Subspace Leakage Analysis and Improved DOA Estimation with Small Sample Size

Mahdi Shaghaghi and Sergiy A. Vorobyov

Abstract

Classical methods of DOA estimation such as the MUSIC algorithm are based on estimating the signal and noise subspaces from the sample covariance matrix. For a small number of samples, such methods are exposed to performance breakdown, as the sample covariance matrix can largely deviate from the true covariance matrix. In this paper, the problem of DOA estimation performance breakdown is investigated. We consider the structure of the sample covariance matrix and the dynamics of the root-MUSIC algorithm. The performance breakdown in the threshold region is associated with the subspace leakage where some portion of the true signal subspace resides in the estimated noise subspace. In this paper, the subspace leakage is theoretically derived. We also propose a two-step method which improves the performance by modifying the sample covariance matrix such that the amount of the subspace leakage is reduced. Furthermore, we introduce a phenomenon named as root-swap which occurs in the root-MUSIC algorithm in the low sample size region and degrades the performance of the DOA estimation. A new method is then proposed to alleviate this problem. Numerical examples and simulation results are given for uncorrelated and correlated sources to illustrate the improvement achieved by the proposed methods. Moreover, the proposed algorithms are combined with the pseudo-noise resampling method to further improve the performance.

Index Terms

Covariance matrix, subspace leakage, DOA estimation, root-MUSIC, root-swap.

M. Shaghaghi is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, T6G 2V4 Canada (e-mail: mahdi.shaghaghi@ualberta.ca). S. A. Vorobyov is with Aalto University, Department of Signal Processing and Acoustics, Finland (e-mail: sergiy.vorobyov@aalto.fi). S. A. Vorobyov is the corresponding author.

I. INTRODUCTION

Classical parameter estimation methods of direction-of-arrival (DOA), Doppler shifts, frequencies, time delays, etc. such as the multiple signal classification (MUSIC) [1], root-MUSIC [2], and estimation of signal parameters via rotational invariance techniques (ESPRIT) [3] are based on estimating the signal and noise subspaces from the sample data covariance matrix. It is well-known that these methods suffer from performance breakdown for a small number of samples or low signal-to-noise ratio (SNR) values where the expected estimation error departs from the Cramér-Rao bound (CRB) [4]. The SNR region at which this phenomenon happens is known as the threshold region.

The fidelity of the sample data covariance matrix to the true data covariance matrix plays a critical role in a successful estimation. At the low SNR and/or small sample size region, the sample data covariance matrix can largely deviate from the true one. There are various methods introduced in the literature which target at improving the estimation of the covariance matrix [5]–[12].

Diagonal loading [5] and shrinkage-based [6] methods improve the estimate of the data covariance matrix by scaling and shifting the eigenvalues of the sample data covariance matrix. However, the eigenvectors are kept unchanged. As a result, the estimated signal and noise projection matrices from the improved covariance matrices are exactly the same as those obtained from the sample data covariance matrix. Therefore, these methods are not really beneficial for the subspace-based parameter estimation algorithms.

Data covariance matrix estimation can be also improved by the means of using forward-backward averaging (FB) [7] and spatial smoothing-based techniques [8]. The effect of FB is known to be equivalent to approximately doubling the number of samples. Thus, the covariance estimate improves accordingly. The spatial smoothing technique can also be interpreted as virtually increasing the number of samples at the cost of averaging over sub-arrays of smaller size compared to the whole array. These approaches can also decorrelate pairs (in case of FB) or more correlated source signals. In [9], techniques from random matrix theory have been developed to improve the performance of the MUSIC algorithm. The introduced method considers the asymptotic situation when both the sample size and the number of array elements tend to infinity at the same rate. It is then inferred that the improved method gives a more accurate description

of the situation when these two quantities are finite and comparable in magnitude. However, the performance of the introduced method is not satisfactory at the small sample size scenario [13].

A more promising approach to remedy the performance breakdown at the threshold region was introduced in [10] and has been further improved in [11] and [12]. These methods are based on a technique called pseudo-noise resampling which uses synthetically generated pseudo-noise to perturb the original noise. The pseudo-noise is added to the observed data, and a new estimate of the covariance matrix is obtained, which leads to new DOA estimates. This process is repeated for a number of times, and the final DOAs are determined based on the bank of the DOA estimates.

In this paper, we tackle the problem of the performance breakdown at the threshold region by considering the structure of the sample data covariance matrix and the dynamics of the root-MUSIC algorithm. It is shown in [14] that the performance breakdown problem is associated with the inter-subspace leakage "whereby a small portion of the true signal eigenvector resides in the sample noise subspace (and vice versa)". The notion of leakage comes originally from the performance assessment strategy based on the first order approximation of the estimation error caused by the perturbed subspace estimate, which happens because of the additive noise contribution [15]–[18]. This approach directly models the leakage of the noise subspace into signal subspace and allows to compute the corresponding perturbation matrix between the components of the subspaces. Here, we formally define the *subspace leakage* notion as a Frobenius norm of the perturbation matrix, and we present its theoretical derivation. We propose a two-step method which improves the performance of the root-MUSIC algorithm by modifying the sample data covariance matrix such that the amount of the subspace leakage is reduced. Furthermore, we introduce a phenomenon named as *root-swap* which occurs in the root-MUSIC algorithm at the threshold region and degrades the performance of the parameter estimation. A new method is then proposed to alleviate this problem.

It will be shown that there are undesirable by-products in the sample data covariance matrix that tend to zero as the number of samples goes to infinity. However, for a limited number of samples, these terms can have significant values leading to a large amount of subspace leakage. One possible approach to remedy the effect of the undesirable components is to consider the *eigenvalue perturbation* caused by these terms. The incorporation of this knowledge into the estimation method can result in better estimates of the signal and noise subspaces. In this paper,

we propose a two-step algorithm in order to reduce the effect of the undesirable terms. The introduced method is based on estimating the parameters at the first step and modifying the covariance matrix using the estimated parameters at the second step. We will theoretically derive the subspace leakage at both steps. Then, it will be shown using numerical examples that the subspace leakage is reduced at the second step leading to better performance.

In the root-MUSIC method, the estimation error of the roots has a variance which is proportional to the variance of noise over the number of samples [19]. Therefore, at the threshold region, the variance of the estimation error can have a significant value which in turn can result in a swap between a root corresponding to a signal source with another root which is not associated with any signal source. We dub this phenomenon as root-swap. Then, a new method is proposed to remedy this problem. The introduced method considers different combinations of the roots as the candidates for the signal sources. These candidates are then evaluated using the stochastic maximum likelihood (SML) function, and the combination that minimizes the objective function is picked up for the parameter estimates.

The rest of the paper is organized as follows. The system model is given and the root-MUSIC algorithm is briefly reviewed in Section II. The two-step and root-swap algorithms are proposed in Section III. Subspace leakage is defined and theoretically derived in Section IV. Numerical examples and simulation results are given in Section V. Section VI concludes the paper. Appendix A gives an approximation for the probability of root-swap, and finally, the details of the subspace leakage derivation for the two-step root-MUSIC algorithm are presented in Appendices B and C.

II. SYSTEM MODEL AND BACKGROUND

An example of a noise-corrupted linear superposition of K undamped exponentials received by M $(M > K)$ antennas is the array processing model. Thus, consider K number of narrowband plane waves impinging on a uniform linear array (ULA) from directions $\theta_1, \theta_2, \cdots, \theta_K$. Without loss of generality, assume $-\pi/2 \leq \theta_1 \leq \theta_2 \leq \cdots \leq \theta_K \leq \pi/2$. The antenna elements are separated from each other by a distance of $d \leq \lambda/2$ where λ is the wavelength of the plane wave impinging on the array. The steering vector of the array $\boldsymbol{a}(\theta) \in \mathbb{C}^{M \times 1}$ is then given as

$$\boldsymbol{a}(\theta) \triangleq \left[1, \; e^{-j2\pi(d/\lambda)\sin(\theta)}, \cdots, \; e^{-j2\pi(M-1)(d/\lambda)\sin(\theta)}\right]^T \tag{1}$$

where $(\cdot)^T$ stands for the transposition operator. At time instant $t \in \mathbb{N}$, the received vector $\boldsymbol{x}(t) \in \mathbb{C}^{M \times 1}$ is given by

$$\boldsymbol{x}(t) = \sum_{i=1}^{K} \boldsymbol{a}(\theta_i) s_i(t) + \boldsymbol{n}(t) \tag{2}$$

where $s_i(t) \in \mathbb{C}$ is the amplitude of the i-th wave (source) and $\boldsymbol{n}(t) \in \mathbb{C}^{M \times 1}$ is the noise vector at time t. By arranging the amplitudes of the sources in the vector $\boldsymbol{s}(t) = [s_1(t), s_2(t), \cdots, s_K(t)]^T \in \mathbb{C}^{K \times 1}$ and forming the Vandermonde matrix $\boldsymbol{A} = [\boldsymbol{a}(\theta_1), \ \boldsymbol{a}(\theta_2), \cdots, \ \boldsymbol{a}(\theta_K)] \in \mathbb{C}^{M \times K}$, the model (2) can be rewritten in matrix-vector form as

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{n}(t). \tag{3}$$

We consider the noise vector $\boldsymbol{n}(t)$ to be independent from the sources and noise vectors at other time instances and to have 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. Considering the system model (3), the data covariance matrix $\boldsymbol{R} \in \mathbb{C}^{M \times M}$ is given by

$$\boldsymbol{R} \triangleq E\left\{\boldsymbol{x}(t)\boldsymbol{x}^H(t)\right\} = \boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^H + \sigma_{\mathrm{n}}^2 \boldsymbol{I}_M \tag{4}$$

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

Let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_M$ be the eigenvalues of $\boldsymbol{R}$ arranged in nondecreasing order, and let $\boldsymbol{g}_1, \ \boldsymbol{g}_2, \cdots, \ \boldsymbol{g}_{M-K}$ be the noise eigenvectors associated with $\lambda_1, \ \lambda_2, \cdots, \ \lambda_{M-K}$ and $\boldsymbol{e}_1, \ \boldsymbol{e}_2, \cdots, \ \boldsymbol{e}_K$ be the signal eigenvectors corresponding to $\lambda_{M-K+1}, \ \lambda_{M-K+2}, \cdots, \ \lambda_M$. Let also $\boldsymbol{G} \in \mathbb{C}^{M \times (M-K)}$ and $\boldsymbol{E} \in \mathbb{C}^{M \times K}$ be defined as $\boldsymbol{G} \triangleq [\boldsymbol{g}_1, \ \boldsymbol{g}_2, \cdots, \ \boldsymbol{g}_{M-K}]$ and $\boldsymbol{E} \triangleq [\boldsymbol{e}_1, \ \boldsymbol{e}_2, \cdots, \ \boldsymbol{e}_K]$. The range spaces of $\boldsymbol{G}$ and $\boldsymbol{E}$ represent the true noise and signal subspaces, 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

$$\widehat{\boldsymbol{R}} \triangleq \frac{1}{N} \sum_{t=1}^{N} \boldsymbol{x}(t)\boldsymbol{x}^H(t) \tag{5}$$

where $\widehat{\boldsymbol{R}} \in \mathbb{C}^{M \times M}$ is the sample data covariance matrix. Consider the eigendecomposition of $\widehat{\boldsymbol{R}}$. Let $\hat{\boldsymbol{g}}_1, \ \hat{\boldsymbol{g}}_2, \cdots, \ \hat{\boldsymbol{g}}_{M-K}$ be the estimated noise eigenvectors and $\hat{\boldsymbol{e}}_1, \ \hat{\boldsymbol{e}}_2, \cdots, \ \hat{\boldsymbol{e}}_K$ be the estimated signal eigenvectors. Form $\widehat{\boldsymbol{G}} \in \mathbb{C}^{M \times (M-K)}$ and $\widehat{\boldsymbol{E}} \in \mathbb{C}^{M \times K}$ by placing the estimated

noise and signal eigenvectors as the columns of $\widehat{\boldsymbol{G}}$ and $\widehat{\boldsymbol{E}}$, respectively. The range spaces of $\widehat{\boldsymbol{G}}$ and $\widehat{\boldsymbol{E}}$ represent the estimations of the noise and signal subspaces, respectively.

Recalling (1) and defining $z \triangleq e^{j2\pi(d/\lambda)\sin(\theta)}$, the steering vector can be rewritten as $\boldsymbol{a}(z) = \left[1, z^{-1}, \cdots, z^{-(M-1)}\right]^{T}$. In the root-MUSIC method, the roots of the equation $\boldsymbol{a}^{T}(z^{-1})\widehat{\boldsymbol{G}}\widehat{\boldsymbol{G}}^{H}\boldsymbol{a}(z) = 0$ which are located inside the unit circle are considered. These roots are sorted based on their distance to the unit circle, and the first K number of the roots which are closer to the unit circle are picked. The estimates of the DOAs denoted by $\hat{\theta}_1$, $\hat{\theta}_2, \cdots$, $\hat{\theta}_K$ are then obtained by multiplying the angles of the selected roots by $\lambda/(2\pi d)$ and taking the inverse sinusoid function of the results.

III. PROPOSED METHODS

A. Two-step root-MUSIC algorithm

Let us start by expanding (5) using (3) as follows

$$
\begin{aligned}
\widehat{\boldsymbol{R}} = &\frac{1}{N} \sum_{t=1}^{N} \left(\boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{n}(t)\right) \left(\boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{n}(t)\right)^{H} \\
= &\boldsymbol{A} \left\{ \frac{1}{N} \sum_{t=1}^{N} \boldsymbol{s}(t)\boldsymbol{s}^{H}(t) \right\} \boldsymbol{A}^{H} + \frac{1}{N} \sum_{t=1}^{N} \boldsymbol{n}(t)\boldsymbol{n}^{H}(t) \\
&+ \boldsymbol{A} \left\{ \frac{1}{N} \sum_{t=1}^{N} \boldsymbol{s}(t)\boldsymbol{n}^{H}(t) \right\} + \left\{ \frac{1}{N} \sum_{t=1}^{N} \boldsymbol{n}(t)\boldsymbol{s}^{H}(t) \right\} \boldsymbol{A}^{H}.
\end{aligned}
\tag{6}
$$

Comparing (6) with (4), it can be observed that the expansion of $\widehat{\boldsymbol{R}}$ consists of four terms while the model for $\boldsymbol{R}$ comprises two summands. The first two terms of $\widehat{\boldsymbol{R}}$ given by (6) can be considered as estimates for the two summands of $\boldsymbol{R}$, which represent the signal and noise components, respectively. The last two terms of $\widehat{\boldsymbol{R}}$ in (6) are undesirable by-products which can be viewed as estimates for the correlation between the signal and noise vectors. In the system model under study, we consider the noise vectors to be zero-mean and also independent of the signal vectors. Therefore, the signal and noise components are uncorrelated to each other. As a result, for a large enough number of samples N, the last two terms in (6) tend to zero. However, the number of available samples can be limited in practical applications. In this case, the last two terms in (6) may have significant values, which causes the estimates of the signal and noise subspaces to deviate from the true signal and noise subspaces.

The main idea of our two-step root-MUSIC algorithm is to modify the sample data covariance matrix at the second step based on the DOA estimates obtained at the first step. The modified covariance matrix is obtained by deducting a scaled version of the estimated undesirable terms from the sample data covariance matrix.

We derive the steps of the proposed method for a general source covariance matrix $\boldsymbol{S}$, so that correlated sources can also be handled by the algorithm. Furthermore, the proposed method can also be beneficial in the case that the assumption on no correlation between the source and noise vectors is not fully met. This is achieved by estimating and removing the correlation terms between the source and noise vectors from the sample data covariance matrix.

The steps of the proposed method are listed in Table I. The algorithm starts by computing the sample data covariance matrix (5). Then, DOAs are estimated using the root-MUSIC algorithm. The superscript $(\cdot)^{(1)}$ refers to the estimation made at the first step. At the second step, the Vandermonde matrix is formed using the available estimates of the DOAs. Then, the amplitudes of the sources are estimated such that the squared norm of the differences between the observations and the estimates are minimized. The corresponding problem is formulated as

$$\hat{\boldsymbol{s}}(t) = \arg\ \min_{\boldsymbol{s}} \|\boldsymbol{x}(t) - \widehat{\boldsymbol{A}}\boldsymbol{s}\|_2^2. \tag{7}$$

The minimization of (7) is performed using the least squares (LS) technique and the corresponding solution is given as

$$\hat{\boldsymbol{s}}(t) = \left(\widehat{\boldsymbol{A}}^H \widehat{\boldsymbol{A}}\right)^{-1} \widehat{\boldsymbol{A}}^H \boldsymbol{x}(t). \tag{8}$$

The noise component is then estimated as the difference between the estimated signal and the observation made by the array, i.e.,

$$\hat{\boldsymbol{n}}(t) = \boldsymbol{x}(t) - \widehat{\boldsymbol{A}}\hat{\boldsymbol{s}}(t). \tag{9}$$

After estimating the signal and noise vectors, the third term in (6) can be found as

$$
\begin{aligned}
\boldsymbol{T} &\triangleq \widehat{\boldsymbol{A}} \left\{ \frac{1}{N} \sum_{t=1}^{N} \hat{\boldsymbol{s}}(t) \hat{\boldsymbol{n}}^H(t) \right\} \\
&= \widehat{\boldsymbol{A}} \left\{ \frac{1}{N} \sum_{t=1}^{N} \left(\widehat{\boldsymbol{A}}^H \widehat{\boldsymbol{A}} \right)^{-1} \widehat{\boldsymbol{A}}^H \boldsymbol{x}(t) \left(\boldsymbol{x}^H(t) - \boldsymbol{x}^H(t) \widehat{\boldsymbol{A}} \left(\widehat{\boldsymbol{A}}^H \widehat{\boldsymbol{A}} \right)^{-1} \widehat{\boldsymbol{A}}^H \right) \right\} \\
&= \widehat{\boldsymbol{P}}_A \left\{ \frac{1}{N} \sum_{t=1}^{N} \boldsymbol{x}(t) \boldsymbol{x}^H(t) \left(\boldsymbol{I}_M - \widehat{\boldsymbol{P}}_A \right) \right\} \\
&= \widehat{\boldsymbol{P}}_A \widehat{\boldsymbol{R}} \widehat{\boldsymbol{P}}_A^{\perp}
\end{aligned}
\tag{10}
$$

where

$$
\widehat{\boldsymbol{P}}_A \triangleq \widehat{\boldsymbol{A}} \left(\widehat{\boldsymbol{A}}^H \widehat{\boldsymbol{A}} \right)^{-1} \widehat{\boldsymbol{A}}^H
\tag{11}
$$

is an estimation for the projection matrix of the signal subspace, and

$$
\widehat{\boldsymbol{P}}_A^{\perp} \triangleq \boldsymbol{I}_M - \widehat{\boldsymbol{P}}_A
\tag{12}
$$

is an estimation for the projection matrix of the noise subspace. The forth term in (6) is equal to the Hermitian of the third term, i.e., $\boldsymbol{T}^H$. Finally, the modified data covariance matrix is obtained by deducting a scaled version of the estimated terms from the initial sample data covariance matrix as follows

$$
\widehat{\boldsymbol{R}}^{(2)} = \widehat{\boldsymbol{R}} - \gamma \left(\boldsymbol{T} + \boldsymbol{T}^H \right).
\tag{13}
$$

The scaling factor γ in (13) is a real number between zero and one. Ideally, the value of γ would be equal to 1 if the estimates of the undesirable terms were perfect. However, estimation errors are inevitable, and therefore, we have introduced γ to deal with the imperfections. The scaling factor γ can be considered as a reliability factor which takes a value close to 1 for an estimate of $\boldsymbol{T}$ with small error and a small value if an estimate of $\boldsymbol{T}$ is erroneous. Given a value for γ, the modified data covariance matrix $\widehat{\boldsymbol{R}}^{(2)}$ is computed and the DOAs are estimated again using the root-MUSIC algorithm.

The value of γ can be fixed to a predetermined value before running the algorithm, or it can be obtained based on the observations. Since γ is a real number between zero and one, we can consider different values for γ taken on a grid (e.g. $\gamma = 0, \ 0.1, \ 0.2, \cdots, \ 1$). For each value of γ, a set of DOA estimates is obtained based on the modified data covariance matrix. Next, we determine which value of γ results in a better estimation. This can be done by choosing a set

of DOA estimates that has a higher likelihood of being the set of true DOAs. In other words, we use the maximum likelihood (ML) criterion to evaluate the quality of the estimated DOAs. Since the system model given in (4) is stochastic, we use the stochastic ML (SML) objective function given by [20]

$$
F_{SML}(\gamma) = \ln \det \left(\widehat{\boldsymbol{P}}_A^{(2)} \widehat{\boldsymbol{R}} \widehat{\boldsymbol{P}}_A^{(2)} + \frac{\operatorname{Tr}\left\{ \widehat{\boldsymbol{P}}_A^{\perp(2)} \widehat{\boldsymbol{R}} \right\}}{M - K} \widehat{\boldsymbol{P}}_A^{\perp(2)} \right)
\tag{14}
$$

where $\operatorname{Tr}\left\{\cdot\right\}$ stands for the trace operator, $\widehat{\boldsymbol{P}}_A^{(2)}$ is an estimation of the projection matrix of the signal subspace obtained from the estimated DOAs based on the modified data covariance matrix and $\widehat{\boldsymbol{P}}_A^{\perp(2)} = \boldsymbol{I}_M - \widehat{\boldsymbol{P}}_A^{(2)}$. The objective function in (14) is evaluated for each value of γ. Then, the set of DOA estimates corresponding to the value of γ that minimizes (14) is chosen as the output of the algorithm.

B. Root-swap root-MUSIC algorithm

Consider the root-MUSIC polynomial $\boldsymbol{a}^T(z^{-1})\boldsymbol{G}\boldsymbol{G}^H\boldsymbol{a}(z)$ which is formed by the noise eigenvectors obtained from the eigendecomposition of the data covariance matrix $\boldsymbol{R}$. This polynomial has K number of roots on the unit circle which correspond to the signal sources. Let these K roots be denoted by $z_1, z_2, \cdots, z_K$ and be referred to as the true signal roots. The polynomial also has additional $M - K - 1$ number of roots inside the unit circle. Let these roots be referred to as the true noise roots and be denoted by $z_{K+1}, z_{K+2}, \cdots, z_{M-1}$.

An estimation for the root-MUSIC polynomial can be formed using the noise eigenvectors obtained from the sample data covariance matrix $\widehat{\boldsymbol{R}}$. Let us assume that in the estimation of the noise and signal subspaces, no subspace swap has occurred [4]. The estimated polynomial is given by $\boldsymbol{a}^T(z^{-1})\widehat{\boldsymbol{G}}\widehat{\boldsymbol{G}}^H\boldsymbol{a}(z)$. This polynomial has $M - 1$ number of roots inside the unit circle. Let $\hat{z}_1, \hat{z}_2, \cdots, \hat{z}_K$ be the roots of the estimated root-MUSIC polynomial which correspond to $z_1, z_2, \cdots, z_K$. We refer to these roots as the estimated signal roots. Furthermore, let $\hat{z}_{K+1}, \hat{z}_{K+2}, \cdots, \hat{z}_{M-1}$ be the roots corresponding to $z_{K+1}, z_{K+2}, \cdots, z_{M-1}$. These roots are referred to as the estimated noise roots.

In the root-MUSIC method, we do not have the knowledge about which of the roots of the estimated root-MUSIC polynomial correspond to the true signal roots. The conventional rule is

TABLE I

TWO-STEP ROOT-MUSIC ALGORITHM

Inputs:

$M,\ d,\ \lambda,\ N,\ K$, and

received vectors $\boldsymbol{x}(1),\ \boldsymbol{x}(2), \cdots,\ \boldsymbol{x}(N)$

Outputs:

Estimates $\hat{\theta}_1^{(2)},\ \hat{\theta}_2^{(2)}, \cdots,\ \hat{\theta}_K^{(2)}$

Step 1:

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

$\left\{ \hat{\theta}_1^{(1)},\ \hat{\theta}_2^{(1)}, \cdots,\ \hat{\theta}_K^{(1)} \right\} \leftarrow \text{root-MUSIC}\left(\widehat{\boldsymbol{R}}, K, d, \lambda \right)$

Step 2:

$\widehat{\boldsymbol{A}} = \left[\boldsymbol{a}\left(\hat{\theta}_1^{(1)}\right),\ \boldsymbol{a}\left(\hat{\theta}_2^{(1)}\right), \cdots,\ \boldsymbol{a}\left(\hat{\theta}_K^{(1)}\right) \right]$

$\widehat{\boldsymbol{P}}_A = \widehat{\boldsymbol{A}} \left(\widehat{\boldsymbol{A}}^H \widehat{\boldsymbol{A}} \right)^{-1} \widehat{\boldsymbol{A}}^H$

$\widehat{\boldsymbol{P}}_A^{\perp} = \boldsymbol{I}_M - \widehat{\boldsymbol{P}}_A$

$\boldsymbol{T} = \widehat{\boldsymbol{P}}_A \widehat{\boldsymbol{R}} \widehat{\boldsymbol{P}}_A^{\perp}$

Determine γ as the minimizer of (14)

$\widehat{\boldsymbol{R}}^{(2)} = \widehat{\boldsymbol{R}} - \gamma \left(\boldsymbol{T} + \boldsymbol{T}^H \right)$

$\left\{ \hat{\theta}_1^{(2)},\ \hat{\theta}_2^{(2)}, \cdots,\ \hat{\theta}_K^{(2)} \right\} \leftarrow \text{root-MUSIC}\left(\widehat{\boldsymbol{R}}^{(2)}, K, d, \lambda \right)$

---

to select K number of the estimated roots which are closer to the unit circle as the estimates for the true signal roots. Then, the DOAs are estimated based on the angles of these roots.

Due to the finiteness of the available samples, the estimated roots obtained from the sample data covariance matrix $\widehat{\boldsymbol{R}}$ deviate from their corresponding true roots obtained from the true data covariance matrix $\boldsymbol{R}$. Let r_i and $\hat{r}_i$ represent the magnitudes of z_i and $\hat{z}_i$ for $1 \leq i \leq M-1$, respectively. Furthermore, let $\Delta r_i \triangleq \hat{r}_i - r_i$ be the difference between the magnitude of the i-th estimated root and the magnitude of the corresponding true root. It is shown in [19] that Δr_i (for the signal roots) has a variance which is proportional to σ_n^2/N. Therefore, Δr_i can have a significant value for a small number of samples and a large value of σ_n^2 (low SNR region). Consequently, there can be a considerable probability that an estimated signal root takes

a smaller magnitude than an estimated noise root. We refer to this phenomenon as a root-swap. The root-swap probability is approximately found in Appendix A as

$$P(\text{root-swap}) \approx 1 - \prod_{k=1}^{K} \prod_{m=K+1}^{M-1} Q\left(\frac{-1 + r_m + \sigma_k\sqrt{M - K - (3/4)}}{\sqrt{\sigma_k^2/4}}\right) \tag{15}$$

where $Q\left(\cdot\right)$ is the tail probability of the standard normal distribution and $\sigma_k^2/4$ is the variance of Δr_k, and it is proportional to σ_{n}^2/N.

In the case that the root-swap happens, selecting the first closest K roots to the unit circle results in picking a noise root instead of a signal root. To deal with this problem, we propose an algorithm that considers different combinations of the roots as candidates for signal roots. The method is dubbed the root-swap root-MUSIC algorithm.

The root-MUSIC polynomial has $M - 1$ number of roots inside the unit circle. Our goal is to find the roots which have a higher likelihood of being associated with the K sources. Consider choosing K number of roots out of the $M - 1$ roots inside the unit circle. There are $N_c \triangleq (M - 1)!/(K!(M - K - 1)!)$ different possible combinations. Let $\Gamma \triangleq \{\Theta_1, \Theta_2, \cdots, \Theta_{N_c}\}$ where Θ_i ($1 \leq i \leq N_c$) is a set containing the DOA estimates obtained from the i-th combination of the roots. Then, the root-swap root-MUSIC method estimates the DOAs as

$$\left\{\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_K\right\} = \arg\ \min_{\Theta \in \Gamma} F_{SML}\left(\Theta\right) \tag{16}$$

where $F_{SML}\left(\Theta\right)$ is the SML function given by

$$F_{SML}(\Theta) = \ln \det\left(\boldsymbol{P}_\Theta \widehat{\boldsymbol{R}} \boldsymbol{P}_\Theta + \frac{\mathrm{Tr}\left\{\boldsymbol{P}_\Theta^\perp \widehat{\boldsymbol{R}}\right\}}{M - K} \boldsymbol{P}_\Theta^\perp\right) \tag{17}$$

and $\boldsymbol{P}_\Theta$ is the signal projection matrix obtained from Θ as

$$\boldsymbol{P}_\Theta \triangleq \boldsymbol{A}(\Theta)\left(\boldsymbol{A}^H(\Theta)\boldsymbol{A}(\Theta)\right)^{-1}\boldsymbol{A}^H(\Theta). \tag{18}$$

The complexity of the introduced root-swap root-MUSIC method can be reduced by pre-eliminating some of the roots. Specifically, let $p \leq K$ roots closest to the unit circle be picked, and let q number of roots closest to the origin (furthest from the unit circle) be ignored. Our task is to choose $K - p$ number of roots out of $M - p - q - 1$ roots. Then, there are $N_{\mathrm{r}} \triangleq (M - p - q - 1)!/((K - p)!(M - K - q - 1)!)$ different possible combinations which is significantly smaller than N_{c}. The rest of the algorithm is the same as above except for that here each combination contains $K - p$ number of roots. Therefore, in order to evaluate the SML function, the fixed p pre-selected roots are added to each combination.

IV. SUBSPACE LEAKAGE

The performance breakdown of the subspace based DOA estimation methods in the threshold region has been associated with the subspace leakage. In this section, we study the subspace leakage for both steps of the proposed two-step root-MUSIC algorithm.

A. Definition

Recall the matrices $\boldsymbol{G}$ and $\boldsymbol{E}$ which are composed of the true noise and signal eigenvectors obtained from the eigendecomposition of the data covariance matrix $\boldsymbol{R}$. 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

$$\boldsymbol{G}\boldsymbol{G}^H + \boldsymbol{E}\boldsymbol{E}^H = \boldsymbol{I}_M \tag{19}$$

or

$$\boldsymbol{P}^\perp + \boldsymbol{P} = \boldsymbol{I}_M \tag{20}$$

where, $\boldsymbol{P}^\perp \triangleq \boldsymbol{G}\boldsymbol{G}^H$ and $\boldsymbol{P} \triangleq \boldsymbol{E}\boldsymbol{E}^H$ are the true projection matrices of the noise and signal subspaces.

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 $\|\boldsymbol{P}^\perp \hat{\boldsymbol{e}}_k\|_2^2$ is almost surely nonzero, which can be viewed as the leakage of $\hat{\boldsymbol{e}}_k$ into the true noise subspace.

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.,

$$\rho \triangleq \frac{1}{K} \sum_{k=1}^{K} \|\boldsymbol{P}^\perp \hat{\boldsymbol{e}}_k\|_2^2. \tag{21}$$

Note that $\boldsymbol{P}^\perp$ is the orthogonal projection matrix. Therefore, ρ can be written as

$$\rho = \frac{1}{K} \sum_{k=1}^{K} \hat{\boldsymbol{e}}_k^H \boldsymbol{P}^\perp \hat{\boldsymbol{e}}_k. \tag{22}$$

Using (20) and some algebra, the expression (22) can be simplified to

$$\rho = \frac{1}{K} \sum_{k=1}^{K} \hat{e}_k^H \left(\boldsymbol{I}_M - \boldsymbol{P} \right) \hat{e}_k$$

$$= 1 - \frac{1}{K} \sum_{k=1}^{K} \mathrm{Tr} \left\{ \hat{e}_k \hat{e}_k^H \boldsymbol{P} \right\}$$

$$= 1 - \frac{1}{K} \mathrm{Tr} \left\{ \left(\sum_{k=1}^{K} \hat{e}_k \hat{e}_k^H \right) \boldsymbol{P} \right\}$$

$$= 1 - \frac{1}{K} \mathrm{Tr} \left\{ \widehat{\boldsymbol{P}} \boldsymbol{P} \right\} \tag{23}$$

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

B. Analysis of two-step root-MUSIC algorithm

The estimated signal and noise projection matrices obtained from the eigendecomposition of the sample data covariance matrix $\widehat{\boldsymbol{R}}$ are deviated from the true signal and noise projection matrices. Let $\Delta \boldsymbol{R} \triangleq \widehat{\boldsymbol{R}} - \boldsymbol{R}$ be the estimation error of the data covariance matrix, and let

$$\boldsymbol{V} \triangleq \boldsymbol{R} - \sigma_{\mathrm{n}}^2 \boldsymbol{I}_M = \boldsymbol{A} \boldsymbol{S} \boldsymbol{A}^H$$

$$= \sum_{k=1}^{K} \left(\lambda_{M-K+k} - \sigma_{\mathrm{n}}^2 \right) \boldsymbol{e}_k \boldsymbol{e}_k^H. \tag{24}$$

Denote the pseudo-inverse of $\boldsymbol{V}$ as $\boldsymbol{V}^\dagger \in \mathbb{C}^{M \times M}$. It is given by

$$\boldsymbol{V}^\dagger = \sum_{k=1}^{K} \frac{1}{\lambda_{M-K+k} - \sigma_{\mathrm{n}}^2} \boldsymbol{e}_k \boldsymbol{e}_k^H. \tag{25}$$

Let ρ_1 and ρ_2 be the subspace leakage due to the error in the estimation of the signal and noise subspaces obtained from $\widehat{\boldsymbol{R}}$ and $\widehat{\boldsymbol{R}}^{(2)}$, respectively. Note that ρ_1 only depends on $\boldsymbol{R}$ and $\Delta \boldsymbol{R}$, and it is not specific to the proposed two-step root-MUSIC algorithm.

It is shown in Appendix B that ρ_1 and its expected value are given by

$$\rho_1 = \frac{1}{K} \mathrm{Tr} \left\{ \boldsymbol{V}^\dagger \Delta \boldsymbol{R} \boldsymbol{P}^\perp \Delta \boldsymbol{R} \boldsymbol{V}^\dagger \right\} \tag{26}$$

and

$$E \left\{ \rho_1 \right\} = \frac{\sigma_{\mathrm{n}}^2 \left(M - K \right)}{NK} \sum_{k=1}^{K} \frac{\lambda_{M-K+k}}{\left(\lambda_{M-K+k} - \sigma_{\mathrm{n}}^2 \right)^2} \tag{27}$$

respectively.

It can be seen from (27) that the expected value of the subspace leakage is proportional to σ_{n}^2/N. Therefore, the amount of the subspace leakage can be significant for a small number of samples or low SNR values. The variance of ρ_1 has also been studied in [21], and it has been shown that Var (ρ_1) is in the order of $1/N^2$.

The subspace leakage at the second step of the two-step root-MUSIC algorithm is computed in Appendix C and is given by

$$\rho_2 = \left(1 - 2\gamma + \gamma^2\right)\rho_1 + \frac{2\left(\gamma - \gamma^2\right)}{K}Re\left\{\mathrm{Tr}\left\{\boldsymbol{V}^\dagger \Delta \boldsymbol{R} \boldsymbol{P}^\perp d\boldsymbol{P}\right\}\right\} + \frac{\gamma^2}{K}\mathrm{Tr}\left\{d\boldsymbol{P}\boldsymbol{P}^\perp d\boldsymbol{P}\right\} \quad (28)$$

where $Re\left\{\cdot\right\}$ stands for the real part operator, and $d\boldsymbol{P}$ is the first order term in the Taylor series expansion of $\widehat{\boldsymbol{P}}_A$ around the true DOAs. It is also shown in Appendix C that the expected value of ρ_2 for a fixed value of γ is given by

$$\begin{aligned}
E\left\{\rho_2\right\} = {}& \left(1 - 2\gamma + \gamma^2\right)E\left\{\rho_1\right\} \\
&+ \frac{2\left(\gamma - \gamma^2\right)\sigma_{\mathrm{n}}^2}{NK}Re\left\{\sum_{k=1}^{K}\frac{\boldsymbol{a}_k^{(1)H}\boldsymbol{P}^\perp\frac{\partial\boldsymbol{A}}{\partial\omega_k}\left(\boldsymbol{A}^H\boldsymbol{A}\right)^{-1}\boldsymbol{A}^H\boldsymbol{V}^\dagger\boldsymbol{R}\boldsymbol{V}^\dagger\boldsymbol{a}_k}{2j\left(\boldsymbol{a}_k^{(1)H}\boldsymbol{P}^\perp\boldsymbol{a}_k^{(1)}\right)}\right\} \\
&+ \frac{\gamma^2\sigma_{\mathrm{n}}^2}{2NK}\sum_{k=1}^{K}\sum_{i=1}^{K}\frac{\mathrm{Tr}\left\{\left(\frac{\partial\boldsymbol{A}}{\partial\omega_k}\right)^H\boldsymbol{P}^\perp\frac{\partial\boldsymbol{A}}{\partial\omega_i}\left(\boldsymbol{A}^H\boldsymbol{A}\right)^{-1}\right\}}{\left(\boldsymbol{a}_k^{(1)H}\boldsymbol{P}^\perp\boldsymbol{a}_k^{(1)}\right)\left(\boldsymbol{a}_i^{(1)H}\boldsymbol{P}^\perp\boldsymbol{a}_i^{(1)}\right)}Re\left\{\boldsymbol{a}_i^H\boldsymbol{V}^\dagger\boldsymbol{R}\boldsymbol{V}^\dagger\boldsymbol{a}_k\boldsymbol{a}_k^{(1)H}\boldsymbol{P}^\perp\boldsymbol{a}_i^{(1)}\right\}
\end{aligned}$$

$$(29)$$

where $\omega_k \triangleq 2\pi(d/\lambda)\sin(\theta_k)$, $\boldsymbol{a}_k$ is a shorthand notation for $\boldsymbol{a}(\theta_k)$, and $\boldsymbol{a}_k^{(1)} \in \mathbb{C}^{M\times 1}$ is defined as

$$\boldsymbol{a}_k^{(1)} \triangleq -\left[0,\ e^{-j\omega_k},\ 2e^{-j2\omega_k},\ \cdots,\ (M-1)e^{-j(M-1)\omega_k}\right]^T. \quad (30)$$

It can be seen in (29) that for $\gamma = 0$, $E\left\{\rho_2\right\}$ reduces to $E\left\{\rho_1\right\}$ as expected, and for $\gamma = 1$, the first two terms in (29) are equal to zero.

V. Numerical Examples and Simulation Results

In this section, the performance of the proposed two-step root-MUSIC and the root-swap root-MUSIC algorithms is investigated and compared with the performance of the unitary root-MUSIC method [22] and the improved unitary root-MUSIC algorithm based on pseudo-noise resampling [12]. We also consider the combination of the proposed methods with the other methods in order to achieve further performance improvement. Compared to the root-MUSIC

method, the unitary root-MUSIC algorithm has a lower computational complexity as it uses the eigendecomposition of a real-valued covariance matrix. Furthermore, the unitary root-MUSIC algorithm has better performance for the case that the sources are correlated. The improved unitary root-MUSIC algorithm based on pseudo-noise resampling increases the estimator complexity, but it is advantageous in removing the outliers, which results in better performance.

We consider $K = 2$ sources impinging on an array of $M = 10$ antenna elements from directions $\theta_1 = 35^\circ \times (\pi/180)$ and $\theta_2 = 37^\circ \times (\pi/180)$. The interelement spacing is set to $d = \lambda/2$ and the number of snapshots is $N = 10$. Each source vector $\boldsymbol{s}(t)$ is considered to be independent from the source vectors at other time instances and to have the circularly-symmetric complex jointly-Gaussian distribution $\mathcal{N}_C(0, \boldsymbol{S})$. The source covariance matrix $\boldsymbol{S}$ is given by

$$\boldsymbol{S} = \sigma_{\mathrm{s}}^2 \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}$$

where $0 \leq r \leq 1$ is the correlation coefficient. The SNR is defined as $\mathrm{SNR} \triangleq 10 \log_{10} (\sigma_{\mathrm{s}}^2/\sigma_{\mathrm{n}}^2)$.

The performance of the proposed algorithms is investigated by considering the subspace leakage, mean squared error (MSE), probability of source resolution, and conditional mean squared error (CMSE). Source resolution is defined as the event when both DOAs are estimated within one degree of their corresponding true values, i.e., the difference between the true value of each DOA and its estimated value is less than $1^\circ \times (\pi/180)$. The CMSE is defined as the expected value of the estimation error conditioned on successful source resolution, i.e., $E\left\{ \sum_{k=1}^{K} \|\hat{\theta}_k - \theta_k\|_2^2 \, \Big| \, \text{successful source resolution} \right\}$. The reason for using the CMSE is to further investigate the accuracy of the algorithms after making successful detection. We estimate the probability of root-swap, subspace leakage, MSE, probability of source resolution, and CMSE using the Monte Carlo method with 10^5 number of trials. Two cases are considered in the simulations: 1) the two sources are uncorrelated, i.e., $r = 0$, and 2) the two sources are correlated with a correlation coefficient of $r = 0.9$.

Let us start by investigating the probability of root-swap in the root-MUSIC algorithm for the case of the uncorrelated sources. The probability of root-swap is estimated using the Monte Carlo simulations. Its approximate value is also obtained using (15). The corresponding curves are shown in Fig. 1. It can be seen that at the low SNR region, the chance that a root-swap occurs is quite significant, which results in the performance breakdown of the root-MUSIC algorithm. This

Fig. 1.   Probability of root-swap and probability of ML failure versus SNR for uncorrelated sources.

problem justifies the need for a method to deal with the root-swap phenomenon. In this paper, we proposed the root-swap root-MUSIC algorithm which instead of picking the roots closer to the unit circle, selects the roots based on the SML criterion. In Fig. 1, we thus also draw a curve which shows the probability that the selected roots by the ML criterion include a noise root. This situation is considered as a breakdown, and therefore, the corresponding probability is called the probability of ML failure. As can be seen, this probability is significantly smaller than the probability of root-swap. As a result, it is expected that the root-swap root-MUSIC algorithm outperforms the conventional root-MUSIC method. This will be shown in the rest of this section.

The subspace leakage in the two-step root-MUSIC algorithm for the case of the uncorrelated sources is investigated next. The expected value of the subspace leakage is estimated using (23) and the Monte Carlo simulations. The approximate value for the subspace leakage is also obtained from the theoretical derivations in (27) and (29). The value of γ is fixed at 0.5. The results are shown in Fig. 2. The solid lines represent the subspace leakage at the first step, and the dashed lines depict the subspace leakage at the second step of the proposed two-step root-MUSIC algorithm. It can be seen that the curves obtained from the simulations are very close to those obtained from our theoretical derivations at high SNR values. At the low SNR region, the curve associated with the theoretical approximation at the second step deviates from the curve obtained by simulations. The reason is that in the derivations, the first order Taylor series expansion is used. More accurate results can be obtained by using higher order Taylor series. However, the computations can become intractable. In Fig. 2, it can be observed from both theoretical and simulation results that the subspace leakage from the modified covariance matrix at the second step is significantly smaller than the subspace leakage from the sample data covariance matrix at the first step. This is achieved by removing the undesirable terms from the sample data covariance matrix leading to an estimate of the signal projection matrix that is closer to the true signal projection matrix, which is equivalent to a lower subspace leakage at the second step.

We next consider the performance of the proposed two-step algorithm when applied to the root-MUSIC [2], unitary root-MUSIC [22], improved unitary root-MUSIC with pseudo-noise resampling [12], root-swap unitary root-MUSIC, and root-swap unitary root-MUSIC with pseudo-noise resampling methods. The unitary root-MUSIC algorithm takes benefit from the forward-backward averaging [7] which is approximately equivalent to doubling the number of samples. For the cases that the pseudo-noise resampling is used, P represents the number of times that the resampling process has been performed. In the figures, the root-MUSIC, unitary root-MUSIC, and root-swap unitary root-MUSIC methods are denoted by R-MUSIC, UR-MUSIC, and RSUR-MUSIC, respectively. The value of the scaling factor γ is obtained by minimizing the SML function as described in the two-step root-MUSIC method. In the root-swap algorithm, the parameters p and q are set to $p = 1$ and $q = 0$, which means the closest root to the unit circle is picked up and paired with other roots one at a time in order to find the pair of DOA estimates that minimizes the SML function. In this case, the number of different possible combinations of the

Fig. 2. Subspace leakage versus SNR for uncorrelated sources. The solid and dashed lines represent the subspace leakage at the first and second steps of the proposed two-step root-MUSIC algorithm, respectively.

roots is $N_r = 8$. The number of samples used for the pseudo-noise resampling method is set to $P = 50$. According to our simulations, using more number of samples would not yield in any considerable improvement in the performance.

The MSE versus SNR performance of the methods tested for the case of the uncorrelated sources is presented in Fig. 3. The corresponding CRB [23] is also shown in the figure. For the R-MUSIC method, the modification of the covariance matrix in the second step of the introduced two-step method shifts the MSE curve by almost half a dB to the left. For the UR-MUSIC method the improvement is more significant and is about one dB. For the rest of the methods, there is no considerable change in the MSE performance. However, as it will be

Fig. 3. MSE versus SNR for uncorrelated sources. The solid and dashed lines are based on the first and second steps of the proposed two-step method, respectively. The methods used in the two-step algorithm are R-MUSIC, UR-MUSIC, and RSUR-MUSIC methods. P is the number of samples used for the pseudo-noise resampling algorithm.

shown in the next figures, the modification of the covariance matrix has benefits in terms of the CMSE performance and probability of source resolution for these methods. It can also be seen from Fig. 3 that the proposed RSUR-MUSIC algorithm performs about 2 dB better than the UR-MUSIC method, while imposing only a small amount of computational complexity for evaluating the SML function for $N_r = 8$ different combinations of the roots. The best performance is achieved by the RSUR-MUSIC algorithm combined with the pseudo-noise resampling method.

Fig. 4 shows probability of source resolution versus SNR for the uncorrelated sources. For the R-MUSIC method, the second step of the two-step algorithm improves the performance by 1 to

Fig. 4. Probability of source resolution versus SNR for uncorrelated sources. The solid and dashed lines are based on the first and second steps of the proposed two-step method, respectively. The methods used in the two-step algorithm are R-MUSIC, UR-MUSIC, and RSUR-MUSIC methods.

2 dB. The rest of the algorithms have almost the same performance with the root-swap based methods slightly outperforming the other algorithms at low SNR values. It is observed that the second step of the two-step algorithm results in about 1 dB improvement in the performance.

Finally, Fig. 5 illustrates the performance of the algorithms tested for the uncorrelated sources in terms of the CMSE. The R-MUSIC method is significantly improved by the two-step method with an improvement ranging from 5 dB at low SNR values to 1 dB at high SNR values. The rest of the algorithms show similar performance, and the application of the two-step method leads to up to 2 dB improvement in the CMSE performance.

Fig. 5. CMSE versus SNR for uncorrelated sources. The solid and dashed lines are based on the first and second steps of the proposed two-step method, respectively. The methods used in the two-step algorithm are R-MUSIC, UR-MUSIC, and RSUR-MUSIC methods.

The results for the case of the correlated sources with $r = 0.9$ are depicted in Figs. 6 to 10. Similar observations are made from these figures as those discussed for the case of the uncorrelated sources. Compared to the uncorrelated case, the performance breakdown occurs at a higher SNR value. This makes the importance of the improved methods more significant, as there is a higher chance that the actual SNR of a system falls in the breakdown region. As seen from the figures for the correlated sources, the proposed methods prove to be helpful in dealing with the performance breakdown problem. The gain obtained by the improved methods is also more significant compared to the case of the uncorrelated sources. For instance, the MSE

Fig. 6.   Probability of root-swap and probability of ML failure versus SNR for correlated sources with $r = 0.9$.

improvement achieved by the two-step root-MUSIC method for the uncorrelated sources is about half a dB, while in the case of the correlated sources, the MSE curve is shifted by more than 2 dB to the left. Similarly, more significant performance gains are obtained for the probability of source resolution and also the CMSE.

VI.   CONCLUSION

The performance breakdown of the subspace based DOA estimation methods in the threshold region where the SNR and/or sample size is low has been studied in this paper. The subspace leakage as the main cause of the performance breakdown was formally defined and theoretically derived. The two-step algorithm has been proposed in order to reduce the amount of subspace

Fig. 7.   Subspace leakage versus SNR for correlated sources with $r = 0.9$. The solid and dashed lines represent the subspace leakage at the first and second steps of the proposed two-step R-MUSIC algorithm, respectively.

leakage. The introduced method is based on estimating the DOAs at the first step and modifying the covariance matrix using the estimated DOAs at the second step. We have theoretically derived the subspace leakage at both steps, and have shown that the subspace leakage is reduced at the second step of the proposed method leading to better performance. The algorithm can also be extended to the third step by further modifying the covariance matrix based on the improved estimates obtained at the second step. We have investigated the performance of the algorithm for further steps through simulations (not included in the paper). However, the achieved improvement is marginal and does not justify the added complexity. The behavior of the root-MUSIC algorithm in the threshold region has been also studied, and a phenomenon called

Fig. 8. MSE versus SNR for correlated sources with $r = 0.9$. The solid and dashed lines are based on the first and second steps of the proposed two-step method, respectively. The methods used in the two-step algorithm are R-MUSIC, UR-MUSIC, and RSUR-MUSIC methods.

root-swap has been observed to contribute to the performance breakdown. Then, an improved method has been introduced to remedy this problem by considering different combinations of the roots and picking up the one that minimizes the SML function. The performance improvement achieved by the proposed methods has also been demonstrated using numerical examples and simulation results. We also combined the proposed algorithms with the previously introduced methods in the literature, which resulted in further improvement in the performance.

Fig. 9. Probability of source resolution versus SNR for correlated sources with $r = 0.9$. The solid and dashed lines are based on the first and second steps of the proposed two-step method, respectively. The methods used in the two-step algorithm are R-MUSIC, UR-MUSIC, and RSUR-MUSIC methods.

APPENDIX A

PROBABILITY OF ROOT-SWAP APPROXIMATION

The root-swap is defined as the event when at least one of the estimated signal roots $\hat{z}_k$ $(1 \leq k \leq K)$ has a smaller magnitude than the magnitude of an estimated noise root $\hat{z}_m$ $(K+1 \leq m \leq M-1)$, i.e., $\hat{r}_k < \hat{r}_m$. Let us denote the probability of the event that $\hat{r}_k < \hat{r}_m$ by p_{km}. The complement of this event represents the case when the k-th estimated signal root has not been swapped with the m-th estimated noise root, and its probability is given by $1 - p_{km}$. Let us denote the probability of root-swap by $P(\text{root-swap})$. The complement of the root-swap event

Fig. 10. CMSE versus SNR for the correlated sources with $r = 0.9$. The solid and dashed lines are based on the first and second steps of the proposed two-step method, respectively. The methods used in the two-step algorithm are R-MUSIC), UR-MUSIC, and RSUR-MUSIC methods.

is the event when none of the estimated signal roots has been swapped with an estimated noise root, and its probability is given by $1 - P(\text{root-swap})$. Assuming that the individual root-swap events are independent from each other, we have

$$1 - P(\text{root-swap}) = \prod_{k=1}^{K} \prod_{m=K+1}^{M-1} \left(1 - p_{km}\right).$$ (31)

In the sequel, we derive an approximation for p_{km}. Noting that $r_k = 1$ for the true signal

roots, we have

$$p_{km} = P\left(\hat{r}_m > \hat{r}_k\right)$$

$$= P\left(\Delta r_m - \Delta r_k > 1 - r_m\right). \tag{32}$$

In order to proceed with the computation of p_{km}, we consider the distributions of Δr_m and Δr_k. It is shown in [19] that Δr_k $(1 \leq k \leq K)$ follows the $-\left(\sigma_k/\sqrt{2}\right)\sqrt{\chi^2\left(2(M-K)-1\right)}$ distribution where $\chi^2\left(\ell\right)$ denotes a chi-square distribution with ℓ degrees of freedom and σ_k^2 is given by

$$\sigma_k^2 = \frac{\sigma_{\mathrm{n}}^2}{N\left(\boldsymbol{a}_k^{(1)H}\boldsymbol{P}^\perp\boldsymbol{a}_k^{(1)}\right)}\sum_{i=1}^{K}\frac{\lambda_{M-K+i}}{\left(\lambda_{M-K+i}-\sigma_{\mathrm{n}}^2\right)^2}\left|\boldsymbol{e}_i^H\boldsymbol{a}_k\right|^2 \tag{33}$$

where $\boldsymbol{P}^\perp$ is the true projection matrix of the noise subspace and $\boldsymbol{a}_k^{(1)}$ is given by (30).

We next consider the distribution of Δr_m. In [19], the distribution of Δr_k is computed using a second order Taylor expansion of the estimated root-MUSIC polynomial around the true signal roots (which are located on the unit circle). The computation of the distribution of Δr_m requires the analysis to be performed around the true noise roots which are located inside the unit circle. The second order expansions of $\boldsymbol{a}(\hat{z}_k)$ and $\boldsymbol{a}^T(\hat{z}_k^{-1})$ around the true signal root z_k are given by [19]

$$\boldsymbol{a}(\hat{z}_k) \approx \boldsymbol{a}_k + j\boldsymbol{a}_k^{(1)}\Delta\omega_k + \boldsymbol{a}_k^{(1)}\Delta r_k$$

$$\boldsymbol{a}^T(\hat{z}_k^{-1}) \approx \boldsymbol{a}_k^H - j\boldsymbol{a}_k^{(1)H}\Delta\omega_k - \boldsymbol{a}_k^{(1)H}\Delta r_k \tag{34}$$

where $\Delta\omega_k$ is the difference between the angle of the k-th estimated root and the angle of the corresponding true root. For the m-th noise root, let $\boldsymbol{a}_m$ be defined as

$$\boldsymbol{a}_m \triangleq \left[1,\ e^{-j\omega_m},\cdots,\ e^{-j(M-1)\omega_m}\right]^T \tag{35}$$

where ω_m is the angle of z_m. Let also $\boldsymbol{a}_m^{(1)}$ be defined similar to (30) with ω_k replaced with ω_m. Then, the second order expansions of $\boldsymbol{a}(\hat{z}_m)$ and $\boldsymbol{a}^T(\hat{z}_m^{-1})$ around the true noise root z_m are given by

$$\boldsymbol{a}(\hat{z}_m) \approx \boldsymbol{R}_m^{-1}\left(\boldsymbol{a}_m + j\boldsymbol{a}_m^{(1)}\Delta\omega_m + \boldsymbol{a}_m^{(1)}\left(\frac{\Delta r_m}{r_m}\right)\right)$$

$$\boldsymbol{a}^T(\hat{z}_m^{-1}) \approx \left(\boldsymbol{a}_m^H - j\boldsymbol{a}_m^{(1)H}\Delta\omega_m - \boldsymbol{a}_m^{(1)H}\left(\frac{\Delta r_m}{r_m}\right)\right)\boldsymbol{R}_m \tag{36}$$

where $\boldsymbol{R}_m$ is a $M \times M$ diagonal matrix with its diagonal elements equal to $1, r_m, \cdots, r_m^{(M-1)}$. Since the Taylor expansion for the steering vectors of the roots on the circle and the expansion for the roots inside the circle, i.e., (34) and (36) have similar structures, it is reasonable to assume that Δr_k and $\Delta r_m / r_m$ also have similar distributions. Then, the variance of Δr_m is in the order of the variance of Δr_k multiplied by r_m^2. Since $r_m < 1$, the variance of Δr_m is smaller than the variance of Δr_k. In order to simplify the computation of p_{km}, we ignore the effect of Δr_m and approximate p_{km} by

$$p_{km} \approx P\left(-\Delta r_k > 1 - r_m\right). \tag{37}$$

This is equivalent to using the probability $P\left(\hat{r}_k < r_m\right)$ as an approximation for p_{km}. Since we have the distribution of Δr_k, we can compute p_{km} using (37). When $M - K \gg 1$, Δr_k follows approximately a normal distribution $\mathcal{N}\left(-\sigma_k \sqrt{M - K - (3/4)}, \sigma_k^2/4\right)$ [19]. Using (37), the probability p_{km} can be approximated by

$$p_{km} \approx Q\left(\frac{1 - r_m - \sigma_k \sqrt{M - K - (3/4)}}{\sqrt{\sigma_k^2/4}}\right). \tag{38}$$

Finally, the approximation of the probability of root-swap $P(\text{root-swap})$ is found by using the approximation (38), the expression (31), and the fact that $Q(-x) = 1 - Q(x)$ as

$$P(\text{root-swap}) \approx 1 - \prod_{k=1}^{K} \prod_{m=K+1}^{M-1} Q\left(\frac{-1 + r_m + \sigma_k \sqrt{M - K - (3/4)}}{\sqrt{\sigma_k^2/4}}\right). \tag{39}$$

It completes the derivation.

APPENDIX B
SUBSPACE LEAKAGE AT THE FIRST STEP

Let us start with the computation of ρ_1. 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 $\text{Tr}\{\boldsymbol{P}\} = K$, the expression (23) for the first step of the two-step root-MUSIC algorithm can be written as

$$\begin{aligned}
\rho_1 &= 1 - \frac{1}{K}\text{Tr}\left\{(\boldsymbol{P} + \Delta \boldsymbol{P})\,\boldsymbol{P}\right\} \\
&= 1 - \frac{1}{K}\left(K + \text{Tr}\{\Delta \boldsymbol{P} \boldsymbol{P}\}\right) \\
&= -\frac{1}{K}\text{Tr}\{\Delta \boldsymbol{P} \boldsymbol{P}\}.
\end{aligned} \tag{40}$$

It is shown in [19] that the series expansion of $\widehat{\boldsymbol{P}}$ based on $\Delta\boldsymbol{R}$ is given by

$$\widehat{\boldsymbol{P}} = \boldsymbol{P} + \delta\boldsymbol{P} + \cdots + \delta^n\boldsymbol{P} + \cdots \tag{41}$$

where

$$\delta\boldsymbol{P} = \boldsymbol{P}^\perp\Delta\boldsymbol{R}\boldsymbol{V}^\dagger + \boldsymbol{V}^\dagger\Delta\boldsymbol{R}\boldsymbol{P}^\perp \tag{42}$$

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

$$
\begin{aligned}
\delta^n\boldsymbol{P} = &-\boldsymbol{P}^\perp\left(\delta^{n-1}\boldsymbol{P}\right)\Delta\boldsymbol{R}\boldsymbol{V}^\dagger + \boldsymbol{P}^\perp\Delta\boldsymbol{R}\left(\delta^{n-1}\boldsymbol{P}\right)\boldsymbol{V}^\dagger \\
&-\boldsymbol{V}^\dagger\Delta\boldsymbol{R}\left(\delta^{n-1}\boldsymbol{P}\right)\boldsymbol{P}^\perp + \boldsymbol{V}^\dagger\left(\delta^{n-1}\boldsymbol{P}\right)\Delta\boldsymbol{R}\boldsymbol{P}^\perp \\
&-\sum_{i=1}^{n-1}\boldsymbol{P}\left(\delta^i\boldsymbol{P}\right)\left(\delta^{n-i}\boldsymbol{P}\right)\boldsymbol{P} \\
&+\sum_{i=1}^{n-1}\boldsymbol{P}^\perp\left(\delta^i\boldsymbol{P}\right)\left(\delta^{n-i}\boldsymbol{P}\right)\boldsymbol{P}^\perp.
\end{aligned} \tag{43}
$$

The following lemma will be further used.

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 proof follows by multiplying $\boldsymbol{P}$ by $\boldsymbol{V}^\dagger$ and then substituting $\boldsymbol{P}$ with $\boldsymbol{E}\boldsymbol{E}^H$ and $\boldsymbol{V}^\dagger$ with (25). ∎

In a similar way to Lemma 1, it can also be shown that

$$\boldsymbol{V}\boldsymbol{V}^\dagger = \boldsymbol{V}^\dagger\boldsymbol{V} = \boldsymbol{P}. \tag{44}$$

Using (40), the series expansion of $\widehat{\boldsymbol{P}}$ in (41), expressions (42) and (43) up to the $\delta^2\boldsymbol{P}$ term, and the facts that $\boldsymbol{P}\boldsymbol{P}^\perp = \boldsymbol{P}^\perp\boldsymbol{P} = \boldsymbol{0}$ and $\boldsymbol{P}\boldsymbol{P} = \boldsymbol{P}$, we can write ρ_1 as

$$\rho_1 = -\frac{1}{K}\mathrm{Tr}\left\{-\boldsymbol{P}\left(\delta\boldsymbol{P}\right)\left(\delta\boldsymbol{P}\right)\right\}. \tag{45}$$

Then, ρ_1 is computed by substituting (42) in (45), using $\boldsymbol{P}^\perp\boldsymbol{P}^\perp = \boldsymbol{P}^\perp$, and Lemma 1 as

$$
\begin{aligned}
\rho_1 &= \frac{1}{K}\mathrm{Tr}\left\{\boldsymbol{P}\left(\boldsymbol{P}^\perp\Delta\boldsymbol{R}\boldsymbol{V}^\dagger + \boldsymbol{V}^\dagger\Delta\boldsymbol{R}\boldsymbol{P}^\perp\right)\left(\boldsymbol{P}^\perp\Delta\boldsymbol{R}\boldsymbol{V}^\dagger + \boldsymbol{V}^\dagger\Delta\boldsymbol{R}\boldsymbol{P}^\perp\right)\right\} \\
&= \frac{1}{K}\mathrm{Tr}\left\{\boldsymbol{P}\boldsymbol{V}^\dagger\Delta\boldsymbol{R}\boldsymbol{P}^\perp\boldsymbol{P}^\perp\Delta\boldsymbol{R}\boldsymbol{V}^\dagger\right\} \\
&= \frac{1}{K}\mathrm{Tr}\left\{\boldsymbol{V}^\dagger\Delta\boldsymbol{R}\boldsymbol{P}^\perp\Delta\boldsymbol{R}\boldsymbol{V}^\dagger\right\}.
\end{aligned} \tag{46}
$$

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 [19].

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

$$E\left\{\Delta\boldsymbol{R}\boldsymbol{A}_1\Delta\boldsymbol{R}\right\} = \frac{1}{N}Tr\left\{\boldsymbol{R}\boldsymbol{A}_1\right\}\boldsymbol{R} \tag{47}$$

and

$$E\left\{Tr\left\{\Delta\boldsymbol{R}\boldsymbol{A}_1\right\}Tr\left\{\Delta\boldsymbol{R}\boldsymbol{A}_2\right\}\right\} = \frac{1}{N}Tr\left\{\boldsymbol{R}\boldsymbol{A}_1\boldsymbol{R}\boldsymbol{A}_2\right\}. \tag{48}$$

Using (46) and (47), the expected value of ρ_1 can be computed as

$$\begin{aligned}
E\left\{\rho_1\right\} &= \frac{1}{K}\mathrm{Tr}\left\{\boldsymbol{V}^\dagger E\left\{\Delta\boldsymbol{R}\boldsymbol{P}^\perp\Delta\boldsymbol{R}\right\}\boldsymbol{V}^\dagger\right\} \\
&= \frac{1}{K}\mathrm{Tr}\left\{\boldsymbol{V}^\dagger\frac{1}{N}\mathrm{Tr}\left\{\boldsymbol{R}\boldsymbol{P}^\perp\right\}\boldsymbol{R}\boldsymbol{V}^\dagger\right\} \\
&= \frac{1}{NK}\mathrm{Tr}\left\{\boldsymbol{P}^\perp\boldsymbol{R}\right\}\mathrm{Tr}\left\{\boldsymbol{V}^\dagger\boldsymbol{V}^\dagger\boldsymbol{R}\right\}.
\end{aligned} \tag{49}$$

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, $\mathrm{Tr}\left\{\boldsymbol{P}^\perp\boldsymbol{R}\right\}$ can be simplified as

$$\begin{aligned}
\mathrm{Tr}\left\{\boldsymbol{P}^\perp\boldsymbol{R}\right\} &= \mathrm{Tr}\left\{\boldsymbol{P}^\perp\left(\boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^H + \sigma_{\mathrm{n}}^2\boldsymbol{I}_M\right)\right\} \\
&= \mathrm{Tr}\left\{\sigma_{\mathrm{n}}^2\boldsymbol{P}^\perp\right\} = \sigma_{\mathrm{n}}^2\mathrm{Tr}\left\{\boldsymbol{I}_M - \boldsymbol{P}\right\} \\
&= \sigma_{\mathrm{n}}^2\left(M - K\right).
\end{aligned} \tag{50}$$

Furthermore, using (25) and the fact that the eigenvectors of $\boldsymbol{R}$ are orthonormal, the product $\boldsymbol{V}^\dagger\boldsymbol{V}^\dagger\boldsymbol{R}$ can be written as

$$\boldsymbol{V}^\dagger\boldsymbol{V}^\dagger\boldsymbol{R} = \sum_{k=1}^{K}\frac{\lambda_{M-K+k}}{\left(\lambda_{M-K+k} - \sigma_{\mathrm{n}}^2\right)^2}\boldsymbol{e}_k\boldsymbol{e}_k^H \tag{51}$$

which results in

$$\mathrm{Tr}\left\{\boldsymbol{V}^\dagger\boldsymbol{V}^\dagger\boldsymbol{R}\right\} = \sum_{k=1}^{K}\frac{\lambda_{M-K+k}}{\left(\lambda_{M-K+k} - \sigma_{\mathrm{n}}^2\right)^2}. \tag{52}$$

Finally, $E\left\{\rho_1\right\}$ is obtained by substituting (50) and (52) in (49) as

$$E\left\{\rho_1\right\} = \frac{\sigma_{\mathrm{n}}^2\left(M - K\right)}{NK}\sum_{k=1}^{K}\frac{\lambda_{M-K+k}}{\left(\lambda_{M-K+k} - \sigma_{\mathrm{n}}^2\right)^2}. \tag{53}$$

APPENDIX C

SUBSPACE LEAKAGE AT THE SECOND STEP

The subspace leakage at the second step of the two-step root-MUSIC algorithm can be obtained through the same steps taken for the computation of ρ_1. Referring to (46), the subspace leakage ρ_2 is given by

$$\rho_2 = \frac{1}{K}\text{Tr}\left\{\boldsymbol{V}^\dagger \Delta \boldsymbol{R}^{(2)} \boldsymbol{P}^\perp \Delta \boldsymbol{R}^{(2)} \boldsymbol{V}^\dagger\right\} \tag{54}$$

where $\Delta \boldsymbol{R}^{(2)} \triangleq \widehat{\boldsymbol{R}}^{(2)} - \boldsymbol{R}$ is the estimation error of the covariance matrix at the second step of the algorithm. Using (13), the estimation error $\Delta \boldsymbol{R}^{(2)}$ is given by

$$\Delta \boldsymbol{R}^{(2)} = \Delta \boldsymbol{R} - \gamma \left(\boldsymbol{T} + \boldsymbol{T}^H\right). \tag{55}$$

Recalling (10), we have $\boldsymbol{T} = \widehat{\boldsymbol{P}}_A \widehat{\boldsymbol{R}} \widehat{\boldsymbol{P}}_A^\perp$.

Consider the first order Taylor series expansion of $\widehat{\boldsymbol{P}}_A$ around the true DOAs given by

$$\widehat{\boldsymbol{P}}_A \approx \boldsymbol{P}_A + d\boldsymbol{P} \tag{56}$$

where $\boldsymbol{P}_A \triangleq \boldsymbol{A}\left(\boldsymbol{A}^H\boldsymbol{A}\right)^{-1}\boldsymbol{A}^H$ is equal to the true signal projection matrix[1], i.e., $\boldsymbol{P}_A = \boldsymbol{P}$, and $d\boldsymbol{P}$ is given by

$$d\boldsymbol{P} = \sum_{k=1}^{K} \frac{\partial \boldsymbol{P}_A}{\partial \omega_k} \Delta\omega_k. \tag{57}$$

Here $\Delta\omega_k \triangleq \hat{\omega}_k - \omega_k$ is the estimation error of ω_k with $\hat{\omega}_k \triangleq 2\pi(d/\lambda)\sin(\hat{\theta}_k)$.

Note that for any square and invertible matrix $\boldsymbol{B}$, the partial derivative of $\boldsymbol{B}^{-1}$ with respect to the variable ω is given by [24]

$$\frac{\partial \boldsymbol{B}^{-1}}{\partial \omega} = -\boldsymbol{B}^{-1}\frac{\partial \boldsymbol{B}}{\partial \omega}\boldsymbol{B}^{-1}. \tag{58}$$

Using (58), the partial derivative $\partial \boldsymbol{P}_A/\partial \omega_k$ can be computed as

$$\begin{aligned}
\frac{\partial \boldsymbol{P}_A}{\partial \omega_k} &= \frac{\partial \boldsymbol{A}}{\partial \omega_k}\left(\boldsymbol{A}^H\boldsymbol{A}\right)^{-1}\boldsymbol{A}^H + \boldsymbol{A}\frac{\partial\left(\boldsymbol{A}^H\boldsymbol{A}\right)^{-1}}{\partial \omega_k}\boldsymbol{A}^H + \boldsymbol{A}\left(\boldsymbol{A}^H\boldsymbol{A}\right)^{-1}\left(\frac{\partial \boldsymbol{A}}{\partial \omega_k}\right)^H \\
&= \frac{\partial \boldsymbol{A}}{\partial \omega_k}\left(\boldsymbol{A}^H\boldsymbol{A}\right)^{-1}\boldsymbol{A}^H - \boldsymbol{A}\left(\boldsymbol{A}^H\boldsymbol{A}\right)^{-1}\left(\left(\frac{\partial \boldsymbol{A}}{\partial \omega_k}\right)^H\boldsymbol{A} + \boldsymbol{A}^H\frac{\partial \boldsymbol{A}}{\partial \omega_k}\right)\left(\boldsymbol{A}^H\boldsymbol{A}\right)^{-1}\boldsymbol{A}^H \\
&\quad + \boldsymbol{A}\left(\boldsymbol{A}^H\boldsymbol{A}\right)^{-1}\left(\frac{\partial \boldsymbol{A}}{\partial \omega_k}\right)^H. \tag{59}
\end{aligned}$$

---

[1]Note that although $\boldsymbol{P}_A$ is equal to $\boldsymbol{P}$, the estimates $\widehat{\boldsymbol{P}}_A$ and $\widehat{\boldsymbol{P}}$ are obtained in different ways and are not essentially equal to each other.

Then, using (20) and $\boldsymbol{P} = \boldsymbol{A} \left(\boldsymbol{A}^H \boldsymbol{A} \right)^{-1} \boldsymbol{A}^H$, the partial derivative $\partial \boldsymbol{P}_A / \partial \omega_k$ is given by

$$\frac{\partial \boldsymbol{P}_A}{\partial \omega_k} = \boldsymbol{P}^\perp \frac{\partial \boldsymbol{A}}{\partial \omega_k} \left(\boldsymbol{A}^H \boldsymbol{A} \right)^{-1} \boldsymbol{A}^H + \boldsymbol{A} \left(\boldsymbol{A}^H \boldsymbol{A} \right)^{-1} \left(\frac{\partial \boldsymbol{A}}{\partial \omega_k} \right)^H \boldsymbol{P}^\perp. \tag{60}$$

The estimation error of ω_k, i.e., $\Delta\omega_k$ in (57), can be written based on $\Delta\boldsymbol{R}$ as [19]

$$\Delta\omega_k = \frac{\boldsymbol{a}_k^{(1)H} \boldsymbol{P}^\perp \Delta\boldsymbol{R} \boldsymbol{V}^\dagger \boldsymbol{a}_k - \boldsymbol{a}_k^H \boldsymbol{V}^\dagger \Delta\boldsymbol{R} \boldsymbol{P}^\perp \boldsymbol{a}_k^{(1)}}{2j \left(\boldsymbol{a}_k^{(1)H} \boldsymbol{P}^\perp \boldsymbol{a}_k^{(1)} \right)}. \tag{61}$$

The first order Taylor series expansion of $\widehat{\boldsymbol{P}}_A^\perp$ is obtained using (12) and (56) as

$$\widehat{\boldsymbol{P}}_A^\perp \approx \boldsymbol{P}_A^\perp - d\boldsymbol{P} \tag{62}$$

where $\boldsymbol{P}_A^\perp \triangleq \boldsymbol{I}_M - \boldsymbol{P}_A$.

The matrix $\boldsymbol{T}$ can be then computed using expressions (10), (56), and (62) with keeping only the first order terms and noting that $\boldsymbol{P}_A = \boldsymbol{P}$, $\boldsymbol{P}_A^\perp = \boldsymbol{P}^\perp$, and $\boldsymbol{P}\boldsymbol{R}\boldsymbol{P}^\perp = \boldsymbol{0}$ as

$$\boldsymbol{T} = \left(\boldsymbol{P}_A + d\boldsymbol{P} \right) \left(\boldsymbol{R} + \Delta R \right) \left(\boldsymbol{P}_A^\perp - d\boldsymbol{P} \right)$$

$$\approx -\boldsymbol{P}\boldsymbol{R}d\boldsymbol{P} + \boldsymbol{P}\Delta R\boldsymbol{P}^\perp + d\boldsymbol{P}\boldsymbol{R}\boldsymbol{P}^\perp. \tag{63}$$

We can now compute ρ_2 using expressions (54), (55), and (63) as

$$\rho_2 = \frac{1}{K} \mathrm{Tr} \Big\{ \boldsymbol{V}^\dagger \left(\Delta\boldsymbol{R} - \gamma \left(\boldsymbol{T} + \boldsymbol{T}^H \right) \right) \boldsymbol{P}^\perp \left(\Delta\boldsymbol{R} - \gamma \left(\boldsymbol{T} + \boldsymbol{T}^H \right) \right) \boldsymbol{V}^\dagger \Big\}$$

$$= \frac{1}{K} \mathrm{Tr} \Big\{ \boldsymbol{V}^\dagger \big(\Delta\boldsymbol{R} - \gamma \big(-\boldsymbol{P}\boldsymbol{R}d\boldsymbol{P} + \boldsymbol{P}\Delta R\boldsymbol{P}^\perp + d\boldsymbol{P}\boldsymbol{R}\boldsymbol{P}^\perp - d\boldsymbol{P}\boldsymbol{R}\boldsymbol{P} + \boldsymbol{P}^\perp\Delta R\boldsymbol{P}$$

$$+ \boldsymbol{P}^\perp\boldsymbol{R}d\boldsymbol{P} \big) \big) \boldsymbol{P}^\perp \big(\Delta\boldsymbol{R} - \gamma \big(-\boldsymbol{P}\boldsymbol{R}d\boldsymbol{P} + \boldsymbol{P}\Delta R\boldsymbol{P}^\perp + d\boldsymbol{P}\boldsymbol{R}\boldsymbol{P}^\perp$$

$$- d\boldsymbol{P}\boldsymbol{R}\boldsymbol{P} + \boldsymbol{P}^\perp\Delta R\boldsymbol{P} + \boldsymbol{P}^\perp\boldsymbol{R}d\boldsymbol{P} \big) \big) \boldsymbol{V}^\dagger \Big\}. \tag{64}$$

Then, using expressions (57), (60), and the fact that $\boldsymbol{P}\boldsymbol{P}^\perp = \boldsymbol{P}^\perp\boldsymbol{P} = \boldsymbol{V}^\dagger\boldsymbol{P}^\perp = \boldsymbol{P}^\perp\boldsymbol{V}^\dagger = \boldsymbol{0}$ to eliminate the terms that equal zero, ρ_2 is computed as

$$\rho_2 = \frac{1}{K} \mathrm{Tr} \Big\{ \boldsymbol{V}^\dagger \big(\Delta\boldsymbol{R} - \gamma \big(-\boldsymbol{P}\boldsymbol{R}d\boldsymbol{P} + \boldsymbol{P}\Delta R\boldsymbol{P}^\perp + d\boldsymbol{P}\boldsymbol{R}\boldsymbol{P}^\perp \big) \big)$$

$$\times \boldsymbol{P}^\perp \big(\Delta\boldsymbol{R} - \gamma \big(-d\boldsymbol{P}\boldsymbol{R}\boldsymbol{P} + \boldsymbol{P}^\perp\Delta R\boldsymbol{P} + \boldsymbol{P}^\perp\boldsymbol{R}d\boldsymbol{P} \big) \big) \boldsymbol{V}^\dagger \Big\}. \tag{65}$$

Expanding the terms in (65) and using the fact that $PV^\dagger = V^\dagger P = V^\dagger$ results in the following expression for ρ_2

$$\rho_2 = \frac{1}{K}\mathrm{Tr}\Big\{ V^\dagger \Delta R P^\perp \Delta R V^\dagger - \gamma\big(-V^\dagger \Delta R P^\perp dP R V^\dagger + V^\dagger \Delta R P^\perp \Delta R V^\dagger$$
$$+ V^\dagger \Delta R P^\perp R dP V^\dagger - V^\dagger R dP P^\perp \Delta R V^\dagger + V^\dagger \Delta R P^\perp \Delta R V^\dagger$$
$$+ V^\dagger dP R P^\perp \Delta R V^\dagger\big) + \gamma^2\big(V^\dagger R dP P^\perp dP R V^\dagger - V^\dagger R dP P^\perp \Delta R V^\dagger$$
$$- V^\dagger R dP P^\perp R dP V^\dagger - V^\dagger \Delta R P^\perp dP R V^\dagger + V^\dagger \Delta R P^\perp \Delta R V^\dagger$$
$$+ V^\dagger \Delta R P^\perp R dP V^\dagger - V^\dagger dP R P^\perp dP R V^\dagger + V^\dagger dP R P^\perp \Delta R V^\dagger$$
$$+ V^\dagger dP R P^\perp R dP V^\dagger\big)\Big\}.$$

(66)

By reordering the terms in (66), the subspace leakage ρ_2 can be further rewritten as

$$\rho_2 = \frac{1}{K}\mathrm{Tr}\Big\{ \big(1 - 2\gamma + \gamma^2\big) V^\dagger \Delta R P^\perp \Delta R V^\dagger + \big(\gamma^2 - \gamma\big)\big(-V^\dagger \Delta R P^\perp dP R V^\dagger$$
$$+ V^\dagger \Delta R P^\perp R dP V^\dagger - V^\dagger R dP P^\perp \Delta R V^\dagger + V^\dagger dP R P^\perp \Delta R V^\dagger\big)$$
$$+ \gamma^2\big(V^\dagger R dP P^\perp dP R V^\dagger - V^\dagger R dP P^\perp R dP V^\dagger - V^\dagger dP R P^\perp dP R V^\dagger$$
$$+ V^\dagger dP R P^\perp R dP V^\dagger\big)\Big\}.$$

(67)

The terms multiplied by $(\gamma^2 - \gamma)$ in (67) can be simplified using expressions (24), (44), and the fact that $P^\perp V = 0$ as

$$-V^\dagger \Delta R P^\perp dP\big(V + \sigma_\mathrm{n}^2 I_M\big)V^\dagger + V^\dagger \Delta R P^\perp\big(V + \sigma_\mathrm{n}^2 I_M\big)dP V^\dagger$$
$$-V^\dagger\big(V + \sigma_\mathrm{n}^2 I_M\big)dP P^\perp \Delta R V^\dagger + V^\dagger dP\big(V + \sigma_\mathrm{n}^2 I_M\big)P^\perp \Delta R V^\dagger$$
$$= -V^\dagger \Delta R P^\perp dP P - P dP P^\perp \Delta R V^\dagger.$$

(68)

In a similar way, the terms multiplied by γ^2 in (67) can be simplified to

$$V^\dagger R dP P^\perp dP\big(V + \sigma_\mathrm{n}^2 I_M\big)V^\dagger - V^\dagger R dP P^\perp\big(V + \sigma_\mathrm{n}^2 I_M\big)dP V^\dagger$$
$$-V^\dagger dP R P^\perp dP\big(V + \sigma_\mathrm{n}^2 I_M\big)V^\dagger + V^\dagger dP R P^\perp\big(V + \sigma_\mathrm{n}^2 I_M\big)dP V^\dagger$$
$$= V^\dagger R dP P^\perp dP P - V^\dagger dP R P^\perp dP P$$
$$= V^\dagger\big(V + \sigma_\mathrm{n}^2 I_M\big)dP P^\perp dP P - V^\dagger dP\big(V + \sigma_\mathrm{n}^2 I_M\big)P^\perp dP P$$
$$= P dP P^\perp dP P$$

(69)

which using the fact that $\boldsymbol{P}^\perp d\boldsymbol{P}\boldsymbol{P}^\perp = \boldsymbol{0}$ (see (57) and (60)) can be further simplified to

$$\boldsymbol{P}d\boldsymbol{P}\boldsymbol{P}^\perp d\boldsymbol{P}\boldsymbol{P} = \left(\boldsymbol{I}_M - \boldsymbol{P}^\perp\right) d\boldsymbol{P}\boldsymbol{P}^\perp d\boldsymbol{P}\left(\boldsymbol{I}_M - \boldsymbol{P}^\perp\right)$$

$$= d\boldsymbol{P}\boldsymbol{P}^\perp d\boldsymbol{P}. \tag{70}$$

Finally, using expressions (46), (67), (68), (70), and Lemma 1, the subspace leakage ρ_2 is computed as

$$\rho_2 = \left(1 - 2\gamma + \gamma^2\right)\rho_1 + \frac{2\left(\gamma - \gamma^2\right)}{K}Re\left\{\mathrm{Tr}\left\{\boldsymbol{V}^\dagger \Delta \boldsymbol{R}\boldsymbol{P}^\perp d\boldsymbol{P}\right\}\right\} + \frac{\gamma^2}{K}\mathrm{Tr}\left\{d\boldsymbol{P}\boldsymbol{P}^\perp d\boldsymbol{P}\right\}. \tag{71}$$

Computation of the expected value of ρ_2 involves finding the expected value of the two trace functions in (71). Using expressions (57) and (60), the expected value of the first trace function in (71) is given by

$$E\left\{\mathrm{Tr}\left\{\boldsymbol{V}^\dagger \Delta \boldsymbol{R}\boldsymbol{P}^\perp d\boldsymbol{P}\right\}\right\} = E\left\{\mathrm{Tr}\left\{\Delta \boldsymbol{R}\sum_{k=1}^{K}\boldsymbol{P}^\perp \frac{\partial \boldsymbol{A}}{\partial \omega_k}\left(\boldsymbol{A}^H \boldsymbol{A}\right)^{-1}\boldsymbol{A}^H \Delta \omega_k \boldsymbol{V}^\dagger\right\}\right\}. \tag{72}$$

Then, by substituting (61) in (72), we have

$$E\left\{\mathrm{Tr}\left\{\boldsymbol{V}^\dagger \Delta \boldsymbol{R}\boldsymbol{P}^\perp d\boldsymbol{P}\right\}\right\} = E\left\{\mathrm{Tr}\left\{\sum_{k=1}^{K}\Delta \boldsymbol{R}\boldsymbol{P}^\perp \frac{\partial \boldsymbol{A}}{\partial \omega_k}\left(\boldsymbol{A}^H \boldsymbol{A}\right)^{-1}\boldsymbol{A}^H \boldsymbol{V}^\dagger\right.\right.$$

$$\left.\left.\times \frac{1}{2j\left(\boldsymbol{a}_k^{(1)H}\boldsymbol{P}^\perp \boldsymbol{a}_k^{(1)}\right)}\left(\boldsymbol{a}_k^{(1)H}\boldsymbol{P}^\perp \Delta \boldsymbol{R}\boldsymbol{V}^\dagger \boldsymbol{a}_k - \boldsymbol{a}_k^H \boldsymbol{V}^\dagger \Delta \boldsymbol{R}\boldsymbol{P}^\perp \boldsymbol{a}_k^{(1)}\right)\right\}\right\}. \tag{73}$$

The order of the summation and trace operator in (73) can be swapped. Moreover, the last two terms can be written using the trace operator as

$$E\left\{\mathrm{Tr}\left\{\boldsymbol{V}^\dagger \Delta \boldsymbol{R}\boldsymbol{P}^\perp d\boldsymbol{P}\right\}\right\} = E\left\{\sum_{k=1}^{K}\frac{1}{2j\left(\boldsymbol{a}_k^{(1)H}\boldsymbol{P}^\perp \boldsymbol{a}_k^{(1)}\right)}\right.$$

$$\times \mathrm{Tr}\left\{\Delta \boldsymbol{R}\boldsymbol{P}^\perp \frac{\partial \boldsymbol{A}}{\partial \omega_k}\left(\boldsymbol{A}^H \boldsymbol{A}\right)^{-1}\boldsymbol{A}^H \boldsymbol{V}^\dagger\right\}\left(\mathrm{Tr}\left\{\Delta \boldsymbol{R}\boldsymbol{V}^\dagger \boldsymbol{a}_k \boldsymbol{a}_k^{(1)H}\boldsymbol{P}^\perp\right\}\right.$$

$$\left.\left. -\mathrm{Tr}\left\{\Delta \boldsymbol{R}\boldsymbol{P}^\perp \boldsymbol{a}_k^{(1)}\boldsymbol{a}_k^H \boldsymbol{V}^\dagger\right\}\right)\right\}. \tag{74}$$

The expression in (74) can be computed using (48) as

$$
E\left\{\mathrm{Tr}\left\{\boldsymbol{V}^{\dagger}\Delta\boldsymbol{R}\boldsymbol{P}^{\perp}d\boldsymbol{P}\right\}\right\} = \frac{1}{N}\sum_{k=1}^{K}\frac{1}{2j\left(\boldsymbol{a}_{k}^{(1)H}\boldsymbol{P}^{\perp}\boldsymbol{a}_{k}^{(1)}\right)}
$$

$$
\times\left(\mathrm{Tr}\left\{\boldsymbol{R}\boldsymbol{P}^{\perp}\frac{\partial\boldsymbol{A}}{\partial\omega_{k}}\left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{H}\boldsymbol{V}^{\dagger}\boldsymbol{R}\boldsymbol{V}^{\dagger}\boldsymbol{a}_{k}\boldsymbol{a}_{k}^{(1)H}\boldsymbol{P}^{\perp}\right\}\right.
$$

$$
\left.-\mathrm{Tr}\left\{\boldsymbol{R}\boldsymbol{P}^{\perp}\frac{\partial\boldsymbol{A}}{\partial\omega_{k}}\left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{H}\boldsymbol{V}^{\dagger}\boldsymbol{R}\boldsymbol{P}^{\perp}\boldsymbol{a}_{k}^{(1)}\boldsymbol{a}_{k}^{H}\boldsymbol{V}^{\dagger}\right\}\right). \qquad (75)
$$

The second trace function in (75) equals zero as $\boldsymbol{V}^{\dagger}\boldsymbol{R}\boldsymbol{P}^{\perp}=\boldsymbol{0}$. Then, expression (75) can be rewritten as

$$
E\left\{\mathrm{Tr}\left\{\boldsymbol{V}^{\dagger}\Delta\boldsymbol{R}\boldsymbol{P}^{\perp}d\boldsymbol{P}\right\}\right\} = \frac{\sigma_{\mathrm{n}}^{2}}{N}\sum_{k=1}^{K}\frac{\boldsymbol{a}_{k}^{(1)H}\boldsymbol{P}^{\perp}\frac{\partial\boldsymbol{A}}{\partial\omega_{k}}\left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{H}\boldsymbol{V}^{\dagger}\boldsymbol{R}\boldsymbol{V}^{\dagger}\boldsymbol{a}_{k}}{2j\left(\boldsymbol{a}_{k}^{(1)H}\boldsymbol{P}^{\perp}\boldsymbol{a}_{k}^{(1)}\right)} \qquad (76)
$$

where we used the equality $\boldsymbol{P}^{\perp}\boldsymbol{R}=\sigma_{\mathrm{n}}^{2}\boldsymbol{P}^{\perp}$.

In a similar way, using expressions (57) and (60), the expected value of the second trace function in (71) is given by

$$
E\left\{\mathrm{Tr}\left\{d\boldsymbol{P}\boldsymbol{P}^{\perp}d\boldsymbol{P}\right\}\right\} =
$$

$$
E\left\{\mathrm{Tr}\left\{\sum_{k=1}^{K}\sum_{i=1}^{K}\boldsymbol{A}\left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{-1}\left(\frac{\partial\boldsymbol{A}}{\partial\omega_{k}}\right)^{H}\boldsymbol{P}^{\perp}\frac{\partial\boldsymbol{A}}{\partial\omega_{i}}\left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{H}\Delta\omega_{k}\Delta\omega_{i}\right\}\right\}. \qquad (77)
$$

Then, by substituting (61) in (77), we have

$$
E\left\{\mathrm{Tr}\left\{d\boldsymbol{P}\boldsymbol{P}^{\perp}d\boldsymbol{P}\right\}\right\} = E\left\{\mathrm{Tr}\left\{\sum_{k=1}^{K}\sum_{i=1}^{K}\boldsymbol{A}\left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{-1}\left(\frac{\partial\boldsymbol{A}}{\partial\omega_{k}}\right)^{H}\boldsymbol{P}^{\perp}\frac{\partial\boldsymbol{A}}{\partial\omega_{i}}\left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{H}\right.\right.
$$

$$
\times\frac{1}{2j\left(\boldsymbol{a}_{k}^{(1)H}\boldsymbol{P}^{\perp}\boldsymbol{a}_{k}^{(1)}\right)}\times\frac{1}{2j\left(\boldsymbol{a}_{i}^{(1)H}\boldsymbol{P}^{\perp}\boldsymbol{a}_{i}^{(1)}\right)}
$$

$$
\times\left(\mathrm{Tr}\left\{\Delta\boldsymbol{R}\boldsymbol{V}^{\dagger}\boldsymbol{a}_{k}\boldsymbol{a}_{k}^{(1)H}\boldsymbol{P}^{\perp}\right\}-\mathrm{Tr}\left\{\Delta\boldsymbol{R}\boldsymbol{P}^{\perp}\boldsymbol{a}_{k}^{(1)}\boldsymbol{a}_{k}^{H}\boldsymbol{V}^{\dagger}\right\}\right)
$$

$$
\left.\left.\times\left(\mathrm{Tr}\left\{\Delta\boldsymbol{R}\boldsymbol{V}^{\dagger}\boldsymbol{a}_{i}\boldsymbol{a}_{i}^{(1)H}\boldsymbol{P}^{\perp}\right\}-\mathrm{Tr}\left\{\Delta\boldsymbol{R}\boldsymbol{P}^{\perp}\boldsymbol{a}_{i}^{(1)}\boldsymbol{a}_{i}^{H}\boldsymbol{V}^{\dagger}\right\}\right)\right\}\right\}
$$

$$
(78)
$$

which is computed using (48) and the fact that $\boldsymbol{P}^{\perp}\boldsymbol{R}\boldsymbol{V}^{\dagger} = \boldsymbol{0}$ as

$$E\left\{\text{Tr}\left\{d\boldsymbol{P}\boldsymbol{P}^{\perp}d\boldsymbol{P}\right\}\right\} = \frac{\sigma_{\text{n}}^2}{2N}\sum_{k=1}^{K}\sum_{i=1}^{K}\frac{\text{Tr}\left\{\left(\frac{\partial\boldsymbol{A}}{\partial\omega_k}\right)^{H}\boldsymbol{P}^{\perp}\frac{\partial\boldsymbol{A}}{\partial\omega_i}\left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{-1}\right\}}{\left(\boldsymbol{a}_k^{(1)H}\boldsymbol{P}^{\perp}\boldsymbol{a}_k^{(1)}\right)\left(\boldsymbol{a}_i^{(1)H}\boldsymbol{P}^{\perp}\boldsymbol{a}_i^{(1)}\right)}$$
$$\times Re\left\{\boldsymbol{a}_i^{H}\boldsymbol{V}^{\dagger}\boldsymbol{R}\boldsymbol{V}^{\dagger}\boldsymbol{a}_k\boldsymbol{a}_k^{(1)H}\boldsymbol{P}^{\perp}\boldsymbol{a}_i^{(1)}\right\}. \tag{79}$$

Finally, the expected value of ρ_2 for a fixed value of γ is obtained using expressions (71), (76), and (79) as

$$E\left\{\rho_2\right\} = \left(1 - 2\gamma + \gamma^2\right)E\left\{\rho_1\right\}$$
$$+ \frac{2\left(\gamma - \gamma^2\right)\sigma_{\text{n}}^2}{NK}Re\left\{\sum_{k=1}^{K}\frac{\boldsymbol{a}_k^{(1)H}\boldsymbol{P}^{\perp}\frac{\partial\boldsymbol{A}}{\partial\omega_k}\left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{H}\boldsymbol{V}^{\dagger}\boldsymbol{R}\boldsymbol{V}^{\dagger}\boldsymbol{a}_k}{2j\left(\boldsymbol{a}_k^{(1)H}\boldsymbol{P}^{\perp}\boldsymbol{a}_k^{(1)}\right)}\right\}$$
$$+ \frac{\gamma^2\sigma_{\text{n}}^2}{2NK}\sum_{k=1}^{K}\sum_{i=1}^{K}\frac{\text{Tr}\left\{\left(\frac{\partial\boldsymbol{A}}{\partial\omega_k}\right)^{H}\boldsymbol{P}^{\perp}\frac{\partial\boldsymbol{A}}{\partial\omega_i}\left(\boldsymbol{A}^{H}\boldsymbol{A}\right)^{-1}\right\}}{\left(\boldsymbol{a}_k^{(1)H}\boldsymbol{P}^{\perp}\boldsymbol{a}_k^{(1)}\right)\left(\boldsymbol{a}_i^{(1)H}\boldsymbol{P}^{\perp}\boldsymbol{a}_i^{(1)}\right)}Re\left\{\boldsymbol{a}_i^{H}\boldsymbol{V}^{\dagger}\boldsymbol{R}\boldsymbol{V}^{\dagger}\boldsymbol{a}_k\boldsymbol{a}_k^{(1)H}\boldsymbol{P}^{\perp}\boldsymbol{a}_i^{(1)}\right\}. \tag{80}$$

It concludes the derivation.

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] A. J. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP 1983)*, Boston, MA, Apr. 1983, pp. 336–339.

[3] 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.

[4] 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.

[5] B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 4, pp. 397–401, Jul. 1988.

[6] Y. Chen, A. Wiesel, Y. C. Eldar, and A. O. Hero, "Shrinkage algorithms for MMSE covariance estimation," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5016–5028, Oct. 2010.

[7] S. U. Pillai and B. H. Known, "Forward/backward spatial smoothing techniques for coherent signal identification," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 1, pp. 8–15, Jan. 1989.

[8] J. E. Evans, J. R. Johnson, and D. F. Sun, "Application of advanced signal processing techniques to angle of arrival estimation in ATC navigation and surveillance systems," MIT Lincoln Lab., Lexington, MA, Jun. 1982.

[9] X. Mestre and M. A. Lagunas, "Modified subspace algorithms for DOA estimation with large arrays," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 598–614, Feb. 2008.

[10] A. B. Gershman and J. F. Böhme, "Improved DOA estimation via pseudorandom resampling of spatial spectrum," *IEEE Signal Process. Lett.*, vol. 4, no. 2, pp. 54–57, Feb. 1997.

[11] V. Vasylyshyn, "Removing the outliers in root-MUSIC via pseudo-noise resampling and conventional beamformer," *Signal Process.*, vol. 93, no. 12, pp. 3423–3429, Dec. 2013.

[12] C. Qian, L. Huang, and H. C. So, "Improved unitary root-MUSIC for DOA estimation based on pseudo-noise resampling," *IEEE Signal Process. Lett.*, vol. 21, no. 2, pp. 140–144, Feb. 2014.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 10, pp. 1783–1795, Oct. 1990.

[21] M. Shaghaghi and S. A. Vorobyov, "Subspace leakage analysis of sample data covariance matrix," Accepted in *IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP 2015)*, Brisbane, Australia, Apr. 2015.

[22] M. Pesavento, A. B. Gershman, and M. Haardt, "Unitary root-MUSIC with a real-valued eigendecomposition: A theoretical and experimental performance study," *IEEE Trans. Signal Process.*, vol. 48, no. 5, pp. 1306–1314, May 2000.

[23] P. Stoica, E. G. Larsson, and A. B. Gershman, "The stochastic CRB for array processing: a textbook derivation," *IEEE Signal Process. Lett.*, vol. 8, no. 5, pp. 148–150, May 2001.

[24] K. B. Petersen and M. S. Pedersen, *The Matrix Cookbook*, Nov. 2012.