Jean-Pierre Delmas
Mohammed Nabil El Korso
Stefano Fortunati
Frédéric Pascal *Editors*

Elliptically Symmetric Distributions in Signal Processing and Machine Learning

Springer

Elliptically Symmetric Distributions in Signal Processing and Machine Learning

Jean-Pierre Delmas · Mohammed Nabil El Korso ·
Stefano Fortunati · Frédéric Pascal
Editors

Elliptically Symmetric Distributions in Signal Processing and Machine Learning

Springer

Editors
Jean-Pierre Delmas
laboratoire Samovar
Institut Polytechnique de Paris, Telecom SudParis
Palaiseau, France

Mohammed Nabil El Korso
Laboratoire des signaux et systèmes
Université Paris-Saclay, CNRS, CentraleSupélec
Gif-sur-Yvette, France

Stefano Fortunati
Laboratoire des signaux et systèmes
Université Paris-Saclay, CNRS, CentraleSupélec & IPSA
Gif-sur-Yvette, France

Frédéric Pascal
Laboratoire des signaux et systèmes
Université Paris-Saclay, CNRS, CentraleSupélec
Gif-sur-Yvette, France

ISBN 978-3-031-52115-7 ISBN 978-3-031-52116-4 (eBook)
https://doi.org/10.1007/978-3-031-52116-4

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

To our families

Preface

Aim of the Book

One of the main goals of Statistical Signal Processing and, more recently, Statistical Learning approaches is to infer/learn information from a set of acquired data. Nowadays, datasets could be more and more complex, e.g., high dimensional, heterogeneous, with missing data. Remarkably, statistical approaches for estimating a set of parameters or testing predefined hypotheses using a collection of noisy measurements have triggered key progress in many fields, such as communication, remote sensing, image processing, radar and sonar processing, seismology, and finance. A critical prerequisite for the inference procedure is the choice of a statistical model for the measured data, i.e., a data distribution. One of the most common assumptions in engineering practice is to assume a Gaussian distribution. While the Gaussianity assumption may be well justified in several scenarios where the Central Limit Theorem can be invoked, everyday practice shows how frequently the acquired data stream can deviate from such a Gaussian modeling. Specifically, extensive analyses have highlighted the impulsive, heavy-tailed nature of the data. One can also mention the potential presence of outliers or missing data in the received dataset. Since it is well known that the performance of Gaussian-based inference procedures can decrease dramatically when the Gaussian assumption is violated, the need for a more general statistical data model takes center stage. The elliptical family of distributions provides a reliable generalization of the Gaussian distribution. The elliptical model can characterize a broad class of heavy- and light-tailed distributions, including the Gaussian one as a particular case. An elliptical distribution is fully characterized by its location vector and scatter/covariance matrix and by an additional functional parameter, called the density generator, that describes the distribution shape, such as the impulsiveness of the data distribution. One can cite, among others, the t-distribution or the generalized Gaussian distribution as part of the elliptical distribution family. Due to their flexibility and ability to model various data behaviors for different and heterogeneous applications, the elliptical distributions have started attracting a lot of attention from the Engineering community, specifically the Signal Processing one.

The aim of this book is thus to provide the interested reader with an up-to-date overview of the recent developments and practical uses of the elliptical model in classical and emerging Engineering areas, such as graph learning, robust clustering, linear shrinkage, information geometry, subspace-based algorithms, and semiparametric and misspecified estimation.

Outline of This Book

This book consists of ten chapters which are organized into an introductory Chap. Background on Real and Complex Elliptically Symmetric Distributions and three parts as follows:

Chapter 1 - Background on Real and Complex Elliptically Symmetric Distributions, by J-P. Delmas.

This chapter presents a short overview of real elliptically symmetric (RES) distributions, complemented by circular and noncircular complex elliptically symmetric 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 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, sample covariance matrix estimate, maximum likelihood estimators, M-estimators, and Tyler's M-estimators. Particular attention is given to the asymptotic (Gaussian) distribution of the M-estimators of the scatter matrix, and some hints about the Slepian–Bangs formula are provided.

Theoretical Developments

Chapter 2 - The Fisher–Rao Geometry of CES Distributions, by F. Bouchard, A. Breloy, A. Collas, A. Renaux, and G. Ginolhac;

Chapter 3 - Linear Shrinkage of Sample Covariance Matrix or Matrices Under Elliptical Distributions: A Review, by E. Ollila;

Chapter 4 - Robust Estimation with Missing Values for Elliptical Distributions, by A. Hippert-Ferrer and M. N. El Korso.

The second part of this book, which contains Chaps. The Fisher–Rao Geometry of CES Distributions–Robust Estimation with Missing Values for Elliptical Distributions, focuses on the theoretical aspects of the elliptical distributions. Although

Chap. Background on Real and Complex Elliptically Symmetric Distributions furnishes an extensive review of elliptical distributions, this part aims at discussing recent advances in the elliptical distributions theory. First, Chap. The Fisher–Rao Geometry of CES Distributions derives the Fisher–Rao geometry of the complex elliptically symmetric distributions. The Fisher metric associated with a Riemannian manifold allows for leveraging many tools from differential geometry, thus offering alternative and somehow more powerful approaches for parameter learning. Then, Chap. Linear Shrinkage of Sample Covariance Matrix or Matrices Under Elliptical Distributions: A Review addresses the fundamental problem of estimator regularization that can be encountered, *e.g.*, in high-dimensional settings. This chapter presents how to shrink the sample covariance matrix under elliptical distributions, providing optimal strategies for learning this shrinkage parameter. Finally, Chap. Robust Estimation with Missing Values for Elliptical Distributions deals with the problem of missing data under an elliptical distribution framework. This chapter reviews various robust estimation methods based on the expectation–maximization algorithm, designed for handling various patterns of missing values. Specifically, robust covariance matrix estimation under RES distribution (with a possible low-rank structure) and robust probabilistic component analysis (possibly under a mixed-effects model) will be studied.

Performance Analysis

Chapter 5 - Semiparametric Estimation in Elliptical Distributions, by S. Fortunati;

Chapter 6 - Estimation and Detection Under Misspecification and Complex Elliptically Symmetric Distributions, by C. D. Richmond and A. S. Bondre;

Chapter 7 - Performance Analysis of Subspace-Based Algorithms in CES Data Models, by J-P. Delmas and H. Abeida.

The third part of this book, which contains Chaps. Semiparametric Estimation in Elliptical Distributions–Performance Analysis of Subspace-Based Algorithms in CES Data Models is devoted to statistical performance analysis of some estimation and detection problems in the context of elliptically symmetric distributions. Chapter Semiparametric Estimation in Elliptical Distributions puts the problem of the joint estimation of the scatter matrix and of the location vector of an elliptically distributed set of observations in the general framework of the semiparametric theory. After having provided an introduction to this fundamental subject in modern robust statistics, the chapter aims at providing a semiparametric lower bound and the related semiparametric efficient estimator of the scatter matrix in the presence of an unknown density generator. After an extensive review of classic Cramér–Rao-type bounds, Chap. Estimation and Detection Under Misspecification and Complex Elliptically Symmetric Distributions aims at discussing their extensions when the assumed probability distribution for the observed data differs from the true distribution. Thus the theory of the misspecified parameter bounds is developed in this chapter. The generalized likelihood ratio test (GLRT) is also derived and its performances are assessed along with its finite sample receiver operating characteristic (ROC) in radar detection under well-specified complex compound Gaussian distributed data. Finally, Chap. Performance Analysis of Subspace-Based Algorithms in CES Data Models focuses on complex elliptically symmetric noisy linear mixture models whose parameters of the mixing matrix characterize its column subspace. This chapter unifies the asymptotic distribution (w.r.t. the number of samples) of subspace-based algorithms adapted to different models of the data and covariance matrix estimates and proves several invariance properties that have impacts on the parameters to be estimated.

Applications to Machine Learning

Chapter 8 - Robust Bayesian Cluster Enumeration for RES Distributions, by F. K. Teklehaymanot, C. Schroth, and M. Muma;

Chapter 9 - FEMDA: A Unified Framework for Discriminant Analysis, by P. Houdouin, M. Jonckheere, and F. Pascal;

Chapter 10 - Learning Graphs from Heavy-Tailed Data (with application to finance data), by J. V. de M. Cardoso, J. Ying, and D. P. Palomar.

The fourth part of this book, which contains Chaps. Robust Bayesian Cluster Enumeration for RES Distributions–Learning Graphs from Heavy-Tailed Data (with application to finance data), presents some supervised and unsupervised machine learning methods based on elliptically symmetric distributed data that advantageously replaces the classic Gaussian distribution to take into account heavy-tailed noise, artifacts, and outliers. First, Chap. Robust Bayesian Cluster Enumeration for RES Distributions revisits some recent developments on robust statistical model-based cluster analysis even when the number of clusters is unknown. In particular, Bayesian robust cluster enumeration criteria are discussed. Such approaches formulate the problem of estimating the number of clusters as a maximization of the posterior probability of the candidate models. Robust clustering algorithms are derived and applied to real-data examples. Chapter FEMDA: A Unified Framework for Discriminant Analysis presents a new approach that overcomes the limitations of the linear and quadratic discriminant analysis under Gaussian assumption. In the considered model, the data in each cluster share the same shape matrix with arbitrary scale parameters with its own elliptically symmetric distribution. The structure of the proposed algorithm involves recursively estimating the parameters of each cluster and then classifying test observations based on the derived decision rule where the unknown scatter and location parameters are estimated by a Bayesian approach with the maximum marginal likelihood estimators. Finally, Chap. Learning Graphs from Heavy-Tailed Data (with application to finance data) investigates the problem of learning graph matrices whose structure follows that of a Laplacian matrix of an undirected weighted graph for which the data-generating process is assumed to be Student's t-distributed. The underlying learning problem is solved via an algorithm based on alternating direction method of multipliers and applied to real-world data.

Audience of This Book

Although the elliptically symmetric distributions were introduced into the statistical literature in the 1970s and a tutorial article in the IEEE Transaction on Signal Processing introduced them to the signal processing community in 2012, no book has been devoted to the elliptically symmetric distributions to the signal processing and/or machine learning community. Since this tutorial article, there has been a proliferation of research, applications, and results in this field. An in-depth book that provides a timely overview, a systematic account, and a foundational theory is essential.

This book is mainly targeted at researchers and graduate students who rely on the theory of probability and statistics to conduct their work in signal processing (e.g., communications, radar, sonar, etc.) and machine learning.

The necessary background which is supposed to take profit from reading this book is sufficient knowledge in probabilities, statistics, analysis, and matrix algebra at the postgraduate level.

Evry, France — Jean-Pierre Delmas
Gif-sur-Yvette, France — Mohammed Nabil El Korso
Ivry-sur-Seine, France — Stefano Fortunati
Gif-sur-Yvette, France — Frédéric Pascal

Contents

Chapter 1
Background on Real and Complex Elliptically Symmetric Distributions

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.

1.1 Introduction

Until 50 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 behaviors. These 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

J.-P. Delmas (✉)
laboratoire Samovar, Institut Polytechnique de Paris, Telecom SudParis, 91120 Palaiseau, France
e-mail: jean-pierre.delmas@it-sudparis.eu

J.-P. Delmas et al. (eds.), *Elliptically Symmetric Distributions in Signal Processing and Machine Learning*, https://doi.org/10.1007/978-3-031-52116-4_1

elliptically symmetric (RES) distributions were equivalently defined in the statistical literature [9, 17, 18, 33] 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 [37] and further fully studied in [55], and in the tutorial paper [49]. Then, the general complex representation of the RES distributions, called noncircular CES (NC-CES) distributions, was introduced in [16].

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 1.2 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 1.3 defines the C-CES and NC-CES distributions as complex representations of the RES distributions. Section 1.4 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 Sect. 1.5. The joint estimation of the symmetry center and scatter matrix is briefly discussed in Sect. 1.6 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, Sect. 1.7 briefly concludes this chapter.

The following notation 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. E(.), $|.|$, and Tr(.) are the expectation, determinant, and trace operators, respectively. $\mathbf{I}$ is the identity matrix with the appropriate dimension. $\text{vec}(\mathbf{A})$ 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 $\text{vec}(\mathbf{A})$ 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 $\text{vec}(\mathbf{C}^T) = \mathbf{K}\text{vec}(\mathbf{C})$ and $\text{vec}(\mathbf{A}) = \mathbf{D}v(\mathbf{A})$ 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, $\Gamma(u) \stackrel{\text{def}}{=} \int_0^\infty t^{u-1}e^{-t}dt$ 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 $\text{Gam}(k, \theta)$ 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.

1.2 Definition of the Real Elliptically Symmetric Distributions

1.2.1 Characteristic Function

After earlier works on this topic, RES distributions were formalized by [33] and further studied by [9, 17, 18]. They were first defined as affine transformations of spherically distributed r.v. Then, by the uniqueness theorem (see, e.g., [56, pp. 346–351]), they were alternatively defined by their characteristic functions.

Definition 1.1 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 $\phi_r(.) : \mathbb{R}^+ \rightarrow \mathbb{R}$ called *symmetry center*, *scatter matrix*, and *characteristic generator*, respectively, such that the characteristic function of $\mathbf{x}$ is of the form

$$\Phi_x(\mathbf{t}) \stackrel{\text{def}}{=} \mathrm{E}[\exp(i\mathbf{t}^T\mathbf{x})] = \exp(i\mathbf{t}^T\boldsymbol{\mu})\phi_r(\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}), \qquad \mathbf{t} \in \mathbb{R}^m. \tag{1.1}$$

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 Sect. 1.4.1), 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.

1.2.2 Stochastic Representation

Equivalently to the definition (1.1), the RES distributed r.v. $\mathbf{x}$ can be defined from an affine function

$$\mathbf{x} \stackrel{\text{def}}{=} \boldsymbol{\mu} + \mathbf{A}\mathbf{x}_s \tag{1.2}$$

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 spherical distributions are defined equivalently in the following [17, Chap. 2].

Definition 1.2 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

$$\Phi_{x_s}(\mathbf{t}) = \phi_r(\|\mathbf{t}\|^2), \ \mathbf{t} \in \mathbb{R}^k. \tag{1.3}$$

- For every $\mathbf{h} \in \mathbb{R}^k$, $\mathbf{h}^T\mathbf{x}_s =_d \|\mathbf{h}\| x_{s_i}$ with $\mathbf{x}_s = (x_{s_1}, \ldots, x_{s_i}, \ldots, x_{s_k})^T$.
- 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

$$p(\mathbf{x}_s) = g_r(\|\mathbf{x}_s\|^2), \tag{1.4}$$

where

$$\delta_{r,k} \stackrel{\text{def}}{=} \int_0^\infty t^{k/2-1} g_r(t)dt = \frac{\Gamma(k/2)}{\pi^{k/2}}, \tag{1.5}$$

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

$$\mathbf{x}_s =_d \sqrt{\mathcal{Q}_{r,k}}\mathbf{u}_{r,k}. \tag{1.6}$$

Consequently, from (1.2) and (1.3), we find the characteristic function (1.1) 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 (1.2) and (1.6), we obtain the following.

Theorem 1.1 $\mathbf{x}$ *is* $\text{RES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi_r)$ *distributed, i.f.f. it admits the following stochastic full-rank representation*

$$\mathbf{x} =_d \boldsymbol{\mu} + \sqrt{\mathcal{Q}_{r,k}}\mathbf{A}\mathbf{u}_{r,k} = \boldsymbol{\mu} + \mathcal{R}_{r,k}\mathbf{A}\mathbf{u}_{r,k}. \tag{1.7}$$

The r.v. $\mathcal{Q}_{r,k}$ and $\mathcal{R}_{r,k} \stackrel{\text{def}}{=} \sqrt{\mathcal{Q}_{r,k}}$ are the second-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 $\mathbf{x} \sim \text{RES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, F_{R_r})$ and, equivalently to Definition 1.1, we retrieve the scale ambiguity in the couple $(\mathbf{A}, \mathcal{R}_{r,k})$ in (1.7). Theorem 1.1 provides an obvious mechanism to generate r.v. $\mathbf{x} \sim \text{RES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, F_{R_r})$: 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

[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 Sect. 1.5).

$\mathbf{n}_{r,k}$ 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 1.1:

$$\mathcal{Q}_{r,k} =_d (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^\#(\mathbf{x} - \boldsymbol{\mu}). \tag{1.8}$$

1.2.3 The Absolutely Continuous Case

From (1.2) 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 (1.6), 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., $\text{rank}(\boldsymbol{\Sigma}) = m$, $\mathbf{A}$ is a non-singular $m \times m$ square matrix. From the one-to-one mapping (1.2) between $\mathbf{x}$ and $\mathbf{x}_s$, and the p.d.f. (1.4), the p.d.f. of $\mathbf{x}$ on $\mathbb{R}^m$ can be expressed as

$$p(\mathbf{x}) = |\boldsymbol{\Sigma}|^{-1/2} g_r[(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})]. \tag{1.9}$$

We note that unlike the notation used in, e.g., [49], (1.9) does not explicitly include the usual p.d.f. normalizing constant. This is a consequence of the definition of $g_r(.)$ introduced in (1.4).

Alternatively, if in (1.9) $g_r(.)$: $\mathbb{R}^+ \mapsto \mathbb{R}^+$ is an arbitrary function, called *density generator*, such that

$$\delta_{r,m,g} \stackrel{\text{def}}{=} \int_0^\infty t^{m/2-1} g_r(t) dt < \infty, \tag{1.10}$$

$\delta_{r,m,g}$ depends now on g_r and the couple $(\boldsymbol{\Sigma}, g_r(.))$ in (1.9) does not uniquely identify the distribution of $\mathbf{x}$ because $(c^2\boldsymbol{\Sigma}, c^m g_r(.\ c^2))$ gives the same distribution.

We adopt the notation $\text{RES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$ instead of $\text{RES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$. 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 $\boldsymbol{\Sigma}$. This justifies the terminology of symmetrical elliptical distributions. Furthermore, the stochastic representation (1.7) reduces to

$$\mathbf{x} =_d \boldsymbol{\mu} + \sqrt{\mathcal{Q}_{r,m}} \boldsymbol{\Sigma}^{1/2} \mathbf{u}_{r,m} = \boldsymbol{\mu} + \mathcal{R}_{r,m} \boldsymbol{\Sigma}^{1/2} \mathbf{u}_{r,m}. \tag{1.11}$$

Here too, note that $(\boldsymbol{\Sigma}, \mathcal{Q}_{r,m})$ and $(c^2\boldsymbol{\Sigma}, c^{-2}\mathcal{Q}_{r,m})$ give the same distribution of $\mathbf{x}$. Equation (1.11) implies that (1.8) simplifies to

$$\mathcal{Q}_{r,m} =_d (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}). \tag{1.12}$$

From (1.9), (1.11), and (1.12), the p.d.f. of $\mathcal{Q}_{r,m}$ and $\mathcal{R}_{r,m}$ are, respectively,

$$p(q) = \delta_{r,m}^{-1} q^{m/2-1} g_r(q) \ \text{ and } \ p(r) = 2\delta_{r,m}^{-1} r^{m-1} g_r(r^2). \tag{1.13}$$

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 $\text{RES}_m(\mathbf{0}, \mathbf{I}, \phi_r)$ distributions, where the explicit (but somewhat involved) form of ϕ_r can be found in [18, Theorem 2.51].

1.2.4 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 [65]. 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, 54, 66]) and as *scale mixtures of normal distributions* in the statistics literature [4, 25]. These distributions are defined by their stochastic representation.

Definition 1.3 An r.v. $\mathbf{x} \in \mathbb{R}^m$ is said to have a real CG distribution (RCG) if it admits the following representation:

$$\mathbf{x} =_d \boldsymbol{\mu} + \sqrt{\tau_r}\mathbf{n}, \tag{1.14}$$

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 $\mathbf{n} \sim \mathbb{R}N_m(\mathbf{0}, \mathbf{\Sigma})$, called the *speckle*. The r.v. $\sqrt{\tau_r}$ is often called *mixing variable* with mixing distribution in the statistical literature (see e.g., [25]). We write $\mathbf{x} \sim \text{RCG}_m(\boldsymbol{\mu}, \mathbf{\Sigma}, F_\tau)$ to denote this case.

Note that (1.14) can be rewritten as

$$\mathbf{x} =_d \boldsymbol{\mu} + \sqrt{\tau_r}\mathbf{A}\mathbf{n}_0, \tag{1.15}$$

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 $s \stackrel{\text{def}}{=} \|\mathbf{n}_0\|^2 \sim \chi_k^2 = \text{Gam}(k/2, 2)$ 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 (1.14) can also be written as

$$\mathbf{x} =_d \boldsymbol{\mu} + \mathcal{R}_{r,k}\mathbf{A}\mathbf{u}_{r,k}, \tag{1.16}$$

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

$$p(q) = \int_0^\infty \frac{1}{\Gamma(k/2)(2\tau)^{k/2}} q^{k/2-1} \exp(-q/(2\tau)) dF_\tau(\tau). \tag{1.17}$$

The characteristic function of $\Phi_{x_s}(\mathbf{t})$ of a RCG distributed r.v. $\mathbf{x}$ defined by (1.14) is given straightforwardly by

$$\begin{aligned}\Phi_{x_s}(\mathbf{t}) &= \exp(i\mathbf{t}^T\boldsymbol{\mu}) \int_0^\infty \mathrm{E}(\exp(i\mathbf{t}^T\sqrt{\tau}\mathbf{n})/\tau) dF_\tau(\tau) \\ &= \exp(i\mathbf{t}^T\boldsymbol{\mu}) \int_0^\infty \exp\left(-\frac{\tau}{2}\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}\right) dF_\tau(\tau) \\ &= \exp(i\mathbf{t}^T\boldsymbol{\mu})\phi_r(\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}),\end{aligned} \tag{1.18}$$

where the characteristic generator

$$\phi_r(u) = \int_0^\infty \exp\left(-\frac{\tau}{2}u\right) dF_\tau(\tau) \tag{1.19}$$

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 (1.14). 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

$$p(\mathbf{x}) = (2\pi)^{-m/2}|\boldsymbol{\Sigma}|^{-1/2} \int_0^\infty \tau^{-m/2} \exp\left(-\frac{1}{2\tau}(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right) dF_\tau(\tau). \tag{1.20}$$

Note that the p.d.f. (1.20) can always be written in the form (1.9) with the density generator

$$g_r(t) = (2\pi)^{-m/2} \int_0^\infty \tau^{-m/2} \exp\left(-\frac{t}{2\tau}\right) dF_\tau(\tau), \tag{1.21}$$

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 (1.21) reduces to the density generator (1.69) 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.

1.3 Definition of the Complex Elliptically Symmetric Distributions

1.3.1 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 [47] if the r.v. $\bar{\mathbf{x}} \stackrel{\text{def}}{=} (\text{Re}(\mathbf{x})^T, \text{Im}(\mathbf{x})^T)^T \in \mathbb{R}^{2m}$ 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 $\bar{\mathbf{t}}^T \bar{\boldsymbol{\mu}} = \text{Re}(\mathbf{t}^H \boldsymbol{\mu})$, $\bar{\mathbf{t}}^T \bar{\boldsymbol{\Sigma}} \bar{\mathbf{t}} = \frac{1}{2}\widetilde{\mathbf{t}}^H (\mathbf{M}\bar{\boldsymbol{\Sigma}}\mathbf{M}^H)\widetilde{\mathbf{t}} = \frac{1}{4}\widetilde{\mathbf{t}}^H \widetilde{\boldsymbol{\Sigma}}\widetilde{\mathbf{t}}$ with $\bar{\mathbf{t}} \stackrel{\text{def}}{=} (\text{Re}(\mathbf{t})^T, \text{Im}(\mathbf{t})^T)^T$, $\bar{\boldsymbol{\mu}} \stackrel{\text{def}}{=} (\text{Re}(\boldsymbol{\mu})^T, \text{Im}(\boldsymbol{\mu})^T)^T$, $\widetilde{\mathbf{t}} \stackrel{\text{def}}{=} (\mathbf{t}^T, \mathbf{t}^H)^T$ and $\widetilde{\boldsymbol{\Sigma}} \stackrel{\text{def}}{=} 2\mathbf{M}\bar{\boldsymbol{\Sigma}}\mathbf{M}^H$ of rank k structured as

$$\widetilde{\boldsymbol{\Sigma}} = \begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{\Omega} \\ \boldsymbol{\Omega}^* & \boldsymbol{\Sigma}^* \end{pmatrix}, \tag{1.22}$$

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 (1.1).

Theorem 1.2 *The characteristic function of an NC-CES distributed r.v.* $\mathbf{x} \in \mathbb{C}^m$ *is of the form*

$$\Phi_x(\mathbf{t}) \stackrel{\text{def}}{=} \text{E}[\exp(i\bar{\mathbf{t}}^T \bar{\mathbf{x}})] = \exp(i\text{Re}(\mathbf{t}^H \boldsymbol{\mu}))\phi_c\left(\frac{1}{2}\widetilde{\mathbf{t}}^H \widetilde{\boldsymbol{\Sigma}}\widetilde{\mathbf{t}}\right) \qquad \mathbf{t} \in \mathbb{C}^m, \tag{1.23}$$

where

$$\phi_c(u) \stackrel{\text{def}}{=} \phi_r\left(\frac{1}{2}u\right) \tag{1.24}$$

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 $\mathbf{x} \sim \text{CES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Omega}, \phi_c)$. In the particular case where $\boldsymbol{\Omega} = \mathbf{0}$, the rank of $\widetilde{\boldsymbol{\Sigma}}$ is even with $\text{rank}(\widetilde{\boldsymbol{\Sigma}}) = 2\text{rank}(\boldsymbol{\Sigma}) = k$ and $\mathbf{x}$ is C-CES distributed [37, 49, 55]. 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 $\mathbf{x} \sim \text{CES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi_c)$. Moreover, from (1.23), we get:

Theorem 1.3 *The characteristic function of a C-CES distributed r.v.* $\mathbf{x} \in \mathbb{C}^m$ *is of the form*

$$\Phi_x(\mathbf{t}) = \exp(i\mathrm{Re}(\mathbf{t}^H\boldsymbol{\mu}))\phi_c(\mathbf{t}^H\boldsymbol{\Sigma}\mathbf{t}), \qquad \mathbf{t} \in \mathbb{C}^m, \tag{1.25}$$

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.*

1.3.2 Stochastic Representation

From the definition of the NC-CES distribution, a simple complex-valued extension of the stochastic representation (1.7) 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 $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^H$ and $\boldsymbol{\Omega} = \mathbf{A}\boldsymbol{\Delta}_\kappa\mathbf{A}^T$ where $\boldsymbol{\Delta}_\kappa = \mathrm{Diag}(\kappa_1, \ldots, \kappa_k)$ is a real diagonal matrix with non-negative real entries $(\kappa_i)_{i=1,\ldots,k}$ [28, Corollary 4.6.12(b)]. Furthermore, it has been proved in [16] that $0 \le \kappa_i \le 1$. 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 [46, Sect. IV.C].

Theorem 1.4 $\mathbf{x}$ *is* $\mathrm{CES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Omega}, \phi_c)$ *distributed i.f.f. it admits the following stochastic representation*

$$\mathbf{x} =_d \boldsymbol{\mu} + \sqrt{\mathcal{Q}_{c,k}}\mathbf{A}[\boldsymbol{\Delta}_1\mathbf{u}_{c,k} + \boldsymbol{\Delta}_2\mathbf{u}^*_{c,k}], \tag{1.26}$$

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

In the particular case of C-CES distributions, $\boldsymbol{\Omega} = \mathbf{0}$, which is equivalent to $\boldsymbol{\Delta}_\kappa = \mathbf{0}$, i.e., $\boldsymbol{\Delta}_1 = \mathbf{I}$ and $\boldsymbol{\Delta}_2 = \mathbf{0}$ and consequently the stochastic representation (1.26) reduces to the well known stochastic representation reported in [49]:

$$\mathbf{x} =_d \boldsymbol{\mu} + \sqrt{\mathcal{Q}_{c,k}}\mathbf{A}\mathbf{u}_{c,k}. \tag{1.27}$$

Note that similarly to the RES distribution, the C-CES distribution can be defined from the affine function (1.2), 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 ($\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^H$) of the scatter matrix [37, 55]. 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 (1.2), we find the characteristic function (1.25) which defines the C-CES distribution since $\Phi_x(\mathbf{t}) \stackrel{\text{def}}{=} \mathrm{E}[\exp(i\mathrm{Re}(\mathbf{t}^H\mathbf{x}))] = \exp(i\mathrm{Re}(\mathbf{t}^H\boldsymbol{\mu}))\Phi_{x_s}(\mathbf{A}^H\mathbf{t}) = \exp(i\mathrm{Re}(\mathbf{t}^H\boldsymbol{\mu}))\phi_c(\mathbf{t}^H\boldsymbol{\Sigma}\mathbf{t})$.

1.3.3 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 $\text{rank}(\widetilde{\boldsymbol{\Sigma}}) = 2m$ and using the identities $(\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}})^T \bar{\boldsymbol{\Sigma}}^{-1} (\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}}) = (\widetilde{\mathbf{x}} - \widetilde{\boldsymbol{\mu}})^H \widetilde{\boldsymbol{\Sigma}}^{-1} (\widetilde{\mathbf{x}} - \widetilde{\boldsymbol{\mu}})$ and $|\bar{\boldsymbol{\Sigma}}| = 2^{-2m} |\widetilde{\boldsymbol{\Sigma}}|$, the p.d.f. (1.9) becomes

$$p(\mathbf{x}) = |\widetilde{\boldsymbol{\Sigma}}|^{-1/2} g_c \left[\frac{1}{2} (\widetilde{\mathbf{x}} - \widetilde{\boldsymbol{\mu}})^H \widetilde{\boldsymbol{\Sigma}}^{-1} (\widetilde{\mathbf{x}} - \widetilde{\boldsymbol{\mu}}) \right], \tag{1.28}$$

where $g_c(t)$ is defined by

$$g_c(t) \stackrel{\text{def}}{=} 2^m g_r(2t), \tag{1.29}$$

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

$$\delta_{c,m} \stackrel{\text{def}}{=} \int_0^{\infty} t^{m-1} g_c(t) dt = \delta_{r,2m,g} < \infty. \tag{1.30}$$

We note that, in this case, (1.26) 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

$$\mathcal{Q}_{c,m} = \frac{1}{2} \mathcal{Q}_{r,2m} \tag{1.31}$$

and

$$\mathcal{Q}_{c,m} =_d \frac{1}{2} (\widetilde{\mathbf{x}} - \widetilde{\boldsymbol{\mu}})^H \widetilde{\boldsymbol{\Sigma}}^{-1} (\widetilde{\mathbf{x}} - \widetilde{\boldsymbol{\mu}}). \tag{1.32}$$

For C-CES distributed $\mathbf{x}$, (1.28) reduces to

$$p(\mathbf{x}) = |\boldsymbol{\Sigma}|^{-1} g_c[(\mathbf{x} - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})], \tag{1.33}$$

and (1.27) implies

$$\mathcal{Q}_{c,m} =_d (\mathbf{x} - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}). \tag{1.34}$$

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

$$p(q) = \delta_{c,m}^{-1} q^{m-1} g_c(q) \;\text{ and }\; p(r) = 2\delta_{c,m}^{-1} r^{2m-1} g_c(r^2). \tag{1.35}$$

1.3.4 The Subclass of Compound Gaussian Distributions

In many engineering applications, only the circular complex case is considered. An 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 1.3) where the associated extended scatter matrix (1.22) 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 (1.3) gives the following stochastic representation

$$\mathbf{x} =_d \boldsymbol{\mu} + \sqrt{\tau_c}\mathbf{n}, \tag{1.36}$$

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 Sect. 1.2.4 can be deduced from this real to complex representation. In particular with $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^H$ of rank k: $\mathbf{n}_0 \sim \mathbb{C}N_k(\mathbf{0}, \mathbf{I})$, (1.16) with now $\mathcal{R}_{c,k} \stackrel{\text{def}}{=} \sqrt{\tau_c s}$ where $s \stackrel{\text{def}}{=} \|\mathbf{n}_0\|^2 \sim \frac{1}{2}\chi^2_{2k} = \text{Gam}(k, 1)$. 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

$$p(q) = \int_0^\infty \frac{1}{\Gamma(k)\tau^k} q^{k-1} \exp(-q/\tau) dF_\tau(\tau) \tag{1.37}$$

and (1.20) becomes

$$p(\mathbf{x}) = \pi^{-m} |\boldsymbol{\Sigma}|^{-1} \int_0^\infty \tau^{-m} \exp\left(-\frac{1}{\tau} (\mathbf{x} - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right) dF_\tau(\tau). \tag{1.38}$$

1.4 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 $\boldsymbol{\Omega}$ defined in (1.22) 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 [20, 24]).

1.4.1 Moments

From the full-rank stochastic representation (1.7), it is clear that $\mathbf{x}$ admits pth-order moments i.f.f. $\text{E}(\mathcal{R}_k^p) < \infty$. Using the characteristic function (1.1), $\text{E}(\mathcal{R}_k^p) < \infty$ 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 $\text{E}(x_1^{p_1} x_2^{p_2} \ldots x_m^{p_m}) = \frac{1}{i^p} \frac{\partial^p \Psi_x(\mathbf{t})}{\partial t_1^{p_1} \partial t_2^{p_2} \ldots \partial t_m^{p_m}}|_{\mathbf{t}=\mathbf{0}}$ with $p = \sum_{i=1}^m p_i$ and $\mathbf{x} = (x_1, x_2, \ldots, x_m)^T$.

Assuming that the correspondent moments are finite, one has

$$\mathrm{E}(\mathbf{x}) = \boldsymbol{\mu} \tag{1.39}$$

$$\mathrm{Cov}(\mathbf{x}) = \frac{\mathrm{E}(\mathcal{R}_k^2)}{k}\boldsymbol{\Sigma} = \frac{\mathrm{E}(\mathcal{Q}_k)}{k}\boldsymbol{\Sigma} = -2\phi'(\mathbf{0})\boldsymbol{\Sigma}. \tag{1.40}$$

In particular for RCG distributions, (1.40) becomes

$$\mathrm{Cov}(\mathbf{x}) = \mathrm{E}(\tau)\boldsymbol{\Sigma}. \tag{1.41}$$

We see from (1.40) and (1.41) that under finite second-order moment assumption, the covariance of $\mathbf{x}$ does not necessarily coincide with the scatter matrix $\boldsymbol{\Sigma}$, 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 $\mathbf{V} \stackrel{\text{def}}{=} \frac{1}{s(\boldsymbol{\Sigma})}\boldsymbol{\Sigma}$ 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., $s(\boldsymbol{\Sigma}) \stackrel{\text{def}}{=} \mathrm{Tr}(\boldsymbol{\Sigma})/m$ leading to the following shape matrix $\mathbf{V} \stackrel{\text{def}}{=} \frac{m}{\mathrm{Tr}(\boldsymbol{\Sigma})}\boldsymbol{\Sigma}$. One can also find $s(\boldsymbol{\Sigma}) \stackrel{\text{def}}{=} |\boldsymbol{\Sigma}|^{1/m}$, leading to unit determinant for the shape matrix.

Otherwise, (1.40) and (1.41) can be used to resolve the scale ambiguity of the couple $(\boldsymbol{\Sigma}, \phi(.))$ in the definition (1.1) of the RES distribution by fixing the constraint on the characteristic generator $\phi(.)$

$$\mathrm{E}(\mathcal{R}_k^2) = \mathrm{E}(\mathcal{Q}_k) = k = \mathrm{rank}(\boldsymbol{\Sigma}) \ \text{ or } \ \mathrm{E}(\tau) = 1 \ \text{ for RCG distributions,} \tag{1.42}$$

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

When the r.v. $\mathbf{x}$ is absolutely continuous w.r.t. Lebesgue measure on $\mathbb{R}^m$ and $k = m$, (1.10) implies that (1.42) is equivalent to the following constraint on the density generator $g(.)$;

$$\delta_{m+2,g}/\delta_{m,g} = m. \tag{1.43}$$

If $\mathrm{E}(\mathcal{Q}_m)$ is not finite, rather imposing $\mathrm{Median}(\mathcal{R}_m) = 1$, (i.e., from (1.13) 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 it avoids any finite moment assumptions. Similarly for RCG distributions, if $\mathrm{E}(\tau)$ is not finite, the constraint $\mathrm{Median}(\tau) = 1$ (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 $\sigma_{i_1,i_2,\ldots,i_\ell} \stackrel{\text{def}}{=} \mathrm{E}[(x_{i_1} - \mu_{i_1})(x_{i_2} - \mu_{i_2})\ldots(x_{i_\ell} - \mu_{i_\ell})]$ with $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_m)^T$ and $(i_1, i_2, \ldots, i_\ell) \in \{1, \ldots, m\}^\ell$. By symmetry, all odd-order central moments are zero, provided that the corresponding moments do exist. As for fourth-order moments, if $\mathrm{E}(\mathcal{R}_m^4) < \infty$ (with $k = m$), they all satisfy the identity [50, p. 2]

$$\sigma_{i,j,k,\ell} = (\kappa+1)(\sigma_{i,j}\sigma_{k,\ell} + \sigma_{i,k}\sigma_{j,\ell} + \sigma_{i,\ell}\sigma_{j,k}), \quad (i,j,k,\ell) \in \{1,\dots,m\}^4, \tag{1.44}$$

where

$$\kappa \stackrel{\text{def}}{=} \frac{1}{3}\left(\frac{\sigma_{i,i,i,i}}{\sigma_{i,i}^2} - 3\right) = \frac{m}{m+2}\frac{\mathrm{E}(\mathcal{R}_m^4)}{(\mathrm{E}(\mathcal{R}_m^2))^2} - 1 = \frac{\phi''(0)}{(\phi'(0))^2} - 1. \tag{1.45}$$

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: $\sigma_{i,i,i,i} = 3\mathrm{E}(\tau^2)$ and $\sigma_{i,i} = \mathrm{E}(\tau)$ and thus

$$\kappa = \frac{\mathrm{Var}(\tau)}{[\mathrm{E}(\tau)]^2}. \tag{1.46}$$

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].

1.4.2 *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 (1.1) of $\mathbf{x}$ by

$$\Phi_y(\mathbf{t}) = \exp(i\mathbf{t}^T\mathbf{b})\Phi_x(\mathbf{B}^T\mathbf{t}) = \exp(i\mathbf{t}^T(\mathbf{B}\boldsymbol{\mu} + \mathbf{b}))\phi(\mathbf{t}^T\mathbf{B}\mathbf{\Sigma}\mathbf{B}^T\mathbf{t}), \quad \mathbf{t} \in \mathbb{R}^n. \tag{1.47}$$

Consequently, $\mathbf{y}$ is $\mathrm{RES}_n(\mathbf{B}\boldsymbol{\mu} + \mathbf{b}, \mathbf{B}\mathbf{\Sigma}\mathbf{B}^T, \phi)$-distributed, and thus the class of $\mathrm{RES}_m(\boldsymbol{\mu}, \mathbf{\Sigma}, \phi)$ distributions is closed under affine transformations. Note that the parameters are transformed as $(\boldsymbol{\mu}, \mathbf{\Sigma}) \mapsto (\mathbf{B}\boldsymbol{\mu} + \mathbf{b}, \mathbf{B}\mathbf{\Sigma}\mathbf{B}^T)$, 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 $\mathbf{x}$ (1.7), we derive

$$\mathbf{y} =_d \mathbf{B}\boldsymbol{\mu} + \mathbf{b} + \mathcal{R}_k(\mathbf{B}\mathbf{A})\mathbf{u}_k, \tag{1.48}$$

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 (1.7) and by its characteristic function (1.1), we get from (1.47) a stochastic full-rank representation of $\mathbf{y}$

$$\mathbf{y} =_d \mathbf{B}\boldsymbol{\mu} + \mathbf{b} + \mathcal{R}_\ell \mathbf{C}\mathbf{u}_\ell, \tag{1.49}$$

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 $\mathbf{x} = (x_1, \ldots, x_m)^T = (\mathbf{x}_1^T, \mathbf{x}_2^T)^T$, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_m)^T = (\boldsymbol{\mu}_1^T, \boldsymbol{\mu}_2^T)^T$ 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 $x_i \sim \mathrm{RES}_1(\mu_i, \sigma_{i,i}, \phi)$, $i = 1, \ldots, m$. Of course, their stochastic representations (1.48) and (1.49) 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

$$\mathbf{x}_1 =_d \boldsymbol{\mu}_1 + \mathcal{R}_k \mathbf{A}_1 \mathbf{u}_k \tag{1.50}$$

is a stochastic full-rank representation of the marginal $\mathbf{x}_1$. As for the univariate marginal x_i, the stochastic representations (1.48) and (1.49) reduce to

$$x_i =_d \mu_i + \mathcal{R}_k \mathbf{a}_i^T \mathbf{u}_k =_d \mu_i + \mathcal{R}_1 \|\mathbf{a}_i\| u_1, \tag{1.51}$$

where $\mathbf{a}_i$ is the ith 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 (1.50) and (1.51) of arbitrary univariate or multivariate marginal r.v. $\mathbf{x}_1$ and x_i reduce to

$$\mathbf{x}_1 =_d \boldsymbol{\mu}_1 + \mathcal{R}_{m_1} \boldsymbol{\Sigma}_{1,1}^{1/2} \mathbf{u}_{m_1} \quad \text{and} \quad x_i =_d \mu_i + \mathcal{R}_1 \sigma_{i,i}^{1/2} u_1, \tag{1.52}$$

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 [18, Corollary p. 59]

$$\mathcal{R}_{m_1} =_d \mathcal{R}_m \, \beta_{\frac{m_1}{2}, \frac{m_2}{2}}. \tag{1.53}$$

In (1.53), 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., $\mathbf{x} \sim \mathrm{RES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$, (1.53) 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 $\mathcal{R}_m$ [18, rel. 2.5.15]

$$p_{m_1}(r) = \frac{2r^{m_1-1}}{\mathrm{B}(\frac{m_1}{2}, \frac{m_2}{2})} \int_r^{+\infty} t^{-(m-2)}(t^2 - r^2)^{\frac{m_2}{2}-1} p_m(t)dt \tag{1.54}$$

and to the density generator $g(.)$ of $\mathbf{x}$ using $p_m(r)$ given by (1.13)

$$p_{m_1}(r) = \frac{2\pi^{m/2} r^{m_1-1}}{\Gamma(\frac{m_1}{2})\Gamma(\frac{m_2}{2})} \int_{r^2}^{+\infty} (t - r^2)^{\frac{m_2}{2}-1} g(t)dt. \tag{1.55}$$

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 (1.13) $p_{m_1}(r) = 2\delta_{m_1}^{-1} r^{m_1-1} g_{m_1|m}(r^2)$

$$g_{m_1|m}(u) = \delta_{m_2}^{-1} \int_u^{+\infty} (t - u)^{\frac{m_2}{2}-1} g(t)dt. \tag{1.56}$$

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

$$p_{m_1|m}(\mathbf{x}_1) = |\boldsymbol{\Sigma}_{1,1}|^{-1/2} g_{m_1|m}[(\mathbf{x}_1 - \boldsymbol{\mu}_1)^T \boldsymbol{\Sigma}_{1,1}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1)]. \tag{1.57}$$

$$p_{i|m}(x_i) = \frac{1}{\sqrt{\sigma_{i,i}}} g_{1|m}\left(\frac{(x_i - \mu_i)^2}{\sigma_{i,i}}\right). \tag{1.58}$$

Note that there is no guarantee that the integral in (1.56) 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 Sect. 1.2.4, it is clear from the stochastic representation (1.14) that

$$\mathbf{x}_1 =_d \boldsymbol{\mu}_1 + \sqrt{\tau}\mathbf{n}_1, \tag{1.59}$$

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 (1.56) reduces thanks to (1.21) to

$$g_{m_1|m}(t) = (2\pi)^{-m_1/2} \int_0^{\infty} \tau^{-m_1/2} \exp\left(-\frac{t}{2\tau}\right) dF_\tau(\tau). \tag{1.60}$$

In fact, this property characterizes the RCG distributions. This point is specified by a consistency result [32, Theorem 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. $\mathbf{x} = (x_1, \ldots, x_m)^T$ 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 [13].

1.4.3 Conditional Distributions

Let $\mathbf{x} = (\mathbf{x}_1^T, \mathbf{x}_2^T)^T \sim \text{RES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$ 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 Sect. 1.4.2. For the sake of simplicity, we consider here only the full-rank case ($k = m$); see e.g., [9] and [20, Theorem 7, Corollary 8] for more general statements. In this case, it is proved [17, Theorem 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 $\text{RES}_{m_2}(\boldsymbol{\mu}_{2|1}, \boldsymbol{\Sigma}_{2|1}, \phi_{2|1}) = \text{RES}_{m_2}(\boldsymbol{\mu}_{2|1}, \boldsymbol{\Sigma}_{2|1}, F_{2|1})$ distributed with

$$\boldsymbol{\mu}_{2|1} = \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{2,1}\boldsymbol{\Sigma}_{1,1}^{-1}(\mathbf{x}_1^0 - \boldsymbol{\mu}_1), \tag{1.61}$$

$$\boldsymbol{\Sigma}_{2|1} = \boldsymbol{\Sigma}_{2,2} - \boldsymbol{\Sigma}_{2,1}\boldsymbol{\Sigma}_{1,1}^{-1}\boldsymbol{\Sigma}_{1,2}, \tag{1.62}$$

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

$$(\mathbf{x}_2|\mathbf{x}_1 = \mathbf{x}_1^0) =_d \boldsymbol{\mu}_{2|1} + \mathcal{R}_{2|1}\mathbf{A}_{2|1}\mathbf{u}_{m_2}, \tag{1.63}$$

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

$$\mathcal{R}_{2|1} =_d [\mathcal{R}_m^2 - (\mathbf{x}_1^0 - \boldsymbol{\mu}_1)^T\boldsymbol{\Sigma}_{1,1}^{-1}(\mathbf{x}_1^0 - \boldsymbol{\mu}_1)]^{1/2}|\mathbf{x}_1 = \mathbf{x}_1^0. \tag{1.64}$$

Otherwise, if the conditional covariance $\text{Cov}(\mathbf{x}_2|\mathbf{x}_1 = \mathbf{x}_1^0)$ exists, its expressions can be derived from (1.40) and we get

$$\text{Cov}(\mathbf{x}_2|\mathbf{x}_1 = \mathbf{x}_1^0) = \frac{1}{m_2}\left(\text{E}[\mathcal{R}_m^2|\mathbf{x}_1 = \mathbf{x}_1^0] - (\mathbf{x}_1^0 - \boldsymbol{\mu}_1)^T\boldsymbol{\Sigma}_{1,1}^{-1}(\mathbf{x}_1^0 - \boldsymbol{\mu}_1)\right)\boldsymbol{\Sigma}_{2|1}. \tag{1.65}$$

We note that the expressions of the conditional symmetry center $\boldsymbol{\mu}_{2|1}$ and conditional scatter matrix $\boldsymbol{\Sigma}_{2|1}$ 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

$$p(\mathbf{x}_2/\mathbf{x}_1^0) = \frac{1}{p(\mathbf{x}_1^0)}\int_0^\infty p(\mathbf{x}_2/\mathbf{x}_1^0, \tau)p(\mathbf{x}_1^0/\tau)dF_\tau(\tau), \tag{1.66}$$

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

$$p(\mathbf{x}_2/\mathbf{x}_1^0) = \frac{(2\pi)^{-m_2/2}|\boldsymbol{\Sigma}_{2|1}|^{-1/2}}{\int_0^\infty \tau^{-m_1/2}\exp(-\frac{1}{2\tau}d^2_{\boldsymbol{\Sigma}_{1,1}}(\mathbf{x}_1^0,\boldsymbol{\mu}_1))dF_\tau(\tau)}$$
$$\times\int_0^\infty \tau^{-m/2}\exp\left(-\frac{1}{2\tau}d^2_{\boldsymbol{\Sigma}_{2|1}}(\mathbf{x}_2,\boldsymbol{\mu}_{2|1})\right)\exp\left(-\frac{1}{2\tau}d^2_{\boldsymbol{\Sigma}_{1,1}}(\mathbf{x}_1^0,\boldsymbol{\mu}_1)\right)dF_\tau(\tau). \quad (1.67)$$

Comparing (1.67) to (1.9), we check that $\mathbf{x}_2|\mathbf{x}_1 = \mathbf{x}_1^0$ is $\text{RES}_{m_2}(\boldsymbol{\mu}_{2|1},\boldsymbol{\Sigma}_{2|1},F_{2|1})$ distributed. Moreover, comparing (1.67) to (1.20), 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 $\mathbf{x}_1$.

1.4.4 Summation Stability

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

$$\Phi_y(\mathbf{t}) = \prod_{i=1}^n \Phi_{x_i}(\mathbf{t}) = \exp\left(i\mathbf{t}^T(\sum_{i=1}^n \boldsymbol{\mu}_i)\right)\prod_{i=1}^n \phi_i(\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}) = \exp(i\mathbf{t}^T\boldsymbol{\mu})\phi(\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}), \quad (1.68)$$

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 (1.1). Consequently, the sum $\mathbf{y}$ is RES-distributed too [20]. Similarly, for independent univariate r.v. $x_1,\ldots,x_i,\ldots,x_n$ but with arbitrary scatters σ_i, where $x_i \sim \text{RES}_1(\mu_i,\sigma_i,\phi_i)$, 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 $y \sim \text{RES}_1(\mu,1,\phi)$. The sum y is then symmetrically distributed w.r.t. μ.

However, it is worth underlying that if the r.v. $\mathbf{x}_1,\ldots,\mathbf{x}_i,\ldots,\mathbf{x}_n$ belong to the same family of RES distribution (i.e., with $\phi_i = \phi$, $i = 1,\ldots,n$), the sum is not of the same family except for the so-called elliptical α-stable distributions [45], (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 $\mathbf{x}_1,\ldots,\mathbf{x}_i,\ldots,\mathbf{x}_n$ 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 [30, Theorem 4.2]. If the full-rank stochastic representations (1.7) $\mathbf{x}_i =_d \boldsymbol{\mu}_i + \mathcal{R}_k^i\mathbf{A}\mathbf{u}_k^i$, $i = 1,2$ 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 $\text{RES}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$-distributed, where the expression of ϕ is given in [30, Theorem 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.

1.5 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 $\text{rank}(\boldsymbol{\Sigma}) = m$. 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. (1.9) where the characteristic and density generators, and the texture (1.14) are simply denoted here by respectively ϕ, g and τ, knowing that C and NC complex-valued definitions (1.28) and (1.33) are simply deduced from the real-valued definition with double dimension (1.24), (1.29). For C-CES distributions, interested readers can consult [49].

1.5.1 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

$$\phi(u) = \exp\left(-\frac{u}{2}\right) \text{ and } g(t) = \frac{1}{(2\pi)^{m/2}} \exp\left(-\frac{t}{2}\right). \tag{1.69}$$

Using (1.13), 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

$$p(q) = \frac{\delta_m^{-1}}{(2\pi)^{m/2}} q^{m/2-1} \exp\left(-\frac{q}{2}\right) = \frac{1}{2^{m/2}\Gamma(m/2)} q^{m/2-1} \exp\left(-\frac{q}{2}\right), \; q \in \mathbb{R}^+, \tag{1.70}$$

where $\delta_m = \frac{\Gamma(m/2)}{\pi^{m/2}}$ from (1.5). It is immediate to verify that the p.d.f. in (1.70) 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