EURASIP Journal on Advances
in Signal Processing

RESEARCH

Open Access

CrossMark

Matched, mismatched, and robust scatter matrix estimation and hypothesis testing in complex *t*-distributed data

Stefano Fortunati[1,2]* , Fulvio Gini[1,2] and Maria S. Greco[1,2]

Abstract

Scatter matrix estimation and hypothesis testing are fundamental inference problems in a wide variety of signal processing applications. In this paper, we investigate and compare the matched, mismatched, and robust approaches to solve these problems in the context of the complex elliptically symmetric (CES) distributions. The matched approach is when the estimation and detection algorithms are tailored on the correct data distribution, whereas the mismatched approach refers to the case when the scatter matrix estimator and the decision rule are derived under a model assumption that is not correct. The robust approach aims at providing good estimation and detection performance, even if suboptimal, over a large set of possible data models, irrespective of the actual data distribution. Specifically, due to its central importance in both the statistical and engineering applications, we assume for the input data a complex *t*-distribution. We analyze scatter matrix estimators derived under the three different approaches and compare their mean square error (MSE) with the constrained Cramér-Rao bound (CCRB) and the constrained misspecified Cramér-Rao bound (CMCRB). In addition, the detection performance and false alarm rate (FAR) of the various detection algorithms are compared with that of the clairvoyant optimum detector.

Keywords: Covariance matrix estimation, Complex elliptically symmetric distribution, Detection problem, Constrained Cramér-Rao bound, Misspecified Cramér-Rao bound

1 Introduction

This paper deals with two common inference problems in radar signal processing, namely the estimation of the disturbance covariance matrix and the adaptive detection of a radar target. In addition to the radar detection, the covariance matrix estimation is a fundamental prerequisite for a lot of applications in many different areas: the direction of arrival (DOA) estimation in array processing [1], the principal component analysis (PCA) [2], and the portfolio optimization in finance [3], just to name a few. We put the covariance estimation and the adaptive detection problems in the more general context of the scatter matrix estimation and hypothesis testing in the complex elliptically symmetric (CES) distribution family. CES distributions constitute a wide class of distributions that includes the complex Gaussian,

generalized Gaussian, the *K*-distribution, complex *t*-distribution and all the compound Gaussian distributions as special cases. Due to their flexibility and their capability to model a plethora of different data behavior, they are widely applied in many areas, such as radar, sonar, and communications [4, 5]. A CES distribution is completely characterized by the mean value **γ**, the scatter (or shape) matrix **Σ**, and the density generator *g*. Given a particular CES distribution, its density generator could depends 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 **γ** and **Σ**.

Specifically, because of its generality, several aspects should be taken into account when making inference on the CES class. The first aspect concerns the existence, convergence, and computational complexity of optimal algorithms tailored (matched) to a particular CES distribution at hand. Think for example of the problem of the joint estimation of the mean value, of the scatter matrix,

* Correspondence: stefano.fortunati@iet.unipi.it
[1]Dipartimento di Ingegneria dell'Informazione, University of Pisa, Via G. Caruso 16, Pisa 56122, Italy
[2]CNIT RaSS (Radar and Surveillance System) National Laboratory, Pisa, Italy

Fortunati *et al. EURASIP Journal on Advances in Signal Processing* (2016) 2016:123

Page 2 of 16

and of the extra parameters that characterize the density generator. As pointed out in, e.g., [6, 7], a joint maximum likelihood (ML) estimation of all these unknown quantities would encounter computational difficulties and convergence (or even existence) issues. To overcome this problem, one has to rely on suboptimal, computationally inexpensive and easy to implement estimators [8]. A different alternative could be to assume a simpler model, e.g., a Gaussian distribution, for the data behavior that allows one for an easy derivation of optimal (but generally mismatched) estimators or detection rules [9]. This consideration leads directly to another issue, namely the *robustness to misspecification*. In particular, it would be of interest to know whether the inference methods based on an assumed CES distribution can achieve "good" performance even if the data follow a different and, in general, more involved CES model. Finally, as direct consequence of the previous considerations, this analysis culminates in the possibility to derive and implement *robust* inference algorithms with good performance over the whole class of CES distributions, even if not optimal under any nominal model.

Following the line of the previous discussion, in this paper, we investigate and compare the matched, mismatched, and robust approaches for inference methods in complex *t*-distributed data. We focus on the multivariate complex *t*-distribution, since it has long been recognized by several authors from both the statistical (see, e.g., [6] and the references therein) and the signal processing communities (see, e.g., [10–13]) as a suitable and flexible model able to describe the heavy-tailed behavior of the measurements in many practical applications (e.g., radar detection).

The paper is organized in two parts. In the first part, we investigate the scatter matrix estimation problem. The second part deals with adaptive detection algorithms. In particular, in the first part, we investigate the performance loss in the scatter matrix estimation when the unknown extra parameters of the *t*-distribution are replaced with low computational complexity estimates obtained via the method of moments (MoM). This represents the *matched case*. The mean square error (MSE) of the "matched" estimators is compared with the constrained Cramér-Rao bound (CCRB). Then, we address the *mismatched case* where, following the approach discussed in our recent work [14], the performance of the mismatched ML (MML) scatter matrix estimator derived under Gaussian assumption is evaluated and its MSE compared with the constrained misspecified Cramér-Rao bound (CMRB) [15]. Finally, the *min-max robust* (among the whole CES class) constrained Tyler (C-Tyler) estimator [16] is introduced and its performance compared with the CCRB and the other previously derived estimators.

The second part of the paper focuses on the detection performance of three detection algorithms: the *linear threshold detector* (LTD) [12], i.e., the *matched* generalized likelihood ratio test (GLRT) *detector* for complex *t*-distributed data; Kelly's detector [17], i.e., the *GLRT detector* derived under the misspecified Gaussian distribution; and the *adaptive normalized matched filter* (ANMF), that represents the robust detector among the CES class. The ANMF has been derived and analyzed by many authors under different names (see, e.g., [4, 18–23]). The three detectors are compared in terms of (*i*) constant false alarm rate (CFAR) property with respect to (w.r.t.) the scatter matrix and the extra parameters estimation and (*ii*) receiver operating characteristic (ROC) curves.

The remainder of the paper is organized as follows. In Section 2, a brief review of the main properties of the CES distribution class and of the complex *t*-distribution is provided. In Section 3, the scatter matrix estimation problem is introduced and the application of the matched, mismatched, and robust approaches extensively analyzed. In Section 4, the hypothesis testing problem in complex *t*-distributed data is investigated. Section 5 collects the simulation results, while Section 6 summarizes our conclusions.

2 The CES distribution class, the compound-Gaussian subclass, and the complex *t*-distribution

The aim of this section is to provide a brief overview of the CES distribution class and makes no claim to completeness. For more comprehensive and detailed discussions, we refer the readers to the excellent works [4] and [5].

A complex N-dimensional random vector $\mathbf{x}_m$ is CES distributed, in shorthand notation $\mathbf{x}_m \sim CE_N(\boldsymbol{\gamma}, \boldsymbol{\Sigma}, g)$, if its probability density function (pdf) is of the form

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

where g is the density generator, $c_{N,g}$ 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, a constraint on the scatter matrix needs to be imposed. In the rest of the paper, we choose to impose the following constraint: $\mathrm{tr}(\boldsymbol{\Sigma}) = N$. 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 $\boldsymbol{\Sigma} = N/\mathrm{tr}(\mathbf{M}) \cdot \mathbf{M}$. For some CES distributions, the un-normalized covariance matrix $\mathbf{M}$ does not exist, but the scatter matrix $\boldsymbol{\Sigma}$ is still well-defined. This is the case, for example, for all the CES distributions that do not have finite second-order moments (e.g., the Cauchy distribution) [5]. Based upon the stochastic representation theorem [5], any $\mathbf{x}_m \sim CE_N(\boldsymbol{\gamma}, \boldsymbol{\Sigma}, g)$ with $\mathrm{rank}(\boldsymbol{\Sigma}) = k \leq N$ admits the stochastic representation

Fortunati *et al. EURASIP Journal on Advances in Signal Processing* (2016) 2016:123

Page 3 of 16

$\mathbf{x}_m = _d \boldsymbol{\gamma} + R\mathbf{Pu}$, where the non-negative random variable (r.v.) $R \triangleq \sqrt{Q}$, the so-called modular variate, is a real, non-negative random variable, $\mathbf{u} \sim U(\mathbb{CS}^k)$ is a k-dimensional random vector uniformly distributed on the unit hyper-sphere $\mathbb{CS}^k$ with $k - 1$ topological dimensions such that $\mathbf{u}^H\mathbf{u} = 1$, R and $\mathbf{u}$ are independent, and $\boldsymbol{\Sigma} = \mathbf{PP}^H$ is a factorization of $\boldsymbol{\Sigma}$, where $\mathbf{P}$ is a Nxk matrix and rank($\mathbf{P}$) = k. It is easy to show that the random vector $\mathbf{u}$ is strictly related to a complex normal distribution $CN(\mathbf{0,I})$, where $\mathbf{I}$ defines the identity matrix. In fact, if $\mathbf{w} \sim CN(\mathbf{0,I})$, then $\mathbf{u} = _d\mathbf{w}/\|\mathbf{w}\|$ [5]. Using this property, the stochastic decomposition can be recast as $\mathbf{x}_m = _d \boldsymbol{\gamma} + R\mathbf{Pu} = _d \boldsymbol{\gamma} + (R/\sqrt{Q_w})\mathbf{Pw}$ where $Q_w \triangleq \|\mathbf{w}\|^2 \sim Gam(N, 1)$. Moreover, Q_w is independent of $\mathbf{w}$, and $E\{Q_w\} = N$, and $E\{Q_w^2\} = N(N + 1)$ [5]. In the following, we assume that $\boldsymbol{\Sigma}$ is real and full-rank, i.e., rank($\mathbf{P}$) = rank($\boldsymbol{\Sigma}$) = N. An important remark is in order: 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 $\boldsymbol{\Sigma}$ are related by $\mathbf{M} = \sigma^2\boldsymbol{\Sigma}$.

An important subclass of the CES distributions is the compound-Gaussian (CG) distributions [24]. In particular, a CG-distributed random vector $\mathbf{x}_m \sim CG_N(\boldsymbol{\gamma}, \boldsymbol{\Sigma}, p_\tau)$ admits the following stochastic representation $\mathbf{x}_m = _d\boldsymbol{\gamma} + \sqrt{\tau}\mathbf{Pw} = _d\sqrt{\tau \cdot Q_w}\mathbf{Pu}$, where, as before, $\mathbf{w} \sim CN(\mathbf{0,I})$, $\mathbf{u} \sim U(\mathbb{CS}^N)$, and $Q_w \sim Gam(N, 1)$. Usually, the positive real random variable τ is called the *texture* and the complex Gaussian random vector $\mathbf{n} = _d\mathbf{Pw}$ is called the *speckle*. It can be noted that a CES-distributed random vector belongs to the subclass of the CG-distributed random vector if and only if the square of its modular variate R^2 can be written as a random scaled gamma distribution, i.e., $R^2 = \tau \cdot Q_w$ and then $R^2|\tau \sim Gam(N, \tau)$ [5].

At this point, we can introduce the complex t-distribution. A complex N-dimensional zero-mean ($\boldsymbol{\gamma} = 0$) random vector $\mathbf{x}_m$ is complex t-distributed if its pdf can be expressed as

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

(2)

where λ and η are the shape and scale parameters, respectively. It is easy to show that the pdf in Eq. (2) is of the form given in Eq. (1) where the density generator can be expressed as $g(t) = (\lambda/\eta + t)^{-(N+\lambda)}$ [5]. Moreover, it can be also shown that it admits a CG representation [24]. 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 [12, 14]:

$$\sigma^2 = E_p\{Q_\Sigma\}/N = \lambda/\eta(\lambda-1).$$

(3)

Before passing to discuss the scatter matrix estimation problem in complex t-distributed data, few remarks are needed. In the rest of the paper, we always assume that:

i) The dataset $\mathbf{x} = \{\mathbf{x}_m\}_{m=1}^M$ is composed of M independent and identically distributed (IID) N-dimensional, zero-mean, complex t-distributed random vectors.
ii) The scatter matrix $\boldsymbol{\Sigma}$ is a *real* and *full rank* matrix.

It must be underlined that the second assumption is quite strong and not always verified in radar/sonar applications. It is well-known in fact that, if the power spectral density (PSD) of the disturbance is not symmetric around a central frequency, the autocorrelation function of the complex envelope of the data is complex valued and consequently also the scatter matrix (see, e.g., [25]). However, working with complex matrices would require the use of more sophisticated mathematical tools, i.e., the so-called Wirtinger calculus [26], but this general approach falls outside the scope of the paper. The case of complex scatter matrix will be considered in future works.

3 Scatter matrix estimation

This section deals with the scatter matrix estimation from a set of IID complex t-distributed data. As discussed in the previous section, we investigate three different approaches: the matched, the mismatched, and the robust cases. We also provide the relative performance bounds, i.e., the CCRB and the CMCRB.

3.1 The matched case for complex *t*-distributed data

In this section, we discuss and derive two matched estimators of the scatter matrix $\boldsymbol{\Sigma}$ and of the extra parameters λ and η by assuming to know perfectly the correct data model, i.e., the complex t-distribution. Building upon previous results, we investigate the performance of the following two estimators: (1) the constrained maximum likelihood (CML) estimator of $\boldsymbol{\Sigma}$ which uses the method of moments (MoM) estimates of λ and η and (2) a recursive (suboptimal) estimator of $\boldsymbol{\Sigma}$, λ, and η.

3.1.1 The constrained ML-MoM estimator (CML-MoM)

The CML estimator of the scatter matrix for t-distributed data is given by the solution of the following fixed-point equation [5, 27]:

$$\hat{\boldsymbol{\Sigma}}_{ML} = \frac{N+\lambda}{M}\sum_{m=1}^{M}\frac{\mathbf{x}_m\mathbf{x}_m^H}{\mathbf{x}_m^H\hat{\boldsymbol{\Sigma}}_{ML}^{-1}\mathbf{x}_m + \lambda/\eta}, \quad \text{s.t. } \text{tr}\big(\hat{\boldsymbol{\Sigma}}_{ML}\big) = N. \tag{4}$$

As we can see, Eq. (4) involves the unknown shape and scale parameter of the t-distribution. To estimate them, we use the low computational complexity (but suboptimal) MoM estimators. The MoM method consists of equating the experimental moments with the corresponding theoretical ones in order to obtain an estimate of the unknown parameters of interest. In particular, given a random variable r whose pdf depends on some unknown parameters, one needs to firstly evaluate analytically the moments $m_k \triangleq E\{r^k\}$, i.e., the expected values of powers of the random variable under consideration, and, secondly, equate the obtained expressions (that will depend on the unknown parameters) with the corresponding sample estimates of the moments, i.e., $\hat{m}_k \triangleq \sum_{m=1}^{M} r_m^k$, where $\{r_m\}_{m=1}^{M}$ are M realizations of the random variable r.

In the problem at hand, we need to estimate the shape and scale parameters, λ and η, of the complex t-distribution. In order to do this, we may apply the MoM method by considering the moments of the amplitude of each entry of the data vector $\mathbf{x}_m$, i.e., $r_n \triangleq |[\mathbf{x}_m]_n|$, $n = 1,...,N$. To evaluate the moments of the amplitude r_n, we can exploit the decomposition given in Theorem 5 of [5] and the fact that the t-distribution is a CG distribution. In particular, we have that the amplitude

$$r_n \triangleq |x_{m,n}| = R|[\mathbf{P}\mathbf{u}_m]_n| = \frac{R}{\sqrt{Q_w}}|[\mathbf{P}\mathbf{w}_m]_n|$$
$$= \sqrt{\tau}|[\mathbf{n}_m]_n| \triangleq \sqrt{\tau}|n_{m,n}| \tag{5}$$

is distributed according to [5] as

$$p_{r_n}(r) = 2\eta r\Big(1 + \frac{\eta}{\lambda}r^2\Big)^{-(\lambda+1)}u(r), \tag{6}$$

where $u(\cdot)$ is the unit step function. It is easy to verify through direct calculation on the pdf in (6) that the moments of order k of r_n are given by

$$m_k \triangleq E\{r_n^k\} = \Big(\frac{\lambda}{\eta}\Big)^{k/2}\frac{\Gamma(k/2+1)\Gamma(\lambda-k/2)}{\Gamma(\lambda)}, \quad \lambda > \frac{k}{2}. \tag{7}$$

Finally, by applying the classical MoM approach using the fourth-order and the second-order moments of r_n, the parameters λ and η can be estimated as follows [28]:

$$\hat{\lambda}_{\text{MoM}} = 2\frac{\hat{m}_4/\hat{m}_2^2-1}{\hat{m}_4/\hat{m}_2^2-2}, \quad \hat{\eta}_{\text{MoM}} = \frac{\hat{\lambda}_{\text{MoM}}}{\hat{m}_2\big(\hat{\lambda}_{\text{MoM}}-1\big)}, \tag{8}$$

where

$$\hat{\mu} = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}x_{m,n}, \tag{9}$$

$$\hat{m}_2 = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}|x_{m,n}-\hat{\mu}|^2, \text{ and}$$

$$\hat{m}_4 = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}|x_{m,n}-\hat{\mu}|^4$$

are the sample estimates of the moments. Due to existence issues of the fourth-order moment, we constrain the estimator of λ to be larger than 2, i.e., $\hat{\lambda}_{\text{MoM}} > 2$. Finally, the following iterative approach [27] can be used to solve Eq. (4):

$$\begin{cases} \hat{\boldsymbol{\Sigma}}_{\text{CML}}^{(0)} = \mathbf{I} \\ \mathbf{S}_{\text{CML}}^{(k+1)} = \sum_{m=1}^{M}\frac{\mathbf{x}_m\mathbf{x}_m^H}{\mathbf{x}_m^H\big(\hat{\boldsymbol{\Sigma}}_{\text{CML}}^{(k)}\big)^{-1}\mathbf{x}_m + \hat{\lambda}_{\text{MoM}}/\hat{\eta}_{\text{MoM}}} \\ \hat{\boldsymbol{\Sigma}}_{\text{CML}}^{(k+1)} = N\mathbf{S}_{\text{CML}}^{(k+1)}/\text{tr}\big(\mathbf{S}_{\text{CML}}^{(k+1)}\big) \end{cases} \tag{10}$$

It can be noted that the constraint on the trace of $\hat{\boldsymbol{\Sigma}}_{\text{CML}}^{(k)}$ has to be imposed at each iteration.

3.1.2 The constrained recursive ML-WMoM estimator (CML-WMoM)

In this section, we propose an improvement of the CML-MoM estimator of Eq. (10). It should be noted first that the moment sample estimators $\hat{m}_2$ and $\hat{m}_4$ have been derived in (9) under the assumption that the $M \cdot N$ data samples are IID, although the entries $\{x_{\overline{m},n}\}_{n=1}^{N}$ with $\overline{m}\in\{1,...,M\}$ are correlated. Applying directly this method to a set of t-distributed random vectors leads to a suboptimal approach. However, from Eq. (5), it is easy to show that each data vector can be whitened (W) in order to have uncorrelated entries, being $\mathbf{n}$ from Eq. (5) a Gaussian random vector. Since both the scatter matrix and the shape and scale parameters are unknown, we rely on a recursive procedure to estimate them jointly [8]:

1. Initialization ($k = 0$)

$$\hat{\boldsymbol{\Sigma}}_W^{(0)} = \mathbf{I}, \tag{11}$$

2. $(k + 1)$th iteration (for $k = 0,...,K$)

$$\tilde{\mathbf{x}}_m = \big(\hat{\boldsymbol{\Sigma}}_W^{(k)}\big)^{-1/2}\mathbf{x}_m, \tag{12}$$

Fortunati *et al. EURASIP Journal on Advances in Signal Processing* (2016) 2016:123

Page 5 of 16

$$\begin{cases} \hat{\mu}^{(k)} = \dfrac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\tilde{x}_{m,n} \\[4pt] \mu_2^{(k)} = \dfrac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left|\tilde{x}_{m,n}-\hat{\mu}^{(k)}\right|^2 \xRightarrow{\text{eq. (9)}} \lambda_{WMoM}^{(k)} > 2, \eta_{WMoM}^{(k)}, \\[4pt] \mu_4^{(k)} = \dfrac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left|\tilde{x}_{m,n}-\hat{\mu}^{(k)}\right|^4 \end{cases}$$

$$(13)$$

$$\begin{cases} \mathbf{S}_W^{(k+1)} = \displaystyle\sum_{m=1}^{M}\dfrac{\mathbf{x}_m\mathbf{x}_m^H}{\mathbf{x}_m^H\left(\hat{\mathbf{\Sigma}}_W^{(k)}\right)^{-1}\mathbf{x}_m + \hat{\lambda}_{WMoM^{(k)}}/\hat{\eta}_{WMoM}^{(k)}}. \\[8pt] \hat{\mathbf{\Sigma}}_W^{(k+1)} = N\mathbf{S}_W^{(k+1)}/\mathrm{tr}\left(\mathbf{S}_W^{(k+1)}\right) \end{cases}$$

$$(14)$$

Even if based on more accurate considerations about the marginal pdf of the entries of $\mathbf{x}_m$, the proposed recursive constrained ML-whitened MoM (CML-WMoM) estimator is itself a suboptimal algorithm. Moreover, the convergence of the recursive procedure is not guaranteed.

3.1.3 The constrained Cramér-Rao bound (CCRB)

This section provides a concise review on the derivation for the constrained Cramér-Rao bound (CCRB) for the estimation of $\mathbf{\theta} \triangleq \left[\mathrm{vecs}(\mathbf{\Sigma})^T \quad \lambda \quad \eta\right]^T$ in complex t-distributed data where the vecs-operator is the "symmetric" counterpart of the standard vec-operator that maps a symmetric $N \times N$ matrix $\mathbf{\Sigma}$ in an l-dimensional vector (where $l = N(N+1)/2$ whose entries are the elements of the lower (or upper) triangular sub-matrix of $\mathbf{\Sigma}$. Following our previous results presented in [8] and [29], we have that the unconstrained Fisher information matrix (FIM) is given by

$$\mathbf{F}_{\mathbf{\theta}} = \mathbf{T}_2^T \begin{bmatrix} \mathbf{F}_{\mathbf{\Sigma}} & \mathbf{F}_c \\ \mathbf{F}_c^T & \mathbf{F}_{\lambda,\eta} \end{bmatrix} \mathbf{T}_2, \tag{15}$$

where

$$\mathbf{F}_{\mathbf{\Sigma}} = -\frac{1}{N+\lambda+1}\mathrm{vec}(\mathbf{\Sigma}^{-1})\mathrm{vec}(\mathbf{\Sigma}^{-1})^T + \frac{N+\lambda}{N+\lambda+1}\mathbf{\Sigma}^{-1}\otimes\mathbf{\Sigma}^{-1}, \tag{16}$$

$$\left[\mathbf{F}_{\lambda,\eta}\right]_{1,1} = \sum_{k=0}^{N-1}\frac{1}{(\lambda+k)^2} + \frac{(\lambda+1)(N+\lambda)}{\lambda(N+\lambda+1)} - \frac{2N}{\lambda(\lambda+N)} - 1, \tag{17}$$

$$\left[\mathbf{F}_{\lambda,\eta}\right]_{2,2} = \frac{N\lambda}{\eta^2(N+\lambda+1)}, \tag{18}$$

$$\left[\mathbf{F}_{\lambda,\eta}\right]_{1,2} = \left[\mathbf{F}_{\lambda,\eta}\right]_{2,1} = \frac{\lambda N}{\eta\lambda(N+\lambda)(N+\lambda+1)}, \tag{19}$$

$$\mathbf{F}_c = \left[-\frac{1}{(N+\lambda)(N+\lambda+1)}\mathrm{vec}(\mathbf{\Sigma}^{-1}) \quad -\frac{\lambda}{\eta(N+\lambda+1)}\mathrm{vec}(\mathbf{\Sigma}^{-1})\right], \tag{20}$$

$$\mathbf{T}_i = \begin{bmatrix} \mathbf{D}_N & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_i \end{bmatrix}, \tag{21}$$

where $\mathbf{I}_i$ is the identity matrix of dimension $i \times i$ and $\mathbf{D}_N$ is the so-called duplication matrix of order N [30]. The duplication matrix is implicitly defined as the unique $N^2 \times l$ matrix that satisfies the following equality: $\mathbf{D}_N\mathrm{vecs}(\mathbf{A}) = \mathrm{vec}(\mathbf{A})$ for any symmetric matrix $\mathbf{A}$.

As discussed before, due to the ambiguity between power and scatter matrix $\mathbf{\Sigma}$, the parameters λ and η are identifiable (i.e., they can be estimated from the data) only by putting a constraint on $\mathbf{\Sigma}$, e.g., $\mathrm{tr}(\mathbf{\Sigma}) = N$. For this reason, a constrained version of the Cramér-Rao bound needs to be derived [31, 32]. To this end, the continuously differentiable constraint $\mathrm{tr}(\mathbf{\Sigma}) = N$ can be rewritten as

$$f(\mathbf{\theta}) = \sum_{i\in I}\mathrm{vecs}(\mathbf{\Sigma})_i - N = 0, \tag{22}$$

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

$$I = \left\{i\middle| i = 1 + N(j-1) - \frac{(j-1)(j-2)}{2}, j = 1, ..., N\right\}. \tag{23}$$

Following [32], we define the $(l+2)$-dimensional gradient vector as

$$\nabla f(\mathbf{\theta}) = \frac{\partial f(\mathbf{\theta})}{\partial \mathbf{\theta}^T} = \left[\frac{\partial\sum_{i\in I}\mathrm{vecs}(\mathbf{\Sigma})_i}{\partial\mathrm{vecs}(\mathbf{\Sigma})^T} \quad 0 \quad 0\right] = \left[\mathbf{1}_I^T \quad 0 \quad 0\right], \tag{24}$$

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

$$[\mathbf{1}_I]_i = \begin{cases} 1 & i\in I \\ 0 & \text{otherwise} \end{cases}. \tag{25}$$

The gradient $\nabla f(\mathbf{\theta})$ has clearly full row rank, and hence, there exists a matrix $\mathbf{U} \in \mathbb{R}^{(l+2)\times(l+1)}$ 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 $\mathbf{U}^T\mathbf{U} = \mathbf{I}$. The matrix $\mathbf{U}$ can be obtained numerically by evaluating, e.g., using the singular value decomposition (SVD), the $l+1$ orthonormal eigenvectors associated to the zero eigenvalue of $\nabla f(\mathbf{\theta})$. Finally, the CCRB on the estimation of $\mathbf{\theta}$ can be expressed as (Theorem 1 in [32])

$$\mathrm{CCRB}(\mathbf{\theta}) = \mathbf{U}\left(\mathbf{U}^T\mathbf{F}_{\mathbf{\theta}}\mathbf{U}\right)^{-1}\mathbf{U}^T. \tag{26}$$

3.2 The mismatched case

In the matched case, the true data model and the model assumed to derive a joint estimator of the scatter matrix

Fortunati *et al. EURASIP Journal on Advances in Signal Processing* (2016) 2016:123

Page 6 of 16

and of the shape and scale parameters are the same; that is, the model is correctly specified. However, a certain amount of mismatch is often inevitable in practice. Among others, the model mismatch can be due to an imperfect knowledge of the true data model or to the need to fulfill some operative constraints on the estimation algorithm (processing time, simple hardware implementation, and so on). In other words, even if the true but involved model is known, in order to derive a simple (mismatched) estimator for practical exploitation, one could decide to assume a simpler model, e.g., a Gaussian distribution. In our recent work [14], we investigated the behavior of the ML estimator of the scatter matrix in CES-distributed data under mismatched conditions, i.e., the mismatched ML (MML) estimator. Moreover, the existence of a lower bound on the error covariance matrix of a certain class of mismatched estimators has been investigated as well (see also [33]). In particular, it has been shown that the asymptotic distribution of the MML estimator is a Gaussian one, whose mean value is the minimizer (also called *pseudo-true parameter vector*) of the Kullback-Leibler (KL) divergence between the true and the assumed data distributions and the covariance matrix is given by the so-called Huber "sandwich" matrix. For brevity, we refer the reader to the recent papers [14] and [33] and references therein for a more comprehensive and insightful review of these topics.

In this paper, we consider the following mismatched scenario: we assume a complex Gaussian 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 IID and each one is distributed according to a complex normal multivariate pdf, which also belongs to the CES family:

$$f_X(\mathbf{x}_m; \boldsymbol{\theta}) \triangleq f_X(\mathbf{x}_m; \boldsymbol{\Sigma}, \sigma^2) = \frac{1}{(\pi\sigma^2)^N |\boldsymbol{\Sigma}|} \exp\left(\frac{-\mathbf{x}_m^H \boldsymbol{\Sigma}^{-1} \mathbf{x}_m}{\sigma^2}\right). \tag{27}$$

The covariance matrix is $\mathbf{M} = E\{\mathbf{x}_m \mathbf{x}_m^H\} = \sigma^2 \boldsymbol{\Sigma}$, where $\mathrm{tr}(\boldsymbol{\Sigma}) = N$ and σ^2 are the power. Hence, the parameter vector to be estimated can be expressed as $\boldsymbol{\theta} = \begin{bmatrix} \mathrm{vecs}(\boldsymbol{\Sigma})^T & \sigma^2 \end{bmatrix}^T \in \Theta$. However, the true data are distributed according to the complex t-distribution $p_X(\mathbf{x}_m; \overline{\boldsymbol{\theta}}) \triangleq p_X(\mathbf{x}_m; \overline{\boldsymbol{\Sigma}}, \lambda, \eta)$ of Eq. (2), where $\overline{\boldsymbol{\theta}} = \begin{bmatrix} \mathrm{vecs}(\overline{\boldsymbol{\Sigma}})^T & \lambda & \eta \end{bmatrix}^T \in T$ is the true parameter vector and $\overline{\boldsymbol{\Sigma}}$ is the true scatter matrix that could be different to the scatter matrix $\boldsymbol{\Sigma}$ of the assumed Gaussian distribution. A point need to be clearly highlighted: in the mismatched case, the parameter space Θ that parameterizes the assumed distribution and the (possibly inaccessible and unknown) parameter space T that parameterizes the true distribution may be different. In the case at hand, for example, $T \subset \mathbb{R}^l \times (0, \infty) \times (0, \infty)$ while $\Theta \subset$

$\mathbb{R}^l \times (0, \infty)$ where $\times$ indicates the Cartesian product and $l = N(N+1)/2$ as before. Moreover, the constraint on the trace of the scatter matrix limits both the true and assumed parameter vector to belong to two lower dimensional smooth manifolds $\widetilde{T} = \{\overline{\boldsymbol{\theta}} \in T | \mathrm{tr}(\overline{\boldsymbol{\Sigma}}) = N\}$ and $\widetilde{\Theta} = \{\boldsymbol{\theta} \in \Theta | \mathrm{tr}(\boldsymbol{\Sigma}) = N\}$, respectively.

3.2.1 The constrained MML (CMML) estimator
In order to obtain an estimation of $\boldsymbol{\theta}$, we apply the ML method, so what we get under mismatched conditions is the so-called MML estimator [34, 35]:

$$\hat{\boldsymbol{\theta}}_{\mathrm{MML}}(\mathbf{x}) \triangleq \underset{\theta \in \Theta}{argmax} \ln f_X(\mathbf{x}; \theta) = \underset{\theta \in \Theta}{argmax} \sum_{m=1}^{M} \ln f_X(\mathbf{x}_m; \theta), \tag{28}$$

where $\mathbf{x}_m \sim p_X(\mathbf{x}_m; \overline{\boldsymbol{\theta}})$ and $\mathrm{tr}(\boldsymbol{\Sigma}) = \mathrm{tr}(\overline{\boldsymbol{\Sigma}}) = N$. It can be shown (see [14, 33–35] and the references therein) that the MML estimator converges *almost surely* (*a.s.*) to $\boldsymbol{\theta}_0$, i.e., the vector that minimizes the KL divergence between $p_X(\mathbf{x}_m; \overline{\boldsymbol{\theta}})$ and $f_X(\mathbf{x}_m; \boldsymbol{\theta})$:

$$\hat{\boldsymbol{\theta}}_{\mathrm{MML}}(\mathbf{x}) \underset{M \to \infty}{\overset{a.s.}{\to}} \boldsymbol{\theta}_0, \tag{29}$$

$$\boldsymbol{\theta}_0 = \underset{\theta \in \Theta}{argmin} \{D(p \| f_\theta)\} = \underset{\theta \in \Theta}{argmin} \{-E_p\{\ln f_X(\mathbf{x}_m; \theta)\}\}, \tag{30}$$

where

$$D(p \| f_{\boldsymbol{\theta}}) \triangleq E_p\left\{\ln\left(\frac{p_X(\mathbf{x}_m; \overline{\boldsymbol{\theta}})}{f_X(\mathbf{x}; \boldsymbol{\theta})}\right)\right\} = \int \ln\left(\frac{p_X(\mathbf{x}_m; \overline{\boldsymbol{\theta}})}{f_X(\mathbf{x}; \boldsymbol{\theta})}\right) p_X(\mathbf{x}_m; \overline{\boldsymbol{\theta}}) d\mathbf{x}. \tag{31}$$

The assumption of a complex normal model is motivated by the fact that the MML estimator for the joint estimation of the scatter $\boldsymbol{\Sigma}$ matrix and σ^2 can be easily derived. The log-likelihood function in fact can be expressed as

$$L(\boldsymbol{\theta}) = \sum_{m=1}^{M} \ln f_X(\mathbf{x}_m; \boldsymbol{\theta}) = -NM\ln\sigma^2 - M\ln|\boldsymbol{\Sigma}| - \sum_{m=1}^{M} \mathbf{x}_m^H \boldsymbol{\Sigma}^{-1} \mathbf{x}_m / \sigma^2. \tag{32}$$

Then, the MML estimator can be obtained by maximizing $L(\boldsymbol{\theta})$ subject to the linear constraint $\mathrm{tr}(\boldsymbol{\Sigma}) = N$. To do this, we do not rely on the Lagrange multiplier method, but we follow a different, yet equivalent (at least asymptotically), procedure [36]: we first derive the unconstrained MML estimator and then we project it on the lower dimensional manifold $\overline{\Theta}$ by imposing the constraint. Specifically, the MML estimator is the solution of the following problem:

Fortunati *et al. EURASIP Journal on Advances in Signal Processing* (2016) 2016:123

Page 7 of 16

$$\begin{cases} \dfrac{\partial L(\boldsymbol{\theta})}{\partial \sigma^2} = -\dfrac{NM}{\sigma^2} + \dfrac{1}{\sigma^4}\sum_{m=1}^{M}\mathbf{x}_m^H\boldsymbol{\Sigma}^{-1}\mathbf{x}_m = 0 \\[2mm] \dfrac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\Sigma}} = -M\boldsymbol{\Sigma}^{-1} + \dfrac{\boldsymbol{\Sigma}^{-1}}{\sigma^2}\sum_{m=1}^{M}\mathbf{x}_m\mathbf{x}_m^H\boldsymbol{\Sigma}^{-1} = \mathbf{0} \\[2mm] \text{s.t.}\quad \text{tr}(\boldsymbol{\Sigma}) = N \end{cases} \tag{33}$$

Before solving (33), a remark is in order. In the second equation, the derivative of log-likelihood function $L(\boldsymbol{\theta})$ is taken with respect to all N^2 elements of the scatter matrix $\boldsymbol{\Sigma}$. Since $\boldsymbol{\Sigma}$ is a symmetric matrix, some derivatives are redundant. On the other hand, this approach has the advantage to allow for a simple and compact calculation of the matrix derivative. A more formal approach is discussed in [30], and it involves the use of the duplication matrix introduced in Section 3.1.3. However, since our approach and the one proposed in [30] lead to the same result, we chose to exploit the simplest of the methods. Solving (33), we have

$$\begin{cases} \hat{\sigma}^2_{\text{MML}} = \dfrac{1}{NM}\sum_{m=1}^{M}\mathbf{x}_m^H\hat{\boldsymbol{\Sigma}}_{\text{MML}}^{-1}\mathbf{x}_m \\[3mm] \hat{\boldsymbol{\Sigma}}_{\text{MML}} = \dfrac{N}{\sum_{m=1}^{M}\mathbf{x}_m^H\hat{\boldsymbol{\Sigma}}_{\text{MML}}^{-1}\mathbf{x}_m}\sum_{m=1}^{M}\mathbf{x}_m\mathbf{x}_m^H \\[3mm] \text{s.t.}\quad \text{tr}\left(\hat{\boldsymbol{\Sigma}}_{\text{MML}}\right) = N \end{cases}$$

$$\tag{34}$$

Hence, imposing the constraint, we get the constrained MML (CMML) estimators of σ^2 and $\boldsymbol{\Sigma}$:

$$\begin{cases} \hat{\sigma}^2_{\text{CMML}} = \dfrac{1}{NM}\sum_{m=1}^{M}\mathbf{x}_m^H\hat{\boldsymbol{\Sigma}}_{\text{CMML}}^{-1}\mathbf{x}_m \\[3mm] \hat{\boldsymbol{\Sigma}}_{\text{CMML}} = \dfrac{N}{\sum_{m=1}^{M}\mathbf{x}_m^H\mathbf{x}_m}\sum_{m=1}^{M}\mathbf{x}_m\mathbf{x}_m^H \end{cases} \tag{35}$$

Now, we need to find the vector $\boldsymbol{\theta}_0 = \left[\text{vecs}(\boldsymbol{\Sigma}_0)^T\ \sigma_0^2\right]^T$ that minimizes the KL divergence between $p_X\left(\mathbf{x}_m;\overline{\boldsymbol{\theta}}\right)$ and $f_X(\mathbf{x}_m;\boldsymbol{\theta})$. This vector is the convergence point of the MML estimator in Eq. (35). To this end, we have to solve the following system:

$$\begin{cases} \dfrac{\partial D(p\|f_{\boldsymbol{\theta}})}{\partial \sigma^2} = -E_p\left\{\dfrac{\partial \ln f_X(\mathbf{x}_m;\boldsymbol{\Sigma},\sigma^2)}{\partial \sigma^2}\right\} = 0 \\[3mm] \dfrac{\partial D(p\|f_{\boldsymbol{\theta}})}{\partial \boldsymbol{\Sigma}} = -E_p\left\{\dfrac{\partial \ln f_X(\mathbf{x}_m;\boldsymbol{\Sigma},\sigma^2)}{\partial \boldsymbol{\Sigma}}\right\} = \mathbf{0} \end{cases} \tag{36}$$

The first equation immediately provides

$$\dfrac{\partial D(p\|f_{\boldsymbol{\theta}})}{\partial \sigma^2} = E_p\left\{\dfrac{\partial}{\partial \sigma^2}\left(N\ln\sigma^2 + \dfrac{\mathbf{x}_m^H\boldsymbol{\Sigma}^{-1}\mathbf{x}_m}{\sigma^2}\right)\right\} = \dfrac{N}{\sigma^2} - \dfrac{E_p\{Q_{\boldsymbol{\Sigma}}\}}{\sigma^4} = 0, \tag{37}$$

where $Q_{\boldsymbol{\Sigma}}\triangleq\mathbf{x}_m^H\boldsymbol{\Sigma}^{-1}\mathbf{x}_m$. By solving (37), we get $\sigma_0^2 = \dfrac{E_p\{Q_{\boldsymbol{\Sigma}}\}}{N}$.

The derivative of the KL divergence with respect to $\boldsymbol{\Sigma}$ is instead given by [14]

$$\dfrac{\partial D(p\|f_{\boldsymbol{\theta}})}{\partial \boldsymbol{\Sigma}} = \boldsymbol{\Sigma}^{-1} - \dfrac{E\{Q_{\boldsymbol{\Sigma}}\}}{N\sigma^2}\boldsymbol{\Sigma}^{-1}\overline{\boldsymbol{\Sigma}}\boldsymbol{\Sigma}^{-1} = \mathbf{0},\ \ \text{tr}(\boldsymbol{\Sigma}) = \text{tr}(\overline{\boldsymbol{\Sigma}}) = N \tag{38}$$

whose solution is $\boldsymbol{\Sigma}_0 = \dfrac{E\{Q_{\boldsymbol{\Sigma}_0}\}}{N\sigma^2}\overline{\boldsymbol{\Sigma}}$. Putting together the two solutions, we finally get

$$\sigma_0^2 = \dfrac{E\{Q_{\boldsymbol{\Sigma}_0}\}}{N}\ \text{ and }\ \boldsymbol{\Sigma}_0 = \dfrac{E\{Q_{\boldsymbol{\Sigma}_0}\}}{N\sigma_0^2}\overline{\boldsymbol{\Sigma}} = \overline{\boldsymbol{\Sigma}}, \text{where } \text{tr}(\boldsymbol{\Sigma}_0) = \text{tr}(\overline{\boldsymbol{\Sigma}}) = N,$$

$$\tag{39}$$

and

$$\sigma_0^2 = \dfrac{E\{Q_{\boldsymbol{\Sigma}_0}\}}{N} = \dfrac{E\{Q_{\overline{\boldsymbol{\Sigma}}}\}}{N} = \dfrac{E_p\left\{\mathbf{x}_m^H\overline{\boldsymbol{\Sigma}}^{-1}\mathbf{x}_m\right\}}{N} = \overline{\sigma}^2 = \dfrac{\lambda}{\eta(\lambda-1)}, \tag{40}$$

where $\overline{\sigma}^2$ is the true statistical power of the data. Equations (39) and (40) show that the MML estimator converges *a.s.* to the true parameter vector $\hat{\boldsymbol{\theta}}_{\text{CMML}}(\mathbf{x}) \underset{M\to\infty}{\overset{a.s.}{\to}} \boldsymbol{\theta}_0 = \left[\text{vecs}(\overline{\boldsymbol{\Sigma}})^T\ \overline{\sigma}^2\right]^T$, i.e.,

$$\hat{\sigma}^2_{\text{CMML}}(\mathbf{x}) \underset{M\to\infty}{\overset{a.s.}{\to}} \overline{\sigma}^2 = \lambda/\eta(\lambda-1),\quad \hat{\boldsymbol{\Sigma}}_{\text{CMML}}(\mathbf{x}) \underset{M\to\infty}{\overset{a.s.}{\to}} \overline{\boldsymbol{\Sigma}}, \tag{41}$$

so it provides consistent estimates for both the scatter matrix and the power of the true data model. From a practical point of view, this means that we can use the simpler mismatched estimator based on the Gaussian model assumption to estimate the scatter matrix and the average power of a set of complex *t*-distributed data since it converges to the true required quantities. The analysis of the performance loss of the mismatched estimator in Eq. (35) is reported in the next section.

3.2.2 The constrained misspecified Cramér-Rao bound (CMCRB)

In the classical matched estimation framework, the performance of any unbiased estimator can be assessed by comparing its error covariance matrix with the Cramér-Rao bound. It is natural to ask if it is possible to establish a lover bound on the estimation performance also in the mismatched estimation framework. In his seminal working paper [37], Vuong proposed the misspecified Cramér-Rao bound (MCRB) and showed that it represents a lower bound on the error covariance matrix of any unbiased (in a proper sense) estimator of a deterministic parameter vector under misspecification of the true data model. Recently, the MCRB has been rediscovered and deeply investigated in [14] and [33]. In particular, in [14], the MCRB on the estimation 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 the result for the case of

Fortunati *et al. EURASIP Journal on Advances in Signal Processing* (2016) 2016:123

Page 8 of 16

unknown power, i.e., when the power σ^2 and the scatter matrix $\boldsymbol{\Sigma}$ are unknown and jointly estimated. In order to do this, we exploit our recent derivation of the constrained MCRB (CMCRB) [15]. In particular, in [15], it was shown that the general expression for the CMCRB, for MS-unbiased and consistent estimators (for the definition of *MS*-unbiasedness, we refer the reader to [14, 15] and [37]), is given by

$$\text{MCRB}(\boldsymbol{\theta}_0) = M^{-1}\mathbf{U}\left(\mathbf{U}^T\mathbf{A}_{\boldsymbol{\theta}_0}\mathbf{U}\right)^{-1}\mathbf{U}^T\mathbf{B}_{\boldsymbol{\theta}_0}\mathbf{U}\left(\mathbf{U}^T\mathbf{A}_{\boldsymbol{\theta}_0}\mathbf{U}\right)^{-1}\mathbf{U}^T, \quad \boldsymbol{\theta}_0 \in \overline{\Theta},$$
$$(42)$$

where the entries of the matrices $\mathbf{A}_{\boldsymbol{\theta}_0}$ and $\mathbf{B}_{\boldsymbol{\theta}_0}$ are defined in [14, 15, 33, 37] as

$$[\mathbf{A}_{\boldsymbol{\theta}_0}]_{ij} \triangleq \left[E_p\left\{\nabla_{\boldsymbol{\theta}_0}\nabla_{\boldsymbol{\theta}_0}^T \ln f_X(\mathbf{x}_m;\theta_0)\right\}\right]_{ij} = E_p\left\{\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ln f_X(\mathbf{x}_m;\theta)\Big|_{\theta=\theta_0}\right\},$$
$$(43)$$

$$[\mathbf{B}_{\boldsymbol{\theta}_0}]_{ij} \triangleq \left[E_p\left\{\nabla_{\boldsymbol{\theta}_0}\ln f_X(\mathbf{x}_m;\theta_0)\nabla_{\boldsymbol{\theta}_0}^T\ln f_X(\mathbf{x}_m;\theta_0)\right\}\right]_{ij}$$
$$= E_p\left\{\frac{\partial\ln f_X(\mathbf{x}_m;\theta)}{\partial\theta_i}\Big|_{\theta=\theta_0}\cdot\frac{\partial\ln f_X(\mathbf{x}_m;\theta)}{\partial\theta_j}\Big|_{\theta=\theta_0}\right\},$$
$$(44)$$

and where the columns of matrix $\mathbf{U}$ is defined as in Section 3.1.3 as an orthonormal basis for the null space of full-rank Jacobian matrix of the constraints. In the following, we specialize this general expression for the case study at hand.

Evaluation of the matrix $\mathbf{A}_{\boldsymbol{\theta}_0}$. Matrix $\mathbf{A}_{\boldsymbol{\theta}_0}$ can be decomposed in the following blocks:

$$\mathbf{A}_{\boldsymbol{\theta}_0} = \mathbf{T}_1^T\begin{bmatrix}\mathbf{A}_{\overline{\Sigma}} & \mathbf{A}_c \\ \mathbf{A}_c^T & A_{\overline{\sigma}^2}\end{bmatrix}\mathbf{T}_1.$$
$$(45)$$

where $\mathbf{T}_i$ has been defined in Eq. (21). Following the procedure in [38], we have

$$\mathbf{A}_{\overline{\Sigma}} = -\overline{\Sigma}^{-1}\otimes\overline{\Sigma}^{-1}.$$
$$(46)$$

$$A_{\overline{\sigma}^2} = E_p\left\{\frac{\partial^2}{\partial^2\overline{\sigma}^2}\ln f_X(\mathbf{x}_m;\theta_0)\right\} = E_p\left\{\frac{\partial^2}{\partial^2\overline{\sigma}^2}\left(-N\ln\overline{\sigma}^2-\frac{Q_{\overline{\Sigma}}}{\overline{\sigma}^2}\right)\right\}$$
$$= \frac{N}{\overline{\sigma}^4}-2\frac{E_p\{Q_{\overline{\Sigma}}\}}{\overline{\sigma}^6} = -\frac{N}{\overline{\sigma}^4},$$
$$(47)$$

$$[\mathbf{A}_c]_{i,1} = E_p\left\{\frac{\partial^2\ln f_X(\mathbf{x}_m;\theta_0)}{\partial\overline{\sigma}^2\partial\text{vec}(\overline{\Sigma})_i}\right\} = -\frac{1}{\overline{\sigma}^4}E_p\left\{\mathbf{x}_m^H\overline{\Sigma}^{-1}\mathbf{A}_i\overline{\Sigma}^{-1}\mathbf{x}_m\right\}$$
$$= -\frac{1}{\overline{\sigma}^2}\text{tr}\left(\overline{\Sigma}^{-1}\mathbf{A}_i\right) = -\frac{1}{\overline{\sigma}^2}\text{vec}\left(\overline{\Sigma}^{-1}\right)^T\text{vec}(\mathbf{A}_i)$$
$$(48)$$

where $\mathbf{A}_i = \partial\Sigma/\partial\theta_i$ is a symmetric 0–1 matrix. From (48), we get $\mathbf{A}_c = -\frac{1}{\overline{\sigma}^2}\text{vec}\left(\overline{\Sigma}^{-1}\right)$.

Evaluation of the matrix $\mathbf{B}_{\boldsymbol{\theta}_0}$. Matrix $\mathbf{B}_{\boldsymbol{\theta}_0}$ can be decomposed in the following blocks:

$$\mathbf{B}_{\boldsymbol{\theta}_0} = \mathbf{T}_1^T\begin{bmatrix}\mathbf{B}_{\overline{\Sigma}} & \mathbf{B}_c \\ \mathbf{B}_c^T & B_{\overline{\sigma}^2}\end{bmatrix}\mathbf{T}_1.$$
$$(49)$$

As before, following the procedure in [38], we get

$$\mathbf{B}_{\overline{\Sigma}} = \frac{1}{\lambda-2}\text{vec}\left(\overline{\Sigma}^{-1}\right)\text{vec}\left(\overline{\Sigma}^{-1}\right)^T + \frac{\lambda-1}{\lambda-2}\overline{\Sigma}^{-1}\otimes\overline{\Sigma}^{-1}.$$
$$(50)$$

$$B_{\overline{\sigma}^2} = E_p\left\{\left[\frac{\partial\ln f_X(\mathbf{x}_m;\theta_0)}{\partial\overline{\sigma}^2}\right]^2\right\} = E_p\left\{\left[\frac{\partial}{\partial\overline{\sigma}^2}\left(-N\ln\overline{\sigma}^2-\frac{Q_{\overline{\Sigma}}}{\overline{\sigma}^2}\right)\right]^2\right\}$$
$$= \frac{N^2}{\overline{\sigma}^4}-2\frac{E_p\{Q_{\overline{\Sigma}}\}}{\overline{\sigma}^6}+\frac{E_p\{Q_{\overline{\Sigma}}^2\}}{\overline{\sigma}^8} = \frac{N^2}{\overline{\sigma}^4}-2\frac{N^2}{\overline{\sigma}^4}+\frac{N(N+1)(\lambda-1)}{\overline{\sigma}^4(\lambda-2)}$$
$$= \frac{N(N+\lambda-1)}{\overline{\sigma}^4(\lambda-2)}.$$
$$(51)$$

$$[\mathbf{B}_c]_{i,1} = E_p\left\{\frac{\partial\ln f_X(\mathbf{x}_m;\theta_0)}{\partial\overline{\sigma}^2}\cdot\frac{\partial\ln f_X(\mathbf{x}_m;\theta_0)}{\partial\text{vec}(\overline{\Sigma})_i}\right\}$$
$$= E_p\left\{\left(\frac{E_p\{Q_{\overline{\Sigma}}\}}{\overline{\sigma}^4}-\frac{N}{\overline{\sigma}^2}\right)\left(\frac{1}{\overline{\sigma}^2}\mathbf{x}_m^H\overline{\Sigma}^{-1}\mathbf{A}_i\overline{\Sigma}^{-1}\mathbf{x}_m-\text{tr}\left(\overline{\Sigma}^{-1}\mathbf{A}_i\right)\right)\right\}$$
$$= -\frac{N}{\overline{\sigma}^2}\text{tr}\left(\overline{\Sigma}^{-1}\mathbf{A}_i\right)+\frac{1}{\overline{\sigma}^6}E_p\left\{\mathbf{x}_m^H\overline{\Sigma}^{-1}\mathbf{x}_m\mathbf{x}_m^H\overline{\Sigma}^{-1}\mathbf{A}_i\overline{\Sigma}^{-1}\mathbf{x}_m\right\}$$
$$= \left(-\frac{N}{\overline{\sigma}^2}+\frac{(N+1)(\lambda-1)}{\overline{\sigma}^2(\lambda-2)}\right)\text{tr}\left(\overline{\Sigma}^{-1}\mathbf{A}_i\right) = \frac{N+\lambda-1}{\overline{\sigma}^2(\lambda-2)}\text{tr}\left(\overline{\Sigma}^{-1}\mathbf{A}_i\right).$$
$$(52)$$

Hence, we get $\mathbf{B}_c = \frac{N+\lambda-1}{\overline{\sigma}^2(\lambda-2)}\text{vec}\left(\overline{\Sigma}^{-1}\right)$. Finally, some clarifications on the matrix $\mathbf{U}$ in the mismatched case need to be done. As for the matched case, the constraint on the trace of the scatter matrix can be rewritten as $f(\boldsymbol{\theta}) = \sum_{i\in I}\text{vecs}(\boldsymbol{\Sigma})_i - N = 0$, where I is the set of indices in (23). Hence, exactly as in Section 3.1.3, the $l + 1$-dimensional row gradient vector of the constraint is

$$\nabla f(\boldsymbol{\theta}) = \frac{\partial f(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}^T} = \left[\frac{\partial\sum_{i\in I}\text{vecs}(\boldsymbol{\Sigma})_i}{\partial\text{vecs}(\boldsymbol{\Sigma})^T} \quad 0\right] = \begin{bmatrix}\mathbf{1}_I^T & 0\end{bmatrix},$$
$$(53)$$

where $\mathbf{1}_I^T$ is given in Eq. (25), but unlike Eq. (24) where the two last zero entries were due to the scale and shape parameters, here after $\mathbf{1}_I^T$, we have only a zero entry relative to the power σ^2. To close this section, we note that $\mathbf{U}\in\mathbb{R}^{(l+1)\times l}$ in (42) is the matrix whose columns form an orthonormal basis for the null space of $\nabla f(\boldsymbol{\theta})$ in (53).

3.3 The robust approach

Unlike previous scenarios, where the estimators of the scatter matrix have been derived by assuming the correct *t*-distributed data model (matched case) or the simpler, but different from the true one, Gaussian data

Fortunati *et al. EURASIP Journal on Advances in Signal Processing* (2016) 2016:123

Page 9 of 16

model (mismatched case), we now focus on robust estimation, i.e., we aim at finding an estimator that does not assume any specific model for the data. A robust estimator is supposed to provide good estimation performance over a large set of different models (in the application discussed here, the set of CES models), even if not optimal under any nominal (matched or mismatched) one. Because of its generality, a robust estimator of the scatter matrix over the CES distributions will not rely on any additional estimates of unknown extra parameters, as it is for the matched ML estimator in Eq. (4) that depends on the estimates of λ and η.

There is a vast literature on robust estimation of the scatter matrix of CES-distributed data. In particular, it can be shown that the so-called constrained Tyler's (C-Tyler) fixed-point estimator is the most robust scatter matrix estimator in min-max sense over the CES class [5, 16, 39–41]. The C-Tyler estimator can be obtained as the recursive solution of the following fixed-point matrix equation:

$$\mathbf{\Sigma} = \frac{N}{M} \sum_{m=1}^{M} \frac{\mathbf{x}_m \mathbf{x}_m^H}{\mathbf{x}_m^H \mathbf{\Sigma}^{-1} \mathbf{x}_m}, \quad \mathrm{tr}(\mathbf{\Sigma}) = N. \tag{54}$$

To solve Eq. (54), we use the following iterative approach [40]:

$$\begin{cases} \hat{\mathbf{\Sigma}}_T^{(0)} = \mathbf{I} \\ \mathbf{S}_T^{(k+1)} = \sum_{m=1}^{M} \frac{\mathbf{x}_m \mathbf{x}_m^H}{\mathbf{x}_m^H \left(\hat{\mathbf{\Sigma}}_T^{(k)}\right)^{-1} \mathbf{x}_m} \\ \hat{\mathbf{\Sigma}}_T^{(k+1)} = N \mathbf{S}_T^{(k+1)} / \mathrm{tr}\left(\mathbf{S}_T^{(k+1)}\right) \end{cases} \tag{55}$$

for $k = 1,...,K$, where K is the number of iterations. It can be noted that in (55), there is a normalization on the trace of $\hat{\mathbf{\Sigma}}_T^{(k)}$ at every step of the iterative procedure to impose the constraint on the trace. Asymptotic consistency and unbiasedness properties are discussed in [5] and [16]. It is worth noting that the performance of the C-Tyler estimator can be assessed by comparing its error covariance matrix on the estimation of $\mathbf{\Sigma}$ with the CCRB derived in (26).

4 Hypothesis testing problem for target detection

After having discussed the three approaches for the scatter matrix estimation in *t*-distributed data, we can introduce the classical radar detection problem. In particular, we address the problem of detecting a complex signal vector **s** in the received data $\mathbf{x} = \mathbf{s} + \mathbf{c}$ where **c** represents the unobserved complex noise/clutter random vector. The target signal **s** is modelled as $\mathbf{s} = \alpha \mathbf{p}$ where **p** (generally called target vector response or Doppler steering vector) is the transmitted known radar pulse vector and $\alpha = \gamma e^{j\phi} \in \mathbb{C}$ is an unknown signal parameter accounting

for both channel propagation effects and the target back-scattering. α can be modelled as an unknown deterministic parameter or as a random variable depending on the application at hand. When modelled as a random quantity, α is assumed to be a circular Gaussian random variable $\alpha \sim CN\left(0, \sigma_\alpha^2\right)$ where the amplitude γ is Rayleigh distributed and the phase ϕ is uniformly distributed in $[0, 2\pi)$ and independent of γ. More general target models are the so-called Swerling models [42]. Regarding the complex noise vector **c**, it has been successfully modelled as a zero-mean CES-distributed random vector with covariance matrix $\mathbf{M} = \sigma^2 \mathbf{\Sigma}$, where $\mathbf{\Sigma}$ and σ^2 represent the unknown scatter matrix and the unknown statistical noise power. In particular, **c** is modelled as a complex *t*-distributed random vector [12, 13, 24].

The target detection problem can be expressed as a composite binary hypothesis testing problem

$$H_0 : |\alpha| = 0 \quad \text{vs.} \quad H_1 : |\alpha| > 0, \tag{56}$$

or, more explicitly as

$$\begin{cases} H_0 : \mathbf{x} = \mathbf{c} & \mathbf{x}_m = \mathbf{c}_m, m = 1, ..., M \\ H_1 : \mathbf{x} = \alpha \mathbf{p} + \mathbf{c} & \mathbf{x}_m = \mathbf{c}_m, m = 1, ..., M \end{cases} \tag{57}$$

where the secondary data $\{\mathbf{x}_m\}_{m=1}^{M}$ can be used to estimate the scatter matrix.

4.1 The matched case and the linear threshold detector

In [12], a GLRT (with respect to the unknown signal parameter α) has been derived as

$$\Lambda_{\mathrm{LTD}} \equiv \Lambda_{\mathrm{LTD}}(\mathbf{x}, \mathbf{\Sigma}, \lambda, \eta) = \frac{\left|\mathbf{p}^H \mathbf{\Sigma}^{-1} \mathbf{x}\right|^2}{\left(\mathbf{p}^H \mathbf{\Sigma}^{-1} \mathbf{p}\right)\left(\mathbf{x}^H \mathbf{\Sigma}^{-1} \mathbf{x} + \lambda/\eta\right)}, \tag{58}$$

where LTD stands for linear threshold detector. For an extensive discussion on the statistical properties of the Λ_{LTD}, we refer the reader to [12]. Of course, the LTD in Eq. (58) cannot be directly used in practical applications since it relies on the exact knowledge of the scatter matrix $\mathbf{\Sigma}$ and of the shape and scale parameters, λ and η, respectively. In order to estimate these quantities, we can use the matched algorithms discussed in Section 3. In particular, using the CML-MoM algorithms of Section 3.1.1 and the CML-WMoM algorithm of Section 3.1.2, we obtain two adaptive LTD detectors:

$$\begin{aligned} \Lambda_{\mathrm{LTD\text{-}CML\text{-}MoM}} &\equiv \Lambda_{\mathrm{LTD}}\left(\mathbf{x}, \hat{\mathbf{\Sigma}}_{\mathrm{CML}}, \hat{\lambda}_{\mathrm{MoM}}, \hat{\eta}_{\mathrm{MoM}}\right) \\ &= \frac{\left|\mathbf{p}^H \hat{\mathbf{\Sigma}}_{\mathrm{CML}}^{-1} \mathbf{x}\right|^2}{\left(\mathbf{p}^H \hat{\mathbf{\Sigma}}_{\mathrm{CML}}^{-1} \mathbf{p}\right)\left(\mathbf{x}^H \hat{\mathbf{\Sigma}}_{\mathrm{CML}}^{-1} \mathbf{x} + \hat{\lambda}_{\mathrm{MoM}}/\hat{\eta}_{\mathrm{MoM}}\right)}, \end{aligned} \tag{59}$$

and

$$\Lambda_{\text{LTD-CML-WMoM}} \equiv \Lambda_{\text{LTD}}\left(\mathbf{x}, \hat{\boldsymbol{\Sigma}}_W, \hat{\lambda}_{\text{WMoM}}, \hat{\eta}_{\text{WMoM}}\right)$$
$$= \frac{\left|\mathbf{p}^H \hat{\boldsymbol{\Sigma}}_W^{-1} \mathbf{x}\right|^2}{\left(\mathbf{p}^H \hat{\boldsymbol{\Sigma}}_W^{-1} \mathbf{p}\right)\left(\mathbf{x}^H \hat{\boldsymbol{\Sigma}}_W^{-1} \mathbf{x} + \hat{\lambda}_{\text{WMoM}}/\hat{\eta}_{\text{WMoM}}\right)}. \tag{60}$$

4.2 The mismatched case and the Kelly's GLRT

Following the mismatched approach used in Section 3.2, in this section, we analyze a sort of mismatched detection algorithm. In particular, as in Section 3.2, we assume that the noise vectors in (57), i.e., $\{\mathbf{x}_m\}_{m=1}^M$, are distributed according to the complex normal pdf of Eq. (27), while the true pdf is given by Eq. (2), i.e., they are complex t-distributed random vectors. It is well-known that, under Gaussian assumption, the GLRT (with respect to both the target parameter α and the noise covariance matrix $\mathbf{M} = \sigma^2 \boldsymbol{\Sigma}$) is given by Kelly's detector:

$$\Lambda_{\text{Kelly's GLRT}} \equiv \Lambda_{\text{Kelly's GLRT}}\left(\mathbf{x}, \hat{\mathbf{M}}\right)$$
$$= \frac{\left|\mathbf{p}^H \hat{\mathbf{M}}^{-1} \mathbf{x}\right|^2}{\left(\mathbf{p}^H \hat{\mathbf{M}}^{-1} \mathbf{p}\right)\left(1 + M^{-1}\mathbf{x}^H \hat{\mathbf{M}}^{-1} \mathbf{x}\right)}, \tag{61}$$

where $\hat{\mathbf{M}}$ is the sample covariance matrix (SCM), which is the ML estimator of the covariance matrix for Gaussian-distributed random vectors [17]. It is immediate to show that in our mismatched framework, the SCM also represents the MML estimator derived in (34)

$$\hat{\mathbf{M}} = \frac{1}{M}\sum_{m=1}^M \mathbf{x}_m \mathbf{x}_m^H = \hat{\sigma}_{\text{MML}}^2 \hat{\boldsymbol{\Sigma}}_{\text{MML}}. \tag{62}$$

Using the equality in Eq. (62), Kelly's GLRT can be expressed as

$$\Lambda_{\text{Kelly's GLRT}} \equiv \Lambda_{\text{Kelly's GLRT}}\left(\mathbf{x}, \hat{\sigma}_{\text{MML}}^2, \hat{\boldsymbol{\Sigma}}_{\text{MML}}\right)$$
$$= \frac{\left|\mathbf{p}^H \hat{\boldsymbol{\Sigma}}_{\text{MML}}^{-1} \mathbf{x}\right|^2}{\left(\mathbf{p}^H \hat{\boldsymbol{\Sigma}}_{\text{MML}}^{-1} \mathbf{p}\right)\left(\hat{\sigma}_{\text{MML}}^2 + M^{-1}\mathbf{x}^H \hat{\boldsymbol{\Sigma}}_{\text{MML}}^{-1} \mathbf{x}\right)}. \tag{63}$$

We note, in passing, that the Kelly's GLRT emerges also in detection problems involving CES-distributed data. In particular, in [43], it is shown that the Kelly's GLRT is a robust detector over a wide subclass of CES data distributions. However, it must be noted that the clutter model assumed in [43] is different from the IID model in Eq. (57). The model adopted here corresponds to what Raghavan and Pulsone in [44] called the "independent model," whereas the one considered in [43] corresponds to the so-called dependent model.

4.3 The robust approach and the ANMF

Finally, in this section, we discuss a robust detection algorithm under CES-distributed data vectors. For the reason we explain below, it is reasonable to choose as robust detector the normalized matched filter (NMF), proposed, e.g., in [18–23], as

$$\Lambda_{\text{NMF}} \equiv \Lambda_{\text{NMF}}(\mathbf{x}, \boldsymbol{\Sigma}) = \frac{\left|\mathbf{p}^H \boldsymbol{\Sigma}^{-1} \mathbf{x}\right|^2}{\left(\mathbf{p}^H \boldsymbol{\Sigma}^{-1} \mathbf{p}\right)\left(\mathbf{x}^H \boldsymbol{\Sigma}^{-1} \mathbf{x}\right)}, \tag{64}$$

where for the moment, the data scatter matrix $\boldsymbol{\Sigma}$ is assumed to be perfectly known.

An important feature of the detector in (64) is the invariance under scalar multiplies of $\mathbf{x}$. In particular, the distribution of the test statistic Λ_{NMF} under the hypothesis H_0 is independent of the unknown average noise power σ^2 or the functional form of the particular CES distribution of the noise, i.e., the NMF is a distribution-free detector under H_0. The proof of this property can be found in [5]. Moreover, it can be shown that $\Lambda_{\text{NMF}}|H_0$ follows a beta distribution:

$$\Lambda_{\text{NMF}}|H_0 \sim \text{beta}(\lambda; 1, N-1), \tag{65}$$

where $\text{beta}(x; \alpha, \beta) = (x^{\alpha-1}(1-x)^{\beta-1})/\text{B}(\alpha, \beta)$, N is the dimension of the data vector, and $\text{B}(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$.

It is clear that the NMF cannot be used in practical applications where the scatter matrix $\boldsymbol{\Sigma}$ of the data vectors is generally unknown. In order to overcome this limitation, an adaptive NMF (ANMF) detector can be derived by substituting to $\boldsymbol{\Sigma}$ its min-max (over the CES distributions) robust estimate, i.e., the C-Tyler estimator $\hat{\boldsymbol{\Sigma}}_T$:

$$\Lambda_{\text{ANMF-C-Tyler}} \equiv \Lambda_{\text{ANMF-C-Tyler}}\left(\mathbf{x}, \hat{\boldsymbol{\Sigma}}_T\right) \tag{66}$$
$$= \frac{\left|\mathbf{p}^H \hat{\boldsymbol{\Sigma}}_T^{-1} \mathbf{x}\right|^2}{\left(\mathbf{p}^H \hat{\boldsymbol{\Sigma}}_T^{-1} \mathbf{p}\right)\left(\mathbf{x}^H \hat{\boldsymbol{\Sigma}}_T^{-1} \mathbf{x}\right)}$$

As a consequence of the consistency of the Tyler's estimator, the resulting adaptive test statistic Λ_{ANMF} will have approximately a beta(1,N-1) distribution for sufficiently large M, i.e., Λ_{ANMF} is asymptotically CFAR w.r.t. $\boldsymbol{\Sigma}$, as desired [5]. Further discussions on the asymptotic properties of the Λ_{ANMF} can be found in [45] and [46].

5 Simulation results

In this section, we integrate through extensive numerical simulations, the theoretical findings on scatter matrix estimation and adaptive detection discussed in previous sections. In all the simulation results reported here, the true scatter matrix is assumed to be $[\boldsymbol{\Sigma}]_{i,j} = \rho^{|i-j|}$, for $i, j = 1,2,\dots,N$. Note that ρ is the clutter one-lag correlation coefficient that is assumed to be real. Under this assumption, $\boldsymbol{\Sigma}$ is real, as well. To exploit this assumption, in all numerical simulations, we took the real part of the scatter matrix estimators. The extension to the more general case of complex scatter matrix will be investigated in future works.

Fortunati *et al. EURASIP Journal on Advances in Signal Processing* (2016) 2016:123

Page 11 of 16

5.1 Estimation performance

In this section, we compare the estimation performance of the matched CML-MoM estimator, the recursive CML-WMoM estimator, and the robust C-Tyler estimator with the CCRB, while the performance of the mismatched CMML estimator are compared with the CMCRB. In order to have a global performance index (i.e., an index that is able to take into account the errors made in the estimation of all the covariance entries), we define ε as the Frobenius norm of the mean square error (MSE) matrix of the estimator [47]:

$$\varepsilon \triangleq \left\| E\left\{ \left(\text{vecs}(\hat{\Sigma}) - \text{vecs}(\Sigma)\right)\left(\text{vecs}(\hat{\Sigma}) - \text{vecs}(\Sigma)\right)^T \right\} \right\|_F,$$
(67)

where $\hat{\Sigma}$ is an estimate of the true covariance matrix Σ (e.g., $\varepsilon_{\text{C-Tyler}}, \varepsilon_{\text{CML-MoM}}, \varepsilon_{\text{CML-WMoM}},$ and $\varepsilon_{\text{CMML}}$) and $\|A\|_F = \sqrt{\text{tr}(A^T A)}$ is the Frobenius norm of matrix A. As performance bounds, the following quantity are plotted:

$$\varepsilon_{\text{CCRB}} \triangleq \|\text{CCRB}(\Sigma)\|_F, \quad \varepsilon_{\text{CMCRB}} \triangleq \|\text{CMCRB}(\Sigma)\|_F.$$
(68)

The accuracy on the estimate of the shape λ and scale η parameters in the matched case and of the average power σ^2 in the mismatched case is measured through their MSE, which is compared with the CCRB and the CMCRB, respectively. To calculate the estimation accuracy, we run 10^5 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, 2, 3, and 4). Simulation parameters: $\rho = 0.8, N = 16, \lambda = 3, \eta = 1, K = 4$.
2. Estimation accuracy as function of the shape parameter λ (Figs. 5, 6, 7, and 8). Simulation parameters: $\rho = 0.8, N = 16, M = 10\,N, \eta = 1, K = 4$.
3. Estimation accuracy as function of the one-lag correlation coefficient ρ (Figs. 9, 10, 11, and 12). Simulation parameters: $N = 16, M = 10\,N, \lambda = 3, \eta = 1, K = 4$.

Based on the numerical analysis, we observe that:

- Regarding the scatter matrix estimation, the robust C-Tyler estimator is an "almost" efficient estimator, even if it is not the most efficient estimator for t-distributed data, in fact when λ increases, the other two estimators achieve better performance. The MSE $\varepsilon_{\text{C-Tyler}}$ is close to the CCRB especially for small λ (see Figs. 1, 5, and 9). In particular, its performance is robust, i.e., it is not affected by the value of the shape parameter λ (see Fig. 5), even if it is not efficient for large λ.
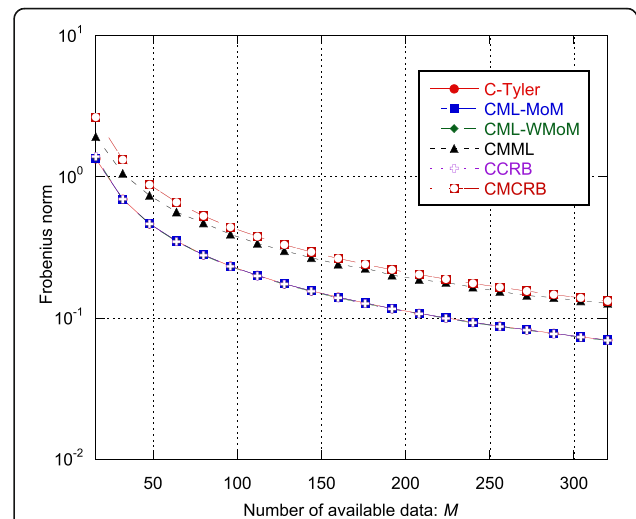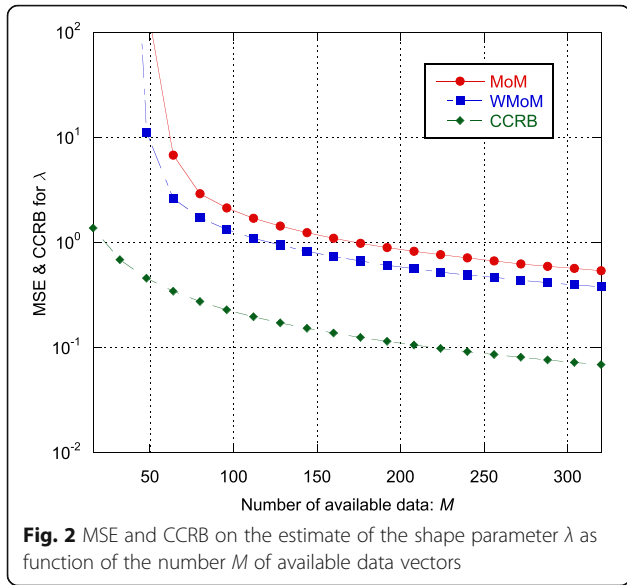

Fig. 1 MSE indices $\varepsilon_{\text{C-Tyler}}, \varepsilon_{\text{CML-MoM}}, \varepsilon_{\text{CML-WMoM}},$ and $\varepsilon_{\text{CMML}}$ and bounds $\varepsilon_{\text{CCRB}}$ and $\varepsilon_{\text{CMCRB}}$ as function of the number M of available data vectors

- Regarding the CMML estimator, it always achieves the CMCRB, both for the scatter matrix estimation and for the estimation of the average power (see Figs. 1, 5, 8, 9, and 12). The CMML presents a small bias on the estimation of the scatter matrix and then, $\hat{\Sigma}_{\text{CMML}}$ is not a *MS*-unbiased estimator [9] (at least in the finite sample regime). For this reason, $\varepsilon_{\text{CMML}}$ is in general slightly below the CMCRB. The loss in estimation accuracy due to the mismatch is particularly high for extremely heavy-tailed data, i.e., when λ is close to 0 (see Fig. 5). When $\lambda \to 0$, the CMCRB rapidly increases while the CCRB is quite independent of λ. On the other hand, when $\lambda \to \infty$, the CMCRB and the CCRB coincide, as expected, and the performance of the CMML estimator converge to that of the CML-MoM and CML-WMoM estimators.
- Quite surprisingly, even if the MoM-based estimators fail to provide an accurate estimate of λ as it increases (see Fig. 6), the MSE of the CML-MoM and CML-WMoM estimators achieve the CCRB, as shown in Fig. 5.
- Regarding the estimation of λ and η, the recursive WMoM estimator always outperforms the classical MoM estimator (see Figs. 2, 3, 7, 10, and 11), even though it does not achieve the CCRB. In particular, as shown in Fig. 10, the MSE of the WMoM is independent from the value of ρ, while this is not the case for the MSE of the classical MoM estimator. This desirable behavior of the WMoM estimator is due to the whitening operation that makes each entry of the data vectors mutually uncorrelated, as discussed in Section 3.1.

Fig. 2 MSE and CCRB on the estimate of the shape parameter λ as function of the number M of available data vectors

Fig. 4 MSE of the CMML estimator of σ^2 and CMCRB as function of the number M of available data vectors
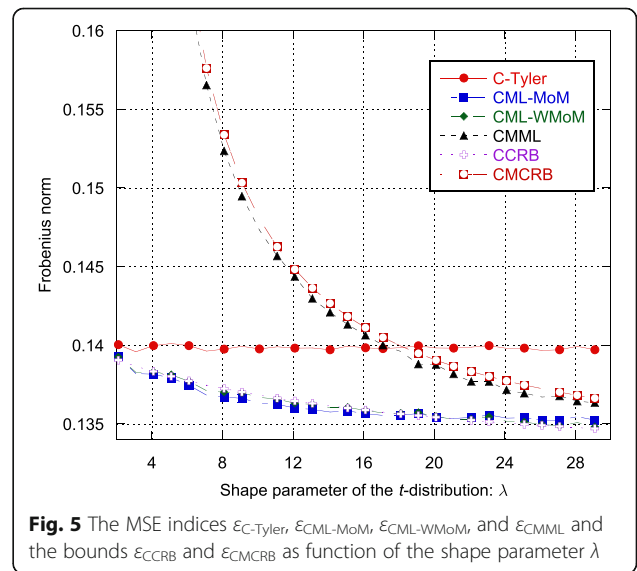
5.2 Detection performance

In this section, the detection performance of the matched LTD detector, which exploits either the CML-MoM or the CML-WMoM estimators, the mismatched Kelly's GLRT, and the robust ANMF, that relies on the robust C-Tyler's estimator, are investigated. In particular, we analyze:

1. The probability of false alarm (P_{FA}) as function of the one-lag coefficient ρ (Fig. 13). This allows us to verify the CFAR property of the $\Lambda_{\text{LTD-CML-MoM}}$ (Eq. (59)), $\Lambda_{\text{LTD-CML-WMoM}}$ (Eq. (60)), Λ_{Kelly} (Eq. (63)), and $\Lambda_{\text{ANMF-C-Tyler}}$ (Eq. (66)) w.r.t. the correlation shape. Simulation parameters: $N = 16$,
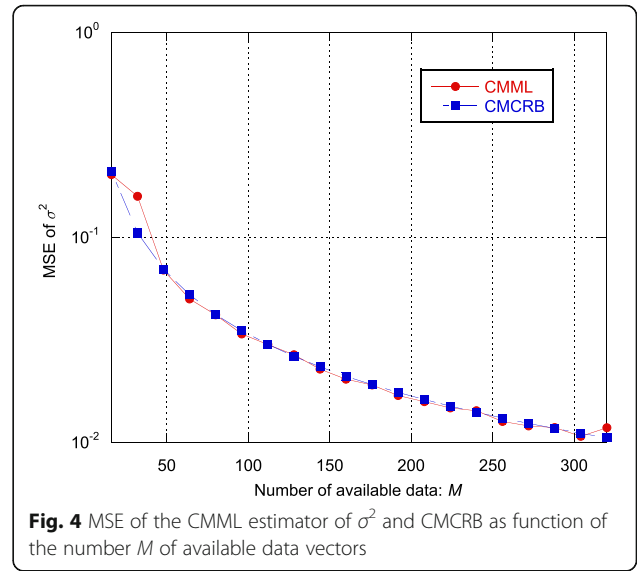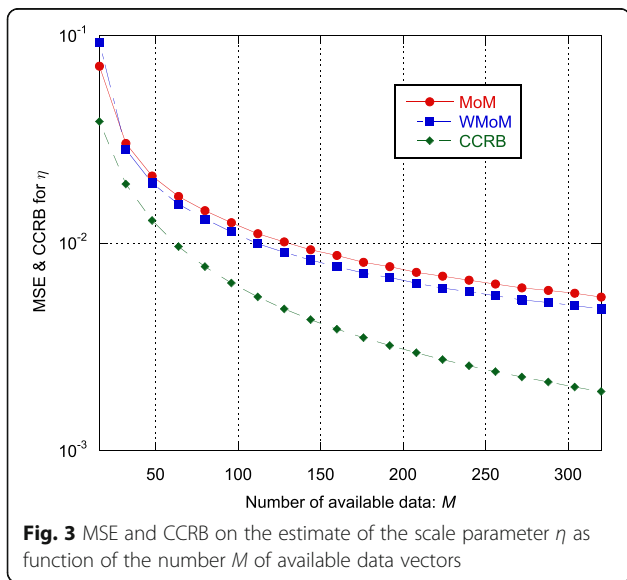
$M = 3\,N$, $\lambda = 3$, $\eta = 1$, $K = 4$. The detection thresholds have been set to achieve a nominal P_{FA} of 10^{-3}.
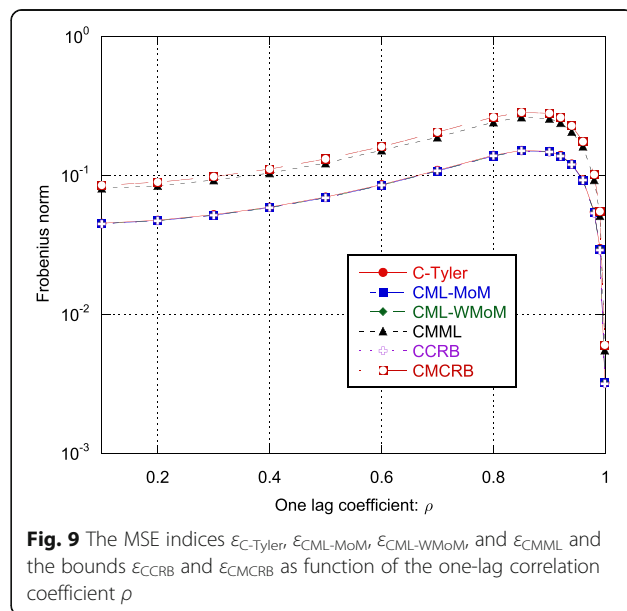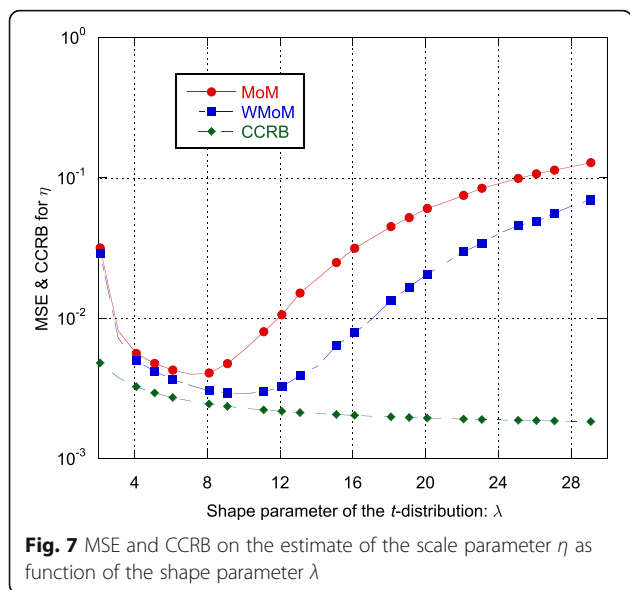
2. The probability of false alarm (P_{FA}) as function of the shape parameter λ of the true complex t-distribution, i.e., for different spikiness levels (Fig. 14). This is important, since it highlights the CFARness of the four detectors w.r.t. the non-Gaussianity level of the data. Simulation parameters: $N = 16$, $M = 3\,N$, $\rho = 0.8$, $\eta = 1$, $K = 4$. The detection thresholds have been set to achieve a nominal P_{FA} of 10^{-3}.

3. The receiver operating characteristic (ROC) curves (Fig. 15). The simulation parameters are the following: $N = 16$, $M = 3\,N$, $\rho = 0.8$, $\lambda = 3$, $\eta = 1$, and $K = 4$. Moreover, $\alpha \sim CN\left(0, \sigma_\alpha^2\right)$ where σ_α^2 is set to have signal to noise power ratio (SNR) equal to 3 dB.

Fig. 3 MSE and CCRB on the estimate of the scale parameter η as function of the number M of available data vectors

Fig. 5 The MSE indices $\varepsilon_{\text{C-Tyler}}$, $\varepsilon_{\text{CML-MoM}}$, $\varepsilon_{\text{CML-WMoM}}$, and $\varepsilon_{\text{CMML}}$ and the bounds $\varepsilon_{\text{CCRB}}$ and $\varepsilon_{\text{CMCRB}}$ as function of the shape parameter λ

Fig. 6 MSE and CCRB on the estimate of the shape parameter λ as function of λ

Fig. 8 MSE of the CMML estimator of σ^2 and CMCRB as function of the shape parameter $\boldsymbol{\lambda}$
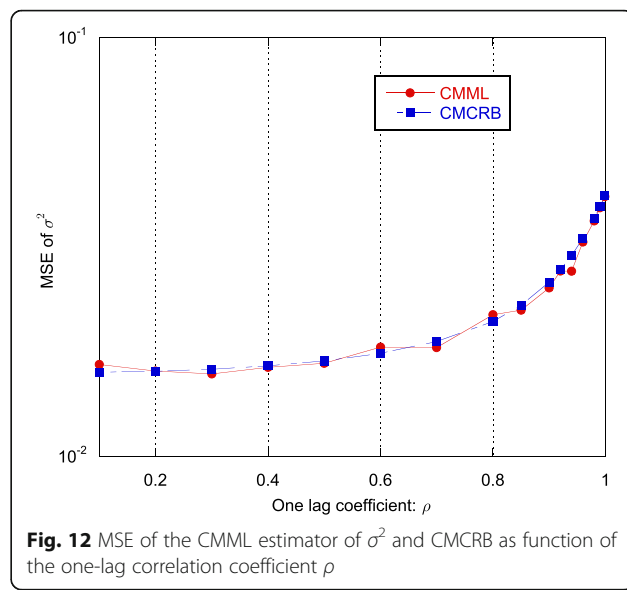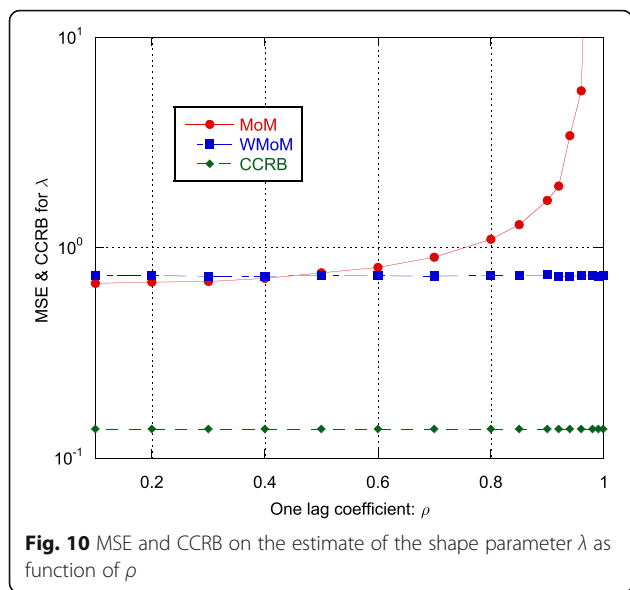
As we can see from Fig. 13, all the analyzed detectors are CFAR with respect to ρ. Their P_{FA} curves are constant and close to the nominal value 10^{-3}. A different behavior can be observed in Fig. 14, where the P_{FA} curves have been evaluated as function of λ. It can be noted that only $\Lambda_{\text{ANMF-C-Tyler}}$ is a CFAR detector w.r.t. the data spikiness, while the P_{FA} of the other detectors change with λ. Finally, in Fig. 15, the ROC curves of $\Lambda_{\text{LTD-CML-MoM}}$, $\Lambda_{\text{LTD-CML-WMoM}}$, $\Lambda_{\text{Kelly's GLRT}}$, and $\Lambda_{\text{ANMF-C-Tyler}}$ are shown. For the sake of comparison, we evaluated also the ROC of the clairvoyant optimum detector for the t-distributed data, i.e., the Λ_{LTD} in Eq. (58) where $\boldsymbol{\Sigma}$, λ, and η are perfectly known. As we can see, the performance of $\Lambda_{\text{LTD-CML-MoM}}$, $\Lambda_{\text{LTD-CML-}}$

$_{\text{WMoM}}$, and $\Lambda_{\text{ANMF-C-Tyler}}$ are close to that of the clairvoyant detector Λ_{LTD}, while $\Lambda_{\text{Kelly's GLRT}}$ undergoes some detection loss for relatively low value of the P_{FA}. In particular, the fact that the performance of the robust $\Lambda_{\text{ANMF-C-Tyler}}$ is close to the one of the matched detectors, $\Lambda_{\text{LTD-CML-MoM}}$ and $\Lambda_{\text{LTD-CML-WMoM}}$, suggests that the detection loss due to the robustness is small. However, it must be highlighted again that we are considering a particular scenario in which the clutter covariance matrix is assumed to be real and full rank. Moreover, due to the high computational load of the Monte Carlo simulations, the detection performance of the proposed detectors has been evaluated only for a P_{FA} greater that 10^{-5}. It would be very useful to investigate

Fig. 7 MSE and CCRB on the estimate of the scale parameter η as function of the shape parameter λ

Fig. 9 The MSE indices $\varepsilon_{\text{C-Tyler}}$, $\varepsilon_{\text{CML-MoM}}$, $\varepsilon_{\text{CML-WMoM}}$, and $\varepsilon_{\text{CMML}}$ and the bounds $\varepsilon_{\text{CCRB}}$ and $\varepsilon_{\text{CMCRB}}$ as function of the one-lag correlation coefficient ρ

Fig. 10 MSE and CCRB on the estimate of the shape parameter λ as function of ρ

Fig. 12 MSE of the CMML estimator of σ^2 and CMCRB as function of the one-lag correlation coefficient ρ
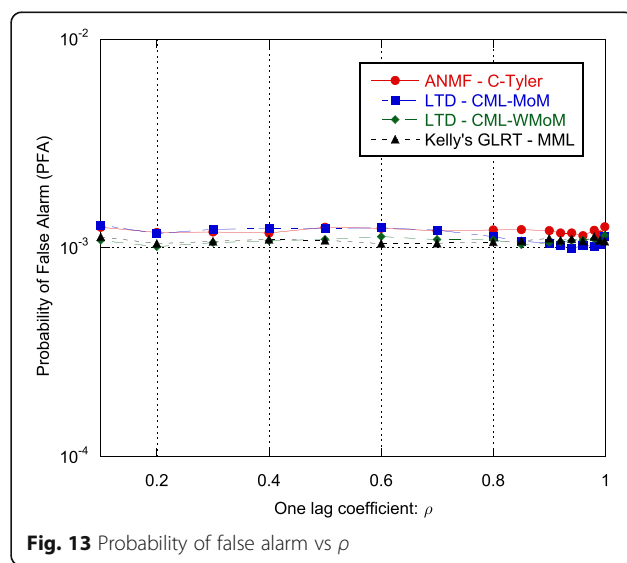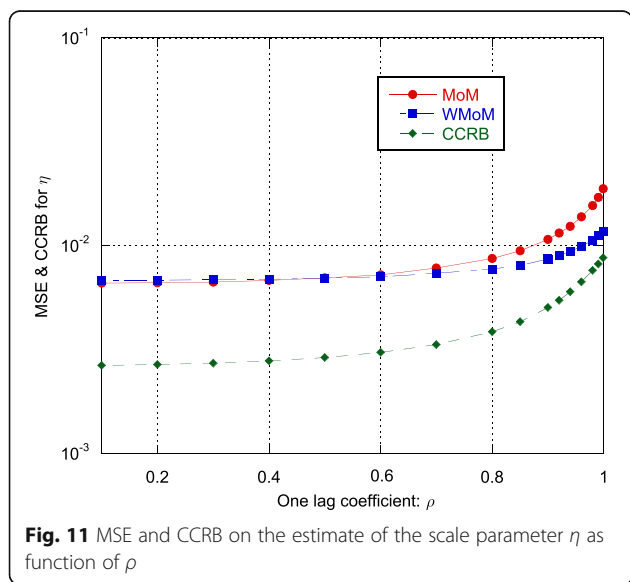
the detection performance at an operative value of P_{FA}, e.g., below 10^{-5}.

6 Conclusions

This paper focused on two inference problems, the scatter matrix estimation and the adaptive detection of radar targets in complex t-distributed data. Three different approaches have been investigated and compared: the matched, the mismatched, and the robust approaches. Regarding the classical matched approach, we analyzed the performance of the CML estimator for the scatter matrix, when the shape and scale parameters are estimated through the low-complexity and suboptimal MoM method (CML-MoM) and a recursive improvement of it (CML-WMoM). We found that both the

CML-MoM and the CML-WMoM estimators achieve the CCRB, while the CML-WMoM estimator outperforms the CML-MoM for the estimation of the shape and scale parameters. Then, the previous two estimators have been adapted to implement the LTD, which is the GLRT decision rule in t-distributed data. Numerical simulations show that the performance of the adaptive LTD are very close to the clairvoyant LTD detector, but it is not CFAR w.r.t. the variation of data spikyness. Regarding the mismatched approach, we proved that the CMML estimator derived under the assumption of Gaussian-distributed data converges almost surely to the true scatter matrix and to the true (t-distributed) data power, so it can be applied for inference problems that require the knowledge of these two quantities. Moreover,
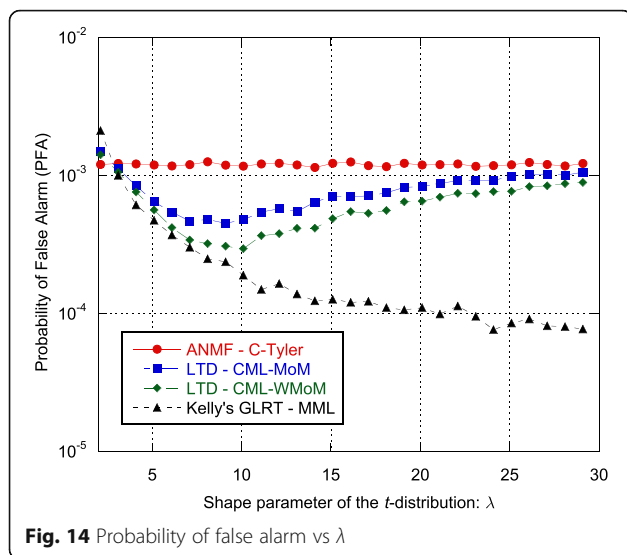

Fig. 11 MSE and CCRB on the estimate of the scale parameter η as function of ρ

Fig. 13 Probability of false alarm vs ρ

Fortunati *et al. EURASIP Journal on Advances in Signal Processing* (2016) 2016:123

Page 15 of 16

Fig. 14 Probability of false alarm vs λ

its efficiency with respect to the CMCRB has been shown and its performance loss with respect to the matched case discussed. The CMML estimator of the scatter matrix has been used in the "mismatched" Kelly's GLRT. Numerical simulations proved that Kelly's GLRT is not CFAR and presents large performance loss for small values of P_{FA}. Finally, the min-max robust C-Tyler scatter matrix estimator and the adaptive version of the robust NMF detector, that exploits the C-Tyler estimator, have been introduced and analyzed. In particular, our numerical results demonstrated that Tyler's estimator is an "almost" efficient estimator w.r.t. the CCRB and its estimation accuracy is independent on the value of the shape parameter. More importantly, the resulting ANMF is CFAR w.r.t. the shape parameter, i.e., w.r.t. the level of data spikiness, and has only a small detection loss
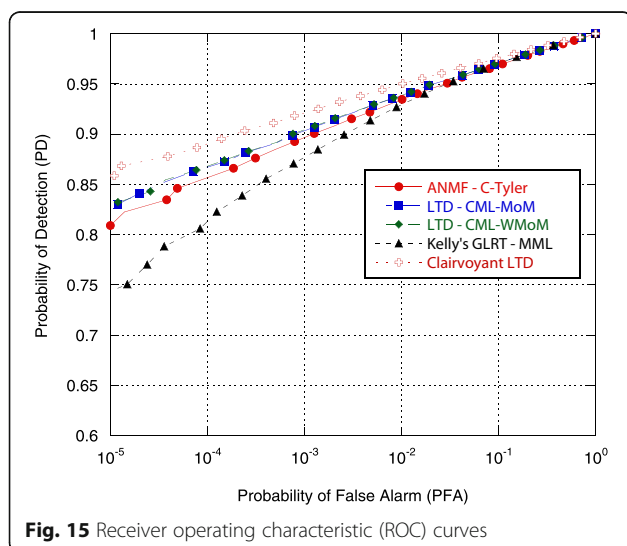

Fig. 15 Receiver operating characteristic (ROC) curves

w.r.t. the clairvoyant LTD. To summarize, the results discussed in this paper show that the robust approach, thanks to its generality, robustness to misspecification, and small estimation and detection losses, seems to be the a good choice in practical applications.

Competing interests
The authors declare that they have no competing interests.

References
1. H Krim, M Viberg, Two decades of array signal processing research: the parametric approach. IEEE Signal Process. Mag. **13**(4), 67–94 (1996)
2. J Karhunen, J Joutsensalo, Generalizations of principal component analysis, optimization problems, and neural networks. Neural Netw. **8**(4), 549–562 (1995)
3. O Ledoit, M Wolf, Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. J. Empir. Finance **10**(5), 603–621 (2003)
4. CD Richmond, *Adaptive array signal processing and performance analysis in non-Gaussian environments*. Ph.D. thesis (MIT, Cambridge, MA, 1996). Available at: https://dspace.mit.edu/handle/1721.1/11005
5. E Ollila, DE Tyler, V Koivunen, VH Poor, Complex elliptically symmetric distributions: survey, new results and applications. IEEE Trans. Signal Process. **60**(11), 5597–5625 (2012)
6. KL Lange, RJA Little, JMG Taylor, Robust statistical modeling using the *t* distribution. J. Am. Stat. Assoc. **84**(408), 881–896 (1989)
7. JP Romano, AF Siegel, *Counterexamples in probability and statistics*, Wadsworth and Brooks/Cole, Monterey, CA
8. S Fortunati, MS Greco, F Gini, The impact of unknown extra parameters on scatter matrix estimation and detection performance in complex *t*-distributed data, *IEEE Workshop on Statistical Signal Processing 2016 (SSP)*, (IEE, Palma de Mallorca, 2016), pp 1–4. doi:10.1109/SSP.2016.7551738
9. S Fortunati, F Gini, MS Greco, On scatter matrix estimation in the presence of unknown extra parameters: mismatched scenario, *EUSIPCO 2016*, (Budapest, Hungary, 2016)
10. A Balleri, A Nehorai, J Wang, Maximum likelihood estimation for compound-Gaussian clutter with inverse gamma texture. IEEE Trans. Aerosp. Electron. Syst. **43**(2), 775–779 (2007)
11. J Wang, A Dogandzic, A Nehorai, Maximum likelihood estimation of compound-Gaussian clutter and target parameters. IEEE Trans. Signal Process. **54**(10), 3884–3898 (2006)
12. KJ Sangston, F Gini, MS Greco, Coherent radar target detection in heavy-tailed compound-Gaussian clutter. IEEE Trans. Aerosp. Electron. Syst. **48**(1), 64–77 (2012)
13. A Younsi, M Greco, F Gini, A Zoubir, Performance of the adaptive generalised matched subspace constant false alarm rate detector in non-Gaussian noise: an experimental analysis. IET Radar Sonar Navigation **3**(3), 195–202 (2009)
14. S Fortunati, F Gini, MS Greco, The misspecified Cramér-Rao bound and its application to the scatter matrix estimation in complex elliptically symmetric distributions. IEEE Trans. Signal Process. **64**(9), 2387–2399 (2016)
15. S Fortunati, F Gini, MS Greco, The constrained misspecified Cramér-Rao bound. IEEE Signal Process. Lett. **23**(5), 718–721 (2016)
16. D Tyler, A distribution-free M-estimator of multivariate scatter. Ann. Stat. **15**(1), 234–251 (1987)
17. EJ Kelly, An adaptive detection algorithm. IEEE Trans. Aerosp. Electron. Syst. **AES-22**(2), 115–127 (1986)
18. LL Scharf, B Friedlander, Matched subspace detectors. IEEE Trans. Signal Process. **42**(8), 2146–2157 (1994)
19. F Gini, A cumulant-based adaptive technique for coherent radar detection in a mixture of K-distributed clutter and Gaussian disturbance. IEEE Trans. Signal Process. **45**(6), 1507–1519 (1997)
20. E Conte, A De Maio, G Ricci, Recursive estimation of the covariance matrix of a compound-Gaussian process and its application to adaptive CFAR detection. IEEE Trans. Signal Process. **50**(8), 1908–1915 (2002)
21. F Gini, Sub-optimum coherent radar detection in a mixture of k-distributed and Gaussian clutter. IEE Proc. Part F **144**(1), 39–48 (1997)
22. F Gini, MV Greco, A Farina, Clairvoyant and adaptive signal detection in non-Gaussian clutter: a data-dependent threshold interpretation. IEEE Trans. on Signal Processing **47**(6), 1522–1531 (1999)

23. F Gini, MV Greco, A suboptimum approach to adaptive coherent radar detection in compound-Gaussian clutter. IEEE Trans. Aerosp. Electron. Syst. **35**(3), 1095–1104 (1999)
24. E Ollila, DE Tyler, V Koivunen, HV Poor, Compound-Gaussian clutter modelling with an inverse Gaussian texture distribution. IEEE Signal Processing Letters **19**(12), 876–879 (2012)
25. A Farina, F Gini, MV Greco, L Verrazzani, High resolution sea clutter data: statistical analysis of recorded live data. IEE Proc. Radar Sonar Navigation **144**(3), 121–130 (1997)
26. R Remmer, *Theory of complex functions* (Springer-Verlag, New York, NY, USA, 1991)
27. MS Greco, F Gini, Cramér-Rao lower bounds on covariance matrix estimation for complex elliptically symmetric distributions. IEEE Trans. Signal Process. **61**(24), 6401–6409 (2013)
28. P Stinco, MS Greco, F Gini, *Adaptive detection in compound-Gaussian clutter with inverse-gamma texture*. IEEE CIE International Conference on Radar 2011, 2011, pp. 434–437
29. MS Greco, S Fortunati, F Gini, *Naive, robust or fully-adaptive: an estimation problem for CES distributions*. 8-th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2014), 2014, pp. 457–460
30. JR Magnus, H Neudecker, The commutation matrix: some properties and applications. Ann. Stat. **7**, 381–394 (1979)
31. JD Gorman, AO Hero, Lower bounds for parametric estimation with constraints. IEEE Trans. Inf. Theory **6**(6), 1285–1301 (1990)
32. P Stoica, BC Ng, On the Cramer-Rao bound under parametric constraints. IEEE Signal Processing Letters **5**(7), 177–179 (1998)
33. CD Richmond, LL Horowitz, Parameter bounds on estimation accuracy under model misspecification. IEEE Trans. on Signal Processing **63**(9), 2263–2278 (2015)
34. PJ Huber, *The behavior of maximum likelihood estimates under nonstandard conditions*. Proc. of the Fifth Berkeley Symposium in Mathematical Statistics and Probability (University of California Press, Berkley, 1967)
35. H White, Maximum likelihood estimation of misspecified models. Econometrica **50**, 1–25 (1982)
36. CC Heyde, R Morton, On constrained quasi-likelihood estimation. Biometrika **80**(4), 755–61 (1993)
37. QH Vuong, Cramér-Rao bounds for misspecified models, *Working paper 652, Division of the Humanities and Social Sciences* (Caltech, 1986). Available at: https://www.hss.caltech.edu/content/cramer-rao-bounds-misspecified-models
38. MS Greco, S Fortunati, F Gini, Maximum likelihood covariance matrix estimation for complex elliptically symmetric distributions under mismatched conditions. Signal Process. **104**, 381–386 (2014)
39. RA Maronna, Robust M-estimators of multivariate location and scatter. Ann. Stat. **4**(1), 51–67 (1976)
40. F Gini, MS Greco, Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter. Signal Process. **82**(12), 1847–1859 (2002)
41. F Pascal, Y Chitour, J Ovarlez, P Forster, P Larzabal, Covariance structure maximum-likelihood estimates in compound Gaussian noise: existence and algorithm analysis. IEEE Trans. Signal Process. **56**(1), 34–48 (2008)
42. M Skolnik, *Introduction to radar systems: third edition* (McGraw-Hill, New York, 2001)
43. CD Richmond, A note on non-Gaussian adaptive array detection and signal parameter estimation. IEEE Trans. on Signal Processing **3**(8), 251–252 (1996)
44. NB Pulsone, RS Raghavan, Analysis of an adaptive CFAR detector in non-Gaussian interference. IEEE Trans. Aerosp. Electron. Syst. **35**(3), 903–916 (1999)
45. F Pascal, JP Ovarlez, Asymptotic detection performance of the robust ANMF, *23rd European Signal Processing Conference* (EUSIPCO), 2015, Nice, 2015, pp. 524-528
46. JP Ovarlez, F Pascal, A Breloy, Asymptotic detection performance analysis of the robust adaptive normalized matched filter, *IEEE CAMSAP 2015*, (Cancun, 2015), pp. 137-140
47. F Gini, JH Michels, Performance analysis of two covariance matrix estimators in compound-Gaussian clutter. IEE Proceedings Part-F **146**(3), 133–140 (1999)