On the Recursions of Robust COMET Algorithm for Convexly Structured Shape Matrix

B. Mériaux*, C. Ren*, A. Breloy†, M.N. El Korso†, P. Forster‡ and J.-P. Ovarlez*§

*SONDRA, CentraleSupélec, Université Paris-Saclay, F-91190 Gif-sur-Yvette, France
†LEME EA 4416, Université Paris-Nanterre, F-92410 Ville d'Avray, France
‡SATIE, Université Paris-Nanterre, F-94230 Cachan, France
§DEMR, ONERA, Université Paris-Saclay, F-91123 Palaiseau, France

Abstract—This paper addresses robust estimation of structured shape (normalized covariance) matrices. Shape matrices most often own a particular structure depending on the application of interest and taking this structure into account improves estimation accuracy. In the framework of robust estimation, we introduce a recursive robust shape matrix estimation technique based on Tyler's M-estimate for convexly structured shape matrices. We prove that the proposed estimator is consistent, asymptotically efficient and Gaussian distributed and we notice that it reaches its asymptotic regime faster as the number of recursions increases. Finally, in the particular wide spreaded case of Hermitian persymmetric structure, we study the convergence of the recursions of the proposed algorithm.

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

I. INTRODUCTION

Most of the adaptive signal processing algorithms require a Covariance Matrix (CM) estimation. In addition to its Hermitian symmetry and positive definiteness, the CM may exhibit a particular structure related to the application of interest. For example, a linear array that is symmetrically spaced w.r.t. the phase center leads to the Hermitian persymmetric structure of the CM [1]. Another example is the Toeplitz structure for uniform linear arrays. Taking into account this structure in the estimation scheme leads to a better estimation accuracy since it decreases the degrees of freedom in the estimation problem [2]. In the Gaussian framework, this challenge has been extensively studied. Notably, the Covariance Matching Estimation Technique (COMET) has been proposed in [3]. The latter is computationally less intensive than Maximum Likelihood (ML) estimation and still provides asymptotically efficient CM estimates. However, COMET is based on the Sample Covariance Matrix (SCM) estimate, thus it is sensitive to outliers. In a context of robust CM estimation, the class of circular Complex Elliptically Symmetric distributions (CES) turns out to be particularly suitable to model spiky radar clutter measurements [4]–[6]. Within this framework, a distribution free estimator of the scatter matrix is derived in [7] and referred to as Tyler's M-estimator. Furthermore, the normalization of zero mean CES distributed data leads to

the common Complex Angular Elliptical (CAE) distribution. Several robust methods have been proposed to leverage Tyler's estimator formulation [7] in the context of structured shape matrices [8]–[13]. A COnvexly ConstrAined (COCA) shape matrix estimator has been recently proposed in [10]. Iterative Majorization-Minimization algorithms for the computation of structured CM estimates are developed in [11] and a robust extension of COMET, named RCOMET, has been derived in [13]. The references [8], [9], [12] considered the problem of robust shape matrix estimation with symmetric structures.

In this paper, we propose a Recursive version of RCOMET (R-RCOMET) based on Tyler's M-estimate and COMET criterion for convexly structured shape matrices. We conduct a theoretical analysis of the asymptotic performance of the proposed estimator. We also compare the non-asymptotic behavior with the RCOMET method. Finally, we analyse theoretically the convergence of the recursions in the particular case of the Hermitian persymmetric structure.

In the following, convergence in distribution and in probability are respectively denoted by $\xrightarrow{\mathcal{L}}$ and $\xrightarrow{\mathcal{P}}$. $\mathbf{A}^T$ (respectively $\mathbf{A}^H$ and $\mathbf{A}^*$) stands for the transpose (respectively conjugate transpose and conjugate) matrix. The vec-operator $\text{vec}(\mathbf{A})$ stacks all columns of $\mathbf{A}$ into a vector. The identity matrix of size m is referred to as $\mathbf{I}_m$. The matrix $\mathbf{J}_m$ denotes the m-dimensional antidiagonal matrix, having 1 as nonzero element. The operator $\otimes$ refers to the Kronecker matrix product and finally, the subscript "e" refers to the true value.

This paper is organized as follows. In section II, a brief review on CAE distribution, Tyler's M-estimate and RCOMET procedure is presented. Section III focuses on the proposed algorithm and its performance analysis. We also analyse the convergence of the recursions in the case of the Hermitian persymmetric structure. Some simulations results in Section IV illustrate the theoretical analysis.

II. BACKGROUND AND PROBLEM SETUP

A. Complex Angular Elliptical Distribution

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

$\mathbf{y} \sim \mathcal{U}_m(\mathbf{M})$. The probability density function of the vector $\mathbf{y}$ w.r.t. spherical measure [10] reads

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

where the matrix $\mathbf{M}$ is defined up to an arbitrary scale factor. To avoid scaling ambiguity, $\mathbf{M}$ is normalized according to $\mathrm{Tr}(\mathbf{M}) = m$. We refer to $\mathbf{M}$ as the shape matrix of $\mathbf{y}$.

B. Tyler's M-estimator

From a set of N i.i.d. CAE distributed data, $\mathbf{y}_n \sim \mathcal{U}_m(\mathbf{M})$, $n = 1, \ldots, N$ with $N > m$, Tyler's M-estimate is the unstructured ML-estimate of the shape matrix, given by the solution of the following fixed-point equation [7]:

$$\widehat{\mathbf{M}}_{\mathrm{FP}} = \frac{m}{N} \sum_{n=1}^{N} \frac{\mathbf{y}_n \mathbf{y}_n^H}{\mathbf{y}_n^H \widehat{\mathbf{M}}_{\mathrm{FP}}^{-1} \mathbf{y}_n} \triangleq \mathcal{H}\left(\widehat{\mathbf{M}}_{\mathrm{FP}}\right) \tag{2}$$

Existence and uniqueness up to a scale factor of the above equation solution have been studied in [14]. In the following, the scale ambiguity is removed by fixing in the latter solution the same constraint as for the shape matrix, i.e. $\mathrm{Tr}\left[\widehat{\mathbf{M}}_{\mathrm{FP}}\right] = m$. The solution $\widehat{\mathbf{M}}_{\mathrm{FP}}$ is obtained by an iterative algorithm, $\mathbf{M}_{k+1} = \mathcal{H}(\mathbf{M}_k)$ with the normalization on the trace, which converges to $\widehat{\mathbf{M}}_{\mathrm{FP}}$, for any initialization point [7], [15]. Furthermore, $\widehat{\mathbf{M}}_{\mathrm{FP}}$ is a consistent, unbiased estimator of $\mathbf{M}$ and its asymptotic distribution is given by [5], [15]:

$$\sqrt{N}\mathrm{vec}\left(\widehat{\mathbf{M}}_{\mathrm{FP}} - \mathbf{M}\right) \xrightarrow{d} \mathbb{CN}\left(\mathbf{0}, \boldsymbol{\Sigma}, \boldsymbol{\Sigma}\mathbf{K}_m\right) \tag{3}$$

$$\text{with} \quad \begin{cases} \boldsymbol{\Sigma} = \dfrac{m+1}{m}\boldsymbol{\Gamma}_{\mathbf{M}}\left(\mathbf{M}^T \otimes \mathbf{M}\right)\boldsymbol{\Gamma}_{\mathbf{M}}^H \\ \boldsymbol{\Gamma}_{\mathbf{M}} = \left(\mathbf{I}_{m^2} - \dfrac{1}{m}\mathrm{vec}(\mathbf{M})\,\mathrm{vec}(\mathbf{I}_m)^H\right) \end{cases}$$

where $\mathbf{K}_m$ is the commutation matrix, which satisfies $\mathbf{K}_m\mathrm{vec}(\mathbf{A}) = \mathrm{vec}\left(\mathbf{A}^T\right)$ [16].

C. Problem Setup and RCOMET Algorithm

Let us consider N i.i.d. CAE distributed observations, $\mathbf{y}_n \sim \mathcal{U}_m(\mathbf{M}_{\mathrm{e}})$, $n = 1, \ldots, N$ with $N > m$. The shape matrix belongs to the convex subset $\mathscr{S}$ of Hermitian positive-definite matrices and there exists a one-to-one differentiable mapping $\boldsymbol{\mu} \mapsto \mathcal{M}(\boldsymbol{\mu})$ from $\mathbb{R}^p$ to $\mathscr{S}$. The vector $\boldsymbol{\mu}$ is the unknown parameter of interest, with exact value $\boldsymbol{\mu}_{\mathrm{e}}$ and $\mathbf{M}_{\mathrm{e}} = \mathcal{M}(\boldsymbol{\mu}_{\mathrm{e}})$. We recall that $\mathrm{Tr}[\mathbf{M}_{\mathrm{e}}] = m$. The RCOMET estimate, $\widehat{\boldsymbol{\mu}}_0$, of $\boldsymbol{\mu}_{\mathrm{e}}$ is obtained by [13]

$$\widehat{\boldsymbol{\mu}}_0 = \arg\min_{\alpha, \boldsymbol{\mu}} \mathrm{Tr}\left[\left\{\left(\widehat{\mathbf{M}}_{\mathrm{FP}} - \alpha\mathcal{M}(\boldsymbol{\mu})\right)\widehat{\mathbf{M}}_{\mathrm{FP}}^{-1}\right\}^2\right] \tag{4}$$

$$= \arg\min_{\alpha, \boldsymbol{\mu}} \left\|\mathbf{I}_m - \widehat{\mathbf{M}}_{\mathrm{FP}}^{-1/2}[\alpha\mathcal{M}(\boldsymbol{\mu})]\widehat{\mathbf{M}}_{\mathrm{FP}}^{-1/2}\right\|_{\mathrm{F}}^2$$

with $\alpha > 0$ and satisfying $\mathrm{Tr}[\mathcal{M}(\widehat{\boldsymbol{\mu}}_0)] = m$. The minimization of the strictly convex criterion in (4) w.r.t $\alpha\mathcal{M}(\boldsymbol{\mu})$ over $\mathscr{S}$ is a convex problem that admits a unique solution. Finally, the one-to-one mapping and the constraint on the trace yield a unique solution for $\boldsymbol{\mu}$. The RCOMET estimator $\widehat{\boldsymbol{\mu}}_0$ (respectively $\mathcal{M}(\widehat{\boldsymbol{\mu}}_0)$) is a consistent estimator of $\boldsymbol{\mu}_{\mathrm{e}}$ (respectively

---

Algorithm 1 R-RCOMET

Require: N i.i.d. data, $\mathbf{y}_n \sim \mathcal{U}_m(\mathbf{M}_{\mathrm{e}})$ with $N > m$, any $K \geq 1$ given
1: Compute $\widehat{\mathbf{M}}_{\mathrm{FP}}$ from $\mathbf{y}_1, \ldots, \mathbf{y}_N$ with (2)
2: Initialize $\widehat{\boldsymbol{\mu}}_0$ with (4)
3: **for** $k = 1$ to K **do**
4:     Compute $\widehat{\boldsymbol{\mu}}_k$ from (6)
5: **end for**
6: **return** $\widehat{\boldsymbol{\mu}}_K$

$\mathcal{M}(\boldsymbol{\mu}_{\mathrm{e}})$) and asymptotically efficient and Gaussian distributed [13]. Specifically

$$\sqrt{N}(\widehat{\boldsymbol{\mu}}_0 - \boldsymbol{\mu}_{\mathrm{e}}) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{CRB}_{\mathrm{CAE}}) \tag{5}$$

where $\mathbf{CRB}_{\mathrm{CAE}}$, denoting the Cramér-Rao Bound (CRB), is detailed in [13].

Although asymptotically efficient, RCOMET algorithm requires a substantial sample support to reach its asymptotic regime. In this paper, we propose a recursive version of RCOMET, for which we conduct a theoretical analysis of its asymptotic performance; we also notice that the latter are achieved at lower sample support. Furthermore, we analyse the convergence behavior for a particular strucuture: the Hermitian persymmetric one.

III. Recursive RCOMET Procedure

A. Algorithm

In the RCOMET objective (4), $\widehat{\mathbf{M}}_{\mathrm{FP}}$ plays both the role of a target together with a metric specification through $\widehat{\mathbf{M}}_{\mathrm{FP}}^{-1}$. Splitting these roles can lead naturally to a recursive formulation where the weighting is refined at each step. For a finite and given number of steps, K, we obtain the R-RCOMET estimate for $\boldsymbol{\mu}$, denoted by $\widehat{\boldsymbol{\mu}}_K$ and achieved at the k-th iteration by solving for $k \in [\![1, K]\!]$

$$\widehat{\boldsymbol{\mu}}_k = \arg\min_{\alpha, \boldsymbol{\mu}} \mathrm{Tr}\left[\left\{\left(\widehat{\mathbf{M}}_{\mathrm{FP}} - \alpha\mathcal{M}(\boldsymbol{\mu})\right)\mathcal{M}(\widehat{\boldsymbol{\mu}}_{k-1})^{-1}\right\}^2\right] \tag{6}$$

with $\widehat{\boldsymbol{\mu}}_0$ given by (4) and such that $\mathrm{Tr}[\mathcal{M}(\widehat{\boldsymbol{\mu}}_K)] = m$. The R-RCOMET algorithm is recapped in the box Algorithm 1. In practice, we can use a more elaborated stopping rule, e.g., a combination of $k \leq K_{\max}$ and $\|\widehat{\boldsymbol{\mu}}_{k+1} - \widehat{\boldsymbol{\mu}}_k\| \leq \varepsilon_{\mathrm{tol}}\|\widehat{\boldsymbol{\mu}}_k\|$.

B. Asymptotic Analysis

First, we introduce an intermediary estimator, $\widehat{\boldsymbol{\mu}}$, of $\boldsymbol{\mu}_{\mathrm{e}}$, for which its asymptotic performance is studied. Then, we deal with the R-RCOMET asymptotic performance.

Lemma 1. Let $\widehat{\boldsymbol{\mu}}$ be the solution obtained by

$$\widehat{\boldsymbol{\mu}} = \arg\min_{\alpha, \boldsymbol{\mu}} \underbrace{\mathrm{Tr}\left[\left\{\left(\widehat{\mathbf{M}}_{\mathrm{FP}} - \alpha\mathcal{M}(\boldsymbol{\mu})\right)\widehat{\mathbf{M}}^{-1}\right\}^2\right]}_{\mathrm{d}\left(\widehat{\mathbf{M}}_{\mathrm{FP}}, \widehat{\mathbf{M}}, \alpha\mathcal{M}(\boldsymbol{\mu})\right)} \tag{7}$$

such that $\mathrm{Tr}[\mathcal{M}(\widehat{\boldsymbol{\mu}})] = m$ and where $\widehat{\mathbf{M}}$ refers to any consistent estimator of $\mathbf{M}_{\mathrm{e}}$ up to a scale factor, i.e., $\widehat{\mathbf{M}} \xrightarrow{\mathcal{P}} \kappa\mathbf{M}_{\mathrm{e}}$, $\kappa > 0$. Then,

$$\widehat{\boldsymbol{\mu}} \xrightarrow{\mathcal{P}} \boldsymbol{\mu}_{\mathrm{e}} \quad \text{and} \quad \sqrt{N}(\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}_{\mathrm{e}}) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \mathbf{CRB}_{\mathrm{CAE}})$$

Proof. The estimate $\widehat{\boldsymbol{\mu}}$ is a function of $\widehat{\mathbf{M}}_{\mathrm{FP}}$ and $\widehat{\mathbf{M}}$ that we denote by:

$$\widehat{\boldsymbol{\mu}} = g\left(\widehat{\mathbf{M}}_{\mathrm{FP}}, \widehat{\mathbf{M}}\right)$$

Function $g(\cdot)$ satisfies $g\left(\mathbf{M}_{\mathrm{e}}, \kappa\mathbf{M}_{\mathrm{e}}\right) = \boldsymbol{\mu}_{\mathrm{e}}$ since $\mathrm{d}\left(\mathbf{M}_{\mathrm{e}}, \kappa\mathbf{M}_{\mathrm{e}}, \mathcal{M}(\boldsymbol{\mu}_{\mathrm{e}})\right) = 0$ for any $\kappa > 0$. Moreover, for a smooth parameterization $\mathcal{M}(\boldsymbol{\mu})$, $g(\cdot)$ is differentiable and thus continuous. Then, the consistency of $\widehat{\mathbf{M}}_{\mathrm{FP}}$ w.r.t $\mathbf{M}_{\mathrm{e}}$ [15] (respectively $\widehat{\mathbf{M}}$ w.r.t. $\kappa\mathbf{M}_{\mathrm{e}}$, for any $\kappa > 0$) and the continuity of g imply $\widehat{\boldsymbol{\mu}} = g\left(\widehat{\mathbf{M}}_{\mathrm{FP}}, \widehat{\mathbf{M}}\right) \xrightarrow{\mathcal{P}} g(\mathbf{M}_{\mathrm{e}}, \kappa\mathbf{M}_{\mathrm{e}}) = \boldsymbol{\mu}_{\mathrm{e}}$. Consequently $\mathcal{M}(\widehat{\boldsymbol{\mu}}) \xrightarrow{\mathcal{P}} \mathcal{M}(\boldsymbol{\mu}_{\mathrm{e}})$

We can rewrite (7) as the following concentrated function

$$\widehat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \underbrace{\widehat{\boldsymbol{\eta}}_{\mathrm{FP}}^H \widehat{\mathbf{W}}^{-1/2} \prod_{\widehat{\mathbf{W}}^{-1/2}\boldsymbol{\eta}(\boldsymbol{\mu})}^{\perp} \widehat{\mathbf{W}}^{-1/2} \widehat{\boldsymbol{\eta}}_{\mathrm{FP}}}_{f_{\widehat{\mathbf{M}},\widehat{\mathbf{M}}_{\mathrm{FP}}}(\boldsymbol{\mu})} \quad (8)$$

where $\prod_{\widehat{\mathbf{W}}^{-1/2}\boldsymbol{\eta}(\boldsymbol{\mu})}^{\perp} = \mathbf{I} - \dfrac{\widehat{\mathbf{W}}^{-1/2}\boldsymbol{\eta}(\boldsymbol{\mu})\boldsymbol{\eta}(\boldsymbol{\mu})^H\widehat{\mathbf{W}}^{-1/2}}{\boldsymbol{\eta}(\boldsymbol{\mu})^H\widehat{\mathbf{W}}\boldsymbol{\eta}(\boldsymbol{\mu})}$ is the orthogonal projector onto $\widehat{\mathbf{W}}^{-1/2}\boldsymbol{\eta}(\boldsymbol{\mu})$ and with $\boldsymbol{\eta}(\boldsymbol{\mu}) = \mathrm{vec}(\mathcal{M}(\boldsymbol{\mu}))$, $\widehat{\mathbf{W}} = \widehat{\mathbf{M}}^T \otimes \widehat{\mathbf{M}}$ and $\widehat{\boldsymbol{\eta}}_{\mathrm{FP}} = \mathrm{vec}\left(\widehat{\mathbf{M}}_{\mathrm{FP}}\right)$. It follows from the Delta method [17, Chapter 3] that

$$\sqrt{N}\left(\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}_{\mathrm{e}}\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Gamma}_{\boldsymbol{\mu}}\right) \quad (9)$$

where $\boldsymbol{\Gamma}_{\boldsymbol{\mu}} = \mathbf{H}(\boldsymbol{\mu}_{\mathrm{e}})^{-1}\mathbf{R}_{\infty}\mathbf{H}(\boldsymbol{\mu}_{\mathrm{e}})^{-H}$ in which [18]

$$\begin{cases} \mathbf{H}(\boldsymbol{\mu}_{\mathrm{e}}) = \kappa^{-2} \left.\dfrac{\partial^2 f_{\mathbf{M}_{\mathrm{e}},\mathbf{M}_{\mathrm{e}}}(\boldsymbol{\mu})}{\partial\boldsymbol{\mu}\partial\boldsymbol{\mu}^T}\right|_{\boldsymbol{\mu}_{\mathrm{e}}} \\[2ex] \mathbf{R}_{\infty} = \kappa^{-4}\mathbb{E}\left[\left.\dfrac{\partial f_{\mathbf{M}_{\mathrm{e}},\widehat{\mathbf{M}}_{\mathrm{FP}}}(\boldsymbol{\mu})}{\partial\boldsymbol{\mu}}\right|_{\boldsymbol{\mu}_{\mathrm{e}}} \left(\left.\dfrac{\partial f_{\mathbf{M}_{\mathrm{e}},\widehat{\mathbf{M}}_{\mathrm{FP}}}(\boldsymbol{\mu})}{\partial\boldsymbol{\mu}}\right|_{\boldsymbol{\mu}_{\mathrm{e}}}\right)^H\right] \end{cases}$$

Thus, $\boldsymbol{\Gamma}_{\boldsymbol{\mu}}$ is independent of κ. RCOMET estimator being a particular case of the problem (7) (where $\widehat{\mathbf{M}} = \widehat{\mathbf{M}}_{\mathrm{FP}}$ is a consistent estimator of $\mathbf{M}_{\mathrm{e}}$), we finally obtain from (5) that $\boldsymbol{\Gamma}_{\boldsymbol{\mu}} = \mathbf{CRB}_{\mathrm{CAE}}$, which concludes the proof. ∎

Theorem 1. Let $\widehat{\boldsymbol{\mu}}_K$ be the R-RCOMET estimator of $\boldsymbol{\mu}_{\mathrm{e}}$ based on N i.i.d. observations, $\mathbf{y}_n \sim \mathcal{U}(\mathbf{M}_{\mathrm{e}})$. $\widehat{\boldsymbol{\mu}}_K$ is consistent w.r.t. $\boldsymbol{\mu}_{\mathrm{e}}$, asymptotically Gaussian distributed and efficient:

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

Proof. By applying Lemma 1 at each iteration, the theorem follows immediately, since $\mathcal{M}(\widehat{\boldsymbol{\mu}}_k) \xrightarrow{\mathcal{P}} \mathbf{M}_{\mathrm{e}}$ for $k \geq 0$. ∎

Finally, for a finite number of steps K, R-RCOMET yields the same asymptotic performance as RCOMET. The practical interest is that it empirically improves in most cases the performance at low sample support, which can be intuited by noting that the minimized norm is refined at each step. Notice that the fixed-point iterations are heuristic as they do not solve an underlying optimization problem when $K \to \infty$. The estimator exists for a finite number of iterations but the convergence of R-RCOMET when $K \to \infty$ requires a case by case study depending on the structure. In the following, the convergence of R-RCOMET for Hermitian persymmetric structure is established.

C. Convergence for Hermitian Persymmetric Structure

In this subsection, we consider the particular case of the Hermitian persymmetric structure, i.e., matrices which belong to the set $\left\{\mathbf{A} \in \mathbb{C}^{m\times m} | \mathbf{A} = \mathbf{A}^H \text{ and } \mathbf{A} = \mathbf{J}_m\mathbf{A}^T\mathbf{J}_m\right\}$. Let us denote $M_{r,s}$ the (r,s) element of the matrix $\mathcal{M}(\boldsymbol{\mu})$. The natural parameterization of a Hermitian persymmetric matrix, with the minimal number of parameters, consists in stacking the real and imaginary parts of the elements $M_{r,s}$ satisfying $s \geq r$ and $s \leq m + 1 - r$. The length of the vector $\boldsymbol{\mu}$ is equal to $p = \dfrac{m(m+1)}{2}$. Hence, there exists a full column rank matrix $\mathcal{J} \in \mathbb{C}^{m^2\times p}$, which relates the vectorized matrix $\mathcal{M}(\boldsymbol{\mu})$ to $\boldsymbol{\mu}$ as

$$\boldsymbol{\eta}(\boldsymbol{\mu}) = \mathrm{vec}(\mathcal{M}(\boldsymbol{\mu})) = \mathcal{J}\boldsymbol{\mu} \quad (10)$$

The full column rank matrix $\mathcal{J}$ admits a left inverse $\mathcal{J}^{\dagger} = \left(\mathcal{J}^H\mathcal{J}\right)^{-1}\mathcal{J}^H$ verifying $\mathcal{J}^{\dagger}\mathcal{J} = \mathbf{I}_p$ [19].

Proposition 1. Let $\mathbf{W} = \mathbf{A}^T \otimes \mathbf{A}$, where $\mathbf{A} \in \mathbb{C}^{m\times m}$ is Hermitian persymmetric. Then the inverse of the matrix $\mathcal{J}^H\mathbf{W}^{-1}\mathcal{J}$ is $\mathcal{J}^{\dagger}\mathbf{W}\mathcal{J}^{\dagger^H}$.

Proof. See Appendix ∎

Corollary 1. Let $\widehat{\boldsymbol{\mu}}_K$ be the R-RCOMET estimate of $\boldsymbol{\mu}_{\mathrm{e}}$. Then,

$$\forall K \in \mathbb{N}^* \quad \widehat{\boldsymbol{\mu}}_K = \mathcal{J}^{\dagger}\widehat{\boldsymbol{\eta}}_{\mathrm{FP}} \quad (11)$$

Proof. For $K = 1$, the R-RCOMET estimate reads

$$\widehat{\boldsymbol{\mu}}_1 \propto \left(\mathcal{J}^H\widehat{\mathbf{W}}_0^{-1}\mathcal{J}\right)^{-1}\mathcal{J}^H\widehat{\mathbf{W}}_0^{-1}\widehat{\boldsymbol{\eta}}_{\mathrm{FP}}$$

where $\widehat{\mathbf{W}}_0 = \mathcal{M}(\widehat{\boldsymbol{\mu}}_0)^T \otimes \mathcal{M}(\widehat{\boldsymbol{\mu}}_0)$ and $\widehat{\boldsymbol{\mu}}_0$ is given by (4). According to Proposition 1, $\widehat{\boldsymbol{\mu}}_1$ can be rewritten as

$$\begin{aligned} \widehat{\boldsymbol{\mu}}_1 &\propto \mathcal{J}^{\dagger}\widehat{\mathbf{W}}_0\mathcal{J}^{\dagger^H}\mathcal{J}^H\widehat{\mathbf{W}}_0^{-1}\widehat{\boldsymbol{\eta}}_{\mathrm{FP}} \\ &\propto \frac{1}{2}\mathcal{J}^{\dagger}\widehat{\mathbf{W}}_0\left[\mathbf{I}_{m^2} + \mathbf{J}_{m^2}\mathbf{K}_m\right]\widehat{\mathbf{W}}_0^{-1}\widehat{\boldsymbol{\eta}}_{\mathrm{FP}} \quad \text{(cf. Lemma 2)} \\ &\propto \frac{1}{2}\mathcal{J}^{\dagger}\left[\mathbf{I}_{m^2} + \mathbf{J}_{m^2}\mathbf{K}_m\right]\widehat{\boldsymbol{\eta}}_{\mathrm{FP}} = \mathcal{J}^{\dagger}\widehat{\boldsymbol{\eta}}_{\mathrm{FP}} \end{aligned}$$

Hence, to verify $\mathrm{Tr}\left[\mathcal{M}(\widehat{\boldsymbol{\mu}}_1)\right] = m$, we have necessarily

$$\widehat{\boldsymbol{\mu}}_1 = \mathcal{J}^{\dagger}\widehat{\boldsymbol{\eta}}_{\mathrm{FP}} \quad (12)$$

By recurrence, we can show that $\widehat{\boldsymbol{\mu}}_1 = \ldots = \widehat{\boldsymbol{\mu}}_K = \mathcal{J}^{\dagger}\widehat{\boldsymbol{\eta}}_{\mathrm{FP}}$ for any $K \in \mathbb{N}^*$. Indeed, let us assume $\widehat{\boldsymbol{\mu}}_K = \mathcal{J}^{\dagger}\widehat{\boldsymbol{\eta}}_{\mathrm{FP}}$, then

$$\widehat{\boldsymbol{\mu}}_{K+1} \propto \left(\mathcal{J}^H\widehat{\mathbf{W}}_K^{-1}\mathcal{J}\right)^{-1}\mathcal{J}^H\widehat{\mathbf{W}}_K^{-1}\widehat{\boldsymbol{\eta}}_{\mathrm{FP}}$$

where $\widehat{\mathbf{W}}_K = \mathcal{M}(\widehat{\boldsymbol{\mu}}_K)^T \otimes \mathcal{M}(\widehat{\boldsymbol{\mu}}_K)$. According to Proposition 1 and the trace constraint, we finally obtain $\widehat{\boldsymbol{\mu}}_{K+1} = \mathcal{J}^{\dagger}\widehat{\boldsymbol{\eta}}_{\mathrm{FP}}$. Consequently, the R-RCOMET procedure for Hermitian persymmetric matrices converges in only one step, i.e., the resulting outcome is identical for any $K \geq 1$. ∎

$\mathcal{J}^{\dagger}\widehat{\boldsymbol{\eta}}_{\mathrm{FP}}$ can be interpreted as the solution of the Euclidean projection of Tyler's M-estimate onto the subspace of Hermitian persymmetric matrices.

$$\mathcal{J}^{\dagger}\widehat{\boldsymbol{\eta}}_{\mathrm{FP}} = \arg\min_{\alpha,\boldsymbol{\mu}} \|\widehat{\boldsymbol{\eta}}_{\mathrm{FP}} - \alpha\mathcal{J}\boldsymbol{\mu}\|^2 \text{ s.t. } \mathrm{Tr}\left[\mathcal{M}(\boldsymbol{\mu})\right] = m \quad (13)$$

Consequently, surprisingly, the R-RCOMET procedure coincides with the classical Euclidean projection of Tyler's M-estimate in the case of Hermitian persymmetric matrices.

IV. NUMERICAL RESULTS

In this section, we illustrate the results of the previous theoretical analysis and compare performance with state-of-the-art algorithms.

A. Hermitian Persymmetric Structure

First, we consider the Hermitian persymmetric structure. For $m = 8$, we generate 5000 sets of N independent m-dimensional CAE distributed samples. We compare the performance of RCOMET and R-RCOMET, the Persymmetric Fixed-Point (PFP) estimate derived in [8] and the Persymmetric Sample Covariance Matrix (PSCM) estimator. The latter is obtained by substituting Tyler's M-estimate in the PFP estimator with the SCM. Finally, the related CRB, $\mathbf{CRB}_{\mathcal{U}}$, is drawn for the comparison [10], [20].
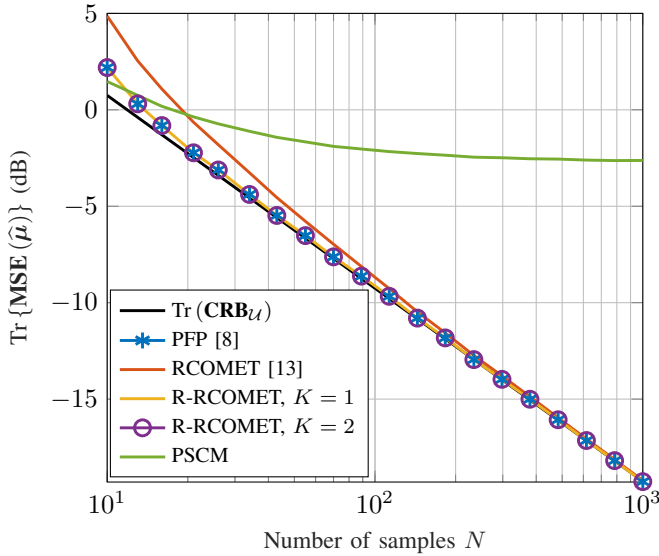

Fig. 1. Comparison on the MSE for Hermitian persymmetric structure

From Fig. 1, we notice that the R-RCOMET estimates reach the CRB as well as the RCOMET estimate, which reflects their asymptotic efficiency. Furthermore, R-RCOMET estimates are identical for 1 and 2 iterations and seems also to coincide with the PFP estimate. The PSCM estimator does not perform well since the SCM computed from CAE distributed data is the normalized SCM and is biased [15].

Remark: Actually, we can show theoretically that the R-RCOMET estimate coincides with the PFP estimate, since $\widehat{\boldsymbol{\mu}}_{\mathrm{PFP}}$ given in [8] can be expressed by

$$\widehat{\boldsymbol{\mu}}_{\mathrm{PFP}} = \mathcal{J}^{\dagger} \left(\mathbf{T}^T \otimes \mathbf{T}^H \right) \mathrm{vec} \left(\frac{1}{2} \left[\mathbf{T}\widehat{\mathbf{M}}_{\mathrm{FP}}\mathbf{T}^H + \mathbf{T}^*\widehat{\mathbf{M}}_{\mathrm{FP}}^T\mathbf{T}^T \right] \right)$$

$$= \frac{1}{2}\mathcal{J}^{\dagger} \left[\mathbf{I}_{m^2} + \mathbf{J}_{m^2}\mathbf{K}_m \right] \widehat{\boldsymbol{\eta}}_{\mathrm{FP}} = \mathcal{J}^{\dagger}\widehat{\boldsymbol{\eta}}_{\mathrm{FP}} \qquad (14)$$

where $\mathbf{T} \in \mathbb{C}^{m \times m}$ is a unitary matrix, defined in [8, Proposition 1] and verifying the relation $\mathbf{T}^T\mathbf{T} = \mathbf{J}_m$.

Again, the R-RCOMET procedure converges in one step in this particular case and it boils down to an Euclidean projection of Tyler's M-estimate for Hermitian persymmetric matrices.

B. Hermitian Toeplitz Structure

Secondly, we examine the Hermitian Toeplitz structure. In this case, the minimal parameterization consists in stacking the real and imaginary parts of the first row of the matrix $\mathcal{M}(\boldsymbol{\mu})$. Furthermore, there exists a full column rank matrix $\mathcal{J}_1 \in \mathbb{C}^{m^2 \times p}$ with $p = 2m - 1$, which relates $\mathcal{M}(\boldsymbol{\mu})$ to $\boldsymbol{\mu}$ as $\mathrm{vec}(\mathcal{M}(\boldsymbol{\mu})) = \mathcal{J}_1\boldsymbol{\mu}$. It is worth noting that the matrix $\mathcal{J}_1$ for Hermitian Toeplitz matrices differs from $\mathcal{J}$ for Hermitian persymmetric matrices since the minimal number of parameters is different on these cases. For $m = 8$, we generate 5000 sets of N independent m-dimensional CAE distributed samples. We compare the performance of RCOMET and R-RCOMET for different numbers of recursions K, the Euclidean projection of Tyler's M-estimate as well as the COCA shape estimator introduced in [10]. The standard semi-definite program solver, CVX, is used to compute this estimator [21]. The related CRB, $\mathbf{CRB}_{\mathcal{U}}$, is also drawn for the comparison [13].
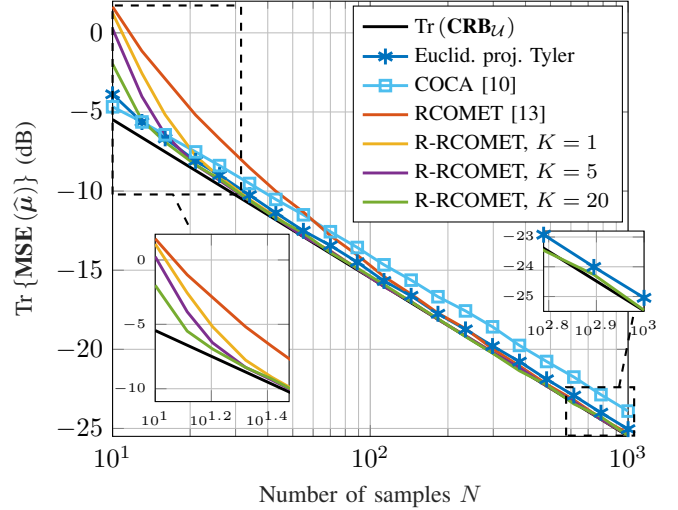

Fig. 2. Comparison on the MSE for Hermitian Toeplitz structure

From Fig. 2, we verify that the R-RCOMET estimates reach the CRB as well as the RCOMET estimate. As already stated, we observe that the CRB is reached faster with R-RCOMET than RCOMET, especially when the number of recursions, K increases. However, in this case R-RCOMET does not coincide with the Euclidean projection of Tyler's M-estimate, which is not asymptotically efficient. Finally, COCA estimator is consistent (as shown in [10]) but not asymptotically efficient. This estimator shows its interest at low sample support, however it suffers from a heavy computational cost. R-RCOMET allows for an interesting performance-computational cost trade-off in this context. Indeed, with the simulations running in Matlab R2017a on E3-1270 v5 CPU, the average calculation time

is 50.40s for COCA (respectively 0.17s for R-RCOMET) for $N = 1000$.

V. Conclusion

In this paper, we have introduced a recursive version of RCOMET for convexly structured shape matrix estimation. We have shown that the proposed R-RCOMET method is consistent, asymptotically unbiased, efficient and Gaussian distributed. In addition, the latter empirically performs better at lower sample support than RCOMET. In the particular case of the Hermitian persymmetric structure, we studied the convergence of the recursions and we related R-RCOMET to the classical Euclidean projection of Tyler's M-estimate.

VI. Appendix

This appendix presents the proof of Proposition 1, which requires the following lemma.

Lemma 2. Let $\boldsymbol{\mathcal{J}} \in \mathbb{C}^{m^2 \times p}$ be the matrix defined in (10). Its left inverse is the matrix $\boldsymbol{\mathcal{J}}^\dagger \in \mathbb{C}^{p \times m^2}$ such that $\boldsymbol{\mathcal{J}}^\dagger = \left(\boldsymbol{\mathcal{J}}^H \boldsymbol{\mathcal{J}}\right)^{-1} \boldsymbol{\mathcal{J}}^H$ and $\boldsymbol{\mathcal{J}}^\dagger \boldsymbol{\mathcal{J}} = \mathbf{I}_p$. Then, we have the equality

$$\boldsymbol{\mathcal{J}} \boldsymbol{\mathcal{J}}^\dagger = \frac{1}{2} \left(\mathbf{I}_{m^2} + \mathbf{J}_{m^2} \mathbf{K}_m\right) \tag{15}$$

Proof. Let us introduce $\mathbf{P}_{\boldsymbol{\mathcal{J}}} = \boldsymbol{\mathcal{J}} \boldsymbol{\mathcal{J}}^\dagger$ and $\mathbf{Q} = \frac{1}{2} \left(\mathbf{I}_{m^2} + \mathbf{J}_{m^2} \mathbf{K}_m\right)$. On the one hand, $\mathbf{P}_{\boldsymbol{\mathcal{J}}}$ is a projection matrix onto the image of $\boldsymbol{\mathcal{J}}$ and has a rank equal to $p = \frac{m(m+1)}{2}$. On the other hand, $\mathbf{Q}$ is also a projection matrix onto the image of $\boldsymbol{\mathcal{J}}$. Furthermore, we have $\mathbf{Q}\mathbf{P}_{\boldsymbol{\mathcal{J}}} = \mathbf{P}_{\boldsymbol{\mathcal{J}}}$. To conclude that $\mathbf{Q} = \mathbf{P}_{\boldsymbol{\mathcal{J}}}$, we need to show that $\text{rank}\,(\mathbf{Q}) = p$. Since $\mathbf{K}_m$ is full rank, we have

$$\text{rank}\,(\mathbf{Q}) = \text{rank}\,(\mathbf{Q}\mathbf{K}_m) = \text{rank}\,(\mathbf{K}_m + \mathbf{J}_{m^2})$$

Let us analyse the columns of the matrix $\mathbf{B} = \mathbf{K}_m + \mathbf{J}_{m^2} \in \mathbb{R}^{m^2 \times m^2}$. To this end, let be $\mathbf{e}_i$ the i-th column unit vector of order m^2. Thus, for $(k,\ell) \in [\![1;m]\!]^2$, we have for $i = k + (\ell - 1)m \in [\![1; m^2]\!]$

$$\mathbf{B}\mathbf{e}_i = \mathbf{e}_{\ell+(k-1)m} + \mathbf{e}_{m^2-i+1} = \mathbf{e}_{\ell+(k-1)m} + \mathbf{e}_{m(m-\ell)+m-k+1}$$

The columns of $\mathbf{B}$ have either only one non-zero coordinate, equal to 2 (m different possibilities), or only two non-zero coordinates, equals to 1 $\left(\frac{m(m-1)}{2}\text{ different possibilities}\right)$. Finally, we obtain $\text{rank}\,(\mathbf{Q}) = \text{rank}\,(\mathbf{B}) = m + \frac{m(m-1)}{2} = p$, hence the relation (15). In addition, by right multiplying (15) with $\boldsymbol{\mathcal{J}}$, we obtain $\boldsymbol{\mathcal{J}} = \mathbf{J}_{m^2} \mathbf{K}_m \boldsymbol{\mathcal{J}}$. ∎

Proof of Proposition 1. Let us introduce $\mathbf{X} = \boldsymbol{\mathcal{J}}^H \mathbf{W}^{-1} \boldsymbol{\mathcal{J}}$ and $\mathbf{Y} = \boldsymbol{\mathcal{J}}^\dagger \mathbf{W} \boldsymbol{\mathcal{J}}^{\dagger H}$. We must show that $\mathbf{X}\mathbf{Y} = \mathbf{Y}\mathbf{X} = \mathbf{I}_p$. First according to Lemma 2, we have

$$\mathbf{X}\mathbf{Y} = \boldsymbol{\mathcal{J}}^H \mathbf{W}^{-1} \boldsymbol{\mathcal{J}} \boldsymbol{\mathcal{J}}^\dagger \mathbf{W} \boldsymbol{\mathcal{J}}^{\dagger H} = \frac{1}{2} \boldsymbol{\mathcal{J}}^H \left[\mathbf{I}_{m^2} + \mathbf{W}^{-1} \mathbf{J}_{m^2} \mathbf{K}_m \mathbf{W}\right] \boldsymbol{\mathcal{J}}^{\dagger H}$$

Then, it is easy to show that the matrix $\mathbf{W}^{-1}$ is persymmetric thus $\mathbf{W}^{-1} \mathbf{J}_{m^2} \mathbf{K}_m \mathbf{W} = \mathbf{J}_{m^2} \mathbf{K}_m$. Finally, we obtain

$$\mathbf{X}\mathbf{Y} = \frac{1}{2} \boldsymbol{\mathcal{J}}^H \left[\mathbf{I}_{m^2} + \mathbf{J}_{m^2} \mathbf{K}_m\right] \boldsymbol{\mathcal{J}}^{\dagger H} = \mathbf{I}_p = \mathbf{Y}\mathbf{X}.$$

∎

References

[1] M. Haardt, M. Pesavento, F. Röemer, and M. N. El Korso, "Subspace methods and exploitation of special array structures," in *Array and Statistical Signal Processing*, ser. Academic Press Library in Signal Processing. Elsevier, Jan. 2014, vol. 3, ch. 15, pp. 651–717.

[2] T. A. Barton and D. R. Fuhrmann, "Covariance structures for multidimensional data," *Multidimensional Systems and Signal Processing*, vol. 4, no. 2, pp. 111–123, 1993.

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

[4] K. D. Ward, C. J. Baker, and S. Watts, "Maritime surveillance radar. part 1: Radar scattering from the ocean surface," *IEE Proceedings F (Communications, Radar and Signal Processing)*, vol. 137, no. 2, pp. 51–62, Apr. 1990.

[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, Nov. 2012.

[6] K. J. Sangston, F. Gini, and M. S. Greco, "Adaptive detection of radar targets in compound-Gaussian clutter," in *Proc. of IEEE Radar Conference*, May, 2015, pp. 587–592.

[7] D. E. Tyler, "A distribution-free M-estimator of multivariate scatter," *The Annals of Statistics*, vol. 15, no. 1, pp. 234–251, 1987.

[8] G. Pailloux, P. Forster, J.-P. Ovarlez, and F. Pascal, "Persymmetric adaptive radar detectors," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 4, pp. 2376–2390, Oct. 2011.

[9] Y. Gao, G. Liao, S. Zhu, X. Zhang, and D. Yang, "Persymmetric adaptive detectors in homogeneous and partially homogeneous environments," *IEEE Transactions on Signal Processing*, vol. 62, no. 2, pp. 331–342, 2014.

[10] 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, Oct. 2014.

[11] 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. 14, no. 64, pp. 3576–3590, Jul. 2016.

[12] I. Soloveychik, D. Trushin, and A. Wiesel, "Group symmetric robust covariance estimation," *IEEE Transactions on Signal Processing*, vol. 62, no. 1, pp. 244–257, Jan. 2016.

[13] B. Mériaux, C. Ren, M. N. El Korso, A. Breloy, and P. Forster, "Robust-COMET for covariance estimation in convex structures: algorithm and statistical properties," in *Proc. of IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Dec. 2017, pp. 1–5.

[14] F. Pascal, Y. Chitour, J.-P. Ovarlez, P. Forster, and P. Larzabal, "Covariance structure maximum-likelihood estimates in compound-Gaussian noise: Existence and algorithm analysis," *IEEE Transactions on Signal Processing*, vol. 56, no. 1, pp. 34–48, Jan. 2008.

[15] 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, Jun. 2008.

[16] J. R. Magnus and H. Neudecker, "The commutation matrix: some properties and applications," *The Annals of Statistics*, vol. 7, no. 2, pp. 381–394, Mar. 1979.

[17] A. W. Van der Vaart, *Asymptotic Statistics (Cambridge Series in Statistical and Probabilistic Mathematics)*. Cambridge University Press, Jun. 2000, vol. 3.

[18] J.-P. Delmas and H. Abeida, "Survey and some new results on performance analysis of complex-valued parameter estimators," *ELSEVIER Signal Processing*, vol. 111, pp. 210–221, 2015.

[19] R. Penrose, "A generalized inverse for matrices," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 51, no. 3, pp. 406–413, 1955.

[20] P. Stoica and B. C. Ng, "On the Cramér-Rao bound under parametric constraints," *IEEE Signal Processing Letters*, vol. 5, no. 7, pp. 177–179, Jul. 1998.

[21] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.