Robust-COMET for Covariance Estimation in Convex Structures: Algorithm and Statistical Properties

Bruno Mériaux*, Chengfang Ren*, Mohammed Nabil El Korso†, Arnaud Breloy† and Philippe Forster‡

* SONDRA, CentraleSupélec, 91192 Gif-sur-Yvette, France
† Université Paris-Nanterre/LEME, 92410 Ville d'Avray, France
‡ SATIE, ENS Paris-Saclay, CNRS, 94230 Cachan, France

Abstract—This paper deals with structured covariance matrix estimation in a robust statistical framework. Covariance matrices often exhibit a particular structure related to the application of interest and taking this structure into account increases estimation accuracy. Within the framework of robust estimation, the class of circular Complex Elliptically Symmetric (CES) distributions is particularly interesting to handle impulsive and spiky data. Normalized CES random vectors are known to share a common Complex Angular Elliptical distribution. In this context, we propose a Robust Covariance Matrix Estimation Technique (RCOMET) based on Tyler's estimate and COMET criterion for convexly structured matrices. We prove that the proposed estimator is consistent and asymptotically efficient while computationally attractive. Numerical results support the theoretical analysis in a particular application for Hermitian Toeplitz structure.

Index Terms—Robust covariance estimation, elliptical distributions, Tyler's M-estimator, structured covariance matrix.

I. INTRODUCTION

Covariance Matrix (CM) estimation plays a central role in adaptive signal processing. Besides the obvious Hermitian and positive characters, CM's often exhibit a particular structure related to the application of interest: a well-known example is the Toeplitz structure for uniform linear arrays. Taking this structure into account increases estimation accuracy. In a Gaussian context, this problem has been widely investigated. In particular, an estimation procedure known as Covariance Matching Estimation Technique (COMET) [1] is an interesting alternative to Maximum Likelihood (ML) estimation: it allows one to handle easily linear structures and provides asymptotically efficient CM estimates. However, COMET is based on the Sample Covariance Matrix (SCM) estimate: therefore, it is neither robust to departures from Gaussianity nor to outliers. Within the framework of robust CM estimation, the class of circular Complex Elliptically Symmetric distributions (CES) has attracted much attention since the pioneering works of Maronna [2] and Tyler [3]. Indeed, CES distributions encompass a large number of non Gaussian distributions that are met in various applications such as radar or sonar [4]: Generalized Gaussian, compound Gaussian, t-distribution and K-distribution... Within this CES framework, Tyler proposed an unstructured distribution free estimate of the scatter matrix [3] which may be interpreted as the ML estimate of the covariance matrix of normalized CES data [5]. The latter random vectors are known to share a common Complex Angular Elliptical (CAE) distribution [5]. Furthermore estimation of CM's with convex structures has been recently addressed for a CAE distribution [6, 7]. A COnvexly ConstrAined (COCA) CM estimator has been recently proposed in [6] in which the proposed estimator is consistent but suffers from heavy computational cost. Moreover, COCA is not asymptotically efficient. Iterative Majorization-Minimization algorithms for the computation of structured CM estimates are developped in [7]. However, the latter approach may be computationally demanding (except for several very specific structures yielding closed form iterations). In order to fulfill this lack, we propose in this paper a Robust Covariance Matrix Estimation Technique (RCOMET) based on Tyler's estimate and COMET criterion for convexly structured CM matrices. The proposed criterion is convex and its minimization admits a unique solution that can be efficiently computed (notably in a straightforward manner for linear structures). Our main contribution is to show that it yields consistent and asymptotically efficient CM estimates for CAE distributions.

This paper is organized as follows. In section II, the data model is presented. Section III focuses on the proposed algorithm. The performance analysis is also treated. Section IV gives a particular application considering a Hermitian Toeplitz structure with some simulation results. Finally, a brief conclusion is given in Section V.

In what follows, the notation $\overset{d}{=}$ indicates "has the same distribution as". Convergence in distribution and in probability are respectively denoted by $\overset{\mathcal{L}}{\rightarrow}$ and $\overset{\mathcal{P}}{\rightarrow}$. For a matrix $\mathbf{A}$, $|\mathbf{A}|$ and $\text{Tr}(\mathbf{A})$ denote the determinant and the trace of $\mathbf{A}$, $\mathbf{A}^T$ (respectively $\mathbf{A}^H$) stands for the transpose (respectively conjugate transpose) matrix. The vec-operator $\text{vec}(\mathbf{A})$ stacks all columns of $\mathbf{A}$ into a vector. The operator $\otimes$ refers to Kronecker matrix product. The subscript "e" refers to the true value. The notations $\Re$ and $\Im$ denote the real and imaginary parts.

II. PROBLEM SETUP

Let $\mathbf{x} \in \mathbb{C}^m$ be a circular CES random vector [5] with scatter matrix $\mathbf{R}$. If it exists, the covariance matrix of $\mathbf{x}$ is proportional to $\mathbf{R}$. The normalized vector $\mathbf{y} = \dfrac{\mathbf{x}}{\|\mathbf{x}\|}, \mathbf{x} \neq \mathbf{0}$, follows a CAE distribution, denoted by $\mathbf{y} \sim \mathcal{U}(\mathbf{R})$.

The vector $\mathbf{y}$ has the following probability density function (p.d.f.) w.r.t. spherical measure which is the natural Borel measure on the unit complex sphere $\mathbb{CS}^m \triangleq \{\mathbf{z} \in \mathbb{C}^m \mid \|\mathbf{z}\| = 1\}$ [6]

$$p(\mathbf{y} \mid \mathbf{R}) \propto |\mathbf{R}|^{-1} \left(\mathbf{y}^H \mathbf{R}^{-1} \mathbf{y}\right)^{-m} \tag{1}$$

where the shape matrix $\mathbf{R}$ is defined up to an arbitrary scale factor. To avoid scaling ambiguity, $\mathbf{R}$ is normalized according

to $\mathrm{Tr}\left(\mathbf{R}\right) = m$. It is worth noting that $\mathbf{R}$ is not the scaled covariance matrix of $\mathbf{y}$ unless $\mathbf{R} = \mathbf{I}$, though it is the scatter matrix of the underlying CES vector $\mathbf{x}$: that is why we refer to $\mathbf{R}$ as the shape matrix of $\mathbf{y}$. We assume that the latter matrix belongs to a convex subset $\mathscr{S}$ of Hermitian positive-definite matrices, and that there exists a one-to-one differentiable mapping $\boldsymbol{\mu} \mapsto \mathbf{R}(\boldsymbol{\mu})$ from $\mathbb{R}^P$ to $\mathscr{S}$. The vector $\boldsymbol{\mu}$ is the unknown interest parameter with exact value $\boldsymbol{\mu}_{\mathrm{e}}$, and $\mathbf{R}_{\mathrm{e}} = \mathbf{R}(\boldsymbol{\mu}_{\mathrm{e}})$ corresponds to the exact shape matrix.

Considering N i.i.d. CAE distributed observations, $\mathbf{y}_n \sim \mathcal{U}(\mathbf{R}_{\mathrm{e}})$, $n = 1, \ldots, N$, the log-likelihood function is given, up to an additive constant by

$$\mathcal{L}(\mathbf{y}_1, \ldots, \mathbf{y}_N; \boldsymbol{\mu}) = -m \sum_{n=1}^{N} \log\left(\mathbf{y}_n^H \mathbf{R}(\boldsymbol{\mu})^{-1} \mathbf{y}_n\right) - N \log |\mathbf{R}(\boldsymbol{\mu})| \quad (2)$$

The above log-likelihood is a non-convex function of $\mathbf{R}$. Therefore its maximization is a difficult and time consuming problem. To overcome this issue, we propose in the next section a new estimation procedure that provides unique, consistent and asymptotically efficient estimates. Moreover, closed form expressions of the estimates are easily obtained for linear structures of the shape matrix.

III. RCOMET: A ROBUST ASYMPTOTICALLY EFFICIENT COVARIANCE MATCHING ESTIMATE

Let $\widehat{\mathbf{R}} \succ \mathbf{0}$ be an unstructured shape matrix estimate based on N CAE distributed observations. Consider the following function

$$\mathrm{d}\left(\widehat{\mathbf{R}}, \alpha \mathbf{R}(\boldsymbol{\mu})\right) = \mathrm{Tr}\left(\left(\widehat{\mathbf{R}} - \alpha \mathbf{R}(\boldsymbol{\mu})\right) \widehat{\mathbf{R}}^{-1} \left(\widehat{\mathbf{R}} - \alpha \mathbf{R}(\boldsymbol{\mu})\right) \widehat{\mathbf{R}}^{-1}\right), \quad (3)$$

where $\alpha > 0$ is needed for the purpose of theoretical derivations, though the shape matrix is parameterized as $\mathbf{R}(\boldsymbol{\mu})$. Given $\widehat{\mathbf{R}}$, $\mathrm{d}\left(\widehat{\mathbf{R}}, \alpha \mathbf{R}(\boldsymbol{\mu})\right)$ is a convex function of $\alpha \mathbf{R}$. Therefore, for $\mathbf{R} \in \mathscr{S}$ convex set and $\alpha > 0$, the minimization of (3) w.r.t. $\alpha \mathbf{R}$ is a convex problem that admits a unique solution. In addition, the constraint on the trace matrix, avoiding scaling ambiguity, ensures the uniqueness of $\mathbf{R}(\boldsymbol{\mu})$. Finally, the one-to-one mapping yields a unique solution for $\boldsymbol{\mu}$.

Function (3) is a COMET type criterion that has been originally introduced in a Gaussian framework [1]. In that former Gaussian data context, taking for $\widehat{\mathbf{R}}$ the SCM, the minimization of $\mathrm{d}\left(\widehat{\mathbf{R}}, \alpha \mathbf{R}(\boldsymbol{\mu})\right)$ w.r.t. α and $\boldsymbol{\mu}$ yields an efficient estimate for the covariance matrix model $\alpha \mathbf{R}(\boldsymbol{\mu})$.

In this paper, we address the study of criterion (3) in the context of CAE observations. In this non-Gaussian context, we will show that taking Tyler's estimate $\widehat{\mathbf{R}}_{\mathrm{T}}$ in (3) for $\widehat{\mathbf{R}}$ ensures asymptotic statistical efficiency for CAE distribution: $\mathrm{d}\left(\widehat{\mathbf{R}}_{\mathrm{T}}, \alpha \mathbf{R}(\boldsymbol{\mu})\right)$ will be referred to as the RCOMET criterion. Properties of Tyler's estimate are recalled in section III-A, and we will show in III-B that minimizing $\mathrm{d}\left(\widehat{\mathbf{R}}_{\mathrm{T}}, \alpha \mathbf{R}(\boldsymbol{\mu})\right)$ w.r.t. α and $\boldsymbol{\mu}$ leads to an asymptotically efficient estimate of $\boldsymbol{\mu}_{\mathrm{e}}$ for CAE observations. In the following, $\widehat{\boldsymbol{\mu}}$ will be referred as the RCOMET estimate of $\boldsymbol{\mu}$.

A. Tyler's estimate: an overview

The unstructured ML estimate of the shape matrix is known to be Tyler's estimate $\widehat{\mathbf{R}}_{\mathrm{T}}$ [5], it maximizes (2) over the set of Hermitian positive matrices under constraint $\mathrm{Tr}\left(\widehat{\mathbf{R}}_{\mathrm{T}}\right) = m$. Assuming $N > m$, it is well known that $\widehat{\mathbf{R}}_{\mathrm{T}}$ is the unique solution of the fixed point equation

$$\widehat{\mathbf{R}}_{\mathrm{T}} = \frac{m}{N} \sum_{n=1}^{N} \frac{\mathbf{y}_n \mathbf{y}_n^H}{\mathbf{y}_n^H \widehat{\mathbf{R}}_{\mathrm{T}}^{-1} \mathbf{y}_n}, \text{ subject to } \mathrm{Tr}\left(\widehat{\mathbf{R}}_{\mathrm{T}}\right) = m \quad (4)$$

It is obtained by a simple algorithm which converges to $\widehat{\mathbf{R}}_{\mathrm{T}}$ [3, 8]. $\widehat{\mathbf{R}}_{\mathrm{T}}$ is also known to be a consistent, unbiased estimate of $\mathbf{R}_{\mathrm{e}}$. Its asymptotic distribution is related to the Complex Wishart distribution. Let $\widehat{\mathbf{R}}_{\mathrm{W}}$ be a complex Wishart matrix with $\frac{m}{m+1} N$ degrees of freedom and parameter matrix $\mathbf{R}_{\mathrm{e}}$. Then, both random matrices $\sqrt{N}\left(\dfrac{m \widehat{\mathbf{R}}_{\mathrm{T}}}{\mathrm{Tr}\left(\mathbf{R}_{\mathrm{e}}^{-1} \widehat{\mathbf{R}}_{\mathrm{T}}\right)} - \mathbf{R}_{\mathrm{e}}\right)$ and

$\sqrt{N}\left(\dfrac{m \widehat{\mathbf{R}}_W}{\mathrm{Tr}\left(\mathbf{R}_{\mathrm{e}}^{-1} \widehat{\mathbf{R}}_W\right)} - \mathbf{R}_{\mathrm{e}}\right)$ converge to the same asymptotic Gaussian distribution [8]. This result plays a central role in the proof of Theorem 2 at the next section.

B. Consistency and asymptotic efficiency of the RCOMET estimator

This section provides a statistical analysis of the RCOMET estimator $\widehat{\boldsymbol{\mu}}$ which minimizes $\mathrm{d}\left(\widehat{\mathbf{R}}_{\mathrm{T}}, \alpha \mathbf{R}(\boldsymbol{\mu})\right)$ where minimization is carried out w.r.t. $\alpha > 0$ and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_P)^T \in \mathbb{R}^P$. As already noticed, $\widehat{\boldsymbol{\mu}}$ is unique.

Theorem 1. The RCOMET estimator $\widehat{\boldsymbol{\mu}}$ is a consistent estimator of $\boldsymbol{\mu}_{\mathrm{e}}$. Likewise, $\mathbf{R}\left(\widehat{\boldsymbol{\mu}}\right)$ is a consistent estimator of $\mathbf{R}(\boldsymbol{\mu}_{\mathrm{e}})$

Proof: See Appendix A ∎

The following Lemma is needed for proving Theorem 2.

Lemma 1. Let $\mathbf{CRB}_{\mathrm{CAE}}$ be the $P \times P$ Cramér-Rao Bound (CRB) on $\boldsymbol{\mu}$ for $\mathbf{y} \sim \mathcal{U}\left(\mathbf{R}(\boldsymbol{\mu}_{\mathrm{e}})\right)$. Let $\mathbf{CRB}_{\mathrm{G}}$ be the CRB on $\boldsymbol{\mu}$ for $\mathbf{z} \sim \mathbb{CN}\left(\mathbf{0}, \alpha_{\mathrm{e}} \mathbf{R}(\boldsymbol{\mu}_{\mathrm{e}})\right)$ where α_{e}, $\boldsymbol{\mu}_{\mathrm{e}}$ are both unknown. Then

$$\mathbf{CRB}_{\mathrm{CAE}} = \frac{m+1}{m} \mathbf{CRB}_{\mathrm{G}}$$

Proof: See Appendix B ∎

Remark: note that $\boldsymbol{\alpha}_{\mathrm{e}}$ is not part of our model $\mathcal{U}\left(\mathbf{R}(\boldsymbol{\mu}_{\mathrm{e}})\right)$ while it appears in the Gaussian model $\mathbb{CN}\left(\mathbf{0}, \boldsymbol{\alpha}_{\mathrm{e}} \mathbf{R}(\boldsymbol{\mu}_{\mathrm{e}})\right)$.

Theorem 2. Let $\widehat{\boldsymbol{\mu}}$ be the RCOMET estimator of $\boldsymbol{\mu}_{\mathrm{e}}$ based on N i.i.d. observations, $\mathbf{y}_n \sim \mathcal{U}\left(\mathbf{R}(\boldsymbol{\mu}_{\mathrm{e}})\right)$. $\widehat{\boldsymbol{\mu}}$ is asymptotically Gaussian and efficient:

$$\sqrt{N}\left(\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}_{\mathrm{e}}\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\mathbf{0}, \mathbf{CRB}_{\mathrm{CAE}}\right)$$

Proof: See Appendix C ∎

IV. Application and Numerical Results

This section presents the RCOMET algorithm in the particular case of an Hermitian Toeplitz shape matrix. Simulation results are given in order to assess the statistical analysis and to compare performance with a state of the art algorithm.

A. Toeplitz structure

Let $\mathbf{R}_e = \mathbf{R}(\boldsymbol{\mu}_e) \in \mathbb{C}^{m \times m}$ belong to $\mathcal{S}$, the convex subset of Hermitian positive-definite matrices, with Toeplitz structure and trace equal to m. A natural parameterization is as follows:

$$\mathbf{R}(\boldsymbol{\mu}) = \begin{pmatrix} 1 & R_2 & \cdots & R_m \\ R_2^* & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & R_2 \\ R_m^* & \cdots & R_2^* & 1 \end{pmatrix} \text{ and } \boldsymbol{\mu} = \begin{pmatrix} \Re(R_2) \\ \Im(R_2) \\ \vdots \\ \Re(R_m) \\ \Im(R_m) \end{pmatrix} \in \mathbb{R}^{2m-2}$$

Let us introduce $\boldsymbol{\lambda} = \alpha \begin{pmatrix} 1 \\ \boldsymbol{\mu} \end{pmatrix} \in \mathbb{R}^{2m-1}$ and $\mathbf{R}'(\boldsymbol{\lambda}) = \alpha \mathbf{R}(\boldsymbol{\mu})$.

Note that there exists a matrix, $\mathbf{J} \in \mathbb{C}^{m^2 \times 2m-1}$, which relates the vectorized matrix $\mathbf{R}'(\boldsymbol{\lambda})$ to $\boldsymbol{\lambda}$ as

$$\mathbf{r}'(\boldsymbol{\lambda}) = \text{vec}\left(\mathbf{R}'(\boldsymbol{\lambda})\right) = \text{vec}\left(\alpha \mathbf{R}(\boldsymbol{\mu})\right) = \mathbf{J}\boldsymbol{\lambda}$$

Consequently, the RCOMET criterion (3) reads

$$\widehat{\boldsymbol{\lambda}} = \arg\min_{\boldsymbol{\lambda}} \left\| \widehat{\mathbf{W}}_{\mathrm{T}}^{-1/2} \widehat{\mathbf{r}}_{\mathrm{T}} - \widehat{\mathbf{W}}_{\mathrm{T}}^{-1/2} \mathbf{J}\boldsymbol{\lambda} \right\|^2$$

where $\widehat{\mathbf{r}}_{\mathrm{T}} = \text{vec}\left(\widehat{\mathbf{R}}_{\mathrm{T}}\right)$ and $\widehat{\mathbf{W}}_{\mathrm{T}} = \widehat{\mathbf{R}}_{\mathrm{T}}^T \otimes \widehat{\mathbf{R}}_{\mathrm{T}}$. The well known analytical solution gives

$$\widehat{\boldsymbol{\lambda}} = \left(\mathbf{J}^H \widehat{\mathbf{W}}_{\mathrm{T}}^{-1} \mathbf{J}\right)^{-1} \mathbf{J}^H \widehat{\mathbf{W}}_{\mathrm{T}}^{-1} \widehat{\mathbf{r}}_{\mathrm{T}} \text{ thus } \begin{cases} \widehat{\alpha} = \left[\widehat{\boldsymbol{\lambda}}\right]_1 \\ \widehat{\boldsymbol{\mu}} = \frac{1}{\widehat{\alpha}} \left[\widehat{\boldsymbol{\lambda}}\right]_{2:2m-1} \end{cases}$$

In this particular case, using lemma 1 the CRB on $\boldsymbol{\mu}$ for a single observation is easily shown to be

$$\mathbf{CRB}_{\mathrm{CAE}}^{-1} = \frac{m}{m+1}\mathbf{K}^H \mathbf{J}^H \left(\mathbf{W}_e^{-1} - \frac{\text{vec}\left(\mathbf{R}_e^{-1}\right)\text{vec}\left(\mathbf{R}_e^{-1}\right)^H}{m}\right) \mathbf{J}\mathbf{K}$$

where $\mathbf{K}$ is the identity matrix deprived of its first row.

B. Simulations

For $m = 8$, the Toeplitz shape matrix is generated from its first row according to $\{\mathbf{R}_e\}_{1,\ell} = \rho^{|\ell-1|}$, $\ell = 1, \ldots, m$ and $\rho = 0.8 + 0.3i$. We generate 1000 sets of N independent m-dimensional Compound-Gaussian distributed samples, $\mathbf{x}_n \sim \tau \mathbb{C}\mathcal{N}(\mathbf{0}, \mathbf{R}_e)$, $n = 1, \ldots, N$ [5]. These data are normalized, to obtain CAE distributed samples, $\mathbf{y}_n = \frac{\mathbf{x}_n}{\|\mathbf{x}_n\|}$, $n = 1, \ldots, N$, thereby getting rid of the random texture τ which plays no role anymore in what follows.

Our algorithm will be compared to the COCA covariance estimator proposed in [6], which consists in solving the following problem:

$$\underset{\mathbf{R} \in \mathcal{S}, d_i}{\text{minimize}} \left\| \mathbf{R} - \frac{1}{N}\sum_{i=1}^N d_i \mathbf{y}_i \mathbf{y}_i^H \right\|_{\mathrm{F}}$$

$$\text{subject to } \begin{cases} \mathbf{R} \succeq \frac{1}{m} d_i \mathbf{y}_i \mathbf{y}_i^H, \ \forall i = 1, \ldots N \\ d_i > 0, \ \forall i = 1, \ldots N \end{cases}$$

The standard semi-definite program solver, CVX, is used to compute this estimator. The algorithm proposed in [7, Algorithm 3] cannot be directly transposed to a complex-valued matrix and cannot be used here.

Performance of our RCOMET estimator and COCA are compared to the CRB via $\mathbf{MSE}(\widehat{\boldsymbol{\mu}})$ and $\text{Tr}\left(\mathbf{CRB}_{\mathrm{CAE}}\right)$.
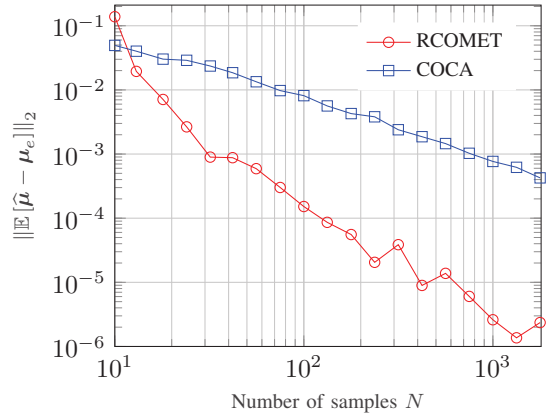

Fig. 1: Bias simulation

Fig. 1 presents the Euclidean norm of the estimated bias for $\widehat{\boldsymbol{\mu}}$ based on 1000 runs for each N. As shown previously, our RCOMET algorithm is asymptotically unbiased with a higher rate than COCA.
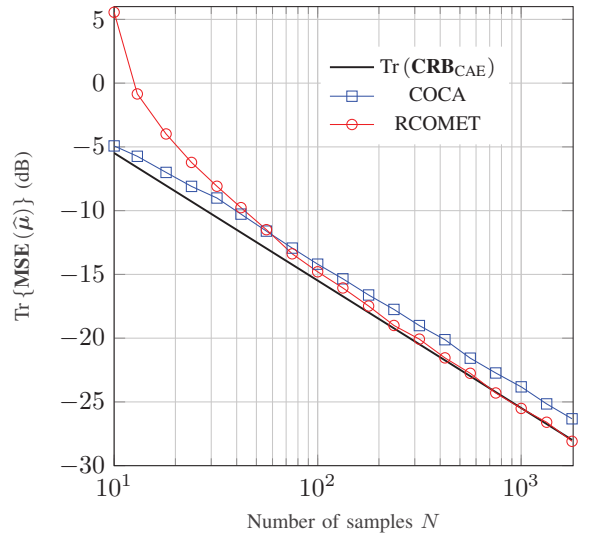

Fig. 2: Efficiency simulation

The asymptotic efficiency of RCOMET is checked on Fig. 2: it gets closer and closer to the CRB as N increases, unlike COCA estimator. However, we can note that COCA performs better at small number of samples, probably due to a bias in finite samples [6].
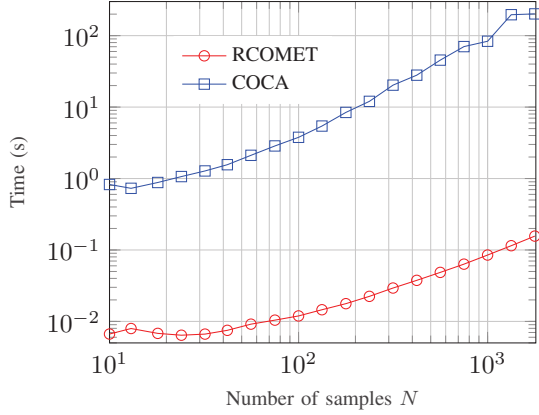

Fig. 3: Average calculation time

The average computing times are reported in Fig. 3. As pointed out in [7], the number of constraints grows linearly in N for the COCA algorithm which becomes computationally prohibitively expensive when N increases: this is not the case for our proposed RCOMET.

V. CONCLUSION

In this paper, we addressed robust structured covariance estimation for convex structures. We proposed a robust extension of the classical Gaussian COMET for CAE distributed observations. The proposed RCOMET method is consistent, asymptotically unbiased and efficient. Numerical simulations confirm the theoretical analysis and the practical interest of this approach.

APPENDIX A
PROOF OF THEOREM 1

Let $(\widehat{\alpha}, \widehat{\boldsymbol{\mu}})$ be the unique values of α and $\boldsymbol{\mu}$ which minimize $\mathrm{d}\left(\widehat{\mathbf{R}}_{\mathrm{T}}, \alpha\mathbf{R}(\boldsymbol{\mu})\right)$: $(\widehat{\alpha}, \widehat{\boldsymbol{\mu}}) = \arg\min_{\alpha, \boldsymbol{\mu}} \mathrm{d}\left(\widehat{\mathbf{R}}_{\mathrm{T}}, \alpha\mathbf{R}(\boldsymbol{\mu})\right)$. $\widehat{\boldsymbol{\mu}}$ is a function of $\widehat{\mathbf{R}}_{\mathrm{T}}$ that we denote by:

$$\widehat{\boldsymbol{\mu}} = g\left(\widehat{\mathbf{R}}_{\mathrm{T}}\right)$$

Function $g(\cdot)$ satisfies $g(\mathbf{R}_{\mathrm{e}}) = \boldsymbol{\mu}_{\mathrm{e}}$ since $\mathrm{d}(\mathbf{R}_{\mathrm{e}}, \mathbf{R}(\boldsymbol{\mu}_{\mathrm{e}})) = 0$. Moreover, for a smooth parameterization $\mathbf{R}(\boldsymbol{\mu})$, $g(\cdot)$ is differentiable and thus continuous. Then, the consistency of $\widehat{\mathbf{R}}_{\mathrm{T}}$ [8] and the continuity of g imply $\widehat{\boldsymbol{\mu}} = g\left(\widehat{\mathbf{R}}_{\mathrm{T}}\right) \xrightarrow{\mathcal{P}} g(\mathbf{R}_{\mathrm{e}}) = \boldsymbol{\mu}_{\mathrm{e}}$. Consequently, $\mathbf{R}(\widehat{\boldsymbol{\mu}}) \xrightarrow{\mathcal{P}} \mathbf{R}(\boldsymbol{\mu}_{\mathrm{e}})$.

APPENDIX B
PROOF OF LEMMA 1

For the Gaussian problem, the Slepian-Bang Formula can be directly applied for the Fisher Information Matrix (FIM):

$$[\mathbf{F}_{\boldsymbol{\mu}\boldsymbol{\mu}}]_{k,\ell} = \mathrm{Tr}\left(\mathbf{R}_{\mathrm{e}}^{-1}\partial_k\mathbf{R}\mathbf{R}_{\mathrm{e}}^{-1}\partial_\ell\mathbf{R}\right) \quad \forall\, k, \ell = 1, \ldots, P$$

$$[\mathbf{F}_{\boldsymbol{\mu}\alpha}]_k = \frac{1}{\alpha}\mathrm{Tr}\left(\mathbf{R}_{\mathrm{e}}^{-1}\partial_k\mathbf{R}\right) \text{ and } \mathbf{F}_{\alpha\alpha} = \frac{m}{\alpha^2}$$

The Schur complement gives us the CRB on $\boldsymbol{\mu}$

$$\forall\, k, \ell = 1, \ldots, P$$

$$\left[\mathbf{CRB}_{\mathrm{G}}^{-1}\right]_{k,\ell} = \mathrm{Tr}\left(\mathbf{R}_{\mathrm{e}}^{-1}\partial_k\mathbf{R}\mathbf{R}_{\mathrm{e}}^{-1}\partial_\ell\mathbf{R}\right) - \frac{\mathrm{Tr}\left(\mathbf{R}_{\mathrm{e}}^{-1}\partial_k\mathbf{R}\right)\mathrm{Tr}\left(\mathbf{R}_{\mathrm{e}}^{-1}\partial_\ell\mathbf{R}\right)}{m}$$

For the CAE problem, by following the same methodology as in [9], we obtain

$$\left[\mathbf{CRB}_{\mathrm{CAE}}^{-1}\right]_{k,\ell} = \frac{m\mathrm{Tr}\left(\mathbf{R}_{\mathrm{e}}^{-1}\partial_k\mathbf{R}\mathbf{R}_{\mathrm{e}}^{-1}\partial_\ell\mathbf{R}\right) - \mathrm{Tr}\left(\mathbf{R}_{\mathrm{e}}^{-1}\partial_k\mathbf{R}\right)\mathrm{Tr}\left(\mathbf{R}_{\mathrm{e}}^{-1}\partial_\ell\mathbf{R}\right)}{m+1}$$

$$= \frac{m}{m+1}\left[\mathbf{CRB}_{\mathrm{G}}^{-1}\right]_{k,\ell} \quad \forall\, k, \ell = 1, \ldots, P$$

Therefore $\mathbf{CRB}_{\mathrm{CAE}} = \frac{m+1}{m}\mathbf{CRB}_{\mathrm{G}}$ which proves the lemma.

APPENDIX C
PROOF OF THEOREM 2

First, remark that $\widehat{\boldsymbol{\mu}}$ is independent of the scaling factor on $\widehat{\mathbf{R}}_{\mathrm{T}}$. Indeed, if $(\widehat{\alpha}, \widehat{\boldsymbol{\mu}}) = \arg\min_{\alpha, \boldsymbol{\mu}} \mathrm{d}\left(\widehat{\mathbf{R}}_{\mathrm{T}}, \alpha\mathbf{R}(\boldsymbol{\mu})\right)$ then $(\lambda\widehat{\alpha}, \widehat{\boldsymbol{\mu}}) = \arg\min_{\alpha, \boldsymbol{\mu}} \mathrm{d}\left(\lambda\widehat{\mathbf{R}}_{\mathrm{T}}, \alpha\mathbf{R}(\boldsymbol{\mu})\right) \quad \forall\lambda \in \mathbb{R}^*$. $\widehat{\boldsymbol{\mu}} = g\left(\widehat{\mathbf{R}}_{\mathrm{T}}\right)$ is thus an homogeneous function of degree 0: $g\left(\lambda\widehat{\mathbf{R}}_{\mathrm{T}}\right) = g\left(\widehat{\mathbf{R}}_{\mathrm{T}}\right) \quad \forall\lambda \in \mathbb{R}^*$.

Let $\widehat{\mathbf{R}}_W$ be a complex Wishart matrix with $\frac{m}{m+1}N$ degrees of freedom and parameter matrix $\mathbf{R}_{\mathrm{e}}$. It is known that

$$\sqrt{N}\left(\frac{m\,\widehat{\mathbf{R}}_{\mathrm{T}}}{\mathrm{Tr}\left(\mathbf{R}_{\mathrm{e}}^{-1}\widehat{\mathbf{R}}_{\mathrm{T}}\right)} - \mathbf{R}_{\mathrm{e}}\right) \text{ and } \sqrt{N}\left(\frac{m\,\widehat{\mathbf{R}}_W}{\mathrm{Tr}\left(\mathbf{R}_{\mathrm{e}}^{-1}\widehat{\mathbf{R}}_W\right)} - \mathbf{R}_{\mathrm{e}}\right)$$

converge to the same asymptotic Gaussian distribution [8]. It follows from the Delta method [10, Chapter 3] and the homogeneity of $g(\cdot)$ that $\sqrt{N}\left(g\left(\widehat{\mathbf{R}}_{\mathrm{T}}\right) - g(\mathbf{R}_{\mathrm{e}})\right)$ and $\sqrt{N}\left(g\left(\widehat{\mathbf{R}}_W\right) - g(\mathbf{R}_{\mathrm{e}})\right)$, where $g(\mathbf{R}_{\mathrm{e}}) = \boldsymbol{\mu}_{\mathrm{e}}$, converge also to the same asymptotic Gaussian distribution. But $g\left(\widehat{\mathbf{R}}_W\right) = \widehat{\boldsymbol{\mu}}_W$ is COMET's estimate of $\boldsymbol{\mu}_{\mathrm{e}}$ based on $K = \frac{m}{m+1}N$ independant complex Gaussian samples $\mathbb{CN}\left(\mathbf{0}, \alpha_{\mathrm{e}}\mathbf{R}(\boldsymbol{\mu}_{\mathrm{e}})\right)$, and it is known [1] that $\widehat{\boldsymbol{\mu}}_W$ is an asymptotically Gaussian efficient estimator. Therefore,

$$\sqrt{K}\left(\widehat{\boldsymbol{\mu}}_W - \boldsymbol{\mu}_{\mathrm{e}}\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\mathbf{0}, \mathbf{CRB}_{\mathrm{G}}\right)$$

$$\sqrt{N}\left(\widehat{\boldsymbol{\mu}}_W - \boldsymbol{\mu}_{\mathrm{e}}\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\mathbf{0}, \frac{m+1}{m}\mathbf{CRB}_{\mathrm{G}}\right) = \mathcal{N}\left(\mathbf{0}, \mathbf{CRB}_{\mathrm{CAE}}\right)$$

$$\sqrt{N}\left(\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}_{\mathrm{e}}\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\mathbf{0}, \mathbf{CRB}_{\mathrm{CAE}}\right)$$

REFERENCES

[1] B. Ottersten, P. Stoica, and R. Roy, "Covariance matching estimation techniques for array signal processing applications," *Digital Signal Processing*, vol. 8, no. 3, pp. 185–210, 1998.

[2] R. A. Maronna, "Robust m-estimators of multivariate location and scatter," *Ann. Statist.*, vol. 4, no. 1, pp. 51–67, 01 1976.

[3] D. E. Tyler, "A distribution-free m-estimator of multivariate scatter," *The Annals of Statistics*, pp. 234–251, 1987.

[4] K. J. Sangston, F. Gini, and M. S. Greco, "Adaptive detection of radar targets in compound-gaussian clutter," in *Radar Conference (Radar-Con), 2015 IEEE*. IEEE, 2015, pp. 0587–0592.

[5] E. Ollila, D. E. Tyler, V. Koivunen, and H. V. Poor, "Complex elliptically symmetric distributions: Survey, new results and applications," *IEEE Transactions on signal processing*, vol. 60, no. 11, pp. 5597–5625, 2012.

[6] I. Soloveychik and A. Wiesel, "Tyler's covariance matrix estimator in elliptical models with convex structure," *IEEE Transactions on Signal Processing*, vol. 62, no. 20, pp. 5251–5259, 2014.

[7] Y. Sun, P. Babu, and D. P. Palomar, "Robust estimation of structured covariance matrix for heavy-tailed elliptical distributions," *IEEE Transactions on Signal Processing*, vol. 64, no. 14, pp. 3576–3590, 2016.

[8] F. Pascal, P. Forster, J.-P. Ovarlez, and P. Larzabal, "Performance analysis of covariance matrix estimates in impulsive noise," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2206–2217, 2008.

[9] O. Besson and Y. I. Abramovich, "On the fisher information matrix for multivariate elliptically contoured distributions," *IEEE Signal Processing Letters*, vol. 20, no. 11, pp. 1130–1133, 2013.

[10] A. W. Van der Vaart, *Asymptotic statistics*. Cambridge university press, 2000, vol. 3.