A Subspace Method for Array Covariance Matrix Estimation

Mostafa Rahmani and George K. Atia, *Member, IEEE,*

Abstract—This paper introduces a subspace method for the estimation of an array covariance matrix. It is shown that when the received signals are uncorrelated, the true array covariance matrices lie in a specific subspace whose dimension is typically much smaller than the dimension of the full space. Based on this idea, a subspace based covariance matrix estimator is proposed. The estimator is obtained as a solution to a semi-definite convex optimization problem. While the optimization problem has no closed-form solution, a nearly optimal closed-form solution is proposed making it easy to implement. In comparison to the conventional approaches, the proposed method yields higher estimation accuracy because it eliminates the estimation error which does not lie in the subspace of the true covariance matrices. The numerical examples indicate that the proposed covariance matrix estimator can significantly improve the estimation quality of the covariance matrix.

Index Terms—Covariance matrix estimation, subspace method, array signal processing, semidefinite optimization.

I. INTRODUCTION

THE estimation of covariance matrices is a crucial component of many signal processing algorithms [1-4]. In many applications, there is a limited number of snapshots and the sample covariance matrix cannot yield the desired estimation accuracy. This covariance matrix estimation error significantly degrades the performance of such algorithms. In some applications, the true covariance matrix has a specific structure. For example, the array covariance matrix of a linear array with equally spaced antenna elements is a Toeplitz matrix when the sources are uncorrelated [5, 6]. Moreover, in some applications [4, 8], the structure of the problem suggests that the underlying true covariance matrix is the Kronecker product of two valid covariance matrices [4, 7].

This side information can be leveraged in covariance matrix estimation to improve the estimation quality. For instance, in [5] a weighted least square estimator for covariance matrices with Toeplitz structures was proposed and it was shown that the resulting covariance matrix can enhance the performance of angle estimation algorithms, such as MUltiple SIgnals Classification (MUSIC) [13]. In [8], covariance matrices with Kronecker structure are investigated and a maximum likelihood based algorithm is introduced.

In addition, the structure of covariance matrices has been exploited in various DOA estimation algorithms, such as the linear structure in [9], and the diagonal structure for the covariance matrix of uncorrelated signals in [10]. Recently

some research works have focused on the application of sparse signal processing in DOA estimation based on the sparse representation of the array covariance matrix. For example, [11] proposes the idea that the eigenvectors of the array covariance matrix have a sparse representation over a dictionary constructed from the steering vectors. In [12, 14], it is shown that when the received signals are uncorrelated, the array covariance matrix has a sparse representation over a dictionary constructed using the atoms, i.e. the correlation vectors. A similar idea is proposed in [15], with the difference that the proposed method does not require choosing a hyper-parameter.

In this paper, we focus on the estimation of array covariance matrices with linear structure. First, we show that when the sources are uncorrelated, the array covariance matrix has a linear structure implying that all possible array covariance matrices can be described by a specific subspace. Based on this idea, a subspace-based covariance matrix estimator is proposed as a solution to a semi-definite convex optimization problem. Furthermore, we propose a nearly optimal closed-form solution for the proposed covariance matrix estimator. Our results show that the proposed method can noticeably improve the covariance matrix estimation quality. Moreover, the closed-form solution is shown to closely approach the optimal performance.

II. ARRAY SIGNAL MODEL

The system model under consideration is a narrowband array system with N antennas. All the signals are assumed to be narrowband with the same center frequency and impinge on the array from the far field. The baseband array output can be expressed as

$$\mathbf{x}(t) = \sum_{i=1}^{p} z_i(t)\mathbf{v}(\theta_i, \phi_i) + \mathbf{n}(t) \tag{1}$$

where $\mathbf{x}(t)$ is the $N \times 1$ array output vector, p is the number of the received signals, $z_i(t)$ is the ith signal, (θ_i, ϕ_i) is the elevation and azimuth arrival angle of the ith signal, $\mathbf{v}(\theta_i, \phi_i)$ is the baseband array response to i^{th} signal and $\mathbf{n}(t)$ is the noise vector. The baseband array response, $\mathbf{v}(\theta_i, \phi_i)$, is called the "steering vector" [13].

If the received signals are uncorrelated, the covariance matrix can be written as

$$\mathbf{R} = \sum_{i=1}^{p} \sigma_i^2 \mathbf{v}(\theta_i, \phi_i)\mathbf{v}^H(\theta_i, \phi_i) + \sigma_n^2 \mathbf{I} \tag{2}$$

where σ_i^2 represents the power of the ith signal, σ_n^2 is the noise variance and $\mathbf{I}$ is the identity matrix.

III. PROPOSED ALGORITHM

We define the "correlation vector" which belongs to direction (θ, ϕ) as follows

$$\mathbf{c}(\theta, \phi) = vec(\mathbf{v}(\theta, \phi)\mathbf{v}^H(\theta, \phi)) \tag{3}$$

where $vec(\bullet)$ is a linear transformation that converts its matrix argument to a vector by stacking the columns of the matrix on top of one another. Consequently, the covariance matrix can be rewritten as

$$vec(\mathbf{R} - \sigma_n^2\mathbf{I}) = \sum_{i=1}^{p} \sigma_i^2 \mathbf{c}(\theta_i, \phi_i) \tag{4}$$

Therefore, $vec(\mathbf{R} - \sigma_n^2\mathbf{I})$ is a linear combination of the correlation vectors of the received signals.

According to (4), $vec(\mathbf{R} - \sigma_n^2\mathbf{I})$ lies in the subspace of the correlation vectors. Hence, if we build the subspace spanned by all possible correlation vectors $\{\mathbf{c}(\theta, \phi)|0 \le \theta \le \pi, 0 \le \phi \le 2\pi\}$, then $vec(\mathbf{R} - \sigma_n^2\mathbf{I})$ completely lies in this subspace. For many array structures, the matrix $(\mathbf{v}(\theta, \phi)\mathbf{v}^H(\theta, \phi))$ inherits some symmetry properties. Accordingly, the correlation vectors cannot span an N^2 dimensional space. For example, when the incoming signals are uncorrelated, the covariance matrix of a uniform linear array is a Toeplitz matrix [5]. It is easy to show that all the $N \times N$ Toeplitz matrices can be described by a $2N - 1$ dimensional space.

The subspace of the correlation vectors $\{\mathbf{c}(\theta, \phi)|0 \le \theta \le \pi, 0 \le \phi \le 2\pi\}$ can be obtained by constructing a positive definite matrix

$$\mathbf{S} = \int_0^{2\pi} \int_0^{\pi} \mathbf{c}(\theta, \phi)\mathbf{c}^H(\theta, \phi)d\theta d\phi \tag{5}$$

where (5) is an element-wise integral. Based on (5), the subspace dimension of the correlation vectors $\{\mathbf{c}(\theta, \phi)|0 \le \theta \le \pi, 0 \le \phi \le 2\pi\}$ is equal to the number of non-zero eigenvalues of the matrix $\mathbf{S}$. Consequently, the subspace of the correlation vectors can be constructed using the eigenvectors which correspond to the non-zero eigenvalues.
Fig. 1 shows the eigenvalues of $\mathbf{S}$ for a square planar array with 16 elements (the horizontal and vertical space between the elements is half a wavelength). One can observe that the number of non-zero eigenvalues is equal to 49. Therefore, for this array, the subspace of the correlation vectors can be constructed from the 49 eigenvectors corresponding to the non-zero eigenvalues. Note that for a 16-element linear array, we observe 31 non-zeros eigenvalues because the covariance matrix is a Toeplitz matrix[5]. For some array structures such as circular array, we may not observe zero eigenvalues but our investigation has shown that the subspace of the correlation vectors can be effectively approximated using the dominant eigenvectors (the eigenvectors corresponding to the dominant eigenvalues). Therefore, if we construct the matrix $\mathbf{Q}$ whose columns form a basis for the correlation vectors subspace, we can rewrite the covariance matrix as

$$vec(\mathbf{R} - \sigma_n^2\mathbf{I}) = \mathbf{Qa}. \tag{6}$$

Hence, we can choose the columns of $\mathbf{Q}$ as the eigenvectors corresponding to the non-zero eigenvalues (or the dominant
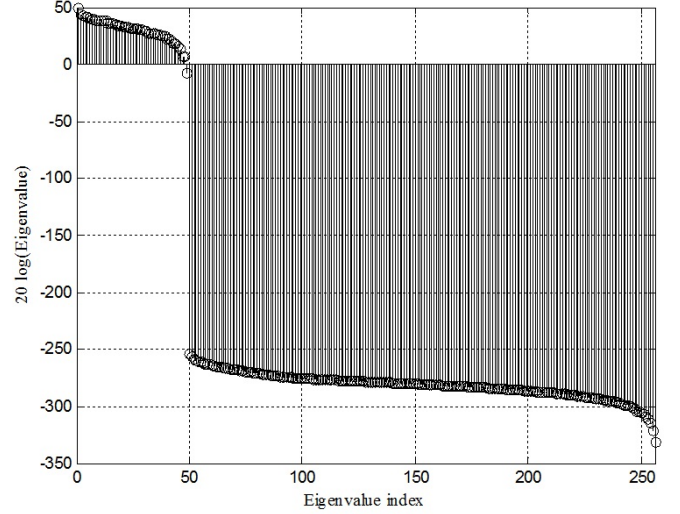

Fig. 1: Eigenvalues of the matrix S

eigenvectors). By imposing the linear structure constraint (6) to the covariance matrix estimation problem, we can significantly improve the estimation quality. Some works have studied covariance matrices with linear structures. For example, a weighted least-square estimator was proposed in [5] based on the linear structure for Toeplitz covariance matrices. However, the Toeplitz structure is restricted to linear arrays and the resulting matrix is not guaranteed to be positive definite.

A. Subspace Based Covariance Matrix Estimation

Based on (4) and (6), the estimated covariance matrix should lie in the subspace spanned by the columns of $\mathbf{Q}$. We are going to estimate $\mathbf{R}_s$ which is defined as

$$\mathbf{R}_s = \mathbf{R} - \sigma_n^2\mathbf{I} \tag{7}$$

Based on the previous discussion, we propose the following optimization problem

$$
\begin{aligned}
\min_{\mathbf{R}_s} \quad & \| (\hat{\mathbf{R}} - \sigma_n^2\mathbf{I}) - \mathbf{R}_s \|_2^2 \\
\text{subject to} \quad & (\mathbf{I} - \mathbf{Q}(\mathbf{Q}^H\mathbf{Q})^{-1}\mathbf{Q}^H)vec(\mathbf{R}_s) = 0 \\
& \mathbf{R}_s \succeq 0
\end{aligned}
\tag{8}
$$

where

$$\hat{\mathbf{R}} = \frac{1}{M} \sum_{m=1}^{M} \mathbf{x}(m)\mathbf{x}^H(m) \tag{9}$$

is the sample covariance matrix and M is the number of time samples. The matrix $(\mathbf{I} - \mathbf{Q}(\mathbf{Q}^H\mathbf{Q})^{-1}\mathbf{Q}^H)$ is the projection matrix on the subspace that is orthogonal to the subspace of the correlations vectors. As such, the first constraint in (8) ensures that the resulting matrix lies in the correlations vectors subspace. The second constraint guarantees that the resulting matrix is positive definite. Note that, (8) is a convex optimization problem and can be solved using standard tools from convex optimization. The proposed method imposes the linear structure using the subspace constraint. If the covariance

matrix is Toeplitz, the subspace constraint enforces the resulting matrix to be Toeplitz. However, the proposed algorithm is not limited to Toeplitz structures and can be used for any linear structure.

The sample covariance matrix (9) can be expressed as

$$\hat{\mathbf{R}} = \sum_{i=1}^{P} \sigma_i^2 \mathbf{v}(\theta_i)\mathbf{v}^H(\theta_i) + \boldsymbol{\Delta} + \sigma_n^2 \mathbf{I}. \tag{10}$$

The second term on the right hand side of (10), $\boldsymbol{\Delta}$, is the unwanted part (estimation error) which tends to zero if we have an infinite number of snapshots. The estimation error has some random behavior and can lie anywhere in the entire space. Since the first constraint in (8) enforces the estimated matrix to lie in the correlation vectors' subspace, it is expected to eliminate the component of estimation error which is not in this subspace. The dimension of the correlation vectors subspace is typically smaller than the entire space dimension N^2. For example, for a 30-element uniform linear array, the dimension of the correlation vectors subspace is equal to 59; while the entire space dimension is 900. Thus, it is conceivable that the proposed method could yield a much better estimation performance in comparison to the sample covariance matrix (9).

B. Near Optimal Closed-form Solution

The proposed optimization problem (8) is an N^2 dimensional optimization problem. Therefore, it may be hard to solve for large arrays. In this section, we derive a closed form near optimal solution which makes our method easy for practical implementation.

According to (4), the covariance matrix should be in the correlation vectors subspace. We define $\mathbf{R}_\perp$ and $\mathbf{R}_\parallel$ as follows

$$vec(\mathbf{R}_\parallel) = \mathbf{Q}(\mathbf{Q}^H\mathbf{Q})^{-1}\mathbf{Q}^H vec(\hat{\mathbf{R}} - \sigma_n^2 \mathbf{I}) \tag{11}$$

$$vec(\mathbf{R}_\perp) = (\mathbf{I} - \mathbf{Q}(\mathbf{Q}^H\mathbf{Q})^{-1}\mathbf{Q}^H)vec(\hat{\mathbf{R}} - \sigma_n^2 \mathbf{I}). \tag{12}$$

Thus, $\mathbf{R}_\perp$ is orthogonal to the correlation vectors subspace and $\mathbf{R}_\parallel$ contains the desired part. Therefore, we rewrite (8) as

$$\begin{aligned} \min_{\mathbf{R}_s} \quad & \| \mathbf{R}_s - \mathbf{R}_\parallel \|_2^2 \\ \text{subject to} \quad & (\mathbf{I} - \mathbf{Q}(\mathbf{Q}^H\mathbf{Q})^{-1}\mathbf{Q}^H)vec(\mathbf{R}_s) = 0 \\ & \mathbf{R}_s \succeq 0 \end{aligned} \tag{13}$$

In the proposed estimator (8), we placed the first constraint to suppress the estimation error which does not lie in the correlation vectors subspace. In (11), we project the sample covariance matrix to the correlation vectors subspace. Thus, we have eliminated the estimation error which does not lie in the correlation vectors subspace. Accordingly, we simplify (13) as follows

$$\begin{aligned} \min_{\mathbf{R}_s} \quad & \| \mathbf{R}_s - \mathbf{R}_\parallel \|_2^2 \\ \text{subject to} \quad & \mathbf{R}_s \succeq 0 \end{aligned} \tag{14}$$

which has a simple closed form solution

$$\hat{\mathbf{R}}_s = (\sum_{i=1}^{q} \lambda_i \boldsymbol{\beta}_i \boldsymbol{\beta}_i^H) \tag{15}$$

where q is the number of positive eigenvalues of $\mathbf{R}_\parallel$, $\{\lambda_i\}_{i=1}^q$ are the positive eigenvalues and $\{\boldsymbol{\beta}_i\}_{i=1}^q$ are their corresponding eigenvectors. Actually, we break the primary optimization problem (8) into two optimization problems. First, we find a matrix in the correlation vectors subspace which is most close to the sample covariance matrix and the resulting matrix is $\mathbf{R}_\parallel$. In the second step, we find the closest positive semi-definite matrix to $\mathbf{R}_\parallel$ and the resulting matrix is given in (15).

IV. SIMULATION RESULTS

In this section, we provide some simulation results to illustrate the performance of the proposed approach. The examples provided include DoA estimation and subspace estimation, which underscores the flexibility of the proposed covariance matrix estimation approach for a broad range of applications. All the curves are based on the average of 500 independent runs.

A. Simulation I (DOA estimation, probability of resolution)

Assume a uniform linear array with $N = 10$ omnidirectional sensors spaced half a wavelength apart. For this array the correlation vectors subspace is a 19 dimensional space since the covariance matrix is Toeplitz. The additive noise is modeled as a complex Gaussian zero-mean spatially and temporally white process with identical variances in each array sensor. In this experiment, we compare the performance of MUSIC when used with the sample covariance matrix and with the proposed covariance matrix estimation method. We also compare its performance with the sparse covariance matrix representation method [14, 12] and the SParse Iterative Covariance-based Estimation approach (SPICE) [15]. We consider two uncorrelated sources located at $45°$ and $50°$ ($90°$ is the direction orthogonal to the array line) and both sources are transmitted with the same power. Fig. 2 shows the probability of resolution (the probability that the algorithm can distinguish these two sources) versus the number of snapshots for one fixed sensor with $\mathrm{SNR} = 0$ dB. It is clear that using the proposed method leads to significant improvement in performance in comparison to using the sample covariance matrix. SPICE [15] is an iterative algorithm, which is based on the sparse representation of the array covariance matrix and requires one matrix inversion in each iteration. One can see that this algorithm fails when we use 20 iterations, however, performs well with 1000 iterations. Nevertheless, for practical purposes it is generally computationally prohibitive to perform 1000 matrix inversion operations. In addition, one can observe that the proposed near optimal solution to (8) yields a close performance to the optimal solution. Fig. 3 displays the probability of resolution against SNR for a fixed training size $M = 500$ snapshots. Roughly, the MUSIC algorithm based on the proposed method is 7 dB better than the MUSIC algorithm based on the sample covariance matrix. In summary, the proposed method yields notable and promising performance even with a small number of snapshots and at low SNR regimes. Furthermore, it is easily implementable using the proposed closed-form solution, which consists of a matrix multiplication and eigen-decomposition.
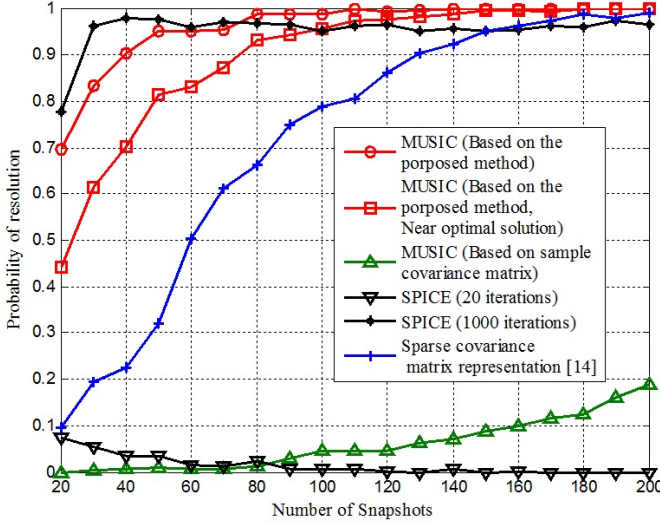
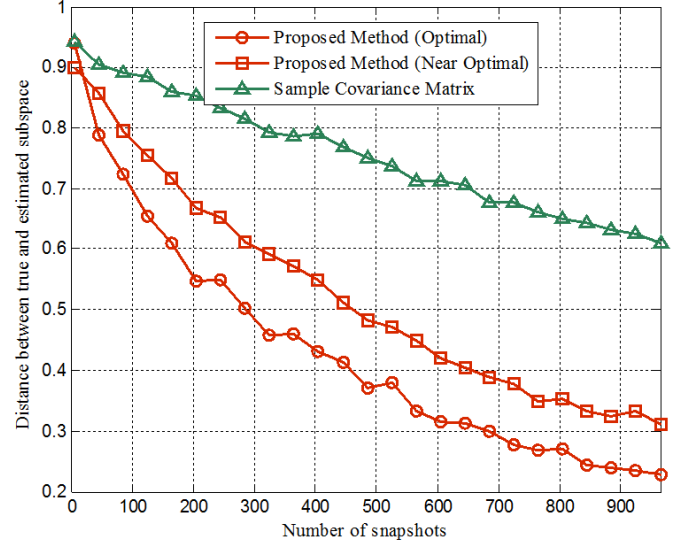Fig. 2: Probability of resolution as a function of the number of snapshots

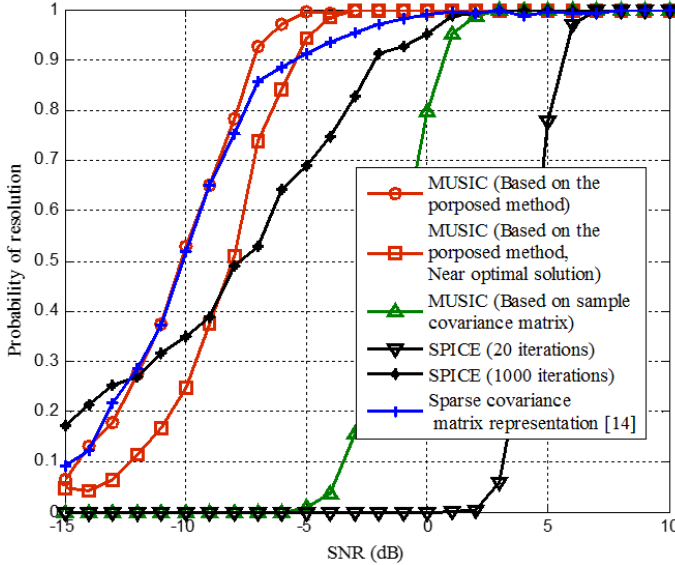Fig. 4: Distance between the true and the estimated signals subspace

Fig. 3: Probability of resolution versus SNR

B. Simulation II (Signals subspace estimation)

The estimation of the subspace of the received signals is an important task in many signal processing algorithms. For example, in the eigen-space based beamforming algorithm [3], the subspace of the received signals is used to make the beamformer robust against the steering vector mismatch. In the MUSIC algorithm, the subspace of the received signals is used to obtain the noise subspace [13]. The subspace of the received signals is usually estimated using the dominant eigenvectors of the estimated covariance matrix. In this simulation, we consider three uncorrelated sources located at $85°$, $90°$ and $95°$ and the sources are received with same signal to noise ratio. To investigate the accuracy of the subspace estimation, we define the distance between two subspaces as follows [16]: Given two matrices $\hat{\mathbf{U}}, \hat{\mathbf{V}} \in \mathbb{R}^{N \times K}$, the distance between the

subspaces spanned by the columns of $\hat{\mathbf{U}}$ and $\hat{\mathbf{V}}$ is defined as

$$dist(\hat{\mathbf{U}}, \hat{\mathbf{V}}) = \parallel \mathbf{U}_{\perp}^{H} \mathbf{V} \parallel_2 = \parallel \mathbf{V}_{\perp}^{H} \mathbf{U} \parallel_2 \qquad (16)$$

where $\mathbf{U}$ and $\mathbf{V}$ are orthonormal bases of the spaces $span(\hat{\mathbf{U}})$ and $span(\hat{\mathbf{V}})$, respectively. Similarly, $\mathbf{U}_{\perp}^{H} \in \mathbb{R}^{N \times (N-K)}$ is an orthonormal basis for the subspace which is orthogonal to $span(\hat{\mathbf{U}})$ and $\mathbf{V}_{\perp}^{H} \in \mathbb{R}^{N \times (N-K)}$ is an orthonormal basis for the subspace which is orthogonal to $span(\hat{\mathbf{V}})$. In addition, $\parallel \mathbf{X} \parallel_2$ denotes the spectral norm of matrix $\mathbf{X}$. Fig. 4 displays the distance between the true signals subspace and the estimated one as a function of the number of snapshots for $\text{SNR} = -6$ dB. We construct the signal subspace using the first three eigenvectors. One can observe that the proposed method exhibits a better rate of convergence. In addition, the performance of the closed-form solution closely approaches the optimal solution.

V. CONCLUSION

In this paper, a subspace method for array covariance matrix estimation was proposed. We have shown that when the received signals are uncorrelated, the covariance matrix lies in the subspace of the correlation vectors. Based on this idea, we posed the estimation problem as a convex optimization problem and enforced a subspace constraint. In addition, a near optimal closed-form solution for the proposed optimization problem was derived. A number of numerical examples demonstrated the notable performance of the proposed approach and its applicability to a wide range of signal processing problems, including but not limited to, DoA estimation and subspace estimation. In contrast to some of the existing approaches, which suffer from drastic performance degradation with limited data and at low SNR regimes, the proposed method showed very graceful degradation in such settings.

REFERENCES

[1] P. Stoica and R. Moses, *Spectral Analysis of Signals*, Prentice Hall, Upper Saddle RiverNJ, 2005.

[2] R. O. Shemidt, *Multiple emitter location and signal parameter estimation*, IEEE Trans. Antennas Propagat., vol. AP-34, pp. 276280, Mar. 1985.

[3] S. A. Vorobyov, *Principles of minimum variance robust adaptive beamforming design*, Elsevier Signal Processing, invited paper, Special Issue: Advances in Sensor Array Processing, vol. 93, no. 12, pp. 3264-3277, Dec. 2013.

[4] K. Yu, M. Bengtsson, B. Ottersten, D. McNamara, and P. Karlsson, *Modeling of wide-band MIMO radio channels based on NLoS indoor measurements*, IEEE Transactions on Vehicular Technology, vol. 53, no. 8, pp. 655665, May 2004.

[5] H. Li, P. Stoica, J. Li, *Computationally Efficient Maximum Likelihood Estimation of Structured Covariance Matrices*, IEEE Transactions on Signal Processing, Vol. 47, pp. 1314 1323, May 1999.

[6] M. Rahmani, M. H. Bastani, *Robust and rapid converging adaptive beamforming via a subspace method for the signal-plus-interferences covariance matrix estimation*, IET Signal Processing, Vol. 8 , Issue 5, pp. 507 520, July 2014.

[7] J. C. de Munck, H. M. Huizenga, L. J. Waldorp, and R. M. Heethaar, *Estimating stationary dipoles from MEG/EEG data contaminated with spatially and temporally correlated background noise*, IEEE Transactions on Signal Processing, vol. 50, no. 7, pp. 15651572, Jul. 2002.

[8] K. Werner, M. Jansson, P. Stoica, *On Estimation of Covariance Matrices with Kronecker Product Structure*, IEEE Transactions on Signal Processing, Vol. 56 , no. 2, pp. 478 491, Feb. 2008.

[9] B. Ottersten, P. Stoica, and R. Roy, *Covariance matching estimation techniques for array signal processing applications*, Digital Signal Processing, vol. 8, pp. 185210, 1998.

[10] B. Gransson, M. Jansson and B. Ottersten, *Spatial and Temporal Frequency Estimation of Uncorrelated Signals Using Subspace Fitting*, in Proc. of 8th IEEE Signal Processing Workshop on Statistical Signal and Array Processing, Corfu, Greece, Jun. 1996, pp. 94-96.

[11] D. Malioutov, M. etin, A. Willsky, *A sparse signal reconstruction perspective for source localization with sensor arrays*, IEEE Trans Signal Process., Vol. 53, pp. 30103022, 2005.

[12] L. Blanco1, M. Njar, *Sparse covariance fitting for direction of arrival estimation*, EURASIP Journal on Advances in Signal Processing, 2012

[13] H.L VanTrees, *Optimum array processing .Part IV of detection, estimation, and modulation theory*, Wiley,2002

[14] J. S. Picard, A. J. Weiss, *Direction finding of multiple emitters by spatial sparsity and linear programming*, 9th International Symposium on Communications and Information Technology, pp. 1258-1262, 2009.

[15] P. Stoica, P. Babu, J. Li, *SPICE: A Sparse Covariance-Based Estimation Method for Array Processing*, IEEE Trans Signal Process., Vol. 59, pp. 629 - 638, 2011.

[16] P. Jain, P. Netrapalli, and S. Sanghavi, *Low-rank matrix completion using alternating minimization*, arXiv preprint arXiv:1212.0467, 2012.