SUBSPACE ESTIMATION USING FACTOR ANALYSIS

Ahmad Mouri Sardarabadi, and Alle-Jan van der Veen

Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology (TU Delft)
e-mail:{A.Mourisardarabadi, A.j.veen}@tudelft.nl

ABSTRACT

Many subspace estimation techniques assume either that the system has a calibrated array or that the noise covariance matrix is known. If the noise covariance matrix is unknown, training or other calibration techniques are used to find it. In this paper another approach to the problem of unknown noise covariance is presented. The complex factor analysis (FA) and a new extended version of this model are used to model the covariance matrix. The steep algorithm for finding the MLE of the model parameters is presented. The Fisher information and an expression for the Cramér–Rao bound are derived. The practical use of the model is illustrated using simulated and experimental data.

***Index Terms*—** Factor analysis, complex factor analysis, subspace estimation, Cramér-Rao bound, maximum-likelihood.

1. INTRODUCTION

Many array processing techniques rely on estimating the signal subspace using eigenvalue decomposition (EVD) of the covariance matrix. This method does not have an inherit model for the noise and its usage for subspace estimation is limited to the case where the noise covariance matrix could approximately be modeled as $\sigma^2\boldsymbol{I}$. In this paper we are interested in estimating the signal subspace even if the noise covariance does not have this special form and is unknown. Also each receiving element of the array could have a different unknown gain. These complications requires a data–model that includes unknown noise powers and is scale–invariant. The FA model has both of these required properties and is used to find the desired subspace.

After its first formulation by Spearman in 1904, the FA model has been used in various fields such as psychology, social sciences, natural science, etc [1, 2, 3]. Variations of the FA model have also been explored for blind source separation and array calibration [4, 5].

Even though the work presented here could be used in various fields, the focus is on its usage for radio–astronomy. Spatial filtering of the strong sources or interference and removing the extended emissions are two possible applications of the FA model in this field. This paper gives some new results needed for extending the FA model to complex case and also extends the model to account for a more general noise models. The MLE of the model parameters is estimated using the steep method. The Fisher information and the Fisher score are presented here could also be used for the scoring algorithm as presented in [6].

2. DATA MODEL

For a system with p receiving elements that is exposed to m sources, a commonly used narrow-band model has the form

$$\boldsymbol{x}(t) = \boldsymbol{A}_0\boldsymbol{s}_0(t) + \boldsymbol{n}(t). \tag{1}$$

Where $\boldsymbol{x}$ is a $p \times 1$ vector of received signals, $\boldsymbol{A}_0$ is the $p \times m$ array-response matrix, $\boldsymbol{s}_0$ is an $m \times 1$ vector of source signals and $\boldsymbol{n}$ is a $p \times 1$ vector representing all the noise contributions in the system.

This data model suffers from some ambiguities that need to be addressed before attempting to estimate the model parameters. Given any invertable matrix $\boldsymbol{Z}$ and any unitary matrix $\boldsymbol{Q}$, the model could be rewritten as

$$\boldsymbol{x}(t) = \boldsymbol{A}_0\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Q}^H\boldsymbol{Z}^{-1}\boldsymbol{s}_0(t) + \boldsymbol{n}(t). \tag{2}$$

It is assumed that the sources and noise contributions are uncorrelated and have proper complex Gaussian distributions $\mathcal{CN}(0, \boldsymbol{R}_{\boldsymbol{s}_0})$ and $\mathcal{CN}(0, \boldsymbol{R}_{\boldsymbol{n}})$ respectively. In this situation the covariance matrix of received signal is given by

$$\boldsymbol{R}_{\boldsymbol{x}} = \boldsymbol{A}_0\boldsymbol{Z}\boldsymbol{Q}\boldsymbol{Q}^H\boldsymbol{Z}^{-1}\boldsymbol{R}_{\boldsymbol{s}_0}\boldsymbol{Z}^{-H}\boldsymbol{Q}\boldsymbol{Q}^H\boldsymbol{Z}^H\boldsymbol{A}_0^H + \boldsymbol{R}_{\boldsymbol{n}}. \tag{3}$$

The matrix $\boldsymbol{Z}$ could always be chosen in such a way that the $\boldsymbol{Z}^{-1}\boldsymbol{R}_{\boldsymbol{s}_0}\boldsymbol{Z}^{-H} = \boldsymbol{I}_m$. By introducing $\boldsymbol{A} = \boldsymbol{A}_0\boldsymbol{Z}\boldsymbol{Q}$ the covariance matrix of the received signal becomes

$$\boldsymbol{R}_{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{A}^H + \boldsymbol{R}_{\boldsymbol{n}}. \tag{4}$$

The only ambiguity left is the choice of the unitary matrix $\boldsymbol{Q}$. In order to choose $\boldsymbol{Q}$, first the case of known noise covariance is considered. Let $\boldsymbol{R}_0 = \boldsymbol{A}_0\boldsymbol{R}_{\boldsymbol{s}_0}\boldsymbol{A}_0^H = \boldsymbol{A}\boldsymbol{A}^H$ and assume that $\boldsymbol{R}_{\boldsymbol{n}}$ is known, then the EVD of the whitened covariance matrix $\boldsymbol{R}_{\boldsymbol{n}}^{-\frac{1}{2}}\boldsymbol{R}_{\boldsymbol{x}}\boldsymbol{R}_{\boldsymbol{n}}^{-\frac{1}{2}} - \boldsymbol{I}_p = \boldsymbol{R}_{\boldsymbol{n}}^{-\frac{1}{2}}\boldsymbol{R}_0\boldsymbol{R}_{\boldsymbol{n}}^{-\frac{1}{2}}$ could be used to estimate $\boldsymbol{A}$. Consider the following eigenvalue problem

$$\boldsymbol{R}_{\boldsymbol{n}}^{-\frac{1}{2}}\boldsymbol{R}_0\boldsymbol{R}_{\boldsymbol{n}}^{-\frac{1}{2}}\boldsymbol{U} = \boldsymbol{U}\boldsymbol{\Gamma}, \tag{5}$$

where $\boldsymbol{\Gamma}$ is a diagonal matrix. If we require $\boldsymbol{U} = \boldsymbol{R}_{\boldsymbol{n}}^{-\frac{1}{2}}\boldsymbol{A}$ then (5) becomes

$$\begin{aligned}\boldsymbol{R}_{\boldsymbol{n}}^{-\frac{1}{2}}\boldsymbol{R}_0\boldsymbol{R}_{\boldsymbol{n}}^{-\frac{1}{2}}\boldsymbol{R}_{\boldsymbol{n}}^{-\frac{1}{2}}\boldsymbol{A} &= \boldsymbol{R}_{\boldsymbol{n}}^{-\frac{1}{2}}\boldsymbol{A}(\boldsymbol{A}^H\boldsymbol{R}_{\boldsymbol{n}}^{-1}\boldsymbol{A}) \\ &= \boldsymbol{R}_{\boldsymbol{n}}^{-\frac{1}{2}}\boldsymbol{A}\boldsymbol{\Gamma}.\end{aligned} \tag{6}$$

One way to make sure that (6) holds, is by setting

$$\boldsymbol{A}^H\boldsymbol{R}_{\boldsymbol{n}}^{-1}\boldsymbol{A} = \boldsymbol{\Gamma}. \tag{7}$$

The matrix $\boldsymbol{Q}$ is now chosen in such a way that (7) holds. With the choice of $\boldsymbol{Q}$ there are no ambiguities left.

Finally we assume that the noise contributions are uncorrelated so that

$$\boldsymbol{R}_{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{A}^H + \boldsymbol{D}. \tag{8}$$

where $\boldsymbol{D}$ is a diagonal matrix.

For the remainder of this paper we refer to (8) as the (classical) FA model.

978-1-4673-1071-0/12/$31.00 ©2012 IEEE

2.1. Extension to Non-diagonal Covariance Matrices

The classical FA model is an especial case of a more general model where the noise covariance could be modeled as $\boldsymbol{R_n} = \boldsymbol{M} \odot \boldsymbol{R_n}$ where $\boldsymbol{M}$ is a symmetric matrix containing only zeros and ones. The FA model then becomes

$$\boldsymbol{R_x} = \boldsymbol{A}\boldsymbol{A}^H + \boldsymbol{M} \odot \boldsymbol{R_n}. \tag{9}$$

We assume that $\boldsymbol{M} = \boldsymbol{I}_p + \boldsymbol{B}$ where $\boldsymbol{B}$ is a symmetric matrix with zeros on its diagonal. If $\boldsymbol{B} = \boldsymbol{0}$ then the classical model is obtained. The choice of this mask depends on the application.

In the follwoing sections we first extend the estimation of the classical FA to complex case and give an algorithm to estimate the model parameters as given by (9).

3. PARAMETER ESTIMATION TECHNIQUES FOR COMPLEX-VALUED DATA

The FA model for real-valued data, is a mature subject and various techniques for the estimation of the model parameters has been suggested in literature[7, 8, 9]. Some of these algorithms extend to complex case very naturally and some need closer look.

Before discussing the various algorithms, the problem definition for FA is examined in more detail. Given N samples from the received signal $\boldsymbol{x}$, we want to find $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{D}}$ based on the sample covariance matrix

$$\hat{\boldsymbol{R}}_{\boldsymbol{x}} = \frac{1}{N} \sum_{i=0}^{N-1} \boldsymbol{x}[i]\boldsymbol{x}[i]^H. \tag{10}$$

This problem becomes more complicated if the the number of sources m is not known in advance. In this situation an estimate, $\hat{m}$, must also be chosen or found. It is also of interest to know the maximum number of sources for which the FA model has any solution at all. In order to find this, we see how many free parameters this model leave us for a given number of p and m. For a complex sample covariance matrix there are p^2 number of parameters. The FA model parameters, $\boldsymbol{A}$ and $\boldsymbol{D}$ have $2pm$ and p free parameters respectively, from which m^2 are fixed by (7). This gives the total of free parameters

$$\begin{aligned} s &= p^2 - 2pm + m^2 - p \\ &= (p-m)^2 - p. \end{aligned} \tag{11}$$

Even though the number of free parameters, s, is different for real and complex case, the interpretation is the same. In [10] the real case is discussed. In both cases if $s < 0$ there is an infinite number of exact solutions and the problem is ill-posed. If $s = 0$ there is one unique solution; this is for example in the case of EVD where $m = p$ and $\boldsymbol{D}$ is known or neglected (s becomes $(p-m)^2 = 0$). The last case is when $s > 0$, in this case it is not necessary that the solution is exact and the parameters are estimated in such a way to optimize a cost function.

In this paper we are interested in the case where $s > 0$. This the case where $m < \lfloor p - \sqrt{p} \rfloor$ and $\boldsymbol{D}$ is unknown. We also assume that the problem up to the matrix $\boldsymbol{Q}$ is (locally) identifiable. Conditions for this type of identifiability are studied by [11, 12]. In the following sections the maximum likelihood (ML) estimation of FA model is discussed.

3.1. Maximum Likelihood Estimator

The aim is to find $\boldsymbol{A}$ and $\boldsymbol{D}$ that maximize the complex log-likelihood function

$$l(\boldsymbol{x}; \boldsymbol{A}, \boldsymbol{D}) = N\left[-\log|\pi^p| + \log|\boldsymbol{R}^{-1}| - \mathrm{tr}(\boldsymbol{R}^{-1}\hat{\boldsymbol{R}})\right]. \tag{12}$$

To achieve this we find the Fisher score and set it equal to zero. The Fisher score for a proper Gaussian distributed signal is given by [13, p.165]

$$\begin{aligned} t_{\theta_j} &= \frac{\partial}{\partial \theta_j^*} l(\boldsymbol{x}; \boldsymbol{\theta}) \\ &= -N\mathrm{tr}\left[\boldsymbol{R}^{-1}\left(\frac{\partial \boldsymbol{R}}{\partial \theta_j}\right)^H\right] + N\mathrm{tr}\left[\boldsymbol{R}^{-1}\left(\frac{\partial \boldsymbol{R}}{\partial \theta_j}\right)^H \boldsymbol{R}^{-1}\hat{\boldsymbol{R}}\right], \end{aligned} \tag{13}$$

where the partial derivatives are Wirtinger derivatives; θ_j is the j–th component of the vector $\boldsymbol{\theta} = [\boldsymbol{a}_1^T, \ldots, \boldsymbol{a}_m^T, \boldsymbol{d}]^T$ where $\boldsymbol{a}_k$ is the k–th column of $\boldsymbol{A}$ and $\boldsymbol{d} = \mathrm{vectdiag}(\boldsymbol{D})$ is a vector containing diagonal elements of $\boldsymbol{D}$; θ^* is the complex conjugate of θ. We have either $\theta_j = a_{ik}$ or $\theta_j = d_i$ where $i = 1, .., p$ and $k = 1, .., m$, so we need to find the partial Wirtinger derivatives $\frac{\partial \boldsymbol{R}}{\partial a_{ik}}$ and $\frac{\partial \boldsymbol{R}}{\partial d_i}$.

Here we give the final results of the derivations

$$\frac{\partial \boldsymbol{R}}{\partial a_{ik}} = \boldsymbol{e}_i \boldsymbol{a}_k^H. \tag{14}$$

and

$$\frac{\partial \boldsymbol{R}}{\partial d_i} = \boldsymbol{e}_i \boldsymbol{e}_i^H. \tag{15}$$

where $\boldsymbol{e}_i$ is a unit vector with entry i equal to 1. The final results for Fisher score in the matrix form become

$$\boldsymbol{T_A} = N\left(-\boldsymbol{R}^{-1}\boldsymbol{A} + \boldsymbol{R}^{-1}\hat{\boldsymbol{R}}\boldsymbol{R}^{-1}\boldsymbol{A}\right). \tag{16}$$

and

$$\boldsymbol{T_D} = N\mathrm{diag}(-\boldsymbol{R}^{-1} + \boldsymbol{R}^{-1}\hat{\boldsymbol{R}}\boldsymbol{R}^{-1}). \tag{17}$$

Thus the Fisher score for the FA model is then

$$\boldsymbol{t_\theta} = [\mathrm{vect}(\boldsymbol{T_A})^T, \mathrm{vectdiag}(\boldsymbol{T_D})^T]^T. \tag{18}$$

In [14] the same expression for the real-valued data is derived.

Even though no closed form solution for $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{D}}$ could be found using (16), (17) or (18), they could be used to derive some interesting properties of the MLE for FA. For a discussion and derivation of these properties the reader is referred to [14, 10].

In the following section we will derive the necessary iterations to approximate the MLE numerically.

3.1.1. Steep

For a function $f(\boldsymbol{Z}, \boldsymbol{Z}^*)$ the direction of the maximum change with respect to $\boldsymbol{Z}$ is give by

$$\nabla_{\boldsymbol{Z}} f(\boldsymbol{Z}, \boldsymbol{Z}^*) = \frac{\partial}{\partial \boldsymbol{Z}^*} f(\boldsymbol{Z}, \boldsymbol{Z}^*) \tag{19}$$

where the notation is adapted from [15].

Using (19), (13), (16) and (17) we have

$$\nabla_{\boldsymbol{A}} l(\boldsymbol{x}; \boldsymbol{A}, \boldsymbol{D}) = \boldsymbol{T_A} \tag{20}$$

and

$$\nabla_{\boldsymbol{D}} l(\boldsymbol{x}; \boldsymbol{A}, \boldsymbol{D}) = \boldsymbol{T_D}. \tag{21}$$

The iteration steps for the steep become

$$\hat{\boldsymbol{A}}_{k+1} = \hat{\boldsymbol{A}}_k + \mu_k \boldsymbol{\Phi}_k \hat{\boldsymbol{A}}_k \tag{22}$$

and

$$\hat{\boldsymbol{D}}_{k+1} = \hat{\boldsymbol{D}}_k + \mu_k(\boldsymbol{I}_p \odot \boldsymbol{\Phi}_k) \tag{23}$$

where

$$\boldsymbol{\Phi}_k = -\boldsymbol{R}_k^{-1} + \boldsymbol{R}_k^{-1}\hat{\boldsymbol{R}}\boldsymbol{R}_k^{-1} \tag{24}$$

and $\boldsymbol{R}_k = \hat{\boldsymbol{A}}_k\hat{\boldsymbol{A}}_k^H + \hat{\boldsymbol{D}}_k$.

The matrix $\boldsymbol{R}_k$ needs to be inverted at each iteration. This could be done efficiently using Woodbury matrix identity

$$\boldsymbol{R}_k^{-1} = \hat{\boldsymbol{D}}_k^{-1} - \hat{\boldsymbol{D}}_k^{-1}\hat{\boldsymbol{A}}_k(\boldsymbol{I}_m + \hat{\boldsymbol{A}}_k^H\hat{\boldsymbol{D}}_k^{-1}\hat{\boldsymbol{A}}_k)^{-1}\hat{\boldsymbol{A}}_k\hat{\boldsymbol{D}}_k^{-1} \tag{25}$$

which requires the inversion of one diagonal matrix and one $m \times m$ matrix. If the matrix $\boldsymbol{Q}$ is also calculated at each iteration then only the inversion of two diagonal matrices is needed.

3.1.2. Steep for Extended FA

For the case of extended FA we need to compute the number of free parameters again. In a system with p receiving elements and m sources satisfying (7) we have

$$s = (p-m)^2 - p - 2k, \tag{26}$$

where k is the number of off-diagonal elements. It is worth noting that the covariance matrix is Hermitian and we only need to estimate the elements below (or above) the diagonal. The factor $2k$ should not be confused with the number of off-diagonal elements of $\boldsymbol{M}$, which is also $2k$, the factor 2 comes from having complex entries on the off-diagonal elements of noise covariance matrix. Again we consider the case where $s > 0$. This limits the number of off-diagonal elements that could be estimated for a given number of sources and vice versa. It is then straightforward to show that we only need to change (23) to adapt the steep algorithm for extended FA.

Following the same procedure used to derive steep in previous section we have

$$\frac{\partial \boldsymbol{R}}{\partial (r_n)_{ij}} = m_{ji}\boldsymbol{e}_j\boldsymbol{e}_i^H. \tag{27}$$

We have $m_{ji} = m_{ij}$ and so we have in the matrix form

$$\nabla_{\boldsymbol{R}_n,\boldsymbol{R}^*{}_n} l(\boldsymbol{x}, \boldsymbol{A}, \boldsymbol{R}_n) = N\boldsymbol{M} \odot (-\boldsymbol{R}^{-1} + \boldsymbol{R}^{-1}\hat{\boldsymbol{R}}\boldsymbol{R}^{-1}). \tag{28}$$

The iteration update for the extended FA model could now be written as

$$\hat{\boldsymbol{R}}_{\boldsymbol{n}\,k+1} = \hat{\boldsymbol{R}}_{\boldsymbol{n}\,k} + \mu_k(\boldsymbol{M} \odot \boldsymbol{\Phi}_k). \tag{29}$$

where $\boldsymbol{\Phi}_k$ is given by (24) and $\boldsymbol{R}_k = \hat{\boldsymbol{A}}_k\hat{\boldsymbol{A}}_k^H + \hat{\boldsymbol{R}}_{\boldsymbol{n}\,k}$.

This concludes the estimation techniques for both classical and extended FA. In the following section the CRB for the classical case is presented.

4. CRAMÉR–RAO BOUND

The Cramér–Rao Bound is the lower bound on the covariance matrix of the estimated parameters. As explained in order to have a unique solution for the FA model the matrix $\boldsymbol{Q}$ needs to be fixed. Choosing this matrix puts constraints on the parameter $\boldsymbol{A}$. As a result the constrained CRB for complex parameters is used here. Following [16] we define

$$\underline{\boldsymbol{\theta}} = \begin{pmatrix} \text{vect}(\boldsymbol{A}) \\ \text{vect}(\boldsymbol{A}^*) \\ \text{vectdiag}(\boldsymbol{D}) \end{pmatrix} \tag{30}$$

as augmented $\boldsymbol{\theta}$.

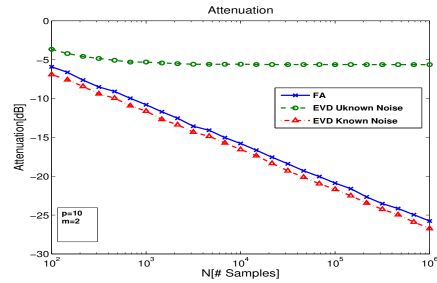

Fig. 1. Spatial Filter attenuation

The augmented CRB for the estimated complex parameters with the constraint $\underline{\boldsymbol{k}}(\underline{\boldsymbol{\theta}}) = \boldsymbol{0}$ is

$$\underline{\boldsymbol{C}}(\underline{\boldsymbol{\theta}}) \geq \boldsymbol{U}\left(\boldsymbol{U}^H\underline{\boldsymbol{F}}\boldsymbol{U}\right)^{-1}\boldsymbol{U}^H, \tag{31}$$

where $\underline{\boldsymbol{F}}$ is the augmented unconstrained complex Fisher information defined in [17] and $\boldsymbol{U}$ is a unitary basis for the null–space of $\boldsymbol{K}$ given by

$$\boldsymbol{K} = \frac{\partial \underline{\boldsymbol{k}}^T}{\partial \underline{\boldsymbol{\theta}}}. \tag{32}$$

Vectorization of (7) gives m^2 equations needed to construct $\underline{\boldsymbol{k}}$.

For the proper Gaussian distribution we can use Bang's formula to find the Fisher information

$$\underline{\boldsymbol{F}} = \boldsymbol{J}^H(\boldsymbol{R}^{-T} \otimes \boldsymbol{R}^{-1})\boldsymbol{J} \tag{33}$$

where

$$\begin{aligned} \boldsymbol{J} &= \frac{\partial \text{vect}(\boldsymbol{R})}{\partial \underline{\boldsymbol{\theta}}^T} \\ &= \left[\frac{\partial \text{vect}(\boldsymbol{R})}{\partial \text{vect}^T(\boldsymbol{A})}, \frac{\partial \text{vect}(\boldsymbol{R})}{\partial \text{vect}^T(\boldsymbol{A}^*)}, \frac{\partial \text{vect}(\boldsymbol{R})}{\partial \text{vectdiag}^T(\boldsymbol{D})}\right] \\ &= [(\boldsymbol{A}^* \otimes \boldsymbol{I}), (\boldsymbol{I} \otimes \boldsymbol{A})\boldsymbol{K}, (\boldsymbol{I} \circ \boldsymbol{I})], \end{aligned} \tag{34}$$

the matrix $\boldsymbol{K}$ is a permutation matrix such that $\boldsymbol{K}\text{vect}(\boldsymbol{A}) = \text{vect}(\boldsymbol{A}^T)$, $\otimes$, $\circ$ and $\odot$ represent Kronecker, Khatri–Rao and Hadamard product respectively. Using the tools developed here it could be shown that the performance of the steep and scoring method reaches CRB for the large number of samples.

5. SIMULATION RESULTS

We simulate a scenario in which the estimated subspace is used to filter the incoming signals with the help of orthogonal projections. The noise matrix is generated randomly in such a way that on each receiving element the SNR is between 0 to -10 dB. We assume to know the number of sources in advance. For this case there are two sources, $m = 2$ and there are ten receiving elements, $p = 10$. We use EVD without knowing the noise matrix, the FA (also without knowing the noise) and EVD with the full knowledge of the noise covariance matrix (whitening is used) and estimate the subspace for different number of samples. The attenuation for each estimation is found using

$$E = \frac{\left\|\hat{\boldsymbol{P}}\boldsymbol{A}\boldsymbol{A}^H\hat{\boldsymbol{P}}\right\|_F}{\left\|\boldsymbol{A}\boldsymbol{A}^H\right\|_F} \tag{35}$$

where $\hat{\boldsymbol{P}} = \hat{\boldsymbol{A}}\left(\hat{\boldsymbol{A}}^H\hat{\boldsymbol{A}}\right)^{-1}\hat{\boldsymbol{A}}^H$.

Fig.1 is the average of 30 MC runs for the estimated attenuations. As expected FA estimates the subspace without knowing the noise covariance very close to EVD method with the full knowledge of the noise. The FA has slightly less attenuation than whitened EVD but that could be expected given the fact that more parameters are estimated.

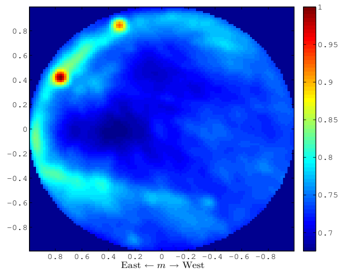

Fig. 2. DFT Imaging withought FA

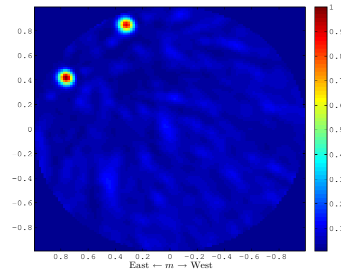

Fig. 3. DFT Imaging with Extended FA

6. EXPERIMENTAL RESULTS

The results presented here illustrate one of the applications of the extended FA model in radio astronomy. We follow discussion in [18, pp.35-40] very closely. In fact the data set used here is the same data set used in [18] from the LOFAR project. The total sky image is made using classical DFT method by combining data from 24 156kHz sub–bands distributed between 45.3 and 67.3 MHz and 10 seconds of integration per channel. **Fig.2** shows the image without being pre–processed with the extended FA. There are two strong sources, Cas A and Cyg A, and there is also a cloud–like radiation from the extended emission in the sky. The extended emission affects the short–baselines that are smaller than 4 wavelengths. This is exploited for modeling this extended emission as noise [18]. The four wavelength criteria was used to create a mask matrix $\boldsymbol{M}_{4\lambda}$ and the subspace of the two strong sources was estimated using extended FA model. **Fig.3** shows the image made using the DFT imaging on $\hat{\boldsymbol{R}}_0 = \hat{\boldsymbol{A}}\hat{\boldsymbol{A}}^{\boldsymbol{H}}$. The estimated subspace gives an accurate estimation for the sources.

7. CONCLUSIONS

It has been shown that the complex factor analysis model is the same model which is commonly used in literature for narrow-band array processing. The CRB and the Fisher score for the case of proper complex Gaussian distribution are found using Wirtinger derivatives. Some of the techniques used in factor analysis for real numbers are extended to the case of the complex numbers. With the help of simulations and experimental data the potential use of FA as a generic signal processing tool has been demonstrated.

8. REFERENCES

[1] C. Spearman, "The proof and measurement of association between two things," *The American Journal of Psychology*, vol. 15, pp. 72–101, Jan 1904.

[2] D. J. Bartholomew, M. Knott, and I. Moustaki, *Latent Variable Models and Factor Analysis: A Unified Approach*. John Wiley and Sons, 2011.

[3] R. A. Reyment and K. G. Jöreskog, *Applied Factor Analysis in the Natural Sciences*. Cambridge University Press, 1996.

[4] S. Schell and W. Gardner, "Maximum likelihood and common factor analysis-based blind adaptive spatial filtering for cyclostationary signals," in *Acoustics, Speech, and Signal Processing, 1993. ICASSP-93., 1993 IEEE International Conference on*, vol. 4, pp. 292 –295 vol.4, april 1993.

[5] A.-J. Boonstra and A.-J. van der Veen, "Gain calibration methods for radio telescope arrays," *IEEE Tr. Signal Processing*, vol. 51, pp. 25–38, 2003.

[6] S. M. Kay, *Fundamentals of Statistical Signal Processing, Estimation theory*, vol. Volume I. Prentice Hall, 1993.

[7] K. G. Jöreskog, "A general approach to confirmatory maximum likelihood factor analysis," *PSYCHOMETRIKA*, vol. 34, no. 2, 1969.

[8] S. Y. Lee, "The gauss-newton algorithm for the weighted least squares factor analysis," *Journal of the Royal Statistical Society*, vol. 27, June 1978.

[9] A.-K. Seghouane, "An iterative projections algorithm for ml factor analysis," in *IEEE Workshop on Machine Learning for Signal Processing*, pp. 333 –338, oct. 2008.

[10] K. Mardia, J. Kent, and J. Bibby, *Multivariate Analysis*. ACADEMIC PRESS, 1997.

[11] T. W. Anderson and H. Rubin, "Statistical inference in factor analysis," *In Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, vol. 5, pp. 111 – 150, 1956.

[12] A. Shapiro, "Identifiability of factor analysis: some results and open problems," *Linear Algebra and its Applications*, vol. 70, no. 0, pp. 1 – 7, 1985.

[13] P. J. Schreier, *Statistical Signal Processing of Complex-Valued Data*. Cambridge University Press, 2010.

[14] T. Anderson, *An Introduction to Multivariate Statistical Analysis*. Wiley, third edition ed., 2003.

[15] A. Hjørungnes, *Complex-Valued Matrix Derivatives with Applications in Signal Processing and Communications*. Cambridge University Press, 2011.

[16] A. K. Jagannatham and B. D. Rao, "Cramer-rao lower bound for constrained complex parameters," *IEEE SIGNAL PROCESSING LETTERS*, vol. 11, NOVEMBER 2004.

[17] E. Ollila, V. Koivunen, and J. Eriksson, "On the Cramér-rao bound for the constrained and unconstrained complex parameters," in *Sensor Array and Multichannel Signal Processing Workshop, 2008. SAM 2008. 5th IEEE*, pp. 414 –418, july 2008.

[18] S. J. Wijnsholds, *Fish-Eye Observing with Phased Array Radio Telescope*. PhD thesis, Delft University of Technology, 2010.