On Scatter Matrix Estimation in the Presence of Unknown Extra Parameters: Mismatched Scenario

Stefano Fortunati[1,2], Fulvio Gini[1,2], Maria S. Greco[1,2]
[1]Dipartimento di Ingegneria dell'Informazione, University of Pisa,
[2]CNIT RaSS (Radar and Surveillance System) National Laboratory, Pisa, Italy
E-mails: {stefano.fortunati,f.gini,m.greco}@iet.unipi.it

Abstract—**In this paper, a Constrained Mismatched Maximum Likelihood (CMML) estimator for the joint estimation of the scatter matrix and the power of Complex Elliptically Symmetric (CES) distributed vectors is derived under misspecified data models. Specifically, this estimator is obtained by assuming a Normal model while the data are sampled from a complex *t*-distribution. The convergence point of such CMML estimator is investigated and its Mean Square Error (MSE) compared with the Constrained Misspecified Cramér-Rao Bound (CMCRB).**

Keywords - Misspecified model; Covariance estimation; Constrained Maximum likelihood; Cramér-Rao Bound.

I. INTRODUCTION

This paper deals with the scatter matrix estimation problem for Complex Elliptically Symmetric (CES) distributed data in the presence of unknown extra parameters. Thanks to their flexibility, the CES distributions represent a reliable data model in many areas such as radar, sonar, and communications [1]. The complex Normal, Generalized Gaussian, *K*-distribution, and the complex *t*-distribution are some examples of probability density functions (pdf) belonging to the CES class. A CES distribution is completely characterized by the mean value $\boldsymbol{\gamma}$, the scatter (or shape) matrix $\boldsymbol{\Sigma}$ and the density generator *h*. Given a particular CES distribution, its density generator could depend on some extra parameters, (e.g. shape and scale parameters for a complex *t*-distribution) that are in general unknown and need to be estimated from the data along with $\boldsymbol{\gamma}$ and $\boldsymbol{\Sigma}$. However, the joint Maximum Likelihood (ML) estimator of all these unknown quantities often encounter computational difficulties and convergence (or even existence) issues. Moreover, even in those cases in which a joint estimator could be successfully derived, a perfect match between the true statistical data model and the one assumed to derive the estimator is often difficult to achieve. For all these reasons, one has to rely on mismatched estimators, i.e. on the estimation algorithm derived under an assumed data distribution different from the true one. In [2], we investigated the scatter matrix estimation problem for CES distributed vectors in the presence of model mismatching, but under the assumptions of a perfect knowledge of the extra parameters. In particular, the performance of the Mismatched ML (MML) estimator has been compared with the Misspecified Cramér-Rao Bound (MCRB) in some relevant study cases.

The main aim of this paper is to generalize the results in [2] to the more realistic case in which both the scatter matrix and the unknown extra parameters of the assumed CES distribution need to be estimated from the data. In particular, we investigate a recurring scenario in radar applications: the true data pdf is a complex *t*-distribution, while the ML estimator of the scatter matrix and of the data power is derived under a Normal model assumption. The choice of the *t*-distribution as true data model has been motivated by experimental evidences (see e.g. [3]) that proved its reliability to model spiky clutter data. On the other hand, many radar systems exploit the Normal model for data inference due to its analytical tractability and the consequent real time implementation of the estimation algorithms based on it.

II. OVERVIEW OF THE CES DISTRIBUTIONS

A complex *N*-dimensional random vector $\mathbf{x}_m$ is CES distributed, i.e. $\mathbf{x}_m \sim CE_N(\boldsymbol{\gamma},\boldsymbol{\Sigma},h)$, if its pdf is of the form:

$$p_X(\mathbf{x}_m) = c_{N,h}|\boldsymbol{\Sigma}|^{-1} h\left((\mathbf{x}_m-\boldsymbol{\gamma})^H \boldsymbol{\Sigma}^{-1}(\mathbf{x}_m-\boldsymbol{\gamma})\right) \tag{1}$$

where *h* is the density generator, $c_{N,h}$ is a normalizing constant, $\boldsymbol{\gamma} \triangleq E\{\mathbf{x}_m\}$, and $\boldsymbol{\Sigma}$ is the normalized (or shape) covariance matrix, also called scatter matrix. Due to the well-known ambiguity between the scatter matrix and the density generator of any CES distribution, we impose the following constraint: tr($\boldsymbol{\Sigma}$)=*N*. As a consequence, if $\mathbf{M} \triangleq E\{(\mathbf{x}_m-\boldsymbol{\gamma})(\mathbf{x}_m-\boldsymbol{\gamma})^H\}$ is the covariance matrix of the random vector $\mathbf{x}_m$, then $\boldsymbol{\Sigma}=N\mathbf{M}/\mathrm{tr}(\mathbf{M})$. It is important to observe that, for some CES distributions, the un-normalized covariance matrix $\mathbf{M}$ does not exist, but the scatter matrix $\boldsymbol{\Sigma}$ is still well defined. Based upon the Stochastic Representation Theorem [1] any $\mathbf{x}_m \sim CE_N(\boldsymbol{\gamma},\boldsymbol{\Sigma},h)$ with rank($\boldsymbol{\Sigma}$)=$k \le N$ admits the stochastic representation $\mathbf{x}_m =_d \boldsymbol{\gamma} + R\mathbf{P}\mathbf{u}$, where the non-negative random variable (r.v.) $R \triangleq \sqrt{Q}$, the so-called *modular variate*, is a real, non-negative random variable, $\mathbf{u}$ is a *k*-dimensional vector uniformly distributed on the unit hyper-sphere with *k*-1 topological dimensions such that $\mathbf{u}^H\mathbf{u}=1$, *R* and $\mathbf{u}$ are independent and $\boldsymbol{\Sigma}=\mathbf{P}\mathbf{P}^H$ is a factorization of $\boldsymbol{\Sigma}$, where $\mathbf{P}$ is a *N*x*k* matrix and rank($\mathbf{P}$)=*k*. In the following, we assume that $\boldsymbol{\Sigma}$ is real and full-rank. For CES distributions, the term $\sigma^2 \triangleq E\{Q\}/N$ can be interpreted as the statistical power of the random vector $\mathbf{x}_m$, i.e.

the covariance matrix $\mathbf{M}$ and the scatter matrix $\mathbf{\Sigma}$ are related by $\mathbf{M}=\sigma^2\mathbf{\Sigma}$.

III. THE MISMATCHED SCENARIO

A. True data model: the complex t-distribution.

The complex t-distribution is completely characterized by its mean vector $\bar{\boldsymbol{\gamma}}$, its scatter matrix $\bar{\mathbf{\Sigma}}$ and its shape and scale parameters, λ and η respectively. In particular, a complex N-dimensional zero-mean ($\bar{\boldsymbol{\gamma}}=0$) random vector $\mathbf{x}_m$ is t-distributed if its pdf can be expressed as [1]:

$$p_X(\mathbf{x}_m;\bar{\mathbf{\Sigma}},\lambda,\eta)=\frac{1}{\pi^N\left|\bar{\mathbf{\Sigma}}\right|}\frac{\Gamma(N+\lambda)}{\Gamma(\lambda)}\left(\frac{\lambda}{\eta}\right)^{\lambda}\left(\frac{\lambda}{\eta}+\mathbf{x}_m^H\bar{\mathbf{\Sigma}}^{-1}\mathbf{x}_m\right)^{-(N+\lambda)} \quad (2)$$

The complex t-distribution has tails heavier than the Normal one for every $\lambda\in(0,+\infty)$, while the limiting case $\lambda\to+\infty$ yields the complex Normal distribution. Moreover, the statistical power is a function of λ and η as follows [2]:

$$\sigma^2=E\{Q_{\Sigma}\}/N=\lambda/\eta(\lambda-1)\,. \quad (3)$$

B. Assumed data model: the complex Normal distribution.

We assume a complex Normal model for the data, i.e. we assume that the M vectors of the available dataset $\mathbf{x}=\{\mathbf{x}_m\}_{m=1}^M$ are, zero-mean, *independent, identically distributed (iid)* and each one is distributed according to a complex Normal multivariate pdf, which also belongs to the CES family:

$$f_X(\mathbf{x}_m;\boldsymbol{\theta})\triangleq f_X(\mathbf{x}_m;\mathbf{\Sigma},\sigma^2)=\frac{1}{(\pi\sigma^2)^N|\mathbf{\Sigma}|}\exp\left(-\frac{\mathbf{x}_m^H\mathbf{\Sigma}^{-1}\mathbf{x}_m}{\sigma^2}\right). \quad (4)$$

The covariance matrix is $\mathbf{M}=E\{\mathbf{x}_m\mathbf{x}_m^H\}=\sigma^2\mathbf{\Sigma}$, where tr($\mathbf{\Sigma}$)=N and σ^2 is the power. Since the statistical inference on the data is built on the assumed model, the parameter vector to be estimated can be expressed as $\boldsymbol{\theta}=[\text{vecs}(\mathbf{\Sigma})^T \quad \sigma^2]^T$, where the vecs-operator is the "symmetric" counterpart of the standard vec-operator that maps a symmetric NxN matrix $\mathbf{\Sigma}$ in a l=N(N+1)/2-dimensional vector whose entries are the elements of the lower (or upper) triangular sub-matrix of $\mathbf{\Sigma}$. The mismatch occurs in the estimation of $\boldsymbol{\theta}$ since the true data are distributed according to the complex t-distribution $p_X(\mathbf{x}_m;\bar{\boldsymbol{\theta}})$ in eq. (2), where $\bar{\boldsymbol{\theta}}\triangleq[\text{vecs}(\bar{\mathbf{\Sigma}})^T \quad \lambda \quad \eta]^T$ is the true parameter vector and $\bar{\mathbf{\Sigma}}$ is the true scatter matrix. It is worth noting that, in the mismatched scenario, the true parameter space and the assumed parameter space are different. Specifically, the true parameter space is $\bar{\boldsymbol{\theta}}\ni\mathrm{T}\subset\mathbb{R}^l\times(0,\infty)\times(0,\infty)$, while the assumed parameter space is $\boldsymbol{\theta}\ni\Theta\subset\mathbb{R}^l\times(0,\infty)$ where $\times$ indicates the Cartesian product. Moreover, the constraint on the trace of $\mathbf{\Sigma}$ limits both the true and assumed parameter vector to belong to two lower dimensional smooth manifolds $\tilde{\mathrm{T}}=\{\bar{\boldsymbol{\theta}}\in\mathrm{T}\,|\,\mathrm{tr}(\bar{\mathbf{\Sigma}})=N\}$ and $\tilde{\Theta}=\{\boldsymbol{\theta}\in\Theta\,|\,\mathrm{tr}(\mathbf{\Sigma})=N\}$, respectively.

IV. THE CONSTRAINED MML ESTIMATOR

In order to obtain an estimate of $\boldsymbol{\theta}$, we apply the ML method under the assumed data model, so what we get under mismatched conditions is the so-called MML estimator [4]:

$$\hat{\boldsymbol{\theta}}_{MML}(\mathbf{x})\triangleq\underset{\boldsymbol{\theta}\in\tilde{\Theta}}{\arg\max}\ln f_X(\mathbf{x};\boldsymbol{\theta})=\underset{\boldsymbol{\theta}\in\tilde{\Theta}}{\arg\max}\sum\nolimits_{m=1}^M\ln f_X(\mathbf{x}_m;\boldsymbol{\theta}) \quad (5)$$

where $\mathbf{x}_m\sim p_X(\mathbf{x}_m;\bar{\boldsymbol{\theta}})$ is given in eq. (2) and $\mathrm{tr}(\mathbf{\Sigma})=\mathrm{tr}(\bar{\mathbf{\Sigma}})=N$ is the constraint. It can be shown (see [2], [4] and references therein) that the MML estimator converges almost surely (*a.s.*) to the so-called *pseudo-true parameter vector* $\boldsymbol{\theta}_0$, i.e. the vector that minimizes the Kullback-Leibler divergence (KLD) between $p_X(\mathbf{x}_m;\bar{\boldsymbol{\theta}})$ and $f_X(\mathbf{x}_m;\boldsymbol{\theta})$:

$$\hat{\boldsymbol{\theta}}_{MML}(\mathbf{x})\underset{M\to\infty}{\overset{a.s.}{\to}}\boldsymbol{\theta}_0\,, \quad (6)$$

$$\boldsymbol{\theta}_0=\underset{\boldsymbol{\theta}\in\tilde{\Theta}}{\arg\min}\left\{D\left(p\|f_{\boldsymbol{\theta}}\right)\right\}=\underset{\boldsymbol{\theta}\in\tilde{\Theta}}{\arg\min}\left\{-E_p\left\{\ln f_X(\mathbf{x}_m;\boldsymbol{\theta})\right\}\right\} \quad (7)$$

where:

$$D\left(p\|f_{\boldsymbol{\theta}}\right)\triangleq\int\ln\left(\frac{p_X(\mathbf{x}_m;\bar{\boldsymbol{\theta}})}{f_X(\mathbf{x}_m;\boldsymbol{\theta})}\right)p_X(\mathbf{x}_m;\bar{\boldsymbol{\theta}})d\mathbf{x}\,. \quad (8)$$

Due to the lack of space, we refer the reader to [2], [4] and references therein for a more comprehensive review of the asymptotic properties of the MML estimator (e.g. its asymptotical Gaussianity). In [2] the MML estimator of the scatter matrix was evaluated for the complex-t distribution when the assumed misspecified distribution is a complex Normal pdf, under the assumption of a-priori known power. Here, we generalize this result for the case of unknown power, i.e. when the power σ^2 and the scatter matrix $\mathbf{\Sigma}$ are unknown and need to be jointly estimated. Due to the well-known ambiguity between the scatter matrix and the density generator of a CES distribution, the power and the scatter matrix are jointly identifiable if and only if a constrain on $\mathbf{\Sigma}$ is established [1]. Even if different constraints can be set on $\mathbf{\Sigma}$, here we choose tr($\mathbf{\Sigma}$)=N since it guarantees the following factorization of the covariance matrix: $\mathbf{M}=\sigma^2\mathbf{\Sigma}$, as discussed in Sect. II.

To derive analytically the constrained MML (CMML) estimator under the mismatched scenario discussed in Sect. III, one has to find the maximum of the log-likelihood function subjected to the linear constraint tr($\mathbf{\Sigma}$)=N. To do this, we do not rely on the Lagrange multiplier method, but we follow a different, yet equivalent, procedure [5]: we first derive the unconstrained MML estimator and then we project it on the (lower dimensional) manifold $\tilde{\Theta}$ by imposing the constraint.

The likelihood function can be expressed as:

$$\begin{aligned}L(\boldsymbol{\theta})&=\sum\nolimits_{m=1}^M\ln f_X(\mathbf{x}_m;\boldsymbol{\theta})\\&=-NM\ln\sigma^2-M\ln|\mathbf{\Sigma}|-\sum\nolimits_{m=1}^M\mathbf{x}_m^H\mathbf{\Sigma}^{-1}\mathbf{x}_m/\sigma^2.\end{aligned} \quad (9)$$

Then, the MML estimator can be obtained by solving the following problem:

$$\begin{cases} \dfrac{\partial L(\boldsymbol{\theta})}{\partial \sigma^2} = -\dfrac{NM}{\sigma^2} + \dfrac{1}{\sigma^4}\sum_{m=1}^{M} \mathbf{x}_m^H \boldsymbol{\Sigma}^{-1}\mathbf{x}_m = 0 \\ \dfrac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\Sigma}} = -M\boldsymbol{\Sigma}^{-1} + \dfrac{\boldsymbol{\Sigma}^{-1}}{\sigma^2}\sum_{m=1}^{M} \mathbf{x}_m \mathbf{x}_m^H \boldsymbol{\Sigma}^{-1} = \mathbf{0} \end{cases}. \tag{10}$$

Then, imposing the constraint, we have:

$$\begin{cases} \hat{\sigma}^2_{MML} = \dfrac{1}{NM}\sum_{m=1}^{M} \mathbf{x}_m^H \hat{\boldsymbol{\Sigma}}^{-1}_{MML}\mathbf{x}_m \\ \hat{\boldsymbol{\Sigma}}_{MML} = \dfrac{N}{\sum_{m=1}^{M} \mathbf{x}_m^H \hat{\boldsymbol{\Sigma}}^{-1}_{MML}\mathbf{x}_m}\sum_{m=1}^{M} \mathbf{x}_m \mathbf{x}_m^H \\ \text{s.t.} \quad \mathrm{tr}(\hat{\boldsymbol{\Sigma}}_{MML}) = N \end{cases}. \tag{11}$$

Hence, we get the CMML estimators of $\boldsymbol{\Sigma}$ and σ^2:

$$\begin{cases} \hat{\sigma}^2_{CMML} = \dfrac{1}{NM}\sum_{m=1}^{M} \mathbf{x}_m^H \hat{\boldsymbol{\Sigma}}^{-1}_{CMML}\mathbf{x}_m \\ \hat{\boldsymbol{\Sigma}}_{CMML} = \dfrac{N}{\sum_{m=1}^{M} \mathbf{x}_m^H \mathbf{x}_m}\sum_{m=1}^{M} \mathbf{x}_m \mathbf{x}_m^H \end{cases}. \tag{12}$$

Now we need to find the vector $\boldsymbol{\theta}_0 = [\mathrm{vecs}(\boldsymbol{\Sigma}_0)^T \quad \sigma_0^2]^T$ that minimizes the KLD between $p_X(\mathbf{x}_m;\bar{\boldsymbol{\theta}})$ and $f_X(\mathbf{x}_m;\boldsymbol{\theta})$. This vector is the convergence point of the CMML estimator in eq. (12). To this end, we have to solve the following system:

$$\begin{cases} \dfrac{\partial D(p\|f_{\boldsymbol{\theta}})}{\partial \sigma^2} = -E_p\left\{\dfrac{\partial \ln f_X(\mathbf{x}_m;\boldsymbol{\Sigma},\sigma^2)}{\partial \sigma^2}\right\} = 0 \\ \dfrac{\partial D(p\|f_{\boldsymbol{\theta}})}{\partial \boldsymbol{\Sigma}} = -E_p\left\{\dfrac{\partial \ln f_X(\mathbf{x}_m;\boldsymbol{\Sigma},\sigma^2)}{\partial \boldsymbol{\Sigma}}\right\} = \mathbf{0} \end{cases}. \tag{13}$$

The first equation immediately provides:

$$\frac{\partial D(p\|f_{\boldsymbol{\theta}})}{\partial \sigma^2} = E_p\left\{\frac{\partial}{\partial \sigma^2}\left(N\ln\sigma^2 + \frac{Q_{\boldsymbol{\Sigma}}}{\sigma^2}\right)\right\} = \frac{N}{\sigma^2} - \frac{E_p\{Q_{\boldsymbol{\Sigma}}\}}{\sigma^4} = 0 \tag{14}$$

where $Q_{\boldsymbol{\Sigma}} \triangleq \mathbf{x}_m^H \boldsymbol{\Sigma}^{-1}\mathbf{x}_m$. Solving eq. (14), we get $\sigma_0^2 = E_p\{Q_{\boldsymbol{\Sigma}}\}/N$. The derivative of the KLD with respect to $\boldsymbol{\Sigma}$ is given by [2]:

$$\frac{\partial D(p\|f_{\boldsymbol{\theta}})}{\partial \boldsymbol{\Sigma}} = \boldsymbol{\Sigma}^{-1} - \frac{E\{Q_{\boldsymbol{\Sigma}}\}}{N\sigma^2}\boldsymbol{\Sigma}^{-1}\bar{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1} = \mathbf{0}, \ \mathrm{tr}(\boldsymbol{\Sigma}) = \mathrm{tr}(\bar{\boldsymbol{\Sigma}}) = N \tag{15}$$

whose solution is: $\boldsymbol{\Sigma}_0 = E_p\{Q_{\boldsymbol{\Sigma}_0}\}\bar{\boldsymbol{\Sigma}}/N\sigma^2$. Putting together the two solutions, we finally get:

$$\boldsymbol{\Sigma}_0 = \frac{E_p\{Q_{\boldsymbol{\Sigma}_0}\}}{N\sigma_0^2}\bar{\boldsymbol{\Sigma}} = \bar{\boldsymbol{\Sigma}}, \quad \text{s.t.}\ \mathrm{tr}(\boldsymbol{\Sigma}_0) = \mathrm{tr}(\bar{\boldsymbol{\Sigma}}) = N, \tag{16}$$

where:

$$\sigma_0^2 = \frac{E_p\{Q_{\boldsymbol{\Sigma}_0}\}}{N} = \frac{E_p\{Q_{\bar{\boldsymbol{\Sigma}}}\}}{N} = \frac{E_p\{\mathbf{x}_m^H \bar{\boldsymbol{\Sigma}}^{-1}\mathbf{x}_m\}}{N} = \frac{\lambda}{\eta(\lambda-1)} = \bar{\sigma}^2, \tag{17}$$

and $\bar{\sigma}^2$ is the true power of the data. Eqs. (16)-(17) show that the CMML estimator converges *a.s.* to the parameter vector $\hat{\boldsymbol{\theta}}_{CMML}(\mathbf{x}) \underset{M\to\infty}{\overset{a.s.}{\to}} \boldsymbol{\theta}_0 = [\mathrm{vecs}(\bar{\boldsymbol{\Sigma}})^T \quad \bar{\sigma}^2]^T$, i.e.:

$$\hat{\sigma}^2_{CMML}(\mathbf{x}) \underset{M\to\infty}{\overset{a.s.}{\to}} \bar{\sigma}^2 = \lambda/\eta(\lambda-1), \quad \hat{\boldsymbol{\Sigma}}_{CMML}(\mathbf{x}) \underset{M\to\infty}{\overset{a.s.}{\to}} \bar{\boldsymbol{\Sigma}}. \tag{18}$$

Hence, it provides consistent estimates for both the scatter matrix and the power of the true data model.

V. CONSTRAINED MISSPECIFIED CRAMÉR-RAO BOUND

The mismatched counterpart of the Cramér-Rao Bound (CRB) for the error covariance matrix of any unbiased (in a proper sense) estimator of a deterministic parameter vector under misspecified models has been originally derived by Q. H. Vuong in [6], and then deeply investigated in [7]. This bound is known as the Misspecified Cramér-Rao Bound (MCRB). Recently, the application of the MCRB in different signal processing problems has been investigated and discussed: see e.g. [7] and [8] for Direction of Arrival (DoA) estimation problems and [2] for inference problems on CES distributed random vectors. In some applications, and in particular in the one discussed here and in [2], one has to deal with additional constraints on the unknown parameter vector. In our recent work [9], we generalized the findings of [10] and [11] on the Constrained CRB in order to obtain a constrained version of the MCRB, i.e. CMCRB. The finding in [9] can be summarized by the following Theorem:

Theorem: *A Constrained MCRB (CMCRB) for any misspecified (MS)-unbiased estimator of* $\boldsymbol{\theta}_0$ *is given by:*

$$\mathrm{CMCRB}(\boldsymbol{\theta}_0) = \mathbf{U}\left(\mathbf{U}^T\mathbf{A}_{\boldsymbol{\theta}_0}\mathbf{U}\right)^{-1}\mathbf{U}^T\mathbf{B}_{\boldsymbol{\theta}_0}\mathbf{U}\left(\mathbf{U}^T\mathbf{A}_{\boldsymbol{\theta}_0}\mathbf{U}\right)^{-1}\mathbf{U}^T, \tag{19}$$

where the (possibly singular) matrices $\mathbf{A}_{\boldsymbol{\theta}_0}$ *and* $\mathbf{B}_{\boldsymbol{\theta}_0}$ *are defined, as in the unconstrained case, as:*

$$\mathbf{A}_{\boldsymbol{\theta}_0} \triangleq E_p\left\{\nabla_{\boldsymbol{\theta}_0}\nabla^T_{\boldsymbol{\theta}_0}\ln f_X(\mathbf{x};\boldsymbol{\theta}_0)\right\}, \tag{20}$$

$$\mathbf{B}_{\boldsymbol{\theta}_0} \triangleq E_p\left\{\nabla_{\boldsymbol{\theta}_0}\ln f_X(\mathbf{x};\boldsymbol{\theta}_0)\nabla^T_{\boldsymbol{\theta}_0}\ln f_X(\mathbf{x};\boldsymbol{\theta}_0)\right\}, \tag{21}$$

$\mathbf{U}$ *is a matrix whose columns form an orthonormal basis of the null-space of the constraint's Jacobian matrix. Moreover, the definition of MS-unbiasedness can be found in* [2, Def. 2].

The proof of this Theorem can be found in [9].

In the following, we provide the explicit expression of the bound in (19) evaluated for the specific mismatched problem at hand: the joint estimation of the power and of the scatter matrix for an assumed Normal model when the data follow the complex *t*-distributed model instead. The parameter vector $\boldsymbol{\theta}_0$ is the one given in eq. (18), i.e. the convergence point of the CMML estimator.

A. Evaluation of $\mathbf{A}_{\boldsymbol{\theta}_0}$

$\mathbf{A}_{\boldsymbol{\theta}_0}$ can be decomposed in the following blocks:

$$\mathbf{A}_{\boldsymbol{\theta}_0} = \mathbf{T}^T\begin{bmatrix}\mathbf{A}_{\bar{\boldsymbol{\Sigma}}} & \mathbf{A}_c \\ \mathbf{A}_c^T & A_{\bar{\sigma}^2}\end{bmatrix}\mathbf{T}, \tag{22}$$

$$\mathbf{T} = \begin{bmatrix}\mathbf{D}_N & \mathbf{0} \\ \mathbf{0} & 1\end{bmatrix}, \tag{23}$$

where $\mathbf{D}_N$ is the so-called Duplication matrix of order N [12]. The duplication matrix is implicitly defined as the unique N^2xl (where $l=N(N+1)/2$) matrix that satisfies the following equality: $\mathbf{D}_N$vecs($\mathbf{A}$)=vec($\mathbf{A}$) for any symmetric matrix $\mathbf{A}$. Following the procedure in [2] and [13], we have:

$$\mathbf{A}_{\bar{\Sigma}} = -\bar{\mathbf{\Sigma}}^{-1} \otimes \bar{\mathbf{\Sigma}}^{-1} \tag{24}$$

$$A_{\bar{\sigma}^2} = E_p\left\{\frac{\partial^2}{\partial^2\bar{\sigma}^2}\ln f_X\left(\mathbf{x}_m;\mathbf{\theta}_0\right)\right\} = -\frac{N}{\bar{\sigma}^4}, \tag{25}$$

$$\mathbf{A}_c = E_p\left\{\frac{\partial^2 \ln f_X\left(\mathbf{x}_m;\mathbf{\theta}_0\right)}{\partial\bar{\sigma}^2\partial\mathrm{vec}(\bar{\mathbf{\Sigma}})}\right\} = -\frac{1}{\bar{\sigma}^2}\mathrm{vec}\left(\bar{\mathbf{\Sigma}}^{-1}\right). \tag{26}$$

B. *Evaluation of* $\mathbf{B}_{\mathbf{\theta}_0}$

$\mathbf{B}_{\mathbf{\theta}_0}$ can be decomposed in the following blocks:

$$\mathbf{B}_{\mathbf{\theta}_0} = \mathbf{T}^T\begin{bmatrix}\mathbf{B}_{\bar{\Sigma}} & \mathbf{B}_c \\ \mathbf{B}_c^T & B_{\bar{\sigma}^2}\end{bmatrix}\mathbf{T}. \tag{27}$$

As before, following [13], we get:

$$\mathbf{B}_{\bar{\Sigma}} = \frac{1}{\lambda-2}\mathrm{vec}\left(\bar{\mathbf{\Sigma}}^{-1}\right)\mathrm{vec}\left(\bar{\mathbf{\Sigma}}^{-1}\right)^T + \frac{\lambda-1}{\lambda-2}\bar{\mathbf{\Sigma}}^{-1}\otimes\bar{\mathbf{\Sigma}}^{-1} \tag{28}$$

$$B_{\bar{\sigma}^2} = E_p\left\{\left[\frac{\partial \ln f_X\left(\mathbf{x}_m;\mathbf{\theta}_0\right)}{\partial\bar{\sigma}^2}\right]^2\right\} = \frac{N(N+\lambda-1)}{\bar{\sigma}^4(\lambda-2)}, \tag{29}$$

$$\mathbf{B}_c = E_p\left\{\frac{\partial \ln f_X(\mathbf{x}_m;\mathbf{\theta}_0)}{\partial\bar{\sigma}^2}\cdot\frac{\partial \ln f_X(\mathbf{x}_m;\mathbf{\theta}_0)}{\partial\mathrm{vec}(\bar{\mathbf{\Sigma}})}\right\} = \frac{N+\lambda-1}{\bar{\sigma}^2(\lambda-2)}\mathrm{vec}(\bar{\mathbf{\Sigma}}^{-1}) \tag{30}$$

C. *Definition of the matrix* $\mathbf{U}$

The linear constraint tr($\mathbf{\Sigma}$)=N can be rewritten as:

$$f(\mathbf{\theta}) = \sum_{i\in I}\mathrm{vecs}(\mathbf{\Sigma})_i - N = 0, \tag{31}$$

where I is the set of the indices of the diagonal entries of $\mathbf{\Sigma}$ that can be explicitly described as:

$$I = \left\{i \,\middle|\, i = 1 + N(j-1) - \frac{(j-1)(j-2)}{2}, j = 1,\dots,N\right\}. \tag{32}$$

Following [10], we define the (l+1)-dim gradient vector as:

$$\nabla f(\mathbf{\theta}) = \frac{\partial f(\mathbf{\theta})}{\partial\mathbf{\theta}^T} = \left[\frac{\partial\sum_{i\in I}\mathrm{vecs}(\mathbf{\Sigma})_i}{\partial\mathrm{vecs}(\mathbf{\Sigma})^T} \quad 0\right] = \left[\mathbf{1}_I^T \quad 0\right], \tag{33}$$

where $\mathbf{1}_I$ is a l-dim column vector defined as:

$$\left[\mathbf{1}_I\right]_i = \begin{cases}1 & i\in I \\ 0 & \text{otherwise}\end{cases}. \tag{34}$$

The gradient $\nabla f(\mathbf{\theta})$ has clearly full row rank and hence there exists a matrix $\mathbf{U}\in\mathbb{R}^{(l+1)\times l}$ whose columns form an orthonormal basis for the null space of $\nabla f(\mathbf{\theta})$, that is $\nabla f(\mathbf{\theta})\mathbf{U} = \mathbf{0}$ where $\mathbf{U}^T\mathbf{U}=\mathbf{I}$. Finally, it can be noted that the matrix $\mathbf{U}$ can be obtained by evaluating the orthonormal eigenvectors associated to the zero eigenvalue of $\nabla f(\mathbf{\theta})$.

VI. PERFORMANCE ANALYSIS

In this section we compare the performance of the CMML estimator with the CMCRB for the estimation of the scatter matrix $\bar{\mathbf{\Sigma}}$ and the power $\bar{\sigma}^2$. To this end, we define the following performance indices for the estimation accuracy:

$$\varepsilon_{\hat{\Sigma}_{CMML}} \triangleq \left\| E\left\{\mathrm{vecs}(\hat{\mathbf{\Sigma}}_{CMML}-\bar{\mathbf{\Sigma}})\mathrm{vecs}(\hat{\mathbf{\Sigma}}_{CMML}-\bar{\mathbf{\Sigma}})^T\right\}\right\|_F, \tag{35}$$

$$MSE\left(\hat{\sigma}^2_{CMML}\right) \triangleq E\left\{\left(\hat{\sigma}^2_{CMML}-\bar{\sigma}^2\right)^2\right\}. \tag{36}$$

Correspondingly, the following performance bounds are calculated and plotted:

$$\varepsilon_{\mathrm{CMCRB}} \triangleq \left\|\mathrm{CMCRB}(\bar{\mathbf{\Sigma}})\right\|_F, \varepsilon_{\mathrm{CCRB}} \triangleq \left\|\mathrm{CCRB}(\bar{\mathbf{\Sigma}})\right\|_F, \mathrm{CMCRB}(\bar{\sigma}),$$

where the CCRB($\bar{\mathbf{\Sigma}}$) is the first "top-left" submatrix of the Cramér-Rao bound on the estimation of the true parameter vector $\bar{\mathbf{\theta}} = [\mathrm{vecs}(\bar{\mathbf{\Sigma}})^T \quad \lambda \quad \eta]^T$ under matched condition (i.e. when the assumed distribution is itself a complex t-distribution) evaluated in [14]. The comparison between the CMML estimator and the CMCRB on one hand and the (matched) CCRB on the other hand allows us to quantify the loss in estimation accuracy due to the mismatch between the true and the assumed models. The true scatter matrix is assumed to be of the form: $[\bar{\mathbf{\Sigma}}]_{i,j} = \rho^{|i-j|}$, where ρ is the one-lag correlation coefficient. To calculate the performance of the CMML estimators, we run 10^4 Monte Carlo trials. The simulation results have been organized as follows:

1) Estimation accuracy as function of the number M of available data vectors (Figs. 1 and 2). The simulation parameters are: ρ=0.8, N=16, λ=3, η=1.

2) Estimation accuracy as function of the shape parameter λ (Figs. 3 and 4). The simulation parameters are: ρ=0.8, N=16, M=10N, η=1.

From simulated analysis on the CMML estimator, it can be noted that the loss in estimation accuracy due to the mismatch is not too high and seems to be always bounded, except for the case of extremely heavy-tailed data, i.e. when λ is close to 0 (see Fig. 3). In particular, when $\lambda\to 0$, the CMCRB rapidly increases while the CCRB is quite independent of the value of the shape parameter. On the other hand, when $\lambda\to+\infty$, i.e. when the t-distribution tends to the Normal one, the CMCRB and the CCRB tend to coincide, as expected. It can be noted from Fig.1 that $\varepsilon_{\hat{\Sigma}_{CMML}}$ can goes below $\varepsilon_{\mathrm{CMCRB}}$. This is probably due to the presence of a slight bias in the CMML estimation of the scatter matrix. The CMML for the estimation of the power σ^2 results to be an efficient estimator w.r.t. the CMCRB.

VII. CONCLUSIONS

In some operative situations, one could be forced to assume a tractable data model instead of the true (and possibly unknown) one. In this paper, we have proved that, in the presence of t-distributed data, the CMML estimator based on the simpler Normal model converges almost surely to the true scatter matrix and to the true data power, so it can be applied for inference problems that require the knowledge of these two quantities, e.g. radar detection problems. Moreover, we have shown that the CMML estimator is an efficient estimator w.r.t. the CMCRB. Future works aim at generalize this results among the CES class and to investigate the impact of the mismatch on the detection performance.

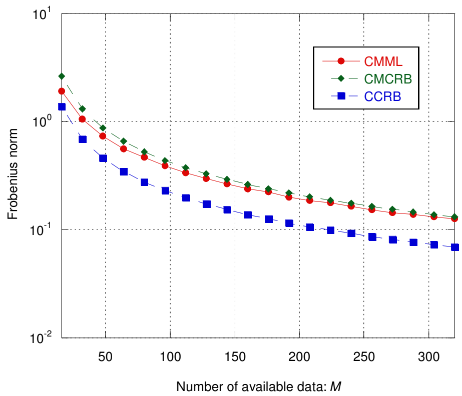

Fig. 1 – $\varepsilon_{\hat{\Sigma}_{CMML}}$, $\varepsilon_{\text{CMCRB}}$ and $\varepsilon_{\text{CCRB}}$ as function of M.

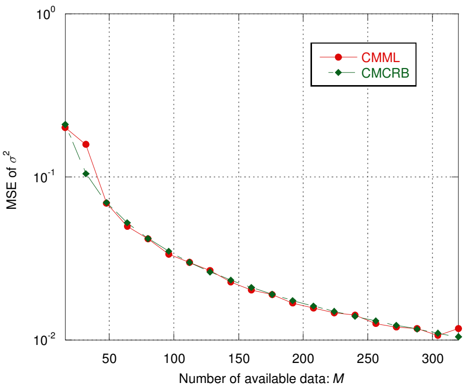

Fig. 2 – MSE of σ^2 and CMCRB as function of M.

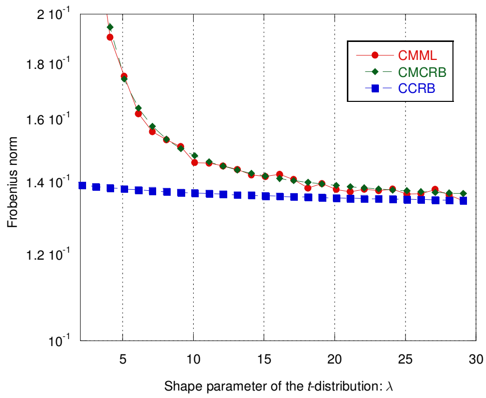

Fig. 3 – $\varepsilon_{\hat{\Sigma}_{CMML}}$, $\varepsilon_{\text{CMCRB}}$ and $\varepsilon_{\text{CCRB}}$ as function of λ.

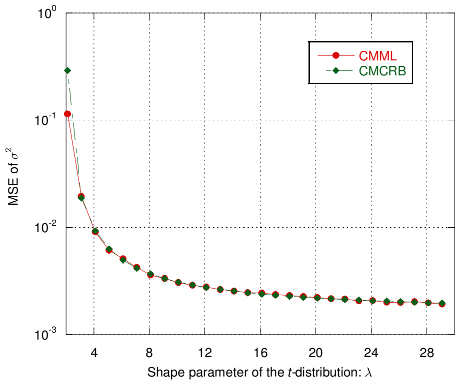

Fig. 4 – MSE of σ^2 and CMCRB as function of λ.

REFERENCES

[1] E. Ollila, D.E. Tyler, V. Koivunen, V.H. Poor, “Complex Elliptically Symmetric Distributions: Survey, New Results and Applications,” *IEEE Trans. on Signal Processing*, vol. 60, no. 11, pp.5597-5625, 2012.

[2] S. Fortunati, F. Gini, M. S. Greco, “The Misspecified Cramér-Rao Bound and its Application to the Scatter Matrix estimation in Complex Elliptically Symmetric distributions,” *IEEE Trans. on Signal Processing*, vol. 64, no. 9, pp. 2387-2399, May 2016.

[3] A.Younsi, M. Greco, F. Gini, and A. Zoubir, “Performance of the adaptive generalised matched subspace constant false alarm rate detector in non-Gaussian noise: An experimental analysis,” *IET Radar Sonar Navig*., vol. 3, no. 3, pp. 195–202, 2009.

[4] H. White, “Maximum likelihood estimation of misspecified models”, *Econometrica* vol. 50, no. 1, pp. 1-25, January 1982.

[5] C.C. Heyde, R. Morton, “On constrained quasi-likelihood estimation,” *Biometrika*, vol. 80, no. 4, pp. 755-61, 1993.

[6] Q. H. Vuong, “Cramér-Rao bounds for misspecified models,” *Working paper 652, Division of the Humanities and Social Sciences, Caltech*, October 1986. Available at: https://www.hss.caltech.edu/content/cramer-rao-bounds-misspecified-models.

[7] C. D. Richmond, L. L. Horowitz, "Parameter Bounds on Estimation Accuracy Under Model Misspecification," *IEEE Trans. on Signal Processing*, vol. 63, no. 9, pp. 2263-2278, May 1, 2015.

[8] C. Ren, M. N. El Korso, J. Galy, E. Chaumette, P. Larzabal, A. Renaux, "Performances bounds under misspecification model for MIMO Radar application", *EUSIPCO 2015*, Nice, France.

[9] S. Fortunati, F. Gini, M. S. Greco, “The Constrained Misspecified Cramér-Rao Bound,” *IEEE Signal Processing Letters*, vol. 23, no. 5, pp. 718-721, May 2016.

[10] P. Stoica, B. C. Ng, “On the Cramer-Rao bound under parametric constraints,” *IEEE Signal Processing Letters*, Vol. 5, no. 7, pp. 177-179, July 1998.

[11] T. J. Moore, R. J. Kozick, B. M. Sadler, "The Constrained Cramér–Rao Bound From the Perspective of Fitting a Model," *IEEE Signal Processing Letters*, vol.14, no.8, pp.564-567, Aug. 2007.

[12] J. R. Magnus, H. Neudecker, “The commutation matrix: some properties and applications,” *The Annals of Statistics*, vol. 7, pp. 381-394, 1979.

[13] M. S. Greco, S. Fortunati, F. Gini, “Maximum likelihood covariance matrix estimation for complex elliptically symmetric distributions under mismatched conditions,” *Signal Processing*, vol. 104, pp. 381-386, November 2014.

[14] M. S. Greco, S. Fortunati, F. Gini, “Naive, robust or fully-adaptive: An estimation problem for CES distributions,” *IEEE SAM* 2014, pp.457-460, June , 22-25, 2014.