly second-order non-circular sources and their performance analysis,” *IEEE Trans. Signal Process.*, vol. 62, no. 18, pp. 4824–4838, Sep. 2014.

[13] X. Mestre, “Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates,” *IEEE Trans. Inform. Theory*, vol. 54, no. 11, pp. 5113–5129, Nov. 2008.

[14] M. Shaghaghi and S. A. Vorobyov, “Iterative root-MUSIC algorithm for DOA estimation,” in *Proc. 5th Inter. Workshop Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP 2013)*, The Friendly Island, Saint Martin, Dec. 2013, pp. 53–56.

[15] H. Krim, P. Forster, and J. G. Proakis, “Operator approach to performance analysis of root-MUSIC and root-min-norm,” *IEEE Trans. Signal Process.*, vol. 40, no. 7, pp. 1687–1696, Jul. 1992.

[16] D. Maiwald and D. Kraus, “Calculation of moments of complex Wishart and complex inverse Wishart distributed matrices,” *IEE Proc. Radar, Sonar Navig.*, vol. 147, no. 4, pp. 162–168, Aug. 2000.

[17] M. Shaghaghi, *Parameter estimation in low-rank models from small sample size and undersampled data: DOA and spectrum estimation*, Ph.D. thesis, University of Alberta, 2014.