Publication V

Esa Ollila and Visa Koivunen. 2003. Robust antenna array processing using *M*-estimators of pseudo-covariance. In: Gong Ke and Zhisheng Niu (editors). Proceedings of the 14th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2003). Beijing, China. 7-10 September 2003, volume 3, pages 2659-2663.

Robust Antenna Array Processing Using M-estimators of Pseudo-Covariance

Esa Ollila and Visa Koivunen
Signal Processing Laboratory,
Helsinki University of Technology,
P.O.Box 3000, FIN-02015 HUT, Finland

***Abstract*— This paper addresses the problem of antenna array processing in non-Gaussian noise and interference conditions. Such conditions arise due to man-made interference in indoor and outdoor mobile communication channels as well as in military communications. In this paper M-estimators of the array (pseudo-)covariance matrix based upon complex data set are introduced. Estimates of the noise and signal subspaces based on M-estimators are then used to robustify the subspace direction of arrival (DOA) estimation methods. In addition, eigenvalues based on M-estimators are used in MDL criterion, thus yielding a robust signal detection method. The reliable performance of the proposed methods are shown by simulations.**

I. Introduction

Multi-antenna transceivers are the key element in beyond 3G communications with high spectral efficiency (MIMO systems), and smart antennas capable of interference cancellation. A starting point for many multi-antenna receiver algorithms is the array output covariance matrix. In particular, noise and signal subspaces spanned by the eigenvectors of the array covariance matrix need to be estimated in many high resolution DOA, channel and symbol sequence estimation methods [1]. Conventional sample covariance matrix computed from spatial snapshots is optimal under Gaussian distribution assumption but it gives poor performance in non-Gaussian and impulsive noise environments. In the latter case, it is well known that if the covariance matrix is estimated in a non-robust manner, statistics (such as eigenvalues and eigenvectors) based on it are unreliable and far from optimal. Obtaining robust estimates of the (pseudo-)covariance matrix is crucial since noise and interference in indoor and outdoor mobile communication channels as well as in sonar and radar signals have been shown to be contaminated by non-Gaussian noise and interference. See e.g. Middeleton [2] and Williams and Johnson [3] and references therein. Several authors have proposed robust DOA estimation methods for non-Gaussian environments, see e.g. [4], [3], [5].

For data in $\mathbb{R}^k$ robust alternatives for the sample covariance estimate are the M-estimators of multivariate scatter matrix introduced in [6] and later studied by several authors, see e.g. [7], [8], [9], [10]. In this paper, we extend the concept of M-estimators of scatter matrix (pseudo-covariance matrix) for data in complex vector space $\mathbb{C}^k$. Maximum Likelihood (ML-)estimators of the scatter for the complex multivariate t-distribution with $\nu > 0$ degrees of freedom are introduced in Section II-A. This model describes a wide variety of heavy-tailed noise models of practical interest including Cauchy noise. In Section II-B, the complex M-estimators of the scatter are introduced. ML-estimators for the complex t-distribution are members in this class. We also introduce complex analogues of Huber's [7] and Tyler's [8] M-estimators. In Section III, the sensor array signal model and subspace and stochastic ML DOA estimation methods are reviewed. In Section IV, simulation studies demonstrating the reliable performance of the subspace DOA estimation and MDL number of signals estimation methods associated with the proposed M-estimators are given. We also demonstrate the reliable performance of the stochastic ML method based on complex multivariate t-distribution as the array output model distribution. Section V concludes the paper.

II. Robust estimators of the pseudo-covariance

A. ML-estimators

Consider a data set $\mathbf{z}_1, \ldots, \mathbf{z}_n$ in $\mathbb{C}^k$ and a set $\mathcal{P}_k$ of all $k \times k$ positive definite hermitian matrices. We will use superscript H to denote the hermitian transpose and $j = \sqrt{-1}$ is the imaginary unit. Sample mean $\bar{\mathbf{z}} = \text{ave}\{\mathbf{z}_i\}$ and covariance matrix $S = \text{ave}\{(\mathbf{z}_i - \bar{\mathbf{z}})(\mathbf{z}_i - \bar{\mathbf{z}})^H\}$ are the ML-estimators of the location $\mu \in \mathbb{C}^k$ and scatter $\Sigma \in \mathcal{P}_k$, respectively, if $\mathbf{z}_i$'s are i.i.d. random vectors from the complex k-variate normal (Gaussian) distribution, denoted by $\mathbb{C}N_k(\mu, \Sigma)$. The complex normal distribution belongs to a class of complex elliptically symmetric (CES) distributions introduced in [11]. Here we discuss the ML-estimation of the multivariate location μ and scatter Σ when the underlying CES distribution is complex multivariate t-distribution defined as follows.

Definition 1: A random vector $\mathbf{z} \in \mathbb{C}^k$ is said to have a complex multivariate t-distribution with $\nu > 0$ degrees of freedom, denoted by $\mathbb{C}t_{k,\nu}(\mu, \Sigma)$, if its probability density function (pdf) is of the form

$$f(\mathbf{z}; \mu, \Sigma) = c|\Sigma|^{-1}\{1 + 2s/\nu\}^{-(2k+\nu)/2},$$

where $s = (\mathbf{z}-\mu)^H\Sigma^{-1}(\mathbf{z}-\mu)$ and c is a normalizing constant not dependent on the parameters $\mu \in \mathbb{C}^k$ and $\Sigma \in \mathcal{P}_k$.

Note that f is real and non-negative as $\Sigma \in \mathcal{P}_k$ (consequently $s \geq 0$ and $|\Sigma| > 0$). The complex multivariate t-distribution is analogous to the real multivariate t-distribution. Indeed, the real $2k$-variate vector composed of real and imaginary parts of $\mathbf{z}$ has a certain real multivariate t-distribution in $\mathbb{R}^{2k}$. The multivariate t-distribution with $\nu = 1$ is called the

multivariate complex Cauchy distribution. Cauchy distribution is a prominent robust heavy-tailed alternative for the Gaussian which is obtained in the limit $\nu \to \infty$. The complex multivariate t-distributions are thus useful for studying robustness of multivariate statistics as a decrement of ν yields a (t-)distribution with an increased heaviness of the tails.

The parameter μ is the symmetry center of the distribution and equals the expectation $E(\mathbf{z})$ for $\nu > 2$. By symmetry we mean that $\mathbf{z}-\mu$ has the same distribution as $\exp(j\theta)(\mathbf{z}-\mu)$ for all $\theta \in \mathbb{R}$. Furthermore, the covariance matrix of the k-variate t-distribution is $\mathrm{Cov}(\mathbf{z}) = \{\nu/(\nu-2)\}\Sigma$ for $\nu > 2$, whereas the Cauchy distribution ($\nu = 1$) has infinite variances. Thus the scatter matrix Σ is proportional to covariance matrix (when it exist), and may therefore also be termed as pseudo-covariance matrix. Note that in the case of Gaussian distribution ($\nu = \infty$), $\mathrm{Cov}(\mathbf{z}) = \Sigma$. A random vector $\mathbf{z}$ from the complex multivariate t-distribution can be generated as follows: for independent $\mathbf{y} \sim \mathbb{C}N_k(\mu, \Sigma)$ and $s \sim \chi^2_\nu$, $\mathbf{z} = (s/\nu)^{-1/2}\mathbf{y}$ has $\mathbb{C}t_{k,\nu}(\mu, \Sigma)$ distribution.

Analogous to the real case [10], one can show that the ML-estimates $\hat{\mu} \in \mathbb{C}^k$ and $\hat{\Sigma} \in \mathcal{P}_k$ of the location μ and scatter Σ are the joint solutions to the estimating equations:

$$\hat{\mu} = \mathrm{ave}\{u_\nu(s_i)\mathbf{z}_i\}/\mathrm{ave}\{u_\nu(s_i)\} \quad (1)$$

$$\hat{\Sigma} = \mathrm{ave}\{u_\nu(s_i)(\mathbf{z}_i - \hat{\mu})(\mathbf{z}_i - \hat{\mu})^H\}, \quad (2)$$

where

$$s_i = (\mathbf{z}_i - \hat{\mu})^H \hat{\Sigma}^{-1}(\mathbf{z}_i - \hat{\mu}) \quad (3)$$

and

$$u_\nu(s) = \frac{2k + \nu}{\nu + 2s}.$$

In this paper we will consider the scatter-only case, that is, we suppose that the location μ is known or fixed, and without loss of generality we assume $\mu = \mathbf{0}$. Note that the DOA estimation problem may be considered as a scatter-only problem (zero mean signals and noise). Nonetheless, computing the ML-estimate $\hat{\mu}$ is useful in some other applications where μ is of interest, thus serving as a robust alternative for the sample mean vector.

Algorithm: In the scatter-only problem, the ML-estimate of scatter thus satisfy

$$\hat{\Sigma} = \mathrm{ave}\{u_\nu(\mathbf{z}_i^H \hat{\Sigma}^{-1}\mathbf{z}_i)\mathbf{z}_i\mathbf{z}_i^H\}. \quad (4)$$

Given an initial estimate $\Sigma_0 \in \mathcal{P}_k$, define

$$\Sigma_{m+1} = \mathrm{ave}\{u_\nu(\mathbf{z}_i^H \Sigma_m^{-1}\mathbf{z}_i)\mathbf{z}_i\mathbf{z}_i^H\}. \quad (5)$$

Kent and Tyler [10] have shown that under mild regularity conditions on the data (e.g. $n > k$ and the data is not too concentrated in low-dimensional subspaces) the iterations Σ_{m+1} converge to the unique solution $\hat{\Sigma}$ of (4). The authors of [10] consider the real case only, but the complex case follows similarly. This also follows from the fact that $\hat{\Sigma}$ can be calculated using real arithmetic as follows. Replace the n vectors $\mathbf{z}_i = \mathbf{x}_i + j\mathbf{y}_i$ ($\mathbf{x}, \mathbf{y} \in \mathbb{R}^k$) in $\mathbb{C}^k$ by the $2n$ vectors $\tilde{\mathbf{z}}_i$ in $\mathbb{R}^{2k}$:

$$\tilde{\mathbf{z}}_i = (\mathbf{x}_i^T, \mathbf{y}_i^T)^T \quad \text{for } i = 1, \ldots, n,$$
$$\tilde{\mathbf{z}}_i = (-\mathbf{y}_i^T, \mathbf{x}_i^T)^T \quad \text{for } i = n+1, \ldots, 2n.$$

Thus we double the sample space and also the dimension. Based on the real data set $\tilde{\mathbf{z}}_1, \ldots, \tilde{\mathbf{z}}_{2n}$ calculate the real $2k \times 2k$ positive definite symmetric (PDS) M-estimate of scatter matrix $\tilde{\Sigma}$ which satisfy

$$\tilde{\Sigma} = \mathrm{ave}\{\tilde{u}_\nu(\tilde{\mathbf{z}}_i^T \tilde{\Sigma}^{-1}\tilde{\mathbf{z}}_i)\tilde{\mathbf{z}}_i\tilde{\mathbf{z}}_i^T\}, \quad (6)$$

where $\tilde{u}_\nu(s) = u_\nu(s/2)$. Given an initial real PDS $2k \times 2k$-matrix $\tilde{\Sigma}_0$, the iterations

$$\tilde{\Sigma}_{m+1} = \mathrm{ave}\{\tilde{u}_\nu(\tilde{\mathbf{z}}_i^T \tilde{\Sigma}_m^{-1}\tilde{\mathbf{z}}_i)\tilde{\mathbf{z}}_i\tilde{\mathbf{z}}_i^T\}$$

converge to the solution $\tilde{\Sigma}$ of (6) under mild regularity conditions [10]. It is immediate to verify that $\tilde{\Sigma}$ is of the form

$$\tilde{\Sigma} = \begin{pmatrix} \tilde{\Sigma}_1 & -\tilde{\Sigma}_2 \\ \tilde{\Sigma}_2 & \tilde{\Sigma}_1 \end{pmatrix},$$

where $\tilde{\Sigma}_1$ and $\tilde{\Sigma}_2$ are real PDS $k \times k$-matrices. The complex ML-estimate of scatter $\hat{\Sigma}$ can then be formed from components of $\tilde{\Sigma}$, that is, $\hat{\Sigma} = 2(\tilde{\Sigma}_1 + j\tilde{\Sigma}_2)$.

In practical implementation of the proposed algorithm (5), we used as a termination point $\varepsilon_m := \|I - \Sigma_{m-1}^{-1}\Sigma_m\| < \varepsilon$, where $\|\cdot\|$ is some matrix norm and ε is some predetermined tolerance level, for example $\varepsilon = 0.001$. The computational complexity is $O(n_{iter} n k^2)$, where n_{iter} is the number of iteration. Naturally, n_{iter} is strongly dependent on the choice of the tolerance level ε, but not significantly (in our experience based on simulations) on the choice of the initial estimate $\Sigma_0 \in \mathcal{P}_k$.

On the choice of ν: The choice of ν, and consequently the weight function $u_\nu()$, merits some discussion. For $\nu \gg 1$, the complex t-distribution for which the estimators are based on, is close to complex Gaussian distribution (having slightly heavier tails). However, since in the DOA estimation problem we wish to achieve good robustness, a small value such as $1 \leq \nu < 5$ needs to be chosen. We suggest to use $\nu = 1$ (corresponding to Cauchy distribution) which yields the best safeguard against outliers. If the underlying distribution is in fact the nominal $\mathbb{C}N_k(\mu, \Sigma)$ instead of $\mathbb{C}t_{k,\nu}(\mu, \Sigma)$, then $\hat{\Sigma}$ is not a consistent to Σ but to $\sigma\Sigma$, where $\sigma \in \mathbb{R}^+$ is the solution of

$$E[u_\nu(R/2\sigma)R/\sigma] = 2k \quad (7)$$

and R is a random deviate from a chi-square distribution with $2k$ degrees of freedom. Thereby, σ can be solved numerically since (7) can be written as a simple integral. Then the estimator $(1/\sigma)\hat{\Sigma}$ is a consistent estimator of the covariance matrix $\mathrm{Cov}[\mathbf{z}_i] = \Sigma$ in the Gaussian case. However, in many applications, the covariance matrix is needed only up to a multiplicative scalar factor and therefore the calculation of the consistency factor σ is indeed not necessary.

B. M-estimators

Since the ML-estimators are the solutions to a set of implicit equations (1) and (2), they clearly belong to a class of M-estimators introduced by Maronna [6], and to a more restricted class of redescending M-estimators studied by Kent and Tyler [10]. The authors of [6] and [10] consider the real case only, but in an analogous fashion we may define M-estimators for complex data set as follows.

Definition 2: The simultaneous M-estimators $\hat{\mu} \in \mathbb{C}^k$ and $\hat{\Sigma} \in \mathcal{P}_k$ of the location μ and scatter Σ based upon a sample $\mathbf{z}_1, \ldots, \mathbf{z}_n$ in $\mathbb{C}^k$ are the solutions to the estimating equations:

$$\hat{\mu} = \mathrm{ave}\{u_1(s_i)\mathbf{z}_i\}/\mathrm{ave}\{u_1(s_i)\}$$
$$\hat{\Sigma} = \mathrm{ave}\{u_2(s_i)(\mathbf{z}_i - \hat{\mu})(\mathbf{z}_i - \hat{\mu})^H\},$$

where s_i is given by (3) and u_1 and u_2 are real valued functions.

As noted earlier, we will focus on this paper to the scatter-only case and assume $\mu = \mathbf{0}$. Thus the M-estimate of scatter $\hat{\Sigma}$ is the solution to

$$\hat{\Sigma} = \mathrm{ave}\{u_2(\mathbf{z}_i^H \hat{\Sigma}^{-1} \mathbf{z}_i)\mathbf{z}_i \mathbf{z}_i^H\}. \tag{8}$$

Similarly as in Section II-A, $\hat{\Sigma}$ can be calculated using real arithmetics from the matrix components of the real $2k \times 2k$ PDS M-estimator of scatter $\tilde{\Sigma}$ satisfying

$$\tilde{\Sigma} = \mathrm{ave}\{\tilde{u}_2(\tilde{\mathbf{z}}_i^T \tilde{\Sigma}^{-1} \tilde{\mathbf{z}}_i)\tilde{\mathbf{z}}_i \tilde{\mathbf{z}}_i^T\}, \tag{9}$$

where $\tilde{u}_2(s) = u_2(s/2)$. Obtaining the above relation with the real M-estimator $\tilde{\Sigma}$ and the complex M-estimator $\hat{\Sigma}$ has theoretical and computational motivation: for existence and uniqueness of the solution $\hat{\Sigma}$ to (8) it is sufficient that the data $\tilde{\mathbf{z}}_1, \ldots, \tilde{\mathbf{z}}_{2n}$ satisfies the conditions (see [6], [9]) needed for the existence and uniqueness of the solution $\tilde{\Sigma}$ to (9).

Huber's M-estimator. The complex analogue of Huber's M-estimator of scatter, denoted by $M(q)$, can be defined with the weight function

$$u_2(s) = \begin{cases} 1/\beta & \text{for } s \le c^2 \\ c^2/s\beta & \text{for } s > c^2 \end{cases}$$

where c is a tuning constant defined so that $q = F_{\chi^2_{2k}}(2c^2)$. The scaling factor β is defined as

$$\beta = F_{\chi^2_{2(k+1)}}(2c^2) + c^2(1-q)/k,$$

where $F_{\chi^2_k}$ denotes the cdf of real chi-square distribution with k degrees of freedom. The scaling factor makes the scatter estimator consistent for $\mathrm{Cov}(\mathbf{z})$ in the complex Gaussian case. The algorithm described in Huber [7], Section 8, can be modified for the complex case.

Tyler's M-estimator. The complex analoque of Tyler's [8] M-estimator is obtained with $u_2(s) = k/s$. Complex Tyler's scatter $\hat{\Sigma}$ thus satisfies

$$\hat{\Sigma} = k\,\mathrm{ave}\{\mathbf{z}_i \mathbf{z}_i^H / (\mathbf{z}_i^H \hat{\Sigma}^{-1} \mathbf{z}_i)\}. \tag{10}$$

The solution is also the ML-estimate of scatter for the complex angular central Gaussian distribution as was noted by Kent [12]. Note also that with $\nu = 0$, $u_\nu()$ gives Tyler's weight function (but naturally $\nu = 0$ does not correspond to any t-distribution). The solution $\hat{\Sigma}$ can be calculated as follows. Given an initial estimate $\Sigma_0 \in \mathcal{P}_k$, define

$$\Sigma_{m+1} = k\,\mathrm{ave}\{\mathbf{z}_i \mathbf{z}_i^H / (\mathbf{z}_i^H \Sigma_m^{-1} \mathbf{z}_i)\}.$$

As in the real case [10], under mild regularity conditions on the data (e.g. $n > k$, $\mathbf{z}_i \neq 0$), the iterations Σ_{m+1} converge to the solution $\hat{\Sigma}$ of (10) which is unique up to scaling.

III. DOA ESTIMATION

A. Signal Model

Consider an array of k elements and p sources impinging on the array, where $k > p$. Assume that the incoming signals $s_1(t), \ldots, s_p(t)$ arriving at distinct DOAs $\theta_1, \ldots, \theta_p$ at time t are from point sources at far field, narrowband and non-coherent. As a result, at time t, we read a $k \times 1$ dimensional array output (snapshot) $\mathbf{z}(t) \in \mathbb{C}^k$ which is a weighted linear combination of the signal sources $\mathbf{s}(t) = [s_1(t), \ldots, s_p(t)]^T \in \mathbb{C}^p$ corrupted by additive noise $\mathbf{n}(t) \in \mathbb{C}^k$, that is,

$$\mathbf{z}(t) = A(\theta)\mathbf{s}(t) + \mathbf{n}(t),$$

where $A(\theta)$ is the $k \times p$ array steering matrix parametrized by the vector of DOAs $\theta = (\theta_1, \ldots, \theta_p)^T$. The noise is assumed to be zero mean, spatially white with variance σ^2 and uncorrelated with the signals. $A(\theta)$ is assumed to be full rank. The observed data $\mathbf{z}_1 = \mathbf{z}(t_1), \ldots, \mathbf{z}_n = \mathbf{z}(t_n)$ sampled at discrete time instants $t_1, \ldots, t_n$ are then modeled as i.i.d. random vectors.

Observe that in the noiseless case the measurements lie in a p-dimensional signal subspace spanned by the columns of $A(\theta)$. But since white noise with diagonal scatter matrix is present any collection $\mathbf{z}_1, \ldots, \mathbf{z}_{k+1}$ of measurements will with probability one span all $\mathbb{C}^k$ yielding the so called low-rank signal in full-rank noise data model. The assumptions of $k > p$ and spatially white noise are restrictions needed by the subspace DOA methods. In some cases the noncoherence assumption of the signals can be relaxed [1].

B. Subspace Methods

Subspace methods [1] employ the structure of the covariance matrix of $\mathbf{z}(t)$,

$$\mathrm{Cov}[\mathbf{z}(t)] = A(\theta)\Omega A(\theta)^H + \sigma^2 I, \tag{11}$$

where $\sigma^2 I = \mathrm{Cov}[\mathbf{n}(t)]$ and $\Omega = \mathrm{Cov}[\mathbf{s}(t)]$ are the noise and signal covariance matrices respectively. Consequently the $k - p$ smallest eigenvalues of $\mathrm{Cov}[\mathbf{z}(t)]$ are equal to σ^2 and the corresponding eigenvectors $\gamma_{p+1}, \ldots, \gamma_k$ are orthogonal to the columns of $A(\theta)$. These eigenvectors span the *noise subspace* and the eigenvectors $\gamma_1, \ldots, \gamma_p$ corresponding to p largest eigenvalues span the signal subspace (the column space of $A(\theta)$).

The subspace DOA estimation methods are based on different properties of the signal and noise subspaces. Basically, subspace methods needs to solve the following two problems:

1) Find an estimate of the signal subspace $\Gamma_S = [\gamma_1 \cdots \gamma_p]$ or noise subspace $\Gamma_N = [\gamma_{p+1} \cdots \gamma_k]$
2) Find $\hat{\theta}$ which best matches the derived criterion, for example, find $\hat{\theta}$ such that distance between subspace $A(\hat{\theta})$ and the estimated subspace $\hat{\Gamma}_S$ is minimal in some sense.

Commonly, the subspace methods differ only in how they approach Problem 2 since the estimates $\hat{\Gamma}_N$ or $\hat{\Gamma}_S$ are obtained from eigenvectors of $S = \mathrm{ave}\{\mathbf{z}_i\mathbf{z}_i^H\}$. Problem 1, however, is in a way more crucial: no matter how clever criterion is used or how distances between subspaces are measured in Problem 2, the DOA estimate $\hat{\theta}$ will be unreliable if the estimates of the subspaces $\hat{\Gamma}_S$ or $\hat{\Gamma}_N$ are unreliable.

To obtain an estimate $\hat{\Gamma}_N$ or $\hat{\Gamma}_S$, we need to estimate the covariance matrix only up to a scalar factor. We may plug in the estimates of $\hat{\Gamma}_S$ or $\hat{\Gamma}_N$ based on M-estimators to existing subspace methods for solving Problem 2. This approach has been used also by other authors. For example, Visuri, Koivunen and Oja [5] employed spatial sign covariance matrix and Williams and Johnson [3] used ML estimation.

C. Stochastic Maximum Likelihood

In the classical stochastic ML (SML) approach [1] the noise and the signal distribution are modeled as complex Gaussian, in which case $\mathbf{z}(t) \sim \mathbb{C}N_k(\mathbf{0}, \Sigma)$ (recall that $\Sigma = \mathrm{Cov}[\mathbf{z}(t)]$) and the signal parameters θ, $\Omega \in \mathcal{P}_p, \sigma^2 \in \mathbb{R}^+$ and consequently Σ are found by solving

$$\{\hat{\theta}, \hat{\Omega}, \hat{\sigma}^2\} = \arg\min_{\theta,\Omega,\sigma^2} \{\log[\det(\Sigma)] + Tr[\Sigma^{-1}S]\},$$

where $S = \mathrm{ave}\{\mathbf{z}_i\mathbf{z}_i^H\}$. We may then construct the scatter matrix estimate by

$$\hat{\Sigma}_{\mathrm{ml}} = A(\hat{\theta})\hat{\Omega}A^H(\hat{\theta}) + \hat{\sigma}^2 I. \tag{12}$$

Although optimal (if the model is correct), the drawback of the SML method is that it leads to a difficult multidimensional nonlinear optimization problem. To obtain robust estimates of DOAs, one could use the complex multivariate t-distribution of Definition 1 as a (robust) array output model distribution. Thus, if we model $\mathbf{z}(t) \sim \mathbb{C}t_{k,\nu}(\mathbf{0}, \Sigma)$, the ML estimate of the signal parameters are found from

$$\{\hat{\theta}, \hat{\Omega}, \hat{\sigma}^2\} = \arg\min_{\theta,\Omega,\sigma^2} \big\{ \log[\det(\Sigma)] + \{(2k+\nu)/2\}\mathrm{ave}\{\log(1 + 2\mathbf{z}_i^H\Sigma^{-1}\mathbf{z}_i/\nu)\}\big\}.$$

See also Williams and Johnson [3] who used heavy-tailed array output distributions together with the SML approach.

IV. Simulation Examples

A. Subspace DOA Estimation

We now compare the quality of the MUSIC pseudospectrum associated with each estimate of the scatter matrix. A 8-element ($k = 8$) Uniform Linear Array (ULA) with interelement spacing equal to $\lambda/2$ was used. Two uncorrelated signals ($p = 2$) with SNR 20 dB are impinging on the array from DOA's $\theta_1 = -2^o$ and $\theta_2 = 2^o$ (broadside). Under these assumptions the covariance matrix (when it exists) is $\mathrm{Cov}[\mathbf{z}(t)] = E(|s_1(t)|^2)A(\theta)A(\theta)^H + E(|n_1(t)|^2)I$. In our study $n = 300$ snapshots are generated from complex Gaussian $\mathbb{C}N_k(\mathbf{0}, \Sigma)$ and complex Cauchy $\mathbb{C}t_{k,1}(\mathbf{0}, \Sigma)$. Recall that Σ is proportional to $\mathrm{Cov}[\mathbf{z}(t)]$ (when it exists).

We estimated the noise subspace Γ_N needed by MUSIC using: **a)** classical sample estimator $S = \mathrm{ave}\{\mathbf{z}_i\mathbf{z}_i^H\}$, **b)** ML-estimate of scatter (4) for the Cauchy distribution (i.e. $\nu = 1$), denoted by t_1M-estimator, **c)** Huber's M-estimator $M(q)$ with choice $q = 0.9$, **d)** ML estimate $\hat{\Sigma}_{\mathrm{ml}}$ (12) assuming the underlying CES distribution of the snapshots (i.e. complex Gaussian and Cauchy) is known. We assumed that the number of signals is known. Note that the SML method yields the DOA estimates directly. However, in our simulation study, also the MUSIC pseudospectrum based on $\hat{\Sigma}_{\mathrm{ml}}$ is included to facilitate the comparison with the suboptimal but more practical subspace methods based on scatter matrix estimators.

Figure 1 depicts the MUSIC pseudospectrum associated with the various estimates of the scatter matrix for five simulated complex Gaussian and Cauchy snapshots. The pseudospectrum using ML-estimate $\hat{\Sigma}_{\mathrm{ml}}$ which exploits the structure of the covariance matrix (hermitian Toepliz) and the underlying model distribution is, as expected, the best in the sense that the signals are more accurately resolved. We notice that all the estimators are able to resolve the two sources in the Gaussian case, and that the scatter estimators, S, t_1M-estimator and Huber's $M(0.9)$-estimator, perform comparably. In the Cauchy case, however, the classical sample estimator S is not able to resolve the sources. The robust M-estimators, however, yield reliable estimates of the DOAs.

B. Estimating number of signals

We now compare the performance of the scatter estimators in resolving the number of sources using the Minimum Description Length (MDL) criteria (see [13]). In other words, the eigenvalues based on different scatter estimators are plugged in to the MDL criterion. Recall that classically the estimates of the eigenvalues are based on the sample covariance matrix. The ULA contains $k = 8$ sensors with half a wavelength interelement spacing and the DoAs are at $\theta_1 = -5^o$ and $\theta_2 = 5^o$. Two uncorrelated Gaussian signals $s_1(t)$ and $s_2(t)$ with equal power 20 dB are impinging on the array. The components $n_1(t), \ldots, n_k(t)$ of the additive noise $\mathbf{n}(t)$ are modeled as i.i.d. with complex symmetric α-stable (SαS) distribution [4] with dispersion $\gamma = 1$ and values α ranging from $\alpha = 1$ (complex Cauchy noise) to $\alpha = 2$ (complex Gaussian noise). Simulations results are based on 200 Monte Carlo runs with $n = 300$ as the sample size. Figure 2 depicts the relative proportion of resolved sources using MDL criterion based on the S and t_1M-estimator. Tyler's and Huber's estimators are omitted in the Figure since percentage of the resolved sources differed from those of t_1M-estimator by at most 2% for all cases of α-values. We notice that MDL criterion based on t_1M-estimator is able to maintain 100% correct estimates of the number of signals for $1.4 \leq \alpha \leq 2$ and has 87% correct detection as its worst case for $\alpha = 1$. The performance of the

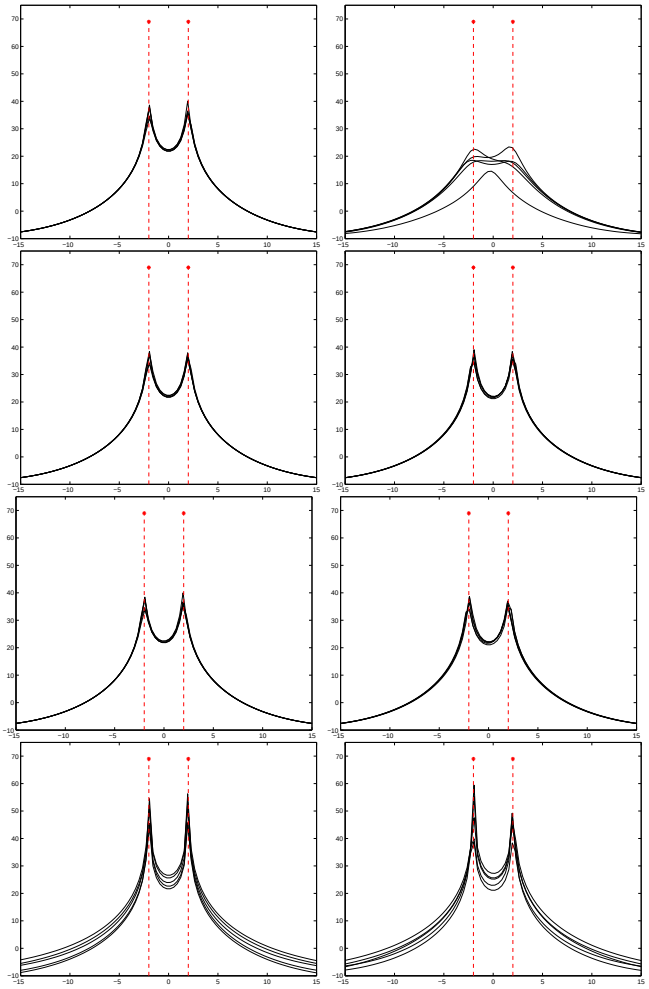

Fig. 1. MUSIC pseudospectrum based on S (first row), t_1M-estimator (second row), $M(0.9)$-estimator (third row) and $\hat{\Sigma}_{\mathrm{ml}}$ (fourth row) for five simulated data sets generated from complex normal (first column) and Cauchy (second column) distribution.

classical MDL employing the sample covariance matrix S is disastrous: it is able to estimate the number of signals reliably only for $\alpha = 2$, i.e. the Gaussian case.

V. Conclusion

In this paper we introduced M-estimators of scatter (i.e. pseudo-covariance) for complex data set observed by receiver antenna array. Simple iterative algorithm may be used for all of the considered M-estimators. The ML-estimator corresponding to complex t-distribution and Tyler's scatter estimator are especially convenient since, unlike the Huber's M-estimator, they do not require any tuning constants. Our simulation studies demonstrated that number of signals and subspace DOA estimation methods based on M-estimators perform comparably to optimal methods under nominal Gaussian conditions but have a superior performance under heavy-tailed non-Gaussian environments. Furthermore, out simulation study demonstrated that the complex multivariate t-distribution is also convenient as a model distribution for the SML approach yielding robust

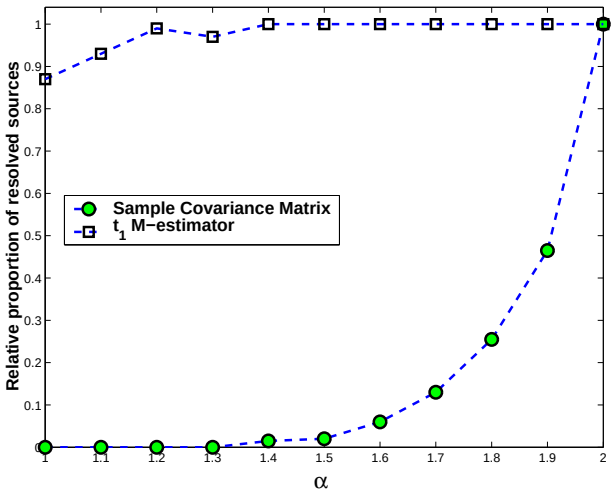

Fig. 2. Relative proportion of resolved sources using MDL criterion based on the sample covariance matrix and the t_1M-estimator.

DOA estimates. Robustness of the considered M-estimators follows from the fact that, unlike the regular mean and covariance, they possess a bounded influence function and a positive breakdown point [14]. More theoretical treatment of the complex M-estimators and CES distributions can be found in [14].

References

[1] H. Krim and M. Viberg, "Two decades of array signal processing: the parametric approach," *IEEE Signal Processing Mag.*, vol. 13, no. 4, pp. 67–94, 1996.

[2] D. Middleton, "Man-made noise in urban environments and transportation systems: Models and measurements," *IEEE Trans. Commun.*, vol. 21, pp. 1232–1241, 1973.

[3] D. B. Williams and D. H. Johnson, "Robust estimation of structured covariance matrices," *IEEE Trans. Signal Processing*, vol. 41, no. 9, pp. 2891–2905, 1993.

[4] P. Tsakalides and C. L. Nikias, "The robust covariation based MUSIC (roc-MUSIC) algorithm for bearing estimation in impulsive noise environments," *IEEE Trans. Signal Processing*, vol. 44, no. 7, pp. 1623–1633, 1995.

[5] S. Visuri, H. Oja, and V. Koivunen, "Subspace-based direction of arrival estimation using nonparametric statistics," *IEEE Trans. Signal Processing*, vol. 49, no. 9, pp. 2060–2073, 2001.

[6] R. A. Maronna, "Robust M-estimators of multivariate location and scatter," *Ann. Statist.*, vol. 5, no. 1, pp. 51–67, 1976.

[7] P. J. Huber, *Robust Statistics*. New York: Wiley, 1981.

[8] D. E. Tyler, "A distribution-free M-estimator of multivariate scatter," *Ann. Statist.*, vol. 15, no. 1, pp. 234–251, 1987.

[9] ——, "Some results on the existence, uniqueness, and computation of the M-estimates of multivariate location and scatter," *SIAM J. Stat. Sci. Comp.*, vol. 9, pp. 354–362, 1988.

[10] J. T. Kent and D. E. Tyler, "Redescending M-estimates of multivariate location and scatter," *Ann. Statist.*, vol. 19, no. 4, pp. 2102–2119, 1991.

[11] P. R. Krishnaiah and J. Lin, "Complex elliptically symmetric distributions," *Comm. Statist. - Th. and Meth.*, vol. 15, pp. 3693–3718, 1986.

[12] J. T. Kent, "Data analysis for shapes and images," *J. Statist. Plann. Inference*, vol. 57, pp. 181–193, 1997.

[13] T. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, no. 2, pp. 387–392, 1985.

[14] E. Ollila, S. Visuri, and V. Koivunen, "On the estimation of location and scatter of complex elliptically symmetric distributions," *Technometrics*, submitted for publication.