Performance analysis of subspace-based algorithms in CES data models

Jean-Pierre Delmas and Habti Abeida

Abstract Subspace-based algorithms that exploit the orthogonality between a sample subspace and a parameter-dependent subspace have proved very useful in many applications in signal processing. The statistical performance of these subspace-based algorithms depends on the deterministic and stochastic statistical model of the noisy linear mixture of the data, the estimate of the projector associated with different estimates of the scatter/covariance of the data, and the algorithm that estimates the parameters from the projector. This chapter presents various complex circular (C-CES) and non-circular (NC-CES) elliptically symmetric models of the data, as well as different associated non-robust and robust covariance estimators. These estimators include the sample covariance matrix (SCM), the maximum likelihood (ML) estimate, robust M-estimates, Tyler's M-estimate and the sample sign covariance matrix (SSCM). The asymptotic distributions of these estimators are also derived. This enables us to unify the asymptotic distribution of subspace projectors adapted to the different models of the data and demonstrate various invariance properties that have impacts on the parameters to be estimated. Particular attention is paid to the comparison between the projectors derived from Tyler's M-estimate and SSCM. Finally, this study investigates the asymptotic distributions of parameter estimates characterized by the principal subspace derived from the distributions of subspace-based parameter estimates. In particular, the efficiency with respect to the stochastic and semiparametric Cramér-Rao bound is considered.

Jean-Pierre Delmas
Samovar, Telecom SudParis, Institut Polytechnique de Paris, 91120 Palaiseau, France, e-mail: jean-pierre.delmas@it-sudparis.eu

Habti Abeida
Electrical Engineering, University of Taif, Al-Haweiah, 21974, Saudi Arabia, email: h.abeida@tu.edu.sa

1 Introduction

Noisy linear mixtures of signals in which the parameter of interest is characterized by the range space of the mixing matrix are very common in many applications, including array processing and linear system identification (see e.g., [1, 29, 38]). To get rid of the nuisance parameters, subspace-based estimates obtained by exploiting the orthogonality between a sample subspace and a parameter-dependent subspace have been exploited since the seminal paper [46] that introduces the multiple signal classification (MUSIC) algorithm for direction of arrival (DOA) estimation. These methods are always the object of active research in many applications (see e.g., [14, 30,56]), with generally many possible algorithms (see e.g., [24] for special structures of the mixing matrix). In these noisy linear mixtures, two statistical models have been commonly used [49]. If the signals in the mixture are nonrandom, but rather unknown deterministic parameters, the model is called deterministic or conditional. Otherwise, they are random and the model is a stochastic or unconditional model. Subspace-based algorithms associated with these two models have been intensely studied in the circular complex Gaussian framework (see e.g., [1, 14, 24, 29, 30, 38, 46, 49, 56] and references therein). But this framework is often insufficient for non-Gaussian heavy-tailed distributed data that are well modeled by circular (C-CES) or non-circular complex (NC-CES) elliptically symmetric distributions.

The aim of this chapter is to unify and complement various deterministic and stochastic CES models of the data and asymptotic distributions of the associated projectors derived from different estimates of the covariance matrix of the parametric noisy linear mixture data presented in the literature. The asymptotic distribution (w.r.t. the number of measurements) of the sample covariance matrix (SCM), maximum likelihood (ML), robust M, Tyler's M, and sample sign covariance matrix (SSCM) estimate of the covariance are considered. This allows us to derive the asymptotic distributions of the associated projectors, whose covariances have a unified structure, and consequently to prove several invariance properties. Particular attention is paid to the comparison between the projectors derived from Tyler's M-estimate and SSCM, yielding that the performance of Tyler's M-estimate is much better than that of SSCM, except for high dimensional data and not too small dimension of the principal subspace w.r.t. this data dimension. In that case, they are close and where the SSCM can be advantageously be used instead of Tyler's M-estimate for its lower computational complexity. Finally, asymptotic distributions of estimates of parameters characterized by the principal subspace derived from the distributions of projectors are studied, in particular, the efficiency with respect to the stochastic and semiparametric Cramér-Rao bound.

The rest of this chapter is organized as follows. Section 2 specifies the deterministic and stochastic C-CES and NC-CES distributed noisy linear mixture models, and the different associated parameterized mixing matrices. Subspace-based algorithms are interpreted in Section 3 as compositions of mappings. The asymptotic distributions of different non-robust and robust covariance estimates adapted to C-CES and NC-CES distributed data models are presented in Section 4. This allows us to derive the asymptotic distribution of the associated projectors and to prove different properties in Section 5. Asymptotic distributions of subspace-based parameter estimates are studied in Section 6. Finally, the chapter is concluded in Section 7.

The notations used in this chapter are those presented in chapter 1 to which some specific notations are used in the following. In particular $\text{RES}_m(\boldsymbol{\mu}, \mathbf{R}, g)$, C-$\text{CES}_m(\boldsymbol{\mu}, \mathbf{R}, g)$, NC-$\text{CES}_m(\boldsymbol{\mu}, \mathbf{R}, \mathbf{C}, g)$, $\mathbb{C}N_m(\boldsymbol{\mu}, \mathbf{R})$ and $\mathbb{C}N_m(\boldsymbol{\mu}, \mathbf{R}, \mathbf{C})$ denote the real elliptically symmetric (RES) distribution, C-CES, NC-CES, circular and non-circular Gaussian distributions of dimension m with finite second-order moments, respectively, where $\boldsymbol{\mu}$, $\mathbf{R}$ and $\mathbf{C}$ are the mean, the covariance and complementary covariance matrices, respectively, and g the density generator. The matrix $\mathbf{J}$ denotes the $2m \times 2m$ exchange matrix $\begin{pmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{pmatrix}$. ${}_2F_1(a, b, c, x)$ is the Gauss hypergeometric functions with ${}_2F_1(a, b, c, x) = \frac{1}{B(b, c-b)} \int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tx)^{-a} dt$ for $c > b > 0$ and $|x| < 1$ where $B(x, y)$ is defined in chapter 1. Finally, throughout the chapter, circularity is defined only up to the second order.

2 Noisy linear mixture model

Consider the following general noisy linear mixture model[1]

$$\mathbf{x}_i = \mathbf{A}\mathbf{s}_i + \mathbf{n}_i \in \mathbb{C}^m, \quad i = 1, ..., n, \tag{1}$$

where $(\mathbf{x}_i)_{i=1}^n$ are independent observations, $\mathbf{s}_i$ and $\mathbf{n}_i$ represent a signal of interest and an additive measurement noise, respectively, which are assumed to be zero-mean mutually and uncorrelated. $\mathbf{n}_i$ is assumed to be complex circular spatially uncorrelated with $\mathrm{E}(\mathbf{n}_i \mathbf{n}_i^H) = \sigma_n^2 \mathbf{I}$ and $\mathrm{E}(\mathbf{n}_i \mathbf{n}_i^T) = \mathbf{0}$.

Deterministic and stochastic parametric data models have been commonly used to model the distribution of $(\mathbf{s}_i, \mathbf{n}_i)$, where $\mathbf{n}_i$ is complex circular Gaussian distributed [49]. These two statistical data models are extended here within the framework of CES distributions.

2.1 Deterministic CES data model

In the conditional or deterministic model, the signal sequence $(\mathbf{s}_i)_{i=1,..,n}$ is conditioned from an independent zero-mean Gaussian process. As explained in [49], the sequence $(\mathbf{s}_i)_{i=1,..,n}$ is the same here in all the realizations of the random data $(\mathbf{x}_i)_{i=1,..,n}$. For complex-valued $\mathbf{s}_i$ with arbitrary circularity, we assume that $\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^n \mathbf{s}_i \mathbf{s}_i^H = \mathbf{R}_{s,\infty}$ exists and is also positive definite. The law of large number implies that:

[1] $\mathbf{x}_i$ can represent the sampled complex baseband amplitudes of the data at the output of m passband antennas.

$$\mathbf{R}_{x,\infty} \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i\mathbf{x}_i^H = \mathbf{A}\mathbf{R}_{s,\infty}\mathbf{A}^H + \sigma_n^2\mathbf{I}. \tag{2}$$

For strictly non-circular complex (also called rectilinear) valued $\mathbf{s}_i$, i.e. satisfying the condition

$$s_{i,k} = r_{i,k}e^{i\phi_k}, k = 1, .., p \text{ where } r_{i,k} \text{ are real-valued}, \tag{3}$$

and let $\mathbf{r}_i \stackrel{\text{def}}{=} (r_{i,1}, ..., r_{i,p})^T$ where $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathbf{r}_i\mathbf{r}_i^T = \mathbf{R}_{r,\infty}$ exists and is also positive definite. The phases ϕ_k associated with each k-th source are assumed fixed, but unknown during the observation. To take into account this property (3) of the signals $s_{i,k}$, we consider the extended observation $\widetilde{\mathbf{x}}_i \stackrel{\text{def}}{=} [\mathbf{x}_i^T, \mathbf{x}_i^H]^T = \widetilde{\mathbf{A}}_r\mathbf{r}_i + [\mathbf{n}_i^T, \mathbf{n}_i^H]^T$ which leads as in (2) to

$$\mathbf{R}_{\tilde{x},\infty} \stackrel{\text{def}}{=} \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^H = \widetilde{\mathbf{A}}_r\mathbf{R}_{r,\infty}\widetilde{\mathbf{A}}_r^H + \sigma_n^2\mathbf{I}, \tag{4}$$

where $\widetilde{\mathbf{A}}_r \stackrel{\text{def}}{=} \begin{bmatrix} \mathbf{A}\boldsymbol{\Delta} \\ \mathbf{A}^*\boldsymbol{\Delta}^* \end{bmatrix}$ with $\boldsymbol{\Delta} \stackrel{\text{def}}{=} \text{Diag}(e^{i\phi_1}, ..., e^{i\phi_p})$. In this deterministic model $(\mathbf{s}_i)_{i=1,..,n}$ or $(\mathbf{r}_i)_{i=1,..,n}$ and $(\phi_1, ..., \phi_p)$ are unknown deterministic parameters. However, the noise $\mathbf{n}_i$ is assumed C-CES distributed with density generator g_n. Consequently, the distribution of the observed data $\mathbf{x}_i$ is either C-CES$_m(\mathbf{A}\mathbf{s}_i, \sigma_n^2\mathbf{I}, g_n)$ or C-CES$_m(\mathbf{A}\boldsymbol{\Delta}\mathbf{r}_i, \sigma_n^2\mathbf{I}, g_n)$ distributed, for complex-valued $\mathbf{s}_i$ with arbitrary circularity or rectilinear, respectively.

2.2 Stochastic Gaussian data model

In the unconditional or stochastic model, both $\mathbf{s}_i$ and $\mathbf{n}_i$ are usually assumed Gaussian distributed and independent of each other. $\mathbf{s}_i$ is here either circular, rectilinear or non-circular and non-rectilinear complex-valued. In the complex circular case, $\mathbf{x}_i$ are $\mathbb{C}N_m(\mathbf{0}, \mathbf{R}_x)$ distributed, where the covariance $\mathbf{R}_x$ is given by

$$\mathbf{R}_x \stackrel{\text{def}}{=} \text{E}(\mathbf{x}_i\mathbf{x}_i^H) = \mathbf{A}\mathbf{R}_s\mathbf{A}^H + \sigma_n^2\mathbf{I}, \tag{5}$$

where $\mathbf{R}_s \stackrel{\text{def}}{=} \text{E}(\mathbf{s}_i\mathbf{s}_i^H)$ is positive definite.

In the complex rectilinear case, the signals $s_{i,k}$, $k = 1, .., p$, satisfy constraint (3). In this case, the distribution of $\mathbf{x}_i$ is characterized by the covariance of the extended observation $\widetilde{\mathbf{x}}_i$ given by

$$\mathbf{R}_{\tilde{x}} \stackrel{\text{def}}{=} \text{E}(\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^H) = \widetilde{\mathbf{A}}_r\mathbf{R}_r\widetilde{\mathbf{A}}_r^H + \sigma_n^2\mathbf{I}, \tag{6}$$

with $\mathbf{R}_r \stackrel{\text{def}}{=} \text{E}(\mathbf{r}_i\mathbf{r}_i^T)$, and thus, $\mathbf{x}_i$ are $\mathbb{C}N_m(\mathbf{0}, \mathbf{R}_x, \mathbf{C}_x)$ distributed with

$$\mathbf{R}_x = \mathbf{A}\boldsymbol{\Delta}\mathbf{R}_r\boldsymbol{\Delta}^*\mathbf{A}^H + \sigma_n^2\mathbf{I} \text{ and } \mathbf{C}_x = \mathbf{A}\boldsymbol{\Delta}\mathbf{R}_r\boldsymbol{\Delta}\mathbf{A}^T. \tag{7}$$

For the arbitrary non-circular case, the distribution of $\mathbf{x}_i$ is also characterized by the extended covariance matrix $\mathbf{R}_{\tilde{x}}$ given by

$$\mathbf{R}_{\tilde{x}} = \widetilde{\mathbf{A}}_c\mathbf{R}_{\tilde{s}}\widetilde{\mathbf{A}}_c^H + \sigma_n^2\mathbf{I}, \tag{8}$$

with $\mathbf{R}_{\tilde{s}} \stackrel{\text{def}}{=} \mathrm{E}(\widetilde{\mathbf{s}}_i\widetilde{\mathbf{s}}_i^H) = \begin{pmatrix} \mathbf{R}_s & \mathbf{C}_s \\ \mathbf{C}_s^* & \mathbf{R}_s^* \end{pmatrix}$, where $\widetilde{\mathbf{s}}_i \stackrel{\text{def}}{=} [\mathbf{s}_i^T, \mathbf{s}_i^H]^T$ and $\widetilde{\mathbf{A}}_c \stackrel{\text{def}}{=} \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^* \end{pmatrix}$. Thus $\mathbf{x}_i$ are $\mathbb{C}N_m(\mathbf{0}, \mathbf{R}_x, \mathbf{C}_x)$ distributed with

$$\mathbf{R}_x = \mathbf{A}\mathbf{R}_s\mathbf{A}^H + \sigma_n^2\mathbf{I} \text{ and } \mathbf{C}_x = \mathbf{A}\mathbf{C}_s\mathbf{A}^T, \tag{9}$$

where $\mathbf{C}_s \stackrel{\text{def}}{=} \mathrm{E}(\mathbf{s}_i\mathbf{s}_i^T)$ is a symmetric complex matrix.

2.3 Stochastic CES data model

A first extension of the stochastic Gaussian data model consists in modeling the independent signals $\mathbf{s}_i$ and $\mathbf{n}_i$ by CES distributions to take into account possible heavy-tailed (with respect to the Gaussian one) signals. For circular and non-circular (both rectilinear and non-rectilinear signal $\mathbf{s}_i$) cases, $\mathbf{s}_i$ is respectively C-CES$_p(\mathbf{0}, \mathbf{R}_s, g_s)$ and NC-CES$_p(\mathbf{0}, \mathbf{R}_s, \mathbf{C}_s, g_s)$ distributed[2], and $\mathbf{n}_i$ is C-CES$_m(\mathbf{0}, \sigma_n^2\mathbf{I}, g_n)$ distributed. It is worth noting here that in this stochastic data model $\mathbf{x}_i$ is not CES distributed (except for Gaussian distributions) because this family of distributions is not closed under summations.

To take advantage of robust covariance matrix estimates available in the context of CES distributions, the CES distribution has been preferred over the Gaussian distribution to model the data $\mathbf{x}_i$ in many DOA finding and beamforming processing (see e.g., [17,39,40,42]). In this case, the distributions of $\mathbf{s}_i$ and $\mathbf{n}_i$ are generally not specified, but only their second-order statistics are imposed by setting the structured covariance in (5) or extended covariance matrices in (6) and (8). More specifically, in the case of circular complex and non-circular complex signals $\mathbf{s}_i$, the observations $\mathbf{x}_i$ are C-CES$_m(\mathbf{0}, \mathbf{R}_x, g_x)$ and NC-CES$_m(\mathbf{0}, \mathbf{R}_x, \mathbf{C}_x, g_x)$ distributed, respectively. Moreover, as a particular case of this modeling, the complex compound Gaussian distribution which is a subclass of CES distributions that was used to model the clutter in radar [22] has also been used in robust and sparse M-estimation of DOA [37, eq. 1] in the form of the model:

$$\mathbf{x}_i = \mathbf{A}\mathbf{s}_i + \mathbf{n}_i \stackrel{\text{def}}{=} \sqrt{\tau_i}(\mathbf{A}\mathbf{s}'_i + \mathbf{n}'_i), \tag{10}$$

[2] Note that for rectilinear $\mathbf{s}_i$, this modeling is equivalent to $\mathbf{r}_i$ being RES$_p(\mathbf{0}, \mathbf{R}_r, g_s)$ distributed.

where the so-called texture $\tau_i > 0$ (with $\mathrm{E}(\tau_i) = 1$) is independent of $(\mathbf{s}'_i, \mathbf{n}'_i)$ which are complex Gaussian distributed. In this case $\mathbf{s}_i$ and $\mathbf{n}_i$ are uncorrelated but not independent, unlike $\mathbf{s}'_i$ and $\mathbf{n}'_i$ which are independent.

2.4 Parameterized mixing matrix

Since the complex-valued signals $\mathbf{s}_i$, can be either circular, non-circular and non-rectilinear or rectilinear signals, together with the dependence of (1) on $m \times p$ mixing matrix $\mathbf{A}$ and on the parameter of interest $\boldsymbol{\theta} \in \mathbb{R}^k$, leads us to distinguish the following two considered parameterized cases:

- For circular, and non-circular and non-rectilinear complex-valued signals $\mathbf{s}_i$, $\boldsymbol{\theta}$ is characterized by the subspace generated by the columns of the full column rank matrix $\mathbf{A}$ with $p < m$. We will use the parameterizations

$$\mathbf{B}(\boldsymbol{\theta}) \stackrel{\text{def}}{=} \mathbf{A} \tag{11}$$

  in the circular case and

$$\mathbf{B}(\boldsymbol{\theta}) \stackrel{\text{def}}{=} \widetilde{\mathbf{A}}_c = \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^* \end{pmatrix} \tag{12}$$

  in the non-circular and non-rectilinear case.
- For rectilinear complex-valued signals $\mathbf{s}_i$, $\boldsymbol{\theta}$ is characterized by the subspace generated by the columns of the full column rank $2m \times p$ extended mixing matrix

$$\mathbf{B}(\boldsymbol{\theta}) \stackrel{\text{def}}{=} \widetilde{\mathbf{A}}_r = \begin{bmatrix} \mathbf{A}\boldsymbol{\Delta} \\ \mathbf{A}^*\boldsymbol{\Delta}^* \end{bmatrix} \text{ with } p < 2m. \tag{13}$$

This low-rank signal in full-rank noise data model (1) encompasses many far or near-field, narrow or wide-band DOA models with scalar or vector-sensors for an arbitrary number of parameters per source $s_{i,k}$ (with $\mathbf{s}_i \stackrel{\text{def}}{=} (s_{i,1}, ..s_{i,k}, .., s_{i,p})^T$) and many other models as the bandlimited SISO, SIMO [38] and MIMO [1] channel models. For example, parametrization (13) can be applied for DOA estimation modeling with rectilinear or strictly second-order sources and for SIMO channels estimation modeling with BPSK or MSK symbols [13] where $\boldsymbol{\theta}$ represents both the localization parameters (azimuth, elevation, range) and the phase of the sources, and the real and imaginary parts of channel impulse response coefficients, respectively. Whereas, parametrization (12) is used for DOA modeling with generally non-circular and non-rectilinear complex sources.

3 Subspace-based estimation approaches

Since the parameter of interest $\boldsymbol{\theta}$ is characterized by the subspace generated by the columns of the full column rank matrices $\mathbf{A}$, $\widetilde{\mathbf{A}}_c$ or $\widetilde{\mathbf{A}}_r$, a simple way to get rid of the nuisance parameters, is to consider subspace-based algorithms as the following mapping:

$$(\mathbf{x}_1, .., \mathbf{x}_i, ..., \mathbf{x}_n) \longmapsto \widehat{\mathbf{R}} \overset{\text{EVD}}{\longmapsto} \widehat{\boldsymbol{\Pi}} = \boldsymbol{\Pi}(\widehat{\mathbf{R}}) \overset{\text{alg}}{\longmapsto} \widehat{\boldsymbol{\theta}} = \text{alg}(\widehat{\boldsymbol{\Pi}}), \tag{14}$$

where $\widehat{\mathbf{R}}$ can be either any estimate $\widehat{\mathbf{R}}_x$ of $\mathbf{R}_x \overset{\text{def}}{=} \text{E}(\mathbf{x}_i\mathbf{x}_i^H)$ or any estimate $\widehat{\mathbf{R}}_{\tilde{x}}$ of $\mathbf{R}_{\tilde{x}} \overset{\text{def}}{=} \text{E}(\tilde{\mathbf{x}}_i\tilde{\mathbf{x}}_i^H)$, and $\widehat{\boldsymbol{\Pi}}$ denotes the orthogonal projection matrix $\widehat{\boldsymbol{\Pi}}_x$ [resp., $\widehat{\boldsymbol{\Pi}}_{\tilde{x}}$] associated with the so-called noise subspace of $\widehat{\mathbf{R}}_x$ [resp., $\widehat{\mathbf{R}}_{\tilde{x}}$]. I.e., if $\widehat{\mathbf{R}}_x = [\widehat{\mathbf{V}}_s, \widehat{\mathbf{V}}_n]\widehat{\boldsymbol{\Lambda}}[\widehat{\mathbf{V}}_s, \widehat{\mathbf{V}}_n]^H$ denotes the EVD of $\widehat{\mathbf{R}}_x$ where $\widehat{\boldsymbol{\Lambda}}$ gathers the eigenvalues of $\widehat{\mathbf{R}}_x$ in decreasing order, $\widehat{\mathbf{V}}_s$ and $\widehat{\mathbf{V}}_n$ are $m \times p$ and $m \times (m-p)$ unitary matrices, respectively, $\widehat{\boldsymbol{\Pi}}_x$ is given by $\widehat{\mathbf{V}}_n\widehat{\mathbf{V}}_n^H$. The functional dependence $\widehat{\boldsymbol{\theta}} = \text{alg}(\widehat{\boldsymbol{\Pi}})$ constitutes an extension of the mapping

$$\boldsymbol{\Pi}(\boldsymbol{\theta}) \overset{\text{def}}{=} \mathbf{I} - \mathbf{B}(\boldsymbol{\theta})[\mathbf{B}^H(\boldsymbol{\theta})\mathbf{B}(\boldsymbol{\theta})]^{-1}\mathbf{B}^H(\boldsymbol{\theta}) \overset{\text{alg}}{\longmapsto} \boldsymbol{\theta}, \tag{15}$$

in the neighborhood of $\boldsymbol{\Pi}(\boldsymbol{\theta})$ with $\mathbf{B}(\boldsymbol{\theta})$ can either be $\mathbf{A}$, $\tilde{\mathbf{A}}_r$ or $\tilde{\mathbf{A}}_c$. Each extension $\text{alg}(.)$ specifies a particular subspace-based algorithm. Conventional MUSIC algorithm [46] based on $\widehat{\boldsymbol{\Pi}}_x$ and non-circular MUSIC algorithms [2] based on $\widehat{\boldsymbol{\Pi}}_{\tilde{x}}$ for parametrization (6) can be seen as examples in DOA estimation. Among these algorithms are the asymptotically minimum variance (AMV) algorithm or the asymptotically best consistent estimators (ABC) introduced by Porat and Friedlander [45] and Stoica *et al* [48], respectively. This algorithm minimizes the covariance matrix of the asymptotic distribution of the estimate $\widehat{\boldsymbol{\theta}}$ among all the estimates $\text{alg}(\widehat{\boldsymbol{\Pi}})$ (14) satisfying $\text{alg}(\boldsymbol{\Pi}(\boldsymbol{\theta})) = \boldsymbol{\theta}$ and a regularity condition (see (74)). The associated covariance is a lower bound that plays *the* role of benchmark against competing algorithms "alg". According to mapping (14), the statistical properties of the estimate $\widehat{\boldsymbol{\theta}}$ depends on both the choice of the covariance estimate $\widehat{\mathbf{R}}$ and that of the subspace-based algorithm "alg".

Note that direct ML approaches have been extensively studied to estimate DOA parameters from the data. But they generally require non-convex multidimensional minimizations except for a single source for which the deterministic and stochastic ML algorithms coincide with the conventional Capon algorithm (see e.g., [12, chap. 16]).

4 Asymptotic distributions of covariance estimates

Analyzing the asymptotic performance of the subspace-based estimation approaches in (14) first requires determining the asymptotic distribution of different covariance estimators $\widehat{\mathbf{R}}$ adapted to the different data models presented in Section 2.

4.1 Deterministic data model

We only consider in this model the sample covariance matrix (SCM) estimates $\widehat{\mathbf{R}} = \frac{1}{n}\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i^H$ for complex-valued of arbitrary circularity signals $\mathbf{s}_i$ and extended SCM estimates $\widehat{\mathbf{R}} = \frac{1}{n}\sum_{i=1}^n \widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^H$ for complex rectilinear-valued signals $\mathbf{s}_i$. Under finite fourth-order moments of $\mathbf{n}_i$, we get for complex-valued of arbitrary circularity signals $\mathbf{s}_i$:

$$\mathrm{E}(\mathrm{vec}(\widehat{\mathbf{R}})) = (\mathbf{A}^* \otimes \mathbf{A})[\frac{1}{n}\sum_i^n \mathbf{s}_i^* \otimes \mathbf{s}_i] + \sigma_n^2 \mathrm{vec}(\mathbf{I}), \tag{16}$$

$$\mathrm{Cov}(\mathrm{vec}(\widehat{\mathbf{R}})) = \frac{1}{n}\left([\mathbf{A}^*(\frac{1}{n}\sum_i^n \mathbf{s}_i\mathbf{s}_i^H)\mathbf{A}^T] \otimes \sigma_n^2\mathbf{I} + \sigma_n^2\mathbf{I} \otimes [\mathbf{A}(\frac{1}{n}\sum_i^n \mathbf{s}_i\mathbf{s}_i^H)\mathbf{A}^H]\right) + \frac{1}{n}\mathrm{Cov}(\mathbf{n}_i^* \otimes \mathbf{n}_i), \tag{17}$$

where $\mathrm{Cov}(\mathbf{n}_i^* \otimes \mathbf{n}_i) = \sigma_n^4[(1+\kappa_n)\mathbf{I} + \kappa_n \mathrm{vec}(\mathbf{I})\mathrm{vec}^T(\mathbf{I})]$ with

$$\kappa_n = \frac{\mathrm{E}(\mathcal{Q}_n^2)}{m(m+1)} - 1 \tag{18}$$

is the kurtosis parameter[3] of $\mathbf{n}_i$ and $\mathcal{Q}_n$ is its 2nd-order modular variate. Using the Liapounov central limit theorem (CLT) for independent non identically distributed r.v. $\mathbf{x}_i^* \otimes \mathbf{x}_i$ (see e.g. [31, Th. 2.7.1]) and the Slutsky theorem (see e.g. [31, Th. 5.1.6]), we get the following convergence in distribution [3], [5].

$$\sqrt{n}(\mathrm{vec}(\widehat{\mathbf{R}}) - \mathrm{vec}(\mathbf{R})) \to_d \mathbb{C}N_{m^2}(\mathbf{0}, \mathbf{R}_{r_x}, \mathbf{R}_{r_x}\mathbf{K}), \tag{19}$$

with $\mathbf{R} \stackrel{\mathrm{def}}{=} \mathbf{R}_{x,\infty}$ defined in (2), $\mathbf{K}$ defined in chapter 1 and

$$\mathbf{R}_{r_x} = \mathbf{A}^*\mathbf{R}_{s,\infty}\mathbf{A}^T \otimes \sigma_n^2\mathbf{I} + \sigma_n^2\mathbf{I} \otimes \mathbf{A}\mathbf{R}_{s,\infty}\mathbf{A}^H + \sigma_n^4[(1+\kappa_n)\mathbf{I} + \kappa_n\mathrm{vec}(\mathbf{I})\mathrm{vec}^T(\mathbf{I})], \tag{20}$$

[3] Note that for C and NC-CES$_m$ distributions, the kurtosis parameter $\kappa_{c,m}$ is defined by the kurtosis parameter of the associated RES$_{2m}$ distribution $\kappa_{r,2m} = \frac{\mathrm{E}(\mathcal{Q}_{r,2m}^2)}{2m(2m+2)} - 1 = \frac{\mathrm{E}(\mathcal{Q}_{c,m}^2)}{m(m+1)} - 1$ with $\mathcal{Q}_{c,m} = \frac{1}{2}\mathcal{Q}_{r,2m}$ where $\mathcal{Q}_{c,m}$ and $\mathcal{Q}_{r,2m}$ are the 2nd-order modular variates of the associated C and NC-CES$_m$, and RES$_{2m}$ distributions (see chapter 1).

where $\mathbf{R}_{s,\infty}$ is defined in Section 2.1. Similarly, we obtain for complex rectilinear-valued $\mathbf{s}_i$ that

$$\sqrt{n}(\mathrm{vec}(\widehat{\mathbf{R}}) - \mathrm{vec}(\mathbf{R})) \to_d \mathbb{C}N_{4m^2}(\mathbf{0}, \mathbf{R}_{r_{\tilde{x}}}, \mathbf{R}_{r_{\tilde{x}}}\mathbf{K}) \tag{21}$$

with $\mathbf{R} \stackrel{\text{def}}{=} \mathbf{R}_{\tilde{x},\infty}$ defined in (4) and

$$\begin{aligned}\mathbf{R}_{r_{\tilde{x}}} &= [\mathbf{I} + \mathbf{K}(\mathbf{J}\otimes\mathbf{J})][(\widetilde{\mathbf{A}}_r^*\mathbf{R}_{r,\infty}\widetilde{\mathbf{A}}_r^T \otimes \sigma_n^2\mathbf{I}) + (\sigma_n^2\mathbf{I}\otimes\widetilde{\mathbf{A}}_r\mathbf{R}_{r,\infty}\widetilde{\mathbf{A}}_r^H) \\ &\quad + \sigma_n^4(1+\kappa_n)\mathbf{I}] + \sigma_n^4\kappa_n\mathrm{vec}(\mathbf{I})\mathrm{vec}^T(\mathbf{I}).\end{aligned} \tag{22}$$

4.2 Stochastic data model

We consider here two cases. In the first one, $\mathbf{s}_i$ and $\mathbf{n}_i$ are both CES distributed for which only the SCM estimate is considered. In this case, $\mathbf{x}_i$ is not CES distributed. In the second one, $\mathbf{x}_i$ is CES distributed and many robust covariance estimates are considered.

4.2.1 SCM estimators for both CES distributed $\mathbf{s}_i$ and $\mathbf{n}_i$

Let's start first with the SCM estimates $\widehat{\mathbf{R}} = \frac{1}{n}\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i^H$ and $\widehat{\mathbf{R}} = \frac{1}{n}\sum_{i=1}^n \widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^H$ for respectively circular and complex of arbitrary circularity signals $\mathbf{s}_i$, where $\mathbf{s}_i$ and $\mathbf{n}_i$ are both CES distributed and have finite fourth-order moments. By applying the classic CLT to the r.v. $\mathbf{x}_i^*\otimes\mathbf{x}_i$ and $\widetilde{\mathbf{x}}_i^*\otimes\widetilde{\mathbf{x}}_i$, we get, respectively, the following asymptotic distribution of $\widehat{\mathbf{R}}$ which did not appear in the literature:

$$\sqrt{n}(\mathrm{vec}(\widehat{\mathbf{R}}) - \mathrm{vec}(\mathbf{R})) \to_d \mathbb{C}N_{m^2}(\mathbf{0}, \mathbf{R}_{r_x}, \mathbf{R}_{r_x}\mathbf{K}), \tag{23}$$

$$\sqrt{n}(\mathrm{vec}(\widehat{\mathbf{R}}) - \mathrm{vec}(\mathbf{R})) \to_d \mathbb{C}N_{4m^2}(\mathbf{0}, \mathbf{R}_{r_{\tilde{x}}}, \mathbf{R}_{r_{\tilde{x}}}\mathbf{K}), \tag{24}$$

with $\mathbf{R} \stackrel{\text{def}}{=} \mathbf{R}_x$ defined in (5) and $\mathbf{R} \stackrel{\text{def}}{=} \mathbf{R}_{\tilde{x}}$ defined in (6) and (8), respectively, and

$$\begin{aligned}\mathbf{R}_{r_x} &= (\mathbf{R}_x^*\otimes\mathbf{R}_x) + \kappa_s[(\mathbf{A}^*\mathbf{R}_s^*\mathbf{A}^T)\otimes(\mathbf{A}\mathbf{R}_s\mathbf{A}^H) \\ &\quad + \mathrm{vec}(\mathbf{A}\mathbf{R}_s\mathbf{A}^H)\mathrm{vec}^H(\mathbf{A}\mathbf{R}_s\mathbf{A}^H)] + \sigma_n^4\kappa_n[\mathbf{I} + \mathrm{vec}(\mathbf{I})\mathrm{vec}^T(\mathbf{I})],\end{aligned} \tag{25}$$

$$\begin{aligned}\mathbf{R}_{r_{\tilde{x}}} &= [\mathbf{I} + \mathbf{K}(\mathbf{J}\otimes\mathbf{J})][(\mathbf{R}_{\tilde{x}}^*\otimes\mathbf{R}_{\tilde{x}}) + \kappa_s(\widetilde{\mathbf{A}}_c^*\mathbf{R}_{\tilde{s}}^*\widetilde{\mathbf{A}}_c^T)\otimes(\widetilde{\mathbf{A}}_c\mathbf{R}_{\tilde{s}}\widetilde{\mathbf{A}}_c^H) \\ &\quad + \kappa_s\mathrm{vec}(\widetilde{\mathbf{A}}_c\mathbf{R}_{\tilde{s}}\widetilde{\mathbf{A}}_c^H)\mathrm{vec}^H(\widetilde{\mathbf{A}}_c\mathbf{R}_{\tilde{s}}\widetilde{\mathbf{A}}_c^H) + \sigma_n^4\kappa_n\mathbf{I}] \\ &\quad + \sigma_n^4\kappa_n\mathrm{vec}(\mathbf{I})\mathrm{vec}^T(\mathbf{I}),\end{aligned} \tag{26}$$

where κ_s is the kurtosis parameter of $\mathbf{s}_i$ given by $\kappa_s = \frac{\mathrm{E}(\mathcal{Q}_s^2)}{p(p+1)} - 1$ and where $\mathcal{Q}_s$ is the 2nd-order modular variate of $\mathbf{s}_i$. Note that (24) and (26) remain valid for complex rectilinear signals $\mathbf{s}_i$ if κ_s is replaced by $\kappa_r = \frac{\mathrm{E}(\mathcal{Q}_r^2)}{p(p+2)} - 1$ where $\mathcal{Q}_r$ is the 2nd-order modular variate of $\mathbf{r}_i$ which is $\text{RES}_p(\mathbf{0}, \mathbf{R}_r, g_s)$ distributed. Furthermore, in this case $\widetilde{\mathbf{A}}_c \mathbf{R}_{\tilde{s}} \widetilde{\mathbf{A}}_c^H$ reduces to $\widetilde{\mathbf{A}}_r \mathbf{R}_r \widetilde{\mathbf{A}}_r^H$.

4.2.2 Covariance estimates for CES distributed $\mathbf{x}_i$

Now consider the cases where the observations $\mathbf{x}_i$ are $\text{C-CES}_m(\mathbf{0}, \mathbf{R}_x, g_x)$ and $\text{NC-CES}_m(\mathbf{0}, \mathbf{R}_x, \mathbf{C}_x, g_x)$ distributed. For these distributions, many covariance estimators have been proposed in the literature. This study examines the asymptotic distributions of various estimators, including the SCM, ML and M, Tyler's and SSCM estimators.

4.2.2.1 SCM estimators

Under finite fourth-order moments of $\mathbf{x}_i$, applying again the classic CLT to the r.v. $\mathbf{x}_i^* \otimes \mathbf{x}_i$ and $\widetilde{\mathbf{x}}_i^* \otimes \widetilde{\mathbf{x}}_i$, we get the convergences in distribution (23) and (24) for the SCM estimate $\widehat{\mathbf{R}}$ with also $\mathbf{R} \stackrel{\text{def}}{=} \mathbf{R}_x$ defined in (5) and $\mathbf{R} \stackrel{\text{def}}{=} \mathbf{R}_{\tilde{x}}$ defined in (6) and (8), respectively, and with now $\mathbf{R}_{r_x}$ and $\mathbf{R}_{r_{\tilde{x}}}$ are given by [3], [5]:

$$\mathbf{R}_{r_x} = \sigma_1(\mathbf{R}_x^* \otimes \mathbf{R}_x) + \sigma_2 \text{vec}(\mathbf{R}_x)\text{vec}^H(\mathbf{R}_x), \tag{27}$$

$$\mathbf{R}_{r_{\tilde{x}}} = \sigma_1[\mathbf{I} + \mathbf{K}(\mathbf{J} \otimes \mathbf{J})](\mathbf{R}_{\tilde{x}}^* \otimes \mathbf{R}_{\tilde{x}}) + \sigma_2 \text{vec}(\mathbf{R}_{\tilde{x}})\text{vec}^H(\mathbf{R}_{\tilde{x}}), \tag{28}$$

with $\sigma_1 = 1 + \kappa_x$ and $\sigma_2 = \kappa_x$, where κ_x is the kurtosis parameter of $\mathbf{x}_i$ given by $\kappa_x = \frac{\text{E}(\mathcal{Q}_x^2)}{m(m+1)} - 1$ where $\mathcal{Q}_x$ is the 2nd-order modular variate of $\mathbf{x}_i$. Note that thanks to the linear one to one mapping $\bar{\mathbf{x}}_i \longleftrightarrow \widetilde{\mathbf{x}}_i$ with $\bar{\mathbf{x}}_i \stackrel{\text{def}}{=} (\text{Re}(\mathbf{x}_i)^T, \text{Im}(\mathbf{x}_i)^T)^T \in \mathbb{R}^{2m}$, (27) and (28) can be directly deduced from the convergence in distribution of the SCM estimate $\frac{1}{n}\sum_{i=1}^n \bar{\mathbf{x}}_i \bar{\mathbf{x}}_i^T$ of $\mathbf{R}_{\bar{x}} \stackrel{\text{def}}{=} \text{E}(\bar{\mathbf{x}}_i \bar{\mathbf{x}}_i^T)$, given by [43, p. 5]

$$\sqrt{n}(\text{vec}(\widehat{\mathbf{R}}) - \text{vec}(\mathbf{R}_{\bar{x}})) \to_d \mathbb{R}N_{4m^2}(\mathbf{0}, \mathbf{R}_{r_{\bar{x}}}), \tag{29}$$

with

$$\mathbf{R}_{r_{\bar{x}}} = (1 + \kappa_{\bar{x}})(\mathbf{I} + \mathbf{K})(\mathbf{R}_{\bar{x}} \otimes \mathbf{R}_{\bar{x}}) + \kappa_{\bar{x}} \text{vec}(\mathbf{R}_{\bar{x}})\text{vec}^T(\mathbf{R}_{\bar{x}}), \tag{30}$$

where $\kappa_{\bar{x}} = \kappa_x$ by definition of the kurtosis κ_x of the complex-valued r.v. $\mathbf{x}_i$ given in footnote 4. For heavier tails than Gaussian distributions, we have $\kappa_x > 0$ and can be very large without any upper bound, and consequently, the SCM estimator can be a very poor estimator of the covariance matrix.

4.2.2.2 ML estimators

To take into account the particular CES distribution of $\mathbf{x}_i$, the ML estimator is often considered as the benchmark estimator. From the p.d.f. of the complex-valued data $\mathbf{x}_i$: $p(\mathbf{x}) = |\mathbf{R}_{\tilde{x}}|^{-1/2} g_x(\frac{1}{2}\widetilde{\mathbf{x}}^H \mathbf{R}_{\tilde{x}}^{-1} \widetilde{\mathbf{x}})$ which reduces to $p(\mathbf{x}) = |\mathbf{R}_x|^{-1} g_x(\mathbf{x}^H \mathbf{R}_x^{-1} \mathbf{x})$ in the circular case (see chapter 1), the ML estimate $\widehat{\mathbf{R}}$ of $\mathbf{R}_x$ and $\mathbf{R}_{\tilde{x}}$ are respectively solutions of the implicit equations

$$\widehat{\mathbf{R}} = \frac{1}{n}\sum_{i=1}^{n} \varphi_x(\mathbf{x}_i^H \widehat{\mathbf{R}}^{-1}\mathbf{x}_i)\mathbf{x}_i\mathbf{x}_i^H \text{ and } \widehat{\mathbf{R}} = \frac{1}{n}\sum_{i=1}^{n} \varphi_x\left(\frac{1}{2}\widetilde{\mathbf{x}}_i^H \widehat{\mathbf{R}}^{-1}\widetilde{\mathbf{x}}_i\right)\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^H, \tag{31}$$

with $\varphi_x(t) \stackrel{\text{def}}{=} -\frac{1}{g_x(t)}\frac{dg_x(t)}{dt}$. Under existence, uniqueness and usual regularity conditions, the ML estimate of $\mathbf{R}_x$ and $\mathbf{R}_{\tilde{x}}$ are generally asymptotically efficient with a speed of convergence in $\sqrt{n}$. Thus, we get the convergences in distribution (23) and (24), with also $\mathbf{R} \stackrel{\text{def}}{=} \mathbf{R}_x$ defined in (5) and $\mathbf{R} \stackrel{\text{def}}{=} \mathbf{R}_{\tilde{x}}$ defined in (6) and (8), respectively, where $\mathbf{R}_{r_x}$ and $\mathbf{R}_{r_{\tilde{x}}}$ can be deduced from the matrix Slepian-Bangs formula [4, 7, 9, 23] associated with the single parameters $\mathbf{R}_x$ and $\mathbf{R}_{\tilde{x}}$ and are still given by (27) and (28) with

$$\sigma_1 = \frac{m(m+1)}{\mathrm{E}[\mathcal{Q}_x^2 \varphi_x^2(\mathcal{Q}_x)]} \text{ and } \sigma_2 = -\frac{2\sigma_1(1-\sigma_1)}{1+2m(1-\sigma_1)}. \tag{32}$$

4.2.2.3 M estimator

Since the ML estimator may be drastically affected by the presence of outliers or when the data distribution deviates slightly from the CES distribution of the model, robust estimators of the covariance of the data have been proposed. Among the different families of robust estimators, we focus our attention on the class of M-estimators introduced by Maronna [36] for RES distributions. Applied to $\mathrm{RES}_{2m}(\mathbf{0}, \mathbf{R}_{\bar{x}}, g_{\bar{x}})$ distributed data $\bar{\mathbf{x}}_i$, with $g_{\bar{x}}(t) = 2^{-m} g_x(\frac{t}{2})$ (see Chapter 1), these M-estimators were defined as solutions $\widehat{\mathbf{R}}$ of the implicit equation

$$\widehat{\mathbf{R}} = \frac{1}{n}\sum_{i=1}^{n} u_{\bar{x}}(\bar{\mathbf{x}}_i^T \widehat{\mathbf{R}}^{-1}\bar{\mathbf{x}}_i)\bar{\mathbf{x}}_i\bar{\mathbf{x}}_i^T, \tag{33}$$

where $u_{\bar{x}}(.)$ is any real-valued weight function on $[0,\infty)$ not related to a particular RES distribution. Under sufficient conditions (called Maronna conditions), it was proved in [36, Th. 4], the existence and uniqueness of solution $\widehat{\mathbf{R}}$ of (33). It has also been proved in [36, Th. 2] the existence and uniqueness of the solution $\mathbf{V}$ of

$$\mathbf{V} = \mathrm{E}[u_{\bar{x}}(\bar{\mathbf{x}}_i^T \mathbf{V}^{-1}\bar{\mathbf{x}}_i)\bar{\mathbf{x}}_i\bar{\mathbf{x}}_i^T]. \tag{34}$$

Sufficient conditions are also given in [36, Th. 5] to ensure the strong consistency of the estimate $\widehat{\mathbf{R}}$ solution of (33) to the solution $\mathbf{V}$ of (34). $\mathbf{V}$ is related to $\mathbf{R}_{\bar{x}}$ by $\mathbf{V} = c^{-1}\mathbf{R}_{\bar{x}}$ where c is the unique solution of $\mathrm{E}[c\mathcal{Q}_{\bar{x}} u_{\bar{x}}(c\mathcal{Q}_{\bar{x}})] = 2m$. [36, Appendix 3]. This is equivalent to the solution of

$$\mathrm{E}[c\mathcal{Q}_x u_x(c\mathcal{Q}_x)] = m, \tag{35}$$

for CES distributions [42, (rel (46)]. Using a general result on M-estimators given in [25, Sec. 4], Maronna proved in [36, Th. 6] the asymptotic gaussianity of $\widehat{\mathbf{R}}$. Then, using the affine invariance property of any M-estimators and the general structure of the covariance of radial random matrices, the covariance of the asymptotic distribution (called also asymptotic error covariance) of $\widehat{\mathbf{R}}$ was specified in [50, Appendix 2].

Extensions of M-estimators to C-CES distributed data were introduced in [39] and later studied and used in various signal processing application (see [42] and references therein). Then, it was extended to NC-CES distributed data [5]. Since $\mathbf{x}_i$ are NC-CES$_m(\mathbf{0}, \mathbf{R}_x, \mathbf{C}_x, g_x)$ distributed by definition if $\bar{\mathbf{x}} = \frac{1}{\sqrt{2}}\mathbf{M}^H\widetilde{\mathbf{x}}_i$ (where $\mathbf{M} \stackrel{\text{def}}{=} \frac{1}{\sqrt{2}}\begin{pmatrix} \mathbf{I} & i\mathbf{I} \\ \mathbf{I} & -i\mathbf{I} \end{pmatrix}$ is unitary) are RES$_{2m}(\mathbf{0}, \mathbf{R}_{\bar{x}}, g_{\bar{x}})$ distributed, (33) is equivalent to (36) where $u_x(t) \stackrel{\text{def}}{=} u_{\bar{x}}(\frac{t}{2})$ (see Chapter 1) and where $\widehat{\mathbf{R}}$ now denotes the M-estimate of $\mathbf{R}_{\tilde{x}} = 2\mathbf{M}\mathbf{R}_{\bar{x}}\mathbf{M}^H$:

$$\widehat{\mathbf{R}} = \frac{1}{n}\sum_{i=1}^{n} u_x\left(\frac{1}{2}\widetilde{\mathbf{x}}_i^H\widehat{\mathbf{R}}^{-1}\widetilde{\mathbf{x}}_i\right)\widetilde{\mathbf{x}}_i\widetilde{\mathbf{x}}_i^H. \tag{36}$$

Moreover in the particular complex circular case where $\mathbf{R}_{\tilde{x}} = \begin{pmatrix} \mathbf{R}_x & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_x^* \end{pmatrix}$, imposing $\widehat{\mathbf{R}}$ in (36) to be block diagonal structured as $\mathbf{R}_{\tilde{x}}$, the M-estimate $\widehat{\mathbf{R}}$ of $\mathbf{R}_x$ introduced in [39] is derived as solution of:

$$\widehat{\mathbf{R}} = \frac{1}{n}\sum_{i=1}^{n} u_x(\mathbf{x}_i^H\widehat{\mathbf{R}}^{-1}\mathbf{x}_i)\mathbf{x}_i\mathbf{x}_i^H. \tag{37}$$

Thus, all the properties of the M-estimate $\widehat{\mathbf{R}}$ of $\mathbf{R}_{\bar{x}}$ are transferred to the M-estimates $\widehat{\mathbf{R}}$ of $\mathbf{R}_x$ and $\mathbf{R}_{\tilde{x}}$ for C-CES and NC-CES [5] distributions, respectively. In particular, we get the convergences in distribution (23) and (24) where $\mathbf{R} \stackrel{\text{def}}{=} c^{-1}\mathbf{R}_x$ in the circular case and $\mathbf{R} \stackrel{\text{def}}{=} c^{-1}\mathbf{R}_{\tilde{x}}$ in the complex of arbitrary circularity case, and where $\mathbf{R}_{r_x}$ and $\mathbf{R}_{r_{\tilde{x}}}$ are given by (27) and (28), respectively, with [42], [35]

$$\sigma_1 = \frac{\mathrm{E}[\mathcal{Q}_x^2 u_x^2(c\mathcal{Q}_x)]}{m(m+1)(1+[m(m+1]^{-1}c_u)^2} \tag{38}$$

and

$$\sigma_2 = \frac{\mathrm{E}[(\mathcal{Q}_x u_x(c\mathcal{Q}_x) - mc_u)^2]}{(m+c_u)^2} - \frac{\sigma_1}{m} \tag{39}$$

with $c_u \stackrel{\text{def}}{=} \mathrm{E}[c^2\mathcal{Q}_x^2 u_x^{'}(c\mathcal{Q}_x)]$ with $u_x^{'}(t) \stackrel{\text{def}}{=} du_x(t)/dt$. Note that for $u_x(t) = \varphi_x(t)$, the M-estimate gives the ML estimate for which $c = 1$. Using the identity $\mathrm{E}[\mathcal{Q}_x^2\varphi_x^{'}(\mathcal{Q}_x)] = \mathrm{E}[\mathcal{Q}_x^2\varphi_x^2(\mathcal{Q}_x)] - m(m+1)$ proved in [5, Appendix] with $\varphi_x^{'}(t) \stackrel{\text{def}}{=} d\varphi_x(t)/dt$, (38)and (39) reduce to (32).

4.2.2.4 Tyler's M estimator

Tyler's M estimator was introduced in [51] for RES distributions as a solution of (33) with the weight function $u_{\bar{x}}(t) = \frac{2m}{t}$. It became very popular in signal processing applications because it enjoys important properties among which to have a distribution that does not depend on the RES distribution of the data. Extensions of this estimator to C-CES [42], [44] and to NC-CES distributions [5] are straightforward using the real $\bar{\mathbf{x}}_i \in \mathbb{R}^{2m}$ representation of CES distributed $\mathbf{x}_i \in \mathbb{C}^m$. They are given by the solution $\widehat{\mathbf{R}}$ of the implicit equation

$$\widehat{\mathbf{R}} = \frac{2m}{n} \sum_{i=1}^{n} \frac{\widetilde{\mathbf{x}}_i \widetilde{\mathbf{x}}_i^H}{\widetilde{\mathbf{x}}_i^H \widehat{\mathbf{R}}^{-1} \widetilde{\mathbf{x}}_i} \tag{40}$$

constrained to $\mathrm{Tr}(\widehat{\mathbf{R}}) = 2m$ for NC-CES distributed $\mathbf{x}_i$, which reduces to the solution $\widehat{\mathbf{R}}$ of

$$\widehat{\mathbf{R}} = \frac{m}{n} \sum_{i=1}^{n} \frac{\mathbf{x}_i \mathbf{x}_i^H}{\mathbf{x}_i^H \widehat{\mathbf{R}}^{-1} \mathbf{x}_i} \tag{41}$$

constrained to $\mathrm{Tr}(\widehat{\mathbf{R}}) = m$ for C-CES distributed $\mathbf{x}_i$. Here too, all the properties of the Tyler's M-estimate of $\mathbf{R}_{\bar{x}}$ are transferred to the Tyler's M-estimates of $\mathbf{R}_x$ and $\mathbf{R}_{\tilde{x}}$ for C-CES and NC-CES distributions, respectively. In particular, we get the convergences in distribution (23) and (24) where $\mathbf{R} \stackrel{\text{def}}{=} \frac{m}{\mathrm{Tr}(\mathbf{R}_x)} \mathbf{R}_x$ in the complex circular case and $\mathbf{R} \stackrel{\text{def}}{=} \frac{m}{\mathrm{Tr}(\mathbf{R}_x)} \mathbf{R}_{\tilde{x}}$ in the complex of arbitrary circularity case, and where $\mathbf{R}_{r_x}$ and $\mathbf{R}_{r_{\tilde{x}}}$ are given by (27) and (28), respectively, with [51]:

$$\sigma_1 = \left(\frac{m}{\mathrm{Tr}(\mathbf{R}_x)}\right)^2 \left(1 + \frac{1}{m}\right) \quad \text{and} \quad \sigma_2 = -\left(\frac{m}{\mathrm{Tr}(\mathbf{R}_x)}\right)^2 \frac{1}{m} \left(1 + \frac{1}{m}\right). \tag{42}$$

4.2.2.5 SSCM estimator

The SSCM estimator is another distribution-free estimator of $\mathbf{R}_x$ and $\mathbf{R}_{\tilde{x}}$ that is easier to compute than Tyler's M estimator. It was first introduced [10, 32] under various names and then studied [16, 33, 34, 41, 54, 55] in the context of RES distributions. It was proved in particular that the expectations of the SSCM and SCM share the same eigenvectors with different eigenvalues with the same multiplicity and with a monotone one-to-one but rather complicated correspondence [16]. As for the covariance matrices of the SSCM and SCM estimates, it was proved that they are similarly structured. They also share the same eigenvectors with different eigenvalues [33, 34]. C-CES [8] and NC-CES [6] extensions of the definition and properties of the SSCM given in the RES context are also straightforward. The SSCM estimate $\widehat{\mathbf{R}}$ of $\mathbf{R}_x$ and $\mathbf{R}_{\tilde{x}}$ are given by

$$\widehat{\mathbf{R}} = \frac{1}{n} \sum_{i=1}^{n} \frac{\mathbf{x}_i \mathbf{x}_i^H}{\|\mathbf{x}_i\|^2} \quad \text{and} \quad \widehat{\mathbf{R}} = \frac{1}{n} \sum_{i=1}^{n} \frac{\widetilde{\mathbf{x}}_i \widetilde{\mathbf{x}}_i^H}{\|\widetilde{\mathbf{x}}_i\|^2}, \tag{43}$$

for C-CES and NC-CES distributed $\mathbf{x}_i$, respectively. Under only finite second-order moments of $\mathbf{x}_i$, applying the classic CLT to r.v. $\frac{\mathbf{x}_i^* \otimes \mathbf{x}_i}{\|\mathbf{x}_i\|^2}$ and $\frac{\widetilde{\mathbf{x}}_i^* \otimes \widetilde{\mathbf{x}}_i}{\|\widetilde{\mathbf{x}}_i\|^2}$, gives the asymptotic distributions of $\widehat{\mathbf{R}}$ which is also given by (23) and (24) where now

$$\mathbf{R} \stackrel{\text{def}}{=} \sum_{k=1}^{m} \chi_k \mathbf{v}_k \mathbf{v}_k^H \text{ and } \mathbf{R} \stackrel{\text{def}}{=} \sum_{k=1}^{2m} \widetilde{\chi}_k \widetilde{\mathbf{v}}_k \widetilde{\mathbf{v}}_k^H, \tag{44}$$

respectively, where $\sum_{k=1}^{m} \lambda_k \mathbf{v}_k \mathbf{v}_k^H$ and $\sum_{k=1}^{2m} \widetilde{\lambda}_k \widetilde{\mathbf{v}}_k \widetilde{\mathbf{v}}_k^H$ denote respectively the EVD of $\mathbf{R}_x$ and $\mathbf{R}_{\tilde{x}}$, and where closed-form expressions of the eigenvalues χ_k and $\widetilde{\chi}_k$ are given by [6, rel. (11), (12)]. The covariances $\mathbf{R}_{r_x}$ and $\mathbf{R}_{r_{\tilde{x}}}$ of the asymptotic distributions (23) and (24) share, respectively, the same eigenvectors as (27) and (28), whose expressions of the different eigenvalues $\gamma_{k,\ell}$, $\gamma_{k,\ell} - \chi_k \chi_\ell$ and $\widetilde{\gamma}_{k\ell}$, $\widetilde{\gamma}_{k,\ell} - \widetilde{\chi}_k \widetilde{\chi}_\ell$ in the EVD [6, rel. (15-16)] of $\mathbf{R}_{r_x}$ and $\mathbf{R}_{r_{\tilde{x}}}$ are given by [6, rel. (17-20)].

5 Asymptotic distributions of subspace projector estimates

The asymptotic distributions of covariance estimates enable us to infer certain properties of the corresponding subspace projector estimates.

5.1 Asymptotic inadmissibility of subspace projector estimates

Tyler's M-estimate is the ML of $\mathbf{R}_x$ when the data are real-valued angular central Gaussian distribution distributed [52]. It is similarly straightforward to prove that Tyler's M-estimates $\widehat{\mathbf{R}}_x^{\text{Ty}}$ and $\widehat{\mathbf{R}}_{\tilde{x}}^{\text{Ty}}$ are also the ML of $\mathbf{R}_x$ and $\mathbf{R}_{\tilde{x}}$, for complex circular [27] and non-circular [6] angular central Gaussian distributed data whose p.d.f. with respect to the Lebesgue measure on complex unit sphere $\mathbb{C}S^m$ are respectively given by

$$p(\mathbf{x}) = \frac{\Gamma(m)}{2\pi^m} |\mathbf{R}_x|^{-1} (\mathbf{x}^H \mathbf{R}_x^{-1} \mathbf{x})^{-m} \tag{45}$$

and

$$p(\mathbf{x}) = \frac{2^m \Gamma(m)}{2\pi^m} |\mathbf{R}_{\tilde{x}}|^{-1/2} (\widetilde{\mathbf{x}}^H \mathbf{R}_{\tilde{x}}^{-1} \widetilde{\mathbf{x}})^{-m}, \tag{46}$$

using the linear one to one mapping $\bar{\mathbf{x}}_i \longleftrightarrow \widetilde{\mathbf{x}}_i$. Thus, by the invariance property of the ML estimator, its associated orthogonal projector $\widehat{\boldsymbol{\Pi}}_x^{\text{Ty}} = \boldsymbol{\Pi}(\widehat{\mathbf{R}}_x^{\text{Ty}})$ (resp., $\widehat{\boldsymbol{\Pi}}_{\tilde{x}}^{\text{Ty}} = \boldsymbol{\Pi}(\widehat{\mathbf{R}}_{\tilde{x}}^{\text{Ty}})$) (14) is the ML of the orthogonal projector $\boldsymbol{\Pi}_x$ (resp., $\boldsymbol{\Pi}_{\tilde{x}}$). This implies that for complex angular central Gaussian distributed data, the following relationship between the covariance of the asymptotic distribution of the projector based on ML, Tyler's M, and SSCM estimates

$$\mathbf{R}_{\pi_x}^{\text{ML}} = \mathbf{R}_{\pi_x}^{\text{Ty}} \leq \mathbf{R}_{\pi_x}^{\text{SSCM}} \text{ and } \mathbf{R}_{\pi_{\tilde{x}}}^{\text{ML}} = \mathbf{R}_{\pi_{\tilde{x}}}^{\text{Ty}} \leq \mathbf{R}_{\pi_{\tilde{x}}}^{\text{SSCM}}. \tag{47}$$

Consequently, thanks to the free distribution property of Tyler's M-estimate and SSCM estimate in the C-CES and NC-CES family, where is added the circular and non-circular complex angular central Gaussian distribution [51], we can extend to arbitrary C-CES and NC-CES with finite second-order moments distributed data, the matrix inequalities between covariance of the asymptotic distribution of the projector based on Tyler's M and SSCM estimates (47). This gives:

$$\mathbf{R}_{\pi_x}^{\mathrm{Ty}} \leq \mathbf{R}_{\pi_x}^{\mathrm{SSCM}} \text{ and } \mathbf{R}_{\pi_{\tilde{x}}}^{\mathrm{Ty}} \leq \mathbf{R}_{\pi_{\tilde{x}}}^{\mathrm{SSCM}}. \tag{48}$$

Furthermore it was proved in [6] that when $\mathbf{R}_x \to \lambda\mathbf{I}$ and $\mathbf{R}_{\tilde{x}} \to \lambda\mathbf{I}$, inequalities (48) approach equalities. These inequalities (48) show that the estimator $\widehat{\mathbf{\Pi}}_x^{\mathrm{Ty}}$ (resp., $\widehat{\mathbf{\Pi}}_{\tilde{x}}^{\mathrm{Ty}}$) asymptotically dominates the estimator $\widehat{\mathbf{\Pi}}_x^{\mathrm{SSCM}}$ (resp., $\widehat{\mathbf{\Pi}}_{\tilde{x}}^{\mathrm{SSCM}}$) in the sense of the mean squared error. This property of asymptotic inadmissibility of the projector associated with the SSCM proved firstly for RES distributed data in [33] and [34], extends to arbitrary C-CES and NC-CES distributed data. However, since the SCM is very sensitive to heavy-tailed CES distributions, $\mathbf{R}_{\pi_x}^{\mathrm{SSCM}}$ and $\mathbf{R}_{\pi_{\tilde{x}}}^{\mathrm{SSCM}}$ is often bounded above by $\mathbf{R}_{\pi_x}^{\mathrm{SCM}}$ and $\mathbf{R}_{\pi_{\tilde{x}}}^{\mathrm{SCM}}$, respectively, for such distributions. An example of such behavior is given in Fig.2 and conditions for which the performance of Tyler's M-estimate and SSCM are close are specified in Section 5.3.

5.2 Asymptotic distributions of projector estimates

From the asymptotic distributions (19), (21), (23) and (24) of the different estimates $\widehat{\mathbf{R}}$ adapted to the different models presented in Section 2, we note that $\widehat{\mathbf{R}}$ converges in probability to the matrices $\mathbf{R}$. All these matrices are structured as

$$\mathbf{R} = \mathbf{S} + \sigma^2\mathbf{I}, \tag{49}$$

where $\mathrm{Span}(\mathbf{S}) = \mathrm{Span}(\mathbf{B}(\boldsymbol{\theta}))$, where $\mathbf{B}(\boldsymbol{\theta})$ denotes the mixing matrices

$$\widehat{\mathbf{R}} = \mathbf{R} + \delta(\mathbf{R}) \overset{\mathrm{EVD}}{\mapsto} \widehat{\mathbf{\Pi}} = \mathbf{\Pi}(\boldsymbol{\theta}) + \delta(\mathbf{\Pi}) \tag{50}$$

for orthogonal projectors [26] (see also the operator approach in [28]) applied to $\mathbf{\Pi}(\boldsymbol{\theta})$ associated with the noise subspace of $\mathbf{R}$,

$$\delta(\mathbf{\Pi}) = -\mathbf{\Pi}(\boldsymbol{\theta})\delta(\mathbf{R})\mathbf{S}^{\#} - \mathbf{S}^{\#}\delta(\mathbf{R})\mathbf{\Pi}(\boldsymbol{\theta}) + o(\delta(\mathbf{R}), \tag{51}$$

the asymptotic behaviors of $\widehat{\mathbf{\Pi}}$ and $\widehat{\mathbf{R}}$ are directly related. A standard theorem of continuity (see e.g., [47, p. 122]) (called also Delta-method) on regular functions of asymptotically Gaussian statistics applies and we get

$$\sqrt{n}(\mathrm{vec}(\widehat{\mathbf{\Pi}}) - \mathrm{vec}(\mathbf{\Pi}(\boldsymbol{\theta}))) \to_d \mathbb{C}N_{m^2}(\mathbf{0}, \mathbf{R}_{\pi_x}, \mathbf{R}_{\pi_x}\mathbf{K}) \tag{52}$$

$$\sqrt{n}(\mathrm{vec}(\widehat{\mathbf{\Pi}}) - \mathrm{vec}(\mathbf{\Pi}(\boldsymbol{\theta}))) \to_d \mathbb{C}N_{4m^2}(\mathbf{0}, \mathbf{R}_{\pi_{\tilde{x}}}, \mathbf{R}_{\pi_{\tilde{x}}}\mathbf{K}), \tag{53}$$

for circular and non-circular complex-valued $\mathbf{s}_i$, respectively, where $\boldsymbol{\Pi}(\boldsymbol{\theta})$ is given by (15) with its associated $\mathbf{B}(\boldsymbol{\theta})$ and where

$$\mathbf{R}_{\pi_x} = [(\mathbf{S}_x^{T\#} \otimes \mathbf{\Pi}_x) + (\mathbf{\Pi}_x^T \otimes \mathbf{S}_x^{\#})]\mathbf{R}_{r_x}[(\mathbf{S}_x^{T\#} \otimes \mathbf{\Pi}_x) + (\mathbf{\Pi}_x^T \otimes \mathbf{S}_x^{\#})], \quad (54)$$
$$\mathbf{R}_{\pi_{\tilde{x}}} = [(\mathbf{S}_{\tilde{x}}^{T\#} \otimes \mathbf{\Pi}_{\tilde{x}}) + (\mathbf{\Pi}_{\tilde{x}}^T \otimes \mathbf{S}_{\tilde{x}}^{\#})]\mathbf{R}_{r_{\tilde{x}}}[(\mathbf{S}_{\tilde{x}}^{T\#} \otimes \mathbf{\Pi}_{\tilde{x}}) + (\mathbf{\Pi}_{\tilde{x}}^T \otimes \mathbf{S}_{\tilde{x}}^{\#})], \quad (55)$$

where $\mathbf{R}_{r_x}$ and $\mathbf{R}_{r_{\tilde{x}}}$ are given by (20), (25), (27) and by (22), (26), (28), respectively, and where each of the two matrices ($\mathbf{S}_x$, $\mathbf{\Pi}_x$) and ($\mathbf{S}_{\tilde{x}}$, $\mathbf{\Pi}_{\tilde{x}}$) are the matrices $\mathbf{S}$ and $\mathbf{\Pi}(\boldsymbol{\theta})$ associated with circular and non-circular complex-valued $\mathbf{s}_i$, respectively.

Then plugging expressions (20), (25), (27) of $\mathbf{R}_{r_x}$ and (22), (26), (28) of $\mathbf{R}_{r_{\tilde{x}}}$ in (54) and (55), and using $\mathbf{\Pi}_x\mathbf{S}_x = \mathbf{0}$ and $\mathbf{\Pi}_{\tilde{x}}\mathbf{S}_{\tilde{x}} = \mathbf{0}$, the following theorem extending [15, Th/IV.1] is proved:

Theorem 1 *The covariance matrices $\mathbf{R}_{\pi_x}$ and $\mathbf{R}_{\pi_{\tilde{x}}}$ of the asymptotic distributions* (52) *and* (53) *of the different projector estimates $\widehat{\mathbf{\Pi}}$ associated with the different data models presented in Section 2 have an unified structure given by*

$$\mathbf{R}_{\pi_x} = (\mathbf{U}^T \otimes \boldsymbol{\Pi}(\boldsymbol{\theta})) + (\boldsymbol{\Pi}^T(\boldsymbol{\theta}) \otimes \mathbf{U}), \quad (56)$$
$$\mathbf{R}_{\pi_{\tilde{x}}} = [\mathbf{I} + \mathbf{K}(\mathbf{J} \otimes \mathbf{J})][(\mathbf{U}^T \otimes \boldsymbol{\Pi}(\boldsymbol{\theta})) + (\boldsymbol{\Pi}^T(\boldsymbol{\theta}) \otimes \mathbf{U})]. \quad (57)$$

Here $\boldsymbol{\Pi}(\boldsymbol{\theta})$ are the projection matrices $\sum_{k=p+1}^{m} \mathbf{v}_k\mathbf{v}_k^H$, $\sum_{k=p+1}^{2m} \widetilde{\mathbf{v}}_k\widetilde{\mathbf{v}}_k^H$ and $\sum_{k=2p+1}^{2m} \widetilde{\mathbf{v}}_k\widetilde{\mathbf{v}}_k^H$ on the noise subspace (i.e., on the orthogonal complement of the range of $\mathbf{A}$, $\widetilde{\mathbf{A}}_r$ and $\widetilde{\mathbf{A}}_c$), associated with circular, rectilinear, and non-rectilinear and non-circular complex-valued $\mathbf{s}_i$, respectively. On the other hand, the matrices $\mathbf{U}$ depend on the covariance estimates $\widehat{\mathbf{R}}$ studied in Section 4.

For the deterministic model and the stochastic model where both $\mathbf{s}_i$ and $\mathbf{n}_i$ are CES distributed, $\mathbf{U}$ takes the common form

$$\mathbf{U} = \sigma_n^2\mathbf{S}^{\#}\mathbf{R}\mathbf{S}^{\#} + \kappa_n\sigma_n^4(\mathbf{S}^{\#})^2 = \sum_{k=1}^{p} \frac{\sigma_n^2(\lambda_k + \kappa_n\sigma_n^2)}{(\lambda_k - \sigma_n^2)^2}\mathbf{v}_k\mathbf{v}_k^H, \quad (58)$$

with $\mathbf{R} \stackrel{\text{def}}{=} \mathbf{R}_{x,\infty} = \mathbf{R}_x = \sum_{k=1}^{m} \lambda_k\mathbf{v}_k\mathbf{v}_k^H$ and $\mathbf{S} \stackrel{\text{def}}{=} \mathbf{A}\mathbf{R}_{s,\infty}\mathbf{A}^H = \mathbf{A}\mathbf{R}_s\mathbf{A}^H$ for $\mathbf{s}_i$ deterministic of arbitrary circularity and circular stochastic. Similarly for rectilinear $\mathbf{s}_i$, (58) also applies where $\mathbf{R}_{x,\infty}$, $\mathbf{R}_x$, $\mathbf{A}\mathbf{R}_{s,\infty}\mathbf{A}^H$, $\mathbf{A}\mathbf{R}_s\mathbf{A}^H$, λ_k and $\mathbf{v}_k$ are replaced by $\mathbf{R}_{\tilde{x},\infty}$, $\mathbf{R}_{\tilde{x}}$, $\widetilde{\mathbf{A}}_r\mathbf{R}_{r,\infty}\widetilde{\mathbf{A}}_r^H$, $\widetilde{\mathbf{A}}_r\mathbf{R}_r\widetilde{\mathbf{A}}_r^H$, $\widetilde{\lambda}_k$ and $\widetilde{\mathbf{v}}_k$ respectively. Furthermore, for non-circular and non-rectilinear stochastic $\mathbf{s}_i$ (58) still applies where $\mathbf{R} \stackrel{\text{def}}{=} \mathbf{R}_{\tilde{x}}$, $\mathbf{S} \stackrel{\text{def}}{=} \widetilde{\mathbf{A}}_c\mathbf{R}_{\tilde{s}}\widetilde{\mathbf{A}}_c^H$ and p, λ_k and $\mathbf{v}_k$ are replaced, respectively, by $2p$, $\widetilde{\lambda}_k$ and $\widetilde{\mathbf{v}}_k$.

For the stochastic model, where $\mathbf{x}_i$ is CES distributed, we get for SCM, ML, M and Tyler's estimator:

$$\mathbf{U} = \vartheta\left(\sum_{k=1}^{p} \frac{\lambda_k\sigma_n^2}{(\lambda_k - \sigma_n^2)^2}\mathbf{v}_k\mathbf{v}_k^H\right) \text{ and } \mathbf{U} = \vartheta\left(\sum_{k=1}^{p} \frac{\widetilde{\lambda}_k\sigma_n^2}{(\widetilde{\lambda}_k - \sigma_n^2)^2}\widetilde{\mathbf{v}}_k\widetilde{\mathbf{v}}_k^H\right), \quad (59)$$

for circular and rectilinear $\mathbf{s}_i$, respectively, with:

$$\vartheta \stackrel{\text{def}}{=} \vartheta_{\text{SCM}} = 1 + \kappa_x \quad \text{for the SCM estimate,} \tag{60}$$

$$\vartheta \stackrel{\text{def}}{=} \vartheta_{\text{ML}} = \sigma_1 \quad \text{for the ML estimate given by (32),} \tag{61}$$

$$\vartheta \stackrel{\text{def}}{=} \vartheta_{\text{M}} = c^2\sigma_1 \quad \text{for the } M\text{-estimate where } \sigma_1 \text{ and } c \text{ are given by (38) and (35), respectively,} \tag{62}$$

$$\vartheta \stackrel{\text{def}}{=} \vartheta_{\text{Ty}} = 1 + \frac{1}{m} \quad \text{for the Tyler's } M\text{-estimate.} \tag{63}$$

For non-circular and non-rectilinear $\mathbf{s}_i$, the second relation of (59) also applies by replacing p by $2p$.

For the SSCM estimator, $\mathbf{U}$ is given by

$$\mathbf{U} = \sum_{k=1}^{p} \frac{\gamma_k}{(\chi_k - \chi)^2}\mathbf{v}_k\mathbf{v}_k^H \quad \text{and} \quad \mathbf{U} = \sum_{k=1}^{p} \frac{\widetilde{\gamma}_k}{(\widetilde{\chi}_k - \widetilde{\chi})^2}\widetilde{\mathbf{v}}_k\widetilde{\mathbf{v}}_k^H, \tag{64}$$

for circular $\mathbf{s}_i$ where $\chi \stackrel{\text{def}}{=} \chi_{p+1} = \chi_{p+2} = ... = \chi_m$ (see (44)) and $\gamma_k \stackrel{\text{def}}{=} \gamma_{k,p+1} = \gamma_{k,p+2} = ... = \gamma_{k,m}$, and rectilinear $\mathbf{s}_i$ where $\widetilde{\chi} \stackrel{\text{def}}{=} \widetilde{\chi}_{p+1} = \widetilde{\chi}_{p+2} = ... = \widetilde{\chi}_{2m}$ (see (44)) and $\widetilde{\gamma}_k \stackrel{\text{def}}{=} \widetilde{\gamma}_{k,p+1} = \widetilde{\gamma}_{k,p+2} = ... = \widetilde{\gamma}_{k,2m}$ are given in [6, rel. (17)-(20)]. For non-circular and non-rectilinear $\mathbf{s}_i$, the second relation of (64) also applies by replacing p by $2p$.

Two remarks are in order from Theorem 1 for the deterministic model and the stochastic model where both $\mathbf{s}_i$ and $\mathbf{n}_i$ are CES distributed:

- The projector estimators have the same asymptotic distribution in both deterministic and stochastic CES distributed models for $\mathbf{s}_i$, regardless of whether $\mathbf{s}_i$ is arbitrary circular or rectilinear. This contrasts with the covariances of the asymptotic distributions of $\mathbf{R}_x$ (20), (25) and $\mathbf{R}_{\tilde{x}}$ (22), (26), which are different. This generalizes classical results on many subspace-based DOA estimators [49] that were proven under the complex circular Gaussian noise framework.
- The asymptotic distributions of the projector estimators remain invariant to the distribution of $\mathbf{s}_i$, regardless of whether $\mathbf{s}_i$ is of arbitrary circularity or is rectilinear. This property also generalizes classical results on subspace-based estimators [11] that were proven under the complex circular Gaussian noise framework.

5.3 Relative efficiency

For circular or non-circular CES distributed data $\mathbf{x}_i$, the covariance matrices of the asymptotic distribution of the projector derived from the SCM, ML, M, and Tyler's estimators are proportional. The proportionality coefficient ϑ (60)-(63) serves as an efficiency index for the estimation of the projector. This proportionality relation only holds for the projector derived from the SSCM estimator when $\lambda_1 = ... = \lambda_p$ and $\widetilde{\lambda}_1 = ... = \widetilde{\lambda}_p$ which implies that $\chi_1 = ... = \chi_p$, $\gamma_1 = ... = \gamma_p$ and $\widetilde{\chi}_1 = ... = \widetilde{\chi}_p$, $\widetilde{\gamma}_1 = ... = \widetilde{\gamma}_p$ in (64), and hence

$$\vartheta \stackrel{\text{def}}{=} \vartheta_{\text{SSCM}} = \frac{\gamma_1}{(\chi_1-\chi)^2}\frac{(\lambda_1-\sigma_n^2)^2}{\lambda_1\sigma_n^2} \text{ and } \vartheta \stackrel{\text{def}}{=} \vartheta_{\text{SSCM}} = \frac{\widetilde{\gamma}_1}{(\widetilde{\chi}_1-\widetilde{\chi})^2}\frac{(\widetilde{\lambda}_1-\sigma_n^2)^2}{\widetilde{\lambda}_1\sigma_n^2} \tag{65}$$

in the circular and rectilinear cases, respectively. For non-circular and non-rectilinear $\mathbf{s}_i$, the second relation of (65) also applies under the condition $\widetilde{\lambda}_1 = ... = \widetilde{\lambda}_{2p}$. To obtain insight into how the asymptotic inadmissibility of the SSCM is affected whenever $\mathbf{R}_x \neq \lambda\mathbf{I}$ and $\mathbf{R}_{\tilde{x}} \neq \lambda\mathbf{I}$ (see discussion after (48)), the following theorem is proved in [6, th. 5]:

Theorem 2 *The asymptotic efficiency of the SSCM estimate relative to Tyler's M-estimate defined by the ratio $r \stackrel{\text{def}}{=} \vartheta_{\text{Ty}}/\vartheta_{\text{SSCM}}$, is given by the closed-form expressions in the circular case under the assumption $\lambda_1 = ... = \lambda_p$ and rectilinear case under the assumption $\widetilde{\lambda}_1 = ... = \widetilde{\lambda}_p$, respectively, by:*

$$r_c = \frac{[{}_2F_1(1, m-p+1, m+2, 1-\rho)]^2}{{}_2F_1(2, m-p+1, m+2, 1-\rho)} \tag{66}$$

and

$$r_r = \frac{[{}_2F_1(1, m-\frac{p}{2}+1, m+2, 1-\widetilde{\rho})]^2}{{}_2F_1(2, m-\frac{p}{2}+1, m+2, 1-\widetilde{\rho})}, \tag{67}$$

where $\rho \stackrel{\text{def}}{=} \frac{\sigma_n^2}{\lambda_1}$ and $\widetilde{\rho} \stackrel{\text{def}}{=} \frac{\sigma_n^2}{\widetilde{\lambda}_1}$, and these ratios are monotonic increasing functions of respectively ρ and $\widetilde{\rho}$ from the intervals (0,1) to (0,1). Furthermore in the neighborhood of $\rho = 1$, $\widetilde{\rho} = 1$ and $\rho = 0$, $\widetilde{\rho}_1 = 0$, we have respectively:

$$r_c = 1 - \frac{(m-p+1)(p+1)}{(m+2)^2(m+3)}(1-\rho)^2 + o(1-\rho)^2, \tag{68}$$

$$r_r = 1 - \left(\frac{(2m-p+2)(p+2)}{4(m+2)^2(m+3)}\right)(1-\widetilde{\rho})^2 + o(1-\widetilde{\rho})^2, \tag{69}$$

and

$$r_c = \begin{cases} o_{m,1}(1) & \text{for } p = 1 \\ (1+\frac{1}{m})(1-\frac{1}{p})(1+o_{m,p}(1)) & \text{for } p > 1 \end{cases} \tag{70}$$

$$r_r = \begin{cases} \widetilde{o}_{m,p}(1) & \text{for } p = 1,2 \\ (1+\frac{1}{m})(1-\frac{2}{p})(1+\widetilde{o}_{m,p}(1)) & \text{for } p > 2 \end{cases}, \tag{71}$$

where first-order expansions of $o_{m,p}(1)$ and $\widetilde{o}_{m,p}(1)$ satisfying $\lim_{\rho\to 0} o_{m,p}(1) = \lim_{\widetilde{\rho}\to 0} \widetilde{o}_{m,p}(1) = 0$ are given by [6, rel. (57-58)]. For non-circular and non-rectilinear $\mathbf{s}_i$, relation (67) and the expansions of r_r also apply under the condition $\widetilde{\lambda}_1 = ... = \widetilde{\lambda}_{2p}$ where p is replaced par $2p$.

It follows from (68) and (69) that the performance of the subspace estimation derived from SSCM and Tyler's M-estimate are very similar for close eigenvalues, and particularly for large values of m and p. It follows, conversely, from (70) and (71), that for well-separated eigenvalues, the performance of the SSCM-based subspace estimation is largely outperformed by those derived from Tyler's M estimate for $p = 1$ and $p = 1, 2$ for C-CES and NC-CES distributed data because r_c and r_r tend to zero as ρ and $\widetilde{\rho}$ tend to zero, respectively. Note, however, that for m and p large, the performance of the subspace estimation derived from SSCM and Tyler's M estimate are very similar because r_c and r_r are equivalent to $(1+\frac{1}{m})(1-\frac{1}{p}) < 1$ and $(1+\frac{1}{m})(1-\frac{2}{p}) < 1$ as ρ and $\widetilde{\rho}$ tend to zero, respectively. Consequently, despite the asymptotic inadmissibility of subspace projectors built from the SSCM estimate, the performance of this estimator and those derived from Tyler's M-estimator are close in particular for large values of m and not too small values of p. Therefore, we can conclude that SSCM estimate is of great interest from the point of view of its lower computational complexity for large values of m.

6 Asymptotic distributions of subspace-based parameter estimates

The asymptotic distributions of projector estimates given in Theorem 1 allow us to derive the asymptotic distribution of subspace-based parameter estimates for any algorithms "alg" (14) which are assumed to be differentiable[4].

6.1 Asymptotic distribution of parameter estimate

Again using a standard theorem of continuity (see e.g., [47, p. 122]) (called also Delta-method), we get:

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \to_d \mathbb{R}N_k(\mathbf{0}, \mathbf{R}^{\text{alg}}_{\hat{\theta}}), \tag{72}$$

where $\mathbf{R}^{\text{alg}}_{\hat{\theta}}$ is given by:

$$\mathbf{R}^{\text{alg}}_{\hat{\theta}} = \mathbf{D}_{\text{alg}} \mathbf{R}_{\pi} \mathbf{D}^{H}_{\text{alg}}, \tag{73}$$

with $\mathbf{R}_{\pi} = \mathbf{R}_{\pi_x}$ [resp., $\mathbf{R}_{\pi} = \mathbf{R}_{\pi_{\tilde{x}}}$] is given by (56) [resp., (57)], for circular [resp., non-circular] $\mathbf{s}_i$, and where $\mathbf{D}_{\text{alg}}$ is the differential (or Jacobian) matrix defined by the relation

$$\widehat{\boldsymbol{\theta}} = \text{alg}(\widehat{\boldsymbol{\Pi}}) = \underbrace{\text{alg}(\boldsymbol{\Pi}(\boldsymbol{\theta}))}_{\boldsymbol{\theta}} + \mathbf{D}_{\text{alg}} \text{vec}(\widehat{\boldsymbol{\Pi}} - \boldsymbol{\Pi}(\boldsymbol{\theta})) + o(\widehat{\boldsymbol{\Pi}} - \boldsymbol{\Pi}(\boldsymbol{\theta})). \tag{74}$$

Using the unified expressions of $\mathbf{R}_{\pi_x}$ and $\mathbf{R}_{\pi_{\tilde{x}}}$ of Theorem 1 where $\mathbf{U}$ is given by (58) and (59), we get the following theorem:

[4] Note that since $\boldsymbol{\Pi}(\boldsymbol{\theta})$ is Hermitian, the mappings "alg" are differentiable w.r.t. $\boldsymbol{\Pi}(\boldsymbol{\theta})$ i.f.f. they are differentiable w.r.t. $(\text{Re}(\boldsymbol{\Pi}(\boldsymbol{\theta})), \text{Im}(\boldsymbol{\Pi}(\boldsymbol{\theta})))$. This last hypothesis being verified by most algorithms "alg".

Theorem 3 *The covariance matrix $\mathbf{R}_{\hat{\theta}}^{\mathrm{alg}}$ of the asymptotic distribution of the parameter estimate $\widehat{\boldsymbol{\theta}}$ given by any subspace-based algorithm "alg" satisfying the previously aforementioned assumptions, derived from the SCM, ML, M and Tyler's M covariance estimate, breaks down as follows*

$$\mathbf{R}_{\hat{\theta}}^{\mathrm{alg}} = \vartheta \mathbf{R}_{\hat{\theta}}^{\mathrm{G,alg}} + \kappa_n \mathbf{R}_{\hat{\theta}}^{'\mathrm{alg}}, \tag{75}$$

where $\mathbf{R}_{\hat{\theta}}^{\mathrm{G,alg}}$ is the covariance matrix $\mathbf{R}_{\hat{\theta}}^{\mathrm{alg}}$ derived for Gaussian distributed data $\mathbf{x}_i$ and in the second term $\mathbf{R}_{\hat{\theta}}^{'\mathrm{alg}}$ is a positive definite matrix obtained only in the deterministic and stochastic model where $\mathbf{n}_i$ is non-Gaussian C-CES distributed. $\vartheta = 1$ in the deterministic and stochastic model where $\mathbf{n}_i$ is C-CES distributed, ϑ is given by (60)-(63) for CES distributed data $\mathbf{x}_i$ and κ_n is the kurtosis parameter of the noise, which is zero for Gaussian distributed noise.

In particular, for CES distributed $\mathbf{x}_i$,

$$\mathbf{R}_{\hat{\theta}}^{\mathrm{alg}} = \vartheta \mathbf{R}_{\hat{\theta}}^{\mathrm{G,alg}} \tag{76}$$

and all the analytical theorems concerning the asymptotic distributions of subspace-based parameter estimates derived in the Gaussian framework extend with a simple multiplicative term ϑ, especially in DOA estimation.

For the SSCM estimate, (76) also applies where ϑ is given by (65) but only in the case $\lambda_1 = ... = \lambda_p$ and $\widetilde{\lambda}_1 = ... = \widetilde{\lambda}_p$. In contrast, for an arbitrary spectrum of eigenvalues, there is no longer a direct relationship between $\mathbf{R}_{\hat{\theta}}^{\mathrm{alg}}$ and $\mathbf{R}_{\hat{\theta}}^{\mathrm{G,alg}}$.

In the deterministic and stochastic model where $\mathbf{n}_i$ is non-Gaussian C-CES distributed, the structure of $\mathbf{R}_{\hat{\theta}}^{\mathrm{alg}}$ is affected due to the additive term $\mathbf{R}_{\hat{\theta}}^{',\mathrm{alg}}$. Furthermore, $\mathbf{R}_{\hat{\theta}}^{\mathrm{alg}} > \mathbf{R}_{\hat{\theta}}^{\mathrm{G,alg}}$ and $\mathbf{R}_{\hat{\theta}}^{\mathrm{alg}} < \mathbf{R}_{\hat{\theta}}^{\mathrm{G,alg}}$ for super-Gaussian ($\kappa_n > 0$) and sub-Gaussian ($\kappa_n < 0$) distributed noise, respectively.

Besides, the covariance of the asymptotic distribution of the estimate $\widehat{\boldsymbol{\theta}}$ given by the AMV algorithm [45], [48] takes the particular expression [5]:

$$\mathbf{R}_{\hat{\theta}}^{\mathrm{AMV}} = (\boldsymbol{\Pi}^{'H}(\boldsymbol{\theta}) \mathbf{R}_{\pi}^{\#} \boldsymbol{\Pi}'(\boldsymbol{\theta}))^{-1} \tag{77}$$

where $\boldsymbol{\Pi}'(\boldsymbol{\theta}) \stackrel{\mathrm{def}}{=} \frac{d\mathrm{vec}(\boldsymbol{\Pi}(\boldsymbol{\theta}))}{d\boldsymbol{\theta}}$. It has been proved in [5] that the AMV estimates $\widehat{\boldsymbol{\theta}}$ derived from the projector estimates $\widehat{\boldsymbol{\Pi}}$ built on the ML estimate of $\mathbf{R}_x$ and $\mathbf{R}_{\tilde{x}}$ for respectively stochastic C-CES and NC-CES distributed data $\mathbf{x}_i$ are asymptotically efficient, i.e., the covariance matrices $\mathbf{R}_{\hat{\theta}}^{\mathrm{ML,AMV}}$ of their asymptotic distributions reach the stochastic Cramér-Rao bound (CRB) of the parameter $\boldsymbol{\theta}$ when the density generator g is known. Consequently the following theorem is deduced from (48), (73) and (77):

Theorem 4 *For CES distributed data $\mathbf{x}_i$, the covariance of the Gaussian asymptotic distribution of the estimated parameter $\widehat{\boldsymbol{\theta}}$ derived for any subspace-based algorithm "alg" built on the SSCM is bounded below by those built on Tyler's M estimate. These two covariance matrices being themselves bounded below by the CRB.*

$$n \times \mathrm{CRB}(\boldsymbol{\theta}) = (\boldsymbol{\Pi}'^{H}(\boldsymbol{\theta})\mathbf{R}_{\pi}^{\mathrm{ML}\#}\boldsymbol{\Pi}'(\boldsymbol{\theta}))^{-1}$$
$$= \mathbf{R}_{\hat{\theta}}^{\mathrm{ML,AMV}} \leq \mathbf{R}_{\hat{\theta}}^{\mathrm{ML,alg}} \leq \mathbf{R}_{\hat{\theta}}^{\mathrm{Ty,alg}} \leq \mathbf{R}_{\hat{\theta}}^{\mathrm{SSCM,alg}}, \qquad (78)$$

where $\mathbf{R}_{\pi}^{\mathrm{ML}}$ *denotes the covariance of the asymptotic distribution of* $\widehat{\boldsymbol{\Pi}}$ *built on the ML estimate of* $\mathbf{R}_x$ *and* $\mathbf{R}_{\tilde{x}}$ *for respectively C-CES and NC-CES distributed data* $\mathbf{x}_i$*, and* $\mathbf{R}_{\hat{\theta}}^{\mathrm{ML,alg}}$, $\mathbf{R}_{\hat{\theta}}^{\mathrm{Ty,alg}}$ *and* $\mathbf{R}_{\hat{\theta}}^{\mathrm{SSCM,alg}}$ *denote the covariance of the asymptotic distribution of the parameter* $\widehat{\boldsymbol{\theta}}$ *estimated by the algorithm "alg" built on the ML, Tyler's* M *and SSCM covariance estimate, respectively.*

Theorem 4 proves that the AMV subspace-based estimators built on Tyler's M-estimator of the covariance matrix are not efficient. To obtain a truly robust efficient subspace-based estimator, one has to find M-estimators with an appropriate weight function $u_x(.)$ such that ϑ_{M} be close or equal to ϑ_{ML}. Note that similarly to the projector estimate, there is no general order relation between $\mathbf{R}_{\hat{\theta}}^{\mathrm{SSCM,alg}}$ and $\mathbf{R}_{\hat{\theta}}^{\mathrm{SCM,alg}}$. However, since the SCM is very sensitive to heavy-tailed CES distributions, $\mathbf{R}_{\hat{\theta}}^{\mathrm{SSCM,alg}}$ can be bounded above by $\mathbf{R}_{\hat{\theta}}^{\mathrm{SCM,alg}}$ for such distributions.

Considering now the stochastic CRB in (78), the following common closed-form expression has been proved in [5]:

$$\mathrm{CRB}(\boldsymbol{\theta}) = \sigma_1 \frac{\sigma_n^2}{2}\left[\mathrm{Re}\left(\frac{d\mathbf{a}_{\theta}^H}{d\boldsymbol{\theta}}(\mathbf{H}^T \otimes \boldsymbol{\Pi}(\boldsymbol{\theta}))\frac{d\mathbf{a}_{\theta}}{d\boldsymbol{\theta}}\right)\right]^{-1}, \qquad (79)$$

where σ_1 is given by (32). $\mathbf{a}_{\theta} \stackrel{\mathrm{def}}{=} \mathrm{vec}(\mathbf{A})$, $\mathbf{H} \stackrel{\mathrm{def}}{=} \mathbf{R}_s^H\mathbf{A}^H\mathbf{R}_x^{-1}\mathbf{A}\mathbf{R}_s$ and $\boldsymbol{\Pi}(\boldsymbol{\theta}) \stackrel{\mathrm{def}}{=} \boldsymbol{\Pi}_x(\boldsymbol{\theta})$ in circular case, $\mathbf{a}_{\theta} \stackrel{\mathrm{def}}{=} \mathrm{vec}(\widetilde{\mathbf{A}}_r)$, $\mathbf{H} \stackrel{\mathrm{def}}{=} \mathbf{R}_r\widetilde{\mathbf{A}}_r^H\widetilde{\mathbf{R}}_{\tilde{x}}^{-1}\widetilde{\mathbf{A}}_r\mathbf{R}_r$ and $\boldsymbol{\Pi}(\boldsymbol{\theta}) \stackrel{\mathrm{def}}{=} \boldsymbol{\Pi}_{\tilde{x}}(\boldsymbol{\theta})$ in the rectilinear case[5] associated with the structured extended covariance (6) and $\mathbf{a}_{\theta} \stackrel{\mathrm{def}}{=} \mathrm{vec}(\mathbf{A})$, $\mathbf{H} \stackrel{\mathrm{def}}{=} \left(\mathbf{R}_s\mathbf{A}^H, \mathbf{C}_s\mathbf{A}^T\right)\widetilde{\mathbf{R}}_{\tilde{x}}^{-1}\begin{pmatrix}\mathbf{A}\mathbf{R}_s \\ \mathbf{A}^*\mathbf{C}_s^*\end{pmatrix}$ and $\boldsymbol{\Pi}(\boldsymbol{\theta}) \stackrel{\mathrm{def}}{=} \boldsymbol{\Pi}_x(\boldsymbol{\theta})$ in the non-circular and no-rectilinear case associated with the structured extended covariance (8).

In general, this stochastic CRB is upper bounded by the semiparametric CRB introduced by [18] when the density generator g is considered as an infinite-dimensional unknown nuisance parameter. This semiparametric CRB has been studied for RES and C-CES distributions in [20] and [19], respectively. In particular, a closed-form expression of this bound has been derived in [17] for the DOA parameter of C-CES distributed observations. By slightly modifying and extending the proof given in the support document of [19] to general C-CES and NC-CES distributed noisy linear mixture models (1), it has been proved in [5], that this semiparametric CRB coincides with the stochastic CRB given by (79). We note that this property is very specific to the parameter of interest characterized by the column space of the mixing matrix. This property is explained by the fact that this column space does not depend on the density generator g. It is important, however, to note that if the AMV subspace-based estimator derived from the ML estimate of the covariance of the

[5] Note that in this case $\frac{d\mathbf{a}_{\theta}^H}{d\boldsymbol{\theta}}(\mathbf{H}^T \otimes \boldsymbol{\Pi}(\boldsymbol{\theta}))\frac{d\mathbf{a}_{\theta}}{d\boldsymbol{\theta}}$ is real-valued.

data $\mathbf{x}_i$ is asymptotically efficient w.r.t. the stochastic CRB, it is no longer asymptotically efficient w.r.t. the semiparametric CRB because this ML M-estimate requires the knowledge of the density generator g. To obtain asymptotically semiparametric efficient subspace-based estimator of $\boldsymbol{\theta}$, the AMV estimator would have to be built on an asymptotically semiparametric efficient estimator $\widehat{\mathbf{R}}$ of the covariance of the data in (14) like the one proposed in [21].

6.2 Numerical illustrations

To illustrate the asymptotic distributions of subspace-based parameter estimates, we focus on the conventional and non-circular MUSIC-based DOA estimation algorithm [2]. We consider that two uncorrelated circular or rectilinear sources of equal power σ_s^2, are impinging on an uniform linear array with m sensors for which $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2]$ with $\mathbf{a}_k = (1, e^{i\theta_k}, \ldots, e^{i(m-1)\theta_k})^T$, where $\theta_k = \pi \sin \alpha_k$, with α_k are the DOAs relative to the normal of array broadside. The SNR is defined as SNR = σ_s^2/σ_n^2. The Jacobian $\mathbf{D}_{\text{music}}$ is given for the conventional MUSIC algorithm by (see [2] for the non-circular MUSIC algorithm):

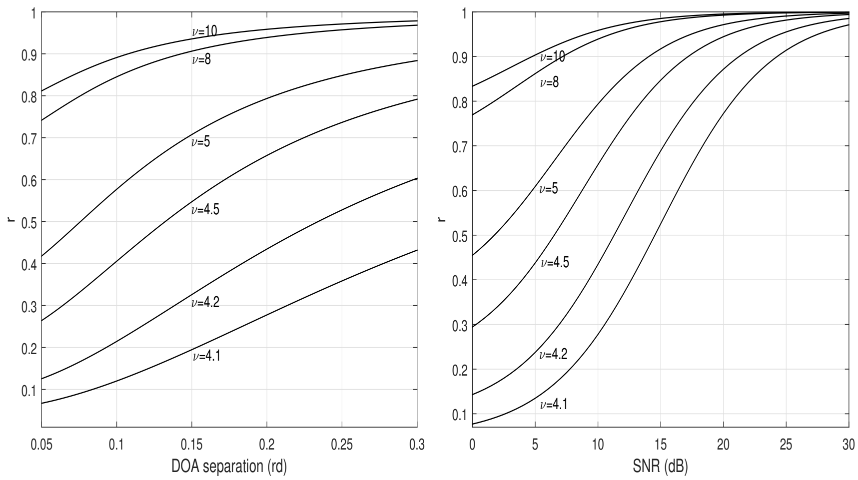
$$\mathbf{D}_{\text{music}} = \begin{pmatrix} \mathbf{d}_1^T \\ \mathbf{d}_2^T \end{pmatrix} \quad \text{with} \quad \mathbf{d}_k^T = \frac{-1}{\alpha_{k,k}} (\mathbf{a}_k^{'\,T} \otimes \mathbf{a}_k^H + \mathbf{a}_k^T \otimes \mathbf{a}_k^{'\,H}) \tag{80}$$

and $\alpha_{k,k} \stackrel{\text{def}}{=} 2\mathbf{a}_k^{'H} \boldsymbol{\Pi}_x \mathbf{a}_k'$ where $\mathbf{a}_k' \stackrel{\text{def}}{=} d\mathbf{a}_k/d\theta_k$. In these circular and rectilinear scenarios, the variances of the asymptotic distribution of the estimates of θ_1 and θ_2 are equal. Thus, we only consider the accuracy of source 1 in the sequel.

In this illustration, we choose the Student t-distribution that belongs to the subclass of the compound Gaussian distributions (See Chapter 1). It is parameterized by the degree of freedom ν ($0 < \nu < \infty$) which controls the tails of the distribution that are heavier than the Gaussian ones (obtained for $\nu \to \infty$). This distribution has first, second and fourth-order moments if, respectively $\nu > 1$, $\nu > 2$ and $\nu > 4$ with $\vartheta_{\text{SCM}} = 1 + \kappa_n = \frac{\nu-2}{\nu-4}$ with $\nu > 4$ and $\vartheta_{\text{ML}} = \sigma_1 = \frac{m+1+\nu/2}{m+\nu/2}$. We also remind the reader that $\vartheta_{\text{Ty}} = 1 + 1/m$ for Tyler's M estimate (see (63)).

In the first setup, the sources $s_{i,k}$ are circular Gaussian distributed and the noise $\mathbf{n}_i$ is circular complex Student t-distributed with $m = 6$ and the SCM is used. Fig.1a and 1b show the theoretical ratio $r \stackrel{\text{def}}{=} [\mathbf{R}_{\hat{\theta}}^{\text{G,music}}]_{1,1}/[\mathbf{R}_{\hat{\theta}}^{\text{S,music}}]_{1,1}$ (where $[\mathbf{R}_{\hat{\theta}}^{\text{S,music}}]$ is given by (75)) of the variances of the asymptotic distribution of $\hat{\theta}_1$ for Gaussian distributed and Student t-distributed noise for different values of the parameter ν, respectively w.r.t. the DOA separation $\Delta\theta = |\theta_2 - \theta_1|$, and w.r.t. to the SNR. We see that the performance deteriorates strongly for heavy-tailed noise distributions (i.e., when $\nu \to 4$) w.r.t. the Gaussian distribution and that degradation increases for small DOA separation and small SNR.

In the second setup, we assume that the measurement $\mathbf{x}_i$ are circular complex Student t-distributed. We compare now the variances of the asymptotic distribu-



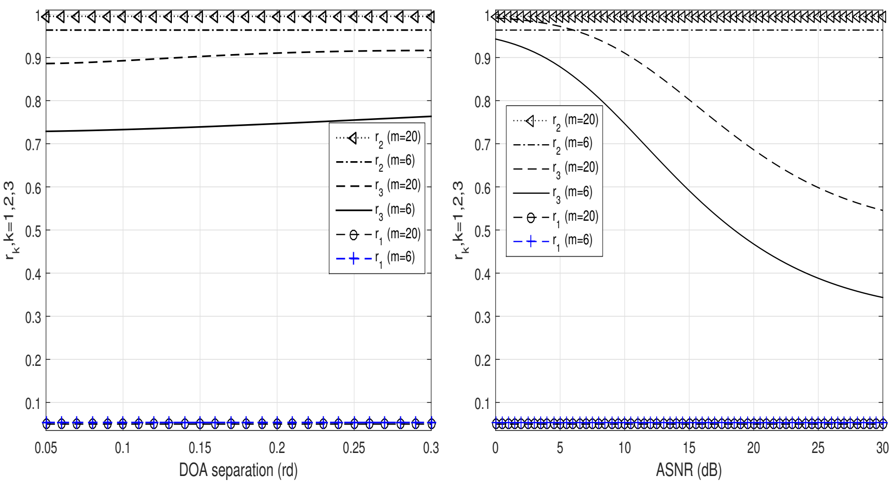
(a) r as a function of $\Delta\theta$ for SNR = 10dB (b) r as a function of SNR for $\Delta\theta = 0.2$rd

Fig. 1 Theoretical ratio $r \stackrel{\text{def}}{=} [\mathbf{R}_{\hat{\theta}}^{\text{G,music}}]_{1,1}/[\mathbf{R}_{\hat{\theta}}^{\text{S,music}}]_{1,1}$ for several values of ν.

tion of $\widehat{\theta}_1$ estimated from the SCM, ML and Tyler's M covariance estimate and SSCM, respectively denoted by $[\mathbf{R}_{\hat{\theta}}^{\text{SCM,music}}]_{1,1}$, $[\mathbf{R}_{\hat{\theta}}^{\text{ML,music}}]_{1,1}$, $[\mathbf{R}_{\hat{\theta}}^{\text{Ty,music}}]_{1,1}$ and $[\mathbf{R}_{\hat{\theta}}^{\text{SSCM,music}}]_{1,1}$. The first three variances are proportional to $[\mathbf{R}_{\hat{\theta}}^{\text{G,music}}]_{1,1}$ given for $\mathbf{x}_i$ circular Gaussian distributed (76), but $[\mathbf{R}_{\hat{\theta}}^{\text{SSCM,music}}]_{1,1}$ is no longer proportional because

$$\lambda_1 = m\sigma_s^2\left(1+\frac{\sin(m\Delta\theta/2)}{m\sin(\Delta\theta/2)}\right)+\sigma_n^2 \neq \lambda_2 = m\sigma_s^2\left(1-\frac{\sin(m\Delta\theta/2)}{m\sin(\Delta\theta/2)}\right)+\sigma_n^2. \tag{81}$$

Consequently the ratios $r_1 \stackrel{\text{def}}{=} [\mathbf{R}_{\hat{\theta}}^{\text{ML,music}}]_{1,1}/[\mathbf{R}_{\hat{\theta}}^{\text{SCM,music}}]_{1,1} = \vartheta_{\text{ML}}/\vartheta_{\text{SCM}}$ and $r_2 \stackrel{\text{def}}{=} [\mathbf{R}_{\hat{\theta}}^{\text{ML,music}}]_{1,1}/[\mathbf{R}_{\hat{\theta}}^{\text{Ty,music}}]_{1,1} = \vartheta_{\text{ML}}/\vartheta_{\text{Ty}}$ depend only on m and ν, in contrast to $r_3 \stackrel{\text{def}}{=} [\mathbf{R}_{\hat{\theta}}^{\text{ML,music}}]_{1,1}/[\mathbf{R}_{\hat{\theta}}^{\text{SSCM,music}}]_{1,1}$ that also depend on $\Delta\theta$ and SNR. Fig.2a and 2b show the theoretical ratio r_1, r_2 and r_3 for $\nu = 4.1$ and two values of m, respectively w.r.t. the DOA separation $\Delta\theta = |\theta_2 - \theta_1|$ where the array SNR (ASNR) (defined by $m\sigma_s^2/\sigma_n^2$ [53, Chap.9]) is fixed at 10dB and w.r.t. to the ASNR where $\Delta\theta = |\theta_2 - \theta_1| = 0.2$rd. We clearly see that the performance of the SCM-based estimate is very poor unlike that of Tyler's M-based estimate whose performance is very close to that of the ML-based estimate for heavy-tailed noise distributions. As for the SSCM-based estimate, its performance is degraded compared to that of Tyler's M-based estimate, and this degradation increases when the ASNR increases and $\Delta\theta$ decreases. This behavior is then explained by an increase in the differences between eigenvalues λ_1, λ_2 and σ_n^2 consistently with the comment following expression (48). But the SSCM-based estimate largely outperforms the SCM-based estimate.



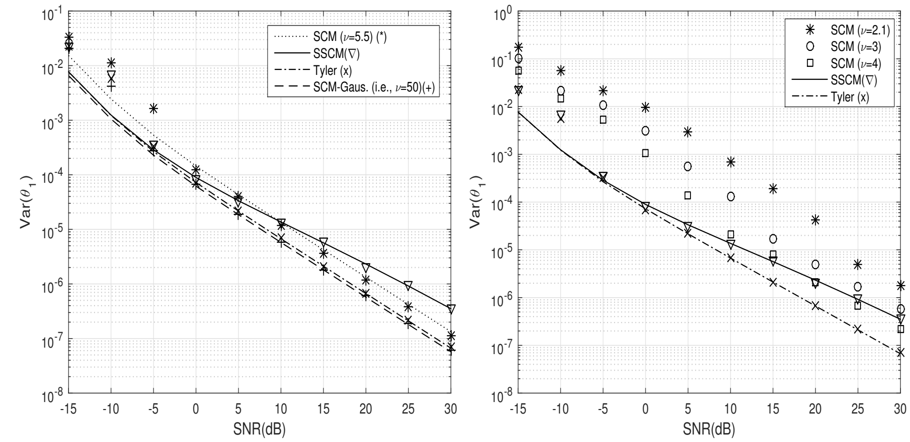
(a) r as a function of $\varDelta\theta$ for ASNR = 10dB (b) r as a function of ASNR for $\varDelta\theta = 0.2$rd

Fig. 2 Theoretical ratios $r_1 \stackrel{\text{def}}{=} [\mathbf{R}^{\text{ML,music}}_{\hat{\theta}}]_{1,1}/[\mathbf{R}^{\text{SCM,music}}_{\hat{\theta}}]_{1,1}$, $r_2 \stackrel{\text{def}}{=} [\mathbf{R}^{\text{ML,music}}_{\hat{\theta}}]_{1,1}/[\mathbf{R}^{\text{Ty,music}}_{\hat{\theta}}]_{1,1}$ and $r_3 \stackrel{\text{def}}{=} [\mathbf{R}^{\text{ML,music}}_{\hat{\theta}}]_{1,1}/[\mathbf{R}^{\text{SSCM,music}}_{\hat{\theta}}]_{1,1}$ for two values of m for $\nu = 4.1$.

In the third setup, the measurement $\mathbf{x}_i$ are complex non-circular Student t-distributed where $\mathbf{R}_{\tilde{x}}$ is given by (8) with $(\phi_1, \phi_2) = (\frac{\pi}{3}, \frac{2\pi}{3})$rd and $m = 6$. Fig.3a and 3b compare the theoretical asymptotic variance $\frac{1}{n}[\mathbf{R}^{\text{music}}_{\hat{\theta}}]_{1,1}$ and MSEs of non-circular MUSIC algorithms based on SCM, SSCM and Tyler's M estimate versus SNR. Note first that for $\nu \in [2,4]$, we get $\vartheta_{\text{Ty}} = 7/6$ and $\vartheta_{\text{ML}} \in [8/7, 9/8]$. So the asymptotic variance of Tyler's M estimator and the CRB are too close to be distinguishable in Fig.3. These figures also show that the theoretical asymptotic variances given by the non-circular MUSIC algorithms based on SSCM and Tyler's M estimates from (73), are very close to each other and to their MSE for a weak SNR and for $\nu > 4$. On the other hand, for $2 < \nu \leq 4$, for which the fourth-order moments of the data do not exist, and hence the asymptotic distribution of the non-circular MUSIC estimates based on the SCM is not available, the associated MSE increases strongly when ν approaches 2, for which the data are no longer of second-order.

7 Conclusion

The aim of this chapter is to unify the different performance analyses of subspace-based algorithms in C-CES and NC-CES data models within the same framework. In particular, general closed-form expressions are given for the covariances of the asymptotic distribution of different subspace projector estimates and the associated subspace-based parameter estimates. This result allows us to prove various invari-



(a) case $\nu > 4$ (b) case $2 < \nu \leq 4$

Fig. 3 Theoretical asymptotic variances $\frac{1}{n}[\mathbf{R}_{\hat{\theta}}^{\mathrm{SCM,music}}]_{1,1}$, $\frac{1}{n}[\mathbf{R}_{\hat{\theta}}^{\mathrm{Ty,music}}]_{1,1}$, $\frac{1}{n}[\mathbf{R}_{\hat{\theta}}^{\mathrm{SSCM,music}}]_{1,1}$ and MSEs (with 2000 Monte Carlo runs) of non-circular MUSIC algorithm , versus SNR where $T = 500$ and $\varDelta\theta = 0.2$rd for either $\nu > 4$ or $2 < \nu \leq 4$ and $m = 6$.

ance properties and general inequalities between covariances of subspace-based parameter estimates, including the stochastic and semiparametric CRB. Finally, note that the presented asymptotic distributions have been obtained w.r.t. the number of measurements. But the main shortcoming of these asymptotic distributions is that they provide good approximations of the variances of the estimates only when the sample size n is sufficiently large w.r.t. the data dimension m. However, when n is comparable to m, it is necessary to use random matrix theory tools to develop new subspace-based algorithms and their associated asymptotic distributions within the CES framework.

References

1. Abed-Meraim, K., Hua, Y.: Blind identification of multi-input multi-output system using minimum noise subspace. IIEEE Trans. Signal Process. **45**(1), 254–258 (1997)
2. Abeida, H., Delmas, J.P.: MUSIC-like estimation of direction of arrival for noncircular sources. IEEE Trans. Signal Process. **54**(7), 2678–2690 (2006)
3. Abeida, H., Delmas, J.P.: Robustness of subspace-based algorithms with respect to the distribution of the noise: Application to DOA estimations. Signal Processing **164**, 313–319 (2019)
4. Abeida, H., Delmas, J.P.: Slepian-Bangs formula and Cramér-Rao bound for circular and non-circular complex elliptical symmetric distributions. IEEE Signal Processing Letters **26**, 1561–1565 (2019)
5. Abeida, H., Delmas, J.P.: Efficiency of subspace-based estimators for elliptical symmetric distributions. Signal Processing **174** (2020)

6. Abeida, H., Delmas, J.P.: Performance of subspace-based algorithms associated with the sample sign covariance matrix. Digital Signal Processing **131** (2022)
7. Abeida, H., Delmas, J.P.: Slepian-Bangs formulas for parameterized density generator of elliptically symmetric distributions. Signal Processing **205** (2023)
8. Bausson, S., Pascal, F., Forster, P., Ovarlez, J.P., Larzabal, P.: First- and second-order moments of the normalized sample covariance matrix of spherically invariant random vectors. IEEE Signal Processing Letters **14**(6), 425–428 (2007)
9. Besson, O., Abramovich, Y.I.: On the Fisher information matrix for multivariate elliptically contoured distributions. IEEE Signal Processing Letters **20**(11), 1130–1133 (2013)
10. Conte, E., Lops, M., Ricci, G.: Adaptive radar detection in compound-gaussian clutter. In: EUSIPCO, pp. 526–529. Edinburgh, Scotland (1994)
11. Delmas, J.P.: Asymptotic performance of second-order algorithms. IEEE Transactions Signal Process. **50**(1), 49–57 (2002)
12. Delmas, J.P.: Performance bounds and statistical analysis of DOA estimation. Academic Press Library in Signal Processing (2013)
13. Delmas, J.P., Comon, P., Meurisse, Y.: Performance limits of alphabet diversities for FIR SISO channel identification. IEEE Transactions Signal Process. **57**(1), 73–82 (2009)
14. Di Claudio, E.D., Parisi, R., Jacovitti, G.: Space time MUSIC: consistent signal subspace estimation for wideband sensor arrays. IEEE Trans. Signal Process. **66**(10), 2685–2699 (2018)
15. Draskovic, G., Breloy, A., Pascal, F.: On the asymptotics of Maronna's robust PCA. IEEE Trans. Signal Process. **67**(19), 4964–4975 (2019)
16. Durre, A., Tyler, D.E., Vogel, D.: On the eigenvalues of the spatial sign covariance matrix in more than two dimensions. Statistics and Probability Letters **111**, 80–85 (2016)
17. Fortunati, S., Gini, F., Greco, M.S., Zoubir, A.M.: Semiparametric stochastic CRB for DOA estimation in elliptical data model. In: EUSIPCO. Coruña, Spain (2019)
18. Fortunati, S., Gini, F., Greco, M.S., Zoubir, A.M., Rangaswamy, M.: A fresh look at the semiparametric Cramér-Rao bound. In: EUSIPCO. Rome, Italy (2018)
19. Fortunati, S., Gini, F., Greco, M.S., Zoubir, A.M., Rangaswamy, M.: Semiparametric CRB and Slepian-Bangs formulas for complex elliptically symmetric distributions. IEEE Trans. Signal Process. **67**(20), 5352–5364 (2019)
20. Fortunati, S., Gini, F., Greco, M.S., Zoubir, A.M., Rangaswamy, M.: Semiparametric inference and lower bounds for real elliptically symmetric distributions. IEEE Trans. Signal Process. **67**(1), 164–177 (2019)
21. Fortunati, S., Renaux, A., Pascal, F.: Robust semiparametric efficient estimators in complex elliptically symmetric distributions. IEEE Trans. Signal Process. **68**, 5001–5015 (2020)
22. Gini, F., Greco, M.V.: Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter. Signal Processing **82**, 1847–1859 (2002)
23. Greco, M., Gini, F.: Cramer-Rao lower bounds on covariance matrix estimation for complex elliptically symmetric distributions. IEEE Trans. Signal Process. **61**, 6401–6409 (2013)
24. Haardt, M., Pesavento, M., Roemer, F., El Korso, M.N.: Subspace methods and exploitation of special array structures. pp. 651–717. Academic Press Library in Signal Processing, Elsevier (M. Viberg, ed.) (2014)
25. Huber, P.J.: The behavior of maximum likelihood estimates under nonstandard conditions. Proc. of the Fifth Berkeley Symposium in Mathematical Statistics and Probability, Berkley: University of California Press (1967)
26. Kato, T.: Perturbation Theory for Linear Operators. Springer Berlin (1995)
27. Kent, J.T.: Data analysis for shapes and images. J. Statist. Plann. Interference **57**(2), 181–197 (1997)
28. Krim, H., Forster, P., Proakis, G.: Operator approach to performance analysis of root-music and root-min-norm. EEE Trans. Signal Process. **40**(7), 1687–1696 (1992)
29. Krim, H., Viberg, M.: Two decades of array signal processing research: The parametric approach. IEEE Signal Processing Mag. **13**, 67–94 (1996)
30. Ladaycia, H., Abed-Meraim, K., Mokraoui, A., Belouchrani, A.: Efficient semi-blind subspace channel estimation for MIMO-OFDM system. In: EUSIPCO. Rome, Italy (2018)

31. Lehmann, E.L.: Elements of large sample theory. Springer texts in statistics (2004)
32. Locantore, M., et al: Robust principle component analysis for functional data. Test (8), 1–73 (1999)
33. Magyar, A.F.: The efficiencies of the spatial median and spatial sign covariance matrix for elliptically symmetric distributions. Ph.D. thesis, New Brunswick, State university of New Jersey (2012)
34. Magyar, A.F., Tyler, D.E.: The asymptotic inadmissibility of the spatial sign covariance matrix for the elliptically symmetric distributions. Biometrika **101**(3), 3673–688 (2014)
35. Mahot, M., Pascal, F., Forster, P., Ovarlez, J.P.: Asymptotic properties of robust complex covariance matrix estimates. IEEE Trans. Signal Process. **61**(13), 3348–3356 (2013)
36. Maronna, R.: Robust M-estimators of multivariate location and scatter. The annals of statistics **4**(1), 51–67 (1976)
37. Mecklenbräuker, C.F., Gerstoft, P., Ollila, E.: DOA M-estimation using sparse bayesian learning. In: ICASSP. Singapore (2022)
38. Moulines, E., Duhamel, P., Cardoso, J.F., Mayrargue, S.: Subspace methods for the blind identification FIR filters. IEEE Trans. Signal Process. **43**(2), 516–525 (1995)
39. Ollila, E., Koivunen, V.: Robust antenna array processing using M-estimators of pseudo covariance. In: 14th lnternational Symposium on Personal lndoor and Mobile Radio Communication. Toronto, Canada (2003)
40. Ollila, E., Koivunen, V.: Influence function and asymptotic efficiency of scatter matrix based array processors: Case MVDR beamformer. IEEE Trans. Signal Process. **57**(1), 247–259 (2009)
41. Ollila, E., Oja, H., Croux, C.: The affine equivariant sign covariance matrix: asymptotic behavior and efficiencies. Journal of Multivariate Analysis **87**, 328–355 (2003)
42. Ollila, E., Tyler, D.E., Koivunen, V., Poor, H.V.: Complex elliptically symmetric distributions: Survey, new results and applications. IEEE Trans. Signal Process. **60**(11), 5597–5625 (2012)
43. Paindaveine, D.: Elliptical symmetry. In book Wiley Statistics Reference Online (2014)
44. Pascal, F., Forster, P., Ovarlez, J.P., Larzabal, P.: Performance analysis of covariance matrix estimates in impulsive noise. IEEE Trans. on Signal Process. **56**(6), 2206–2216 (2008)
45. Porat, B., Friedlander, B.: Asymptotic accuracy of ARMA parameter estimation methods based on sample covariances. In: 7th IFAC/IFORS Symposium on Identification and System Parameter Estimation, York. York (1985)
46. Schmidt, R.O.: Multiple emitter location and signal parameter estimation. IEEE Trans. Antennas Propagation **34**, 276–280 (1986)
47. Serfling, R.J.: Approximation Theorems of Mathematical Statistics. John Wiley and Sons (1980)
48. Stoica, P., Friedlander, B., Söderström, T.: An approximate maximum approach to ARMA spectral estimation. In: 24th IEEE Conference on Decision and Control. Fort Lauderdale (1985)
49. Stoica, P., Nehorai, A.: Performances study of conditional and unconditional direction of arrival estimation. IEEE Trans. Signal Process. **38**(10), 1783–1795 (1990)
50. Tyler, D.E.: Radial estimates and the test for sphericity. Biometrika **69**(2), 429 (1982)
51. Tyler, D.E.: A distribution-free M-estimator of multivariate scatter. The Annals of Statistics **15**(1), 234–251 (1987)
52. Tyler, D.E.: Statistical analysis for the angular central gaussian distribution on the sphere. Biometrika **74**(3), 579–589 (1987)
53. Van Trees, H.L.: Optimum array processing Part IV of detection, estimation, and modulation theory. Wiley Interscience, John Wiley and Sons, New York (2002)
54. Visuri, S., Koivunen, V., Oja, H.: Sign and rank covariance matrices. Journal of Statistical Planning and Inference **91**, 557–575 (2000)
55. Visuri, S., Oja, H., Koivunen, V.: Subspace-based direction of arrival estimation using nonparametric statistics. IEEE Trans. Signal Process. **49**(9), 2060–2073 (2001)
56. Zuo, W., Xin, J., Ohmori, H., Zheng, N., Sano, A.: Subspace-based algorithms for localization and tracking of multiple near-field sources. IEEE Journal of selected topics in Signal Processing **13**(1) (2019)