Study of Multi-Step Knowledge-Aided Iterative ESPRIT for Direction Finding

Silvio F. B. Pinto [1] and Rodrigo C. de Lamare [1,2]
Center for Telecommunications Studies (CETUC)
[1] Pontifical Catholic University of Rio de Janeiro, RJ, Brazil.
[2] Department of Electronics, University of York, UK
Emails: silviof@cetuc.puc-rio.br, delamare@cetuc.puc-rio.br

***Abstract*—In this work, we propose a subspace-based algorithm for DOA estimation which iteratively reduces the disturbance factors of the estimated data covariance matrix and incorporates prior knowledge which is gradually obtained on line. An analysis of the MSE of the reshaped data covariance matrix is carried out along with comparisons between computational complexities of the proposed and existing algorithms. Simulations focusing on closely-spaced sources, where they are uncorrelated and correlated, illustrate the improvements achieved.**

I. Introduction

In array signal processing, direction-of-arrival (DOA) estimation is a key task in a broad range of important applications including radar and sonar systems, wireless communications and seismology [1]. Traditional high-resolution methods for DOA estimation such as the multiple signal classification (MUSIC) method [2], the root-MUSIC algorithm [3], the estimation of signal parameters via rotational invariance techniques (ESPRIT) [4] and subspace techniques [5], [6], [7], [8], [9], [10], [11], [12], [26], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [37], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56],[57], [58], [59] exploit the eigenstructure of the input data matrix. These techniques may fail for reduced data sets or low signal-to-noise ratio (SNR) levels where the expected estimation error is not asymptotic to the Cramér-Rao bound (CRB) [60]. The accuracy of the estimates of the covariance matrix is of fundamental importance in parameter estimation. Low levels of SNR or short data records can result in significant divergences between the true and the sample data covariance matrices. In practice, only a modest number of data snapshots is available and when the number of snapshots is similar to the number of sensor array elements, the estimated and the true subspaces can differ significantly. Several approaches have been developed with the aim of enhancing the computation of the covariance matrix [61]-[70] and for dealing with large sensor-array systems large [71], [72], [73], [74], [75], [76], [77], [78], [79], [80], [81], [82], [83], [84], [85], [86], [87], [97], [88], [89], [90], [91], [92], [93], [94], [95], [96], [98], [99], [100], [101], [102], [103], [104], [105], [106], [111], [108], [109], [110], [111], [112], [113], [114], [115], [118], [117], [118], [119], [120].

Diagonal loading [61] and shrinkage [62], [63], [64] techniques can enhance the estimate of the data covariance matrix by weighing and individually increasing its diagonal by a real constant. Nevertheless, the eigenvectors remain the same, which leads to unaltered estimates of the signal and noise projection matrices obtained from the enhanced covariance matrix. Additionally, an improvement of the estimates of the covariance matrix can be achieved by employing forward/backward averaging and spatial smoothing approaches [65], [66]. The former leads to twice the number of the original samples and its corresponding enhancement. The latter extracts the array covariance matrix as the average of all covariance matrices from its sub-arrays, resulting in a greater number of samples. Both techniques are employed in signal decorrelation. An approach to improve MUSIC dealing with the condition in which the number of snapshots and the sensor elements approach infinity was presented in [67]. Nevertheless, this technique is not that effective for reduced number of snapshots. Other approaches to deal with reduced data sets or low SNR levels [68], [70] consist of reiterating the procedure of adding pseudo-noise to the observations which results in new estimates of the covariance matrix. Then, the set of solutions is computed from previously stored DOA estimates. In [121], two aspects resulting from the computation of DOAs for reduced data sets or low SNR levels have been studied using the root-MUSIC technique. The first aspect dealt with the probability of estimated signal roots taking a smaller magnitude than the estimated noise roots, which is an anomaly that leads to wrong choices of the closest roots to the unit circle. To mitigate this problem, different groups of roots are considered as potential solutions for the signal sources and the most likely one is selected [122]. The second aspect previously mentioned, shown in [123], refers to the fact that a reduced part of the true signal eigenvectors exists in the sample noise subspace (and vice-versa). Such coexistence has been expressed by a Frobenius norm of the related irregularity matrix and introduced its mathematical foundation. An iterative technique to enhance the efficacy of root-MUSIC by reducing this anomaly making use of the gradual reshaping of the sample data covariance matrix has been reported. Inspired by the work in [121], we have developed an ESPRIT-based method known as Two-Step KAI-ESPRIT (TS-ESPRIT) [124], which combines that modifications of the sample data covariance matrix with the use of prior knowledge [125]-[131] about the covariance matrix of a set of impinging signals to enhance the estimation accuracy in the finite sample size region. In practice, this prior knowledge could be from the signals coming from known base stations or from static users in a system. TS-ESPRIT determines the value of a correction factor that reduces the undesirable terms in the estimation of the signal and noise subspaces in an iterative process, resulting in better estimates.

In this work [132], [133], we present the Multi-Step KAI ESPRIT (MS-KAI-ESPRIT) approach that refines the covariance matrix of the input data via multiple steps of reduction of its undesirable terms. This work presents the MS-KAI-ESPRIT in further detail, an analysis of the mean squared error (MSE) of the data covariance matrix free of undesired terms (side effects), a more accurate study of the computational complexity and a comprehensive study of MS-KAI-ESPRIT and other competing techniques for scenarios with both uncorrelated and correlated signals. Unlike TS-ESPRIT, which makes use of only one iteration and available known DOAs, MS-KAI-ESPRIT employs multiple iterations and obtains prior knowledge on line. At each iteration of MS-KAI-ESPRIT, the initial Vandermonde matrix is updated by replacing an increasing number of steering vectors of initial estimates with their corresponding refined versions. In other words, at each iteration, the knowledge obtained on line is updated, allowing the direction finding algorithm to correct the sample covariance matrix estimate, which yields more accurate estimates.

In summary, this work has the following contributions:

- The proposed MS-KAI-ESPRIT technique.
- An MSE analysis of the covariance matrix obtained with the proposed MS-KAI-ESPRIT algorithm.
- A comprehensive performance study of MS-KAI-ESPRIT and competing techniques.

This paper is organized as follows. Section II describes the system model. Section III presents the proposed MS-KAI-ESPRIT algorithm. In section IV, an analytical study of the MSE of the data covariance matrix free of side-effects is carried out together with a study of the computational complexity of the proposed and competing algorithms. In Section V, we present and discuss the simulation results. Section VI concludes the paper.

II. System Model

Let us assume that P narrowband signals from far-field sources impinge on a uniform linear array (ULA) of M $(M > P)$ sensor elements from directions $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_P]^T$. We also consider that the sensors are spaced from each other by a distance $d \leq \frac{\lambda_c}{2}$, where λ_c is the signal wavelength, and that without loss of generality, we have $\frac{-\pi}{2} \leq \theta_1 \leq \theta_2 \leq \ldots \leq \theta_P \leq \frac{\pi}{2}$.

The ith data snapshot of the M-dimensional array output vector can be modeled as

$$\boldsymbol{x}(i) = \boldsymbol{A}\, s(i) + \boldsymbol{n}(i), \qquad i = 1, 2, \ldots, N, \tag{1}$$

where $\boldsymbol{s}(i) = [s_1(i), \ldots, s_P(i)]^T \in \mathbb{C}^{P\times 1}$ represents the zero-mean source data vector, $\boldsymbol{n}(i) \in \mathbb{C}^{M\times 1}$ is the vector of white circular complex Gaussian noise with zero mean and variance σ_n^2, and N denotes the number of available snapshots.

The Vandermonde matrix $\boldsymbol{A}(\boldsymbol{\Theta}) = [\boldsymbol{a}(\theta_1), \ldots, \boldsymbol{a}(\theta_P)] \in \mathbb{C}^{M\times P}$, known as the array manifold, contains the array steering vectors $\boldsymbol{a}(\theta_j)$ corresponding to the nth source, which can be expressed as

$$\boldsymbol{a}(\theta_n) = [1, e^{j2\pi\frac{d}{\lambda_c}\sin\theta_n}, \ldots, e^{j2\pi(M-1)\frac{d}{\lambda_c}\sin\theta_n}]^T, \tag{2}$$

where $n = 1, \ldots, P$. Using the fact that $\boldsymbol{s}(i)$ and $\boldsymbol{n}(i)$ are modeled as uncorrelated linearly independent variables, the $M \times M$ signal covariance matrix is calculated by

$$\boldsymbol{R} = \mathbb{E}\left[\boldsymbol{x}(i)\boldsymbol{x}^H(i)\right] = \boldsymbol{A}\,\boldsymbol{R}_{ss}\boldsymbol{A}^H + \sigma_n^2\boldsymbol{I}_M, \tag{3}$$

where the superscript H and $\mathbb{E}[\cdot]$ in $\boldsymbol{R}_{ss} = \mathbb{E}[\boldsymbol{s}(i)\boldsymbol{s}^H(i)]$ and in $\mathbb{E}[\boldsymbol{n}(i)\boldsymbol{n}^H(i)] = \sigma_n^2\boldsymbol{I}_M$ denote the Hermitian transposition and the expectation operator and $\boldsymbol{I}_M$ stands for the M-dimensional identity matrix. Since the true signal covariance matrix is unknown, it must be estimated and a widely-adopted approach is the sample average formula given by

$$\hat{\boldsymbol{R}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}(i)\boldsymbol{x}^H(i), \tag{4}$$

whose estimation accuracy is dependent on N.

III. Proposed MS-KAI-ESPRIT Algorithm

In this section, we present the proposed MS-KAI-ESPRIT algorithm and detail its main features. We start by expanding (4) using (1) as derived in [121]:

$$\begin{aligned}\hat{\boldsymbol{R}} &= \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{A}\,s(i)+\boldsymbol{n}(i))\,(\boldsymbol{A}\,s(i)+\boldsymbol{n}(i))^H \\ &= \boldsymbol{A}\left\{\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{s}(i)\boldsymbol{s}^H(i)\right\}\boldsymbol{A}^H + \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{n}(i)\boldsymbol{n}^H(i) \; + \\ &\underbrace{\boldsymbol{A}\left\{\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{s}(i)\boldsymbol{n}^H(i)\right\} + \left\{\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{n}(i)\boldsymbol{s}^H(i)\right\}\boldsymbol{A}^H}_{"undesirable terms"}\end{aligned} \tag{5}$$

The first two terms of $\hat{\boldsymbol{R}}$ in (5) can be considered as estimates of the two summands of $\boldsymbol{R}$ given in (3), which represent the signal and the noise components, respectively. The last two terms in (5) are undesirable side effects, which can be seen as estimates for the correlation between the signal and the noise vectors. The system model under study is based on noise vectors which are zero-mean and also independent of the signal vectors. Thus, the signal and noise components are uncorrelated to each other. As a consequence, for a large enough number of samples N, the last two terms of (5) tend to zero. Nevertheless, in practice the number of available samples can be limited. In such situations, the last two terms in (5) may have significant values, which causes the deviation of the estimates of the signal and the noise subspaces from the true signal and noise subspaces.

The key point of the proposed MS-KAI-ESPRIT algorithm is to modify the sample data covariance matrix estimate at each iteration by gradually incorporating the knowledge provided by the newer Vandermonde matrices which progressively embody the refined estimates from the preceding iteration. Based on these updated Vandermonde matrices, refined estimates of the projection matrices of the signal and noise subspaces are calculated. These estimates of projection matrices associated with the initial sample covariance matrix estimate and the reliability factor are employed to reduce its side effects and allow the algorithm to choose the set of estimates that has the highest likelihood of being the set of the true DOAs. The modified covariance matrix is computed by computing a scaled version of the undesirable terms of $\hat{\boldsymbol{R}}$, as pointed out in (5).

The steps of the proposed algorithm are listed in Table I. The algorithm starts by computing the sample data covariance matrix (4). Next, the DOAs are estimated using the ESPRIT algorithm. The superscript $(\cdot)^{(1)}$ refers to the estimation task performed in the first step. Now, a procedure consisting of $n = 1 : P$ iterations starts by forming the Vandermonde matrix using the DOA estimates. Then, the amplitudes of the sources are estimated such that the square norm of the differences between the observation vector and the vector containing estimates and the available known DOAs is minimized. This problem can be formulated [121] as:

$$\hat{\boldsymbol{s}}(i) = \arg\min_{\boldsymbol{s}} \parallel \boldsymbol{x}(i) - \hat{\boldsymbol{A}}\mathbf{s} \parallel_2^2 . \quad (6)$$

The minimization of (6) is achieved using the least squares technique and the solution is described by

$$\hat{\boldsymbol{s}}(i) = (\hat{\mathbf{A}}^H \, \hat{\mathbf{A}})^{-1} \, \hat{\mathbf{A}} \, \boldsymbol{x}(i) \quad (7)$$

The noise component is then estimated as the difference between the estimated signal and the observations made by the array, as given by

$$\hat{\boldsymbol{n}}(i) = \boldsymbol{x}(i) \; - \; \hat{\boldsymbol{A}} \, \hat{\boldsymbol{s}}(i). \quad (8)$$

After estimating the signal and noise vectors, the third term in (5) can be computed as:

$$\begin{aligned}
\boldsymbol{V} &\triangleq \hat{\boldsymbol{A}} \left\{ \frac{1}{N} \sum_{i=1}^{N} \hat{\mathbf{s}}(i) \hat{\mathbf{n}}^H(i) \right\} \\
&= \hat{\boldsymbol{A}} \left\{ \frac{1}{N} \sum_{i=1}^{N} (\hat{\mathbf{A}}^H \, \hat{\mathbf{A}})^{-1} \hat{\mathbf{A}}^H \boldsymbol{x}(i) \right. \\
&\left. \times (\boldsymbol{x}^H(i) - \boldsymbol{x}^H(i) \hat{\mathbf{A}} (\hat{\mathbf{A}}^H \hat{\mathbf{A}})^{-1} \, \hat{\mathbf{A}}^H) \right\} \\
&= \hat{\mathbf{Q}}_A \left\{ \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{x}(i) \boldsymbol{x}^H(i) \, \left(\mathbf{I}_M \; - \; \hat{\mathbf{Q}}_A \right) \right\} \\
&= \hat{\mathbf{Q}}_A \, \hat{\mathbf{R}} \, \hat{\mathbf{Q}}_A^{\perp},
\end{aligned} \quad (9)$$

where

$$\hat{\mathbf{Q}}_A \triangleq \hat{\mathbf{A}} \, (\hat{\mathbf{A}}^H \, \hat{\mathbf{A}})^{-1} \, \hat{\mathbf{A}}^H \quad (10)$$

is an estimate of the projection matrix of the signal subspace, and

$$\hat{\mathbf{Q}}_A^{\perp} \triangleq \mathbf{I}_M \; - \; \hat{\mathbf{Q}}_A \quad (11)$$

is an estimate of the projection matrix of the noise subspace.

Next, as part of the procedure consisting of $n = 1 : P$ iterations, the modified data covariance matrix $\hat{\mathbf{R}}^{(n+1)}$ is obtained by computing a scaled version of the estimated terms from the initial sample data covariance matrix as given by

$$\hat{\mathbf{R}}^{(n+1)} = \hat{\mathbf{R}} \; - \; \mu \, (\mathbf{V}^{(n)} \; + \; \mathbf{V}^{(n)H}), \quad (12)$$

where the superscript $(\cdot)^{(n)}$ refers to the n^{th} iteration performed. The scaling or reliability factor μ increases from 0 to 1 incrementally, resulting in modified data covariance matrices. Each of them gives origin to new estimated DOAs also denoted by the superscript $(\cdot)^{(n+1)}$ by using the ESPRIT algorithm, as briefly described ahead.

In this work, the rank *P* is assumed to be known, which is an assumption frequently found in the literature. Alternatively, the rank *P* could be estimated by model-order selection schemes such as Akaike´s Information Theoretic Criterion (AIC) [144] and the Minimum Descriptive Length (MDL) Criterion [145].

In order to estimate the signal and the orthogonal subspaces from the data records, we may consider two approaches [146], [147]: the direct data approach and the covariance approach. The direct data approach makes use of singular value decomposition(SVD) of the data matrix $\mathbf{X}$, composed of the ith data snapshot (1) of the M-dimensional array data vector:

$$\begin{aligned}
\boldsymbol{X} =&[\boldsymbol{x}(1), \boldsymbol{x}(2), \ldots, \boldsymbol{x}(N)] \\
=&\boldsymbol{A}[\boldsymbol{s}(1), \boldsymbol{s}(2), \ldots, \boldsymbol{s}(N)] + [\boldsymbol{n}(1), \boldsymbol{n}(2), \ldots, \boldsymbol{n}(N)] \\
=&\boldsymbol{A}(\boldsymbol{\Theta}) \; \boldsymbol{S} + \; \boldsymbol{N} \; \in \mathbb{C}^{M \times N}
\end{aligned} \quad (13)$$

Since the number of the sources is assumed known or can be estimated by AIC[144] or MDL[145] , as previously mentioned, we can write $\mathbf{X}$ as:

$$\boldsymbol{X} \;\; = \left[\begin{array}{cc} \hat{\mathbf{U}}_s & \hat{\mathbf{U}}_n \end{array} \right] \left[\begin{array}{cc} \hat{\mathbf{\Gamma}}_s & 0 \\ 0 & \hat{\mathbf{\Gamma}}_n \end{array} \right] \left[\begin{array}{c} \hat{\mathbf{U}}_s^H \\ \hat{\mathbf{U}}_n^H \end{array} \right], \quad (14)$$

where the diagonal matrices $\hat{\mathbf{\Gamma}}_s$ and $\hat{\mathbf{\Gamma}}_n$ contain the P largest singular values and the $M - P$ smallest singular values, respectively. The estimated signal subspace $\hat{\mathbf{U}}_s \in \mathbb{C}^{M \times P}$ consists of the singular vectors corresponding to $\hat{\mathbf{\Gamma}}_s$ and the orthogonal subspace $\hat{\mathbf{U}}_n \in \mathbb{C}^{M \times (M-P)}$ is related to $\hat{\mathbf{\Gamma}}_n$. If the signal subspace is estimated a rank-*P* approximation of the SVD can be applied.

The covariance approach applies the eigenvalue decomposition (EVD) of the sample covariance matrix (4), which is related to the data matrix (13):

$$\hat{\boldsymbol{R}} = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{x}(i) \boldsymbol{x}^H(i) = \frac{1}{N} \boldsymbol{X}\boldsymbol{X}^H \; \in \mathbb{C}^{M \times M} \quad (15)$$

Then, the EVD of (15) can be carried out as follows:

$$\hat{\mathbf{R}} \;\; = \left[\begin{array}{cc} \hat{\mathbf{U}}_s & \hat{\mathbf{U}}_n \end{array} \right] \left[\begin{array}{cc} \hat{\mathbf{\Lambda}}_s & 0 \\ 0 & \hat{\mathbf{\Lambda}}_n \end{array} \right] \left[\begin{array}{c} \hat{\mathbf{U}}_s^H \\ \hat{\mathbf{U}}_n^H \end{array} \right], \quad (16)$$

where the diagonal matrices $\hat{\mathbf{\Lambda}}_{\mathbf{s}}$ and $\hat{\mathbf{\Lambda}}_{\mathbf{n}}$ contain the *P* largest and the *M-P* smallest eigenvalues, respectively. The estimated signal subspace $\hat{\mathbf{U}}_s \in \mathbb{C}^{M \times P}$ corresponding to $\hat{\mathbf{\Gamma}}_{\mathbf{s}}$ and the orthogonal subspace $\hat{\mathbf{U}}_n \in \mathbb{C}^{M \times (M-P)}$ complies with $\hat{\Gamma}_n$. If the signal subspace is estimated a rank-P approximation of the EVD can be applied. With infinite precision arithmetic, both SVD and EVD can be considered equivalent. However, as in practice, finite precision arithmetic is employed, 'squaring' the data to obtain the Gramian $\boldsymbol{X}\boldsymbol{X}^H$ (15) can result in round-off errors and overflow. These are potential problems to be aware when using the covariance approach.

Now, we can briefly review ESPRIT. We start by forming a twofold subarray configuration, as each row of the array steering matrix $\boldsymbol{A}(\boldsymbol{\Theta})$ corresponds to one sensor element of the antenna array. The subarrays are specified by two $(s \times M)$-dimensional selection matrices $\mathbf{J_1}$ and $\mathbf{J_2}$ which choose s elements of the M existing sensors, respectively, where s is in the range $P \leq s < M$. For maximum overlap, the matrix $\mathbf{J_1}$ selects the first $s = M - 1$ elements and the matrix $\mathbf{J_2}$ selects the last $s = M - 1$ rows of $\boldsymbol{A}(\boldsymbol{\Theta})$.

Since the matrices $\mathbf{J_1}$ and $\mathbf{J_2}$ have now been computed, we can estimate the operator $\mathbf{\Psi}$ by solving the approximation of the shift invariance equation (17) given by

$$\mathbf{J}_1\,\hat{\mathbf{U}}_s\,\mathbf{\Psi} \;\approx\; \mathbf{J}_2\,\hat{\mathbf{U}}_s. \tag{17}$$

where $\hat{U}_s$ is obtained in (16).

Using the least squares (LS) method, which yields

$$\hat{\mathbf{\Psi}} = \arg\min_{\mathbf{\Psi}} \parallel \mathbf{J}_2\hat{\mathbf{U}}_s - \mathbf{J}_1\hat{\mathbf{U}}_s\mathbf{\Psi} \parallel_F = \left(\mathbf{J}_1\,\hat{\mathbf{U}}_s\right)^{\dagger}\mathbf{J}_2\hat{\mathbf{U}}_s, \tag{18}$$

where $\|\cdot\|_F$ denotes the Frobenius norm and $(\cdot)^{\dagger}$ stands for the pseudo-inverse.

Lastly, the eigenvalues λ_i of $\hat{\mathbf{\Psi}}$ contain the estimates of the spatial frequencies γ_i computed as:

$$\gamma_i \;=\; \arg\left(\lambda_i\right), \tag{19}$$

so that the DOAs can be calculated as:

$$\hat{\theta}_i \;=\; \arcsin\left(\frac{\gamma_i\,\lambda_c}{2\pi\,d}\right) \tag{20}$$

where for (19) and (20) $\mathrm{i} = 1,\cdots,\mathrm{P}$.

Then, a new Vandermonde matrix $\hat{\mathbf{B}}^{(n+1)}$ is formed by the steering vectors of those refined estimates of the DOAs. By using this updated matrix, it is possible to compute the refined estimates of the projection matrices of the signal $\hat{\mathbf{Q}}_B^{(n+1)}$ and the noise $\hat{\mathbf{Q}}_B^{(n+1)\perp}$ subspaces.

Next, employing the refined estimates of the projection matrices, the initial sample data matrix, $\hat{\boldsymbol{R}}$, and the number of sensors and sources, the stochastic maximum likelihood objective function $U^{(n+1)}(\mu)$ [122] is computed for each value of μ at the n^{th} iteration, as follows:

$$U^{(n+1)}(\mu) = \ln\det\Big(\hat{\mathbf{Q}}_B^{(n+1)}\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_B^{(n+1)} + \frac{\mathrm{Trace}\{\hat{\mathbf{Q}}_B^{\perp\,(n+1)}\,\hat{\mathbf{R}}\}}{\mathrm{M}-\mathrm{P}}\,\hat{\mathbf{Q}}_B^{(n+1)\perp}\Big). \tag{21}$$

The previous computation selects the set of unavailable DOA estimates that have a higher likelihood at each iteration. Then, the set of estimated DOAs corresponding to the optimum value of μ that minimizes (21) also at each n^{th} iteration is determined. Finally, the output of the proposed MS-KAI-ESPRIT algorithm is formed by the set of the estimates obtained at the P^{th} iteration, as described in Table I.

TABLE I
PROPOSED MS-KAI-ESPRIT ALGORITHM

Inputs:
$M,\ d,\ \lambda,\ N,\ P$
Received vectors $\boldsymbol{x}(1),\ \boldsymbol{x}(2),\cdots,\ \boldsymbol{x}(N)$
Outputs:
Estimates $\hat{\theta}_1^{(n+1)}(\mu\,opt),\ \hat{\theta}_2^{(n+1)}(\mu\,opt),\cdots,\ \hat{\theta}_P^{(n+1)}(\mu\,opt)$

First step:
$\hat{\mathbf{R}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{x}(i)\boldsymbol{x}^H(i)$
$\{\hat{\theta}_1^{(1)},\ \hat{\theta}_2^{(1)},\cdots,\hat{\theta}_P^{(1)}\}\ \underleftarrow{ESPRIT}\ (\hat{\mathbf{R}},P,d,\lambda)$
$\hat{\mathbf{A}}^{(1)} = \left[\mathbf{a}(\hat{\theta}_1^{(1)}),\mathbf{a}(\hat{\theta}_2^{(1)}),\cdots,\mathbf{a}(\hat{\theta}_P^{(1)})\right]$
Second step:
for $n = 1 : P$
$\hat{\mathbf{Q}}_A^{(n)} = \hat{\mathbf{A}}^{(n)}\,(\hat{\mathbf{A}}^{(n)H}\,\hat{\mathbf{A}}^{(n)})^{-1}\,\hat{\mathbf{A}}^{(n)H}$
$\hat{\mathbf{Q}}_A^{(n)\perp} = \mathbf{I}_M - \hat{\mathbf{Q}}_A^{(n)}$
$\mathbf{V}^{(n)} = \hat{\mathbf{Q}}_A^{(n)}\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A^{(n)\perp}$
for $\mu = 0 : \iota\ : 1$
$\hat{\mathbf{R}}^{(n+1)} = \hat{\mathbf{R}} - \mu\,(\mathbf{V}^{(n)} + \mathbf{V}^{(n)H})$
$\{\hat{\theta}_1^{(n+1)},\ \hat{\theta}_2^{(n+1)},\cdots,\hat{\theta}_P^{(n+1)}\}\ \underleftarrow{ESPRIT}\ (\hat{\mathbf{R}}^{(n+1)},\ P,d,\lambda)$
$\hat{\mathbf{B}}^{(n+1)} = \left[\mathbf{a}(\hat{\theta}_1^{(n+1)}),\mathbf{a}(\hat{\theta}_2^{(n+1)}),\cdots,\mathbf{a}(\hat{\theta}_P^{(n+1)})\right]$
$\hat{\mathbf{Q}}_B^{(n+1)} = \hat{\mathbf{B}}^{(n+1)}\,(\hat{\mathbf{B}}^{(n+1)H}\,\hat{\mathbf{B}}^{(n+1)})^{-1}\,\hat{\mathbf{B}}^{(n+1)H}$
$\hat{\mathbf{Q}}_B^{(n+1)\perp} = \mathbf{I}_M - \hat{\mathbf{Q}}_B^{(n+1)}$
$U^{(n+1)}(\mu) = \ln\det(\cdot)\,,$
$(\cdot) = \left(\hat{\mathbf{Q}}_B^{(n+1)}\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_B^{(n+1)} + \frac{\mathrm{Trace}\{\hat{\mathbf{Q}}_B^{\perp\,(n+1)}\,\hat{\mathbf{R}}\}}{\mathrm{M}-\mathrm{P}}\,\hat{\mathbf{Q}}_B^{(n+1)\perp}\right)$
$\mu_{\mathrm{opt}}^{(n+1)} = \arg\min\ U^{(n+1)}(\mu)$
$\mathrm{DOAs}^{(n+1)} = (*)\,,$
$(*) = \{\hat{\theta}_1^{(n+1)}(\mu\,opt),\ \hat{\theta}_2^{(n+1)}(\mu\,opt),\cdots,\ \hat{\theta}_P^{(n+1)}(\mu\,opt)\}$
$\hat{\mathbf{A}}^{(n+1)} = \left\{\mathbf{a}(\hat{\theta}_{\{1,\cdots,n\}}^{(n+1)})\right\}\bigcup\left\{\mathbf{a}(\hat{\theta}_{\{1,\cdots,P\}-\{1,\cdots,n\}}^{(1)})\right\}$
end for
end for

IV. Analysis

In this section, we carry out an analysis of the MSE of the data covariance matrix free of side effects along with a study of the computational complexity of the proposed MS-KAI-ESPRIT and existing direction finding algorithms.

A. MSE Analysis

In this subsection we show that at the first of the P iterations, the MSE of the data covariance matrix free of side effects $\hat{\mathbf{R}}^{(n+1)}$ is less than or equal to the MSE of that of the original one $\hat{\mathbf{R}}$. This can be formulated as:

$$\mathrm{MSE}\left(\hat{\mathbf{R}}^{(n+1)}\right) \le \mathrm{MSE}\left(\hat{\mathbf{R}}\right) \tag{22}$$

or, alternatively, as

$$\mathrm{MSE}\left(\hat{\mathbf{R}}^{(n+1)}\right) - \mathrm{MSE}\left(\hat{\mathbf{R}}\right) \le 0 \tag{23}$$

The proof of this inequality is provided in the Appendix.

B. Computational Complexity Analysis

In this section, we evaluate the computational cost of the proposed MS-KAI-ESPRIT algorithm which is compared to the following classical subspace methods: ESPRIT [4], MUSIC [2], Root-MUSIC [3], Conjugate Gradient (CG) [138], Auxiliary Vector Filtering (AVF) [139] and TS-ESPRIT [124]. The ESPRIT and MUSIC-based methods use the Singular Value Decomposition (SVD) of the sample covariance matrix (4). The computational complexity of MS-KAI-ESPRIT in terms of number of multiplications and additions is depicted in Table II, where $\tau = \frac{1}{\iota} + 1$. The increment ι is defined in Table I. As can be seen, for this specific configuration used in the simulations V MS-KAI-ESPRIT shows a relatively high computational burden with $\mathcal{O}(P\tau(3M^3 + 8MN^2))$, where τ is typically an integer that ranges from 1 to 20. It can be noticed that for the configuration used in the simulations

$(P = 4, M = 40, N = 25)$ $3M^3$ and $8MN^2$ are comparable, resulting in two dominant terms. It can also be seen that the number of multiplications required by the proposed algorithm is more significant than the number of additions. For this reason, in Table III, we computed only the computational burden of the previously mentioned algorithms in terms of multiplications for the purpose of comparisons. In that table, Δ stands for the search step.

Next, we will evaluate the influence of the number of sensor elements on the number of multiplications based on the specific configuration described in Table II. Supposing $\mathrm{P} = 4$ narrowband signals impinging a ULA of M sensor elements and $\mathrm{N} = 25$ available snapshots, we obtain Fig. 1. We can see the main trends in terms of computational cost measured in multiplications of the proposed and analyzed algorithms. By examining Fig. 1, it can be noticed that in the range $M = [20\ 70]$ sensors, the curves describing the exact number of multiplications in MS-KAI-ESPRIT and AVF tend to merge. For $M = 40$, this ratio tends to 1, i.e. the number of multiplications are almost equiva

TABLE II
COMPUTATIONAL COMPLEXITY

MS-KAI -ESPRIT (Proposed)	Multiplications
	$P\,\tau[\frac{10}{3}\mathrm{M}^3 + \mathrm{M}^2(3\mathrm{P}+2)$ $+\mathrm{P}^2(\frac{17}{2}\mathrm{P}+\frac{1}{2})]$
	$+\mathrm{P}\,[2\mathrm{M}^3 + \mathrm{M}^2(\mathrm{P}) + \mathrm{M}($ $+2\mathrm{M}^2(\mathrm{P}) + \mathrm{M}(\mathrm{P}^2 - \mathrm{P} +$
	Additions
	$P\,\tau[\frac{10}{3}\mathrm{M}^3 + \mathrm{M}^2(3\mathrm{P}-1)$ $+\mathrm{P}(8\mathrm{P}^2 - 2\mathrm{P} - \frac{5}{2})]$
	$+\mathrm{P}\,[2\mathrm{M}^3 + \mathrm{M}^2(\mathrm{P}-2) +$ $+2\mathrm{M}^2(\mathrm{P}) + \mathrm{M}(\mathrm{P}^2 - 4\mathrm{P} + 8\mathrm{N}^2) + \mathrm{P}(8\mathrm{P}^2 - \mathrm{P} - 2)$

TABLE III
COMPUTATIONAL COMPLEXITY - OTHER ALGORITHMS

Algorithm	Multiplications
MUSIC [2]	$\frac{180}{\Delta}[\mathrm{M}^2 + \mathrm{M}(2-\mathrm{P}) - \mathrm{P}] + 8\mathrm{MN}^2$
root-MUSIC[3]	$2\mathrm{M}^3 - \mathrm{M}^2\mathrm{P} + 8\mathrm{MN}^2$
AVF [139]	$\frac{180}{\Delta}[\mathrm{M}^2(3\mathrm{P}+1) + \mathrm{M}(4\mathrm{P}-2) + \mathrm{P} + 2]$ $+\mathrm{M}^2\mathrm{N}$
CG [138]	$\frac{180}{\Delta}[\mathrm{M}^2(\mathrm{P}+1) + \mathrm{M}(6\mathrm{P}+2) + \mathrm{P} + 1] + \mathrm{M}^2\mathrm{N}$
ESPRIT[4]	$2\mathrm{M}^2\mathrm{P} + \mathrm{M}(\mathrm{P}^2 - 2\mathrm{P} + 8\mathrm{N}^2) + 8\mathrm{P}^3 - \mathrm{P}^2$
TS-ESPRIT [124]*	$\tau[3\mathrm{M}^3 + \mathrm{M}^2(3\mathrm{P}+2) + \mathrm{M}(\frac{5}{2}\mathrm{P}^2 - \frac{3}{2}\mathrm{P} + 8\mathrm{N}^2)$ $+\mathrm{P}^2(\frac{17}{2}\mathrm{P}+\frac{1}{2}) + 1]$ $+[2\mathrm{M}^3 + \mathrm{M}^2(3\mathrm{P}) + \mathrm{M}(\frac{5}{2}\mathrm{P}^2 - \frac{3}{2}\mathrm{P} + 8\mathrm{N}^2)$ $+\mathrm{P}^2(\frac{17}{2}\mathrm{P}+\frac{1}{2})]$

V. SIMULATIONS

In this section, we examine the performance of the proposed MS-KAI-ESPRIT in terms of probability of resolution and RMSE and compare them to the standard ESPRIT [4], the Iterative ESPRIT (IESPRIT), which is also developed here by combining the approach in [121] that exploits knowledge of the structure of the covariance matrix and its perturbation terms, the Conjugate Gradient (CG) [138], the Root-MUSIC

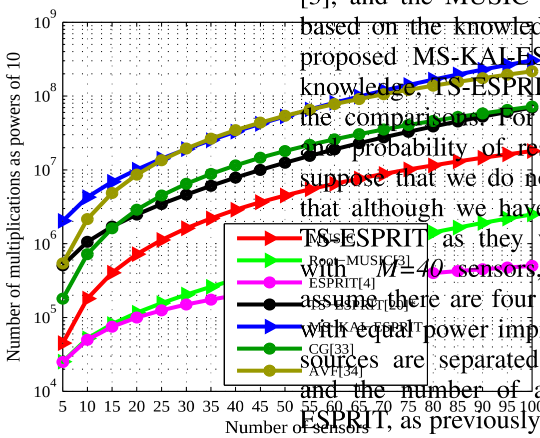


Fig. 1. Number of multiplications as powers of 10 versus number of sensors for $P = 4$, $N = 25$.

[3], and the MUSIC [2] algorithms. Despite TS-ESPRIT is based on the knowledge of available known DOAS and the proposed MS-KAI-ESPRIT does not have access to prior knowledge, TS-ESPRIT is plotted with the aim of illustrating the comparisons. For a fair comparison in terms of RMSE and probability of resolution of all studied algorithms, we suppose that we do not have prior knowledge, that is to say that although we have available known DOAs, we compute TS-ESPRIT as they were unavailable. We employ a ULA with M=40 sensors, inter-element spacing $\Delta = \frac{\lambda_c}{2}$ and assume there are four uncorrelated complex Gaussian signals with equal power impinging on the array. The closely-spaced sources are separated by 2.4^o, at $(10.2^o, 12.6^o, 15^o, 17.4^o)$, and the number of available snapshots is N=25. For TS-ESPRIT, as previously mentioned, we presume a priori knowledge of the last true DOAS $(15^o, 17.4^o)$

In Fig. 2, we show the probability of resolution versus SNR. We take into account the criterion [141], in which two sources with DOA θ_1 and θ_2 are said to be resolved if their respective estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ are such that both $\left|\hat{\theta}_1 - \theta_1\right|$ and $\left|\hat{\theta}_2 - \theta_2\right|$ are less than $|\theta_1 - \theta_2|/2$. The proposed MS-KAI-ESPRIT algorithm outperforms IESPRIT developed here, based on [121], and the standard ESPRIT [4] in the range between -6 and $5dB$ and MUSIC [2] from -6 to $8.5dB$. MS-KAI-ESPRIT also outperforms CG [138] and Root-Music [3] throughout the whole range of values. The poor performance of the latter could be expected from the results for two closed signals obtained in [121]. When compared to TS-ESPRIT, which as previously discussed, was supposed to have the best performance, the proposed MS-KAI-ESPRIT algorithm is outperformed by the former only in the range between -6 and $-2dB$. From this last point to $20dB$ its performance is superior or equal to the other algorithms.

In Fig. 3, it is shown the RMSE in dB versus SNR, where the term CRB refers to the square root of the deterministic Cramér-Rao bound [142]. The RMSE is defined as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{L\,P}\sum_{l=1}^{L}\sum_{p=1}^{P}(\theta_p - \hat{\theta}_p(l))^2}, \tag{24}$$

where L is the number of trials.

The results show the superior performance of MS-KAI-ESPRIT in the range between -2.5 and 5 dB. From this last point to 20 dB, MS-KAI-ESPRIT, IESPRIT, ESPRIT and TS-ESPRIT have similar performance. The only range in which MS-KAI-ESPRIT is outperformed lies in the range between -6 and -2.5 dB. From this last point to 20 dB its performance is better or similar to the others.

Fig. 2. Probability of resolution versus SNR
$M = 40$, $N = 25$, $L = 100$ runs

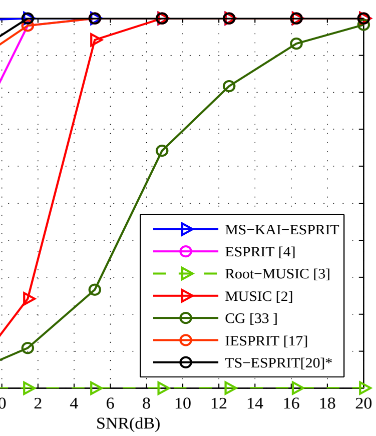
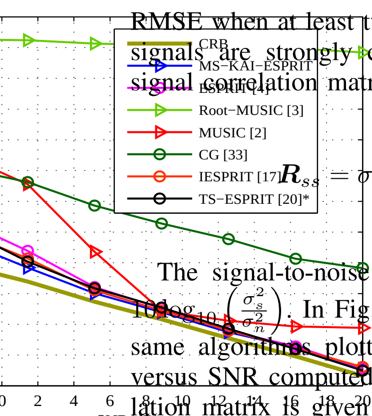


Fig. 3. RMSE and the square root of CRB versus SNR with $P = 4$ uncorrelated sources, $M = 40$, $N = 25$, $L = 100$ runs

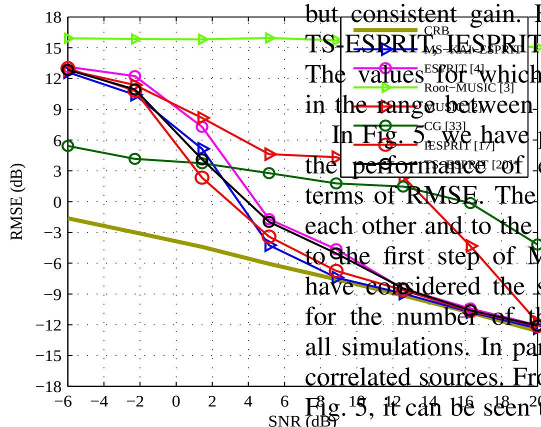
Now, we focus on the performance of MS-KAI-ESPRIT under more severe conditions, i.e., we analyze it in terms of RMSE when at least two of the four equal-powered Gaussian signals are strongly correlated, as shown in the following signal correlation matrix $\boldsymbol{R}_{ss}$ (25):



$$\boldsymbol{R}_{ss} = \sigma_s^2 \begin{bmatrix} 1 & 0.9 & 0.6 & 0 \\ 0.9 & 1 & 0.4 & 0.5 \\ 0.6 & 0.4 & 1 & 0 \\ 0 & 0.5 & 0 & 1 \end{bmatrix}. \quad (25)$$

The signal-to-noise ratio (SNR
$10\log_{10}\left(\frac{\sigma_s^2}{\sigma_n^2}\right)$. In Fig. 4, we can s
same algorithms plotted in Fig. 3
versus SNR computed after 250 ru
lation matrix is given by (25). As
performance of MS-KAI-ESPRIT

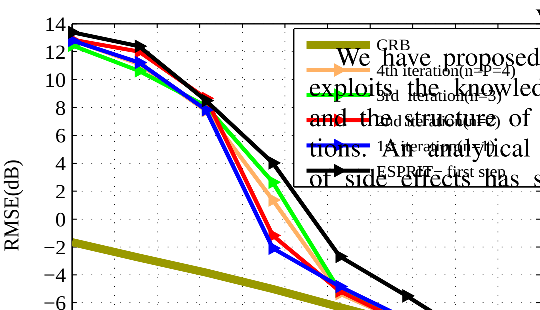
Fig. 4. RMSE and the square root of CRB versus SNR with $P = 4$ correlated sources, $M = 40$, $N = 25$, $L = 250$ runs



between 4.0 and 12 dB , which can be considered a small but consistent gain. From 12dB to 20dB MS-KAI-ESPRIT, TS-ESPRIT, IESPRIT and ESPRIT have similar performance. The values for which MS-KAI-ESPRIT is outperformed are in the range between -6.0 and 4.0dB.

In Fig. 5 we have provided further simulations to illustrate the performance of each iteration of MS-KAI ESPRIT in terms of RMSE. The resulting iterations can be compared to each other and to the the original ESPRIT, which corresponds to the first step of MS-KAI ESPRIT. For this purpose, we have considered the same scenario employed before, except for the number of the trials, which is $L = 200$ runs for all simulations. In particular, we have considered the case of correlated sources. From Fig. 6, which is a magnified detail of Fig. 5, it can be seen that the estimates become more accurate with the increase of iterations.

Fig. 5. RMSE for each iteration of MS-KAI ESPRIT,original ESPRIT and CRB versus SNR with $P = 4$ correlated sources, $M = 40$, $N = 25$, $L = 200$ runs



VI. Conclusions

We have proposed the MS-KAI-ESPRIT algorithm which exploits the knowledge of source signals obtained on line and the structure of the covariance matrix and its perturbations. An analytical study of the MSE of this matrix free of side effects has shown that it is less or equal than the

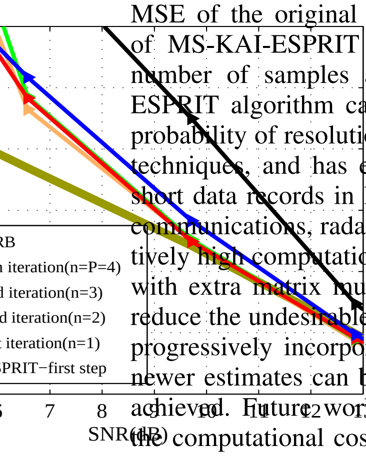


Fig. 6. RMSE for each iteration of MS-KAI ESPRIT,original ESPRIT and CRB versus SNR with $P = 4$ correlated sources, $M = 40$, $N = 25$, $L = 200$ runs -magnification

MSE of the original matrix, resulting in better performance of MS-KAI-ESPRIT especially in scenarios where limited number of samples are available. The proposed MS-KAI-ESPRIT algorithm can obtain significant gains in RMSE or probability of resolution performance over previously reported techniques, and has excellent potential for applications with short data records in large-scale antenna systems for wireless communications, radar and other large sensor arrays. The relatively high computational burden required, which is associated with extra matrix multiplications, the increment ι applied to reduce the undesirable side effects and the iterations needed to progressively incorporate the knowledge obtained on line as newer estimates can be justified for the superior performance achieved. Future work will consider approaches to reducing the computational cost.

Appendix

Here, we prove the inequality (23) described in Section IV-A. We start by expressing the MSE of the original data covariance matrix (4) as:

$$\mathrm{MSE}\left(\hat{\mathbf{R}}\right) = \mathbb{E}\left[\|\hat{\mathbf{R}} - \mathbf{R}\|_F^2\right]. \tag{26}$$

where $\mathbf{R}$ is the true covariance matrix . Similarly, the MSE of the data covariance matrix free of side effects $\hat{\mathbf{R}}^{(n+1)}$ can be expressed for the first iteration $n = 1$ by making use of (12), as follows

$$\begin{aligned}
&\mathrm{MSE}\left(\hat{\mathbf{R}}^{(n+1)}\right)|_{n=1} \\
&= \mathrm{MSE}\left(\hat{\mathbf{R}}^{(2)}\right) = \mathbb{E}\left[\|\hat{\mathbf{R}}^{(2)} - \mathbf{R}\|_F^2\right] \\
&= \mathbb{E}\left[\|\hat{\mathbf{R}} - \mu\left(\mathbf{V}^{(1)} + \mathbf{V}^{(1)H}\right) - \mathbf{R}\|_F^2\right] \\
&= \mathbb{E}\left[\|\left(\hat{\mathbf{R}} - \mathbf{R}\right) - \mu\left(\mathbf{V}^{(1)} + \mathbf{V}^{(1)H}\right)\|_F^2\right]
\end{aligned} \tag{27}$$

where for the sake of simplicity, from now on we omit the superscript $^{(1)}$, which refers to the first iteration. In order to expand the result in (27), we make use of the following proposition:

Lemma 1: The squared Frobenius norm of the difference between any two matrices $\mathbf{A} \in \mathbb{C}^{m\times m}$ and $\mathbf{B} \in \mathbb{C}^{m\times m}$ is given by

$$\|\mathbf{A} - \mathbf{B}\|_F^2 = \|\mathbf{A}\|_F^2 + \|\mathbf{B}\|_F^2 - \left(\mathrm{Tr}\,\mathbf{A}^H\mathbf{B} + \mathrm{Tr}\,\mathbf{A}\mathbf{B}^H\right) \tag{28}$$

Proof of Lemma 1:
The Frobenius norm of any $\mathbf{D} \in \mathbb{C}^{m\times m}$ matrix is defined [1] as

$$\|\mathbf{D}\|_F = \left(\sum_{i=1}^{m}\sum_{j=1}^{m}|d_{ij}|^2\right)^{\frac{1}{2}} = \left[\mathrm{Tr}\left(\mathbf{D}^H\mathbf{D}\right)\right]^{\frac{1}{2}} \tag{29}$$

We express $\mathbf{D}$ as a difference between two matrices $\mathbf{A}$ and $\mathbf{B}$, both also $\in \mathbb{C}^{m\times m}$. Making use of Lemma1 and the properties of the trace, we obtain

$$\begin{aligned}
&\|\mathbf{A} - \mathbf{B}\|_F^2 = \mathrm{Tr}\left[(\mathbf{A} - \mathbf{B})^H(\mathbf{A} - \mathbf{B})\right] \\
&= \mathrm{Tr}\left[\left(\mathbf{A}^H - \mathbf{B}^H\right)(\mathbf{A} - \mathbf{B})\right] \\
&= \mathrm{Tr}\left[\left(\mathbf{A}^H\mathbf{A}\right) - \mathrm{Tr}\left(\mathbf{A}^H\mathbf{B}\right) - \mathrm{Tr}\left(\mathbf{B}^H\mathbf{A}\right) + \mathrm{Tr}\left(\mathbf{B}^H\mathbf{B}\right)\right] \\
&= \|\mathbf{A}\|_F^2 + \|\mathbf{B}\|_F^2 - \left(\mathrm{Tr}\,\mathbf{A}^H\mathbf{B} + \mathrm{Tr}\,\mathbf{A}\mathbf{B}^H\right),
\end{aligned} \tag{30}$$

which is the desired result.

Now, assuming that the true $\mathbf{R}$ [134] and the data covariance matrices $\hat{\mathbf{R}}$ [134] are Hermitian and using (27) combined with Lemma1, the cyclic [135] property of the trace and the linearity [136] property of the expected value, we get

$$\begin{aligned}
\mathrm{MSE}\left(\hat{\mathbf{R}}^{(2)}\right) &= \mathbb{E}\Big\{\|\hat{\mathbf{R}} - \mathbf{R}\|_F^2 + \mu^2\|\mathbf{V} + \mathbf{V}^H\|_F^2 \\
&- \mathrm{Tr}\left[\left(\hat{\mathbf{R}} - \mathbf{R}\right)^H\mu\left(\mathbf{V} + \mathbf{V}^H\right)\right] \\
&- \mathrm{Tr}\left[\mu\left(\mathbf{V} + \mathbf{V}^H\right)^H\left(\hat{\mathbf{R}} + \mathbf{R}\right)\right]\Big\} \\
&= \mathbb{E}\Big\{\|\hat{\mathbf{R}} - \mathbf{R}\|_F^2 + \mu^2\|\mathbf{V} + \mathbf{V}^H\|_F^2 \\
&-\mu\,\mathrm{Tr}\left[\left(\hat{\mathbf{R}} - \mathbf{R}\right)^H\left(\mathbf{V} + \mathbf{V}^H\right)\right] \\
&-\mu\,\mathrm{Tr}\left[\left(\mathbf{V} + \mathbf{V}^H\right)^H\left(\hat{\mathbf{R}} + \mathbf{R}\right)\right]\Big\} \\
&= \mathbb{E}\Big\{\|\hat{\mathbf{R}} - \mathbf{R}\|_F^2 + \mu^2\|\mathbf{V} + \mathbf{V}^H\|_F^2 \\
&-\mu\,\mathrm{Tr}\left[\left(\hat{\mathbf{R}} - \mathbf{R}\right)\left(\mathbf{V} + \mathbf{V}^H\right)\right] \\
&-\mu\,\mathrm{Tr}\left[\left(\mathbf{V}^H + \mathbf{V}\right)\left(\hat{\mathbf{R}} + \mathbf{R}\right)\right]\Big\} \\
&= \mathbb{E}\Big\{\|\hat{\mathbf{R}} - \mathbf{R}\|_F^2 + \mu^2\|\mathbf{V} + \mathbf{V}^H\|_F^2 \\
&-\mu\,\mathrm{Tr}\left[\left(\hat{\mathbf{R}} - \mathbf{R}\right)\left(\mathbf{V} + \mathbf{V}^H\right)\right] \\
&-\mu\,\mathrm{Tr}\left[\left(\hat{\mathbf{R}} + \mathbf{R}\right)\left(\mathbf{V} + \mathbf{V}^H\right)\right]\Big\} \\
&= \mathbb{E}\left\{\|\hat{\mathbf{R}} - \mathbf{R}\|_F^2\right\} + \mu^2\mathbb{E}\left\{\|\mathbf{V} + \mathbf{V}^H\|_F^2\right\} \\
&- 2\mu\mathbb{E}\left\{\mathrm{Tr}\left[\left(\hat{\mathbf{R}} - \mathbf{R}\right)\left(\mathbf{V} + \mathbf{V}^H\right)\right]\right\} \\
&= \mathrm{MSE}\left(\hat{\mathbf{R}}\right) + \mu^2\mathbb{E}\left\{\|\mathbf{V} + \mathbf{V}^H\|_F^2\right\} \\
&- 2\mu\mathbb{E}\left\{\mathrm{Tr}\left[\left(\hat{\mathbf{R}} - \mathbf{R}\right)\left(\mathbf{V} + \mathbf{V}^H\right)\right]\right\}
\end{aligned} \tag{31}$$

By moving the first summand of (31) to its first element, we obtain the intended expression for the difference between the $MSEs$ of the data covariance matrix free of perturbations and the original one, i.e.:

$$\text{MSE}\left(\hat{\mathbf{R}}^{(n+1)}\right)|_{n=1} - \text{MSE}\left(\hat{\mathbf{R}}\right) = \mu^2 \mathbb{E}\left\{\|\mathbf{V}+\mathbf{V}^H\|_F^2\right\} \\ -2\mu\mathbb{E}\left\{\text{Tr}\left[\left(\hat{\mathbf{R}}-\mathbf{R}\right)\left(\mathbf{V}+\mathbf{V}^H\right)\right]\right\}. \quad (32)$$

Now, we expand the expressions inside braces of the second member of (32) individually. We start with the first summand

$$\|\mathbf{V}+\mathbf{V}^H\|_F^2 = \|\mathbf{V}\|_F^2 + \|\mathbf{V}^H\|_F^2 + \text{Tr}\left(\mathbf{V}^H\mathbf{V}^H\right) + \\ \text{Tr}\left(\left(\mathbf{V}^H\right)^H\mathbf{V}\right) \\ = \|\mathbf{V}\|_F^2 + \|\mathbf{V}^H\|_F^2 + \text{Tr}\left(\mathbf{V}^H\mathbf{V}^H\right) + \text{Tr}\left(\mathbf{V}\mathbf{V}\right). \quad (33)$$

The equation (33) can be computed by using the projection matrices of the signal and the noise subspaces and the data covariance matrix by using (9), (11), the idempotence [1] [135] of $\hat{\mathbf{Q}}_A$ and the cyclic property [135] of the trace. Starting with the computation of its fourth summand, we have

$$\text{Tr}\left(\mathbf{V}\mathbf{V}\right) = \text{Tr}\left[\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A^{\perp}\right)\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A^{\perp}\right)\right] \\ = \text{Tr}\left[\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\left(\mathbf{I_M}-\hat{\mathbf{Q}}_A\right)\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\left(\mathbf{I_M}-\hat{\mathbf{Q}}_A\right)\right] \\ = \text{Tr}\left[\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}-\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right) \\ \left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}-\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right)\right] \\ = \text{Tr}\left[\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}-\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A \\ -\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}+\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right] \\ = \text{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\right)-\text{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right) \\ -\text{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\right) \\ +\text{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right) \\ = \text{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\right)-\text{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\right) \\ -\text{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\right)+\text{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\right) = 0. \quad (34)$$

Taking into account that the data covariance matrix $\hat{\mathbf{R}}$ and the estimate of the projection matrix of the noise subspace $\hat{\mathbf{Q}}_A^{\perp}$ are Hermitian, we can evaluate the third summand of (33) as follows:

$$\text{Tr}\left(\mathbf{V}^H\mathbf{V}^H\right) = \text{Tr}\left[\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A^{\perp}\right)^H\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A^{\perp}\right)^H\right] \\ = \text{Tr}\left\{\left[\left(\hat{\mathbf{Q}}_A^{\perp}\right)^H\,\hat{\mathbf{R}}^H\,\hat{\mathbf{Q}}_A^H\right]\left[\left(\hat{\mathbf{Q}}_A^{\perp}\right)^H\,\hat{\mathbf{R}}^H\,\hat{\mathbf{Q}}_A^H\right]\right\} \\ = \text{Tr}\left\{\left[\hat{\mathbf{Q}}_A^{\perp}\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right]\left[\hat{\mathbf{Q}}_A^{\perp}\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right]\right\} \\ = \text{Tr}\left\{\left[\left(\mathbf{I_M}-\hat{\mathbf{Q}}_\mathbf{A}\right)\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right]\left[\left(\mathbf{I_M}-\hat{\mathbf{Q}}_\mathbf{A}\right)\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right]\right\} \\ = \text{Tr}\left\{\left[\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A-\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right]\left[\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A-\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right]\right\} \\ = \text{Tr}\left\{\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A-\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A \\ -\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A+\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right\} \\ = \text{Tr}\left(\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right)-\text{Tr}\left(\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right) \\ -\text{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right)+\text{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right) \\ = \text{Tr}\left(\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\right)-\text{Tr}\left(\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\right) \\ -\text{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\right)+\text{Tr}\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\right) = 0. \quad (35)$$

By using (29), we can expand the first and the second summands of (33) as follows:

$$\|\mathbf{V}\|_F^2 + \|\mathbf{V}^H\|_F^2 = \text{Tr}\left(\mathbf{V}^H\mathbf{V}\right) + \text{Tr}\left(\left(\mathbf{V}^H\right)^H\mathbf{V}^H\right) \\ = \text{Tr}\left(\mathbf{V}^H\mathbf{V}\right) + \text{Tr}\left(\mathbf{V}\mathbf{V}^H\right) \\ = \text{Tr}\left(\mathbf{V}\mathbf{V}^H\right) + \text{Tr}\left(\mathbf{V}\mathbf{V}^H\right) = 2\,\text{Tr}\left(\mathbf{V}\mathbf{V}^H\right). \quad (36)$$

Equation (36) can be expressed in terms of the projection matrices of the signal and the noise subspaces and the data covariance, in a similar way as for the third and fourth summands of (33), as follows:

$$2\,\text{Tr}\left(\mathbf{V}\mathbf{V}^H\right) = 2\,\text{Tr}\left[\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A^{\perp}\right)\left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A^{\perp}\right)^H\right] \\ = 2\,\text{Tr}\left\{\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\left(\mathbf{I_M}-\hat{\mathbf{Q}}_A\right)\left[\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\left(\mathbf{I_M}-\hat{\mathbf{Q}}_A\right)\right]^H\right\} \\ = 2\,\text{Tr}\left\{\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}-\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\right)\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}-\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\right)^H\right\} \\ = 2\,\text{Tr}\left\{\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}\hat{\mathbf{Q}}_A-\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}} \\ -\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A+\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\} \\ = 2\left\{\text{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\right)-\text{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right) \\ -\text{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\right)+\text{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right)\right\} \\ = 2\left\{\text{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}\right)-\text{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right) \\ -\text{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right)+\text{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right)\right\} \\ = 2\left\{\text{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}\right)-\text{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right)\right\} \quad (37)$$

From (33), (34), (35), (36) and (37), we obtain the first summand of (32), as follows:

$$\mu^2\mathbb{E}\left\{\|\mathbf{V}+\mathbf{V}^H\|_F^2\right\} = 2\mu^2\mathbb{E}\left\{\text{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}\right) \\ -\text{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right)\right\} \quad (38)$$

In order to finish the expansion of the expressions inside braces of the second member of (32), now we deal with its second summand, in which we make use of the cyclic property [135] of the trace and the idempotence [1] [135] of $\hat{\mathbf{Q}}_A$.

$$\begin{aligned}
&\operatorname{Tr}\left[\left(\hat{\mathbf{R}}-\mathbf{R}\right)\left(\mathbf{V}+\mathbf{V}^H\right)\right] = \left\{\operatorname{Tr}\left(\hat{\mathbf{R}}-\mathbf{R}\right)\right.\\
&\left.\left[\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A^{\perp} + \left(\hat{\mathbf{Q}}_A\,\hat{\mathbf{R}}\,\hat{\mathbf{Q}}_A^{\perp}\right)^H\right]\right\}\\
&= \operatorname{Tr}\left\{\left(\hat{\mathbf{R}}-\mathbf{R}\right)\left[\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\left(\mathbf{I_M}-\hat{\mathbf{Q}}_A\right)\right.\right.\\
&\left.\left.+\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\left(\mathbf{I_M}-\hat{\mathbf{Q}}_A\right)\right)^H\right]\right\}\\
&= \operatorname{Tr}\left\{\left(\hat{\mathbf{R}}-\mathbf{R}\right)\left[\hat{\mathbf{Q}}_A\hat{\mathbf{R}}-\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\right.\right.\\
&\left.\left.+\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}-\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\right)^H\right]\right\}\\
&= \operatorname{Tr}\left\{\left(\hat{\mathbf{R}}-\mathbf{R}\right)\left[\hat{\mathbf{Q}}_A\hat{\mathbf{R}}-\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A+\hat{\mathbf{R}}\hat{\mathbf{Q}}_{\mathbf{A}}-\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\right]\right\}\\
&= \operatorname{Tr}\left\{\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}+\hat{\mathbf{R}}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{\mathbf{A}}-2\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\right.\\
&\left.-\mathbf{R}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}-\mathbf{R}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{\mathbf{A}}+2\mathbf{R}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\right\}\\
&= \operatorname{Tr}\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}+\operatorname{Tr}\hat{\mathbf{R}}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{\mathbf{A}}-2\operatorname{Tr}\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\\
&-\operatorname{Tr}\mathbf{R}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}-\operatorname{Tr}\mathbf{R}\hat{\mathbf{R}}\hat{\mathbf{Q}}_{\mathbf{A}}+2\operatorname{Tr}\mathbf{R}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\\
&= \operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}+\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}-2\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\\
&-\operatorname{Tr}\mathbf{R}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}-\operatorname{Tr}\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{R}}+2\operatorname{Tr}\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\\
&= 2\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}-2\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}-\operatorname{Tr}\mathbf{R}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\\
&-\operatorname{Tr}\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{R}}+2\operatorname{Tr}\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\\
&= 2\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}-2\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}-\operatorname{Tr}\mathbf{R}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\\
&-\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{R}}+2\operatorname{Tr}\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}
\end{aligned} \tag{39}$$

By using (39), we can straightforwardly write the second summand of the second member of (32) in terms of the projection matrices of the signal and the noise subspaces and the data covariance matrix as follows:

$$\begin{aligned}
&-2\mu\mathbb{E}\left\{\operatorname{Tr}\left[\left(\hat{\mathbf{R}}-\mathbf{R}\right)\left(\mathbf{V}\,+\,\mathbf{V}^H\right)\right]\right\}\\
&= -2\mu\mathbb{E}\left\{2\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}-2\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}-\operatorname{Tr}\mathbf{R}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right.\\
&\left.-\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{R}}+2\operatorname{Tr}\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\}\\
&= -4\mu\mathbb{E}\left\{\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}-\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\}\\
&-2\mu\left\{-\operatorname{Tr}\mathbb{E}\left[\mathbf{R}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right]-\operatorname{Tr}\mathbb{E}\left[\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{R}}\right]\right.\\
&\left.+2\operatorname{Tr}\mathbb{E}\left[\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right]\right\}\\
&= -4\mu\mathbb{E}\left\{\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}-\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\}\\
&-2\mu\left\{-\operatorname{Tr}R\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbb{E}\left[\hat{\mathbf{R}}\right]-\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R}\mathbb{E}\left[\hat{\mathbf{R}}\right]\right.\\
&\left.+2\operatorname{Tr}\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\mathbb{E}\left[\hat{\mathbf{R}}\right]\right\}
\end{aligned} \tag{40}$$

Now, by using (38) and (40), and assuming that $\mathbb{E}\left[\hat{\mathbf{R}}\right]$ is an unbiased estimate of $\hat{\mathbf{R}}$, i.e., $\mathbb{E}\left[\hat{\mathbf{R}}\right]=\mathbf{R}$, we can rewrite (32) as follows:

$$\begin{aligned}
&\operatorname{MSE}\left(\hat{\mathbf{R}}^{(n+1)}\right)\big|_{n=1}-\operatorname{MSE}\left(\hat{\mathbf{R}}\right)=\mu^2\mathbb{E}\left\{\|\mathbf{V}+\mathbf{V}^H\|_F^2\right\}\\
&-2\mu\mathbb{E}\left\{\operatorname{Tr}\left[\left(\hat{\mathbf{R}}-\mathbf{R}\right)\left(\mathbf{V}\,+\,\mathbf{V}^H\right)\right]\right\}\\
&=2\mu^2\mathbb{E}\left\{\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}-\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\}\\
&-4\mu\mathbb{E}\left\{\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}-\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\}\\
&-2\mu\left\{-\operatorname{Tr}\mathbf{R}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R}-\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R}\mathbf{R}\right.\\
&\left.+2\operatorname{Tr}\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\mathbf{R}\right\}\\
&=2\mu^2\mathbb{E}\left\{\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}-\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\}\\
&-4\mu\mathbb{E}\left\{\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}-\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\}\\
&-2\mu\left\{-2\operatorname{Tr}\mathbf{R}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R}+2\operatorname{Tr}\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\mathbf{R}\right\}\\
&=2\mu^2\mathbb{E}\left\{\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}-\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\}\\
&-4\mu\mathbb{E}\left\{\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}-\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\}\\
&-4\mu\left\{\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R}\mathbf{R}-\operatorname{Tr}\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\mathbf{R}\right\}\\
&=\left(2\mu^2-4\mu\right)\mathbb{E}\left\{\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}-\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right\}\\
&-4\mu\left\{\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R}\mathbf{R}-\operatorname{Tr}\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\mathbf{R}\right\}
\end{aligned} \tag{41}$$

Next, we will discuss equation (41). For this purpose, we assume that the estimate of the projection matrix of the signal subspace $\hat{\mathbf{Q}}_A$ [1], the true $\mathbf{R}$ [134] and the data covariance matrices $\hat{\mathbf{R}}$ [134] are Hermitian. For the next steps we will make use of the following Theorem which is proved in [137]:

Theorem 1: For two Hermitian matrices $\mathbf{A}$ and $\mathbf{B}$ of the same order,

$$\operatorname{Tr}\left(\mathbf{A}\mathbf{B}\right)^{2^k} \leq \operatorname{Tr}\left(\mathbf{A}^{2^k}\mathbf{B}^{2^k}\right), \tag{42}$$

where k is in integer.

By replacing $\mathbf{A}$ with $\hat{\mathbf{Q}}_A$ and $\mathbf{B}$ with $\hat{\mathbf{R}}$ in (42) and also considering $k=1$, we have

$$\begin{aligned}
&\operatorname{Tr}\left(\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\right)^2 \leq \operatorname{Tr}\left(\hat{\mathbf{Q}}_A^2\hat{\mathbf{R}}^2\right)\\
&\therefore \operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}} \leq \operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}\\
&\Rightarrow \operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{R}}-\operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{R}}\hat{\mathbf{Q}}_A\hat{\mathbf{R}} \geq 0
\end{aligned} \tag{43}$$

Similarly, making $\mathbf{A}=\hat{\mathbf{Q}}_A$ and $\mathbf{B}=\mathbf{R}$ for $k=1$, we obtain

$$\begin{aligned}
&\operatorname{Tr}\left(\hat{\mathbf{Q}}_A\mathbf{R}\right)^2 \leq \operatorname{Tr}\left(\hat{\mathbf{Q}}_A^2\mathbf{R}^2\right)\\
&\therefore \operatorname{Tr}\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\mathbf{R} \leq \operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R}\mathbf{R}\\
&\Rightarrow \operatorname{Tr}\hat{\mathbf{Q}}_A\hat{\mathbf{Q}}_A\mathbf{R}\mathbf{R}-\operatorname{Tr}\hat{\mathbf{Q}}_A\mathbf{R}\hat{\mathbf{Q}}_A\mathbf{R} \geq 0
\end{aligned} \tag{44}$$

Next, we analyze the behavior of the expressions -4μ and $\left(2\mu^2-4\mu\right)$ based on the reliability factor $\mu \in [0\ 1]$, as defined in (12). In order to illustrate the case being studied, we assume that both expressions are continuous functions as depicted in Fig. 7. It can be seen in it that in the range $[0\ 1]$ both

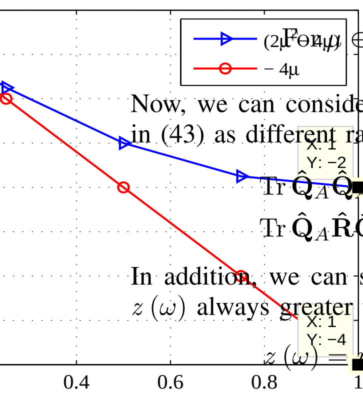


Fig. 7. Behavior of $(2\mu^2 - 4\mu)$ and -4μ for $\mu \in [0\ 1]$

expressions assume values f$(\mu) \leq 0$, i.e.:

$$\text{For } \mu \in [0\ 1] : \begin{cases} (2\mu^2 - 4\mu) \leq 0 \\ -4\mu \leq 0 \end{cases} \tag{45}$$

Now, we can consider the traces which form the subtraction in (43) as different random variables $y(\omega)$ and $x(\omega)$, i.e.:

$$\left.\begin{array}{l} \mathrm{Tr}\, \hat{\mathbf{Q}}_A \hat{\mathbf{Q}}_A \hat{\mathbf{R}} \hat{\mathbf{R}} = y(\omega) \\ \mathrm{Tr}\, \hat{\mathbf{Q}}_A \hat{\mathbf{R}} \hat{\mathbf{Q}}_A \hat{\mathbf{R}} = x(\omega) \end{array}\right\}, \ \forall \omega \ \in \Omega. \tag{46}$$

In addition, we can suppose that there is a random variable $z(\omega)$ always greater than zero, i.e., $z(\omega) \geq 0$, so that

$$z(\omega) = y(\omega) - x(\omega) \geq 0, \ \forall \omega \ \in \Omega \tag{47}$$

Taking the expectation of (47) and applying its properties of linearity and monotonicity [136], [140], we obtain

$$\mathbb{E}[z(\omega)] = \mathbb{E}[y(\omega) - x(\omega)] \geq 0, \tag{48}$$

which, by making use of (46), results in

$$\begin{aligned} &\mathbb{E}[z(\omega)] = \mathbb{E}[y(\omega) - x(\omega)] \\ &= \mathbb{E}\left\{ \mathrm{Tr}\, \hat{\mathbf{Q}}_A \hat{\mathbf{Q}}_A \hat{\mathbf{R}} \hat{\mathbf{R}} - \mathrm{Tr}\, \hat{\mathbf{Q}}_A \hat{\mathbf{R}} \hat{\mathbf{Q}}_A \hat{\mathbf{R}} \right\} \geq 0 \end{aligned} \tag{49}$$

Next, we can combine the inequalities (45) with (49) to compute the second member of (41), for $\mu \in [0\ 1]$.

For its first summand, we combine (45) and (49), as follows:

$$\begin{cases} \mathbb{E}\left\{ \mathrm{Tr}\, \hat{\mathbf{Q}}_A \hat{\mathbf{Q}}_A \hat{\mathbf{R}} \hat{\mathbf{R}} - \mathrm{Tr}\, \hat{\mathbf{Q}}_A \hat{\mathbf{R}} \hat{\mathbf{Q}}_A \hat{\mathbf{R}} \right\} \geq 0 \\ (2\mu^2 - 4\mu) \leq 0, \ \mu \in [0\ 1], \end{cases} \tag{50}$$

to obtain in a straightforward way

$$(2\mu^2 - 4\mu)\, \mathbb{E}\left\{ \mathrm{Tr}\, \hat{\mathbf{Q}}_A \hat{\mathbf{Q}}_A \hat{\mathbf{R}} \hat{\mathbf{R}} - \mathrm{Tr}\, \hat{\mathbf{Q}}_A \hat{\mathbf{R}} \hat{\mathbf{Q}}_A \hat{\mathbf{R}} \right\} \leq 0 \tag{51}$$

Similarly, we can compute its second member, by combining (45) and (44), as described by

$$\begin{cases} \mathrm{Tr}\, \hat{\mathbf{Q}}_A \hat{\mathbf{Q}}_A \mathbf{R} \mathbf{R} - \mathrm{Tr}\, \hat{\mathbf{Q}}_A \hat{\mathbf{R}} \hat{\mathbf{Q}}_A \hat{\mathbf{R}} \geq 0 \\ -4\mu \leq 0, \ \mu \in [0\ 1], \end{cases} \tag{52}$$

to obtain also straightforwardly the expression given by

$$-4\mu \left\{ \mathrm{Tr}\, \hat{\mathbf{Q}}_A \hat{\mathbf{Q}}_A \mathbf{R} \mathbf{R} - \mathrm{Tr}\, \hat{\mathbf{Q}}_A \mathbf{R} \hat{\mathbf{Q}}_A \mathbf{R} \right\} \leq 0 \tag{53}$$

By combining the inequalities (51) and (53) with (41), we have

$$\begin{aligned} &\mathrm{MSE}\left(\hat{\mathbf{R}}^{(n+1)}\right)\big|_{n=1} - \mathrm{MSE}\left(\hat{\mathbf{R}}\right) \\ &= \underbrace{(2\mu^2 - 4\mu)\, \mathbb{E}\left\{ \mathrm{Tr}\, \hat{\mathbf{Q}}_A \hat{\mathbf{Q}}_A \hat{\mathbf{R}} \hat{\mathbf{R}} - \mathrm{Tr}\, \hat{\mathbf{Q}}_A \hat{\mathbf{R}} \hat{\mathbf{Q}}_A \hat{\mathbf{R}} \right\}}_{\leq 0} \\ &\underbrace{-4\mu \left\{ \mathrm{Tr}\, \hat{\mathbf{Q}}_A \hat{\mathbf{Q}}_A \mathbf{R} \mathbf{R} - \mathrm{Tr}\, \hat{\mathbf{Q}}_A \mathbf{R} \hat{\mathbf{Q}}_A \mathbf{R} \right\}}_{\leq 0} \end{aligned} \tag{54}$$

$$\therefore\ \mathrm{MSE}\left(\hat{\mathbf{R}}^{(n+1)}\right)\big|_{n=1} - \mathrm{MSE}\left(\hat{\mathbf{R}}\right) \leq 0 \tag{55}$$

which is the desired result.

References

[1] H. L. Van Trees, *Detection, Estimation, and Modulation, Part IV, Optimum Array Processing*, John Wiley &Sons, 2002.

[2] R. Schmidt, "Multiple emitter location and signal parameter estimation" *IEEE Trans on Antennas and Propagation*, vol.34, No.3, Mar 1986, pp 276-280.

[3] A. J. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in Proc. ICASSP, Boston, MA, Apr. 1983, pp. 336–339.

[4] R. Roy and T. Kailath, "Estimation of signal parameters via rotational invariance techniques", *IEEE Trans. Acoust., Speech., Signal Processing*, vol. 37, July 1989, pp 984-995.

[5] L. L. Scharf and D. W. Tufts, "Rank reduction for modeling stationary signals," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-35, pp. 350-355, March 1987.

[6] A. M. Haimovich and Y. Bar-Ness, "An eigenanalysis interference canceler," *IEEE Trans. on Signal Processing*, vol. 39, pp. 76-84, Jan. 1991.

[7] D. A. Pados and S. N. Batalama "Joint space-time auxiliary vector filtering for DS/CDMA systems with antenna arrays" *IEEE Transactions on Communications*, vol. 47, no. 9, pp. 1406 - 1415, 1999.

[8] J. S. Goldstein, I. S. Reed and L. L. Scharf "A multistage representation of the Wiener filter based on orthogonal projections" *IEEE Transactions on Information Theory*, vol. 44, no. 7, 1998.

[9] Y. Hua, M. Nikpour and P. Stoica, "Optimal reduced rank estimation and filtering," IEEE Transactions on Signal Processing, pp. 457-469, Vol. 49, No. 3, March 2001.

[10] M. L. Honig and J. S. Goldstein, "Adaptive reduced-rank interference suppression based on the multistage Wiener filter," *IEEE Transactions on Communications*, vol. 50, no. 6, June 2002.

[11] E. L. Santos and M. D. Zoltowski, "On Low Rank MVDR Beamforming using the Conjugate Gradient Algorithm", *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, 2004.

[12] Q. Haoli and S.N. Batalama, "Data record-based criteria for the selection of an auxiliary vector estimator of the MMSE/MVDR filter", *IEEE Transactions on Communications*, vol. 51, no. 10, Oct. 2003, pp. 1700 - 1708.

[13] R. C. de Lamare and R. Sampaio-Neto, "Reduced-Rank Adaptive Filtering Based on Joint Iterative Optimization of Adaptive Filters", *IEEE Signal Processing Letters*, Vol. 14, no. 12, December 2007.

[14] Z. Xu and M.K. Tsatsanis, "Blind adaptive algorithms for minimum variance CDMA receivers," *IEEE Trans. Communications*, vol. 49, No. 1, January 2001.

[15] R. C. de Lamare and R. Sampaio-Neto, "Low-Complexity Variable Step-Size Mechanisms for Stochastic Gradient Algorithms in Minimum Variance CDMA Receivers", *IEEE Trans. Signal Processing*, vol. 54, pp. 2302 - 2317, June 2006.

[16] C. Xu, G. Feng and K. S. Kwak, "A Modified Constrained Constant Modulus Approach to Blind Adaptive Multiuser Detection," *IEEE Trans. Communications*, vol. 49, No. 9, 2001.

[17] Z. Xu and P. Liu, "Code-Constrained Blind Detection of CDMA Signals in Multipath Channels," *IEEE Sig. Proc. Letters*, vol. 9, No. 12, December 2002.

[18] R. C. de Lamare and R. Sampaio Neto, "Blind Adaptive Code-Constrained Constant Modulus Algorithms for CDMA Interference Suppression in Multipath Channels", *IEEE Communications Letters*, vol 9. no. 4, April, 2005.

[19] L. Landau, R. C. de Lamare and M. Haardt, "Robust adaptive beamforming algorithms using the constrained constant modulus criterion," IET Signal Processing, vol.8, no.5, pp.447-457, July 2014.

[20] R. C. de Lamare, "Adaptive Reduced-Rank LCMV Beamforming Algorithms Based on Joint Iterative Optimisation of Filters", *Electronics Letters*, vol. 44, no. 9, 2008.

[21] R. C. de Lamare and R. Sampaio-Neto, "Adaptive Reduced-Rank Processing Based on Joint and Iterative Interpolation, Decimation and Filtering", *IEEE Transactions on Signal Processing*, vol. 57, no. 7, July 2009, pp. 2503 - 2514.

[22] R. C. de Lamare and Raimundo Sampaio-Neto, "Reduced-rank Interference Suppression for DS-CDMA based on Interpolated FIR Filters", *IEEE Communications Letters*, vol. 9, no. 3, March 2005.

[23] R. C. de Lamare and R. Sampaio-Neto, "Adaptive Reduced-Rank MMSE Filtering with Interpolated FIR Filters and Adaptive Interpolators", *IEEE Signal Processing Letters*, vol. 12, no. 3, March, 2005.

[24] R. C. de Lamare and R. Sampaio-Neto, "Adaptive Interference Suppression for DS-CDMA Systems based on Interpolated FIR Filters with Adaptive Interpolators in Multipath Channels", *IEEE Trans. Vehicular Technology*, Vol. 56, no. 6, September 2007.

[25] R. C. de Lamare, "Adaptive Reduced-Rank LCMV Beamforming Algorithms Based on Joint Iterative Optimisation of Filters," Electronics Letters, 2008.

[26] R. C. de Lamare and R. Sampaio-Neto, "Reduced-rank adaptive filtering based on joint iterative optimization of adaptive filters", *IEEE Signal Process. Lett.*, vol. 14, no. 12, pp. 980-983, Dec. 2007.

[27] R. C. de Lamare, M. Haardt, and R. Sampaio-Neto, "Blind Adaptive Constrained Reduced-Rank Parameter Estimation based on Constant Modulus Design for CDMA Interference Suppression", *IEEE Transactions on Signal Processing*, June 2008.

[28] M. Yukawa, R. C. de Lamare and R. Sampaio-Neto, "Efficient Acoustic Echo Cancellation With Reduced-Rank Adaptive Filtering Based on Selective Decimation and Adaptive Interpolation," IEEE Transactions on Audio, Speech, and Language Processing, vol.16, no. 4, pp. 696-710, May 2008.

[29] R. C. de Lamare and R. Sampaio-Neto, "Reduced-rank space-time adaptive interference suppression with joint iterative least squares algorithms for spread-spectrum systems," *IEEE Trans. Vehi. Technol.*, vol. 59, no. 3, pp. 1217-1228, Mar. 2010.

[30] R. C. de Lamare and R. Sampaio-Neto, "Adaptive reduced-rank equalization algorithms based on alternating optimization design techniques for MIMO systems," *IEEE Trans. Vehi. Technol.*, vol. 60, no. 6, pp. 2482-2494, Jul. 2011.

[31] R. C. de Lamare, L. Wang, and R. Fa, "Adaptive reduced-rank LCMV beamforming algorithms based on joint iterative optimization of filters: Design and analysis," Signal Processing, vol. 90, no. 2, pp. 640-652, Feb. 2010.

[32] R. Fa, R. C. de Lamare, and L. Wang, "Reduced-Rank STAP Schemes for Airborne Radar Based on Switched Joint Interpolation, Decimation and Filtering Algorithm," *IEEE Transactions on Signal Processing*, vol.58, no.8, Aug. 2010, pp.4182-4194.

[33] L. Wang and R. C. de Lamare, "Low-Complexity Adaptive Step Size Constrained Constant Modulus SG Algorithms for Blind Adaptive Beamforming", *Signal Processing*, vol. 89, no. 12, December 2009, pp. 2503-2513.

[34] L. Wang and R. C. de Lamare, "Adaptive Constrained Constant Modulus Algorithm Based on Auxiliary Vector Filtering for Beamforming," IEEE Transactions on Signal Processing, vol. 58, no. 10, pp. 5408-5413, Oct. 2010.

[35] L. Wang, R. C. de Lamare, M. Yukawa, "Adaptive Reduced-Rank Constrained Constant Modulus Algorithms Based on Joint Iterative Optimization of Filters for Beamforming," *IEEE Transactions on Signal Processing*, vol.58, no.6, June 2010, pp.2983-2997.

[36] L. Wang, R. C. de Lamare and M. Yukawa, "Adaptive reduced-rank constrained constant modulus algorithms based on joint iterative optimization of filters for beamforming", IEEE Transactions on Signal Processing, vol.58, no. 6, pp. 2983-2997, June 2010.

[37] L. Wang and R. C. de Lamare, "Adaptive constrained constant modulus algorithm based on auxiliary vector filtering for beamforming", IEEE Transactions on Signal Processing, vol. 58, no. 10, pp. 5408-5413, October 2010.

[38] R. Fa and R. C. de Lamare, "Reduced-Rank STAP Algorithms using Joint Iterative Optimization of Filters," IEEE Transactions on Aerospace and Electronic Systems, vol.47, no.3, pp.1668-1684, July 2011.

[39] Z. Yang, R. C. de Lamare and X. Li, "L1-Regularized STAP Algorithms With a Generalized Sidelobe Canceler Architecture for Airborne Radar," IEEE Transactions on Signal Processing, vol.60, no.2, pp.674-686, Feb. 2012.

[40] Z. Yang, R. C. de Lamare and X. Li, "Sparsity-aware space–time adaptive processing algorithms with L1-norm regularisation for airborne radar", IET signal processing, vol. 6, no. 5, pp. 413-423, 2012.

[41] Neto, F.G.A.; Nascimento, V.H.; Zakharov, Y.V.; de Lamare, R.C., "Adaptive re-weighting homotopy for sparse beamforming," in Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European , vol., no., pp.1287-1291, 1-5 Sept. 2014

[42] Almeida Neto, F.G.; de Lamare, R.C.; Nascimento, V.H.; Zakharov, Y.V.,"Adaptive reweighting homotopy algorithms applied to beamforming," IEEE Transactions on Aerospace and Electronic Systems, vol.51, no.3, pp.1902-1915, July 2015.

[43] L. Wang, R. C. de Lamare and M. Haardt, "Direction finding algorithms based on joint iterative subspace optimization," IEEE Transactions on Aerospace and Electronic Systems, vol.50, no.4, pp.2541-2553, October 2014.

[44] S. D. Somasundaram, N. H. Parsons, P. Li and R. C. de Lamare, "Reduced-dimension robust capon beamforming using Krylov-subspace techniques," IEEE Transactions on Aerospace and Electronic Systems, vol.51, no.1, pp.270-289, January 2015.

[45] S. Xu and R.C de Lamare, , *Distributed conjugate gradient strategies for distributed estimation over sensor networks*, Sensor Signal Processing for Defense SSPD, September 2012.

[46] S. Xu, R. C. de Lamare, H. V. Poor, "Distributed Estimation Over Sensor Networks Based on Distributed Conjugate Gradient Strategies", IET Signal Processing, 2016 (to appear).

[47] S. Xu, R. C. de Lamare and H. V. Poor, *Distributed Compressed Estimation Based on Compressive Sensing*, IEEE Signal Processing letters, vol. 22, no. 9, September 2014.

[48] S. Xu, R. C. de Lamare and H. V. Poor, "Distributed reduced-rank estimation based on joint iterative optimization in sensor networks," in Proceedings of the 22nd European Signal Processing Conference (EUSIPCO), pp.2360-2364, 1-5, Sept. 2014

[49] S. Xu, R. C. de Lamare and H. V. Poor, "Adaptive link selection strategies for distributed estimation in diffusion wireless networks," in Proc. IEEE International Conference onAcoustics, Speech and Signal Processing (ICASSP), , vol., no., pp.5402-5405, 26-31 May 2013.

[50] S. Xu, R. C. de Lamare and H. V. Poor, "Dynamic topology adaptation for distributed estimation in smart grids," in Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2013 IEEE 5th International Workshop on , vol., no., pp.420-423, 15-18 Dec. 2013.

[51] S. Xu, R. C. de Lamare and H. V. Poor, "Adaptive Link Selection Algorithms for Distributed Estimation", EURASIP Journal on Advances in Signal Processing, 2015.

[52] N. Song, R. C. de Lamare, M. Haardt, and M. Wolf, "Adaptive Widely Linear Reduced-Rank Interference Suppression based on the Multi-Stage Wiener Filter," IEEE Transactions on Signal Processing, vol. 60, no. 8, 2012.

[53] N. Song, W. U. Alokozai, R. C. de Lamare and M. Haardt, "Adaptive Widely Linear Reduced-Rank Beamforming Based on Joint Iterative Optimization," IEEE Signal Processing Letters, vol.21, no.3, pp. 265-269, March 2014.

[54] R.C. de Lamare, R. Sampaio-Neto and M. Haardt, "Blind Adaptive Constrained Constant-Modulus Reduced-Rank Interference Suppression Algorithms Based on Interpolation and Switched Decimation," *IEEE Trans. on Signal Processing*, vol.59, no.2, pp.681-695, Feb. 2011.

[55] Y. Cai, R. C. de Lamare, "Adaptive Linear Minimum BER Reduced-Rank Interference Suppression Algorithms Based on Joint and Iterative Optimization of Filters," IEEE Communications Letters, vol.17, no.4, pp.633-636, April 2013.

[56] R. C. de Lamare and R. Sampaio-Neto, "Sparsity-Aware Adaptive Algorithms Based on Alternating Optimization and Shrinkage," IEEE Signal Processing Letters, vol.21, no.2, pp.225,229, Feb. 2014.

[57] J. Steinwandt, R. C. de Lamare and M. Haardt, "Beamspace direction finding based on the conjugate gradient and the auxiliary vector filtering algorithms", Signal Processing, vol. 93, no. 4, April 2013, pp. 641-651.

[58] L. Wang, R. C. de Lamare and M. Haardt, "Direction finding algorithms based on joint iterative subspace optimization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 4, pp. 2541-2553, October 2014.

[59] L. Qiu, Y. Cai, R. C. de Lamare and M. Zhao, "Reduced-Rank DOA Estimation Algorithms Based on Alternating Low-Rank Decomposition," *IEEE Signal Processing Letters*, vol. 23, no. 5, pp. 565-569, May 2016.

[60] 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.

[61] 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.

[62] 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.

[63] H. Ruan and R. C. de Lamare, "Robust Adaptive Beamforming Using a Low-Complexity Shrinkage-Based Mismatch Estimation Algorithm," IEEE Signal Processing Letters, vol. 21, no. 1, pp. 60-64, Jan. 2014.

[64] H. Ruan and R. C. de Lamare, "Robust Adaptive Beamforming Based on Low-Rank and Cross-Correlation Techniques," IEEE Transactions on Signal Processing, vol. 64, no. 15, pp. 3919-3932, Aug.1, 1 2016.

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

[66] 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. Lexington, MA, USA: MIT Lincoln Lab., June 1982.

[67] 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.

[68] 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.

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

[70] 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.

[71] R. C. de Lamare, "Massive MIMO Systems: Signal Processing Challenges and Future Trends", Radio Science Bulletin, December 2013.

[72] W. Zhang, H. Ren, C. Pan, M. Chen, R. C. de Lamare, B. Du and J. Dai, "Large-Scale Antenna Systems With UL/DL Hardware Mismatch: Achievable Rates Analysis and Calibration", IEEE Trans. Commun., vol.63, no.4, pp. 1216-1229, April 2015.

[73] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, no. 3, pp. 439-441, May 1983.

[74] R. C. de Lamare and A. Alcaim, "Strategies to improve the performance of very low bit rate speech coders and application to a 1.2 kb/s codec" IEE Proceedings- Vision, image and signal processing, vol. 152, no. 1, February, 2005.

[75] P. Clarke and R. C. de Lamare, "Joint Transmit Diversity Optimization and Relay Selection for Multi-Relay Cooperative MIMO Systems Using Discrete Stochastic Algorithms," *IEEE Communications Letters*, vol.15, no.10, pp.1035-1037, October 2011.

[76] P. Clarke and R. C. de Lamare, "Transmit Diversity and Relay Selection Algorithms for Multirelay Cooperative MIMO Systems" *IEEE Transactions on Vehicular Technology*, vol.61, no. 3, pp. 1084-1098, October 2011.

[77] Y. Cai, R. C. de Lamare, and R. Fa, "Switched Interleaving Techniques with Limited Feedback for Interference Mitigation in DS-CDMA Systems," IEEE Transactions on Communications, vol.59, no.7, pp.1946-1956, July 2011.

[78] Y. Cai, R. C. de Lamare, D. Le Ruyet, "Transmit Processing Techniques Based on Switched Interleaving and Limited Feedback for Interference Mitigation in Multiantenna MC-CDMA Systems," IEEE Transactions on Vehicular Technology, vol.60, no.4, pp.1559-1570, May 2011.

[79] T. Wang, R. C. de Lamare, and P. D. Mitchell, "Low-Complexity Set-Membership Channel Estimation for Cooperative Wireless Sensor Networks," IEEE Transactions on Vehicular Technology, vol.60, no.6, pp.2594-2607, July 2011.

[80] T. Wang, R. C. de Lamare and A. Schmeink, "Joint linear receiver design and power allocation using alternating optimization algorithms for wireless sensor networks," *IEEE Trans. on Vehi. Tech.*, vol. 61, pp. 4129-4141, 2012.

[81] R. C. de Lamare, "Joint iterative power allocation and linear interference suppression algorithms for cooperative DS-CDMA networks", IET Communications, vol. 6, no. 13 , 2012, pp. 1930-1942.

[82] T. Peng, R. C. de Lamare and A. Schmeink, "Adaptive Distributed Space-Time Coding Based on Adjustable Code Matrices for Cooperative MIMO Relaying Systems", *IEEE Transactions on Communications*, vol. 61, no. 7, July 2013.

[83] T. Peng and R. C. de Lamare, "Adaptive Buffer-Aided Distributed Space-Time Coding for Cooperative Wireless Networks," IEEE Transactions on Communications, vol. 64, no. 5, pp. 1888-1900, May 2016.

[84] J. Gu, R. C. de Lamare and M. Huemer, "Buffer-Aided Physical-Layer Network Coding with Optimal Linear Code Designs for Cooperative Networks," IEEE Transactions on Communications, 2018.

[85] K. Zu, R. C. de Lamare, "Low-Complexity Lattice Reduction-Aided Regularized Block Diagonalization for MU-MIMO Systems", IEEE. Communications Letters, Vol. 16, No. 6, June 2012, pp. 925-928.

[86] K. Zu, R. C. de Lamare, "Low-Complexity Lattice Reduction-Aided Regularized Block Diagonalization for MU-MIMO Systems", IEEE. Communications Letters, Vol. 16, No. 6, June 2012.

[87] K. Zu, R. C. de Lamare and M. Haart, "Generalized design of low-complexity block diagonalization type precoding algorithms for multiuser MIMO systems", IEEE Trans. Communications, 2013.

[88] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electronic Letters*, vol. 7, Mar. 1971.

[89] C. T. Healy and R. C. de Lamare, "Decoder-optimised progressive edge growth algorithms for the design of LDPC codes with low error floors", *IEEE Communications Letters*, vol. 16, no. 6, June 2012, pp. 889-892.

[90] A. G. D. Uchoa, C. T. Healy, R. C. de Lamare, R. D. Souza, "LDPC codes based on progressive edge growth techniques for block fading channels", *Proc. 8th International Symposium on Wireless Communication Systems (ISWCS)*, 2011, pp. 392-396.

[91] A. G. D. Uchoa, C. T. Healy, R. C. de Lamare, R. D. Souza, "Generalised Quasi-Cyclic LDPC codes based on progressive edge growth techniques for block fading channels", *Proc. International Symposium Wireless Communication Systems (ISWCS)*, 2012, pp. 974-978.

[92] A. G. D. Uchoa, C. T. Healy, R. C. de Lamare, R. D. Souza, "Design of LDPC Codes Based on Progressive Edge Growth Techniques for Block Fading Channels", *IEEE Communications Letters*, vol. 15, no. 11, November 2011, pp. 1221-1223.

[93] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Trans. Commun.*, vol. 20, Aug. 1972.

[94] K. Zu, R. C. de Lamare and M. Haardt, "Multi-branch tomlinson-harashima precoding for single-user MIMO systems," in Smart Antennas (WSA), 2012 International ITG Workshop on , vol., no., pp.36-40, 7-8 March 2012.

[95] K. Zu, R. C. de Lamare and M. Haardt, "Multi-Branch Tomlinson-Harashima Precoding Design for MU-MIMO Systems: Theory and Algorithms," IEEE Transactions on Communications, vol.62, no.3, pp.939,951, March 2014.

[96] L. Zhang, Y. Cai, R. C. de Lamare and M. Zhao, "Robust Multi-branch Tomlinson–Harashima Precoding Design in Amplify-and-Forward MIMO Relay Systems," IEEE Transactions on Communications, vol.62, no.10, pp.3476,3490, Oct. 2014.

[97] W. Zhang et al., "Widely Linear Precoding for Large-Scale MIMO with IQI: Algorithms and Performance Analysis," IEEE Transactions on Wireless Communications, vol. 16, no. 5, pp. 3298-3312, May 2017.

[98] B. Hochwald, C. Peel and A. Swindlehurst, "A vector-perturbation technique for near capacity multiantenna multiuser communication - Part II: Perturbation," *IEEE Trans. Commun.*, vol. 53, no. 3, Mar. 2005.

[99] C. B. Chae, S. Shim and R. W. Heath, "Block diagonalized vector perturbation for multiuser MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4051 - 4057, Nov. 2008.

[100] R. C. de Lamare, R. Sampaio-Neto, "Adaptive MBER decision feedback multiuser receivers in frequency selective fading channels", *IEEE Communications Letters*, vol. 7, no. 2, Feb. 2003, pp. 73 - 75.

[101] A. Rontogiannis, V. Kekatos, and K. Berberidis," A Square-Root Adaptive V-BLAST Algorithm for Fast Time-Varying MIMO Channels," *IEEE Signal Processing Letters*, Vol. 13, No. 5, pp. 265-268, May 2006.

[102] R. C. de Lamare, R. Sampaio-Neto, A. Hjorungnes, "Joint iterative interference cancellation and parameter estimation for CDMA systems", *IEEE Communications Letters*, vol. 11, no. 12, December 2007, pp. 916 - 918.

[103] Y. Cai and R. C. de Lamare, "Adaptive Space-Time Decision Feedback Detectors with Multiple Feedback Cancellation", *IEEE Transactions on Vehicular Technology*, vol. 58, no. 8, October 2009, pp. 4129 - 4140.

[104] J. W. Choi, A. C. Singer, J Lee, N. I. Cho, "Improved linear soft-input soft-output detection via soft feedback successive interference cancellation," *IEEE Trans. Commun.*, vol.58, no.3, pp.986-996, March 2010.

[105] R. C. de Lamare and R. Sampaio-Neto, "Blind adaptive MIMO receivers for space-time block-coded DS-CDMA systems in multipath channels using the constant modulus criterion," IEEE Transactions on Communications, vol.58, no.1, pp.21-27, January 2010.

[106] R. Fa, R. C. de Lamare, "Multi-Branch Successive Interference Cancellation for MIMO Spatial Multiplexing Systems", *IET Communications*, vol. 5, no. 4, pp. 484 - 494, March 2011.

[107] R.C. de Lamare and R. Sampaio-Neto, "Adaptive reduced-rank equalization algorithms based on alternating optimization design techniques for MIMO systems," IEEE Trans. Veh. Technol., vol. 60, no. 6, pp. 2482-2494, July 2011.

[108] P. Li, R. C. de Lamare and R. Fa, "Multiple Feedback Successive Interference Cancellation Detection for Multiuser MIMO Systems," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2434 - 2439, August 2011.

[109] R.C. de Lamare, R. Sampaio-Neto, "Minimum mean-squared error iterative successive parallel arbitrated decision feedback detectors for

DS-CDMA systems," IEEE Trans. Commun., vol. 56, no. 5, May 2008, pp. 778-789.

[110] R.C. de Lamare, R. Sampaio-Neto, "Minimum mean-squared error iterative successive parallel arbitrated decision feedback detectors for DS-CDMA systems," IEEE Trans. Commun., vol. 56, no. 5, May 2008.

[111] R.C. de Lamare and R. Sampaio-Neto, "Adaptive reduced-rank equalization algorithms based on alternating optimization design techniques for MIMO systems," IEEE Trans. Veh. Technol., vol. 60, no. 6, pp. 2482-2494, July 2011.

[112] P. Li, R. C. de Lamare and J. Liu, "Adaptive Decision Feedback Detection with Parallel Interference Cancellation and Constellation Constraints for Multiuser MIMO systems", IET Communications, vol.7, 2012, pp. 538-547.

[113] J. Liu, R. C. de Lamare, "Low-Latency Reweighted Belief Propagation Decoding for LDPC Codes," IEEE Communications Letters, vol. 16, no. 10, pp. 1660-1663, October 2012.

[114] C. T. Healy and R. C. de Lamare, "Design of LDPC Codes Based on Multipath EMD Strategies for Progressive Edge Growth," IEEE Transactions on Communications, vol. 64, no. 8, pp. 3208-3219, Aug. 2016.

[115] P. Li and R. C. de Lamare, Distributed Iterative Detection With Reduced Message Passing for Networked MIMO Cellular Systems, IEEE Transactions on Vehicular Technology, vol.63, no.6, pp. 2947-2954, July 2014.

[116] A. G. D. Uchoa, C. T. Healy and R. C. de Lamare, "Iterative Detection and Decoding Algorithms For MIMO Systems in Block-Fading Channels Using LDPC Codes," IEEE Transactions on Vehicular Technology, 2015.

[117] R. C. de Lamare, "Adaptive and Iterative Multi-Branch MMSE Decision Feedback Detection Algorithms for Multi-Antenna Systems", *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, October 2013.

[118] A. G. D. Uchoa, C. T. Healy and R. C. de Lamare, "Iterative Detection and Decoding Algorithms for MIMO Systems in Block-Fading Channels Using LDPC Codes," IEEE Transactions on Vehicular Technology, vol. 65, no. 4, pp. 2735-2741, April 2016.

[119] Y. Cai, R. C. de Lamare, B. Champagne, B. Qin and M. Zhao, "Adaptive Reduced-Rank Receive Processing Based on Minimum Symbol-Error-Rate Criterion for Large-Scale Multiple-Antenna Systems," in IEEE Transactions on Communications, vol. 63, no. 11, pp. 4185-4201, Nov. 2015.

[120] Z. Shao, R. C. de Lamare and L. T. N. Landau, "Iterative Detection and Decoding for Large-Scale Multiple-Antenna Systems with 1-Bit ADCs," IEEE Wireless Communications Letters, 2018.

[121] M. Shaghaghi and S. A. Vorobyov, "Subspace leakage analysis and improved DOA estimation with small sample size", IEEE Trans. Signal Process., vol. 63, no.12, pp 3251-3265, Jun.2015.

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

[123] 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.

[124] S. F. B. Pinto, R. C. de Lamare, Two-Step Knowledge-aided Iterative ESPRIT Algorithm, Twenty First ITG Workshop on Smart Antennas, 15-17 March 2017, Berlin, Germany.

[125] W. L. Melvin and J. R. Guerci, "Knowledge-aided signal processing: a new paradigm for radar and other advanced sensors," IEEE Transactions on Aerospace and Electronic Systems, vol. 42, no. 3, pp. 983–996, July 2006.

[126] W. L. Melvin and G. A. Showman, "An approach to knowledge-aided covariance estimation," IEEE Transactions on Aerospace and Electronic Systems, vol. 42, no. 3, pp.1021–1042, July 2006.

[127] J. S. Bergin, C. M. Teixeira, P. M. Techau, and J. R. Guerci, "Improved clutter mitigation performance using knowledge-aided space-time adaptive processing," IEEE Transactions on Aerospace and Electronic Systems, vol. 42, no. 3, pp. 997–1009, July 2006.

[128] J. Steinwandt, R. C. de Lamare and M. Haardt, "Knowledge-aided direction finding based on Unitary ESPRIT," 2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), Pacific Grove, CA, 2011, pp. 613-617.

[129] P. Stoica, J. Li, X. Zhu, and J. R. Guerci, "On using a priori knowledge in space-time adaptive processing," IEEE Transactions on Signal Processing, vol. 56, no. 6, pp. 2598-2602, June 2008.

[130] G. Bouleux, P. Stoica, and R. Boyer, "An optimal prior knowledge-based DOA estimation method," in 17th European Signal Processing Conference (EUSIPCO), Aug. 2009, pp. 869-873.

[131] P. Stoica, J. Li, X. Zhu, and J. R. Guerci, "On using a priori knowledge in spacetime adaptive processing," IEEE Transactions on Signal Processing, vol. 56, no. 6, pp. 2598–2602, June 2008.

[132] S. F. B. Pinto, R. C. de Lamare, "Multi-Step Knowledge-Aided Iterative ESPRIT for Direction Finding", submitted to 22nd International Conference on Digital Signal Processing, 23-25 August 2017 (DSP 2017), London, United Kingdom.

[133] S. F. B. Pinto and R. C. de Lamare, "Multi-Step Knowledge-Aided Iterative ESPRIT: Design and Analysis," IEEE Transactions on Aerospace and Electronic Systems, 2018.

[134] Simon Haykin, Adaptive Filter Theory, fourth edition, 2003.

[135] F.A. Graybill, Matrices with Applications in Statistics, Wadsworth Publishing Company, Inc., Second Edition, 1983.

[136] Alan F. Karr, Probability, Springer-Verlag NY, Second Edition, 1993.

[137] Da-Wei Chang, "A Matrix Trace Inequality for Products of Hermitian Matrices", Journal of Mathematical Analysis and Applications 237, pp.721-725, 1999.

[138] H.Semira, H.Belkacemi, S.Marcos, "High-resolution source localization algorithm based on the conjugate gradient", EURASIP Journal on Advances in Signal Processing, 2007(2)(2007)1–9.

[139] R.Grover, D.A.Pados, M.J. Medley, "Subspace direction finding with an auxiliary-vector basis", IEEE Transactions on Signal Processing, 55 (2) (2007) pp.758–763.

[140] J.P.A. Almeida, J.M.P. Fortes, W.A. Finamore, Probability, Random Variables and Stochastic Processes, PUC-RIO/Interciencia, 2008.

[141] P.Stoica and A.B.Gershman, "Maximum-likelihood DOA estimation by data-supported grid search", IEEE Signal Processing Letters, vol. 6, no. 10, pp. 273- 275, Oct 1999.

[142] P.Stoica and Arye Nehorai, "MUSIC, maximum Likelihood, and Cramer-Rao Bound", IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 37,no. 5, pp. 720- 741, May 1989.

[143] J.C.Liberti Jr, Theodore S. rappaport, "Smart antennas for Wireless Communications: IS-95 and Third Generation CDMA Applications", Chapter 9, pp 253-284, Prentice Hall, 1999.

[144] S.V.Schell, W.A. Gardner, "High Resolution Direction Finding", Chapter 17, K. Bose and C.R. Rao, pp 755-817,1993.

[145] J.Rissanen, "Modeling by the Shortest Data Description", Automatica, Vol.14, pp 465-471,1978.

[146] F. Li, R. J. Vaccaro, "Analysis of Min-Norm and MUSIC with arbitrary array geometry", *IEEE Transactions on Aerospace and Electronic Systems*, vol.26, no. 6, pp 976-985, 1990.

[147] M. Haardt, "Efficient one-, two-, and multidimensional high-resolution array signal processing", Ph.D. dissertation, Munich University of Technology, Shaker Verlag, 1996.