HAL open science

Background on real and complex elliptically symmetric distributions

Jean-Pierre Delmas

To cite this version:

Jean-Pierre Delmas. Background on real and complex elliptically symmetric distributions. Springer. Elliptically symmetric distributions in Signal Processing, inPress. hal-04217510

HAL Id: hal-04217510

https://hal.science/hal-04217510

Submitted on 25 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Background on real and complex elliptically symmetric distributions

Preprint chapter to be published in the book "Elliptically symmetric distributions in Signal Processing" by Springer

Jean-Pierre Delmas

Abstract This chapter presents a short overview of real elliptically symmetric (RES) distributions, complemented by circular complex elliptically symmetric (C-CES) and noncircular CES (NC-CES) distributions as complex representations of RES distributions. These distributions are both an extension of the multivariate Gaussian distribution and a multivariate extension of univariate symmetric distributions. They are equivalently defined through their characteristic functions and their stochastic representations, which naturally follow from the spherically symmetric distributions after affine transformations. Particular attention is paid to the absolutely continuous case and to the subclass of compound Gaussian distributions. Results related to moments, affine transformations, marginal and conditional distributions, and summation stability are also presented. Some well-known instances of RES distributions are provided with their main properties. Finally, the estimation of the symmetry center and scatter matrix is briefly discussed through the sample mean (SM), sample covariance matrix (SCM) estimate, maximum estimate (ML), M-estimators, and Tyler's M-estimators. Particular attention will be paid to the asymptotic Gaussianity of the M-estimators of the scatter matrix. To conclude, some hints about the Slepian-Bangs formula are provided.

1 Introduction

Until fifty years ago, most of the procedures in multivariate analysis were developed under the Gaussian assumption, mainly for mathematical convenience. However, in many applications, Gaussianity is a poor approximation of reality. As a consequence, elliptically symmetric distributions have been widely used in various applications due to their flexibility and capability to better model various data behavior. These

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

distributions form a natural extension of the Gaussian one by allowing for both heavier-than-Gaussian and lighter-than-Gaussian tails while maintaining the elliptical geometry of the underlying equidensity (when it exists) contours. These real elliptically symmetric (RES) distributions were equivalently defined in the statistical literature [9, 16, 17, 32] through their characteristic functions and their stochastic representations, which naturally follow from the spherically symmetric distributions after affine transformations. A first systematic treatment of circular complex elliptically symmetric (C-CES) distributions was provided in the engineering literature [36] and further fully studied in [54], and in the tutorial paper [48]. Then, the general complex representation of the RES distributions, called noncircular CES (NC-CES) distributions, was introduced in [15].

The aim of this chapter is twofold. At first, a short overview of RES, C-CES, and NC-CES distributions (as complex representations of the RES distributions) is introduced with the aim of providing a common background for the other chapters of this book. Secondly, the main definitions and properties of these distributions are listed and shortly discussed.

This chapter is organized as follows. Section 2 defines the RES distributions equivalently through their characteristic functions and to their stochastic representations. Particular attention is paid to the absolutely continuous case and to the subclass of compound Gaussian distributions. Section 3 defines the C-CES and NC-CES distributions as complex representations of the RES distributions. Section 4 presents basic properties related to moments, affine transformations, marginal and conditional distributions, and summation stability. Then some well-known instances of RES distributions are provided with their main properties in Section 5. The joint estimation of the symmetry center and scatter matrix is briefly discussed in Section 6 through the SM and SCM estimators, ML, M-estimators, and Tyler's M-estimators, asymptotic Gaussian distribution of scatter M-estimators, and Slepian-Bangs formula. Finally, Section 7 briefly concludes this chapter.

The following notations are used throughout this chapter. Matrices and vectors are represented by bold upper case and bold lower case characters, respectively. Vectors are by default in column orientation, while the superscripts T, H, $*$, and $\#$ stand for transpose, conjugate transpose, conjugate, and Moore Penrose inverse, respectively. $(\mathbf{a})_k$ and $(\mathbf{A})_{k,\ell}$ denote the k and (k,ℓ)-th element of the vector $\mathbf{a}$ and the matrix $\mathbf{A}$, respectively. $\mathrm{E}(.)$, $|.|$, and $\mathrm{Tr}(.)$ are the expectation, determinant, and trace operators, respectively. $\mathbf{I}$ is the identity matrix with the appropriate dimension. $\mathrm{vec}(\mathbf{A})$ denotes the "vectorization" operator that turns a matrix $\mathbf{A}$ into a vector by stacking the columns of the matrix one below another and $v(\mathbf{A})$ is the vector that is obtained from $\mathrm{vec}(\mathbf{A})$ by eliminating all supradiagonal elements of $\mathbf{A}$. These vectors are used in conjunction with the Kronecker product $\mathbf{A}\otimes\mathbf{B}$ as the block matrix whose (i,j) block element is $a_{i,j}\mathbf{B}$, and with the commutation matrix $\mathbf{K}$ and the duplication matrix $\mathbf{D}$ of appropriate dimension such that $\mathrm{vec}(\mathbf{C}^T) = \mathbf{K}\mathrm{vec}(\mathbf{C})$ and $\mathrm{vec}(\mathbf{A}) = \mathbf{D}v(\mathbf{A})$ where $\mathbf{A}$ is symmetric. The acronyms r.v., p.d.f. and c.d.f. for respectively random variable, probability density function and cumulative distribution function are used. Finally, $\Gamma(u) \stackrel{\mathrm{def}}{=} \int_0^\infty t^{u-1}e^{-t}dt$ is the Gamma function with $\Gamma(k) = (k-1)!$ for

$k \in \mathbb{N}$, $B(k,\ell)$ denotes the Beta function with $B(k,\ell) = \frac{\Gamma(k)\Gamma(\ell)}{\Gamma(k+\ell)}$ and $\mathrm{Gam}(k,\theta)$ is the Gamma distribution of scale θ with p.d.f. $p(x) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} \exp(-x/\theta)$. $x =_d y$, $x_n \to_d D$ and $x \sim D$ mean that the r.v. x and y have the same distribution, the sequence of r.v. x_n converges in distribution to D and x follows the distribution D, respectively. The subscripts r and c are used to refer to the real and complex data cases, respectively.

2 Definition of the real elliptically symmetric distributions

2.1 Characteristic function

After earlier works on this topic, RES distributions were formalized by [32] and further studied by [9, 16, 17]. They were first defined as affine transformations of spherically distributed r.v. Then, by the uniqueness theorem (see, e.g., [55, pp. 346-351], they were alternatively defined by their characteristic functions.

Definition 1 An r.v. $\mathbf{x} \in \mathbb{R}^m$ is said to have a RES distribution if there exists a vector $\boldsymbol{\mu} \in \mathbb{R}^m$, an $m \times m$ symmetric positive semi-definite matrix $\boldsymbol{\Sigma}$ of rank $k \leq m$ and a function $\phi_r(.) : \mathbb{R}^+ \to \mathbb{R}$ called *symmetry center*, *scatter matrix* and *characteristic generator*, respectively, such that the characteristic function of $\mathbf{x}$ is of the form

$$\Phi_x(\mathbf{t}) \stackrel{\text{def}}{=} \mathrm{E}[\exp(i\mathbf{t}^T\mathbf{x})] = \exp(i\mathbf{t}^T\boldsymbol{\mu})\phi_r(\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}), \qquad \mathbf{t} \in \mathbb{R}^m. \tag{1}$$

We shall write $\mathbf{x} \sim \mathrm{RES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi_r)$ and note that the couple $(\boldsymbol{\Sigma}, \phi_r(.))$ does not uniquely identify the distribution of $\mathbf{x}$ because $(c^2\boldsymbol{\Sigma}, \phi_r(./c^2))$ gives the same distribution. This scale ambiguity is easily avoided by restricting the function $\phi_r(.)$ in a suitable way (e.g., by fixing a moment as it is explained in Section 4.1), or by putting a constraint on the scatter matrix $\boldsymbol{\Sigma}$ (e.g., $\mathrm{Tr}(\boldsymbol{\Sigma}) = m$). Note that for $m = 1$, these distributions coincide with the class of one-dimensional symmetric distributions w.r.t. the symmetry center.

2.2 Stochastic representation

Equivalently to the definition (1), the RES distributed r.v. $\mathbf{x}$ can be defined from an affine function

$$\mathbf{x} \stackrel{\text{def}}{=} \boldsymbol{\mu} + \mathbf{A}\mathbf{x}_s \tag{2}$$

of a k-dimensional spherically distributed r.v. $\mathbf{x}_s$, where $\mathbf{A} \in \mathbb{R}^{m\times k}$ is any square root ($\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^T$) of the scatter matrix $\boldsymbol{\Sigma}$ of rank k, and thus full column rank. Such spherically distributions are defined equivalently in the following [16, Chap.2]

Definition 2 An r.v. $\mathbf{x}_s \in \mathbb{R}^k$ is spherically distributed i.f.f.

- $\mathbf{x}_s =_d \mathcal{O}\mathbf{x}_s$ for arbitrary real-valued k-dimensional orthonormal matrix $\mathcal{O}$,
- There exists a function $\phi_r(.) : \mathbb{R}^+ \mapsto \mathbb{R}$, such that the characteristic function of $\mathbf{x}_s$ is given by
$$\Phi_{x_s}(\mathbf{t}) = \phi_r(\|\mathbf{t}\|^2),\ \mathbf{t} \in \mathbb{R}^k, \tag{3}$$
- For every $\mathbf{h} \in \mathbb{R}^k$, $\mathbf{h}^T\mathbf{x}_s =_d \|\mathbf{h}\| x_{s_i}$ with $\mathbf{x}_s = (x_{s_1}, .., x_{s_i}, .., x_{s_k})^T$,
- If $\mathbf{x}_s$ is absolutely continuous w.r.t. Lebesgue measure on $\mathbb{R}^k$, there exists a function[1] $g_r(.)$: $\mathbb{R}^+ \mapsto \mathbb{R}^+$ such that the p.d.f. of $\mathbf{x}_s$ w.r.t. this measure is of the form
$$p(\mathbf{x}_s) = g_r(\|\mathbf{x}_s\|^2), \tag{4}$$
where
$$\delta_{r,k} \stackrel{\text{def}}{=} \int_0^\infty t^{k/2-1} g_r(t)dt = \frac{\Gamma(k/2)}{\pi^{k/2}}, \tag{5}$$
ensuring that $p(\mathbf{x}_s)$ integrates to one.
- There exists a non-negative r.v. $\mathcal{Q}_{r,k}$, and $\mathbf{u}_{r,k}$ that are independent where $\mathbf{u}_{r,k}$ is uniformly distributed on the unit real k-sphere ($\mathbf{u}_{r,k} \sim U(\mathbb{R}S^k)$) such that
$$\mathbf{x}_s =_d \sqrt{\mathcal{Q}_{r,k}}\mathbf{u}_{r,k}. \tag{6}$$

Consequently from (2) and (3), we find the characteristic function (1) which defines the RES distribution since

$$\Phi_x(\mathbf{t}) = \exp(i\mathbf{t}^T\boldsymbol{\mu})\Phi_{x_s}(\mathbf{A}^T\mathbf{t}) = \exp(i\mathbf{t}^T\boldsymbol{\mu})\phi_r(\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}).$$

From (2) and (6), we obtain the following

Theorem 1 $\mathbf{x}$ *is* $\mathrm{RES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi_r)$ *distributed, i.f.f. it admits the following stochastic full-rank representation*

$$\mathbf{x} =_d \boldsymbol{\mu} + \sqrt{\mathcal{Q}_{r,k}}\mathbf{A}\mathbf{u}_{r,k} = \boldsymbol{\mu} + \mathcal{R}_{r,k}\mathbf{A}\mathbf{u}_{r,k}. \tag{7}$$

The r.v. $\mathcal{Q}_{r,k}$ and $\mathcal{R}_{r,k} \stackrel{\text{def}}{=} \sqrt{\mathcal{Q}_{r,k}}$ are the 2nd-order modular and modular (or generating) variates of the r.v. $\mathbf{x}$, respectively. We note that there is a one-to-one mapping between the characteristic generator ϕ_r and the c.d.f. F_{R_r} of $\mathcal{R}_{r,k}$ (called generating c.d.f.). Thus we can also write $\mathbf{x} \sim \mathrm{RES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, F_{R_r})$ and, equivalently to Definition 1, we retrieve the scale ambiguity in the couple $(\mathbf{A}, \mathcal{R}_{r,k})$ in (7). Theorem 1 provides an obvious mechanism to generate r.v. $\mathbf{x} \sim \mathrm{RES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, F_{R_r})$: it only involves generating $\mathcal{R}_{r,k}$ according to its c.d.f. F_{R_r} and $\mathbf{u}_{r,k} = \frac{\mathbf{n}_{r,k}}{\|\mathbf{n}_{r,k}\|}$ where $\mathbf{n}_{r,k}$ is k-dimensional zero-mean Gaussian distributed r. v. with covariance $\mathbf{I}$ ($\mathbf{n}_{r,k} \sim \mathbb{R}N_k(\mathbf{0}, \mathbf{I})$). Moreover, the following important property follows from Theorem 1:
$$\mathcal{Q}_{r,k} =_d (\mathbf{x} - \boldsymbol{\mu})^T\boldsymbol{\Sigma}^{\#}(\mathbf{x} - \boldsymbol{\mu}). \tag{8}$$

[1] both functions $\phi_r(.)$ and $g_r(.)$ are generally parameterized by the dimension k and in practice by a finite-dimensional parameter (see examples in Section 5).

2.3 The absolutely continuous case

From (2) the r.v. $\mathbf{x}$ is absolutely continuous w.r.t. Lebesgue measure on $\mathbb{R}^m$, i.f.f. $\mathbf{x}_s$ is too. From (6), it is immediate to verify that this condition is satisfied i.f.f. $\mathcal{Q}_{r,k}$ or $\mathcal{R}_{r,k}$ is absolutely continuous w.r.t. Lebesgue measure on $\mathbb{R}^+$. In this case, the p.d.f. of $\mathbf{x}$ is defined on the k-dimensional subspace of $\mathbb{R}^m$ spanned by the range space of $\mathbf{A}$.

In the particular case where $k = m$, i.e., rank$(\mathbf{\Sigma}) = m$, $\mathbf{A}$ is a non singular $m \times m$ square matrix. From the one to one mapping (2) between $\mathbf{x}$ and $\mathbf{x}_s$, and the p.d.f. (4), the p.d.f. of $\mathbf{x}$ on $\mathbb{R}^m$ can be expressed as:

$$p(\mathbf{x}) = |\mathbf{\Sigma}|^{-1/2} g_r[(\mathbf{x} - \boldsymbol{\mu})^T \mathbf{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})]. \tag{9}$$

We note that unlike the notation used in, e.g., [48], (9) does not explicitly include the usual p.d.f. normalizing constant. This is a consequence of the definition of $g_r(.)$ introduced in (4).

Alternatively, if in (9) $g_r(.)$: $\mathbb{R}^+ \mapsto \mathbb{R}^+$ is an arbitrary function, called *density generator*, such that

$$\delta_{r,m,g} \stackrel{\text{def}}{=} \int_0^\infty t^{m/2-1} g_r(t) dt < \infty, \tag{10}$$

$\delta_{r,m,g}$ depends now on g_r and the couple $(\mathbf{\Sigma}, g_r(.))$ in (9) does not uniquely identify the distribution of $\mathbf{x}$ because $(c^2\mathbf{\Sigma}, c^m g_r(.\ c^2))$ gives the same distribution.

We adopt the notation $\mathrm{RES}_m(\boldsymbol{\mu}, \mathbf{\Sigma}, g)$ instead of $\mathrm{RES}_m(\boldsymbol{\mu}, \mathbf{\Sigma}, \phi)$. The level sets of $p(\mathbf{x})$ are a family of hyper ellipsoids in $\mathbb{R}^m$ symmetrically centered at $\boldsymbol{\mu}$, where shape and orientation are determined by $\mathbf{\Sigma}$. This justifies the terminology of symmetrical elliptical distributions. Furthermore, the stochastic representation (7) reduces to

$$\mathbf{x} =_d \boldsymbol{\mu} + \sqrt{\mathcal{Q}_{r,m}} \mathbf{\Sigma}^{1/2} \mathbf{u}_{r,m} = \boldsymbol{\mu} + \mathcal{R}_{r,m} \mathbf{\Sigma}^{1/2} \mathbf{u}_{r,m}. \tag{11}$$

Here too, note that $(\mathbf{\Sigma}, \mathcal{Q}_{r,m})$ and $(c^2\mathbf{\Sigma}, c^{-2}\mathcal{Q}_{r,m})$ give the same distribution of $\mathbf{x}$. (11) implies that (8) simplifies to

$$\mathcal{Q}_{r,m} =_d (\mathbf{x} - \boldsymbol{\mu})^T \mathbf{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}). \tag{12}$$

From (9), (11) and (12), the p.d.f. of $\mathcal{Q}_{r,m}$ and $\mathcal{R}_{r,m}$ are respectively

$$p(q) = \delta_{r,m}^{-1} q^{m/2-1} g_r(q) \ \text{ and } \ p(r) = 2\delta_{r,m}^{-1} r^{m-1} g_r(r^2). \tag{13}$$

Finally, we note that the RES distributions do not necessarily possess a p.d.f. w.r.t. Lebesgue measure on $\mathbb{R}^m$ even when $\mathbf{\Sigma}$ is not singular. Such an example is the $U(\mathbb{R}S^m)$ distribution which belongs to $\mathrm{RES}_m(\mathbf{0}, \mathbf{I}, \phi_r)$ distributions, where the explicit (but somewhat involved) form of ϕ_r can be found in [17, the. 2.51].

2.4 The subclass of compound-Gaussian distributions

An important subclass of RES distributions are the compound-Gaussian (CG) distributions, whose circular complex representations (denoted C-CCG) have been widely used in the engineering literature, for example, for modeling radar clutter [64]. An r.v. having CG distributions with zero symmetry center is also referred to as *spherically invariant random vectors* (SIRV) in the engineering literature (see, e.g., in [11, 12, 53, 65]) and as *scale mixtures of normal distributions* in the statistics literature [4] [24]. These distributions are defined by their stochastic representation.

Definition 3 An r.v. $\mathbf{x} \in \mathbb{R}^m$ is said to have a real CG distribution (RCG) if it admits the following representation

$$\mathbf{x} =_d \boldsymbol{\mu} + \sqrt{\tau_r}\mathbf{n}, \tag{14}$$

for some positive real r.v. τ_r with c.d.f. F_τ (not related neither to dimension m nor to rank k), called the *texture* independent of $\mathbf{n} \sim \mathbb{R}N_m(\mathbf{0}, \boldsymbol{\Sigma})$, called the *speckle*. The r.v. $\sqrt{\tau_r}$ is often called *mixing variable* with mixing distribution in the statistical literature (see e.g., [24]). We write $\mathbf{x} \sim \mathrm{RCG}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, F_\tau)$ to denote this case.

Note that (14) can be rewritten as

$$\mathbf{x} =_d \boldsymbol{\mu} + \sqrt{\tau_r}\mathbf{A}\mathbf{n}_0, \tag{15}$$

where $\mathbf{n}_0 \sim \mathbb{R}N_k(\mathbf{0}, \mathbf{I})$. Then by recalling that $\mathbf{n}_0 = \|\mathbf{n}_0\|\mathbf{u}_{r,k}$ with $s \stackrel{\mathrm{def}}{=} \|\mathbf{n}_0\|^2 \sim \chi_k^2 = \mathrm{Gam}(k/2, 2)$ and $\mathbf{u}_{r,k} \sim U(\mathbb{R}S^k)$, and where s and $\mathbf{u}_{r,k}$ are independent. It follows that the stochastic representation (14) can also be written as

$$\mathbf{x} =_d \boldsymbol{\mu} + \mathcal{R}_{r,k}\mathbf{A}\mathbf{u}_{r,k}, \tag{16}$$

where the modular variate $\mathcal{R}_{r,k} \stackrel{\mathrm{def}}{=} \sqrt{\tau_r s}$ and $\mathbf{u}_{r,k}$ are independent. Consequently, the RCG distributions form a subclass of the RES distributions. Furthermore, $\mathbf{x} \sim \mathrm{RES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$ belongs to the set of RCG distributions i.f.f. there exists an r.v. τ_r such that the 2nd-order modular variate $\mathcal{Q}_{r,k} = \mathcal{R}_{r,k}^2$ satisfies $\mathcal{Q}_{r,k} = \tau_r s$, i.e., $\mathcal{Q}_{r,k}$ is a scale mixture of the $\mathrm{Gam}(k/2, 2)$ distribution. This means that the conditional distribution of $\mathcal{Q}_{r,k}$ given $\tau_r = \tau$ is the $\mathrm{Gam}(k/2, 2\tau)$ distribution and thus the p.d.f of $\mathcal{Q}_{r,k}$ is

$$p(q) = \int_0^\infty \frac{1}{\Gamma(k/2)(2\tau)^{k/2}} q^{k/2-1} \exp(-q/(2\tau)) dF_\tau(\tau). \tag{17}$$

The characteristic function of $\Phi_{x_s}(\mathbf{t})$ of a RCG distributed r.v. $\mathbf{x}$ defined by (14) is given straightforwardly by

$$\begin{aligned}\Phi_{x_s}(\mathbf{t}) &= \exp(i\mathbf{t}^T\boldsymbol{\mu})\int_0^\infty \mathrm{E}(\exp(i\mathbf{t}^T\sqrt{\tau}\mathbf{n})/\tau)dF_\tau(\tau)\\ &= \exp(i\mathbf{t}^T\boldsymbol{\mu})\int_0^\infty \exp\left(-\frac{\tau}{2}\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}\right)dF_\tau(\tau)\\ &= \exp(i\mathbf{t}^T\boldsymbol{\mu})\phi_r(\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}) \qquad (18)\end{aligned}$$

where the characteristic generator

$$\phi_r(u) = \int_0^\infty \exp\left(-\frac{\tau}{2}u\right)dF_\tau(\tau) \qquad (19)$$

does not depend on the dimension m nor on the rank k, unlike the RES distributions which are not RCG whose characteristic generator can depend on it.

In the particular case where $\boldsymbol{\Sigma}$ is not singular $(k = m)$, the conditional distribution of $\mathbf{x}$ given $\tau_r = \tau$ is the $\mathbb{R}N_m(\boldsymbol{\mu}, \tau\boldsymbol{\Sigma})$ distribution from (14). Consequently, the distribution of $\mathbf{x}$ is always continuous w.r.t. Lebesgue measure on $\mathbb{R}^m$ and its p.d.f. is given by

$$p(\mathbf{x}) = (2\pi)^{-m/2}|\boldsymbol{\Sigma}|^{-1/2}\int_0^\infty \tau^{-m/2}\exp\left(-\frac{1}{2\tau}(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)dF_\tau(\tau). \qquad (20)$$

Note that the p.d.f. (20) can always be written in the form (9) with the density generator

$$g_r(t) = (2\pi)^{-m/2}\int_0^\infty \tau^{-m/2}\exp\left(-\frac{t}{2\tau}\right)dF_\tau(\tau), \qquad (21)$$

and similarly to the RES distributions, we are faced with scale ambiguity where $(c^2\boldsymbol{\Sigma}, c^{-2}\tau_r, c^m g_r(.c^2))$ gives the same R-CG distribution. Note that (21) reduces to the density generator (69) of the Gaussian distribution when τ_r is a degenerate r.v. putting all the probability at $\tau_r = 1$. Note also that the ϵ-contaminated Gaussian distribution belongs to the class of CG distributions and is obtained when τ_r is a discrete r.v. with $P(\tau_r = a^2) = \epsilon$ and $P(\tau_r = 1) = 1 - \epsilon$, where (a^2, ϵ) are parameters that control the heaviness of the tails as compared to the Gaussian distribution.

3 Definition of the complex elliptically symmetric distributions

3.1 Characteristic function

An r.v. $\mathbf{x} \in \mathbb{C}^m$ is said to have a noncircular complex elliptically symmetric (NC-CES) distribution (also called generalized complex elliptical in [46] if the r.v. $\bar{\mathbf{x}} \stackrel{\text{def}}{=} (\mathrm{Re}(\mathbf{x})^T, \mathrm{Im}(\mathbf{x})^T)^T \in \mathbb{R}^{2m}$ is RES distributed. Denote the symmetry center and the scatter matrix (of rank $k \leq 2m$) of $\bar{\mathbf{x}}$ by $\bar{\boldsymbol{\mu}} \in \mathbb{R}^{2m}$ and $\bar{\boldsymbol{\Sigma}} \in \mathbb{R}^{2m\times 2m}$, respectively.

Using the one-to-one mapping $\bar{\mathbf{x}} \mapsto \widetilde{\mathbf{x}} \stackrel{\text{def}}{=} (\mathbf{x}^T, \mathbf{x}^H)^T = \sqrt{2}\mathbf{M}\bar{\mathbf{x}}$ 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, we obtain $\bar{\mathbf{t}}^T\bar{\boldsymbol{\mu}} = \mathrm{Re}(\mathbf{t}^H\boldsymbol{\mu})$, $\bar{\mathbf{t}}^T\bar{\boldsymbol{\Sigma}}\bar{\mathbf{t}} = \frac{1}{2}\widetilde{\mathbf{t}}^H(\mathbf{M}\bar{\boldsymbol{\Sigma}}\mathbf{M}^H)\widetilde{\mathbf{t}} = \frac{1}{4}\widetilde{\mathbf{t}}^H\widetilde{\boldsymbol{\Sigma}}\widetilde{\mathbf{t}}$ with $\bar{\mathbf{t}} \stackrel{\text{def}}{=} (\mathrm{Re}(\mathbf{t})^T, \mathrm{Im}(\mathbf{t})^T)^T$, $\bar{\boldsymbol{\mu}} \stackrel{\text{def}}{=} (\mathrm{Re}(\boldsymbol{\mu})^T, \mathrm{Im}(\boldsymbol{\mu})^T)^T$, $\widetilde{\mathbf{t}} \stackrel{\text{def}}{=} (\mathbf{t}^T, \mathbf{t}^H)^T$ and $\widetilde{\boldsymbol{\Sigma}} \stackrel{\text{def}}{=} 2\mathbf{M}\bar{\boldsymbol{\Sigma}}\mathbf{M}^H$ of rank k structured as

$$\widetilde{\boldsymbol{\Sigma}} = \begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{\Omega} \\ \boldsymbol{\Omega}^* & \boldsymbol{\Sigma}^* \end{pmatrix}, \tag{22}$$

where $\boldsymbol{\Sigma}$ and $\boldsymbol{\Omega}$ defined from $2\mathbf{M}\bar{\boldsymbol{\Sigma}}\mathbf{M}^H$, are positive semi-definite Hermitian and complex symmetric matrices, respectively. Consequently, we obtain the following theorem by the definition (1)

Theorem 2 *The characteristic function of an NC-CES distributed r.v.* $\mathbf{x} \in \mathbb{C}^m$ *is of the form*

$$\Phi_x(\mathbf{t}) \stackrel{\text{def}}{=} \mathrm{E}[\exp(i\bar{\mathbf{t}}^T\bar{\mathbf{x}})] = \exp(i\mathrm{Re}(\mathbf{t}^H\boldsymbol{\mu}))\phi_c\left(\frac{1}{2}\widetilde{\mathbf{t}}^H\widetilde{\boldsymbol{\Sigma}}\widetilde{\mathbf{t}}\right) \qquad \mathbf{t} \in \mathbb{C}^m, \tag{23}$$

where

$$\phi_c(u) \stackrel{\text{def}}{=} \phi_r\left(\frac{1}{2}u\right) \tag{24}$$

is the characteristic generator, and where $\boldsymbol{\mu} \in \mathbb{C}^m$ *and* $\widetilde{\boldsymbol{\Sigma}} \in \mathbb{C}^{2m\times 2m}$ *denote respectively the symmetric center and the extended scatter matrix of the NC-CES distributed r.v.* $\mathbf{x}$.

We shall write $\mathbf{x} \sim \mathrm{CES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Omega}, \phi_c)$. In the particular case where $\boldsymbol{\Omega} = \mathbf{0}$, the rank of $\widetilde{\boldsymbol{\Sigma}}$ is even with $\mathrm{rank}(\widetilde{\boldsymbol{\Sigma}}) = 2\mathrm{rank}(\boldsymbol{\Sigma}) = k$ and $\mathbf{x}$ is C-CES distributed [36, 48, 54]. The term *circular* is often dropped in the current terminology used in signal processing where the distribution of a C-CES r.v. is usually indicated as $\mathbf{x} \sim \mathrm{CES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi_c)$. Moreover, from (23), we get:

Theorem 3 *The characteristic function of a C-CES distributed r.v.* $\mathbf{x} \in \mathbb{C}^m$ *is of the form*

$$\Phi_x(\mathbf{t}) = \exp(i\mathrm{Re}(\mathbf{t}^H\boldsymbol{\mu}))\phi_c(\mathbf{t}^H\boldsymbol{\Sigma}\mathbf{t}) \qquad \mathbf{t} \in \mathbb{C}^m, \tag{25}$$

where $\boldsymbol{\mu} \in \mathbb{C}^m$ *and* $\boldsymbol{\Sigma} \in \mathbb{C}^{m\times m}$ *denote the symmetric center, and the scatter matrix of the C-CES distributed r.v.* $\mathbf{x}$*, respectively.*

3.2 Stochastic representation

From the definition of the NC-CES distribution, a simple complex-valued extension of the stochastic representation (7) is only possible if the rank of $\widetilde{\boldsymbol{\Sigma}}$, which is equal to the rank of $\bar{\boldsymbol{\Sigma}}$, is even (it is, in particular, the case of the C-CES distribution and the case where $\widetilde{\boldsymbol{\Sigma}}$ is not singular for which $k = 2m$). Let $2k$ be the rank of $\widetilde{\boldsymbol{\Sigma}}$. In

this case, there exists an $m \times k$ full column rank matrix $\mathbf{A}$ such that $\mathbf{\Sigma} = \mathbf{A}\mathbf{A}^H$ and $\mathbf{\Omega} = \mathbf{A}\mathbf{\Delta}_\kappa\mathbf{A}^T$ where $\mathbf{\Delta}_\kappa = \mathrm{Diag}(\kappa_1, \ldots, \kappa_k)$ is a real diagonal matrix with non-negative real entries $(\kappa_i)_{i=1,..,k}$ [27, Corollary 4.6.12(b)]. Furthermore, it has been proved in [15], that $0 \leq \kappa_i \leq 1$. This parameterization allows us to state that the stochastic representation of this distribution, proved in [1], is a multivariate extension of the univariate generation of NC-CES r.v. presented in [45, sec. IV.C].

Theorem 4 $\mathbf{x}$ *is* $\mathrm{CES}_m(\boldsymbol{\mu}, \mathbf{\Sigma}, \mathbf{\Omega}, \phi_c)$ *distributed i.f.f. it admits the following stochastic representation*

$$\mathbf{x} =_d \boldsymbol{\mu} + \sqrt{\mathcal{Q}_{c,k}}\mathbf{A}[\mathbf{\Delta}_1\mathbf{u}_{c,k} + \mathbf{\Delta}_2\mathbf{u}^*_{c,k}], \tag{26}$$

where $\mathcal{Q}_{c,k} \stackrel{\mathrm{def}}{=} \frac{1}{2}\mathcal{Q}_{r,2k}$ *and* $\mathbf{u}_{c,k} \sim U(\mathbb{C}S^k)$ *are independent,* $\mathbf{\Delta}_1 \stackrel{\mathrm{def}}{=} \frac{\mathbf{\Delta}_+ + \mathbf{\Delta}_-}{2}$ *and* $\mathbf{\Delta}_2 \stackrel{\mathrm{def}}{=} \frac{\mathbf{\Delta}_+ - \mathbf{\Delta}_-}{2}$ *where* $\mathbf{\Delta}_+ \stackrel{\mathrm{def}}{=} \sqrt{\mathbf{I} + \mathbf{\Delta}_\kappa}$ *and* $\mathbf{\Delta}_- \stackrel{\mathrm{def}}{=} \sqrt{\mathbf{I} - \mathbf{\Delta}_\kappa}$.

In the particular case of C-CES distributions, $\mathbf{\Omega} = \mathbf{0}$, which is equivalent to $\mathbf{\Delta}_\kappa = \mathbf{0}$, i.e., $\mathbf{\Delta}_1 = \mathbf{I}$ and $\mathbf{\Delta}_2 = \mathbf{0}$ and consequently the stochastic representation (26) reduces to the well known stochastic representation reported in [48]:

$$\mathbf{x} =_d \boldsymbol{\mu} + \sqrt{\mathcal{Q}_{c,k}}\mathbf{A}\mathbf{u}_{c,k}. \tag{27}$$

Note that similarly to the RES distribution, the C-CES distribution can be defined from the affine function (2), where here $\mathbf{x}_s$ is k-dimensional spherically distributed defined by the equality $\mathbf{x}_s =_d \mathcal{U}\mathbf{x}_s$ for arbitrary complex-valued k-dimensional unitary matrix $\mathcal{U}$, and where $\mathbf{A} \in \mathbb{C}^{m\times k}$ is any square root ($\mathbf{\Sigma} = \mathbf{A}\mathbf{A}^H$) of the scatter matrix [36], [54]. Such r.v. $\mathbf{x}_s$ is also characterized by the stochastic representation $\mathbf{x}_s =_d \sqrt{\mathcal{Q}_{c,k}}\mathbf{u}_{c,k}$ where the non-negative r.v. $\mathcal{Q}_{c,k}$, and $\mathbf{u}_{c,k}$ are independent, with $\mathbf{u}_{c,k} \sim U(\mathbb{C}S^k)$. $\mathbf{x}_s$ is also characterized by a characteristic function of the form $\Phi_{x_s}(\mathbf{t}) = \phi_c(\|\mathbf{t}\|^2)$, $\mathbf{t} \in \mathbb{C}^k$. Consequently, from (2), we find the characteristic function (25) which defines the C-CES distribution since $\Phi_x(\mathbf{t}) \stackrel{\mathrm{def}}{=} \mathrm{E}[\exp(i\mathrm{Re}(\mathbf{t}^H\mathbf{x}))] = \exp(i\mathrm{Re}(\mathbf{t}^H\boldsymbol{\mu}))\Phi_{x_s}(\mathbf{A}^H\mathbf{t}) = \exp(i\mathrm{Re}(\mathbf{t}^H\boldsymbol{\mu}))\phi_c(\mathbf{t}^H\mathbf{\Sigma}\mathbf{t})$.

3.3 The absolutely continuous case

The p.d.f. of $\mathbf{x}$ is defined on $\mathbb{C}^m$ i.f.f. $\bar{\mathbf{x}}$ is absolutely continuous w.r.t. Lebesgue measure on $\mathbb{R}^{2m}$. Assuming that $\mathrm{rank}(\widetilde{\mathbf{\Sigma}}) = 2m$ and using the identities $(\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}})^T\bar{\mathbf{\Sigma}}^{-1}(\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}}) = (\widetilde{\mathbf{x}} - \widetilde{\boldsymbol{\mu}})^H\widetilde{\mathbf{\Sigma}}^{-1}(\widetilde{\mathbf{x}} - \widetilde{\boldsymbol{\mu}})$ and $|\bar{\mathbf{\Sigma}}| = 2^{-2m}|\widetilde{\mathbf{\Sigma}}|$, the p.d.f. (9) becomes

$$p(\mathbf{x}) = |\widetilde{\mathbf{\Sigma}}|^{-1/2} g_c\left[\frac{1}{2}(\widetilde{\mathbf{x}} - \widetilde{\boldsymbol{\mu}})^H\widetilde{\mathbf{\Sigma}}^{-1}(\widetilde{\mathbf{x}} - \widetilde{\boldsymbol{\mu}})\right], \tag{28}$$

where $g_c(t)$ is defined by

$$g_c(t) \stackrel{\mathrm{def}}{=} 2^m g_r(2t), \tag{29}$$

and $g_r(t)$ is the density generator associated with the distribution $\mathrm{RES}_{2m}(\bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\Sigma}}, \phi_r)$, which satisfies

$$\delta_{c,m} \stackrel{\text{def}}{=} \int_0^\infty t^{m-1} g_c(t) dt = \delta_{r,2m,g} < \infty. \tag{30}$$

We note that in this case, (26) is written equivalently in the form $\widetilde{\mathbf{x}} =_d \widetilde{\boldsymbol{\mu}} + \widetilde{\boldsymbol{\Sigma}}^{1/2} \widetilde{\mathbf{u}}_{c,m}$ (where $\widetilde{\mathbf{u}}_{c,m} \stackrel{\text{def}}{=} (\mathbf{u}_{c,m}^T, \mathbf{u}_{c,m}^H)^T$), then it follows that

$$\mathcal{Q}_{c,m} = \frac{1}{2} \mathcal{Q}_{r,2m} \tag{31}$$

and

$$\mathcal{Q}_{c,m} =_d \frac{1}{2} (\widetilde{\mathbf{x}} - \widetilde{\boldsymbol{\mu}})^H \widetilde{\boldsymbol{\Sigma}}^{-1} (\widetilde{\mathbf{x}} - \widetilde{\boldsymbol{\mu}}). \tag{32}$$

For C-CES distributed $\mathbf{x}$, (28) reduces to

$$p(\mathbf{x}) = |\boldsymbol{\Sigma}|^{-1} g_c[(\mathbf{x} - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})], \tag{33}$$

and (27) implies

$$\mathcal{Q}_{c,m} =_d (\mathbf{x} - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}). \tag{34}$$

Following the same derivation as for the RES distribution, we get the following p.d.f. of $\mathcal{Q}_{c,m}$ and $\mathcal{R}_{c,m}$

$$p(q) = \delta_{c,m}^{-1} q^{m-1} g_c(q) \ \text{ and } \ p(r) = 2\delta_{c,m}^{-1} r^{2m-1} g_c(r^2). \tag{35}$$

3.4 The subclass of compound Gaussian distributions

In many engineering applications, only the circular complex case is considered. A r.v. $\mathbf{x} \in \mathbb{C}^m$ is said to be circular complex compound Gaussian (C-CCG) distributed if the r.v. $\bar{\mathbf{x}} \in \mathbb{R}^{2m}$ is RCG distributed (see definition 3) where the associated extended scatter matrix (22) is bloc-diagonal $\widetilde{\boldsymbol{\Sigma}} = \begin{pmatrix} \boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}^* \end{pmatrix}$. Consequently from the one to one mapping $\bar{\mathbf{x}} \mapsto \tilde{\mathbf{x}}$, definition (3) gives the following stochastic representation

$$\mathbf{x} =_d \boldsymbol{\mu} + \sqrt{\tau_c} \mathbf{n}, \tag{36}$$

where $\tau_c = 2\tau_r$ with τ_r, independent from $\mathbf{n}$, is associated with the $2m$-dimensional RCG distribution and $\mathbf{n} \sim \mathbb{C}N_m(\mathbf{0}, \boldsymbol{\Sigma})$.

As a consequence, all the properties given in Section 2.4 can be deduced from this real to complex representation. In particular with $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^H$ of rank k: $\mathbf{n}_0 \sim \mathbb{C}N_k(\mathbf{0}, \mathbf{I})$, (16) with now $\mathcal{R}_{c,k} \stackrel{\text{def}}{=} \sqrt{\tau_c s}$ where $s \stackrel{\text{def}}{=} \|\mathbf{n}_0\|^2 \sim \frac{1}{2}\chi_{2k}^2 = \mathrm{Gam}(k, 1)$. Then, the C-CCG distributions form a subclass of the C-CES distributions and $\mathbf{x} \sim$ C-CES$_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$ belongs to the set of C-CCG distributions i.f.f. the p.d.f of $\mathcal{Q}_{c,k}$ is

$$p(q) = \int_0^\infty \frac{1}{\Gamma(k)\tau^k} q^{k-1} \exp(-q/\tau) dF_\tau(\tau) \tag{37}$$

and (20) becomes

$$p(\mathbf{x}) = \pi^{-m} |\mathbf{\Sigma}|^{-1} \int_0^\infty \tau^{-m} \exp\left(-\frac{1}{\tau}(\mathbf{x}-\boldsymbol{\mu})^H \mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right) dF_\tau(\tau). \tag{38}$$

4 Basic properties

In this section and throughout the rest of this chapter, we mainly consider the RES distributions, knowing that the C-CES distributions are only a particular representation of them for even m, where the complementary scatter matrix $\mathbf{\Omega}$ defined in (22) is zero. Consequently we drop the indices r in $\delta_{r,k}$, $\delta_{r,k,g}$, g_r, ϕ_r, $\mathcal{Q}_{r,k}$, $\mathcal{R}_{r,k}$ and $\mathbf{u}_{r,k}$ associated with the v.a. $\mathbf{x}$. We will show that these distributions benefit from most of the properties of the Gaussian distribution except the additive stability, whose conditions are more restrictive (see e.g., the quick surveys in [23], [19]).

4.1 Moments

From the full-rank stochastic representation (7), it is clear that $\mathbf{x}$ admits pth-order moments i.f.f. $\mathrm{E}(\mathcal{R}_k^p) < \infty$. Using the characteristic function (1), $\mathrm{E}(\mathcal{R}_k^p) < \infty$ i.f.f. the characteristic generator $\phi(\mathbf{t})$ is p times differentiable. In this case, the pth-order moments of $\mathbf{x}$ are given by $\mathrm{E}(x_1^{p_1} x_2^{p_2} ... x_m^{p_m}) = \frac{1}{i^p} \frac{\partial^p \Psi_x(\mathbf{t})}{\partial t_1^{p_1} \partial t_2^{p_2} ... \partial t_m^{p_m}}|_{\mathbf{t}=\mathbf{0}}$ with $p = \sum_{i=1}^m p_i$ and $\mathbf{x} = (x_1, x_2, ..., x_m)^T$.

Assuming that the correspondent moments are finite, one has

$$\mathrm{E}(\mathbf{x}) = \boldsymbol{\mu} \tag{39}$$

$$\mathrm{Cov}(\mathbf{x}) = \frac{\mathrm{E}(\mathcal{R}_k^2)}{k}\mathbf{\Sigma} = \frac{\mathrm{E}(\mathcal{Q}_k)}{k}\mathbf{\Sigma} = -2\phi'(\mathbf{0})\mathbf{\Sigma}. \tag{40}$$

In particular for RCG distributions (40) becomes

$$\mathrm{Cov}(\mathbf{x}) = \mathrm{E}(\tau)\mathbf{\Sigma}. \tag{41}$$

We see from (40) and (41) that under finite second-order moment assumption, the covariance of $\mathbf{x}$ does not necessarily coincide with the scatter matrix $\mathbf{\Sigma}$, but these two matrices are proportional. Note that many second-order signal processing methodologies, such as, for example, subspace-based processing (where $k = m$), require an estimate of the covariance only at up to a multiplicative scalar. In this case, the *shape matrix*, defined as a scaled version of the scatter matrix $\mathbf{V} \stackrel{\mathrm{def}}{=} \frac{1}{s(\mathbf{\Sigma})}\mathbf{\Sigma}$ can be adopted to characterize the correlation structure. Even if the choice of the

scale functional is entirely arbitrary, in signal processing literature, the most popular scale is the one on the trace of the scatter matrix, i.e., $s(\mathbf{\Sigma}) \stackrel{\text{def}}{=} \text{Tr}(\mathbf{\Sigma})/m$ leading to the following shape matrix $\mathbf{V} \stackrel{\text{def}}{=} \frac{m}{\text{Tr}(\mathbf{\Sigma})}\mathbf{\Sigma}$. One can also find $s(\mathbf{\Sigma}) \stackrel{\text{def}}{=} |\mathbf{\Sigma}|^{1/m}$, leading to unit determinant for the shape matrix.

Otherwise, (40) and (41) can be used to resolve the scale ambiguity of the couple $(\mathbf{\Sigma}, \phi(.))$ in the definition (1) of the RES distribution by fixing the constraint on the characteristic generator $\phi(.)$

$$\text{E}(\mathcal{R}_k^2) = \text{E}(\mathcal{Q}_k) = k = \text{rank}(\mathbf{\Sigma}) \text{ or } \text{E}(\tau) = 1 \text{ for RCG distributions,} \quad (42)$$

which ensures that $\text{Cov}(\mathbf{x}) = \mathbf{\Sigma}$.

When the r.v. $\mathbf{x}$ is absolutely continuous w.r.t. Lebesgue measure on $\mathbb{R}^m$ and $k = m$, (10) implies that (42) is equivalent to the following constraint on the density generator $g(.)$;

$$\delta_{m+2,g}/\delta_{m,g} = m. \quad (43)$$

If $\text{E}(\mathcal{Q}_m)$ is not finite, rather imposing Median$(\mathcal{R}_m) = 1$, (i.e., from (13) the constraint $2\delta_m^{-1}\int_0^1 r^{m-1}g(r^2)dr = \frac{1}{2}$) is a more appropriate scaling constraint as its avoids any finite moment assumptions. Similarly for RCG distributions, if $\text{E}(\tau)$ is not finite, the constraint Median$(\tau) = 1$ (i.e., $F_\tau(1) = \frac{1}{2}$) can be used. Indeed many RES distributions do not have finite second-order moments.

To consider higher-order multivariate central moments, let us consider $\sigma_{i_1,i_2,..,i_\ell} \stackrel{\text{def}}{=} \text{E}[(x_{i_1}-\mu_{i_1})(x_{i_2}-\mu_{i_2})...(x_{i_\ell}-\mu_{i_\ell})]$ with $\boldsymbol{\mu} = (\mu_1, \mu_2, ..., \mu_m)^T$ and $(i_1, i_2, .., i_\ell) \in \{1, .., m\}^\ell$. By symmetry, all odd-order central moments are zero, provided that the corresponding moments do exist. As for fourth-order moments, if $\text{E}(\mathcal{R}_m^4) < \infty$ (with $k = m$), they all satisfy the identity [49, p. 2]

$$\sigma_{i,j,k,\ell} = (\kappa + 1)(\sigma_{i,j}\sigma_{k,\ell} + \sigma_{i,k}\sigma_{j,\ell} + \sigma_{i,\ell}\sigma_{j,k}), \quad (i, j, k, \ell) \in \{1, ..., m\}^4 \quad (44)$$

where

$$\kappa \stackrel{\text{def}}{=} \frac{1}{3}\left(\frac{\sigma_{i,i,i,i}}{\sigma_{i,i}^2} - 3\right) = \frac{m}{m+2}\frac{\text{E}(\mathcal{R}_m^4)}{(\text{E}(\mathcal{R}_m^2))^2} - 1 = \frac{\phi"(0)}{(\phi'(0))^2} - 1. \quad (45)$$

κ is the *kurtosis* parameter of the marginal r.v. x_i [3, 6]. It usually depends on the dimension m, but, remarkably, the kurtosis of the ith component does not depend on i, nor on the scatter matrix $\mathbf{\Sigma}$. Consequently taken $\mathbf{\Sigma} = \mathbf{I}$, we get for all RCG distributions: $\sigma_{i,i,i,i} = 3\text{E}(\tau^2)$ and $\sigma_{i,i} = \text{E}(\tau)$ and thus

$$\kappa = \frac{\text{Var}(\tau)}{[\text{E}(\tau)]^2}. \quad (46)$$

Consequently the kurtosis parameter does not depend on the dimension m for RCG distributions. Note that $\kappa = 0$ for the Gaussian distribution and that there exist other definitions of the kurtosis parameter or coefficient in the literature as $\kappa \stackrel{\text{def}}{=}$

$\frac{\sigma_{i,i,i,i}}{\sigma_{i,i}^2} - 3$ and $\kappa \stackrel{\text{def}}{=} \frac{\sigma_{i,i,i,i}}{\sigma_{i,i}^2}$. Note also that the kurtosis is bounded below such that $\kappa \geq -2/(m+2)$ [7].

4.2 Affine transformations and marginal distributions

Let $\mathbf{b} \in \mathbb{R}^n$ and $\mathbf{B} \in \mathbb{R}^{n \times m}$, and consider the transformed r.v. $\mathbf{y} = \mathbf{B}\mathbf{x} + \mathbf{b}$. Its characteristic function $\Phi_y(\mathbf{t})$ is deduced from the characteristic function (1) of $\mathbf{x}$ by

$$\Phi_y(\mathbf{t}) = \exp(i\mathbf{t}^T\mathbf{b})\Phi_x(\mathbf{B}^T\mathbf{t}) = \exp(i\mathbf{t}^T(\mathbf{B}\boldsymbol{\mu} + \mathbf{b}))\phi(\mathbf{t}^T\mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^T\mathbf{t}), \quad \mathbf{t} \in \mathbb{R}^n. \tag{47}$$

Consequently, $\mathbf{y}$ is $\mathrm{RES}_n(\mathbf{B}\boldsymbol{\mu} + \mathbf{b}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^T, \phi)$-distributed, and thus the class of $\mathrm{RES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$ distributions is closed under affine transformations. Note that the parameters are transformed as $(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \mapsto (\mathbf{B}\boldsymbol{\mu} + \mathbf{b}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^T)$, which is the usual transformation for the couple (expectation, covariance) (when it exists) for the affine transformation $\mathbf{x} \mapsto \mathbf{B}\mathbf{x} + \mathbf{b}$. Note also that for $n \neq m$, $\mathbf{y}$ may not belong to the same family as that of the r.v. $\mathbf{x}$ because the characteristic generator ϕ may depend on the dimension m.

From the full-rank stochastic representation of $\mathbf{x}$ (7), we derive

$$\mathbf{y} =_d \mathbf{B}\boldsymbol{\mu} + \mathbf{b} + \mathcal{R}_k(\mathbf{B}\mathbf{A})\mathbf{u}_k, \tag{48}$$

which is not necessarily a full-rank stochastic representation of $\mathbf{y}$ because $\mathrm{rank}(\mathbf{B}\mathbf{A}) \leq \min(\mathrm{rank}(\mathbf{B}), \mathrm{rank}(\mathbf{A})) \leq \min(n, k)$. If $\ell \stackrel{\text{def}}{=} \mathrm{rank}(\mathbf{B}\mathbf{A})$, there exist full column rank $n \times \ell$ matrices $\mathbf{C}$ such that $\mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^T = \mathbf{C}\mathbf{C}^T$. And thus, similarly to $\mathbf{x}$, where its distribution is equivalently defined by its stochastic full rank representation (7) and by its characteristic function (1), we get from (47) a stochastic full rank representation of $\mathbf{y}$

$$\mathbf{y} =_d \mathbf{B}\boldsymbol{\mu} + \mathbf{b} + \mathcal{R}_\ell\mathbf{C}\mathbf{u}_\ell, \tag{49}$$

where the non-negative r.v. $\mathcal{R}_\ell$, and $\mathbf{u}_\ell$ are independent, $\mathbf{u}_\ell$ is uniformly distributed on the unit real ℓ-sphere.

Of course, it directly follows that univariate and multivariate marginals of $\mathbf{x}$ also are RES distributed with ϕ remains unchanged. If $\mathbf{x} = (x_1, ..., x_m)^T = (\mathbf{x}_1^T, \mathbf{x}_2^T)^T$, $\boldsymbol{\mu} = (\mu_1, ..., \mu_m)^T = (\boldsymbol{\mu}_1^T, \boldsymbol{\mu}_2^T)^T$ and $\boldsymbol{\Sigma} = (\sigma_{i,j})_{i,j=1}^m = \begin{pmatrix} \boldsymbol{\Sigma}_{1,1} & \boldsymbol{\Sigma}_{1,2} \\ \boldsymbol{\Sigma}_{2,1} & \boldsymbol{\Sigma}_{2,2} \end{pmatrix}$, where $\mathbf{x}_1$ and $\boldsymbol{\mu}_1$ are m_1-dimensional vectors, $\boldsymbol{\Sigma}_{1,1}$ and $\boldsymbol{\Sigma}_{2,2}$ are $m_1 \times m_1$ and $m_2 \times m_2$ matrices, respectively (with $m_1 + m_2 = m$), then $\mathbf{x}_1 \sim \mathrm{RES}_{m_1}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{1,1}, \phi)$, $\mathbf{x}_2 \sim \mathrm{RES}_{m_2}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{2,2}, \phi)$, and $x_i \sim \mathrm{RES}_1(\mu_i, \sigma_{i,i}, \phi)$, $i = 1, ..., m$. Of course, their stochastic representations (48) and (49) also, follow. For example for $m_1 \geq k$, if $\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix}$ where $\mathbf{A}_1$ is a $m_1 \times k$ full column rank matrix

$$\mathbf{x}_1 =_d \boldsymbol{\mu}_1 + \mathcal{R}_k\mathbf{A}_1\mathbf{u}_k \tag{50}$$

is a stochastic full-rank representation of the marginal $\mathbf{x}_1$. As for the univariate marginal x_i, the stochastic representations (48) and (49) reduce to

$$x_i =_d \mu_i + \mathcal{R}_k \mathbf{a}_i^T \mathbf{u}_k =_d \mu_i + \mathcal{R}_1 \|\mathbf{a}_i\| u_1, \tag{51}$$

where $\mathbf{a}_i$ is the i-th column of $\mathbf{A}^T$ (thus $\|\mathbf{a}_i\|^2 = \sigma_{i,i}$), $\mathcal{R}_k$ and $\mathcal{R}_1$ are the modular variates of $\mathbf{x}$ and x_i, respectively, and where u_1 reduces to the uniform discrete r.v. $\{-1, +1\}$.

Now we take a closer look at the marginal distributions when $k = m$. In this case, the stochastic full-rank representation (50) and (51) of arbitrary univariate or multivariate marginal r.v. $\mathbf{x}_1$ and x_i reduce to

$$\mathbf{x}_1 =_d \boldsymbol{\mu}_1 + \mathcal{R}_{m_1} \boldsymbol{\Sigma}_{1,1}^{1/2} \mathbf{u}_{m_1} \text{ and } x_i =_d \mu_i + \mathcal{R}_1 \sigma_{i,i}^{1/2} u_1, \tag{52}$$

where the modular variates $\mathcal{R}_{m_1}$ of $\mathbf{x}_1$ (which includes the modular variates of the univariate marginal r.v. x_i for $m_1 = 1$) and $\mathcal{R}_m$ of $\mathbf{x}$ are related by the relation [17, Corollary p.59]

$$\mathcal{R}_{m_1} =_d \mathcal{R}_m \, \beta_{\frac{m_1}{2}, \frac{m_2}{2}}. \tag{53}$$

In (53) the r.v. $\mathcal{R}_m$ and $\beta_{\frac{m_1}{2}, \frac{m_2}{2}}$ are independent and $\beta^2_{\frac{m_1}{2}, \frac{m_2}{2}} \sim \mathrm{Beta}(\frac{m_1}{2}, \frac{m_2}{2})$.

Moreover in the absolutely continuous case w.r.t. Lebesgue measure on $\mathbb{R}^m$, i.e., $\mathbf{x} \sim \mathrm{RES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$, (53) allows us to relate the p.d.f. $p_{m_1}(r)$ of $\mathcal{R}_{m_1}$ to the p.d.f. $p_m(r)$ of $\mathcal{R}_m$ [17, rel. 2.5.15]

$$p_{m_1}(r) = \frac{2r^{m_1-1}}{\mathrm{B}(\frac{m_1}{2}, \frac{m_2}{2})} \int_r^{+\infty} t^{-(m-2)} (t^2 - r^2)^{\frac{m_2}{2}-1} p_m(t) dt \tag{54}$$

and to the density generator $g(.)$ of $\mathbf{x}$ using $p_m(r)$ given by (13)

$$p_{m_1}(r) = \frac{2\pi^{m/2} r^{m_1-1}}{\Gamma(\frac{m_1}{2})\Gamma(\frac{m_2}{2})} \int_{r^2}^{+\infty} (t - r^2)^{\frac{m_2}{2}-1} g(t) dt. \tag{55}$$

This allows us to derive the density generators $g_{m_1|m}(.)$ of the multivariate and univariate marginal r.v. $\mathbf{x}_1$ and x_i thanks to (13) $p_{m_1}(r) = 2\delta_{m_1}^{-1} r^{m_1-1} g_{m_1|m}(r^2)$

$$g_{m_1|m}(u) = \delta_{m_2}^{-1} \int_u^{+\infty} (t - u)^{\frac{m_2}{2}-1} g(t) dt. \tag{56}$$

Therefore the p.d.f. of the multivariate marginal r.v. $\mathbf{x}_1 \sim \mathrm{RES}_{m_1}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{1,1}, g_{m_1|m})$ and $x_i \sim \mathrm{RES}_1(\mu_i, \sigma_{i,i}, g_{1|m})$ are given respectively by

$$p_{m_1|m}(\mathbf{x}_1) = |\boldsymbol{\Sigma}_{1,1}|^{-1/2} g_{m_1|m}[(\mathbf{x}_1 - \boldsymbol{\mu}_1)^T \boldsymbol{\Sigma}_{1,1}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1)]. \tag{57}$$

$$p_{i|m}(x_i) = \frac{1}{\sqrt{\sigma_{i,i}}} g_{1|m} \left(\frac{(x_i - \mu_i)^2}{\sigma_{i,i}} \right). \tag{58}$$

Note that there is no guarantee that the integral in (56) leads to a closed-form expression when a closed-form expression of $g(u)$ is available, except for the class of RCG distributions for which $g_{m_1|m}(u) \propto g(u)$. In particular for the Gaussian distribution, we obtain $g_{m_1|m}(u) = \frac{1}{(2\pi)^{m_1/2}} \exp(-\frac{u}{2})$.

For the RCG distributions defined in Subsection 2.4, it is clear from the stochastic representation (14) that

$$\mathbf{x}_1 =_d \boldsymbol{\mu}_1 + \sqrt{\tau}\mathbf{n}_1, \tag{59}$$

with $\mathbf{n}_1 \sim \mathbb{R}N_{m_1}(\mathbf{0}, \boldsymbol{\Sigma}_{1,1})$. Consequently, all the marginals belong to the same subclass of RCG distributions with the same c.d.f $F_\tau(.)$ and the density generator (56) reduces thanks to (21) to

$$g_{m_1|m}(t) = (2\pi)^{-m_1/2} \int_0^\infty \tau^{-m_1/2} \exp\left(-\frac{t}{2\tau}\right) dF_\tau(\tau). \tag{60}$$

In fact this property characterizes the RCG distributions. This point is specified by a consistency result [31, Th. 1]. This result states that any univariate and multivariate marginal distribution of an r.v. $\mathbf{x}$ belong to the same family as that of the r.v. $\mathbf{x}$, i.f.f. the RES distribution belongs to the class of RCG distributions. This condition is also equivalent to the characteristic generator ϕ not related to the dimension m. This consistency result shows that not all elliptically symmetric distributions can be used to define random processes. Indeed from Kolmogorov's theorem, only among the elliptically symmetric distributions, the m dimensional RGG distributed r.v. $\mathbf{x} = (x_1, ..., x_m)^T$ for all m can define a unique random process $(x_n)_{n\in\mathbb{N}}$ that is characterize by the distribution of τ, symmetric center and scatter of the process.

4.3 Conditional distributions

Let $\mathbf{x} = (\mathbf{x}_1^T, \mathbf{x}_2^T)^T \sim \mathrm{RES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$ where $\mathbf{x}_1 \in \mathbb{R}^{m_1}$. The conditional distribution of $\mathbf{x}_2$ given $\mathbf{x}_1$ is generally more difficult to describe than its marginal distribution presented in Section 4.2. For the sake of simplicity, we consider here only the full rank case ($k = m$); see e.g. [9] and [19, Th. 7, Cor. 8] for more general statements. In this case, it is proved [16, Th 2.18], that $\mathbf{x}_2$ given $\mathbf{x}_1 = \mathbf{x}_1^0$ (denoted $\mathbf{x}_2|\mathbf{x}_1 = \mathbf{x}_1^0$) is $\mathrm{RES}_{m_2}(\boldsymbol{\mu}_{2|1}, \boldsymbol{\Sigma}_{2|1}, \phi_{2|1}) = \mathrm{RES}_{m_2}(\boldsymbol{\mu}_{2|1}, \boldsymbol{\Sigma}_{2|1}, F_{2|1})$ distributed with

$$\boldsymbol{\mu}_{2|1} = \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{2,1}\boldsymbol{\Sigma}_{1,1}^{-1}(\mathbf{x}_1^0 - \boldsymbol{\mu}_1) \tag{61}$$

$$\boldsymbol{\Sigma}_{2|1} = \boldsymbol{\Sigma}_{2,2} - \boldsymbol{\Sigma}_{2,1}\boldsymbol{\Sigma}_{1,1}^{-1}\boldsymbol{\Sigma}_{1,2}, \tag{62}$$

and where $\phi_{2|1}$ and $F_{2|1}$ correspond respectively to the characteristic generator of $\mathbf{x}_2|\mathbf{x}_1 = \mathbf{x}_1^0$ and the c.d.f. of the conditional modular variate $\mathcal{R}_{2|1}$ defined by the following stochastic full rank representation

$$(\mathbf{x}_2|\mathbf{x}_1 = \mathbf{x}_1^0) =_d \boldsymbol{\mu}_{2|1} + \mathcal{R}_{2|1}\mathbf{A}_{2|1}\mathbf{u}_{m_2}, \tag{63}$$

where $\mathbf{A}_{2|1}$ is an arbitrary square root of $\mathbf{\Sigma}_{2|1}$ (i.e., $\mathbf{\Sigma}_{2|1} = \mathbf{A}_{2|1}\mathbf{A}_{2|1}^T$ and $\mathcal{R}_{2|1}$ and $\mathbf{u}_{m_2}$ are independent. $\mathcal{R}_{2|1}$ is given by

$$\mathcal{R}_{2|1} =_d [\mathcal{R}_m^2 - (\mathbf{x}_1^0 - \boldsymbol{\mu}_1)^T \mathbf{\Sigma}_{1,1}^{-1}(\mathbf{x}_1^0 - \boldsymbol{\mu}_1)]^{1/2} | \mathbf{x}_1 = \mathbf{x}_1^0. \tag{64}$$

Otherwise if the conditional covariance $\mathrm{Cov}(\mathbf{x}_2|\mathbf{x}_1 = \mathbf{x}_1^0)$ exists, its expressions can be derived from (40) and we get

$$\mathrm{Cov}(\mathbf{x}_2|\mathbf{x}_1 = \mathbf{x}_1^0) = \frac{1}{m_2}\left(\mathrm{E}[\mathcal{R}_m^2|\mathbf{x}_1 = \mathbf{x}_1^0] - (\mathbf{x}_1^0 - \boldsymbol{\mu}_1)^T \mathbf{\Sigma}_{1,1}^{-1}(\mathbf{x}_1^0 - \boldsymbol{\mu}_1)\right)\mathbf{\Sigma}_{2|1}. \tag{65}$$

We note that the expressions of the conditional symmetry center $\boldsymbol{\mu}_{2|1}$ and conditional scatter matrix $\mathbf{\Sigma}_{2|1}$ are those obtained for the Gaussian distribution, but the conditional characteristic generator $\phi_{2|1}$ no longer belongs to the same family of RES, except for the Gaussian distribution.

For RCG distributions, the p.d.f. of $\mathbf{x}_2|\mathbf{x}_1 = \mathbf{x}_1^0$, (denoted $p(\mathbf{x}_2/\mathbf{x}_1^0)$, is given by

$$p(\mathbf{x}_2/\mathbf{x}_1^0) = \frac{1}{p(\mathbf{x}_1^0)}\int_0^\infty p(\mathbf{x}_2/\mathbf{x}_1^0, \tau)p(\mathbf{x}_1^0/\tau)dF_\tau(\tau) \tag{66}$$

where $p(\mathbf{x}_1^0/\tau) = (2\pi\tau)^{-m_1/2}|\mathbf{\Sigma}_{1,1}|^{-1/2}\exp(-\frac{1}{2\tau}d^2_{\mathbf{\Sigma}_{1,1}}(\mathbf{x}_1^0, \boldsymbol{\mu}_1))$ and $p(\mathbf{x}_2/\mathbf{x}_1^0, \tau)$ $= (2\pi\tau)^{-m_2/2}|\mathbf{\Sigma}_{2|1}|^{-1/2}\exp(-\frac{1}{2\tau}d^2_{\mathbf{\Sigma}_{2|1}}(\mathbf{x}_2, \boldsymbol{\mu}_{2|1}))$ with $d^2_{\mathbf{\Sigma}_{1,1}}(\mathbf{x}_1^0, \boldsymbol{\mu}_1) \stackrel{\mathrm{def}}{=} (\mathbf{x}_1^0 - \boldsymbol{\mu}_1)^T\mathbf{\Sigma}_{1,1}^{-1}(\mathbf{x}_1^0 - \boldsymbol{\mu}_1)$ and $d^2_{\mathbf{\Sigma}_{2|1}}(\mathbf{x}_2, \boldsymbol{\mu}_{2|1}) \stackrel{\mathrm{def}}{=} (\mathbf{x}_2 - \boldsymbol{\mu}_2)^T\mathbf{\Sigma}_{2|1}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$. This yields

$$p(\mathbf{x}_2/\mathbf{x}_1^0) = \frac{(2\pi)^{-m_2/2}|\mathbf{\Sigma}_{2|1}|^{-1/2}}{\int_0^\infty \tau^{-m_1/2}\exp(-\frac{1}{2\tau}d^2_{\mathbf{\Sigma}_{1,1}}(\mathbf{x}_1^0, \boldsymbol{\mu}_1))dF_\tau(\tau)}$$
$$\times \int_0^\infty \tau^{-m/2}\exp\left(-\frac{1}{2\tau}d^2_{\mathbf{\Sigma}_{2|1}}(\mathbf{x}_2, \boldsymbol{\mu}_{2|1}\right)\exp\left(-\frac{1}{2\tau}d^2_{\mathbf{\Sigma}_{1,1}}(\mathbf{x}_1^0, \boldsymbol{\mu}_1)\right)dF_\tau(\tau). \tag{67}$$

Comparing (67) to (9), we check that $\mathbf{x}_2|\mathbf{x}_1 = \mathbf{x}_1^0$ is $\mathrm{RES}_{m_2}(\boldsymbol{\mu}_{2|1}, \mathbf{\Sigma}_{2|1}, F_{2|1})$ distributed. Moreover, comparing (67) to (20), we see that $\mathbf{x}_2|\mathbf{x}_1 = \mathbf{x}_1^0$ is RCG distributed, but with a differently distributed texture τ than $\mathbf{x}$ and the marginal $\mathbf{x}_1$.

4.4 Summation stability

Consider now the sum $\mathbf{y}$ of n independent r.v. $\mathbf{x}_1, .., \mathbf{x}_i, .., \mathbf{x}_n$ with the same scatter matrix, where $\mathbf{x}_i \sim \mathrm{RES}_m(\boldsymbol{\mu}_i, \mathbf{\Sigma}, \phi_i)$. The characteristic function $\Phi_y(\mathbf{t})$ of $\mathbf{y} = \sum_{i=1}^n \mathbf{x}_i$ is

$$\Phi_y(\mathbf{t}) = \prod_{i=1}^n \Phi_{x_i}(\mathbf{t}) = \exp\left(i\mathbf{t}^T(\sum_{i=1}^n \boldsymbol{\mu}_i)\right)\prod_{i=1}^n \phi_i(\mathbf{t}^T\mathbf{\Sigma}\mathbf{t}) = \exp(i\mathbf{t}^T\boldsymbol{\mu})\phi(\mathbf{t}^T\mathbf{\Sigma}\mathbf{t}), \tag{68}$$

where $\boldsymbol{\mu} \stackrel{\text{def}}{=} \sum_{i=1}^{n} \boldsymbol{\mu}_i$ and $\phi(u) \stackrel{\text{def}}{=} \prod_{i=1}^{n} \phi_i(u)$, which is structured as (1). Consequently, the sum $\mathbf{y}$ is RES-distributed too [19]. Similarly, for independent univariate r.v. $x_1, ..., x_i, ..., x_n$ but with arbitrary scatters σ_i, where $x_i \sim \mathrm{RES}_1(\mu_i, \sigma_i, \phi_i)$, the characteristic function $\Phi_y(t)$ of $y = \sum_{i=1}^{n} x_i$ is given by $\Phi_y(t) = \exp(it\mu)\phi(t^2)$ where $\mu \stackrel{\text{def}}{=} \sum_{i=1}^{n} \mu_i$ and $\phi(u) \stackrel{\text{def}}{=} \prod_{i=1}^{n} \phi_i(\sigma_i u)$ and then $y \sim \mathrm{RES}_1(\mu, 1, \phi)$. The sum y is then symmetrically distributed w.r.t. μ.

However, it is worth underlying that if the r.v. $\mathbf{x}_1, .., \mathbf{x}_i, .., \mathbf{x}_n$ belong to the same family of RES distribution (i.e., with $\phi_i = \phi$, $i = 1, .., n$), the sum is not of the same family except for the so-called elliptical α-stable distributions [44], (i.e., of characteristic functions $\Phi_x(\mathbf{t}) = \exp(i\mathbf{t}^T\boldsymbol{\mu}) \exp(-\frac{1}{2}(\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t})^{\alpha/2})$ with $\alpha \in (0, 2]$), which includes the Gaussian distribution for $\alpha = 2$. Finally note that if the scatter matrices $\boldsymbol{\Sigma}_i$ of $\mathbf{x}_1, .., \mathbf{x}_i, .., \mathbf{x}_n$ are not identical, the summation stability is generally lost for independent multivariate non-Gaussian, RES distributed r.v..

The independence condition of the v.a. $\mathbf{x}_i$ can be relaxed by the following properties proved in [29, th. 4.2]. If the full-rank stochastic representations (7) $\mathbf{x}_i =_d \boldsymbol{\mu}_i + \mathcal{R}_k^i \mathbf{A}\mathbf{u}_k^i$, $i = 1, 2$ with $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^T$ satisfy the condition $(\mathcal{R}_k^1, \mathcal{R}_k^2)$, $\mathbf{u}_k^1, \mathbf{u}_k^2$ are mutually independent, whereas $\mathcal{R}_k^1, \mathcal{R}_k^2$ may be dependent on each other, the sum $\mathbf{y}$ is also $\mathrm{RES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$-distributed, where the expression of ϕ is given in [29, th. 4.2]. ϕ reduces to the product $\phi_1\phi_2$ when $\mathcal{R}_k^1$ and $\mathcal{R}_k^2$ are independent. A natural application of this property is in the context of a multivariate time series.

5 Example of elliptically symmetric distributions

In this section, we present some examples of elliptically symmetric distributions and discuss their main specific properties. Throughout this section, we mainly consider the case of distributions absolutely continuous w.r.t. Lebesgue measure on $\mathbb{R}^m$ with rank$(\boldsymbol{\Sigma}) = m$. Each distribution is defined by its density generator under a functional form parameterized by m and a finite-dimensional parameter. We mainly use their real-valued definition through the p.d.f. (9) where the characteristic and density generators, and the texture (14) are simply denoted here by respectively ϕ, g and τ, knowing that C and NC complex-valued definitions (28) and (33) are simply deduced from the real-valued definition with double dimension (24), (29). For C-CES distributions, interested readers can consult [48].

5.1 Gaussian distribution

The Gaussian distribution is the best-known and widely used distribution among the RES class in classical signal processing applications. Its ubiquity is mainly due to the central limit theorem (CLT) that allows for the use of the Gaussian distribution as a good and handy approximation of the statistical behavior of a set of observations in many practical scenarios.

As a particular case of RES distribution, the Gaussian distribution denoted hereafter by $\mathbb{R}N_m(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, is characterized by a *characteristic* and a *density* generators that can be expressed respectively as:

$$\phi(u) = \exp\left(-\frac{u}{2}\right) \text{ and } g(t) = \frac{1}{(2\pi)^{m/2}} \exp\left(-\frac{t}{2}\right). \tag{69}$$

Using (13), when $\boldsymbol{\Sigma}$ is non-singular, it is immediately verified that the p.d.f. of the second-order modular variate $\mathcal{Q}$ of a Gaussian r.v. is given by:

$$p(q) = \frac{\delta_m^{-1}}{(2\pi)^{m/2}} q^{m/2-1} \exp\left(-\frac{q}{2}\right) = \frac{1}{2^{m/2}\Gamma(m/2)} q^{m/2-1} \exp\left(-\frac{q}{2}\right), \ q \in \mathbb{R}^+, \tag{70}$$

where $\delta_m = \frac{\Gamma(m/2)}{\pi^{m/2}}$ from (5). It is immediate to verify that the p.d.f. in (70) is the one of a central χ^2-distribution with m degrees of freedom, i.e., $\mathcal{Q} \sim \chi^2_m$. Note that this result is perfectly in line with the well-known property of the Gaussian r.v. with mean value $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ whose quadratic form $(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})$ is χ^2_m-distributed.

Using again (13), we can express the p.d.f. of the modular variate $\mathcal{R} \sim \sqrt{\chi^2_m}$ as :

$$p(r) = \frac{1}{2^{m/2-1}\Gamma(m/2)} r^{m-1} \exp\left(-\frac{r^2}{2}\right), \ r \in \mathbb{R}^+. \tag{71}$$

Note that from the real to complex representation given in Section 3, the density generator of complex Gaussian distributions becomes from (29)

$$g(t) = \frac{1}{\pi^m} \exp(-t) \tag{72}$$

and consequently, the p.d.f. of the circular and non-circular Gaussian distribution, denoted respectively by $\mathbb{C}N_m(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\mathbb{C}N_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Omega})$ are given from (33) and (28) by respectively

$$\begin{aligned} p(\mathbf{x}) &= \pi^{-m}|\boldsymbol{\Sigma}|^{-1} \exp[(\mathbf{x}-\boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})], \\ p(\mathbf{x}) &= \pi^{-m}|\widetilde{\boldsymbol{\Sigma}}|^{-1/2} \exp\left[\frac{1}{2}(\widetilde{\mathbf{x}}-\widetilde{\boldsymbol{\mu}})^H \widetilde{\boldsymbol{\Sigma}}^{-1}(\widetilde{\mathbf{x}}-\widetilde{\boldsymbol{\mu}})\right], \end{aligned} \tag{73}$$

where $\boldsymbol{\Sigma}$ and $\boldsymbol{\Omega}$ are the covariance and complementary covariance (or pseudo covariance) of $\mathbf{x}$, respectively.

The Gaussian distribution is usually used as a reference to define heavier-tailed and lighter-tailed distributions in the RES class. Some examples will be given below.

5.2 Student's t-distribution

A popular example of heavy-tailed distribution is the Student t-distribution (or simply t-distribution). The t-distribution belongs to the RES class, and it is characterized by the following density generator [37]:

$$g(u) = \frac{\Gamma(\frac{\nu+m}{2})}{(\nu\pi)^{m/2}\Gamma(\nu/2)} \left(1 + \frac{u}{\nu}\right)^{-\frac{\nu+m}{2}}, \; u \in \mathbb{R}^+ \tag{74}$$

with $0 < \nu < \infty$ *degrees of freedom*. The parameter ν controls the tails of the distribution that are uniformly heavier than the Gaussian ones. In particular, for small values of ν, the sampled r.v. are highly non-Gaussian while, as $\nu \to \infty$, the t-distribution collapses into the Gaussian one. The case $\nu = 1$ is called the Cauchy distribution.

As shown in [37], the second-order modular variate of an m-dimensional t-distributed r.v. satisfies the following property:

$$\mathcal{Q}/m \sim F_{m,\nu}, \tag{75}$$

where $F_{m,\nu}$ denotes the F-distribution with m and ν degrees of freedom [30]. Moreover, using (13), the p.d.f. of the second-order modular variate $\mathcal{Q}$ and of the modular variate $\mathcal{R}$ can be directly derived.

It is worth mentioning that the t-distribution belongs to the subclass of the CG distribution described in Section 3.4. Specifically, t-distributed r.v. can be obtained by generating the texture τ such that its inverse is distributed as a Gamma random variable,

$$\tau^{-1} \sim \mathrm{Gam}(\nu/2, 2/\nu), \tag{76}$$

Since the t-distribution belongs to the subclass of the CG distributions, it holds that the marginals belong to the same family. Hence in particular the univariate marginals x_i are t-distributed with parameter ν, symmetry center $(\boldsymbol{\mu})_i$ and scatter $(\boldsymbol{\Sigma})_{i,i}$ with density generator given by (74) where $m = 1$.

This distribution has first, second and fourth-order moments if, respectively, $\nu > 1$, $\nu > 2$, and $\nu > 4$. In particular, we have $\mathrm{E}(\tau) = \frac{\nu}{2(\nu-2)}$, $\mathrm{E}(\mathcal{Q}) = \frac{m\nu}{\nu-2}$ and $\mathrm{Cov}(\mathbf{x}) = \frac{\nu}{\nu-2}\boldsymbol{\Sigma}$ for $\nu > 2$, while the kurtosis, defined in subsection 4.1 in (45) is $\kappa = \frac{2}{\nu-4}$ for $\nu > 4$.

5.3 Generalized Gaussian distributions

The Generalized Gaussian (GG) distribution is a RES distribution with the remarkable ability to characterize both heavy-tailed and light-tailed (with respect to the Gaussian one) data behavior. The multivariate GG distribution was introduced in [22] as a multivariate generalization of the power exponential distribution. As a member of the RES class, using the parametrization [48, rel. (27)] and (29), the GG

distribution has a density generator given by:

$$g(u) = \frac{s\Gamma(m/2)}{(2\pi)^{m/2} b^{m/2s} \Gamma(m/2s)} \exp\left(-\frac{u^s}{2^s b}\right), \; u \in \mathbb{R}^+ \tag{77}$$

where $s > 0$ and $b > 0$ are two parameters generally called *exponent* (or *shape*) and *scale*, respectively. The scale parameter b is used to ensure that $p(\mathbf{x})$ integrates to 1 in (9), while the exponent parameter s controls the non-Gaussianity. In particular, for $0 < s < 1$, the tails of the GG distribution are heavier with respect to the Gaussian one, while for $s > 1$ they are lighter. For s tending to ∞, this distribution converges to a uniform distribution in an ellipsoid centered on $\boldsymbol{\mu}$. Clearly, for $s = 1$ we get the Gaussian distribution. Otherwise for $s = 1/2$, we get the Laplace distribution (also called double exponential distribution). Note that (77) is consistent with [35] and [50] with different definitions of the scale.

It was proved in [25] that for $s \in (0, 1]$, the GG distributions belong to the subset of CG distributions whose p.d.f. of the mixing variable $\sqrt{\tau}$ is given by [25, (2.2)]. On the contrary for $s > 1$, the GG distributions no longer belong to the CG family and the marginals are no longer GG distributed. In particular, in this case, the univariate marginals of a multivariate GG distribution are not power exponential distributed, i.e., with p.d.f. given by [38]

$$p(x) = \frac{1}{\sigma\Gamma(1 + \frac{1}{2s}) 2^{1+\frac{1}{2s}}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^{2s}\right). \tag{78}$$

Regarding the second-order modular variate $\mathcal{Q}$ of an m-dimensional GG-distributed r.v., it is straightforward to deduce from (13) and (77) that:

$$\mathcal{Q}^s \sim \mathrm{Gam}\left(\frac{m}{2s}, 2^s b\right), \tag{79}$$

We can exploit again (13) to get a closed-form expression of the p.d.f. of $\mathcal{Q}$ and $\mathcal{R}$.

We note that the moments exist at all orders and $\mathrm{Cov}(\mathbf{x}) = \boldsymbol{\Sigma}$ if the parameters s and b are related by the relation $b = [\frac{m}{2}\Gamma(\frac{m}{2s})/\Gamma(\frac{m/2+1}{s})]^s$. Note that in this case the kurtosis κ defined in (45) depends on the dimension m, e.g., for $m = 1$ we get $\kappa = \frac{\Gamma(5/2s)\Gamma(1/2s)}{3[\Gamma(3/2s)]^2} - 1$ which gives $\kappa = 0$ for $s = 1$ (Gaussian case).

5.4 The K-distribution

Another example of RES distribution belonging to the CG-subclass is the K-distribution. An m-dimensional r.v. $\mathbf{x}$ is said to be K-distributed if it has the CG-representation (14) characterized by a Gamma-distributed texture,

$$\tau \sim \frac{1}{2}\mathrm{Gam}(\nu, 1/\nu), \tag{80}$$

where $\nu > 0$ is a *shape* parameter. So the p.d.f. of τ is

$$p(\tau) = \frac{2\nu^{\nu}}{\Gamma(\nu)}(2\tau)^{\nu-1}\exp(-2\nu\tau),\ \tau \in \mathbb{R}^{+} \tag{81}$$

with $\mathrm{E}(\tau) = 1/2$.

By using the definition of CG-distribution, it can be shown that the density generator is given by:

$$g(u) = \frac{\nu^{m/2}}{2^{\nu-1}\pi^{m/2}\Gamma(\nu)}(2\nu u)^{(2\nu-m)/4}K_{(2\nu-m)/2}\left(\sqrt{2\nu u}\right),\ u \in \mathbb{R}^{+}, \tag{82}$$

where $K_{\alpha}(\cdot)$ denotes the modified Bessel function of the second kind of order α. The shape $\nu > 0$ is a parameter that controls the tails of the K-distribution: when $\nu \to 0$, the tails become heavier while for $\nu \to \infty$ the K-distribution collapses onto the Gaussian one. Moreover, using (13), the p.d.f. of the second-order modular variate $\mathcal{Q}$ and of the modular variate $\mathcal{R}$ can be directly derived.

Since the K-distribution belongs to the subclass of the CG distributions, it holds that the marginals belong to the same family. Hence, in particular the univariate marginals x_i are K-distributed with parameter ν, symmetry center $(\boldsymbol{\mu})_i$ and scatter $(\boldsymbol{\Sigma})_{i,i}$ with density generator given by (82) where $m = 1$.

To conclude, we note that the moments of all orders of the K-distribution exist and it can be shown that the kurtosis in (45) is given by $\kappa = \frac{1}{\nu}$.

5.5 Related distribution: the angular central Gaussian distribution

To conclude this section, let us introduce a distribution that has a strong link with the RES family, even if it does not belong to this class. An r.v. $\mathbf{x}$ is said angular central Gaussian (ACG) distributed if it admits the stochastic representation:

$$\mathbf{x} =_{d} \frac{\mathbf{n}}{\|\mathbf{n}\|},\ \text{where } \mathbf{n} \sim \mathbb{R}N_m(\mathbf{0}, \boldsymbol{\Sigma}). \tag{83}$$

For non-singular $\boldsymbol{\Sigma}$, the p.d.f. of this distribution is given by [62]

$$p(\mathbf{x}) = \frac{2\pi^{m/2}}{\Gamma(m/2)}\left|\boldsymbol{\Sigma}\right|^{-1/2}\left(\mathbf{x}^{T}\boldsymbol{\Sigma}^{-1}\mathbf{x}\right)^{-m/2},\ \mathbf{x} \in \mathbb{R}S^{m}. \tag{84}$$

Even if this expression may look similar to the p.d.f. of a centered RES-distributed r.v., it is worth underlying a crucial difference: the density of a RES r.v., given in (9), is defined w.r.t. the Lebesgue measure on $\mathbb{R}^m$, while the density of an ACG r.v. in (84) is defined w.r.t. the Lebesgue measure on real unit sphere $\mathbb{R}S^m$. For this reason, the stochastic representation provided in Sec. 2.2 does not hold for ACG r.v.. It can be noted from (83) or (84), that the parameter $\boldsymbol{\Sigma}$ can only be identified up to a multiplicative scalar factor.

Furthermore, it turns out that if $\mathbf{x}$ is arbitrarily centered RES distributed (not necessarily Gaussian), i.e., if $\mathbf{x} \sim \mathrm{RES}_m(\mathbf{0}, \boldsymbol{\Sigma}, \phi)$, we get from its stochastic representation (7) with $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^T$ of rank k and where $\mathbf{u}_k =_d \frac{\mathbf{n}_0}{\|\mathbf{n}_0\|}$ and $\mathbf{n}_0 \sim \mathbb{R}N_k(\mathbf{0}, \mathbf{I})$

$$\frac{\mathbf{x}}{\|\mathbf{x}\|} =_d \frac{\mathbf{A}\mathbf{u}_k}{\|\mathbf{A}\mathbf{u}_k\|} =_d \frac{\mathbf{A}\mathbf{n}_0}{\|\mathbf{A}\mathbf{n}_0\|} =_d \frac{\mathbf{n}}{\|\mathbf{n}\|}, \text{ where } \mathbf{n} \sim \mathbb{R}N_m(\mathbf{0}, \boldsymbol{\Sigma}). \tag{85}$$

Consequently the projection of $\mathbf{x}$ onto the unit real m-sphere is also ACG-distributed. In its definition (83), $\mathbb{R}N_m(\mathbf{0}, \boldsymbol{\Sigma})$ can be replaced by any $\mathrm{RES}_m(\mathbf{0}, \boldsymbol{\Sigma}, \phi)$ distribution and the term ACG appears to be a slight misnomer.

Finally, note that this class of distribution is closed under standardized linear transformations, i.e., $\mathbf{x} \sim \mathrm{ACG}_m(\mathbf{0}, \boldsymbol{\Sigma})$, then $\mathbf{w} \stackrel{\mathrm{def}}{=} \mathbf{B}\mathbf{x}/\|\mathbf{B}\mathbf{x}\| \sim \mathrm{ACG}_k(\mathbf{0}, \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^T)$ for any nonzero $k \times m$ matrix $\mathbf{B}$.

6 Parameter estimation

We are interested in this Section to the estimation of the symmetry center $\boldsymbol{\mu}$ and scatter matrix $\boldsymbol{\Sigma}$ (assumed invertible here) of RES distributions. Suppose we have an i.i.d. sample $\mathbf{x}_1, .., \mathbf{x}_i, .., \mathbf{x}_n$ of size $n > m$ from a $\mathrm{RES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$ distribution absolutely continuous w.r.t. Lebesgue measure on $\mathbb{R}^m$ under finite second-order moments. The identifiability issue is solved here by imposing constraint (42) which ensures that $\mathrm{Cov}(\mathbf{x}_i) = \boldsymbol{\Sigma}$.

6.1 Sample mean and sample covariance matrix

A natural estimate of the parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the sample mean $\widehat{\boldsymbol{\mu}} \stackrel{\mathrm{def}}{=} \frac{1}{n}\sum_{i=1}^n \mathbf{x}_i$ and the sample covariance matrix (SCM) $\widehat{\boldsymbol{\Sigma}} \stackrel{\mathrm{def}}{=} \frac{1}{n-1}\sum_{i=1}^n (\mathbf{x}_i - \widehat{\boldsymbol{\mu}})(\mathbf{x}_i - \widehat{\boldsymbol{\mu}})^T$, respectively. It is well known that $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Sigma}}$ are unbiased and mutually uncorrelated estimators. Under the particular case of Gaussian data, the two estimators are then independent. $\widehat{\boldsymbol{\mu}}$ is RES distributed with symmetry center $\boldsymbol{\mu}$ and scatter matrix $\frac{1}{n}\boldsymbol{\Sigma}$ by imposing constraint (42) on this distribution which does not necessarily belong to the same RES distribution family as $\mathbf{x}_i$ (see Section 4.4). Under finite fourth-order moments, applying the CLT to $\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i^T$ and $\sum_{i=1}^n \mathbf{x}_i$, it can be shown that $\widehat{\boldsymbol{\Sigma}}$ is asymptotically Gaussian distributed, i.e.,

$$\sqrt{n}\left(\mathrm{vec}(\widehat{\boldsymbol{\Sigma}}) - \mathrm{vec}(\boldsymbol{\Sigma})\right) \to_d \mathbb{R}N_{m^2}\left(\mathbf{0}, \mathbf{R}_{\Sigma_{SCM}}\right) \tag{86}$$

with [49, p. 5]

$$\mathbf{R}_{\Sigma_{SCM}} = (1+\kappa)(\mathbf{I}+\mathbf{K})(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma}) + \kappa\mathrm{vec}(\boldsymbol{\Sigma})\mathrm{vec}^T(\boldsymbol{\Sigma}), \tag{87}$$

where κ is the kurtosis parameter (45) of the $\mathrm{RES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$ distribution. Note that this asymptotic distribution would also be the asymptotic distribution of the estimate $\frac{1}{n}\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^T$ of $\boldsymbol{\Sigma}$ when $\boldsymbol{\mu}$ is known. For heavier tails than Gaussian distributions, $\kappa > 0$ and can be very large without any upper-bound (see e.g. in Section 5.2 when κ approaches 4 for Student's t-distribution) and consequently, the SCM estimate can be a very bad estimator.

6.2 ML estimation

To take into account the particular RES distribution of the data, the maximum likelihood (ML) estimator is often considered as the reference estimator because it is generally (when it exists and is unique) asymptotically efficient with a speed of convergence in $\sqrt{n}$. From Slepian-Bangs formula (111), the Fisher information matrix (FIM) for the parameter $(\boldsymbol{\mu}, v(\boldsymbol{\Sigma}))$ is given by

$$\mathrm{FIM}\begin{pmatrix}\boldsymbol{\mu}\\ v(\boldsymbol{\Sigma})\end{pmatrix} = n\begin{pmatrix} a_0\boldsymbol{\Sigma}^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{D}^T[a_1(\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1}) + a_2\mathrm{vec}(\boldsymbol{\Sigma}^{-1})\mathrm{vec}^T(\boldsymbol{\Sigma}^{-1})]\mathbf{D}\end{pmatrix}, \tag{88}$$

where

$$a_0 = \frac{\mathrm{E}[\mathcal{Q}\varphi^2(\mathcal{Q})]}{m},\ a_1 = \frac{\mathrm{E}[\mathcal{Q}^2\varphi^2(\mathcal{Q})]}{2m(m+2)} \text{ and } a_2 = \frac{1}{4}\left(\frac{\mathrm{E}[\mathcal{Q}^2\varphi^2(\mathcal{Q})]}{m(m+2)} - 1\right), \tag{89}$$

assuming that g is continuously differentiable with $\varphi(t) \stackrel{\mathrm{def}}{=} -2g'(t)/g(t)$. Hence under existence, uniqueness and usual regularity conditions, the ML estimate $(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}})$ is asymptotically Gaussian distributed:

$$\sqrt{n}\begin{pmatrix}\widehat{\boldsymbol{\mu}} - \boldsymbol{\mu}\\ \mathrm{vec}(\widehat{\boldsymbol{\Sigma}}) - \mathrm{vec}(\boldsymbol{\Sigma})\end{pmatrix} \to_d \mathbb{R}N_{m+m^2}\left(\begin{pmatrix}\mathbf{0}\\ \mathbf{0}\end{pmatrix}, \begin{pmatrix}\mathbf{R}_{\mu_{ML}} & \mathbf{0}\\ \mathbf{0} & \mathbf{R}_{\Sigma_{ML}}\end{pmatrix}\right), \tag{90}$$

where $\mathbf{R}_{\mu_{ML}}$ and $\mathbf{R}_{\Sigma_{ML}}$ can be deduced from (88) using the efficiency of the ML estimate:

$$\mathbf{R}_{\mu_{ML}} = \sigma_0\boldsymbol{\Sigma} \text{ and } \mathbf{R}_{\Sigma_{ML}} = \sigma_1(\mathbf{I}+\mathbf{K})(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma}) + \sigma_2\mathrm{vec}(\boldsymbol{\Sigma})\mathrm{vec}^T(\boldsymbol{\Sigma}), \tag{91}$$

with [59]

$$\sigma_0 = \frac{m}{\mathrm{E}[\mathcal{Q}\varphi^2(\mathcal{Q})]},\ \sigma_1 = \frac{m(m+2)}{\mathrm{E}[\mathcal{Q}^2\varphi^2(\mathcal{Q})]} \text{ and } \sigma_2 = -\frac{2\sigma_1(1-\sigma_1)}{2+m(1-\sigma_1)}. \tag{92}$$

Note that this asymptotic distribution was first given in [59] by using the general structure of the covariance of random matrices whose distributions are radial.

The ML estimator of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the vector $\widehat{\boldsymbol{\mu}}$ and the symmetric positive definite matrix $\widehat{\boldsymbol{\Sigma}}$ that minimize the negative log-likelihood function equal from (9) to

$$L(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{n}{2} \log |\boldsymbol{\Sigma}| - \sum_{i=1}^{n} \log \left(g[(\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu})] \right). \tag{93}$$

Setting the derivatives of $L(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ w.r.t. $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ to zero yields the following estimation equations:

$$\mathbf{0} = \frac{1}{n} \sum_{i=1}^{n} \varphi[(\mathbf{x}_i - \widehat{\boldsymbol{\mu}})^T \widehat{\boldsymbol{\Sigma}}^{-1} (\mathbf{x}_i - \widehat{\boldsymbol{\mu}})](\mathbf{x}_i - \widehat{\boldsymbol{\mu}}) \tag{94}$$

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{n} \sum_{i=1}^{n} \varphi[(\mathbf{x}_i - \widehat{\boldsymbol{\mu}})^T \widehat{\boldsymbol{\Sigma}}^{-1} (\mathbf{x}_i - \widehat{\boldsymbol{\mu}})](\mathbf{x}_i - \widehat{\boldsymbol{\mu}})(\mathbf{x}_i - \widehat{\boldsymbol{\mu}})^T. \tag{95}$$

Except for the case where the function φ is constant, the set of implicit equations (94)-(95) does not guarantee neither the existence nor the uniqueness of the ML estimators $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Sigma}}$. Some sufficient conditions to ensure their existence and uniqueness are given in [33]. Note that for the Gaussian distribution, $\varphi(t) = 1$ from (69) and (94)-(95) yield the sample mean $\widehat{\boldsymbol{\mu}} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i$ and the biased sample covariance matrix $\widehat{\boldsymbol{\Sigma}} = \frac{1}{n} \sum_{i=1}^{n} (\mathbf{x}_i - \widehat{\boldsymbol{\mu}})(\mathbf{x}_i - \widehat{\boldsymbol{\mu}})^T$, respectively.

6.3 M-estimators

Since the ML estimator may be drastically affected by the presence of outliers or when the data distribution deviates slightly from the RES distribution of the model, robust estimators have been proposed. Among the different families of robust estimators, in the following we focus our attention on the class of M-estimators. As for the ML estimator, an M-estimator can be obtained form the minimization of a function on the observation with respect to the parameters of interest. A classical example consists in replacing in the negative log-likelihood (93), the function $-\log(g([.])$ by a loss function $\rho(.) : \mathbb{R}^+ \mapsto \mathbb{R}$ (generally not related to g) to form another objective function. If ρ is also continuously differentiable with $u(t) \stackrel{\text{def}}{=} 2\rho'(t)$, an M-estimator is obtained by minimizing this new objective function. By replacing $\varphi(t)$ with $u(t)$ and setting again the derivatives of this objective function to zero, we obtain similar equations to (94) and (95). Some sufficient conditions are also given in [33] to ensure the existence and uniqueness of this M-estimator.

The seminal paper by Maronna [40] defined a more general class of M-estimator by replacing $\varphi(.)$ by the weight functions $u_1([.]^{1/2})$ and $u_2(.)$ in (94) and (95), respectively, i,e.,

$$\mathbf{0} = \frac{1}{n} \sum_{i=1}^{n} u_1 \left([(\mathbf{x}_i - \widehat{\boldsymbol{\mu}})^T \widehat{\boldsymbol{\Sigma}}^{-1} (\mathbf{x}_i - \widehat{\boldsymbol{\mu}})]^{1/2} \right) (\mathbf{x}_i - \widehat{\boldsymbol{\mu}}) \tag{96}$$

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{n} \sum_{i=1}^{n} u_2 \left((\mathbf{x}_i - \widehat{\boldsymbol{\mu}})^T \widehat{\boldsymbol{\Sigma}}^{-1} (\mathbf{x}_i - \widehat{\boldsymbol{\mu}}) \right) (\mathbf{x}_i - \widehat{\boldsymbol{\mu}})(\mathbf{x}_i - \widehat{\boldsymbol{\mu}})^T. \tag{97}$$

The functions $u_1([.]^{1/2})$ and $u_2(.) : \mathbb{R}^+ \mapsto \mathbb{R}$ need not be the same function and hence the more general M-estimators need not be related to a minimization problem. Under sufficient conditions (called Maronna conditions), it is proved in [40, Th. 4], the existence and uniqueness of solution $(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}})$ of (96), (97) but also [40, Th. 2] of $(\mathbf{t}, \mathbf{V})$ solution of

$$\mathbf{0} = \mathrm{E}\left[u_1\left([(\mathbf{x}_1 - \mathbf{t})^T \mathbf{V}^{-1}(\mathbf{x}_1 - \mathbf{t})]^{1/2}\right)(\mathbf{x}_1 - \mathbf{t})\right] \tag{98}$$

$$\mathbf{V} = \mathrm{E}\left[u_2\left((\mathbf{x}_1 - \mathbf{t})^T \mathbf{V}^{-1}(\mathbf{x}_1 - \mathbf{t})\right)(\mathbf{x}_1 - \mathbf{t})(\mathbf{x}_1 - \mathbf{t})^T\right]. \tag{99}$$

Sufficient conditions are also given in [40, Th. 5] to ensure the strong consistency of the estimate $(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}})$ solution of (96), (97) to the solution $(\mathbf{t}, \mathbf{V})$ of (98), (99) with $\mathbf{t} = \boldsymbol{\mu}$ and $\mathbf{V} = \sigma^{-1}\boldsymbol{\Sigma}$, where σ is the unique solution of $\mathrm{E}[\sigma\mathcal{Q}u_2(\sigma\mathcal{Q})] = m$ [59, Appendix 3].

Using a general result on M-estimators given in [28, Sec. 4], Maronna proved in [40, Th. 6] the asymptotic gaussianity of $(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}})$, where $\widehat{\boldsymbol{\mu}}$ and $\widehat{\boldsymbol{\Sigma}}$ are asymptotically independent and where only the covariance of the asymptotic distribution of $\widehat{\boldsymbol{\mu}}$ was specified. 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 of $\widehat{\boldsymbol{\Sigma}}$ was specified in [59, Appendix 2] to get:

$$\sqrt{n}\begin{pmatrix} \widehat{\boldsymbol{\mu}} - \boldsymbol{\mu} \\ \mathrm{vec}(\widehat{\boldsymbol{\Sigma}}) - \mathrm{vec}(\sigma^{-\mathbf{1}}\boldsymbol{\Sigma}) \end{pmatrix} \rightarrow_d \mathbb{R}N_{m+m^2}\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{R}_{\mu_M} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{\Sigma_M} \end{pmatrix}\right), \tag{100}$$

where

$$\mathbf{R}_{\mu_M} = \frac{\alpha}{\beta^2}\mathbf{V} = \frac{\alpha\sigma^{-1}}{\beta^2}\boldsymbol{\Sigma} \tag{101}$$

where $\alpha = \frac{1}{m}\mathrm{E}[\psi_1^2(\sqrt{\sigma\mathcal{Q}})]$ and $\beta = \mathrm{E}[(1-m^{-1})u_1(\sqrt{\sigma\mathcal{Q}}) + m^{-1}\psi_1'(\sqrt{\sigma\mathcal{Q}})]$ with $\psi_1(t) \stackrel{\mathrm{def}}{=} tu_1(t)$, and

$$\mathbf{R}_{\Sigma_M} = \sigma_1(\mathbf{I} + \mathbf{K})(\mathbf{V} \otimes \mathbf{V}) + \sigma_2 \mathrm{vec}(\mathbf{V})\mathrm{vec}^T(\mathbf{V}), \tag{102}$$

where $\sigma_1 = \frac{(m+2)^2 a_1}{(2a_2+m)^2}$ and $\sigma_2 = a_2^{-1}\left[(a_1 - 1) - 2(a_2 - 1)a_1\frac{m+(m+4)a_2}{(2a_2+m)^2}\right]$ with $a_1 = \frac{\mathrm{E}[\psi_2^2(\sigma\mathcal{Q})]}{m(m+2)}$ and $a_2 = \frac{\mathrm{E}[\sigma\mathcal{Q}\psi_2{}'(\sigma\mathcal{Q})]}{m}$ where $\psi_2(t) \stackrel{\mathrm{def}}{=} tu_2(t)$ and $\psi_2'(t) \stackrel{\mathrm{def}}{=} \frac{d\psi_2(t)}{dt}$. Furthermore, using the general structure of the covariance of radial random matrices, it is proved in [59, Th. 1] that

$$\sigma_2 \geq -\frac{2}{m}\sigma_1. \tag{103}$$

Finally note that for $u_1(t) = u_2(t^2) = \varphi(t^2)$, the M-estimate reduces to the ML estimate for which $\sigma = 1$ and (100) reduce to (90).

To the best of our knowledge, when both $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are unknown parameters, it does not exist in the literature an analysis of sufficient conditions under which an algorithm would jointly converge toward the solution of the implicit equations (96), (97). However, when $\boldsymbol{\mu}$ is known, an algorithm which is essentially a fixed-point

algorithm with a scale adjustment made at each iteration, with less severe conditions than that required in [40] was presented in [63]. Without loss of generality, $\boldsymbol{\mu}$ can be taken to $\mathbf{0}$ and the following algorithm is proved in [63] to converge to the unique solution of $\widehat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{i=1}^n u_2[\mathbf{x}_i^T\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{x}_i]\mathbf{x}_i\mathbf{x}_i^T$:

$$\boldsymbol{\Sigma}_0 = \frac{1}{n}\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i^T, \;\; \boldsymbol{\Sigma}_{k+1} = \frac{1}{n}\sum_{i=1}^n u_2[c_k\mathbf{x}_i^T\boldsymbol{\Sigma}_k^{-1}\mathbf{x}_i]\mathbf{x}_i\mathbf{x}_i^T, \tag{104}$$

c_k being the unique positive scalar satisfying $\frac{1}{n}\sum_{i=1}^n \psi_2[c_k\mathbf{x}_i^T\boldsymbol{\Sigma}_k^{-1}\mathbf{x}_i] = m$ with $\psi_2(t) \stackrel{\text{def}}{=} tu_2(t)$.

6.4 Tyler's M-estimator

When $\boldsymbol{\mu}$ is known (equal to $\mathbf{0}$ without loss of generality), the M-estimator proposed by Tyler [61] has become a very popular robust scatter estimator in the signal processing literature. This M-estimator is defined by its weight function $u_2(t) = \frac{m}{t}$ associated with the loss function $\rho(t) = \frac{m}{2}\log t$, leading to the following objective function

$$L_T(\boldsymbol{\Sigma}) = \frac{n}{2}\log|\boldsymbol{\Sigma}| + \frac{m}{2}\sum_{i=1}^n \log(\mathbf{x}_i^T\boldsymbol{\Sigma}^{-1}\mathbf{x}_i) \tag{105}$$

that is minimized by $\widehat{\boldsymbol{\Sigma}}$, solution of the implicit equation

$$\widehat{\boldsymbol{\Sigma}} = \frac{m}{n}\sum_{i=1}^n \frac{\mathbf{x}_i\mathbf{x}_i^T}{\mathbf{x}_i^T\widehat{\boldsymbol{\Sigma}}^{-1}\mathbf{x}_i}. \tag{106}$$

Note that $L_T(c^2\boldsymbol{\Sigma}) = L_T(\boldsymbol{\Sigma}) + a$ where a does not depend on $\boldsymbol{\Sigma}$. Consequently, if $\widehat{\boldsymbol{\Sigma}}$ is solution of (106), $c^2\widehat{\boldsymbol{\Sigma}}$ is it too. The existence proof of solution of (106) given in Maronna [40] does not apply here, but existence and uniqueness (up to a multiplicative constant) under continuous RES distribution in $\mathbb{R}^m$ is proved in [61], by showing that it is the limiting point of the following specific fixed point algorithm: with $\boldsymbol{\Sigma}_0$ arbitrary symmetric positive definite matrix,

$$\boldsymbol{\Sigma}'_{k+1} = \frac{m}{n}\sum_{i=1}^n \frac{\mathbf{x}_i\mathbf{x}_i^T}{\mathbf{x}_i^T\boldsymbol{\Sigma}_k^{-1}\mathbf{x}_i}, \;\; \boldsymbol{\Sigma}_{k+1} = \frac{m}{\mathrm{Tr}(\boldsymbol{\Sigma}'_{k+1})}\boldsymbol{\Sigma}'_{k+1}. \tag{107}$$

Under this same condition, it is also proved in [61] that the solution of (106) standardized so that $\mathrm{Tr}(\widehat{\boldsymbol{\Sigma}}) = m$ is strongly consistent to the symmetric positive definite matrix $\boldsymbol{\Sigma}$ which is also solution of

$$\boldsymbol{\Sigma} = m\mathrm{E}\left[\frac{\mathbf{x}_1\mathbf{x}_1^T}{\mathbf{x}_1^T\boldsymbol{\Sigma}^{-1}\mathbf{x}_1}\right] \;\; \text{with} \;\; \mathrm{Tr}(\boldsymbol{\Sigma}) = m. \tag{108}$$

Furthermore, under continuous RES distribution in $\mathbb{R}^m$, it is proved in [61] that the solution $\widehat{\mathbf{\Sigma}}$ of (106) constrained to $\mathrm{Tr}(\widehat{\mathbf{\Sigma}}) = m$ is asymptotically Gaussian distributed, i.e.,

$$\sqrt{n}\left(\mathrm{vec}(\widehat{\mathbf{\Sigma}}) - \mathrm{vec}(\mathbf{\Sigma})\right) \to_d \mathbb{R}N_{m^2}\left(\mathbf{0}, \mathbf{R}_{\Sigma_{Ty}}\right) \tag{109}$$

with

$$\mathbf{R}_{\Sigma_{Ty}} = \left(1 + \frac{2}{m}\right)(\mathbf{I} + \mathbf{K})(\mathbf{\Sigma} \otimes \mathbf{\Sigma}) - \frac{2}{m}\left(1 + \frac{2}{m}\right)\mathrm{vec}(\mathbf{\Sigma})\mathrm{vec}^T(\mathbf{\Sigma}). \tag{110}$$

Comparing (102) to (110), we see that $\sigma_2 = -\frac{2}{m}\sigma_1$ and therefore we have equality in (103).

Note that when $\mathbf{x}_i$ are SIRV distributed, i.e., $\mathbf{x}_i \sim \mathrm{RCG}_m(\mathbf{0}, \mathbf{\Sigma}, F_{\tau_i})$ where here τ_i are assumed deterministic and unknown, the ML estimate of $\mathbf{\Sigma}$ coincides with Tyler's M-estimate (106). This estimate was introduced in [21] and existence and uniqueness were studied in [51]. Tyler's M-estimator enjoys four interesting properties:

- Objective function (105) is also the negative log-likelihood of i.i.d. data from an $\mathrm{ACG}_m(\mathbf{0}, \mathbf{\Sigma})$ distribution (see (84)). Consequently Tyler's M-estimator of $\mathbf{\Sigma}$ is the ML estimator $\widehat{\mathbf{\Sigma}}$ (satisfying the constraint $\mathrm{Tr}(\widehat{\mathbf{\Sigma}}) = m$) under the $\mathrm{ACG}_m(\mathbf{0}, \mathbf{\Sigma})$ distribution.
- Replacing $\mathbf{x}_i$ by $\frac{\mathbf{x}_i}{\|\mathbf{x}_i\|}$ in (106) does not affect the solution of (106). Since $\frac{\mathbf{x}_i}{\|\mathbf{x}_i\|}$ is $\mathrm{ACG}_m(\mathbf{0}, \mathbf{\Sigma})$ distributed for arbitrary $\mathrm{RES}_m(\mathbf{0}, \mathbf{\Sigma}, g)$ distribution, the distribution of Tyler's M estimator $\widehat{\mathbf{\Sigma}}$ does not depend on the density generator of this RES distribution. This is the reason why Tyler referred to his estimate as a distribution-free estimator [61].
- Suppose $\mathbf{x}_1, .., \mathbf{x}_i, .., \mathbf{x}_n$ are independent where $\mathbf{x}_i \sim \mathrm{RES}_m(\mathbf{0}, c_i^2\mathbf{\Sigma}, g_i)$ with unknown parameters c_i^2 and arbitrary unknown density generators g_i, then the ML of $\mathbf{\Sigma}$ corresponds to Tyler's M-estimate [47, Th. 1].
- Tyler's M-estimator can be considered as the most robust estimator of the scatter matrix of a RES distribution in the sense of minimizing the maximum asymptotic covariance w.r.t. the generator density (see [61, Remark 3.1], and [60, Th. 1].

Also, note that Tyler's M-estimate has been extensively studied in the statistics literature (e.g., [34] has specified the existence proof of solution of (106) and [20] have proved that Tyler's iterative procedure (107) has a linear convergence rate).

We finally note that many complementary results have been carried out on RES distributions (see e.g., [50] for ML estimate of GG scatter, [57] for Tyler's estimate of structured scatter). Furthermore, most of the definitions and results presented in this section have been extended to CES distributions with many new results (see, e.g., [10, 14, 39, 42, 43, 51, 52]).

6.5 Slepian-Bangs formula

We use in this Section the real to complex representation to unify the Slepian-Bangs (SB) for RES, C-CES and NC-CES distributed data. The SB formula provides a convenient way to compute the FIM and thus the Cramer-Rao bound (CRB) on the real-valued parameter $\boldsymbol{\alpha}$ parameterizing and characterizing the couple $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ of elliptically symmetric distributions. We omit this dependence in $\boldsymbol{\alpha}$ to simplify the notations. To derive this formula, it is necessary that the scatter matrix $\boldsymbol{\Sigma}$ is not singular and the second-order moments of $\mathbf{x}$ are finite, which is equivalent to the first-order moments of $\mathcal{Q}$ being finite. Furthermore, to avoid the ambiguity between the scatter matrix and the density generator, we either assume that $\boldsymbol{\Sigma} = \mathrm{Cov}(\mathbf{x})$ or that there is a scale constraint on $\boldsymbol{\Sigma}$. This formula has been derived for the real Gaussian distribution in [56] and [5], then extended to the circular complex Gaussian and non-circular Gaussian case in [58] and [13], respectively. This formula has been extended to C-CES distributions in [8] and [26], and recently to NC-CES distributions [1]. In all these scenarios, the density generator is assumed to be perfectly known. For RES distributed $\mathbf{x}$, all the steps of the proof of the SB formula for C-CES distributions given in [8] apply, and we get the following structured matrix SB formula:

$$\begin{aligned}\mathrm{CRB}^{-1}(\boldsymbol{\alpha}) = a_0 \frac{d\boldsymbol{\mu}^T}{d\boldsymbol{\alpha}^T} \boldsymbol{\Sigma}^{-1} \frac{d\boldsymbol{\mu}}{d\boldsymbol{\alpha}^T} \\ + \left(\frac{d\mathrm{vec}(\boldsymbol{\Sigma})}{d\boldsymbol{\alpha}^T}\right)^T \left(a_1(\boldsymbol{\Sigma}^{-T} \otimes \boldsymbol{\Sigma}^{-1}) + a_2 \mathrm{vec}(\boldsymbol{\Sigma}^{-1})\mathrm{vec}^T(\boldsymbol{\Sigma}^{-1})\right) \\ \left(\frac{d\mathrm{vec}(\boldsymbol{\Sigma})}{d\boldsymbol{\alpha}^T}\right), \quad (111)\end{aligned}$$

where $a_0 = \xi_{r,1,m}$, $a_1 = \frac{1}{2}\xi_{r,2,m}$ and $a_2 = \frac{1}{4}(\xi_{r,2,m} - 1)$ with

$$\xi_{r,1,m} \stackrel{\mathrm{def}}{=} \frac{\mathrm{E}[Q\varphi_r^2(Q)]}{m} \text{ and } \xi_{r,2,m} \stackrel{\mathrm{def}}{=} \frac{\mathrm{E}[Q^2\varphi_r^2(Q)]}{m(m+2)}, \quad (112)$$

where $Q \stackrel{\mathrm{def}}{=} \mathcal{Q}_{r,m}$ and $\varphi_r(t) \stackrel{\mathrm{def}}{=} -\frac{2}{g_{r,m}(t)} \frac{dg_{r,m}(t)}{dt}$.

Note that because $(c^2\boldsymbol{\Sigma}, c^2Q, c^m g_{r,m}(.\ c^2), c^2\varphi_r(.\ c^2))$ gives the same RES distribution, the coefficient $\xi_{r,2,m}$ is free of scale ambiguity in contrast to $\xi_{r,1,m}$ which depends on the scale factor. This is consistent with eq. (111).

This SB formula allows us to directly deduce those of NC-CES distributed data obtained thanks to the real to complex representation. This SB formula is similarly structured where $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, $\frac{d\boldsymbol{\mu}^T}{d\boldsymbol{\alpha}^T}$, $\left(\frac{d\mathrm{vec}(\boldsymbol{\Sigma})}{d\boldsymbol{\alpha}^T}\right)^T$ and $\mathrm{vec}^T(\boldsymbol{\Sigma}^{-1})$ in (111) are replaced by $\widetilde{\boldsymbol{\mu}}$, $\widetilde{\boldsymbol{\Sigma}}$, $\frac{d\widetilde{\boldsymbol{\mu}}^H}{d\boldsymbol{\alpha}^T}$, $\left(\frac{d\mathrm{vec}(\widetilde{\boldsymbol{\Sigma}})}{d\boldsymbol{\alpha}^T}\right)^H$ and $\mathrm{vec}^H(\widetilde{\boldsymbol{\Sigma}}^{-1})$, respectively, where $a_0 = \xi_{c,1,m}$, $a_1 = \frac{\xi_{c,2,m}}{2}$ and $a_2 = \frac{1}{4}(\xi_{c,2,m} - 1)$ with

$$\xi_{c,1,m} \stackrel{\mathrm{def}}{=} \frac{\mathrm{E}[Q\varphi_c^2(Q)]}{m} \text{ and } \xi_{c,2,m} \stackrel{\mathrm{def}}{=} \frac{\mathrm{E}[Q^2\varphi_c^2(Q)]}{m(m+1)}, \quad (113)$$

where $Q \stackrel{\text{def}}{=} \mathcal{Q}_{c,m}$ and $\varphi_c(t) \stackrel{\text{def}}{=} -\frac{1}{g_{c,m}(t)}\frac{dg_{c,m}(t)}{dt}$. On the other hand, the SB formulas for C-CES distributed data can be deduced directly by replacing $\widetilde{\boldsymbol{\Sigma}}$ by $\begin{pmatrix} \boldsymbol{\Sigma} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}^* \end{pmatrix}$, yielding the SB formulas proved in [8] and [26], which are also similarly structured with $a_0 = 2\xi_{c,1,m}$, $a_1 = \xi_{c,2,m}$ and $a_2 = \xi_{c,2,m} - 1$.

Note that (a_0, a_1, a_2) reduces to $(1, 1/2, 0)$, $(2, 1, 0)$ and $(1, 1/2, 0)$ for real, complex noncircular and complex circular Gaussian distributions, respectively. Note also that the decoupling between the parameters $\boldsymbol{\alpha}_1$ of $\boldsymbol{\mu}$ and the parameters $\boldsymbol{\alpha}_2$ of $\boldsymbol{\Sigma}$ when $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ have no parameters in common for Gaussian distributions, extends to any elliptically symmetric distributions with

$$\mathrm{CRB}(\boldsymbol{\alpha}_1) = \left(a_0 \frac{d\boldsymbol{\mu}^T}{d\boldsymbol{\alpha}_1{}^T} \boldsymbol{\Sigma}^{-1} \frac{d\boldsymbol{\mu}}{d\boldsymbol{\alpha}_1{}^T} \right)^{-1} \tag{114}$$

and

$$\mathrm{CRB}(\boldsymbol{\alpha}_2) = \left(\left(\frac{d\mathrm{vec}(\boldsymbol{\Sigma})}{d\boldsymbol{\alpha}_2{}^T} \right)^T \left(a_1 (\boldsymbol{\Sigma}^{-T} \otimes \boldsymbol{\Sigma}^{-1}) + a_2 \mathrm{vec}(\boldsymbol{\Sigma}^{-1}) \mathrm{vec}^T(\boldsymbol{\Sigma}^{-1}) \right) \left(\frac{d\mathrm{vec}(\boldsymbol{\Sigma})}{d\boldsymbol{\alpha}_2{}^T} \right) \right)^{-1}. \tag{115}$$

The coefficients $\xi_{c,1,m}$ and $\xi_{c,2,m}$ are deduced from $\xi_{r,1,m}$ and $\xi_{r,2,m}$ by the general relations

$$\xi_{c,1,m} = \xi_{r,1,2m} \text{ and } \xi_{c,2,m} = \xi_{r,2,2m}. \tag{116}$$

For example, the coefficients $\xi_{c,1,m}$ and $\xi_{c,2,m}$ have been calculated for complex Student's t and complex generalized Gaussian distributions in [8] and [26]. They are given respectively by:

$$\xi_{c,1,m} = \frac{\nu/2}{((\nu/2)-1)} \frac{(\nu/2)+m}{((\nu/2)+m+1)} \text{ and } \xi_{c,2,m} = \frac{(\nu/2)+m}{(\nu/2)+m+1}, \tag{117}$$

$$\xi_{c,1,m} = \frac{\Gamma(2+\frac{m-1}{s})\Gamma(\frac{m+1}{s})}{(\Gamma(1+\frac{m}{s}))^2} \quad \text{and } \xi_{c,2,m} = \frac{m+s}{m+1}. \tag{118}$$

Finally note that the SB formula for elliptically symmetric distributions has been extended when the data model is misspecified by the parametric probabilistic model [18], and when the density generator is considered as an infinite-dimensional nuisance parameter [41] or parameterized by a nuisance parameter [2].

7 Conclusion

The aim of this chapter was to provide a short overview of the main properties of elliptically symmetric distributions and it can be used as a background for all the other

chapters in this book. There is no claim to completeness in the material presented here. The reader interested in deeper discussions and investigations on specific aspects may find the references list useful to this goal. As a last concluding remark, we would like to highlight our choice to focus mainly on the RES distributions. As explained in the chapter, this class can be considered as the most general one since it encompasses the C-CES and NC-CES distributions as special cases. Some effort has been then put into showing explicitly the mapping between the RES class and all its sub-class. We hope that this chapter may represent a reference for the reader helping him to not get lost while going down his path through this book.

References

1. 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)
2. Abeida, H., Delmas, J.P.: Slepian-Bangs formulas for parameterized density generator of elliptically symmetric distributions. Signal Processing **205** (2023)
3. Anderson, T.W.: An Introduction to Multivariate Statistical Analysis. Wiley, New York (1984)
4. Andrews, D., Mallows, C.: Scale mixtures of normal distributions. J. Royal Stat. Soc., Ser. B **36**(1), 99–102 (1974)
5. Bangs, W.J.: Array processing with generalized beamformers. Ph.D. thesis, Yale University, New Haven, CT. (1971)
6. Bentler, P.M.: Some contributions to efficient statistics in structural models: specification and estimation of moments structure. Psychometrika **48**, 493–517 (1983)
7. Bentler, P.M., Berkane, M.: Greatest lower bound to the elliptical theory kurtosis parameter. Biometrika **73**, 240–241 (1986)
8. Besson, O., Abramovich, Y.I.: On the Fisher information matrix for multivariate elliptically contoured distributions. IEEE Signal Processing Letters **20**, 1130–1133 (2013)
9. Cambanis, S., Huang, S., Simons, G.: On the theory of elliptically contoured distributions. Journal of Multivariate Analysis **11**, 368–385 (1981)
10. Chitour, Y., Pascal, F.: Exact maximum likelihood estimates for sirv covariance matrix: Existence and algorithm analysis. IEEE Trans. Signal Processing **56**(10), 4563–4573 (2008)
11. Conte, E., Longo, M.: Characterisation of radar clutter as a spherically invariant random process. Proc. Inst. Electr. Eng. F **134**(2), 191–197 (1981)
12. Conte, E., Lops, M., Ricci, G.: Asymptotically optimum radar detection in compound-gaussian clutter. IEEE Trans. Aerosp. Electron. Syst. **31**(2), 617–624 (1995)
13. Delmas, J.P., Abeida, H.: Stochastic Cramér-Rao bound for noncircular signals with application to DOA estimation. IEEE Trans. Signal Processing **52**, 3192–3199 (2004)
14. Draskovic, G., Pascal, F.: New insights into the statistical properties of M-estimators. IEEE Trans. Signal Processing **66**(16), 4253–4263 (2018)
15. Eriksson, J., Koivunen, V.: Complex random vectors and ICA models: identifiability, uniqueness, and separability. IEEE Trans. Information Theory **52**, 1561–1565 (2006)
16. Fang, K., Kotz, S., Ng, K.W.: Symmetric multivariate and related distributions. Chapman & Hall, New-York (1990)
17. Fang, K., Zhang, Y.: Generalized multivariate analysis. Springer-Verlag, Berlin, Germany (1990)
18. 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 Processing **67**(1), 164–177 (2019)
19. Frahm, G.: Generalized elliptical distribution: theory and applications. Ph.D. thesis, University of Koln, Germany (2014)
20. Franks, W.C., Moitra, A.: Rigorous guarantees for tyler's M-estimator via quantum expansion. In: Proceedings of Machine Learning Research, vol. 125, pp. 1–32 (2020)
21. Gini, F.: Sub-optimum coherent radar detection in a mixture of k-distributed and gaussian clutter. In: Proc. Inst. Electr. Eng. Radar, Sonar Navigat., vol. 144, pp. 39–48 (1997)
22. Gomez-Sanchez-Manzano, E., Gomez-Villegas, M., Marin, J.: A multivariate generalization of the power exponential family of distributions. Comm. Statist. Theory Methods **27**, 589–600 (1998)
23. Gomez-Sanchez-Manzano, E., Gomez-Villegas, M., Marin, J.: A survey on continuous elliptical vector distributions. Rev. Mat. Comput **16**, 345–361 (2003)
24. Gomez-Sanchez-Manzano, E., Gomez-Villegas, M., Marin, J.: Sequences of elliptical distributions and mixtures of normal distributions. J. Multivariate Anal. **97**, 295–310 (2006)
25. Gomez-Sanchez-Manzano, E., Gomez-Villegas, M., Marin, J.: Multivariate exponential power distributions as mixtures of normal distributions with bayesian applications. Communications in Statistics - Theory and Methods **37**(6), 972–985 (2008)

26. Greco, M., Gini, F.: Cramer-Rao lower bounds on covariance matrix estimation for complex elliptically symmetric distributions. IEEE Trans. Signal Processing **61**, 6401–6409 (2013)
27. Horn, R.A., Johnson, C.A.: Matrix Analysis. Cambridge, University Press, New York (1985)
28. 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)
29. Hult, H., Lindskog, F.: Multivariate extremes, aggregation and dependence in elliptical distributions. Advances in Applied Probability **34**, 587–608 (2002)
30. Johnson, N.L., Kotz, S., Balakrishnan, N.: Continuous Univariate Distributions. John Wiley & Sons (1994)
31. Kano, Y.: Consistency property of elliptic probability density functions. Journal Multivariate Analysis **51**(1), 139–147 (1994)
32. Kelker, D.: Distribution theory of spherical distributions and a location-scale parameter generalization. Sankhya, Series A **32**, 419–430 (1970)
33. Kent, J.T., Tyler, D.E.: Redescending M-estimates of multivariate location and scatter. Annals of Statistics **19**(4), 2102–2119 (1991)
34. King, A.D.: Moduli of representation of finite dimension algebras. The Quaterly Journal of Mathematics **45**(4), 515–530 (1994)
35. Kotz, S.: Statistical Distributions in Scientific Work I, ch. Multivariate distributions at a cross road. Dordrecht: Reidel (1968)
36. Krishnaiah, P.R., Lin, J.: Complex elliptically symmetric distributions. Comm. Statist. - Theory. and Methods **15**, 3693–3718 (1986)
37. Lange, K.L., Little, R.J.A., Taylor, J.M.G.: Robust statistical modeling using the t distribution. Journal of the American Statistical Association **84**(408), 881–896 (1989)
38. Lindsey, J.K.: Multivariate elliptically contoured distributions for repeated measurements. Biometrics. **55**, 1277–1280 (1999)
39. Mahot, M., Pascal, F., Forster, P., Ovarlez, J.P.: Asymptotic properties of robust complex covariance matrix estimates. IEEE Trans. Signal Processing **61**(13), 3348–3356 (2013)
40. Maronna, R.: Robust M-estimators of multivariate location and scatter. The annals of statistics **4**(1), 51–67 (1976)
41. Mennad, A., Fortunati, S., El Korso, M.N., Younsi, A., Zoubir, A.M., Renaux, A.: Slepian-Bangs-type formulas and the related misspecified Cramer-Rao bounds for complex elliptically symmetric distributions. Signal Processing **142**, 320–329 (2018)
42. Meriaux, B., C.Ren, El Korso, M.N., Breloy, A., Forster, P.: Asymptotic performance of complex m-estimators for multivariate location and scatter estimation. IEEE Signal Processing Letters **26**(2), 363–371 (2019)
43. Meriaux, B., Ren, C., El Korso, M.N., Breloy, A., Forster, P.: Efficient estimation of scatter matrix with convex structure under t-distribution. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018). Calgary, Alberta, Canada (2018)
44. Nolan, J.P.: Multivariate elliptically contoured stable distributions: theory and estimation. Computational Statistics, Springer **28**, 2067–2089 (2013)
45. Ollila, E., Eriksson, J., Koivunen, V.: Complex elliptically symmetric random variables-generation, characterization, and circularity tests. IEEE Trans. Signal Processing **59**, 58–69 (2011)
46. Ollila, E., Koivunen, V.: Generalized complex elliptical distributions. In: Proc. SAM Workshop. pp. 460-464 (2014)
47. Ollila, E., Tyler, D.E.: Distribution-free detection under complex elliptically symmetric clutter distribution. In: IEEE Sensor Array and Multichannel Signal Processing Workshop - SAM 2012. Hoboken, NJ, USA (2012)
48. Ollila, E., Tyler, D.E., Koivunen, V., Poor, H.V.: Complex elliptically symmetric distributions: Survey, new results and applications. IEEE Trans. Signal Processing **60**(11), 5597–5625 (2012)
49. Paindaveine, D.: Elliptical symmetry. In book Wiley Statistics Reference Online (2014)
50. Pascal, F., Bombrun, L., J.-Y. Tourneret, Berthoumieu, Y.: Parameter estimation for multivariate generalized gaussian distributions. IEEE Trans. Signal Processing **61**(23) (2013)

51. Pascal, F., Chitour, Y., Ovarlez, J.P., Forster, P., Larzabal, P.: Covariance structure maximum likelihood estimates in compound gaussian noise: Existence and algorithm analysis. IEEE Trans. Signal Processing **56**, 34–48 (2008)
52. Pascal, F., Forster, P., Ovarlez, J.P., Larzabal, P.: Performance analysis of covariance matrix estimates in impulsive noise. IEEE Trans. Signal Processing **56**(6), 2206–2216 (2008)
53. Rangaswamy, M., Wainer, D., Ozturk, A.: Non-gaussian random vector identification using spherically invariant random processes. IEEE Trans. Aerospace and Electronic systems **29**(1), 111–124 (1993)
54. Richmond, C.D.: Adaptive array signal processing and performance analysis in non-Gaussian environments. PhD thesis, Massachusetts Institute of Technology (1996)
55. Shorackd, G.R.: Probability for statisticians. Springer, New-York (2000)
56. Slepian, D.: Estimation of signal parameters in the presence of noise. Transactions of the IRE Professional Group on Information Theory 3 **85**, 68–89 (1954)
57. Soloveychik, I., Wiesel, A.: Tyler's covariance matrix estimator in elliptical models with convex structure. IEEE Trans. Signal Processing **62**(20) (2014)
58. Stoica, P., Moses, R.: Spectral Analysis of Signals. Prentice Hall, NJ (2005)
59. Tyler, D.E.: Radial estimates and the test for sphericity. Biometrika **69**(2), 429 (1982)
60. Tyler, D.E.: Robustness and efficiency properties of scatter matrices. Biometrika **70**(2), 411 (1983)
61. Tyler, D.E.: A distribution-free M-estimator of multivariate scatter. The Annals of Statistics **15**(1), 234–251 (1987)
62. Tyler, D.E.: Statistical analysis for the angular central gaussian distribution on the sphere. Biometrika **74**(3), 579–589 (1987)
63. Tyler, D.E.: Some results on the existence, uniqueness, and computation of the M-estimates of multivariate location and scatter. SIAM Journal on Scientific and Statistical Computing **9**, 354 (1988)
64. Ward, K., R.J.A.Tough, Watts, S.: Sea clutter scattering the K distribution and radar performance. IET Radar, Sonar, Navigation and Avionics, London, U.K (2006)
65. Yao, K.: A representation theorem and its applications to spherically invariant random processes. IEEE Trans. Information Theory **19**(5), 600–608 (1973)